aws

Guide de l’utilisateur

Amazon DataZone

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon DataZone: Guide de l’utilisateur

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Les marques et la présentation commerciale d’Amazon ne peuvent être utilisées en relation avec un produit ou un service qui n’est pas d’Amazon, d’une manière susceptible de créer une confusion parmi les clients, ou d’une manière qui dénigre ou discrédite Amazon. Toutes les autres marques commerciales qui ne sont pas la propriété d’Amazon appartiennent à leurs propriétaires respectifs, qui peuvent ou non être affiliés ou connectés à Amazon, ou sponsorisés par Amazon.

Table of Contents

Qu'est-ce qu'Amazon DataZone ? 1
..... 1
Comment Amazon DataZone prend en charge les autres AWS services et s'y intègre ? 2
Comment accéder à Amazon DataZone ? 2
Amazon SageMaker et quand utiliser Amazon par SageMaker rapport à Amazon DataZone 4
Terminologie et concepts 5
DataZone Composants Amazon 5
Que sont les DataZone domaines Amazon ? 6
Quels sont les DataZone projets et les environnements Amazon ? 7
Que sont les DataZone plans Amazon ? 10
Que sont les flux de DataZone production et de publication d'Amazon ? 13
Création des actifs d'inventaire du projet 13
Publication des actifs de l'inventaire du projet dans le DataZone catalogue Amazon 14
Quels sont les processus DataZone d'abonnement et d'expédition d'Amazon ? 15
Les personas des utilisateurs d'Amazon DataZone 16
DataZone Terminologie Amazon 17
Qu'est-ce qui est nouveau ? 27
2024 27
Amazon DataZone lance des règles d'application des métadonnées pour les demandes d'abonnement 27
Les plans AWS de service DataZone personnalisés d'Amazon permettent désormais à Amazon de SageMaker bénéficier d'une nouvelle expérience de configuration pour les projets Amazon DataZone 27
Amazon DataZone lance le AWS CloudFormation support pour les plans AWS de services personnalisés 28
Amazon DataZone lance des unités de domaine et des politiques d'autorisation 28
Amazon DataZone lance des produits de données 28
Amazon DataZone lance une fonctionnalité de contrôle d'accès précise 29
Amazon DataZone lance une fonctionnalité de traçabilité des données 29
Amazon DataZone lance des plans AWS de services personnalisés 30
Améliorations apportées au flux de création de sources de données 30
Amazon DataZone lance l'intégration avec Amazon SageMaker 31
Amazon DataZone lance l'intégration avec AWS le mode d'accès hybride Lake Formation 31
Amazon DataZone lance l'intégration avec AWS Glue Data Quality 31

Mise à disposition générale des recommandations relatives à l'IA pour les descriptions sur Amazon DataZone ... 32
Amazon apporte DataZone des améliorations à l'intégration d'Amazon Redshift ... 32
AWS Support à la formation dans le cloud pour Amazon DataZone ... 33
Ajoutez des responsables IAM directement en tant que membres de projets Amazon DataZone ... 34
Support pour les types d'actifs personnalisés depuis le portail de données ... 34
2023 ... 34
Supprimer le domaine ... 34
Mode hybride ... 35
Éligibilité HIPAA ... 35
Recommandations de l'IA pour les descriptions dans Amazon DataZone (version préliminaire) ... 35
DefaultDataLake amélioration du plan ... 36
Régions prises en charge ... 37
Configuration ... 38
Créez un AWS compte ... 38
Configurer les autorisations IAM requises pour utiliser la console de gestion ... 39
Associer des politiques obligatoires et facultatives à un utilisateur, un groupe ou un rôle pour accéder à la console de gestion ... 40
Créez une politique personnalisée pour les autorisations IAM afin de permettre la création de rôles simplifiée par la console de service de gestion ... 40
Créez une politique personnalisée pour les autorisations de gestion d'un compte associé à un domaine ... 42
(Facultatif) Créez une politique personnalisée pour les autorisations AWS d'Identity Center afin d'ajouter et de supprimer l'accès des utilisateurs SSO et des groupes SSO aux domaines ... 45
(Facultatif) Ajoutez votre principal IAM en tant qu'utilisateur clé pour créer votre domaine à l'aide d'une clé gérée par le client par KMS AWS ... 46
Configurer les autorisations IAM requises pour utiliser le portail de données ... 47
Associer la politique requise à un utilisateur, un groupe ou un rôle pour accéder au portail de données ... 47
Associer la politique requise à un utilisateur, un groupe ou un rôle pour accéder au catalogue ... 49

Associez une politique facultative à un utilisateur, un groupe ou un rôle pour accéder au portail de données ou au catalogue si votre domaine est chiffré à l'aide d'une clé gérée par le client par KMS AWS 50
Configuration de l' AWS IAM Identity Center pour Amazon DataZone 51
Prise en main 53
Guide de démarrage rapide avec des exemples de données AWS Glue 53
Étape 1 - Création du DataZone domaine Amazon et du portail de données 54
Étape 2 - Création du projet de publication 56
Étape 3 - Création de l'environnement 56
Étape 4 - Produire des données pour publication 57
Étape 5 - Collectez les métadonnées à partir de AWS Glue 58
Étape 6 - Organiser et publier la ressource de données 58
Étape 7 - Création du projet pour l'analyse des données 59
Étape 8 - Création d'un environnement pour l'analyse des données 59
Étape 9 - Rechercher dans le catalogue de données et s'abonner aux données 60
Étape 10 - Approuver la demande d'abonnement 60
Étape 11 - Création d'une requête et analyse des données dans Amazon Athena 60
Guide de démarrage rapide avec des exemples de données Amazon Redshift 61
Étape 1 - Création du DataZone domaine Amazon et du portail de données 61
Étape 2 - Création du projet de publication 63
Étape 3 - Création de l'environnement 64
Étape 4 - Produire des données pour publication 65
Étape 5 - Collectez les métadonnées depuis Amazon Redshift 65
Étape 6 - Organiser et publier la ressource de données 66
Étape 7 - Création du projet pour l'analyse des données 66
Étape 8 - Création d'un environnement pour l'analyse des données 67
Étape 9 - Rechercher dans le catalogue de données et s'abonner aux données 68
Étape 10 - Approuver la demande d'abonnement 68
Étape 11 - Création d'une requête et analyse des données dans Amazon Redshift 68
Exemples de scripts pour les tâches courantes 69
Création d'un DataZone domaine Amazon et d'un portail de données 69
Création d'un projet de publication 70
Création d'un profil d'environnement 70
Création d'un environnement 73
Collectez des métadonnées à partir de AWS Glue 74
Organiser et publier un actif de données 76

Rechercher dans le catalogue de données et s'abonner aux données 80
Rechercher des actifs dans le catalogue de données 80
Autres exemples de scripts utiles 83
Domaines et accès utilisateur 85
Création de domaines 85
Modifier des domaines 88
Supprimer des domaines 88
Activer IAM Identity Center pour Amazon DataZone 90
Désactiver IAM Identity Center pour Amazon DataZone 91
Gérer les utilisateurs dans la DataZone console Amazon 92
Gérer les rôles et les utilisateurs IAM 93
Gérer les utilisateurs SSO 94
Gérer les groupes SSO 96
Gérer les autorisations des utilisateurs dans le portail de données 97
Restreindre l'accès à Amazon DataZone 97
Mise à niveau DataZone des domaines Amazon vers des domaines SageMaker unifiés Amazon 98
Considérations à prendre en compte avant de mettre à niveau votre domaine 98
Mettez à niveau votre DataZone domaine Amazon vers un domaine SageMaker unifié Amazon 99
Questions fréquemment posées sur la mise à niveau de DataZone domaines Amazon vers des domaines SageMaker unifiés Amazon 100
Unités de domaine et politiques d'autorisation 102
Création d'unités de domaine 104
Modifier les unités de domaine 105
Supprimer des unités de domaine 105
Gérer les propriétaires d'unités de domaine 106
Attribuer des politiques d'autorisation aux utilisateurs et aux groupes au sein d'une unité de domaine 107
Politique d'adhésion au projet dans la hiérarchie des unités de domaine sur Amazon DataZone 108
Attribuer des politiques d'autorisation aux projets au sein d'une unité de domaine 114
Attribuez des politiques d'autorisation dans les configurations du plan 115
Plans intégrés 118
Activez les plans intégrés dans le AWS compte propriétaire du domaine Amazon DataZone 118

Ajoutez Amazon SageMaker en tant que service de confiance dans le AWS compte propriétaire du DataZone domaine Amazon ... 125
AWS Schémas de service personnalisés ... 126
Activer un plan AWS de service personnalisé ... 127
Création d'un environnement à l'aide d'un plan AWS de service personnalisé ... 127
Créez des actions dans un environnement AWS de service personnalisé ... 129
Ajouter des membres du projet à un environnement AWS de service personnalisé ... 130
Configuration d'une source de données dans un environnement AWS de service ... 130
Configuration d'un objectif d'abonnement dans un environnement AWS de service ... 131
Comptes associés ... 133
Demande d'association avec d'autres AWS comptes ... 133
Fournissez un accès au compte à votre clé KMS gérée par le client ... 134
Accepter une demande d'association de compte provenant d'un DataZone domaine Amazon et activer un plan d'environnement ... 135
Activer un plan d'environnement dans un compte associé AWS ... 136
Ajoutez Amazon SageMaker en tant que service de confiance dans le AWS compte associé ... 142
Rejeter une demande d'association de compte provenant d'un DataZone domaine Amazon ... 142
Supprimer un compte associé sur Amazon DataZone ... 143
Data catalog (Catalogue de données) ... 144
Création d'un glossaire commercial ... 145
Modifier un glossaire commercial ... 146
Supprimer un glossaire professionnel ... 147
Création d'un terme dans un glossaire ... 148
Modifier un terme dans un glossaire ... 149
Supprimer un terme dans un glossaire ... 150
Création d'un formulaire de métadonnées ... 151
Modifier un formulaire de métadonnées ... 152
Supprimer un formulaire de métadonnées ... 152
Création d'un champ dans un formulaire de métadonnées ... 154
Modifier un champ dans un formulaire de métadonnées ... 154
Supprimer un champ dans un formulaire de métadonnées ... 155
Projets et environnements ... 157
Création d'un profil d'environnement ... 158
Modifier un profil d'environnement ... 161
Supprimer un profil d'environnement ... 162
Créez un nouvel environnement ... 163

Modifier un environnement 164
Supprimer un environnement 164
Création d'un nouveau projet. 165
Modifier le projet 166
Déplacer le projet vers une autre unité de domaine 166
Supprimer le projet 167
Quitter le projet 169
Ajouter des membres à un projet 169
Supprimer des membres d'un projet 171
Inventaire et publication des données 172
Configurer les autorisations de Lake Formation pour Amazon DataZone 173
DataZone Intégration d'Amazon au mode hybride de AWS Lake Formation 174
Création de types d'actifs personnalisés 178
Créez et exécutez une source de données pour AWS Glue Data Catalog 182
Création et exécution d'une source de données pour Amazon Redshift 185
Modifier une source de données 189
Supprimer une source de données 190
Publier des actifs dans le catalogue à partir de l'inventaire du projet 190
Publier un actif 191
Gérez l'inventaire et organisez les actifs 192
Joindre des formulaires de métadonnées supplémentaires aux ressources 194
Publier la ressource dans le catalogue après curation 194
Création manuelle d'un actif 195
Dépublier un actif du catalogue 196
Supprimer un actif 196
Lancer manuellement l'exécution d'une source de données 197
Gestion des versions des actifs 198
Qualité des données sur Amazon DataZone 199
Permettre la qualité des données pour les actifs AWS de Glue 200
Permettre la qualité des données pour les types d'actifs personnalisés 201
Utilisation de l'apprentissage automatique et de l'IA générative sur Amazon DataZone 203
Régions prises en charge 204
Étapes d'utilisation de GenAI 205
Support pour les types d'actifs relationnels personnalisés 207
Quotas 207
Lignage des données sur Amazon DataZone 207

Types de nœuds de lignage dans Amazon DataZone 209
Attributs clés des nœuds de lignage 210
Visualisation du lignage des données 211
Autorisation du lignage des données sur Amazon DataZone 212
Exemple d'expérience de lignage de données sur Amazon DataZone 212
Activer le lignage des données dans la console de gestion 213
Utilisation du lignage DataZone de données Amazon par programmation 214
Automatisez le lignage pour le catalogue AWS Glue 215
Automatisez le lignage depuis Amazon Redshift 217
Règles d'application des métadonnées pour la publication 218
Produits de données 220
Création de nouveaux produits de données 221
Publiez des produits de données 221
Modifier des produits de données 222
Dépublier des produits de données 223
Supprimer des produits de données 224
S'abonner à un produit de données 225
Examiner une demande d'abonnement et accorder un abonnement à un produit de données .. 226
Republiez les produits de données 226
Découverte, abonnement et consommation de données 229
Rechercher et consulter des actifs dans le catalogue 230
Demander un abonnement à des actifs 232
Approuver ou rejeter une demande d'abonnement 233
Approbation automatique des demandes d'abonnement 235
Révoquer un abonnement existant 235
Annuler une demande d'abonnement 237
Se désabonner d'un actif 238
Utilisation des rôles IAM existants pour traiter les abonnements Amazon DataZone 239
Accorder l'accès aux AWS Glue Data Catalog actifs gérés 242
Accorder l'accès aux ressources Amazon Redshift gérées 243
Accorder l'accès aux actifs non gérés pour les abonnements approuvés 244
Interrogez des données dans Amazon Athena ou Amazon Redshift 245
Interrogez des données à l'aide d'Amazon Athena 246
Interrogez des données à l'aide d'Amazon Redshift 249
Règles d'application des métadonnées pour les demandes d'abonnement 251

Analysez les données que vous avez souscrites avec des applications d'analyse externes via une connexion JDBC ... 254
RedeemAccessToken Référence d'API ... 255
Contrôle d'accès précis aux données ... 258
Création de filtres de lignes ... 259
Création de filtres de colonnes ... 260
Supprimer les filtres de ligne ou de colonne ... 261
Modifier les filtres de ligne ou de colonne ... 261
Autoriser l'accès à l'aide de filtres ... 262
AWS Tables à colle ... 262
Amazon Redshift ... 263
Événements et notifications ... 264
Événements via la boîte de réception dédiée du portail de DataZone données Amazon ... 264
Événements via le bus EventBridge par défaut d'Amazon ... 272
Sécurité ... 275
Protection des données ... 276
Chiffrement des données ... 277
Chiffrement en transit ... 277
Confidentialité du trafic inter-réseaux ... 278
Le chiffrement des données est au repos pour Amazon DataZone ... 278
Utilisation des points de terminaison VPC d'interface pour Amazon DataZone ... 297
Autorisation sur Amazon DataZone ... 298
Autorisation dans la DataZone console Amazon ... 299
Autorisation sur le DataZone portail Amazon ... 299
DataZone Profils et rôles Amazon ... 299
Contrôle de l'accès ... 300
AWS politiques gérées ... 301
Rôles IAM pour Amazon DataZone ... 325
Informations d'identification temporaires ... 336
Autorisations de principal ... 337
Validation de conformité ... 337
Bonnes pratiques de sécurité ... 337
Implémentation d'un accès sur la base du moindre privilège ... 338
Utilisation des rôles IAM ... 338
Implémentation d'un chiffrement côté serveur dans des ressources dépendantes ... 338
CloudTrail À utiliser pour surveiller les appels d'API ... 339

Utilisation de la RAM sur Amazon DataZone 339
Résilience 339
Résilience des sources de données 340
Résilience des actifs 340
Résilience des types d'actifs et des formulaires de métadonnées 340
Glossaire : résilience 341
Résilience des recherches mondiales 341
Résilience des abonnements 341
Résilience environnementale 341
Résilience du plan environnemental 342
Résilience des projets 342
résilience de la RAM 342
Résilience de la gestion des profils utilisateur 342
Résilience du domaine 342
Sécurité de l'infrastructure sur Amazon DataZone 342
Prévention interservices confuse des adjoints sur Amazon DataZone 343
Analyse de configuration et de vulnérabilité dans for Amazon DataZone 343
Domaines à ajouter à votre liste d'autorisations 343
Surveillance 344
Surveillance des événements 345
CloudTrail journaux 345
DataZone Informations Amazon dans CloudTrail 345
Résolution des problèmes 347
Résolution des problèmes liés aux autorisations de AWS Lake Formation pour Amazon DataZone 347
Résolution des problèmes liés à la liaison des actifs Amazon DataZone Lineage avec des ensembles de données en amont 350
SourceIdentifier sur un nœud de lignée 351
Comment Amazon crée-t-il le DataZone SourceIdentifier à partir de l' OpenLineage événement ? 350
Approche alternative 357
Résolution d'un manque d'amont pour le nœud de traçabilité des actifs 357
Quotas 361
DataZone Quotas Amazon 11
Limites de débit des DataZone API Amazon 362
Historique de la documentation 368

..... cdxi

Qu'est-ce qu'Amazon DataZone ?

Amazon DataZone est un service de gestion des données qui vous permet de cataloguer, de découvrir, de partager et de gérer plus rapidement et plus facilement les données stockées auprès de sources tierces AWS, sur site ou sur site. Avec Amazon DataZone, les administrateurs qui supervisent les actifs de données de l'organisation peuvent gérer et gouverner l'accès aux données à l'aide de contrôles précis. Ces contrôles permettent de garantir un accès avec le bon niveau de privilèges et de contexte. Amazon DataZone permet aux ingénieurs, aux data scientists, aux chefs de produit, aux analystes et aux utilisateurs professionnels de partager et d'accéder facilement aux données au sein d'une organisation afin qu'ils puissent découvrir, utiliser et collaborer pour obtenir des informations basées sur les données.

Amazon vous DataZone aide à fournir des données directement aux utilisateurs finaux et simplifie votre architecture en intégrant des services de gestion des données, notamment Amazon Redshift, Amazon Athena, QuickSight Amazon, Glue AWS , AWS Lake Formation, des sources sur site, des sources tierces, etc.

Rubriques

- Que puis-je faire avec Amazon DataZone ?
- Comment Amazon DataZone prend en charge les autres AWS services et s'y intègre ?
- Comment accéder à Amazon DataZone ?

Que puis-je faire avec Amazon DataZone ?

Avec Amazon DataZone, vous pouvez effectuer les opérations suivantes :

- Gérez l'accès aux données au-delà des limites de l'organisation. Avec Amazon DataZone, vous pouvez garantir que les bonnes données sont accessibles par le bon utilisateur aux bonnes fins, conformément aux règles de sécurité de votre entreprise, sans vous fier à des informations d'identification individuelles. Vous pouvez également garantir la transparence sur l'utilisation des actifs de données et approuver les abonnements aux données à l'aide d'un flux de travail régi. Vous pouvez également surveiller les actifs de données dans l'ensemble des projets grâce à des fonctionnalités d'audit d'utilisation.
- Connectez les travailleurs des données grâce à des données et à des outils partagés pour obtenir des informations commerciales. Avec Amazon DataZone, vous pouvez améliorer l'efficacité de

votre équipe commerciale en collaborant harmonieusement entre les équipes et en fournissant un accès en libre-service aux données et aux outils d'analyse. Vous pouvez utiliser des termes commerciaux pour rechercher, partager et accéder aux données cataloguées stockées chez AWS, sur site ou auprès de fournisseurs tiers. Vous pouvez également en savoir plus sur les données que vous souhaitez utiliser en utilisant les glossaires DataZone commerciaux d'Amazon.

- Automatisez la découverte et le catalogage des données grâce à l'apprentissage automatique. Amazon DataZone vous permet de réduire le temps consacré à la saisie manuelle des attributs de données dans le catalogue de données commerciales. Des données plus riches dans le catalogue de données améliorent également l'expérience de recherche.

Comment Amazon DataZone prend en charge les autres AWS services et s'y intègre ?

Amazon DataZone prend en charge trois types d'intégrations avec d'autres AWS services :

- Sources de données du producteur : vous pouvez publier des actifs de données dans le DataZone catalogue Amazon à partir des données stockées dans les tables et vues AWS Glue Data Catalog et Amazon Redshift. Vous pouvez également publier manuellement des objets depuis Amazon Simple Storage Service (S3) dans le catalogue Amazon DataZone .
- Outils grand public : vous pouvez utiliser les éditeurs de requêtes Amazon Athena ou Amazon Redshift pour accéder à vos actifs de données et les analyser.
- Contrôle d'accès et traitement des commandes : Amazon DataZone prend en charge l'octroi de l'accès aux tables AWS Glue gérées par AWS Lake Formation et aux tables et vues Amazon Redshift. Pour tous les autres actifs de données, Amazon DataZone publie des événements standard liés à vos actions (par exemple, l'approbation donnée à une demande d'abonnement) sur Amazon EventBridge. Vous pouvez utiliser ces événements standard pour intégrer d'autres AWS services ou des solutions tierces pour des intégrations personnalisées.

Comment accéder à Amazon DataZone ?

Vous pouvez accéder DataZone à Amazon de l'une des manières suivantes :

- Console SageMaker de gestion Amazon ou DataZone console Amazon

  Vous pouvez utiliser la console SageMaker de gestion Amazon ou la console DataZone de gestion Amazon pour accéder à vos DataZone domaines, plans et utilisateurs Amazon et les configurer.

Pour plus d'informations, consultez https://console.aws.amazon.com/datazone. Cette console est également utilisée pour créer le portail de DataZone données Amazon.

- Portail DataZone de données Amazon

  Le portail de DataZone données Amazon est une application Web basée sur un navigateur dans laquelle vous pouvez cataloguer, découvrir, gérer, partager et analyser des données en libre-service. Le portail de données peut vous authentifier à l'aide des informations d'identification de votre fournisseur d'identité via AWS IAM Identity Center (successeur du AWS SSO) ou à l'aide de vos informations d'identification IAM. Vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone.

- API DataZone HTTPS Amazon

  Vous pouvez accéder à Amazon par DataZone programmation à l'aide de l'API Amazon DataZone HTTPS, qui vous permet d'envoyer des requêtes HTTPS directement au service. Pour plus d'informations, consultez le Amazon DataZone API Reference.

Amazon SageMaker et quand utiliser Amazon par SageMaker rapport à Amazon DataZone

Amazon SageMaker Catalog, basé sur Amazon DataZone, permet aux utilisateurs de gérer leurs actifs de données de manière centralisée. Vous pouvez cataloguer vos actifs de données, rechercher et découvrir les données, utiliser les fonctionnalités d'intelligence artificielle générative intégrées pour créer des métadonnées, ou vous pouvez simplement poser des questions en langage naturel à Amazon Q Developer pour trouver vos données. Les utilisateurs peuvent définir et appliquer des politiques d'accès de manière cohérente à l'aide d'un modèle d'autorisation unique avec des contrôles d'accès précis centralisés dans Amazon SageMaker Unified Studio. Vous pouvez créer un glossaire métier, étendre vos métadonnées et créer des produits de données qui peuvent être partagés avec de grandes équipes grâce à un contrôle d'accès précis. Vous pouvez également consulter les scores de qualité des données et découvrir le lignage des données de vos actifs de données.

Vous pouvez accéder à Amazon SageMaker Catalog depuis Amazon SageMaker Unified Studio. Unified Studio est une expérience de développement au sein d'Amazon SageMaker qui réunit AWS des services de données, d'analyse, d'intelligence artificielle (IA) et d'apprentissage automatique (ML). Il fournit un endroit pour créer, déployer, exécuter et surveiller les flux de travail à partir d'une interface unique. Cela permet de favoriser la collaboration entre les équipes et de faciliter le développement agile.

DataZone Terminologie et concepts d'Amazon

Amazon DataZone est un service de gestion des données qui vous permet de cataloguer, de découvrir, de partager et de gérer plus rapidement et plus facilement les données stockées auprès de sources tierces AWS, sur site ou auprès de sources tierces. Avec Amazon DataZone, les administrateurs et les responsables de la gestion des données chargés de superviser les actifs de données d'une entreprise peuvent gérer et régir l'accès aux données à l'aide de contrôles précis. Ces contrôles sont conçus pour garantir un accès avec le bon niveau de privilèges et de contexte. Amazon DataZone permet aux ingénieurs, aux data scientists, aux chefs de produit, aux analystes et aux utilisateurs professionnels d'accéder plus facilement aux données au sein d'une organisation afin qu'ils puissent découvrir, utiliser et collaborer pour obtenir des informations basées sur les données.

Lorsque vous débutez avec Amazon DataZone, il est important que vous compreniez ses concepts clés, sa terminologie et ses composants.

Rubriques

- DataZone Composants Amazon
- Que sont les DataZone domaines Amazon ?
- Quels sont les DataZone projets et les environnements Amazon ?
- Que sont les DataZone plans Amazon ?
- Que sont les flux de DataZone production et de publication d'Amazon ?
- Quels sont les processus DataZone d'abonnement et d'expédition d'Amazon ?
- Les personas des utilisateurs d'Amazon DataZone
- DataZone Terminologie Amazon

DataZone Composants Amazon

Amazon DataZone inclut les quatre principaux composants suivants :

- Catalogue de données commerciales : vous pouvez utiliser ce composant pour cataloguer les données de votre organisation en fonction du contexte commercial et permettre ainsi à tous les membres de votre organisation de trouver et de comprendre rapidement les données.
- Publiez et abonnez des flux de travail : vous pouvez utiliser ces flux de travail automatisés pour sécuriser les données entre les producteurs et les consommateurs en libre-service et pour garantir que tous les membres de votre organisation ont accès aux bonnes données aux bonnes fins.

- Projets et environnements
  - Dans Amazon, les DataZone projets sont des regroupements de personnes, d'actifs (données) et d'outils basés sur des cas d'utilisation professionnelle utilisés pour simplifier l'accès aux analyses. AWS Les projets fournissent des zones dans lesquelles les membres du projet peuvent collaborer, échanger des données et partager des actifs. Par défaut, les projets sont confiés de telle sorte que seuls ceux qui y sont explicitement ajoutés puissent accéder aux données et aux outils d'analyse qu'ils contiennent. Les projets gèrent la propriété des actifs produits conformément aux politiques du projet relatives à l'accès des consommateurs de données.
  - Au sein DataZone des projets Amazon, les environnements sont des ensembles de ressources configurées nulles ou plus (par exemple, un compartiment Amazon S3, une AWS Glue base de données ou un groupe de travail Amazon Athena) sur lesquels un ensemble donné de principes IAM (par exemple, les utilisateurs disposant d'autorisations de contributeur) peut opérer.
- Portail de données (en dehors de la console de AWS gestion) : il s'agit d'une application Web basée sur un navigateur dans laquelle différents utilisateurs peuvent accéder pour cataloguer, découvrir, gouverner, partager et analyser des données en libre-service. Le portail de données authentifie les utilisateurs à l'aide d'informations d'identification IAM ou d'informations d'identification existantes auprès de votre fournisseur d'identité via. AWS IAM Identity Center

Que sont les DataZone domaines Amazon ?

Vous pouvez utiliser DataZone les domaines Amazon pour organiser vos actifs, vos utilisateurs et leurs projets. En associant des AWS comptes supplémentaires à vos DataZone domaines Amazon, vous pouvez regrouper vos sources de données. Vous pouvez ensuite publier les ressources issues de ces sources de données dans le catalogue de votre domaine, à l'aide de formulaires de métadonnées et de glossaires qui améliorent l'exhaustivité et la qualité des métadonnées. Vous pouvez également rechercher et parcourir ces ressources pour voir quelles données sont publiées dans le domaine. En outre, vous pouvez rejoindre des projets pour collaborer avec d'autres utilisateurs, vous abonner à des ressources et utiliser des environnements de projet pour accéder à des outils d'analyse, notamment Amazon Athena et Amazon Redshift. Les DataZone domaines Amazon vous offrent la flexibilité nécessaire pour répondre aux besoins en données et en analyse de votre structure organisationnelle, qu'il s'agisse de créer un seul DataZone domaine Amazon pour votre entreprise ou plusieurs DataZone domaines Amazon pour différentes unités commerciales.

Quels sont les DataZone projets et les environnements Amazon ?

Amazon DataZone permet aux équipes et aux utilisateurs d'outils d'analyse de collaborer sur des projets en créant des groupes d'équipes, d'outils et de données basés sur des cas d'utilisation.

- Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de données du DataZone catalogue Amazon. Les membres du projet consomment les actifs du DataZone catalogue Amazon et en produisent de nouveaux à l'aide d'un ou de plusieurs flux de travail analytiques. Les projets soutiennent les activités suivantes au sein du portail de données :
  - Les propriétaires de projets peuvent ajouter des membres dotés des autorisations de propriétaire, de contributeur, de consommateur, de délégué et de spectateur
  - Les membres du projet peuvent être des utilisateurs SSO, des groupes SSO et des utilisateurs IAM
  - Les membres du projet peuvent demander un abonnement aux actifs du catalogue de données

    Les approbations d'abonnement sont fournies aux projets

	Créer sup prime des projet	Créer sup prime des profils de projet	Créer sup prime des profils d'envi neme	Créer sup prime des envirc ments	Ajoute s upprir des memt aux projet	Reche et décou e	Creat de lete metac forms glo ssarie	Créez des exécu s de sourc de donné et ingére des donné	Publie des donné	Dema des abonr ts	Apprc ou rejete les dema d'abo ent	Lire les donné d'abo ent prove d'Ama Athen et d'Ama Redsl	Création de filtres d'actifs
Propr ire	À gérer par le	À gérer par le	À gérer par le	À gérer par le	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui

	Créer supprimer des projets	Créer supprimer des profils de projet	Créer supprimer des profils d'environnement	Créer supprimer des environnements	Ajouter supprimer des membres aux projets	Rechercher et découvrir	Créer de lete métadonnées formulaires glossaire	Créez des exécutions de source de données et ingérer des données	Publier des données	Demander des abonnements	Approuver ou rejeter les demandes d'abonnement	Lire les données d'abonnement provenant d'Amazon Athena et d'Amazon Redshift	Création de filtres d'actifs
	membre de l'unité de domaine	membre de l'unité de domaine	membre de l'unité de domaine	membre de l'unité de domaine									
Participant	À gérer par le membre de l'unité de domaine	À gérer par le membre de l'unité de domaine	À gérer par le membre de l'unité de domaine	À gérer par le membre de l'unité de domaine	Non	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui

	Créer sup prime des projet	Créer sup prime des profils de projet	Créer sup prime des profils d'envi neme	Créer sup prime des envirc ments	Ajoute s upprir des memb aux projet	Reche et décou e	Creat de lete metac forms glo ssarie	Créez des exécu s de sourc de donné et ingére des donné	Publie des donné	Dema des abonr ts	Apprc ou rejete les dema d'abo ent	Lire les donné d'abo ent prove d'Ama Athen et d'Ama Redsl	Création de filtres d'actifs
Conso eur	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	Non	Oui	Non	Non	Non	Oui	Non	Oui	Non
Lecte	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	Non	Oui	Non	Non	Non	Non	Non	Oui	Non

	Créer sup prime des projet	Créer sup prime des profils de projet	Créer sup prime des profils d'envi neme	Créer sup prime des envirc ments	Ajoute s upprir des memb aux projet	Reche et décou e	Creat de lete metac forms glo ssarie	Créez des exécu s de sourc de donné et ingére des donné	Publie des donné	Dema des abonr ts	Apprc ou rejete les dema d'abo ent	Lire les donné d'abo ent prove d'Ama Athen et d'Ama Redsl	Création de filtres d'actifs
Intenc	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	À gérer par le memb de l'unité de doma	Non	Oui	Oui	Oui	Oui	Non	Oui	Oui	Non

- Dans un DataZone projet Amazon, les environnements sont des ensembles de ressources configurées nulles ou plus (par exemple, un Amazon S3, une AWS Glue base de données ou un groupe de travail Amazon Athena), avec un ensemble donné de principes IAM capables d'opérer sur ces ressources. Les environnements sont créés à l'aide de profilés d'environnement, qui sont des ensembles préconfigurés de ressources et de plans qui fournissent des modèles réutilisables pour créer des environnements. Les profils d'environnement définissent des paramètres tels que la région Compte AWS ou la région dans laquelle les environnements sont déployés.

Que sont les DataZone plans Amazon ?

Un plan avec lequel l'environnement est créé définit les AWS outils et services (par exemple, AWS Glue ou Amazon Redshift) que les membres du projet auquel appartient l'environnement peuvent utiliser lorsqu'ils travaillent avec les actifs du catalogue Amazon DataZone .

Dans la version actuelle d'Amazon DataZone, les plans par défaut suivants sont pris en charge :

Nom du plan	Description	Ressources créées
Plan directeur de Data Lake	Permet aux membres DataZone du projet Amazon de lancer les services aux producteurs et aux consommateurs de Data Lake au sein de l'environnement. En tant que consommateur, il permet aux membres DataZone du projet Amazon d'accéder à une copie « en lecture seule » des ressources gérées par Lake Formation directement dans Amazon Athena et dans d'autres moteurs de requêtes pris en charge par Lake Formation. En tant que producteur, il permet aux membres DataZone du projet Amazon de créer de nouvelles tables LakeFormation gérées à l'aide d'Amazon Athena et de les publier dans le catalogue Amazon DataZone.	Permet aux utilisateurs de créer et d'interroger des tables de Lake Formation à l'aide d'Amazon Athena. Groupe de travail Amazon Athena, AWS Glue base de données avec autorisations « lecture seule » sur Lake Formation, autorisations IAM « lecture seule » et accès à Amazon S3 géré par le projet. AWS Glue base de données avec autorisations « créer » et « accorder » à Lake Formation, autorisations IAM « lecture » et « écriture », AWS Glue ETL (extraction, transformation et chargement) avec balisage.
Plan d'entrepôt de données	En tant que consommateur, ce plan permet aux membres DataZone du projet Amazon de se connecter à leurs propres clusters Amazon Redshift pour interroger des magasins de données	Accès à l'éditeur de requêtes Amazon Redshift, accès en « lecture » aux sources de données abonnées depuis le DataZone catalogue Amazon, possibilité de créer des ressources locales dans le

Nom du plan	Description	Ressources créées
	distants et créer et stocker de nouveaux ensembles de données. En tant que producteur, ce plan permet aux membres DataZone du projet Amazon de se connecter à leurs propres clusters Amazon Redshift pour interroger des magasins de données distants, créer de nouveaux ensembles de données et les publier dans le catalogue Amazon. DataZone	cluster Amazon Redshift configuré. Accès à l'éditeur de requêtes Amazon Redshift, accès en « lecture » aux sources de données abonnées depuis le DataZone catalogue Amazon, possibilité de créer et de publier des ressource s à partir du cluster Amazon Redshift configuré.

Nom du plan	Description	Ressources créées
Plan Amazon Sagemaker	Ce plan aide les producteurs de données et les consommateurs à passer facilement SageMaker à Amazon pour collaborer sur des projets d'apprentissage automatique (ML) tout en renforçant la gouvernance de l'accès aux données et aux actifs de machine learning. Grâce à la nouvelle intégration intégrée entre Amazon DataZone et Amazon SageMaker, les consommateurs et les producteurs de données peuvent rationaliser la gouvernance du ML lors de la configuration de l'infrastructure, collaborer sur des initiatives commerciales et gérer facilement les données et les actifs de ML.	Vous pouvez créer un SageMaker domaine Amazon qui permet de rechercher, de souscrire et de publier des données et des actifs de machine learning sur Amazon DataZone. Vous pouvez également vous abonner et publier sur les bases de données AWS Glue et sur la formation des lacs selon la configuration.

Que sont les flux de DataZone production et de publication d'Amazon ?

Création des actifs d'inventaire du projet

Pour pouvoir utiliser Amazon DataZone pour cataloguer vos données, vous devez d'abord les importer (actifs) en tant qu'inventaire de votre projet sur Amazon DataZone. La création d'un inventaire pour un projet rend les actifs accessibles uniquement aux membres de ce projet. Les ressources de l'inventaire du projet ne sont pas accessibles à tous les utilisateurs du domaine,

search/browse sauf si elles sont publiées explicitement. Dans la version actuelle d'Amazon DataZone, vous pouvez ajouter des actifs à l'inventaire du projet de la manière suivante :

- Créez et exécutez des sources de données via le portail de données ou à l'aide d'Amazon DataZone APIs. Dans la version actuelle d'Amazon DataZone, vous pouvez créer et exécuter des sources de données pour AWS Glue et Amazon Redshift. En créant et en exécutant des sources de données AWS Glue ou Amazon Redshift, vous créez des actifs dans l'inventaire d'un projet choisi et vous importez leurs métadonnées techniques depuis les tables de la base de données source ou les entrepôts de données sous forme d'inventaire sur Amazon. DataZone
- À l'aide de APIs, vous pouvez créer des actifs à partir des types de ressources système disponibles (AWS Glue, Amazon Redshift, objets Amazon S3) ou à partir de vos types de ressources personnalisés.
  - Créez des types d'actifs personnalisés dans l'inventaire d'un projet à l'aide d'Amazon DataZone APIs. Les types d'actifs personnalisés peuvent inclure des modèles de machine learning, des tableaux de bord, des tables sur site, etc.
  - Créez des actifs à partir de ces types d'actifs personnalisés à l'aide d'Amazon DataZone APIs.
- Créez manuellement des ressources pour les objets S3 à l'aide du portail DataZone de données Amazon.

Gestion des actifs d'inventaire de votre projet : après avoir créé un inventaire de projet, les propriétaires de données peuvent organiser leurs actifs d'inventaire avec les métadonnées commerciales requises en ajoutant ou en mettant à jour les noms commerciaux (actif et schéma), les descriptions (actif et schéma), lisez-moi, les termes du glossaire (actif et schéma) et les formulaires de métadonnées. Vous pouvez le faire via le portail de données ou en utilisant Amazon DataZone APIs. Chaque modification apportée à votre actif crée une nouvelle version de l'inventaire.

Publication des actifs de l'inventaire du projet dans le DataZone catalogue Amazon

L'étape suivante de l'utilisation d'Amazon DataZone pour cataloguer vos données consiste à rendre les actifs d'inventaire de votre projet accessibles aux utilisateurs du domaine. Vous pouvez le faire en publiant les actifs d'inventaire dans le DataZone catalogue Amazon. Seule la dernière version de la ressource d'inventaire peut être publiée dans le catalogue et seule la dernière version publiée est active dans le catalogue de découverte. Si un actif d'inventaire est mis à jour après sa publication dans le DataZone catalogue Amazon, vous devez le republier explicitement pour que la dernière version figure dans le catalogue de découverte. Dans la version actuelle d'Amazon DataZone, vous

pouvez publier les actifs d'inventaire de votre projet dans le DataZone catalogue Amazon de la manière suivante :

- Publiez manuellement les actifs d'inventaire de votre projet dans le DataZone catalogue Amazon via le portail de données ou via Amazon DataZone APIs.
- Dans le cadre de la création ou de la modification de sources de données, activez les paramètres facultatifs Publish your AWS Glue dans le catalogue ou Publiez vos actifs Amazon Redshift dans le catalogue à utiliser lors des exécutions planifiées ou automatisées des sources de données. Lorsque ce paramètre est activé, l'exécution d'une source de données ajoute des actifs à l'inventaire de votre projet, puis publie également les actifs d'inventaire dans le DataZone catalogue Amazon. Notez que si vous publiez directement, les ressources peuvent ne pas contenir de métadonnées commerciales et seront directement accessibles à tous les utilisateurs du domaine. Vous pouvez utiliser ce paramètre sur vos sources de données via le portail de données ou via Amazon DataZone APIs.

Quels sont les processus DataZone d'abonnement et d'expédition d'Amazon ?

Une fois vos actifs publiés dans le DataZone catalogue Amazon, les utilisateurs de votre domaine peuvent les découvrir, demander et accéder à ces actifs, et continuer à utiliser Amazon DataZone pour gérer, partager et analyser ces actifs.

Les utilisateurs demandent l'accès à une ressource en s'abonnant à cette ressource pour le compte d'un projet. Une fois qu'une demande d'abonnement est créée, les propriétaires de l'actif reçoivent une notification et peuvent examiner la demande d'abonnement et décider de l'approuver ou de la rejeter. Si la demande d'abonnement est approuvée par le propriétaire des données, le projet abonné obtient l'accès à cette ressource.

Une fois qu'une demande d'abonnement est approuvée, Amazon DataZone lance un flux de traitement des abonnements qui ajoute automatiquement l'actif à tous les environnements applicables au sein du projet en créant les subventions nécessaires dans AWS Lake Formation ou Amazon Redshift. Cela permet aux membres du projet abonnés d'interroger la ressource à l'aide de l'un des outils de requête (Amazon Athena ou éditeur de requêtes Amazon Redshift) de leur environnement.

Amazon DataZone peut déclencher cette logique d'expédition automatisée uniquement pour les actifs gérés (cela inclut les tables AWS Glue et les tables et vues Amazon Redshift). Pour

tous les autres types d'actifs (actifs non gérés), Amazon ne DataZone peut pas déclencher automatiquement l'expédition, mais publie un événement sur Amazon Eventbridge avec tous les détails nécessaires dans la charge utile de l'événement afin que vous puissiez créer les subventions nécessaires en dehors d'Amazon. DataZone Amazon fournit DataZone également l'`updateSubscriptionStatus`API qui vous permet de mettre à jour le statut de l'abonnement une fois qu'il est rempli en dehors d'Amazon DataZone afin qu'Amazon DataZone puisse informer les membres du projet qu'ils peuvent commencer à consommer l'actif.

Les personas des utilisateurs d'Amazon DataZone

Les principaux DataZone utilisateurs d'Amazon sont les suivants :

- Administrateurs de domaine responsables de la configuration d'Amazon DataZone en tant que plateforme d'analyse pour leur organisation.

  Dans le contexte d'Amazon DataZone, les administrateurs de domaines installent Amazon DataZone dans les AWS comptes, créent des DataZone domaines Amazon et configurent les associations de AWS comptes et de fournisseurs d'identité avec les DataZone domaines Amazon. Les administrateurs de domaine utilisent également d'autres consoles de AWS service telles que AWS Organization et Service Catalog pour configurer Amazon DataZone.
- Utilisateurs de données qui sont les principaux utilisateurs d'Amazon DataZone (éditeurs d'actifs et abonnés) pour leurs tâches d'analyse et d'apprentissage automatique.

  Les utilisateurs de données incluent les professionnels de l'analyse des données, les scientifiques des données et les utilisateurs du système qui produisent et consomment des actifs de données. Dans le contexte d'Amazon DataZone, les utilisateurs de données créent et rejoignent des projets et des environnements, s'abonnent et consomment des actifs de données à l'aide d'outils d'analyse ou d'apprentissage automatique préconfigurés, et publient les actifs de données de sortie dans le catalogue de DataZone domaines Amazon pour les partager avec d'autres.
- Développeurs de systèmes qui créent des modèles d'infrastructure personnalisés et intègrent Amazon DataZone à des catalogues internes ou à des systèmes de production.

  Dans le contexte d'Amazon DataZone, les développeurs de systèmes élaborent des plans d'environnement (modèles d'infrastructure) ou des Infrastructure-As-Code CI/CD pipelines en tant que fournisseur d'environnement, des pipelines de données pour promouvoir les actifs de données dans différents environnements, des adaptateurs de synchronisation de catalogues et de distribution des subventions d'abonnement à intégrer aux catalogues internes, ou des intégrations

entre Amazon DataZone APIs et les interfaces utilisateur internes ou les systèmes de production si nécessaire.

- Des responsables de la gouvernance des données qui maîtrisent les définitions et les risques liés à la sécurité organisationnelle, à la confidentialité et aux autres politiques de conformité et qui s'assurent que l'utilisation d'Amazon DataZone dans leur entreprise est conforme à ces définitions.

DataZone Terminologie Amazon

Domain

Un DataZone domaine Amazon est l'entité organisatrice qui connecte vos actifs, vos utilisateurs et leurs projets. Avec DataZone les domaines Amazon, vous avez la flexibilité nécessaire pour répondre aux besoins en données et en analyse de votre structure organisationnelle, qu'il s'agisse de créer un seul DataZone domaine Amazon pour votre entreprise ou plusieurs zones de données, des domaines pour différentes unités commerciales ou équipes.

Unité de domaine

Les unités de domaine vous permettent d'organiser facilement vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour configurer un partage de données sécurisé et efficace au sein et entre les unités commerciales de votre organisation, vous pouvez créer des unités de domaine au sein d'Amazon DataZone et permettre à des utilisateurs sélectionnés au sein de chaque unité commerciale de se connecter et de partager leurs actifs dans le catalogue. Les unités de domaine peuvent également être utilisées pour permettre aux propriétaires de ressources, tels que les propriétaires de AWS comptes, de configurer des DataZone autorisations Amazon sur leurs ressources. Les unités de domaine fournissent une autorité déléguée des propriétaires de comptes aux propriétaires d'unités de domaine et elles peuvent configurer des autorisations d'autorisation sur les profils d'environnement (créés à l'aide de configurations de plan), au nom des propriétaires de comptes. Pour de plus amples informations, veuillez consulter Unités de domaine et politiques d'autorisation sur Amazon DataZone.

Politique d'autorisation

Les politiques DataZone d'autorisation Amazon sont un ensemble de contrôles au sein d'Amazon DataZone appliqués à des entités telles que les projets, les plans, les environnements, le glossaire et les formulaires de métadonnées. Ces politiques définissent qui peut créer ces entités et gérer leur cycle de vie sur le DataZone portail Amazon.

Au sein d'une unité de DataZone domaine Amazon, vous pouvez attribuer les politiques d'autorisation suivantes à vos utilisateurs et groupes afin de leur accorder des autorisations spécifiques :

- Politique de création d'unités de domaine
- Politique de création de projets
- Politique d'adhésion au projet
- Politique d'hypothèse de propriété des unités de domaine
- Politique d'hypothèse de propriété du projet

Pour de plus amples informations, veuillez consulter Attribuer des politiques d'autorisation aux utilisateurs et aux groupes au sein d'une unité DataZone de domaine Amazon.

Au sein d'une unité de DataZone domaine Amazon, vous pouvez attribuer les politiques d'autorisation suivantes à vos projets afin de leur accorder des autorisations spécifiques :

- Politique de création de glossaires
- Politique de création de formulaires de métadonnées
- Politique de création de types d'actifs personnalisés

Pour de plus amples informations, veuillez consulter Attribuer des politiques d'autorisation aux projets au sein d'une unité DataZone de domaine Amazon.

Dans une configuration de plan spécifique, vous pouvez attribuer les politiques d'autorisation suivantes aux propriétaires de projets et d'unités de domaine :

- Créez des profils d'environnement à l'aide de ce plan. Cette politique peut être attribuée aux DataZone projets Amazon et les autorise à créer des profils d'environnement à l'aide de ce plan.
- Accordez des autorisations pour créer des profils d'environnement à l'aide de ce plan : cette politique peut être attribuée aux propriétaires d'unités de domaine et les autorise à accorder des autorisations aux projets pour créer des profils d'environnement à l'aide de ce plan.

Pour de plus amples informations, veuillez consulter Attribuez des politiques d'autorisation dans les configurations Amazon DataZone Blueprint.

Compte associé

L'association de vos AWS comptes à des DataZone domaines Amazon vous permet de publier les données de ces AWS comptes dans le DataZone catalogue Amazon et de créer DataZone des projets Amazon pour exploiter vos données sur plusieurs AWS comptes. Les demandes

d'association de comptes ne peuvent être initiées que sur AWS des comptes possédant un DataZone domaine Amazon. Les demandes d'association de comptes ne peuvent être acceptées que par les utilisateurs administratifs des AWS comptes invités. Une fois qu'un AWS compte est associé à un DataZone domaine Amazon, vous pouvez enregistrer vos sources de données telles que le catalogue AWS Glue et Amazon Redshift dans ce compte sur ce domaine. L'association permet également à un AWS compte de créer des DataZone projets et des environnements Amazon.

Un Compte AWS peut être associé à un ou plusieurs DataZone domaines Amazon.

Source de données

Dans Amazon DataZone, vous pouvez utiliser des sources de données pour importer les métadonnées techniques des actifs (données) depuis les bases de données sources ou les entrepôts de données vers Amazon DataZone. Dans la version actuelle d'Amazon DataZone, vous pouvez créer et exécuter des sources de données pour AWS Glue et Amazon Redshift. En créant une source de données, vous établissez une connexion entre Amazon DataZone et la source (AWS Glue Data Catalog ou Amazon Redshift Warehouse) qui vous permet de lire les métadonnées techniques, notamment les noms de tables, les noms de colonnes et les types de données. En créant une source de données, vous lancez également l'exécution initiale de la source de données qui crée de nouvelles ressources ou met à jour des actifs existants sur Amazon DataZone. Lors de la création d'une source de données ou une fois celle-ci créée avec succès, vous avez également la possibilité de définir un calendrier pour les exécutions de votre source de données.

Exécution de la source de données

Dans Amazon DataZone, une exécution de source de données est une tâche qu'Amazon DataZone exécute afin de créer des actifs dans les inventaires de projets et également de publier éventuellement des actifs d'inventaire de projet dans le DataZone catalogue Amazon. Les exécutions de sources de données peuvent être automatisées (lancées lors de la création initiale d'une source de données), planifiées ou manuelles. Les critères de sélection des données vous permettent d'affiner les ensembles de données existants et futurs à intégrer dans les inventaires des projets ou le DataZone catalogue Amazon, ainsi que la fréquence des mises à jour des métadonnées de ces actifs d'inventaire ou de catalogue.

Objectif d'abonnement

Sur Amazon DataZone, les objectifs d'abonnement vous permettent d'accéder aux données auxquelles vous êtes abonné dans le cadre de vos projets. Un objectif d'abonnement indique l'emplacement (par exemple, une base de données ou un schéma) et les autorisations requises

(par exemple, un rôle IAM) qu'Amazon DataZone peut utiliser pour établir une connexion avec les données sources et pour créer les autorisations nécessaires afin que les membres du DataZone projet Amazon puissent commencer à interroger les données auxquelles ils se sont abonnés.

Demande d'abonnement

Sur Amazon DataZone, une demande d'abonnement est un processus qu'un DataZone projet Amazon doit suivre pour avoir accès à un actif spécifique. Les demandes d'abonnement peuvent être approuvées, rejetées, révoquées ou accordées.

Ressource

Dans Amazon DataZone, un actif est une entité qui présente un seul objet de données physique (par exemple, un tableau, un tableau de bord, un fichier) ou un objet de données virtuel (par exemple, une vue).

Type de ressource

Les types d'actifs définissent la manière dont les actifs sont représentés dans le DataZone catalogue Amazon. Un type d'actif définit le schéma d'un type d'actif spécifique. Lorsque des actifs sont créés, ils sont validés par rapport au schéma défini par leur type d'actif (par défaut, la dernière version). Lorsqu'une mise à jour d'un actif a lieu, Amazon DataZone crée une nouvelle version d'actif et permet aux DataZone utilisateurs d'Amazon d'opérer sur toutes les versions d'actifs.

Glossaire commercial

Sur Amazon DataZone, un glossaire commercial est un ensemble de termes commerciaux qui peuvent être associés à des actifs. Un glossaire métier permet de garantir que les mêmes termes et définitions sont utilisés au sein d'une organisation dans ses différentes tâches d'analyse de données.

Les termes d'un glossaire commercial peuvent être ajoutés aux actifs et aux colonnes pour classer ou améliorer l'identification de ces attributs lors de la recherche. Le glossaire peut être sélectionné comme type de valeur pour un champ dans un formulaire de métadonnées associé à une ressource. Lorsqu'un terme particulier est sélectionné comme valeur pour le champ du formulaire de métadonnées d'un actif, les utilisateurs peuvent rechercher le terme du glossaire commercial et trouver les actifs associés.

Type de formulaire de métadonnées

Un type de formulaire de métadonnées est un modèle qui définit les métadonnées collectées et enregistrées lorsque les actifs sont créés sous forme d'inventaire ou publiés dans un DataZone

domaine Amazon. Les types de formulaires de métadonnées peuvent être associés à une ressource de données. Les types de formulaires de métadonnées aident les administrateurs de domaine à définir les formulaires de métadonnées nécessaires pour ce domaine, tels que les informations de conformité, les informations réglementaires ou les classifications. Il permet aux administrateurs de domaine de personnaliser des métadonnées supplémentaires pour leurs actifs. Amazon DataZone propose des types de formulaires de métadonnées système tels que asset-common-details-form -type, column-business-metadata-form -type, glue-table-form-type, glue-view-form-type, redshift-table-form-type, redshift-view-form-type, s3- object-collection-form-type subscription-terms-form-type, et. suggestion-form-type

Formulaire de métadonnées

Dans Amazon DataZone, les formulaires de métadonnées définissent les métadonnées collectées et enregistrées lorsque les actifs sont créés sous forme d'inventaire ou publiés dans un DataZone domaine Amazon. Les définitions des formulaires de métadonnées sont créées dans le domaine du catalogue par un administrateur de domaine. Une définition de formulaire de métadonnées est composée d'une ou de plusieurs définitions de champs, avec prise en charge des types de données booléens, datés, décimaux, entiers, chaînes et valeurs de champs du glossaire commercial.

Un administrateur de domaine applique un formulaire de métadonnées aux actifs de son domaine en ajoutant le formulaire de métadonnées à son domaine. Les éditeurs de ressources fournissent ensuite toutes les valeurs de champ facultatives et obligatoires dans le formulaire de métadonnées.

Project

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale qui impliquent de créer des actifs dans les inventaires de projets et de les rendre ainsi accessibles à tous les membres du projet, puis de publier, de découvrir, de souscrire et de consommer des actifs dans le catalogue Amazon DataZone . Les membres du projet consomment les actifs du DataZone catalogue Amazon et en produisent de nouveaux à l'aide d'un ou de plusieurs flux de travail analytiques. Les membres du projet peuvent être des propriétaires, des contributeurs, des consommateurs, des gérants et des spectateurs.

	Créer/ sup primer des projets	Créer/ sup primer des profils de projet	Créer/ sup primer des profils d'envi nemer	Créer/ sup primer des enviro ments	Ajoute s upprin des memb aux projets	Reche et décou e	Create de lete metad forms/ glo ssarie	Créez des exécu s de source de donné et ingére des donné	Publie des donné	Dema des abonn ts	Appro ou rejeter les demar d'abor ent	Lire les données d'abonnem ent provenant d'Amazon Athena et d'Amazon Redshift
Propri ire	À gérer par le memb de l'unité de domai	À gérer par le memb de l'unité de domai	À gérer par le memb de l'unité de domai	À gérer par le memb de l'unité de domai	Oui	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Partici nt	À gérer par le memb de l'unité de domai	À gérer par le memb de l'unité de domai	À gérer par le memb de l'unité de domai	À gérer par le memb de l'unité de domai	Non	Oui	Oui	Oui	Oui	Oui	Oui	Oui
Consc eur	À gérer par le memb	À gérer par le memb	À gérer par le memb	À gérer par le memb	Non	Oui	Non	Non	Non	Oui	Non	Oui

	Créer/supprimer des projets	Créer/supprimer des profils de projet	Créer/supprimer des profils d'environnement	Créer/supprimer des environments	Ajouter/supprimer des membres aux projets	Rechercher et découvrir	Create de lete metadforms/glossarie	Créez des exécutions de source de données et ingérer des données	Publier des données	Demander des abonnements	Approuver ou rejeter les demandes d'abonnement	Lire les données d'abonnement provenant d'Amazon Athena et d'Amazon Redshift
	de l'unité de domai	de l'unité de domai	de l'unité de domai	de l'unité de domai								
Lecteu	À gérer par le membre de l'unité de domai	À gérer par le membre de l'unité de domai	À gérer par le membre de l'unité de domai	À gérer par le membre de l'unité de domai	Non	Oui	Non	Non	Non	Non	Non	Oui
Intend	À gérer par le membre de l'unité de domai	À gérer par le membre de l'unité de domai	À gérer par le membre de l'unité de domai	À gérer par le membre de l'unité de domai	Non	Oui	Oui	Oui	Oui	Non	Oui	Oui

Les propriétaires de projets peuvent ajouter ou supprimer d'autres utilisateurs en tant que propriétaires ou contributeurs, et ils peuvent modifier ou supprimer des projets. D'autres restrictions imposées aux contributeurs peuvent être définies à l'aide de politiques. Lorsqu'un utilisateur crée un projet, il en devient le premier propriétaire.

Environnement

Un environnement est un ensemble de ressources configurées (par exemple, un compartiment Amazon S3, une AWS Glue base de données ou un groupe de travail Amazon Athena), avec un ensemble donné de principes IAM (avec des autorisations de contributeur attribuées) qui peuvent opérer sur ces ressources. Chaque environnement peut également avoir des utilisateurs principaux autorisés à accéder aux ressources et aux données via un abonnement et un traitement des commandes. Les environnements sont conçus pour stocker des liens exploitables vers des AWS services, des applications externes IDEs et des consoles. Les membres du projet peuvent accéder à des services tels que la console Amazon Athena et bien d'autres via des liens profonds configurés au sein d'un environnement. Les utilisateurs SSO et IAM du projet peuvent être limités à use/access des environnements spécifiques.

Profil environnemental

Sur Amazon DataZone, un profil d'environnement est un modèle que vous pouvez utiliser pour créer des environnements. Les profils d'environnement sont créés à l'aide de plans.

Avec les profils d'environnement, les administrateurs de domaine peuvent encapsuler des plans avec des paramètres préconfigurés, puis les travailleurs des données peuvent créer rapidement un certain nombre de nouveaux environnements en sélectionnant les profils d'environnement existants et en spécifiant les noms des nouveaux environnements. Cela permet aux travailleurs des données de gérer efficacement leurs projets et leurs environnements tout en s'assurant qu'ils respectent les politiques de gouvernance des données appliquées par leurs administrateurs de domaine.

Plan

Un plan avec lequel l'environnement est créé définit les AWS outils et services (par exemple, AWS Glue ou Amazon Redshift) que les membres du projet auquel appartient l'environnement peuvent utiliser lorsqu'ils travaillent avec les actifs du catalogue Amazon DataZone .

Dans la version actuelle d'Amazon, DataZone les plans par défaut suivants sont pris en charge :

- Plan du lac de données
- Plan d'entrepôt de données

- Plan Amazon Sagemaker

Profil de l'utilisateur

Un profil utilisateur représente DataZone les utilisateurs d'Amazon. Amazon DataZone prend en charge à la fois les rôles IAM et les identités SSO pour interagir avec Amazon DataZone Management Console et le portail de données à différentes fins. Les administrateurs de domaine utilisent les rôles IAM pour effectuer le travail administratif initial lié au domaine dans Amazon DataZone Management Console, notamment la création de nouveaux DataZone domaines Amazon, la configuration des types de formulaires de métadonnées et la mise en œuvre de politiques. Les travailleurs des données utilisent leur identité d'entreprise SSO via Identity Center pour se connecter à Amazon DataZone Data Portal et accéder aux projets auxquels ils sont membres.

Profil du groupe

Les profils de groupe représentent des groupes d' DataZone utilisateurs d'Amazon. Les groupes peuvent être créés manuellement ou mappés à des groupes Active Directory de clients professionnels. Sur Amazon DataZone, les groupes ont deux objectifs. Tout d'abord, un groupe peut être mappé à une équipe d'utilisateurs dans l'organigramme, réduisant ainsi le travail administratif du DataZone chef de projet Amazon lorsque de nouveaux employés rejoignent ou quittent une équipe. Ensuite, les administrateurs d'entreprise utilisent des groupes Active Directory pour gérer et mettre à jour les statuts des utilisateurs. Les administrateurs de DataZone domaine Amazon peuvent donc utiliser ces appartenances à des groupes pour mettre en œuvre les politiques de DataZone domaine Amazon.

Administrateur de domaine

Dans Amazon DataZone, le principal IAM qui crée un DataZone domaine Amazon est l'administrateur de domaine par défaut de ce domaine. Les administrateurs de domaine d'Amazon DataZone exécutent les fonctionnalités clés du domaine, notamment la création de domaines, l'affectation d'autres administrateurs de domaine, l'ajout de sources de données et de cibles d'abonnement, la création de projets et d'environnements et l'attribution de propriétaires de projets.

Éditeur

Dans Amazon DataZone, les éditeurs publient des actifs dans le DataZone catalogue Amazon et peuvent modifier les métadonnées des actifs qu'ils publient. Si cette autorisation leur est accordée, les éditeurs peuvent approuver ou rejeter les demandes d'abonnement aux ressources qu'ils ont publiées dans le DataZone catalogue Amazon.

Subscriber

Dans Amazon DataZone, un abonné est un DataZone projet Amazon qui souhaite trouver, accéder et consommer des actifs du DataZone catalogue Amazon.

Compte AWS owner

Dans Amazon DataZone, Compte AWS les propriétaires créent des rôles, des politiques et des autorisations Comptes AWS qui permettent Comptes AWS de les associer à des DataZone domaines Amazon.

Quelles sont les nouveautés d'Amazon DataZone ?

Cette section décrit les nouvelles fonctionnalités et améliorations d'Amazon DataZone par date de sortie.

Rubriques

- 2024
- 2023

2024

Amazon DataZone lance des règles d'application des métadonnées pour les demandes d'abonnement

Publié le 20/11/2024

Les nouvelles règles d'application des métadonnées pour les demandes d'abonnement sur Amazon DataZone renforcent la gouvernance des données en permettant aux propriétaires d'unités de domaine d'établir des exigences claires en matière de métadonnées pour les consommateurs de données, en rationalisant les demandes d'accès et en améliorant la gouvernance des données. Cette fonctionnalité permet aux organisations de s'aligner sur les normes de métadonnées de l'organisation, de mettre en œuvre des flux de travail personnalisés et de fournir une expérience d'accès aux données cohérente et gouvernée. Pour de plus amples informations, veuillez consulter Règles d'application des métadonnées pour les demandes d'abonnement.

Les plans AWS de service DataZone personnalisés d'Amazon permettent désormais à Amazon de SageMaker bénéficier d'une nouvelle expérience de configuration pour les projets Amazon DataZone

Publié le 15/11/2024

Avec Amazon DataZone Custom AWS Service Prints, vous pouvez migrer votre SageMaker domaine Amazon existant vers Amazon DataZone. Grâce à cette fonctionnalité, les administrateurs peuvent désormais configurer DataZone des projets Amazon en important leurs utilisateurs autorisés existants, leurs configurations de sécurité et leurs politiques depuis les SageMaker

domaines Amazon. Pour plus d'informations, voir Configuration SageMaker des actifs (guide de l'administrateur).

Amazon DataZone lance le AWS CloudFormation support pour les plans AWS de services personnalisés

Publié le 9/12/2024

Amazon DataZone a ajouté la AWS CloudFormation prise en charge des plans AWS de service personnalisés. Cette nouvelle fonctionnalité vous permet d' AWS CloudFormation automatiser la création d'environnements dans Amazon DataZone. Grâce à des plans personnalisés, les administrateurs peuvent désormais intégrer facilement Amazon à leurs pipelines de données existants DataZone en utilisant les rôles IAM existants pour publier les actifs de données dans le DataZone catalogue Amazon, en facilitant le partage régi de ces actifs et en améliorant la gouvernance dans l'ensemble de l'infrastructure. Pour plus d'informations, consultez la référence DataZone des types de ressources Amazon.

Amazon DataZone lance des unités de domaine et des politiques d'autorisation

Publié le 08/12/2024

Amazon DataZone introduit un ensemble de nouvelles fonctionnalités de gouvernance des données appelées unités de domaine et politiques d'autorisation qui permettent aux clients de créer une organisation au niveau de l'unité commerciale/de l'équipe et de gérer les politiques en fonction des besoins de leur entreprise. Avec l'ajout d'unités de domaine, les utilisateurs peuvent organiser, créer, rechercher et trouver des actifs de données et des projets associés à des unités commerciales ou à des équipes. Grâce aux politiques d'autorisation, ces utilisateurs d'unités de domaine peuvent définir des politiques d'accès pour créer des projets, des glossaires et utiliser des ressources informatiques au sein d'Amazon DataZone. Pour de plus amples informations, veuillez consulter Unités de domaine et politiques d'autorisation sur Amazon DataZone.

Amazon DataZone lance des produits de données

Publié le 08/05/2024

Amazon DataZone lance des produits de données, qui permettent de regrouper les actifs de données dans des packages autonomes bien définis, adaptés à des cas d'utilisation professionnels

spécifiques. Par exemple, un produit de données d'analyse marketing peut regrouper divers actifs de données, tels que les données de campagnes marketing, les données de pipeline et les données clients. Grâce aux produits de données, les clients peuvent simplifier les processus de découverte et d'abonnement, en les alignant sur les objectifs commerciaux et en réduisant la redondance dans la gestion des actifs individuels. Pour plus d'informations, consultez Produits DataZone de données Amazon.

Amazon DataZone lance une fonctionnalité de contrôle d'accès précise

Publié le 07/02/2024

Amazon DataZone a introduit un contrôle d'accès précis, qui vous permet de contrôler de manière précise vos actifs de données dans le catalogue de données commerciales d'Amazon, à travers les lacs DataZone de données et les entrepôts de données. Grâce à cette nouvelle fonctionnalité, les propriétaires de données peuvent désormais restreindre l'accès à des enregistrements de données spécifiques au niveau des lignes et des colonnes, au lieu d'accorder l'accès à l'ensemble des actifs de données. Par exemple, si vos données contiennent des colonnes contenant des informations sensibles telles que des informations personnelles identifiables (PII), vous pouvez limiter l'accès aux seules colonnes nécessaires, afin de protéger les informations sensibles tout en autorisant l'accès aux données non sensibles. De même, vous pouvez contrôler l'accès au niveau des lignes, en permettant aux utilisateurs de ne voir que les enregistrements correspondant à leur rôle ou à leur tâche. Pour de plus amples informations, consultez Contrôle d'accès précis aux données sur Amazon DataZone.

Amazon DataZone lance une fonctionnalité de traçabilité des données

Publié le 27/06/2024

Amazon DataZone lance le lignage des données en version préliminaire, aidant les clients à visualiser les événements de lignage à partir de systèmes OpenLineage compatibles ou via une API et à suivre le mouvement des données de la source à la consommation. Grâce à la OpenLineage compatibilité DataZone d'Amazon APIs, les administrateurs de domaines et les producteurs de données peuvent capturer et stocker des événements de lignage autres que ceux disponibles sur Amazon DataZone, y compris les transformations dans Amazon S3, AWS Glue et d'autres services. En outre, DataZone les versions d'Amazon font le suivi de chaque événement, ce qui permet aux utilisateurs de visualiser le lignage à tout moment ou de comparer les transformations liées à l'historique d'un actif ou d'une tâche. Cette généalogie historique permet de mieux comprendre l'évolution des données, ce qui est essentiel pour le dépannage, l'audit et la validation de l'intégrité

des actifs de données. Pour de plus amples informations, consultez Lignage des données sur Amazon DataZone.

Amazon DataZone lance des plans AWS de services personnalisés

Publié le 17/06/2024

Avec les plans de AWS service personnalisés, si vous disposez de AWS ressources existantes, notamment des rôles IAM, des lacs de données, des maillages de données, des compartiments Amazon S3 et des clusters Amazon Redshift, vous pouvez désormais spécifier des autorisations pour ces ressources existantes à l'aide de votre propre rôle IAM personnalisé, afin que vos DataZone utilisateurs Amazon puissent tirer parti de la publication et de l'abonnement pour partager et gérer ces ressources. Avec des plans AWS de service personnalisés, DataZone les administrateurs Amazon peuvent configurer des environnements AWS de service en utilisant leurs propres rôles personnalisés. Ils peuvent configurer des liens d'actions pour ces environnements de AWS service et fournir ainsi un accès fédéré à toutes leurs AWS ressources existantes. Ils peuvent également configurer des cibles d'abonnement et des sources de données dans ces environnements AWS de service personnalisés. Les administrateurs peuvent configurer des environnements de AWS service dans leur propre compte de DataZone domaine Amazon ou dans tout compte associé à partir duquel ils souhaitent publier, s'abonner, découvrir ou gérer des données. Pour de plus amples informations, veuillez consulter DataZone Schémas AWS de service personnalisés Amazon.

Améliorations apportées au flux de création de sources de données

Publié le 06/10/2024

Amazon DataZone a apporté des améliorations au flux de création des sources de données afin de simplifier la gestion des accès pour les producteurs de données. Avec ces mises à jour, lorsqu'un producteur de données crée une source de données pour publier ses actifs AWS Glue et Amazon Redshift, Amazon DataZone accorde des autorisations de lecture seule aux membres du projet. Lors de la création d'une source de données AWS Glue, Amazon accorde DataZone automatiquement des autorisations « en lecture seule » au rôle IAM de l'environnement utilisé pour créer la source de données, permettant ainsi l'accès à toutes les tables des bases de données Glue AWS associées. De même, pour les sources de données Amazon Redshift, Amazon DataZone accorde un accès « en lecture seule » à toutes les tables des schémas Amazon Redshift utilisés dans la source de données. Pour plus d'informations, consultez Créez et exécutez une source DataZone de données Amazon pour AWS Glue Data Catalog et Création et gestion d'une source de DataZone données Amazon pour Amazon Redshift.

Amazon DataZone lance l'intégration avec Amazon SageMaker

Publié le 05/06/2024

Amazon DataZone lance l'intégration avec Amazon SageMaker pour aider les producteurs de données et les consommateurs à passer facilement SageMaker à Amazon pour collaborer sur des projets d'apprentissage automatique (ML) tout en renforçant la gouvernance de l'accès aux données et aux actifs de machine learning. Grâce à la nouvelle intégration intégrée entre Amazon DataZone et Amazon SageMaker, les consommateurs et les producteurs de données peuvent rationaliser la gouvernance du machine learning lors de la configuration de l'infrastructure, collaborer sur des initiatives commerciales et gérer facilement les données et les actifs de machine learning. Pour plus d'informations, consultez Blueprints DataZone intégrés à Amazon et Comptes associés sur Amazon DataZone.

Amazon DataZone lance l'intégration avec AWS le mode d'accès hybride Lake Formation

Publié le 04/03/2024

Amazon DataZone a introduit une intégration avec AWS le mode d'accès hybride Lake Formation. Cette intégration vous permet de publier et de partager facilement vos tables AWS Glue via Amazon DataZone, sans avoir à les enregistrer au préalable dans AWS Lake Formation. Pour commencer, les administrateurs activent le paramètre d'enregistrement de la localisation des données dans le `DefaultDataLake` plan de la DataZone console Amazon. Ensuite, lorsqu'un consommateur de données s'abonne à une table AWS Glue gérée via des autorisations IAM, Amazon enregistre d' DataZone abord les emplacements Amazon S3 de cette table en mode hybride, puis accorde l'accès au consommateur de données en gérant les autorisations sur la table via AWS Lake Formation. Cela garantit que les autorisations IAM disponibles continuent d'exister avec les autorisations AWS Lake Formation récemment accordées, sans perturber les flux de travail existants. Pour de plus amples informations, veuillez consulter DataZone Intégration d'Amazon au mode hybride de AWS Lake Formation.

Amazon DataZone lance l'intégration avec AWS Glue Data Quality

Publié le 04/03/2024

Amazon DataZone lance l'intégration avec AWS Glue Data Quality et propose APIs d'intégrer des indicateurs de qualité des données issus de solutions de qualité des données tierces. La nouvelle intégration vous permet de publier automatiquement les scores de AWS Glue Data Quality

dans le catalogue de données Amazon DataZone Business. Amazon DataZone APIs peut être utilisé pour ingérer des indicateurs de qualité provenant de sources tierces. Une fois publiées, les consommateurs de données peuvent facilement rechercher des actifs de données, consulter des indicateurs de qualité détaillés et identifier les contrôles et les règles défaillants, ce qui permet aux entreprises de prendre des décisions éclairées. Pour de plus amples informations, veuillez consulter Qualité des données sur Amazon DataZone.

Mise à disposition générale des recommandations relatives à l'IA pour les descriptions sur Amazon DataZone

Publié le 27/03/2024

Amazon DataZone a annoncé la mise à disposition générale de la nouvelle fonctionnalité générative basée sur l'IA afin d'améliorer la découverte des données, leur compréhension et leur utilisation en enrichissant le catalogue de données commerciales. En un seul clic, les producteurs de données peuvent générer des descriptions et un contexte complets des données commerciales, mettre en évidence les colonnes pertinentes et inclure des recommandations sur des cas d'utilisation analytiques. Le lancement ajoute un support APIs que les producteurs de données peuvent utiliser pour générer par programmation des descriptions pour les actifs. Pour de plus amples informations, veuillez consulter Utilisation de l'apprentissage automatique et de l'IA générative sur Amazon DataZone.

Amazon apporte DataZone des améliorations à l'intégration d'Amazon Redshift

Publié le 21/03/2024

Amazon DataZone a apporté plusieurs améliorations à son intégration avec Amazon Redshift, simplifiant le processus de publication et d'abonnement aux tables et aux vues Amazon Redshift. Ces mises à jour rationalisent l'expérience des producteurs de données et des consommateurs, en leur permettant de créer rapidement des environnements d'entrepôt de données à l'aide d'informations d'identification et de paramètres de connexion préconfigurés fournis par leurs DataZone administrateurs Amazon. En outre, ces améliorations permettent aux administrateurs de mieux contrôler qui peut utiliser les ressources de leurs AWS comptes et des clusters Amazon Redshift, et dans quel but.

- Configuration du plan : une fois que vous avez activé le `DefaultDataWarehouseBlueprint` plan, vous pouvez contrôler quels projets peuvent utiliser le `DefaultDataWarehouseBlueprint`

plan dans votre compte pour créer des profils d'environnement en attribuant la gestion des projets au plan activé. Vous pouvez également créer des ensembles de paramètres en `DefaultDataWarehouseBlueprint` fournissant des paramètres tels que le cluster, la base de données et un AWS secret. Vous pouvez également créer des AWS secrets depuis la DataZone console Amazon.

- Profil d'environnement : lors de la création d'un profil d'environnement, vous pouvez choisir de fournir vos propres paramètres Amazon Redshift ou d'utiliser l'un des ensembles de paramètres de la configuration du plan. Si vous choisissez d'utiliser le jeu de paramètres créé dans la configuration du plan, le AWS secret ne nécessite qu'une `AmazonDataZoneDomain` balise (la `AmazonDataZoneProject` balise n'est requise que si vous choisissez de fournir vos propres ensembles de paramètres dans le profil d'environnement). Dans le profil d'environnement, vous pouvez spécifier une liste de projets autorisés. Seuls les projets autorisés peuvent utiliser ce profil d'environnement pour créer des environnements d'entrepôt de données. Vous pouvez également spécifier les données que les projets autorisés sont autorisés à publier. Actuellement, vous pouvez choisir l'une des options suivantes : 1) Publier à partir de n'importe quel schéma, 2) Publier à partir du schéma d'environnement par défaut, 3) Ne pas autoriser la publication.
- Environnement : les producteurs ou les consommateurs de données peuvent désormais sélectionner un profil d'environnement pour créer des environnements, sans avoir à fournir leurs propres paramètres Amazon Redshift, notamment AWS Secret, cluster, groupe de travail et base de données. Ces paramètres sont transférés vers l'environnement à partir du profil d'environnement. Parallèlement à la création de l'environnement, Amazon crée DataZone désormais également un schéma par défaut pour l'environnement. Les membres du projet ont un accès en lecture et en écriture à ce schéma et peuvent facilement publier les tables créées dans ce schéma dans le catalogue en exécutant la source de données par défaut créée dans le cadre de la création de l'environnement. Les paramètres Amazon Redshift utilisés pour créer un environnement peuvent également être utilisés pour créer de nouvelles sources de données (au lieu que le producteur de données fournisse ses propres paramètres lors de la création de la source de données).

AWS Support à la formation dans le cloud pour Amazon DataZone

Publié le 18/01/2024

Les utilisateurs d'Amazon DataZone peuvent désormais en tirer parti AWS CloudFormation pour modéliser et gérer efficacement une suite de DataZone ressources Amazon. Cette approche facilite le provisionnement cohérent des ressources, tout en permettant la gestion du cycle de vie par le biais

de pratiques liées à l'infrastructure en tant que code. Grâce aux modèles personnalisés, vous pouvez définir avec précision les ressources dont vous avez besoin et leurs interdépendances. Pour plus d'informations, consultez la référence DataZone des types de ressources Amazon.

Ajoutez des responsables IAM directement en tant que membres de projets Amazon DataZone

Publié le 01/05/2024

Vous pouvez désormais ajouter des responsables IAM en tant que membres du projet, même s'ils ne se sont pas encore connectés à Amazon DataZone (exigence précédente). Une fois qu'un administrateur de domaine ou un administrateur informatique a ajouté `iam:GetUser` et `iam:GetRole` au rôle d'exécution du domaine, les propriétaires de projet peuvent ajouter les principaux IAM en tant que membres en fournissant simplement le nom de ressource Amazon (ARN) du rôle IAM ou de l'utilisateur IAM. Le principal IAM doit toujours disposer des autorisations IAM requises pour accéder à Amazon DataZone et celles-ci peuvent être configurées dans la console IAM. Pour de plus amples informations, veuillez consulter Ajouter des membres à un projet.

Support pour les types d'actifs personnalisés depuis le portail de données

Publié le 01/05/2024

La prise en charge des actifs personnalisés permet DataZone à Amazon de cataloguer les actifs via le portail de données pour les données non structurées, notamment les tableaux de bord, les requêtes et les modèles, ce qui vous permet d'ajouter plus facilement des actifs personnalisés directement dans le portail de données, en plus du support d'API précédemment disponible. La possibilité de créer, de mettre à jour et de publier des ressources personnalisées sur Amazon DataZone vous permet de partager, de rechercher, de vous abonner à tout type d'actif et de créer un flux de travail qui assure la gouvernance de ces actifs. Pour de plus amples informations, veuillez consulter Créez des types d'actifs personnalisés sur Amazon DataZone.

2023

Supprimer le domaine

Publié le 27/12/2023

Il s'agit d'une fonctionnalité qui vous permet de supprimer plus facilement vos domaines. Vous pouvez désormais procéder à la suppression du domaine même s'il n'est pas vide (car il contient

des projets, des environnements, des actifs, des sources de données, etc.). Pour de plus amples informations, veuillez consulter Supprimer des DataZone domaines Amazon.

Mode hybride

Publié le 22/12/2023

Amazon DataZone a ajouté la prise en charge du mode hybride AWS Lake Formation. Grâce à ce support, si vous publiez une table AWS Glue sur Amazon DataZone dont l'emplacement AWS S3 est enregistré dans Lake Formation en mode hybride, Amazon DataZone traite cette table comme un actif géré et peut gérer les subventions d'abonnement à cette table. Avant la sortie de cette fonctionnalité, Amazon DataZone traitait ce tableau comme un actif non géré, c'est-à-dire DataZone qu'Amazon ne serait pas en mesure d'accorder des abonnements à ce tableau. Pour de plus amples informations, veuillez consulter Configurer les autorisations de Lake Formation pour Amazon DataZone.

Éligibilité HIPAA

Publié le 14/12/2023

Amazon DataZone est désormais conforme à la loi américaine HIPAA (Health Insurance Portability and Accountability Act) de 1996. Pour consulter la liste des AWS services conformes à la loi HIPAA, voir https://aws.amazon.com/compliance/hipaa-eligible-services-reference/.

Recommandations de l'IA pour les descriptions dans Amazon DataZone (version préliminaire)

Publié le 28/11/2023

AWS annonce la prévisualisation d'une nouvelle fonctionnalité générative basée sur l'IA dans Amazon DataZone afin d'améliorer la découverte des données, leur compréhension et leur utilisation en enrichissant le catalogue de données commerciales. En un seul clic, les producteurs de données peuvent générer des descriptions et un contexte complets des données commerciales, mettre en évidence les colonnes pertinentes et inclure des recommandations sur des cas d'utilisation analytiques. Grâce aux recommandations de l'IA pour les descriptions sur Amazon DataZone, les consommateurs de données peuvent identifier les tables de données et les colonnes nécessaires à l'analyse, ce qui améliore la découvrabilité des données et réduit les back-and-forth communications avec les producteurs de données. La version préliminaire est disponible dans DataZone les domaines Amazon fournis dans les AWS régions suivantes : USA Est (Virginie du Nord), USA

Ouest (Oregon). Pour de plus amples informations, veuillez consulter Utilisation de l'apprentissage automatique et de l'IA générative sur Amazon DataZone.

DefaultDataLake amélioration du plan

Publié le 20/11/2023

Amazon DataZone a ajouté une amélioration au DefaultDataLake plan qui vous permet de mieux contrôler qui peut publier quelles données depuis votre AWS compte. Deux modifications majeures ont été introduites avec le lancement de cette fonctionnalité.

- Dans la console, une fois que vous avez activé le DefaultDataLake plan, vous pouvez contrôler quels projets peuvent utiliser le DefaultDataLake plan dans votre compte pour créer des profils d'environnement en attribuant la gestion des projets au plan activé.
- Le deuxième changement concerne le portail. Si vous créez un profil d'environnement à l'aide du DefaultDataLake plan, vous pouvez également sélectionner les projets autorisés autorisés à utiliser le profil d'environnement pour créer des environnements. Par défaut, tous les projets sont autorisés à utiliser le profil d'environnement du lac de données, mais vous pouvez restreindre le profil d'environnement à des projets spécifiques et également contrôler les données qui peuvent être publiées à l'aide des environnements créés avec le profil.

Pour de plus amples informations, veuillez consulter Création d'un profil d'environnement.

Régions prises en charge par Amazon DataZone

Dans la version actuelle, Amazon DataZone est pris en charge dans les AWS régions suivantes :

- USA Est (Ohio)
- USA Est (Virginie du Nord)
- USA Ouest (Oregon)
- Asie-Pacifique (Mumbai)
- Asie-Pacifique (Séoul)
- Asie-Pacifique (Singapour)
- Asie-Pacifique (Sydney)
- Asie-Pacifique (Tokyo)
- Canada (Centre)
- Europe (Francfort)
- Europe (Irlande)
- Europe (Londres)
- Europe (Paris)
- Europe (Stockholm)
- Amérique du Sud (São Paulo)

Configuration d'Amazon DataZone

Pour configurer Amazon DataZone, vous devez disposer d'un AWS compte et configurer les politiques et autorisations IAM requises pour Amazon DataZone.

Une fois que vous avez configuré vos DataZone autorisations Amazon, il est recommandé de suivre les étapes de la section Getting started qui vous explique comment créer le DataZone domaine Amazon, obtenir l'URL du portail de données et les DataZone flux de travail Amazon de base pour les producteurs et les consommateurs de données.

Rubriques

- Créez un AWS compte
- Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazon
- Configurer les autorisations IAM requises pour utiliser le portail de DataZone données Amazon
- Configuration de l' AWS IAM Identity Center pour Amazon DataZone

Créez un AWS compte

Si vous n'avez pas de AWS compte, procédez comme suit pour en créer un.

Si vous avez une AWS organisation, créez un compte :

1. Connectez-vous à la console de AWS gestion et ouvrez la console Organizations à l'adresse https://console.aws.amazon.com/organizations/.
2. Dans le volet de navigation, sélectionnez AWS Accounts.
3. Choisissez Ajouter un AWS compte.
4. Choisissez Créer un AWS compte et fournissez les informations demandées. Choisissez Créer un AWS compte.

Pour créer un AWS compte

1. Ouvrez l'https://portal.aws.amazon.com/billing/inscription
2. Suivez les instructions en ligne.

   Dans le cadre de la procédure d'inscription, vous recevrez un appel téléphonique et vous saisirez un code de vérification en utilisant le clavier numérique du téléphone.

Lorsque vous créez un AWS compte, un AWS utilisateur root est créé. L'utilisateur root a accès à tous les AWS services et ressources du compte. La meilleure pratique de sécurité consiste à attribuer un accès administratif à un utilisateur administratif, et à utiliser uniquement l'utilisateur racine pour effectuer les tâches nécessitant un accès utilisateur racine.

Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazon

Pour accéder à vos DataZone domaines, plans et utilisateurs Amazon et les configurer, et pour créer le portail de DataZone données Amazon, vous devez utiliser la console de gestion Amazon DataZone .

Vous devez suivre les procédures suivantes afin de configurer les autorisations requises et/ou facultatives pour tout utilisateur, groupe ou rôle souhaitant utiliser la console de DataZone gestion Amazon.

Procédures de configuration des autorisations IAM pour utiliser la console de gestion

- Associer des politiques obligatoires et facultatives à un utilisateur, un groupe ou un rôle pour accéder à DataZone la console Amazon
- Créez une politique personnalisée pour les autorisations IAM afin de permettre à la console de DataZone service Amazon de simplifier la création de rôles
- Créez une politique personnalisée pour les autorisations nécessaires à la gestion d'un compte associé à un DataZone domaine Amazon
- (Facultatif) Créez une politique personnalisée pour les autorisations AWS d'Identity Center afin d'ajouter et de supprimer l'accès des utilisateurs SSO et des groupes SSO aux domaines Amazon DataZone
- (Facultatif) Ajoutez votre principal IAM en tant qu'utilisateur clé pour créer votre DataZone domaine Amazon à l'aide d'une clé gérée par le client via AWS Key Management Service (KMS)

Associer des politiques obligatoires et facultatives à un utilisateur, un groupe ou un rôle pour accéder à DataZone la console Amazon

Procédez comme suit pour associer les politiques personnalisées obligatoires et facultatives à un utilisateur, un groupe ou un rôle. Pour de plus amples informations, veuillez consulter AWS politiques gérées pour Amazon DataZone.

1. Connectez-vous à la console de AWS gestion et ouvrez la console IAM à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le panneau de navigation, choisissez Politiques.
3. Choisissez les politiques suivantes à associer à votre utilisateur, à votre groupe ou à un rôle.
   - Dans la liste des politiques, cochez la case à côté de AmazonDataZoneFullAccess. Vous pouvez utiliser le menu Filtre et la zone de recherche pour filtrer la liste de politiques. Pour de plus amples informations, veuillez consulter AWS politique gérée : AmazonDataZoneFullAccess.
   - (Facultatif) Créez une politique personnalisée pour les autorisations IAM afin de permettre à la console de DataZone service Amazon de simplifier la création de rôles.
   - (Facultatif) Créez une politique personnalisée pour les autorisations AWS Identity Center afin d'ajouter et de supprimer l'accès des utilisateurs SSO et des groupes SSO à votre domaine Amazon DataZone .
4. Sélectionnez Actions, puis Attach (Attacher).
5. Choisissez l'utilisateur, le groupe ou le rôle auquel vous souhaitez associer la politique. Vous pouvez utiliser le menu Filtre et la zone de recherche pour filtrer la liste des entités du principal. Après avoir choisi l'utilisateur, le groupe ou le rôle, choisissez Attacher une politique.

Créez une politique personnalisée pour les autorisations IAM afin de permettre à la console de DataZone service Amazon de simplifier la création de rôles

Suivez la procédure suivante pour créer une politique en ligne personnalisée afin de disposer des autorisations nécessaires pour permettre DataZone à Amazon de créer les rôles nécessaires dans la console AWS de gestion en votre nom.

> **Note**
>
> Pour obtenir des informations sur les meilleures pratiques relatives à la configuration des autorisations permettant la création de rôles de service, consultez le https://docs.aws.amazon.com/IAM/latest/UserGuide/idfichier _roles_create_for-service.html.

1. Connectez-vous à la console de AWS gestion et ouvrez la console IAM à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le volet de navigation, choisissez Utilisateurs ou Groupes d'utilisateurs.
3. Dans la liste, sélectionnez le nom de l'utilisateur ou du groupe auquel intégrer une politique.
4. Sélectionnez l'onglet Autorisations et, si nécessaire, développez la section Politiques d'autorisations.
5. Choisissez Ajouter des autorisations et Créer un lien de politique intégré.
6. Sur l'écran Create Policy, dans la section Policy editor, sélectionnez JSON.

   Créez un document de politique avec les instructions JSON suivantes, puis choisissez Next.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "iam:CreatePolicy",
                "iam:CreateRole"
            ],
            "Resource": [
                "arn:aws:iam::*:policy/service-role/AmazonDataZone*",
                "arn:aws:iam::*:role/service-role/AmazonDataZone*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": "iam:AttachRolePolicy",
            "Resource": "arn:aws:iam::*:role/service-role/AmazonDataZone*",
```

```
            "Condition": {
                "ArnLike": {
                    "iam:PolicyARN": [
                        "arn:aws:iam::aws:policy/AmazonDataZone*",
                        "arn:aws:iam::*:policy/service-role/AmazonDataZone*"
                    ]
                }
            }
        }
    ]
}
```

7. Sur l'écran Révision de la politique, entrez le nom de la politique. Lorsque vous êtes satisfait de la politique, sélectionnez Create policy (Créer une politique). Assurez-vous qu'aucune erreur ne s'affiche dans un cadre rouge en haut de l'écran. Corrigez les erreurs signalées.

Créez une politique personnalisée pour les autorisations nécessaires à la gestion d'un compte associé à un DataZone domaine Amazon

Suivez la procédure ci-dessous pour créer une politique en ligne personnalisée afin de disposer des autorisations nécessaires dans un AWS compte associé pour répertorier, accepter et rejeter les partages de ressources d'un domaine, puis activer, configurer et désactiver les plans d'environnement dans le compte associé. Pour activer la création de rôles simplifiée optionnelle sur la console de DataZone service Amazon disponible lors de la configuration du Blueprint, vous devez également[Créez une politique personnalisée pour les autorisations IAM afin de permettre à la console de DataZone service Amazon de simplifier la création de rôles](#).

> **Note**
>
> Pour obtenir des informations sur les meilleures pratiques relatives à la configuration des autorisations permettant la création de rôles de service, consultez le https://docs.aws.amazon.com/IAM/latest/UserGuide/idfichier _roles_create_for-service.html.

1. Connectez-vous à la console de AWS gestion et ouvrez la console IAM à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le volet de navigation, choisissez Utilisateurs ou Groupes d'utilisateurs.

3. Dans la liste, sélectionnez le nom de l'utilisateur ou du groupe auquel intégrer une politique.
4. Sélectionnez l'onglet Autorisations et, si nécessaire, développez la section Politiques d'autorisations.
5. Choisissez Ajouter des autorisations et Créer un lien de politique intégré.
6. Sur l'écran Create Policy, dans la section Policy editor, sélectionnez JSON. Créez un document de politique avec les instructions JSON suivantes, puis choisissez Next.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "datazone:ListEnvironmentBlueprintConfigurations",
                "datazone:PutEnvironmentBlueprintConfiguration",
                "datazone:GetDomain",
                "datazone:ListDomains",
                "datazone:GetEnvironmentBlueprintConfiguration",
                "datazone:ListEnvironmentBlueprints",
                "datazone:GetEnvironmentBlueprint",
                "datazone:ListAccountEnvironments",
                "datazone:DeleteEnvironmentBlueprintConfiguration"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": [
                "arn:aws:iam::*:role/AmazonDataZone",
                "arn:aws:iam::*:role/service-role/AmazonDataZone*"
            ],
            "Condition": {
                "StringEquals": {
                    "iam:passedToService": "datazone.amazonaws.com"
                }
            }
        },
```

```
        {
            "Effect": "Allow",
            "Action": "iam:AttachRolePolicy",
            "Resource": "arn:aws:iam::*:role/service-role/AmazonDataZone*",
            "Condition": {
                "ArnLike": {
                    "iam:PolicyARN": [
                        "arn:aws:iam::aws:policy/AmazonDataZone*",
                        "arn:aws:iam::*:policy/service-role/AmazonDataZone*"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": "iam:ListRoles",
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "iam:CreatePolicy",
                "iam:CreateRole"
            ],
            "Resource": [
                "arn:aws:iam::*:policy/service-role/AmazonDataZone*",
                "arn:aws:iam::*:role/service-role/AmazonDataZone*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "ram:AcceptResourceShareInvitation",
                "ram:RejectResourceShareInvitation",
                "ram:GetResourceShareInvitations"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:ListAllMyBuckets",
                "s3:ListBucket",
                "s3:GetBucketLocation"
```

```
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "s3:CreateBucket",
            "Resource": "arn:aws:s3:::amazon-datazone*"
        }
    ]
}
```

7. Sur l'écran Révision de la politique, entrez le nom de la politique. Lorsque vous êtes satisfait de la politique, sélectionnez Create policy (Créer une politique). Assurez-vous qu'aucune erreur ne s'affiche dans un cadre rouge en haut de l'écran. Corrigez les erreurs signalées.

(Facultatif) Créez une politique personnalisée pour les autorisations AWS d'Identity Center afin d'ajouter et de supprimer l'accès des utilisateurs SSO et des groupes SSO aux domaines Amazon DataZone

Suivez la procédure ci-dessous pour créer une politique en ligne personnalisée afin de disposer des autorisations nécessaires pour ajouter et supprimer l'accès d'un utilisateur SSO et d'un groupe SSO à votre domaine Amazon. DataZone

1. Connectez-vous à la console de AWS gestion et ouvrez la console IAM à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le volet de navigation, choisissez Utilisateurs ou Groupes d'utilisateurs.
3. Dans la liste, sélectionnez le nom de l'utilisateur ou du groupe auquel intégrer une politique.
4. Sélectionnez l'onglet Autorisations et, si nécessaire, développez la section Politiques d'autorisations.
5. Choisissez Ajouter des autorisations et Créer une politique en ligne.
6. Sur l'écran Create Policy, dans la section Policy editor, sélectionnez JSON.

   Créez un document de politique avec les instructions JSON suivantes, puis choisissez Next.

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sso:GetManagedApplicationInstance",
        "sso:ListProfiles",
        "sso:AssociateProfile",
        "sso:DisassociateProfile",
        "sso:GetProfile"
      ],
      "Resource": "*"
    }
  ]
}
```

7. Sur l'écran Révision de la politique, entrez le nom de la politique. Lorsque vous êtes satisfait de la politique, sélectionnez Create policy (Créer une politique). Assurez-vous qu'aucune erreur ne s'affiche dans un cadre rouge en haut de l'écran. Corrigez les erreurs signalées.

(Facultatif) Ajoutez votre principal IAM en tant qu'utilisateur clé pour créer votre DataZone domaine Amazon à l'aide d'une clé gérée par le client via AWS Key Management Service (KMS)

Avant de pouvoir éventuellement créer votre DataZone domaine Amazon avec une clé gérée par le client (CMK) à partir du service de gestion des AWS clés (KMS), suivez la procédure suivante pour faire de votre principal IAM un utilisateur de votre clé KMS.

1. Connectez-vous à la console de AWS gestion et ouvrez la console KMS à l'adresse https://console.aws.amazon.com/kms/.
2. Pour afficher les clés de votre compte que vous créez et gérez vous-même, dans le volet de navigation, choisissez Clés gérées par le client.

3. Dans la liste des clés KMS, choisissez l'alias ou l'ID de clé de la clé KMS que vous souhaitez examiner.
4. Pour ajouter ou supprimer des utilisateurs clés, et pour autoriser ou interdire à AWS des comptes externes d'utiliser la clé KMS, utilisez les commandes de la section Utilisateurs clés de la page. Les utilisateurs de clé peuvent utiliser la clé KMS dans les opérations de chiffrement, telles que le chiffrement, le déchiffrement, le rechiffrement et la génération de clés de données.

Configurer les autorisations IAM requises pour utiliser le portail de DataZone données Amazon

Le portail de DataZone données Amazon (en dehors de la console de AWS gestion) est une application Web basée sur un navigateur dans laquelle les utilisateurs peuvent cataloguer, découvrir, gérer, partager et analyser des données en libre-service. Le portail de données authentifie les utilisateurs à l'aide d'informations d'identification IAM ou d'informations d'identification existantes auprès de votre fournisseur d'identité via AWS IAM Identity Center.

Vous devez suivre les procédures suivantes afin de configurer les autorisations requises pour tout utilisateur, groupe ou rôle souhaitant utiliser le portail de DataZone données ou le catalogue Amazon :

Procédures de configuration des autorisations IAM pour l'utilisation du portail de données

- Associer la politique requise à un utilisateur, un groupe ou un rôle pour accéder au portail de DataZone données Amazon
- Associer la politique requise à un utilisateur, un groupe ou un rôle pour accéder au DataZone catalogue Amazon
- Associez une politique facultative à un utilisateur, un groupe ou un rôle pour accéder au portail de DataZone données ou au catalogue Amazon si votre domaine est chiffré à l'aide d'une clé gérée par le client par le service de gestion des AWS clés (KMS)

Associer la politique requise à un utilisateur, un groupe ou un rôle pour accéder au portail de DataZone données Amazon

Vous pouvez accéder au portail de DataZone données Amazon en utilisant vos AWS informations d'identification ou vos informations d'identification unique (SSO). Suivez les instructions de la section ci-dessous pour configurer les autorisations requises pour accéder au portail de données avec vos

AWS informations d'identification. Pour plus d'informations sur l'utilisation d'Amazon DataZone avec SSO, consultezConfiguration de l' AWS IAM Identity Center pour Amazon DataZone.

> **Note**
>
> Seuls les principaux IAM du AWS compte de votre domaine peuvent accéder au portail de données du domaine. Les principaux IAM d'autres AWS comptes ne peuvent pas accéder au portail de données du domaine.

Procédez comme suit pour associer la politique requise à un utilisateur, un groupe ou un rôle. Pour de plus amples informations, veuillez consulter AWS politiques gérées pour Amazon DataZone.

1. Connectez-vous à la console de AWS gestion et ouvrez la console IAM à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le volet de navigation, sélectionnez Utilisateurs, Groupes d'utilisateurs ou Rôles.
3. Dans la liste, choisissez le nom de l'utilisateur, du groupe ou du rôle dans lequel vous souhaitez intégrer une politique.
4. Sélectionnez l'onglet Autorisations et, si nécessaire, développez la section Politiques d'autorisations.
5. Choisissez Ajouter des autorisations et Créer un lien de politique intégré.
6. Sur l'écran Create Policy, dans la section Policy editor, sélectionnez JSON. Créez un document de politique avec les instructions JSON suivantes, puis choisissez Next.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "datazone:GetIamPortalLoginUrl"
            ],
            "Resource": [
                "*"
            ]
```

```
        }
    ]
}
```

7. Sur l'écran Révision de la politique, entrez le nom de la politique. Lorsque vous êtes satisfait de la politique, sélectionnez Create policy (Créer une politique). Assurez-vous qu'aucune erreur ne s'affiche dans un cadre rouge en haut de l'écran. Corrigez les erreurs signalées.

Associer la politique requise à un utilisateur, un groupe ou un rôle pour accéder au DataZone catalogue Amazon

> **Note**
>
> Seuls les principaux IAM du AWS compte de votre domaine peuvent accéder au catalogue du domaine. Les principaux IAM d'autres AWS comptes ne peuvent pas accéder au catalogue du domaine.

Vous pouvez accorder à vos identités IAM l'accès au catalogue de votre DataZone domaine Amazon via l'API et le SDK en suivant la procédure suivante. Si vous souhaitez que ces identités IAM aient également accès au portail de DataZone données Amazon, suivez également la procédure ci-dessus pourAssocier la politique requise à un utilisateur, un groupe ou un rôle pour accéder au portail de DataZone données Amazon. Pour de plus amples informations, veuillez consulter AWS politiques gérées pour Amazon DataZone.

1. Connectez-vous à la console de AWS gestion et ouvrez la console IAM à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le panneau de navigation, choisissez Politiques.
3. Dans la liste des politiques, sélectionnez le bouton radio situé à côté de la AmazonDataZoneFullUserAccesspolitique. Vous pouvez utiliser le menu Filtre et la zone de recherche pour filtrer la liste de politiques. Pour de plus amples informations, consultez AWS politique gérée : AmazonDataZoneFullUserAccess.
4. Sélectionnez Actions, puis Attach (Attacher).
5. Choisissez l'utilisateur, le groupe ou le rôle auquel vous souhaitez associer la politique en cochant la case à côté de chaque principal. Vous pouvez utiliser le menu Filtre et la zone de

recherche pour filtrer la liste des entités du principal. Après avoir choisi l'utilisateur, le groupe ou le rôle, choisissez Attacher une politique.

Associez une politique facultative à un utilisateur, un groupe ou un rôle pour accéder au portail de DataZone données ou au catalogue Amazon si votre domaine est chiffré à l'aide d'une clé gérée par le client par le service de gestion des AWS clés (KMS)

Si vous créez votre DataZone domaine Amazon avec votre propre clé KMS pour le chiffrement des données, vous devez également créer une politique intégrée avec les autorisations suivantes et l'associer à vos principaux IAM afin qu'ils puissent accéder au portail de DataZone données ou au catalogue Amazon.

1. Connectez-vous à la console de AWS gestion et ouvrez la console IAM à https://console.aws.amazon.com/iam/l'adresse.
2. Dans le volet de navigation, sélectionnez Utilisateurs, Groupes d'utilisateurs ou Rôles.
3. Dans la liste, choisissez le nom de l'utilisateur, du groupe ou du rôle dans lequel vous souhaitez intégrer une politique.
4. Sélectionnez l'onglet Autorisations et, si nécessaire, développez la section Politiques d'autorisations.
5. Choisissez Ajouter des autorisations et Créer un lien de politique intégré.
6. Sur l'écran Create Policy, dans la section Policy editor, sélectionnez JSON. Créez un document de politique avec les instructions JSON suivantes, puis choisissez Next.

   JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "Statement1",
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt",
                "kms:DescribeKey",
                "kms:GenerateDataKey"
```

```
            ],
            "Resource": [
                "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
            ]
        }
    ]
}
```

7. Sur l'écran Révision de la politique, entrez le nom de la politique. Lorsque vous êtes satisfait de la politique, sélectionnez Create policy (Créer une politique). Assurez-vous qu'aucune erreur ne s'affiche dans un cadre rouge en haut de l'écran. Corrigez les erreurs signalées.

Configuration de l' AWS IAM Identity Center pour Amazon DataZone

> **Note**
>
> AWS Identity Center doit être activé dans la même AWS région que votre DataZone domaine Amazon. À l'heure actuelle, AWS Identity Center ne peut être activé que dans une seule AWS région.

Vous pouvez accéder au portail de DataZone données Amazon en utilisant vos informations d'identification unique (SSO) ou AWS vos informations d'identification. Suivez les instructions de cette section pour configurer AWS IAM Identity Center pour Amazon DataZone. Pour plus d'informations sur l'utilisation d'Amazon DataZone avec vos AWS informations d'identification, consultezConfigurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazon.

Vous pouvez ignorer les procédures décrites dans cette section si AWS IAM Identity Center (successeur de Single Sign-On) est déjà activé et configuré dans la AWS région où vous souhaitez créer votre domaine Amazon DataZone . AWS

Procédez comme suit pour activer AWS IAM Identity Center (successeur de Single Sign-On). AWS

1. Pour activer AWS IAM Identity Center, vous devez vous connecter à la console de AWS gestion à l'aide des informations d'identification de votre compte de gestion AWS Organizations. Vous ne pouvez pas activer IAM Identity Center lorsque vous êtes connecté avec les informations d'identification d'un compte membre d' AWS Organizations. Pour plus d'informations, consultez

la section Création et gestion d'une organisation dans le Guide de l'utilisateur AWS des Organizations.

2. Ouvrez la console AWS IAM Identity Center (qui a succédé à AWS Single Sign-On) et utilisez le sélecteur de région dans la barre de navigation supérieure pour choisir la AWS région dans laquelle vous souhaitez créer votre domaine Amazon. DataZone
3. Sélectionnez Activer.
4. Choisissez votre source d'identité.

   Par défaut, vous disposez d'un magasin IAM Identity Center pour une gestion rapide et facile des utilisateurs. Vous pouvez éventuellement connecter un fournisseur d'identité externe à la place. Dans cette procédure, nous utilisons le magasin IAM Identity Center par défaut.

   Pour plus d'informations, voir Choisir votre source d'identité.
5. Dans le volet de navigation d'IAM Identity Center, choisissez Groups, puis Create group. Entrez le nom du groupe et choisissez Create.
6. Dans le volet de navigation d'IAM Identity Center, sélectionnez Users.
7. Sur l'écran Ajouter un utilisateur, entrez les informations requises et choisissez Envoyer un e-mail à l'utilisateur avec les instructions de configuration du mot de passe. L'utilisateur doit recevoir un e-mail concernant les prochaines étapes de configuration.
8. Choisissez Suivant : Groupes, choisissez le groupe de votre choix, puis choisissez Ajouter un utilisateur. Les utilisateurs devraient recevoir un e-mail les invitant à utiliser le SSO. Dans cet e-mail, ils doivent choisir Accepter l'invitation et définir le mot de passe.

Après avoir créé votre DataZone domaine Amazon, vous pouvez activer AWS Identity Center pour Amazon DataZone et donner accès à vos utilisateurs et groupes SSO. Pour de plus amples informations, veuillez consulter Activer IAM Identity Center pour Amazon DataZone.

Commencer à utiliser Amazon DataZone

Les informations contenues dans cette section vous aideront à commencer à utiliser Amazon DataZone. Si vous utilisez Amazon pour la première DataZone fois, commencez par vous familiariser avec les concepts et la terminologie présentés dansDataZone Terminologie et concepts d'Amazon.

Avant de commencer les étapes de l'un ou l'autre de ces flux de travail de démarrage rapide, vous devez suivre les procédures décrites dans la section Configuration de ce guide. Si vous utilisez un tout nouveau AWS compte, vous devez configurer les autorisations requises pour utiliser la console DataZone de gestion Amazon. Si vous utilisez un AWS compte qui possède des objets AWS Glue Data Catalog existants, vous devez également configurer les autorisations Lake Formation pour Amazon DataZone.

Cette section de mise en route vous présente les flux de travail Amazon DataZone Quickstart suivants :

Rubriques

- Amazon DataZone Quickstart avec les données AWS Glue
- Amazon DataZone QuickStart avec les données Amazon Redshift
- Amazon DataZone QuickStart avec des exemples de scripts

Amazon DataZone Quickstart avec les données AWS Glue

Suivez les étapes de démarrage rapide suivantes pour passer en revue l'ensemble des flux de production et de consommation de données sur Amazon DataZone avec des exemples de données AWS Glue.

Étapes de démarrage rapide

- Étape 1 - Création du DataZone domaine Amazon et du portail de données
- Étape 2 - Création du projet de publication
- Étape 3 - Création de l'environnement
- Étape 4 - Produire des données pour publication
- Étape 5 - Collectez les métadonnées à partir de AWS Glue
- Étape 6 - Organiser et publier la ressource de données

- Étape 7 - Création du projet pour l'analyse des données
- Étape 8 - Création d'un environnement pour l'analyse des données
- Étape 9 - Rechercher dans le catalogue de données et s'abonner aux données
- Étape 10 - Approuver la demande d'abonnement
- Étape 11 - Création d'une requête et analyse des données dans Amazon Athena

Étape 1 - Création du DataZone domaine Amazon et du portail de données

Cette section décrit les étapes de création d'un DataZone domaine Amazon et d'un portail de données pour ce flux de travail.

Suivez la procédure ci-dessous pour créer un DataZone domaine Amazon. Pour plus d'informations sur DataZone les domaines Amazon, consultezDataZone Terminologie et concepts d'Amazon.

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone, connectez-vous, puis choisissez Create domain.

   > **Note**
   >
   > Si vous souhaitez utiliser un DataZone domaine Amazon existant pour ce flux de travail, choisissez Afficher les domaines, puis choisissez le domaine que vous souhaitez utiliser, puis passez à l'étape 2 de création d'un projet de publication.

2. Sur la page Créer un domaine, entrez des valeurs pour les champs suivants :

   - Nom : spécifiez le nom de votre domaine. Dans le cadre de ce flux de travail, vous pouvez appeler ce domaine Marketing.
   - Description : spécifiez une description de domaine facultative.
   - Chiffrement des données : vos données sont cryptées par défaut à l'aide d'une clé qui vous AWS appartient et qui est gérée pour vous. Dans ce cas d'utilisation, vous pouvez conserver les paramètres de chiffrement des données par défaut.

     Pour plus d'informations sur l'utilisation des clés gérées par le client, consultezLe chiffrement des données est au repos pour Amazon DataZone. Si vous utilisez votre propre clé KMS pour le chiffrement des données, vous devez inclure l'instruction suivante dans votre clé par défautAmazonDataZoneDomainExecutionRole.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "Statement1",
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt",
                "kms:DescribeKey",
                "kms:GenerateDataKey"
            ],
            "Resource": [
                "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
            ]
        }
    ]
}
```

- Accès au service : laissez inchangée l'option Utiliser un rôle par défaut sélectionnée par défaut.

> **Note**
>
> Si vous utilisez un DataZone domaine Amazon existant pour ce flux de travail, vous pouvez choisir l'option Utiliser un rôle de service existant, puis choisir un rôle existant dans le menu déroulant.

- Sous Configuration rapide, choisissez Configurer ce compte pour la consommation et la publication de données. Cette option active les DataZone plans Amazon intégrés du lac de données et de l'entrepôt de données, et configure les autorisations, les ressources, un projet par défaut et les profils d'environnement de lac de données et d'entrepôt de données par défaut pour ce compte. Pour plus d'informations sur les DataZone plans Amazon, consultezDataZone Terminologie et concepts d'Amazon.
- Conservez les champs restants sous Détails des autorisations inchangés.

> **Note**
>
> Si vous possédez déjà un DataZone domaine Amazon, vous pouvez choisir l'option Utiliser un rôle de service existant, puis choisir un rôle existant dans le menu déroulant pour le rôle Glue Manage Access, le rôle Redshift Manage Access et le rôle Provisioning.

- Ne modifiez pas les champs situés sous Tags.
- Choisissez Create domain (Créer un domaine).

3. Une fois le domaine créé avec succès, choisissez-le et, sur la page de résumé du domaine, notez l'URL du portail de données pour ce domaine. Vous pouvez utiliser cette URL pour accéder à votre portail de DataZone données Amazon afin d'effectuer les autres étapes de ce flux de travail. Vous pouvez également accéder au portail de données en choisissant Portail de données ouvertes.

> **Note**
>
> Dans la version actuelle d'Amazon DataZone, une fois le domaine créé, l'URL générée pour le portail de données ne peut pas être modifiée.

La création d'un domaine peut prendre plusieurs minutes. Attendez que le statut du domaine soit défini sur Disponible avant de passer à l'étape suivante.

Étape 2 - Création du projet de publication

Cette section décrit les étapes requises pour créer le projet de publication pour ce flux de travail.

1. Une fois que vous aurez terminé l'étape 1 ci-dessus et créé un domaine, vous verrez le message Welcome to Amazon DataZone ! fenêtre. Dans cette fenêtre, choisissez Créer un projet.
2. Spécifiez le nom du projet, par exemple, pour ce flux de travail SalesDataPublishingProject, vous pouvez le nommer, puis laisser les autres champs inchangés, puis choisir Créer.

Étape 3 - Création de l'environnement

Cette section décrit les étapes requises pour créer un environnement pour ce flux de travail.

1. Une fois que vous avez terminé l'étape 2 ci-dessus et créé votre projet, vous verrez la fenêtre Votre projet est prêt à être utilisé. Dans cette fenêtre, choisissez Create environment.
2. Sur la page Créer un environnement, spécifiez ce qui suit, puis choisissez Créer un environnement.
3. Spécifiez les valeurs suivantes :
   - Nom : spécifiez le nom de l'environnement. Pour cette procédure pas à pas, vous pouvez l'`Default data lake environment`appeler.
   - Description : spécifiez une description de l'environnement.
   - Profil d'environnement : choisissez le profil d'DataLakeProfileenvironnement. Cela vous permet d'utiliser Amazon DataZone dans ce flux de travail pour travailler avec des données dans Amazon S3, AWS Glue Catalog et Amazon Athena.
   - Pour cette procédure pas à pas, conservez les autres champs inchangés.
4. Choisissez Create environment.

Étape 4 - Produire des données pour publication

Cette section décrit les étapes requises pour produire des données destinées à être publiées dans ce flux de travail.

1. Une fois que vous avez terminé l'étape 3 ci-dessus, dans votre `SalesDataPublishingProject` projet, dans le panneau de droite, sous Outils d'analyse, sélectionnez Amazon Athena. Cela ouvre l'éditeur de requêtes Athena en utilisant les informations d'identification de votre projet pour l'authentification. Assurez-vous que votre environnement de publication est sélectionné dans le menu déroulant de DataZone l'environnement Amazon et que la `<environment_name>%_pub_db` base de données est sélectionnée comme dans l'éditeur de requêtes.
2. Pour cette procédure pas à pas, vous utilisez le script de requête Create Table as Select (CTAS) pour créer une nouvelle table que vous souhaitez publier sur Amazon. DataZone Dans votre éditeur de requêtes, exécutez ce script CTAS pour créer une `mkt_sls_table` table que vous pouvez publier et rendre disponible pour la recherche et l'abonnement.

```
CREATE TABLE mkt_sls_table AS
```

```
SELECT 146776932 AS ord_num, 23 AS sales_qty_sld, 23.4 AS wholesale_cost, 45.0 as
 lst_pr, 43.0 as sell_pr, 2.0 as disnt, 12 as ship_mode,13 as warehouse_id, 23 as
 item_id, 34 as ctlg_page, 232 as ship_cust_id, 4556 as bill_cust_id
UNION ALL SELECT 46776931, 24, 24.4, 46, 44, 1, 14, 15, 24, 35, 222, 4551
UNION ALL SELECT 46777394, 42, 43.4, 60, 50, 10, 30, 20, 27, 43, 241, 4565
UNION ALL SELECT 46777831, 33, 40.4, 51, 46, 15, 16, 26, 33, 40, 234, 4563
UNION ALL SELECT 46779160, 29, 26.4, 50, 61, 8, 31, 15, 36, 40, 242, 4562
UNION ALL SELECT 46778595, 43, 28.4, 49, 47, 7, 28, 22, 27, 43, 224, 4555
UNION ALL SELECT 46779482, 34, 33.4, 64, 44, 10, 17, 27, 43, 52, 222, 4556
UNION ALL SELECT 46779650, 39, 37.4, 51, 62, 13, 31, 25, 31, 52, 224, 4551
UNION ALL SELECT 46780524, 33, 40.4, 60, 53, 18, 32, 31, 31, 39, 232, 4563
UNION ALL SELECT 46780634, 39, 35.4, 46, 44, 16, 33, 19, 31, 52, 242, 4557
UNION ALL SELECT 46781887, 24, 30.4, 54, 62, 13, 18, 29, 24, 52, 223, 4561
```

Assurez-vous que la table mkt_sls_table est correctement créée dans la section Tables et vues sur le côté gauche. Vous disposez désormais d'un actif de données qui peut être publié dans le DataZone catalogue Amazon.

Étape 5 - Collectez les métadonnées à partir de AWS Glue

Cette section décrit l'étape de collecte des métadonnées à partir de AWS Glue pour ce flux de travail.

1. Une fois que vous avez terminé l'étape 4 ci-dessus, dans le portail de DataZone données Amazon, choisissez le `SalesDataPublishingProject` projet, puis choisissez l'onglet Données, puis choisissez Sources de données dans le panneau de gauche.
2. Choisissez la source créée dans le cadre du processus de création de l'environnement.
3. Choisissez Exécuter à côté du menu déroulant Action, puis cliquez sur le bouton d'actualisation. Une fois l'exécution de la source de données terminée, les actifs sont ajoutés à l' DataZone inventaire Amazon.

Étape 6 - Organiser et publier la ressource de données

Cette section décrit les étapes de conservation et de publication de la ressource de données dans ce flux de travail.

1. Une fois que vous avez terminé l'étape 5 ci-dessus, dans le portail de DataZone données Amazon, choisissez le `SalesDataPublishingProject` projet que vous avez créé à l'étape

précédente, choisissez l'onglet Données, choisissez Données d'inventaire dans le panneau de gauche et recherchez le `mkt_sls_table` tableau.

2. Ouvrez la page de détails de l'`mkt_sls_table`actif pour voir les noms commerciaux générés automatiquement. Cliquez sur l'icône de métadonnées générées automatiquement pour afficher les noms générés automatiquement pour les actifs et les colonnes. Vous pouvez accepter ou rejeter chaque nom individuellement ou choisir Accepter tout pour appliquer les noms générés. Vous pouvez également éventuellement ajouter le formulaire de métadonnées disponible à votre ressource et sélectionner des termes du glossaire pour classer vos données.
3. Choisissez Publier la ressource pour publier la `mkt_sls_table` ressource.

Étape 7 - Création du projet pour l'analyse des données

Cette section décrit les étapes de création du projet pour l'analyse des données. C'est le début des étapes de consommation de données de ce flux de travail.

1. Une fois que vous avez terminé l'étape 6 ci-dessus, dans le portail de DataZone données Amazon, choisissez Create project dans le menu déroulant Project.
2. Sur la page Créer un projet, spécifiez le nom du projet. Par exemple, pour ce flux de travail MarketingDataAnalysisProject, vous pouvez le nommer, laisser les autres champs inchangés, puis choisir Créer.

Étape 8 - Création d'un environnement pour l'analyse des données

Cette section décrit les étapes de création d'un environnement pour l'analyse des données.

1. Une fois que vous avez terminé l'étape 7 ci-dessus, dans le portail de DataZone données Amazon, choisissez le `MarketingDataAnalysisProject` projet, puis choisissez l'onglet Environnements, puis choisissez Create environment.
2. Sur la page Créer un environnement, spécifiez ce qui suit, puis choisissez Créer un environnement.

   - Nom : spécifiez le nom de l'environnement. Pour cette procédure pas à pas, vous pouvez l'`Default data lake environment`appeler.
   - Description : spécifiez une description de l'environnement.
   - Profil d'environnement : choisissez le profil d'DataLakeProfileenvironnement intégré.
   - Pour cette procédure pas à pas, conservez les autres champs inchangés.

Étape 9 - Rechercher dans le catalogue de données et s'abonner aux données

Cette section décrit les étapes de recherche dans le catalogue de données et d'abonnement aux données.

1. Une fois que vous avez terminé l'étape 8 ci-dessus, dans le portail de DataZone données Amazon, choisissez l' DataZoneicône Amazon, et dans le champ de DataZone recherche Amazon, recherchez des actifs de données à l'aide de mots clés (par exemple, « catalogue » ou « ventes ») dans la barre de recherche du portail de données.

   Si nécessaire, appliquez des filtres ou effectuez un tri. Une fois que vous avez localisé la ressource Product Sales Data, vous pouvez la choisir pour ouvrir la page de détails de la ressource.
2. Sur la page de détails de la ressource Catalog Sales Data, choisissez Subscribe.
3. Dans la boîte de dialogue S'abonner, choisissez votre projet MarketingDataAnalysisProjectclient dans la liste déroulante, puis spécifiez le motif de votre demande d'abonnement, puis choisissez S'abonner.

Étape 10 - Approuver la demande d'abonnement

Cette section décrit les étapes d'approbation de la demande d'abonnement.

1. Une fois que vous avez terminé l'étape 9 ci-dessus, dans le portail de DataZone données Amazon, choisissez le SalesDataPublishingProjectprojet avec lequel vous avez publié votre ressource.
2. Cliquez sur l'onglet Données, puis sur Données publiées, puis sur Demandes entrantes.
3. Vous pouvez maintenant voir la ligne correspondant à la nouvelle demande nécessitant une approbation. Choisissez Afficher la demande. Indiquez le motif de l'approbation et choisissez Approuver.

Étape 11 - Création d'une requête et analyse des données dans Amazon Athena

Maintenant que vous avez publié avec succès une ressource dans le DataZone catalogue Amazon et que vous vous y êtes abonné, vous pouvez l'analyser.

1. Sur le portail de DataZone données Amazon, choisissez votre projet MarketingDataAnalysisProjectclient, puis, dans le panneau de droite, sous Outils d'analyse, choisissez le lien de données Query avec Amazon Athena. Cela ouvre l'éditeur de requêtes Amazon Athena en utilisant les informations d'identification de votre projet pour l'authentification. Choisissez l'environnement du MarketingDataAnalysisProjectconsommateur dans le menu déroulant Amazon DataZone Environment de l'éditeur de requêtes, puis choisissez celui de votre projet dans le menu déroulant `<environment_name>%sub_db` de la base de données.
2. Vous pouvez désormais exécuter des requêtes sur la table abonnée. Vous pouvez choisir le tableau dans Tables et vues, puis choisir Aperçu pour afficher l'instruction de sélection sur l'écran de l'éditeur. Exécutez la requête pour voir les résultats.

Amazon DataZone QuickStart avec les données Amazon Redshift

Suivez les étapes de démarrage rapide suivantes pour passer en revue l'ensemble des flux de production et de consommation de données d'Amazon DataZone avec des exemples de données Amazon Redshift.

Étapes de démarrage rapide

- Étape 1 - Création du DataZone domaine Amazon et du portail de données
- Étape 2 - Création du projet de publication
- Étape 3 - Création de l'environnement
- Étape 4 - Produire des données pour publication
- Étape 5 - Collectez les métadonnées depuis Amazon Redshift
- Étape 6 - Organiser et publier la ressource de données
- Étape 7 - Création du projet pour l'analyse des données
- Étape 8 - Création d'un environnement pour l'analyse des données
- Étape 9 - Rechercher dans le catalogue de données et s'abonner aux données
- Étape 10 - Approuver la demande d'abonnement
- Étape 11 - Création d'une requête et analyse des données dans Amazon Redshift

Étape 1 - Création du DataZone domaine Amazon et du portail de données

Suivez la procédure ci-dessous pour créer un DataZone domaine Amazon. Pour plus d'informations sur DataZone les domaines Amazon, consultezDataZone Terminologie et concepts d'Amazon.

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone, connectez-vous, puis choisissez Create domain.

   > **Note**
   >
   > Si vous souhaitez utiliser un DataZone domaine Amazon existant pour ce flux de travail, choisissez Afficher les domaines, puis choisissez le domaine que vous souhaitez utiliser, puis passez à l'étape 2 de création d'un projet de publication.

2. Sur la page Créer un domaine, entrez des valeurs pour les champs suivants :

   - Nom : spécifiez le nom de votre domaine. Dans le cadre de ce flux de travail, vous pouvez appeler ce domaine`Marketing`.
   - Description : spécifiez une description de domaine facultative.
   - Chiffrement des données : vos données sont cryptées par défaut à l'aide d'une clé qui vous AWS appartient et qui est gérée pour vous. Pour cette procédure pas à pas, vous pouvez conserver les paramètres de chiffrement des données par défaut.

     Pour plus d'informations sur l'utilisation des clés gérées par le client, consultezLe chiffrement des données est au repos pour Amazon DataZone. Si vous utilisez votre propre clé KMS pour le chiffrement des données, vous devez inclure l'instruction suivante dans votre clé par défautAmazonDataZoneDomainExecutionRole.

     JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "Statement1",
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt",
                "kms:DescribeKey",
                "kms:GenerateDataKey"
            ],
            "Resource": [
                "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
            ]
```

```
        }
    ]
}
```

- Accès au service : choisissez l'option Utiliser un rôle de service personnalisé, puis choisissez-la dans le AmazonDataZoneDomainExecutionRolemenu déroulant.
- Sous Configuration rapide, choisissez Configurer ce compte pour la consommation et la publication de données. Cette option active les DataZone plans Amazon intégrés du lac de données et de l'entrepôt de données, et configure les autorisations et les ressources requises pour effectuer les autres étapes de ce flux de travail. Pour plus d'informations sur les DataZone plans Amazon, consultezDataZone Terminologie et concepts d'Amazon.
- Conservez les champs restants sous Détails des autorisations et Tags inchangés, puis choisissez Créer un domaine.

3. Une fois le domaine créé avec succès, choisissez-le et, sur la page de résumé du domaine, notez l'URL du portail de données pour ce domaine. Vous pouvez utiliser cette URL pour accéder à votre portail de DataZone données Amazon afin d'effectuer les autres étapes de ce flux de travail.

> **Note**
>
> Dans la version actuelle d'Amazon DataZone, une fois le domaine créé, l'URL générée pour le portail de données ne peut pas être modifiée.

La création d'un domaine peut prendre plusieurs minutes. Attendez que le statut du domaine soit défini sur Disponible avant de passer à l'étape suivante.

Étape 2 - Création du projet de publication

La section suivante décrit les étapes de création du projet de publication dans ce flux de travail.

1. Une fois l'étape 1 terminée, accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de vos informations d'identification unique (SSO) ou AWS IAM.
2. Choisissez Créer un projet, spécifiez le nom du projet. Par exemple, pour ce flux de travail SalesDataPublishingProject, vous pouvez le nommer, laisser les autres champs inchangés, puis choisir Créer.

Étape 3 - Création de l'environnement

La section suivante décrit les étapes de création d'un environnement dans ce flux de travail.

1. Une fois que vous avez terminé l'étape 2, dans le portail de DataZone données Amazon, choisissez le `SalesDataPublishingProject` projet que vous avez créé à l'étape précédente, puis choisissez l'onglet Environnements, puis choisissez Create environment.
2. Sur la page Créer un environnement, spécifiez ce qui suit, puis choisissez Créer un environnement.

   - Nom : spécifiez le nom de l'environnement. Pour cette procédure pas à pas, vous pouvez l'`Default data warehouse environment`appeler.
   - Description : spécifiez une description de l'environnement.
   - Profil d'environnement : choisissez le profil d'DataWarehouseProfileenvironnement.
   - Indiquez le nom de votre cluster Amazon Redshift, le nom de la base de données et l'ARN secret du cluster Amazon Redshift dans lequel vos données sont stockées.

   > **Note**
   >
   > Assurez-vous que votre secret dans AWS Secrets Manager inclut les balises suivantes (clé/valeur) :
   >
   > - Pour le cluster Amazon Redshift - datazone.rs.cluster : <cluster_name:database name>
   >
   >   Pour le groupe de travail Amazon Redshift Serverless - datazone.rs.workgroup : <workgroup_name:database_name>
   > - AmazonDataZoneProject: <projectID>
   > - AmazonDataZoneDomain: <domainID>
   >
   > Pour plus d'informations, consultez Stockage des informations d'identification de base de données dans AWS Secrets Manager.
   > L'utilisateur de base de données que vous indiquez dans le Gestionnaire AWS des Secrets doit disposer des autorisations de superutilisateur.

Étape 4 - Produire des données pour publication

La section suivante décrit les étapes de production des données destinées à être publiées dans ce flux de travail.

1. Une fois l'étape 3 terminée, dans le portail de DataZone données Amazon, choisissez le `SalesDataPublishingProject` projet, puis, dans le panneau de droite, sous Outils d'analyse, choisissez Amazon Redshift. Cela ouvre l'éditeur de requêtes Amazon Redshift en utilisant les informations d'identification de votre projet pour l'authentification.
2. Pour cette procédure pas à pas, vous utilisez le script de requête Create Table as Select (CTAS) pour créer une nouvelle table que vous souhaitez publier sur Amazon. DataZone Dans votre éditeur de requêtes, exécutez ce script CTAS pour créer une `mkt_sls_table` table que vous pouvez publier et rendre disponible pour la recherche et l'abonnement.

   ```
   CREATE TABLE mkt_sls_table AS
   SELECT 146776932 AS ord_num, 23 AS sales_qty_sld, 23.4 AS wholesale_cost, 45.0 as
    lst_pr, 43.0 as sell_pr, 2.0 as disnt, 12 as ship_mode,13 as warehouse_id, 23 as
    item_id, 34 as ctlg_page, 232 as ship_cust_id, 4556 as bill_cust_id
   UNION ALL SELECT 46776931, 24, 24.4, 46, 44, 1, 14, 15, 24, 35, 222, 4551
   UNION ALL SELECT 46777394, 42, 43.4, 60, 50, 10, 30, 20, 27, 43, 241, 4565
   UNION ALL SELECT 46777831, 33, 40.4, 51, 46, 15, 16, 26, 33, 40, 234, 4563
   UNION ALL SELECT 46779160, 29, 26.4, 50, 61, 8, 31, 15, 36, 40, 242, 4562
   UNION ALL SELECT 46778595, 43, 28.4, 49, 47, 7, 28, 22, 27, 43, 224, 4555
   UNION ALL SELECT 46779482, 34, 33.4, 64, 44, 10, 17, 27, 43, 52, 222, 4556
   UNION ALL SELECT 46779650, 39, 37.4, 51, 62, 13, 31, 25, 31, 52, 224, 4551
   UNION ALL SELECT 46780524, 33, 40.4, 60, 53, 18, 32, 31, 31, 39, 232, 4563
   UNION ALL SELECT 46780634, 39, 35.4, 46, 44, 16, 33, 19, 31, 52, 242, 4557
   UNION ALL SELECT 46781887, 24, 30.4, 54, 62, 13, 18, 29, 24, 52, 223, 4561
   ```

   Assurez-vous que la table mkt_sls_table est correctement créée. Vous disposez désormais d'un actif de données qui peut être publié dans le DataZone catalogue Amazon.

Étape 5 - Collectez les métadonnées depuis Amazon Redshift

La section suivante décrit les étapes de collecte de métadonnées depuis Amazon Redshift.

1. Une fois l'étape 4 terminée, dans le portail de DataZone données Amazon, choisissez le `SalesDataPublishingProject` projet, puis l'onglet Données, puis choisissez Sources de données.
2. Choisissez la source créée dans le cadre du processus de création de l'environnement.
3. Choisissez Exécuter à côté du menu déroulant Action, puis cliquez sur le bouton d'actualisation. Une fois l'exécution de la source de données terminée, les actifs sont ajoutés à l' DataZone inventaire Amazon.

Étape 6 - Organiser et publier la ressource de données

La section suivante décrit les étapes de conservation et de publication de la ressource de données dans ce flux de travail.

1. Une fois l'étape 5 terminée, dans le portail de DataZone données Amazon, choisissez le `SalesDataPublishingProject` projet, puis l'onglet Données, choisissez les données d'inventaire et localisez le `mkt_sls_table` tableau.
2. Ouvrez la page de détails de l'`mkt_sls_table`actif pour voir les noms commerciaux générés automatiquement. Cliquez sur l'icône de métadonnées générées automatiquement pour afficher les noms générés automatiquement pour les actifs et les colonnes. Vous pouvez accepter ou rejeter chaque nom individuellement ou choisir Accepter tout pour appliquer les noms générés. Vous pouvez également éventuellement ajouter le formulaire de métadonnées disponible à votre ressource et sélectionner des termes du glossaire pour classer vos données.
3. Choisissez Publier pour publier la `mkt_sls_table` ressource.

Étape 7 - Création du projet pour l'analyse des données

La section suivante décrit les étapes de création du projet pour l'analyse des données dans ce flux de travail.

1. Une fois l'étape 6 terminée, dans le portail de DataZone données Amazon, choisissez Create project.
2. Dans la page Créer un projet, spécifiez le nom du projet. Par exemple, pour ce flux de travail MarketingDataAnalysisProject, vous pouvez le nommer, laisser les autres champs inchangés, puis choisir Créer.

Étape 8 - Création d'un environnement pour l'analyse des données

La section suivante décrit les étapes de création d'un environnement pour l'analyse des données dans ce flux de travail.

1. Une fois l'étape 7 terminée, dans le portail de DataZone données Amazon, choisissez le `MarketingDataAnalysisProject` projet que vous avez créé à l'étape précédente, puis choisissez l'onglet Environnements, puis choisissez Ajouter un environnement.
2. Sur la page Créer un environnement, spécifiez ce qui suit, puis choisissez Créer un environnement.

   - Nom : spécifiez le nom de l'environnement. Pour cette procédure pas à pas, vous pouvez l'`Default data warehouse environment`appeler.
   - Description : spécifiez une description de l'environnement.
   - Profil d'environnement : choisissez le profil d'DataWarehouseProfileenvironnement.
   - Indiquez le nom de votre cluster Amazon Redshift, le nom de la base de données et l'ARN secret du cluster Amazon Redshift dans lequel vos données sont stockées.

     > **Note**
     >
     > Assurez-vous que votre secret dans AWS Secrets Manager inclut les balises suivantes (clé/valeur) :
     >
     > - Pour le cluster Amazon Redshift - datazone.rs.cluster : <cluster_name:database name>
     >
     >   Pour le groupe de travail Amazon Redshift Serverless - datazone.rs.workgroup : <workgroup_name:database_name>
     > - AmazonDataZoneProject: <projectID>
     > - AmazonDataZoneDomain: <domainID>
     >
     > Pour plus d'informations, consultez Stockage des informations d'identification de base de données dans AWS Secrets Manager.
     > L'utilisateur de base de données que vous indiquez dans le Gestionnaire AWS des Secrets doit disposer des autorisations de superutilisateur.

   - Pour cette procédure pas à pas, conservez les autres champs inchangés.

Étape 9 - Rechercher dans le catalogue de données et s'abonner aux données

La section suivante décrit les étapes de recherche dans le catalogue de données et d'abonnement aux données.

1. Une fois l'étape 8 terminée, dans le portail de DataZone données Amazon, recherchez des actifs de données à l'aide de mots clés (par exemple, « catalogue » ou « ventes ») dans la barre de recherche du portail de données.

   Si nécessaire, appliquez des filtres ou effectuez un tri. Une fois que vous avez localisé la ressource Product Sales Data, vous pouvez la choisir pour ouvrir la page de détails de la ressource.
2. Sur la page de détails de la ressource Product Sales Data, choisissez S'abonner.
3. Dans la boîte de dialogue, choisissez votre projet client dans la liste déroulante, indiquez le motif de la demande d'accès, puis choisissez S'abonner.

Étape 10 - Approuver la demande d'abonnement

La section suivante décrit les étapes d'approbation de la demande d'abonnement dans ce flux de travail.

1. Une fois l'étape 9 terminée, dans le portail de DataZone données Amazon, choisissez le SalesDataPublishingProjectprojet avec lequel vous avez publié votre ressource.
2. Choisissez l'onglet Données, puis Données publiées, puis Demandes entrantes.
3. Choisissez le lien de demande de consultation, puis sélectionnez Approuver.

Étape 11 - Création d'une requête et analyse des données dans Amazon Redshift

Maintenant que vous avez publié avec succès une ressource dans le DataZone catalogue Amazon et que vous vous y êtes abonné, vous pouvez l'analyser.

1. Dans le portail de DataZone données Amazon, sur le panneau de droite, cliquez sur le lien Amazon Redshift. Cela ouvre l'éditeur de requêtes Amazon Redshift en utilisant les informations d'identification du projet pour l'authentification.

2. Vous pouvez désormais exécuter une requête (instruction select) sur la table abonnée. Vous pouvez cliquer sur le tableau (three-vertical-dots option) et choisir l'aperçu pour afficher l'instruction sélectionnée sur l'écran de l'éditeur. Exécutez la requête pour voir les résultats.

Amazon DataZone QuickStart avec des exemples de scripts

Vous pouvez accéder à Amazon DataZone via le portail de gestion ou le portail de DataZone données Amazon, ou par programmation à l'aide de l'API Amazon DataZone HTTPS, qui vous permet d'envoyer des requêtes HTTPS directement au service. Cette section contient des exemples de scripts qui invoquent Amazon et DataZone APIs que vous pouvez utiliser pour effectuer les tâches courantes suivantes :

Exemples de scripts

- Création d'un DataZone domaine Amazon et d'un portail de données
- Création d'un projet de publication
- Création d'un profil d'environnement
- Création d'un environnement
- Collectez des métadonnées à partir de AWS Glue
- Organiser et publier un actif de données
- Rechercher dans le catalogue de données et s'abonner aux données
- Rechercher des actifs dans le catalogue de données
- Autres exemples de scripts utiles

Création d'un DataZone domaine Amazon et d'un portail de données

Vous pouvez utiliser l'exemple de script suivant pour créer un DataZone domaine Amazon. Pour plus d'informations sur DataZone les domaines Amazon, consultezDataZone Terminologie et concepts d'Amazon.

```
import sys
import boto3

// Initialize datazone client
region = 'us-east-1'
```

```
dzclient = boto3.client(service_name='datazone', region_name='us-east-1')

// Create DataZone domain
def create_domain(name):
    return dzclient.create_domain(
        name = name,
        description = "this is a description",
        domainExecutionRole = "arn:aws:iam::<account>:role/
AmazonDataZoneDomainExecutionRole",
    )
```

Création d'un projet de publication

Vous pouvez utiliser l'exemple de script suivant pour créer un projet de publication sur Amazon DataZone.

```
// Create Project
def create_project(domainId):
    return dzclient.create_project(
        domainIdentifier = domainId,
        name = "sample-project"
    )
```

Création d'un profil d'environnement

Vous pouvez utiliser les exemples de scripts suivants pour créer un profil d'environnement dans Amazon DataZone.

Cet exemple de charge utile est utilisé lorsque l'`CreateEnvironmentProfile`API est invoquée :

```
Sample Payload
{
    "Content":{
        "project_name": "Admin_project",
        "domain_name": "Drug-Research-and-Development",
        "blueprint_account_region": [
            {
                "blueprint_name": "DefaultDataLake",
                "account_id": ["066535990535",
```

```
                "413878397724",
                "676266385322",
                "747721550195",
                "755347404384"
                ],
                "region": ["us-west-2", "us-east-1"]
            },
            {
                "blueprint_name": "DefaultDataWarehouse",
                "account_id": ["066535990535",
                "413878397724",
                "676266385322",
                "747721550195",
                "755347404384"
                ],
                "region":["us-west-2", "us-east-1"]
            }
        ]
    }
}
```

Cet exemple de script appelle l'`CreateEnvironmentProfile`API :

```
def create_environment_profile(domain_id, project_id, env_blueprints)
        try:
            response = dz.list_environment_blueprints(
                domainIdentifier=domain_id,
                managed=True
            )
            env_blueprints = response.get("items")
            env_blueprints_map = {}
            for i in env_blueprints:
                env_blueprints_map[i["name"]] = i['id']

            print("Environment Blueprint map", env_blueprints_map)
            for i in blueprint_account_region:
                print(i)
                for j in i["account_id"]:
                    for k in i["region"]:
                        print("The env blueprint name is", i['blueprint_name'])
                        dz.create_environment_profile(
```

```
                            description='This is a test environment profile created via
 lambda function',
                            domainIdentifier=domain_id,
                            awsAccountId=j,
                            awsAccountRegion=k,

 environmentBlueprintIdentifier=env_blueprints_map.get(i["blueprint_name"]),
                            name=i["blueprint_name"] + j + k + "_profile",
                            projectIdentifier=project_id
                        )
        except Exception as e:
            print("Failed to created Environment Profile")
            raise e
```

Voici l'exemple de charge utile de sortie une fois que l'`CreateEnvironmentProfile`API est invoquée :

```
{
    "Content":{
        "project_name": "Admin_project",
        "domain_name": "Drug-Research-and-Development",
        "blueprint_account_region": [
            {
                "blueprint_name": "DefaultDataWarehouse",
                "account_id": ["111111111111"],
                "region":["us-west-2"],
                "user_parameters":[
                    {
                        "name": "dataAccessSecretsArn",
                        "value": ""
                    }
                ]
            }
        ]
    }
}
```

Création d'un environnement

Vous pouvez utiliser l'exemple de script suivant pour créer un environnement dans Amazon DataZone.

```
def create_environment(domain_id, project_id,blueprint_account_region ):
         try:
            #refer to get_domain_id and get_project_id for fetching ids using names.
            sts_client = boto3.client("sts")
            # Get the current account ID
            account_id = sts_client.get_caller_identity()["Account"]
            print("Fetching environment profile ids")
            env_profile_map = get_env_profile_map(domain_id, project_id)

            for i in blueprint_account_region:
                for j in i["account_id"]:
                    for k in i["region"]:
                        print(" env blueprint name", i['blueprint_name'])
                        profile_name = i["blueprint_name"] + j + k + "_profile"
                        env_name = i["blueprint_name"] + j + k + "_env"
                        description = f'This is environment is created for
 {profile_name}, Account {account_id} and region {i["region"]}'
                        try:
                            dz.create_environment(
                                description=description,
                                domainIdentifier=domain_id,

 environmentProfileIdentifier=env_profile_map.get(profile_name),
                                name=env_name,
                                projectIdentifier=project_id
                            )
                            print(f"Environment created - {env_name}")
                        except:
                            dz.create_environment(
                                description=description,
                                domainIdentifier=domain_id,

 environmentProfileIdentifier=env_profile_map.get(profile_name),
                                name=env_name,
                                projectIdentifier=project_id,
                                userParameters= i["user_parameters"]
                            )
```

```
                            print(f"Environment created - {env_name}")
        except Exception as e:
            print("Failed to created Environment")
            raise e
```

Collectez des métadonnées à partir de AWS Glue

Vous pouvez utiliser cet exemple de script pour collecter des métadonnées à partir de AWS Glue. Ce script s'exécute selon un calendrier standard. Vous pouvez récupérer les paramètres à partir de l'exemple de script et les rendre globaux. Récupérez le projet, l'environnement et l'ID de domaine à l'aide des fonctions standard. La source de données AWS Glue est créée et exécutée à une heure standard qui peut être mise à jour dans la section cron du script.

```
def crcreate_data_source(domain_id, project_id,data_source_name)
        print("Creating Data Source")
        data_source_creation = dz.create_data_source(
            # Define data source : Customize the data source to which you'd like to
 connect
            # define the name of the Data source to create, example: name
 ='TestGlueDataSource'
            name=data_source_name,
            # give a description for the datasource (optional), example:
 description='This is a dorra test for creation on DZ datasources'
            description=data_source_description,
            # insert the domain identifier corresponding to the domain to which the
 datasource will belong, example: domainIdentifier= 'dzd_6f3gst5jjmrrmv'
            domainIdentifier=domain_id,
            # give environment identifier , example: environmentIdentifier=
 '3weyt6hhn8qcvb'
            environmentIdentifier=environment_id,
            # give corresponding project identifier, example: projectIdentifier=
 '6tl4csoyrg16ef',
            projectIdentifier=project_id,
            enableSetting="ENABLED",
            # publishOnImport used to select whether assets are added to the inventory
 and/or discovery catalog .
            # publishOnImport = True : Assets will be added to project's inventory as
 well as published to the discovery catalog
            # publishOnImport = False : Assets will only be added to project's
 inventory.
```

```
            # You can later curate the metadata of the assets and choose subscription
terms to publish them from the inventory to the discovery catalog.
            publishOnImport=False,
            # Automated business name generation : Use AI to automatically generate
metadata for assets as they are published or updated by this data source run.
            # Automatically generated metadata can be be approved, rejected, or edited
by data publishers.
            # Automatically generated metadata is badged with a small icon next to the
corresponding metadata field.
            recommendation={"enableBusinessNameGeneration": True},
            type="GLUE",
            configuration={
                "glueRunConfiguration": {
                    "dataAccessRole": "arn:aws:iam::"
                    + account_id
                    + ":role/service-role/AmazonDataZoneGlueAccess-"
                    + current_region
                    + "-"
                    + domain_id
                    + "",
                    "relationalFilterConfigurations": [
                        {
                            #
                            "databaseName": glue_database_name,
                            "filterExpressions": [
                                {"expression": "*", "type": "INCLUDE"},
                            ],
                            #    "schemaName": "TestSchemaName",
                        },
                    ],
                },
            },
            # Add metadata forms to the data source (OPTIONAL).
            # Metadata forms will be automatically applied to any assets that are
created by the data source.
            # assetFormsInput=[
            #     {
            #         "content": "string",
            #         "formName": "string",
            #         "typeIdentifier": "string",
            #         "typeRevision": "string",
            #     },
            # ],
            schedule={
```

```
                "schedule": "cron(5 20 * * ? *)",
                "timezone": "UTC",
            },
        )
        # This is a suggested syntax to return values
        #       return_values["data_source_creation"] = data_source_creation["items"]
        print("Data Source Created")


//This is the sample response payload after the CreateDataSource API is invoked:

{
    "Content":{
        "project_name": "Admin",
        "domain_name": "Drug-Research-and-Development",
        "env_name": "GlueEnvironment",
        "glue_database_name": "test",
        "data_source_name" : "test",
        "data_source_description" : "This is a test data source"
    }
}
```

Organiser et publier un actif de données

Vous pouvez utiliser les exemples de scripts suivants pour organiser et publier des actifs de données sur Amazon DataZone.

Vous pouvez utiliser le script suivant pour créer des types de formulaires personnalisés :

```
def create_form_type(domainId, projectId):
    return dzclient.create_form_type(
        domainIdentifier = domainId,
        name = "customForm",
        model = {
            "smithy": "structure customForm { simple: String }"
        },
        owningProjectIdentifier = projectId,
        status = "ENABLED"
    )
```

Vous pouvez utiliser l'exemple de script suivant pour créer des types de ressources personnalisés :

```
def create_custom_asset_type(domainId, projectId):
    return dzclient.create_asset_type(
        domainIdentifier = domainId,
        name = "userCustomAssetType",
        formsInput = {
            "Model": {
                "typeIdentifier": "customForm",
                "typeRevision": "1",
                "required": False
            }
        },
        owningProjectIdentifier = projectId,
    )
```

Vous pouvez utiliser l'exemple de script suivant pour créer des ressources personnalisées :

```
def create_custom_asset(domainId, projectId):
    return dzclient.create_asset(
        domainIdentifier = domainId,
        name = 'custom asset',
        description = "custom asset",
        owningProjectIdentifier = projectId,
        typeIdentifier = "userCustomAssetType",
        formsInput = [
            {
                "formName": "UserCustomForm",
                "typeIdentifier": "customForm",
                "content": "{\"simple\":\"sample-catalogId\"}"
            }
        ]
    )
```

Vous pouvez utiliser l'exemple de script suivant pour créer un glossaire :

```
def create_glossary(domainId, projectId):
```

```
    return dzclient.create_glossary(
        domainIdentifier = domainId,
        name = "test7",
        description = "this is a test glossary",
        owningProjectIdentifier = projectId
    )
```

Vous pouvez utiliser l'exemple de script suivant pour créer un terme de glossaire :

```
def create_glossary_term(domainId, glossaryId):
    return dzclient.create_glossary_term(
        domainIdentifier = domainId,
        name = "soccer",
        shortDescription = "this is a test glossary",
        glossaryIdentifier = glossaryId,
    )
```

Vous pouvez utiliser l'exemple de script suivant pour créer une ressource à l'aide d'un type de ressource défini par le système :

```
def create_asset(domainId, projectId):
    return dzclient.create_asset(
        domainIdentifier = domainId,
        name = 'sample asset name',
        description = "this is a glue table asset",
        owningProjectIdentifier = projectId,
        typeIdentifier = "amazon.datazone.GlueTableAssetType",
        formsInput = [
            {
                "formName": "GlueTableForm",
                "content": "{\"catalogId\":\"sample-catalogId\",\"columns\":
[{\"columnDescription\":\"sample-columnDescription\",\"columnName\":\"sample-
columnName\",\"dataType\":\"sample-dataType\",\"lakeFormationTags\":{\"sample-
key1\":\"sample-value1\",\"sample-key2\":\"sample-value2\"}}],\"compressionType\":
\"sample-compressionType\",\"lakeFormationDetails\":{\"lakeFormationManagedTable
\":false,\"lakeFormationTags\":{\"sample-key1\":\"sample-value1\",\"sample-key2\":
\"sample-value2\"}},\"primaryKeys\":[\"sample-Key1\",\"sample-Key2\"],\"region\":
\"us-east-1\",\"sortKeys\":[\"sample-sortKey1\"],\"sourceClassification\":\"sample-
sourceClassification\",\"sourceLocation\":\"sample-sourceLocation\",\"tableArn\":
```

```
\"sample-tableArn\",\"tableDescription\":\"sample-tableDescription\",\"tableName\":
\"sample-tableName\"}"
            }
        ]
    )
```

Vous pouvez utiliser l'exemple de script suivant pour créer une révision de ressource et y joindre un terme du glossaire :

```
def create_asset_revision(domainId, assetId):
    return dzclient.create_asset_revision(
        domainIdentifier = domainId,
        identifier = assetId,
        name = 'glue table asset 7',
        description = "glue table asset description update",
        formsInput = [
            {
                "formName": "GlueTableForm",
                "content": "{\"catalogId\":\"sample-catalogId\",\"columns\":
[{\"columnDescription\":\"sample-columnDescription\",\"columnName\":\"sample-
columnName\",\"dataType\":\"sample-dataType\",\"lakeFormationTags\":{\"sample-
key1\":\"sample-value1\",\"sample-key2\":\"sample-value2\"}}],\"compressionType\":
\"sample-compressionType\",\"lakeFormationDetails\":{\"lakeFormationManagedTable
\":false,\"lakeFormationTags\":{\"sample-key1\":\"sample-value1\",\"sample-key2\":
\"sample-value2\"}},\"primaryKeys\":[\"sample-Key1\",\"sample-Key2\"],\"region\":
\"us-east-1\",\"sortKeys\":[\"sample-sortKey1\"],\"sourceClassification\":\"sample-
sourceClassification\",\"sourceLocation\":\"sample-sourceLocation\",\"tableArn\":
\"sample-tableArn\",\"tableDescription\":\"sample-tableDescription\",\"tableName\":
\"sample-tableName\"}"
            }
        ],
        glossaryTerms = ["<glossaryTermId:>"]
    )
```

Vous pouvez utiliser l'exemple de script suivant pour publier une ressource :

```
def publish_asset(domainId, assetId):
    return dzclient.create_listing_change_set(
        domainIdentifier = domainId,
```

```
        entityIdentifier = assetId,
        entityType = "ASSET",
        action = "PUBLISH",
    )
```

Rechercher dans le catalogue de données et s'abonner aux données

Vous pouvez utiliser les exemples de scripts suivants pour effectuer des recherches dans le catalogue de données et vous abonner aux données :

```
def search_asset(domainId, projectId, text):
    return dzclient.search(
        domainIdentifier = domainId,
        owningProjectIdentifier = projectId,
        searchScope = "ASSET",
        searchText = text,
    )
```

Vous pouvez utiliser l'exemple de script suivant pour obtenir l'ID de liste de l'actif :

```
def search_listings(domainId, assetName, assetId):
    listings = dzclient.search_listings(
        domainIdentifier=domainId,
        searchText=assetName,
        additionalAttributes=["FORMS"]
    )

    assetListing = None
    for listing in listings['items']:
        if listing['assetListing']['entityId'] == assetId:
            assetListing = listing

    return listing['assetListing']['listingId']
```

Vous pouvez utiliser les exemples de scripts suivants pour créer une demande d'abonnement à l'aide de l'ID de liste :

```
create_subscription_response = def create_subscription_request(domainId, projectId,
 listingId):
    return dzclient.create_subscription_request(
        subscribedPrincipals=[{
            "project": {
                "identifier": projectId
            }
        }],
        subscribedListings=[{
            "identifier": listingId
        }],
        requestReason="Give request reason here."
    )
```

À l'aide de `create_subscription_response` ce qui précède, obtenez le `subscription_request_id`, puis accept/approve l'abonnement à l'aide de l'exemple de script suivant :

```
subscription_request_id = create_subscription_response["id"]

def accept_subscription_request(domainId, subscriptionRequestId):
    return dzclient.accept_subscription_request(
        domainIdentifier=domainId,
        identifier=subscriptionRequestId
    )
```

Rechercher des actifs dans le catalogue de données

Vous pouvez utiliser les exemples de scripts suivants qui utilisent la recherche en texte libre pour rechercher vos actifs de données publiés (listes) dans le DataZone catalogue Amazon.

- L'exemple suivant effectue une recherche par mot clé en texte libre dans le domaine et renvoie toutes les listes correspondant au mot clé « crédit » fourni :

```
aws datazone search-listings \
  --domain-identifier dzd_c1s7uxe71prrtz \
  --search-text "credit"
```

- Vous pouvez également combiner plusieurs mots clés pour affiner davantage le champ de recherche. Par exemple, si vous recherchez tous les actifs de données publiés (listes) contenant des données relatives aux ventes au Mexique, vous pouvez formuler votre requête avec deux mots clés « Mexique » et « ventes ».

```
          aws datazone search-listings \
  --domain-identifier dzd_c1s7uxe71prrtz \
  --search-text "mexico sales"
```

Vous pouvez également rechercher une annonce à l'aide de filtres. Le `filters` paramètre de l' SearchListings API vous permet de récupérer les résultats filtrés du domaine. L'API prend en charge plusieurs filtres par défaut et vous pouvez également combiner deux filtres ou plus et effectuer une opération ET/OU sur eux. La clause filter prend en compte deux paramètres : attribut et valeur. Les attributs de filtre pris en charge par défaut sont `typeName``owningProjectId`, et`glossaryTerms`.

- L'exemple suivant effectue une recherche dans toutes les listes d'un domaine donné à l'aide du `assetType` filtre dans lequel la liste est un type de Redshift Table.

```
           aws datazone search-listings \
--domain-identifier dzd_c1s7uxe71prrtz \
--filters '{"or":[{"filter":
{"attribute":"typeName","value":"RedshiftTableAssetType"}} ]}'
```

- Vous pouvez également combiner plusieurs filtres à l'aide d' AND/OR opérations. Dans l'exemple suivant, vous combinez `typeName` et `project` filtrez.

```
           aws datazone search-listings \
--domain-identifier dzd_c1s7uxe71prrtz \
--filters '{"or":[{"filter":
{"attribute":"typeName","value":"RedshiftTableAssetType"}},  {"filter":
{"attribute":"owningProjectId","value":"cwrrjch7f5kppj"}} ]}'
```

- Vous pouvez même combiner la recherche en texte libre avec des filtres pour obtenir des résultats exacts et les trier en fonction de l'heure de creation/last mise à jour de la liste, comme indiqué dans l'exemple suivant :

```
            aws datazone search-listings \
--domain-identifier dzd_c1s7uxe71prrtz \
--search-text "finance sales" \
--filters '{"or":[{"filter":{"attribute":"typeName","value":"GlueTableViewType"}} ]}'
 \
--sort '{"attribute": "UPDATED_AT", "order":"ASCENDING"}'
```

Autres exemples de scripts utiles

Vous pouvez utiliser les exemples de scripts suivants pour effectuer diverses tâches lorsque vous travaillez avec vos données sur Amazon DataZone.

Utilisez l'exemple de script suivant pour répertorier les DataZone domaines Amazon existants :

```
def list_domains():
    datazone = boto3.client('datazone')
    response = datazone.list_domains(status='AVAILABLE')
    [print("%12s | %16s | %12s | %52s" % (item['id'], item['name'],
 item['managedAccountId'], item['portalUrl'])) for item in response['items']]
    return
```

Utilisez l'exemple de script suivant pour répertorier les DataZone projets Amazon existants :

```
def list_projects(domain_id):
    datazone = boto3.client('datazone')
    response = datazone.list_projects(domainIdentifier=domain_id)
    [print("%12s | %16s " % (item['id'], item['name'])) for item in response['items']]
    return
```

Utilisez l'exemple de script suivant pour répertorier les formulaires de DataZone métadonnées Amazon existants :

```
def list_metadata_forms(domain_id):
    datazone = boto3.client('datazone')
    response = datazone.search_types(domainIdentifier=domain_id,
        managed=False,
        searchScope='FORM_TYPE')
    [print("%16s | %16s | %3s | %8s" % (item['formTypeItem']['name'],
 item['formTypeItem']['owningProjectId'],item['formTypeItem']['revision'],
 item['formTypeItem']['status'])) for item in response['items']]
    return
```

Domaines et accès utilisateur sur Amazon DataZone

Cette section décrit comment créer et gérer les domaines et l'accès des utilisateurs sur Amazon DataZone.

Un DataZone domaine Amazon est l'entité organisatrice qui connecte vos actifs, vos utilisateurs et leurs projets. Avec DataZone les domaines Amazon, vous avez la flexibilité nécessaire pour répondre aux besoins en données et en analyse de votre structure organisationnelle, qu'il s'agisse de créer un seul DataZone domaine Amazon pour votre entreprise ou plusieurs zones de données, des domaines pour différentes unités commerciales ou équipes.

Cette section décrit également la gestion de l'accès des utilisateurs à la DataZone console Amazon et au portail Amazon DataZone.

Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Rubriques

- Création de DataZone domaines Amazon
- Modifier les DataZone domaines Amazon
- Supprimer des DataZone domaines Amazon
- Activer IAM Identity Center pour Amazon DataZone
- Désactiver IAM Identity Center pour Amazon DataZone
- Gérer les utilisateurs dans la DataZone console Amazon
- Gérez les autorisations des utilisateurs sur le portail DataZone de données Amazon
- Restreindre l'accès à Amazon DataZone
- Mise à niveau DataZone des domaines Amazon vers des domaines SageMaker unifiés Amazon

Création de DataZone domaines Amazon

> **Note**
>
> Si vous utilisez Amazon DataZone avec AWS Identity Center pour fournir un accès aux utilisateurs et aux groupes SSO, votre DataZone domaine Amazon doit actuellement se trouver dans la même AWS région que votre instance AWS Identity Center.

Amazon DataZone, un domaine est une entité organisatrice permettant de connecter vos actifs, vos utilisateurs et leurs projets. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour créer un DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales nécessaires à la création d'un domaine.

Amazon a besoin de rôles IAM supplémentaires DataZone pour effectuer des actions au nom des utilisateurs du domaine avec une configuration par défaut. Vous pouvez créer ces rôles IAM à l'avance ou demander à Amazon de les DataZone créer pour vous. Si vous souhaitez qu'Amazon DataZone crée ces rôles IAM pour vous pendant le processus de création du domaine, vous devez alors assumer un rôle IAM avec des autorisations de création de rôles. Consultez Créez une politique personnalisée pour les autorisations IAM afin de permettre à la console de DataZone service Amazon de simplifier la création de rôles. En fonction de vos choix de création de domaine, Amazon DataZone créera jusqu'à quatre nouveaux rôles IAM pour vous : AmazonDataZoneDomainExecutionRole, AmazonDataZoneGlueManageAccessRoleAmazonDataZoneRedshiftManageAccessRole, et AmazonDataZoneProvisioningRole.

Suivez la procédure ci-dessous pour créer un DataZone domaine Amazon.

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et utilisez le sélecteur de région dans la barre de navigation supérieure pour choisir la région appropriée. AWS
2. Choisissez Créer un domaine et entrez des valeurs pour les champs suivants :

   - Nom : spécifiez un nom convivial pour le domaine. Une fois le domaine créé, ce nom ne peut pas être modifié.
   - Description - (facultatif) spécifiez une description de domaine.
   - Chiffrement des données : votre DataZone domaine Amazon, vos métadonnées et vos données de reporting sont chiffrés par le AWS Key Management Service (KMS) à l'aide d'une clé spécifique à votre Amazon DataZone. Utilisez ce champ pour indiquer si vous souhaitez utiliser une clé AWS détenue ou choisir une autre clé AWS KMS.

     Pour plus d'informations sur l'utilisation des clés gérées par le client, consultezLe chiffrement des données est au repos pour Amazon DataZone. Si vous utilisez votre propre clé KMS

pour le chiffrement des données, vous devez inclure l'instruction suivante dans votre clé par défautAmazonDataZoneDomainExecutionRole.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "Statement1",
            "Effect": "Allow",
            "Action": [
                "kms:Decrypt",
                "kms:DescribeKey",
                "kms:GenerateDataKey"
            ],
            "Resource": [
                "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
            ]
        }
    ]
}
```

- Accès au service : choisissez si vous voulez qu'Amazon en DataZone crée et utilise un nouveau DomainExecutionRolepour vous, ou choisissez un rôle IAM existant.
- Configuration rapide : (facultatif) cochez cette case pour démarrer plus rapidement en demandant à Amazon de DataZone configurer votre compte pour la consommation et la publication de données. Amazon DataZone créera trois rôles IAM pour le provisionnement, l'ingestion et la gestion de l'accès aux ressources AWS Glue et Amazon Redshift, créera un nouveau compartiment Amazon S3, créera un DataZone projet Amazon administratif et créera des profils d'environnement pour les plans par défaut du lac de données et de l'entrepôt de données.
- Balises : (facultatif) spécifiez les AWS balises (paires clé/valeur) pour le domaine.
- Une fois le domaine créé avec succès, votre navigateur doit être actualisé pour afficher la page de détails de votre nouveau DataZone domaine Amazon.

Modifier les DataZone domaines Amazon

Dans Amazon DataZone, un domaine est une entité organisatrice permettant de relier vos actifs, vos utilisateurs et leurs projets. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Après avoir créé un DataZone domaine Amazon, vous pouvez le modifier ultérieurement pour : modifier la description, activer IAM Identity Center et ajouter, modifier ou supprimer des clés de balise et leurs valeurs. Pour modifier un DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales nécessaires pour modifier un domaine.

Pour modifier un domaine, procédez comme suit :

1. Connectez-vous à la console de AWS gestion et ouvrez la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Sur la page de détails du domaine, choisissez Modifier.
4. - Modifiez la description.
   - Définissez les paramètres du centre d'identité IAM. Pour en savoir plus sur ces paramètres, consultezConfiguration de l' AWS IAM Identity Center pour Amazon DataZone.
   - Ajoutez, modifiez ou supprimez les clés Tag et leurs valeurs.
5. Une fois que vous avez effectué vos modifications, choisissez Mettre à jour le domaine.

Supprimer des DataZone domaines Amazon

Dans Amazon DataZone, un domaine est une entité organisatrice permettant de relier vos actifs, vos utilisateurs et leurs projets. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

La suppression d'un domaine est définitive. La suppression supprime irrévocablement toutes les DataZone entités Amazon, y compris les sources de données, les projets, les environnements, les actifs, les glossaires et les formulaires de métadonnées. La suppression ne supprime pas les DataZone AWS ressources extérieures à Amazon qu'Amazon DataZone peut vous avoir aidé à créer, telles que les rôles IAM, les compartiments S3, les bases de données AWS Glue et les subventions

d'abonnement via ou LakeFormation Redshift. Si vous n'avez plus besoin de ces ressources, supprimez-les dans le AWS service correspondant.

Pour empêcher quelqu'un de supprimer un domaine de manière malveillante, la suppression d'un domaine nécessite des autorisations administratives IAM pour Amazon DataZone, que vous pouvez configurer avec IAM. Pour empêcher quelqu'un de supprimer un domaine accidentellement, la suppression d'un domaine nécessite un mot de confirmation (dans la DataZone console Amazon).

Pour supprimer un domaine, procédez comme suit :

1. Connectez-vous à la console de AWS gestion et ouvrez la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Choisissez Supprimer et consultez les avertissements d'information.
4. Tapez le texte demandé pour confirmer que vous avez bien compris ces avertissements. Sélectionnez Delete (Supprimer).

Important

La suppression de votre domaine est une action irrévocable qui ne peut être annulée ni par vous ni par vous. AWS

Note

Lorsque vous ou les utilisateurs de votre domaine créez un environnement dans un projet, Amazon DataZone crée AWS des ressources dans votre domaine ou dans les comptes associés afin de vous fournir, ainsi qu'aux utilisateurs de votre domaine, des fonctionnalités. Vous trouverez ci-dessous la liste des AWS ressources qu'Amazon DataZone peut créer pour les projets de votre domaine, ainsi que le nom par défaut. La suppression d'un domaine ne supprime aucune de ces AWS ressources de vos AWS comptes.

- <environmentId>Rôles IAM : datazone_usr_.
- <environmentName>Bases de données Glue : (1) <environmentName>_pub_db-*, (2) _sub_db-*. S'il existe déjà une base de données portant ce nom, Amazon DataZone ajoutera l'ID d'environnement.

- <environmentName>Groupes de travail Athena : -*. S'il existe déjà un groupe de travail portant ce nom, Amazon DataZone ajoutera l'ID d'environnement.
- CloudWatch groupe de journaux : datazone_ <environmentId>

Activer IAM Identity Center pour Amazon DataZone

> **Note**
>
> Pour terminer cette procédure, le centre d'identité AWS IAM doit être activé dans la même AWS région que votre DataZone domaine Amazon.

Vous pouvez fournir aux utilisateurs et aux groupes SSO l'accès à votre portail de DataZone données Amazon à l'aide d' AWS IAM Identity Center. Une fois que vous avez terminéConfiguration de l' AWS IAM Identity Center pour Amazon DataZone, vous pouvez permettre à vos utilisateurs et groupes SSO d'accéder à votre portail de données de DataZone domaine Amazon.

Pour activer l'utilisation d' AWS IAM Identity Center avec votre DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonet Créez une politique personnalisée pour les autorisations IAM afin de permettre à la console de DataZone service Amazon de simplifier la création de rôles pour obtenir les autorisations minimales nécessaires pour activer l'utilisation d'IAM Identity Center avec Amazon DataZone.

Suivez la procédure ci-dessous pour activer l' AWS IAM Identity Center pour Amazon DataZone.

1. Connectez-vous à la console de AWS gestion et ouvrez la console à l'adresse DataZone https://console.aws.amazon.com/datazone.
2. Sélectionnez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Sur la page détaillée du domaine, choisissez Modifier.
   - Cochez la case Activer les utilisateurs dans IAM Identity Center.
   - Choisissez de vous connecter à une instance d'organisation du centre d'identité IAM ou de vous connecter à une instance de compte du centre d'identité IAM.

- Choisissez entre les deux modes d'attribution des utilisateurs. Une fois que votre domaine est mis à jour avec votre sélection, il ne peut pas être modifié ultérieurement.
  - Avec l'attribution d'utilisateurs implicite, tout utilisateur ajouté à votre annuaire IAM Identity Center peut accéder à votre DataZone domaine Amazon.
  - Avec l'attribution d'utilisateurs explicite, vous ajouterez des utilisateurs ou des groupes spécifiques à partir de votre annuaire IAM Identity Center pour leur permettre d'accéder à votre DataZone domaine Amazon. Vous ajouterez et supprimerez ces utilisateurs et groupes ultérieurement dans la DataZone console Amazon.

4. Une fois que vous êtes satisfait de votre sélection, choisissez Mettre à jour le domaine.

Désactiver IAM Identity Center pour Amazon DataZone

La désactivation d' AWS IAM Identity Center pour un DataZone domaine Amazon supprimera l'accès de tous les utilisateurs SSO.

> **Note**
>
> La désactivation d'IAM Identity Center n'interrompra pas la facturation des utilisateurs SSO. Pour arrêter la facturation aux utilisateurs du SSO, vous devez les désactiver dans votre domaine. La facturation se poursuit jusqu'à la fin du mois au cours duquel un utilisateur est désactivé. Pour désactiver des utilisateurs, consultezGérer les utilisateurs dans la DataZone console Amazon.

Vous pouvez fournir aux utilisateurs et aux groupes SSO l'accès à votre portail de DataZone données Amazon à l'aide d' AWS IAM Identity Center. Si vous avez activé AWS IAM Identity Center pour Amazon DataZone, vous pourrez ultérieurement désactiver l'accès pour tous les utilisateurs.

Pour désactiver AWS IAM Identity Center afin de l'utiliser avec votre DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonet Créez une politique personnalisée pour les autorisations IAM afin de permettre à la console de DataZone service Amazon de simplifier la création de rôles pour obtenir les autorisations minimales nécessaires pour désactiver l'utilisation d'IAM Identity Center avec Amazon DataZone.

Suivez la procédure ci-dessous pour désactiver l' AWS IAM Identity Center pour Amazon DataZone.

1. Connectez-vous à la console de AWS gestion et ouvrez la console à l'adresse DataZone https://console.aws.amazon.com/datazone.
2. Sélectionnez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. <regionName><accountId><domainName>Copiez le nom de ressource Amazon (ARN) de votre domaine, qui commence par arn:aws:datazone : ::domain/.
4. Ouvrez la console IAM Identity Center à https://console.aws.amazon.com/singlesignon/l'adresse.
5. Choisissez Applications.
6. Choisissez le domaine pour lequel vous souhaitez désactiver AWS IAM Identity Center, ce qui supprimera l'accès au portail de données du domaine pour tous les utilisateurs de l'authentification unique. Vous pouvez utiliser le menu Filtre et le champ de recherche pour filtrer la liste des applications.
7. Dans le menu Actions, choisissez Désactiver.
8. Les utilisateurs du SSO perdront l'accès au DataZone domaine Amazon.
9. Pour réactiver AWS IAM Identity Center pour le DataZone domaine Amazon, choisissez le domaine pour lequel vous souhaitez réactiver AWS IAM Identity Center, puis dans le menu Actions, choisissez Activer.

Gérer les utilisateurs dans la DataZone console Amazon

Vos utilisateurs peuvent accéder au portail de DataZone données Amazon en utilisant leurs AWS informations d'identification ou leurs informations d'identification unique (SSO). Pour gérer les utilisateurs dans la DataZone console Amazon pour un DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec les autorisations de la console DataZone de gestion Amazon. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales nécessaires pour gérer les utilisateurs dans la DataZone console Amazon.

Rubriques

- Gérer les rôles et les utilisateurs IAM
- Gérer les utilisateurs SSO
- Gérer les groupes SSO

Gérer les rôles et les utilisateurs IAM

Les rôles et utilisateurs IAM sont créés à l'aide AWS d'Identity and Access Management (IAM) et accèdent à vos DataZone domaines Amazon grâce aux autorisations qui leur sont associées via des politiques. Pour de plus amples informations, veuillez consulter Configurer les autorisations IAM requises pour utiliser le portail de DataZone données Amazon. Dans la version actuelle d'Amazon DataZone, un administrateur d'un compte de propriétaire de DataZone domaine Amazon peut créer des profils utilisateur IAM pour les utilisateurs de son propre compte ou pour les utilisateurs des comptes associés. Un administrateur d'un compte de propriétaire de DataZone domaine Amazon peut également définir le statut d'un utilisateur existant sur Assigné ou Non attribué (comme dans le cas d'un utilisateur assigné ou non assigné pour utiliser Amazon DataZone) ou activer ou désactiver un utilisateur existant.

1. Connectez-vous à la console de AWS gestion et ouvrez la console à l'adresse DataZone https://console.aws.amazon.com/datazone.
2. Sélectionnez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Sur la page de détails du domaine, choisissez Gestion des utilisateurs.
4. Pour ajouter un utilisateur IAM dans le compte du propriétaire du DataZone domaine Amazon ou dans le compte associé, choisissez Ajouter, puis sélectionnez Ajouter des utilisateurs IAM.
5. Sur la page Ajouter des utilisateurs, choisissez Compte actuel ou Compte associé, utilisez le champ Rechercher et ajouter des utilisateurs ou des rôles pour rechercher les utilisateurs que vous souhaitez ajouter, puis choisissez Ajouter des utilisateurs.
6. Pour consulter le statut d'un utilisateur IAM existant, sur la page de gestion des utilisateurs, sélectionnez Utilisateurs IAM dans le menu déroulant du type d'utilisateur.

    - La colonne Nom indique l'ARN de l'utilisateur ou du rôle IAM.
    - La colonne Status indique le statut actuel de l'utilisateur ou du rôle IAM dans le domaine.
        - Attribué signifie que l'utilisateur IAM a été autorisé à utiliser Amazon DataZone.
        - Non attribué signifie que l'utilisateur IAM n'a pas été autorisé à utiliser Amazon. DataZone
        - Activé signifie que l'utilisateur ou le rôle IAM a appelé une API, émis une commande (via l'interface de ligne de commande) ou accédé au DataZone portail Amazon pour votre domaine.

- Désactivé signifie que l'utilisateur ou le rôle IAM ne peut plus utiliser Amazon DataZone Data Portal. Pour restreindre l'accès programmatique, voirRestreindre l'accès à Amazon DataZone.

7. Pour désactiver un utilisateur ou un rôle IAM actuellement activé, cochez la case à côté de l'utilisateur et sélectionnez Désactiver dans le menu Actions. Cela empêchera l'utilisateur d'utiliser le portail de DataZone données Amazon. Pour restreindre l'accès programmatique, voirRestreindre l'accès à Amazon DataZone.
8. Pour activer un utilisateur ou un rôle IAM actuellement désactivé, cochez la case à côté de l'utilisateur et sélectionnez Activer dans le menu Actions. L'utilisateur aura accès au portail Amazon DataZone Data si l'utilisateur ou le rôle IAM dispose d'`datazone:GetUserPortalLoginUrl`autorisations.

Gérer les utilisateurs SSO

Les utilisateurs SSO sont créés ou synchronisés avec votre fournisseur d'identité. Pour plus d'informations, consultez Configuration de l' AWS IAM Identity Center pour Amazon DataZone et Activer IAM Identity Center pour Amazon DataZone pour activer et configurer AWS IAM Identity Center pour Amazon DataZone. Vous pouvez consulter la liste des utilisateurs SSO assignés au domaine, ajouter des utilisateurs SSO et supprimer des utilisateurs SSO.

1. Connectez-vous à la console de AWS gestion et ouvrez la console à l'adresse DataZone https://console.aws.amazon.com/datazone.
2. Sélectionnez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Sur la page de détails du domaine, faites défiler la page vers le bas et choisissez Gestion des utilisateurs.
4. Pour le type d'utilisateur, sélectionnez Utilisateurs SSO pour afficher la liste actuelle des utilisateurs SSO qui se sont déjà authentifiés sur le portail de données. Lorsque l'attribution d'utilisateurs implicite est utilisée, les utilisateurs SSO qui ne se sont pas authentifiés auparavant sur le portail de données ne seront pas répertoriés.

   - La colonne Nom indique le nom de l'utilisateur SSO.
   - La colonne Status indique le statut actuel de l'utilisateur SSO dans le domaine.

- Attribué signifie que l'utilisateur SSO a été explicitement affecté au domaine. Par conséquent, l'utilisateur a accès à Amazon DataZone. Ce statut n'est utilisé que lorsque le mode fournisseur d'identité de votre domaine est défini sur une attribution explicite.
- Activé signifie que l'utilisateur SSO a accédé au DataZone portail Amazon pour le domaine. L'activation se fait automatiquement.
- Désactivé signifie que l'accès de l'utilisateur SSO au portail de données du domaine est bloqué.
- Supprimé signifie que l'utilisateur SSO a déjà été affecté au domaine, mais qu'il a été supprimé avant même d'y avoir accédé.

5. Ajoutez des utilisateurs SSO en choisissant Ajouter et Ajouter des utilisateurs. Cette option n'est pas disponible si le domaine est défini sur une attribution d'utilisateur implicite, ce qui signifie que tous les utilisateurs du pool d'identités ont accès au DataZone domaine Amazon.

   - Sur la page Ajouter des utilisateurs, recherchez les alias des utilisateurs que vous souhaitez ajouter. Une liste de correspondances potentielles apparaîtra sous le champ de recherche.
   - Choisissez l'utilisateur que vous souhaitez ajouter. Leur alias apparaîtra sous la forme d'une puce sous le champ de recherche.
   - Lorsque vous êtes satisfait de la liste des utilisateurs que vous souhaitez ajouter, choisissez Ajouter un ou plusieurs utilisateurs.
   - Les utilisateurs sont affectés au DataZone domaine Amazon avec le statut Attribué.
   - Lorsque l'utilisateur a accédé pour la première fois au portail de données du domaine, le statut passe automatiquement à Activé.

6. Supprimez un utilisateur SSO attribué en le sélectionnant et en choisissant Annuler l'attribution dans le menu Actions. Par conséquent, l'utilisateur perdra l'accès au DataZone domaine Amazon. Le statut de l'utilisateur sera affiché comme Non attribué. Cette option n'est pas disponible si le domaine est défini sur une attribution utilisateur implicite.
7. Désactivez un utilisateur SSO activé en le sélectionnant et en choisissant Désactiver dans le menu Actions. Par conséquent, l'accès de l'utilisateur au portail de DataZone données Amazon sera perdu et bloqué. Le statut de l'utilisateur sera indiqué comme Désactivé.
8. Activez un utilisateur SSO désactivé en le sélectionnant et en choisissant Activer dans le menu Actions. Par conséquent, l'utilisateur retrouvera l'accès au portail de DataZone données Amazon. Celui de l'utilisateur s'affichera comme Activé.

Gérer les groupes SSO

Les groupes SSO sont créés ou synchronisés avec votre fournisseur d'identité dans AWS IAM Identity Center. Pour plus d'informations, consultez Configuration de l' AWS IAM Identity Center pour Amazon DataZone et Activer IAM Identity Center pour Amazon DataZone pour activer et configurer AWS IAM Identity Center pour Amazon DataZone. Vous pouvez consulter la liste des groupes SSO attribués au domaine, ajouter des groupes SSO et supprimer des groupes SSO.

1. Connectez-vous à la console de AWS gestion et ouvrez la console à l'adresse DataZone https://console.aws.amazon.com/datazone.
2. Sélectionnez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Sur la page de détails du domaine, faites défiler la page vers le bas et choisissez Gestion des utilisateurs.
4. Pour le type d'utilisateur, sélectionnez Groupes SSO pour afficher la liste actuelle des groupes SSO.

   - La colonne Nom indique le nom du groupe SSO.
   - La colonne Status indique le statut actuel du groupe SSO dans le domaine.
     - Attribué signifie que le groupe SSO a été explicitement attribué au domaine. Par conséquent, tous les utilisateurs du groupe ont accès au portail de données du domaine (sauf si l'utilisateur est désactivé).
     - Non attribué signifie que le groupe SSO a été supprimé du domaine. Les utilisateurs du groupe n'ont pas accès au portail de données du domaine par le biais de leur appartenance à ce groupe.
5. Ajoutez des groupes SSO en choisissant Ajouter et ajouter des groupes. Cette option n'est pas disponible si le domaine est défini sur une attribution d'utilisateur implicite, ce qui signifie que tous les utilisateurs du pool d'identités ont accès au DataZone domaine Amazon, quelle que soit leur appartenance au groupe.

   - Sur la page Ajouter des groupes, recherchez les alias des groupes que vous souhaitez ajouter. Une liste de correspondances potentielles apparaîtra sous le champ de recherche.
   - Choisissez le groupe que vous souhaitez ajouter. Leur alias apparaîtra sous la forme d'une puce sous le champ de recherche.
   - Lorsque vous êtes satisfait de la liste des groupes que vous souhaitez ajouter, choisissez Ajouter un ou plusieurs groupes.

- Les groupes sont affectés au DataZone domaine Amazon avec le statut Assigned.
- Lorsqu'un membre du groupe accède au portail de données du domaine, le statut passe automatiquement à Activé.

6. Supprimez un groupe SSO attribué en le sélectionnant et en choisissant Annuler l'attribution dans le menu Actions. Par conséquent, le groupe perdra l'accès au DataZone domaine Amazon. Le statut du groupe sera indiqué comme Non attribué. Les utilisateurs qui ont obtenu leur accès à Amazon DataZone via leur appartenance à ce groupe perdront leur accès. Cette option n'est pas disponible si le domaine est défini sur une attribution utilisateur implicite.

Gérez les autorisations des utilisateurs sur le portail DataZone de données Amazon

Vous pouvez utiliser le portail de DataZone gestion Amazon pour configurer l'authentification pour les utilisateurs et les rôles IAM, les utilisateurs et les groupes SSO et les utilisateurs SAML. Amazon DataZone attribue un profil utilisateur à chaque utilisateur qui utilise Amazon DataZone.

Les autorisations du profil utilisateur permettant d'utiliser des projets, de créer des entités, etc. sont gérées à l'aide d'unités de domaine et de subventions politiques. Au sein d'un projet spécifique, la désignation des membres du projet (propriétaire, contributeur, vue) détermine l'autorisation d'action.

Restreindre l'accès à Amazon DataZone

Restreindre l'accès programmatique à Amazon DataZone : pour les utilisateurs ou les rôles IAM effectuant des appels d'API programmatiques, l'accès peut être restreint via des politiques IAM. Si vous souhaitez révoquer des informations d'identification à court terme déjà émises pour des rôles, vous pouvez utiliser le mécanisme de révocation de session IAM sur le rôle ou sur la politique de contrôle des services.

Restreindre l'accès au portail de DataZone données Amazon : pour restreindre l'accès au portail de DataZone données Amazon, pour les utilisateurs ou les rôles IAM, les politiques IAM peuvent restreindre l'accès à l'`datazone:GetUserPortalLoginUrl`action. Pour les utilisateurs et les groupes SSO, limitez l'accès au portail de DataZone données Amazon en définissant le statut du profil DataZone utilisateur Amazon sur Désactivé. Si votre domaine est configuré avec une attribution implicite et que l'utilisateur n'a jamais utilisé Amazon auparavant DataZone, vous devrez le supprimer du fournisseur d'identité.

Mise à niveau DataZone des domaines Amazon vers des domaines SageMaker unifiés Amazon

Considérations à prendre en compte avant de mettre à niveau votre domaine

Avant de mettre à niveau votre DataZone domaine Amazon vers un domaine SageMaker unifié Amazon, prenez en compte ces points importants pour garantir le bon déroulement du processus de mise à niveau.

- Le processus de mise à niveau est uniquement disponible via la console AWS de gestion. Actuellement, aucun support d'API n'est proposé pour la mise à niveau de votre domaine. Vous pouvez initialiser le processus de mise à niveau depuis la page des détails du domaine de votre DataZone domaine Amazon.
- Le processus de mise à niveau nécessite la configuration des rôles suivants (vous pouvez sélectionner des rôles existants ou demander à Amazon SageMaker Unified Studio de créer les rôles en votre nom) :
  - Rôle d'exécution du domaine : pour un DataZone domaine Amazon, vous utilisez le AmazonDataZoneDomainExecutionRolerôle requis par Amazon DataZone pour cataloguer, découvrir, gouverner, partager et analyser les données de votre domaine. Avec un domaine SageMaker unifié Amazon, vous devez utiliser le AmazonSageMakerDomainExecutionrôle existant ou en créer un nouveau.
  - Rôle de service de domaine : Amazon DataZone n'a pas besoin de rôle de service de domaine. Avec un domaine SageMaker unifié Amazon, vous devez utiliser le AmazonSageMakerDomainServicerôle existant ou en créer un nouveau. Il s'agit d'un rôle de service pour les actions au niveau du domaine effectuées par Amazon SageMaker Unified Studio.
- Considérations relatives à la propriété du domaine racine :
  - Les utilisateurs IAM ou SSO users/groups peuvent éventuellement être désignés en tant que propriétaires de domaines racine au cours du processus de mise à niveau.
  - Si seuls des rôles IAM sont assignés en tant que propriétaires à l'unité de domaine racine, il est recommandé d'ajouter un utilisateur IAM ou un SSO en tant user/group que propriétaire. Pour plus d'informations, consultez la section Gestion des utilisateurs dans le manuel Amazon DataZone Administrator Guide.
  - Important : les rôles IAM ne peuvent pas se connecter à Amazon SageMaker Unified Studio.

- Comptes associés et modifications apportées au AWS Resource Access Manager (AWS RAM) :
  - Les comptes associés utilisent des partages de ressources provenant de la AWS RAM pour autoriser les actions d'API depuis le compte de domaine racine.
  - Le processus de mise à niveau modifie les autorisations gérées sous-jacentes pour le partage de AWS RAM créé et géré par Amazon DataZone. Les autorisations gérées concernées sont `AWSRAMPermissionsAmazonDatazoneDomainExtendedServiceAccess` et`AWSRAMPermissionsAmazonDatazoneDomainExtendedServiceWithPortalAccess`.
- Modifications apportées à l'abonnement Amazon Q : le niveau d'abonnement Amazon Q sera réglé par défaut sur le niveau gratuit pour le domaine mis à niveau. Les administrateurs de domaine peuvent modifier cela une fois la mise à niveau du domaine terminée.
- Après la mise à niveau, l'`domainVersion`attribut du domaine passe de `V1` à`V2`.

Mettez à niveau votre DataZone domaine Amazon vers un domaine SageMaker unifié Amazon

Vous pouvez suivre la procédure suivante pour mettre à niveau votre DataZone domaine Amazon vers un domaine SageMaker unifié Amazon.

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et utilisez le sélecteur de région dans la barre de navigation supérieure pour choisir la région appropriée. AWS
2. Choisissez le DataZone domaine Amazon que vous souhaitez mettre à niveau et accédez à sa page de détails.
3. Sur la page de détails du domaine, cliquez sur le bouton Commencer situé dans la notification Mettre votre domaine à niveau vers Amazon SageMaker Unified Studio.
4. Sur la page Mettre à niveau votre domaine vers Amazon SageMaker Unified Studio, choisissez Start.
5. Spécifiez ensuite le rôle d'exécution du domaine et les rôles de service de domaine pour le domaine et les propriétaires de l'unité de domaine racine si le DataZone domaine Amazon que vous mettez à niveau n'a pas de propriétaires de type utilisateur IAM, utilisateur/groupe SSO. Choisissez ensuite Mettre à niveau le domaine.

Questions fréquemment posées sur la mise à niveau de DataZone domaines Amazon vers des domaines SageMaker unifiés Amazon

- Quelles propriétés et configurations seront transférées au domaine après la mise à niveau ?

  Toutes les propriétés configurées sur le DataZone domaine Amazon sont transférées vers le domaine SageMaker unifié Amazon mis à niveau. Cela inclut les propriétés de chiffrement des données, les propriétés des applications d'authentification, etc.

- Dois-je configurer à nouveau l'accès à authentification unique (SSO) pour mes utilisateurs ?

  Non. Votre application SSO IAM Identity Center associée au domaine sera transférée vers le domaine SageMaker unifié Amazon mis à niveau. En outre, tout utilisateur ou rôle IAM attribué au domaine sera disponible dans le domaine SageMaker unifié Amazon mis à niveau.

- Puis-je toujours utiliser le DataZone portail Amazon après la mise à niveau ?

  Oui. Après la mise à niveau, le DataZone portail Amazon et Amazon SageMaker Unified Studio seront disponibles pour que les utilisateurs finaux puissent interagir avec eux. Les deux portails resteront ouverts jusqu'à ce qu'un administrateur de domaine désactive le DataZone portail Amazon depuis la console de SageMaker gestion Amazon.

- Est-ce que je verrai les projets et autres entités créés sur le DataZone portail Amazon dans Amazon SageMaker Unified Studio ?

  Oui. La plupart des entités (projets, formulaires de métadonnées, glossaires, unités de domaine) créées via le DataZone portail Amazon seront visibles dans Amazon SageMaker Unified Studio. Les projets reprendront tous les actifs, les formulaires de métadonnées et les glossaires associés aux actifs, les abonnements aux actifs, les membres, etc. Ces projets nécessitent d'interroger les données à partir des éditeurs de requêtes AWS Athena ou Amazon Redshift. Les formulaires de métadonnées et les glossaires apparaîtront dans Amazon SageMaker Unified Studio. Ils peuvent être modifiés depuis Amazon SageMaker et affectés à des ressources issues de projets créés via Amazon SageMaker. Les environnements et les profils d'environnement d'Amazon ne DataZone s'afficheront pas dans Amazon SageMaker Unified Studio. Ces entités ont été remplacées par des profils de SageMaker projet Amazon. Les projets créés dans Amazon SageMaker Unified Studio ne seront pas visibles via le DataZone portail Amazon.

- Qu'advient-il de l'identifiant de domaine et des identifiants de projet après la mise à niveau vers le domaine SageMaker unifié Amazon ?

Tous les identifiants d'entité, y compris le domaine et les projets, resteront les mêmes après la mise à niveau.

- Mes piles AWS CloudFormation (CFN) continueront-elles de fonctionner pour le domaine SageMaker unifié Amazon récemment mis à niveau ?

  Amazon SageMaker Unified Studio utilise la même solution APIs qu'Amazon DataZone. Cependant, certaines modifications de la logique des modèles CFN seront nécessaires. Par exemple, les domaines Amazon DataZone sont distingués des domaines SageMaker unifiés Amazon par un attribut nommé DomainVersion (valeurs V1 | V2).
- Que se passe-t-il lorsque la mise à niveau est annulée ?
  - L'annulation de la mise à niveau fait passer la version du domaine de V2 à V1. Amazon SageMaker Unified Studio ne sera plus accessible. La vue console du domaine reviendra à la DataZone vue Amazon. Les ressources créées avant l'annulation seront conservées tant qu'elles ne sont pas liées à un projet créé à partir d'Amazon SageMaker Unified Studio. La restauration n'est autorisée que lorsqu'aucun projet créé depuis Amazon SageMaker Unified Studio n'est présent.
  - Les paramètres tels que l'abonnement AWS Q seront également conservés après l'annulation.
  - S' VPCs ils ont été créés pour être utilisés par Amazon SageMaker, ils persisteront après la restauration. Les VPC créés par le SageMaker service porteront le tag suivant : Nom = VPC SageMakerUnifiedStudio
  - L'autorisation gérée dans le cadre du partage de ressources RAM ne sera pas annulée. L'autorisation gérée est un sur-ensemble d'Amazon DataZone et d'Amazon SageMaker Unified Studio.
  - Un domaine qui a été annulé peut à nouveau être mis à niveau vers le domaine SageMaker unifié Amazon.

Unités de domaine et politiques d'autorisation sur Amazon DataZone

Utilisez les unités de domaine pour organiser facilement vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour configurer un partage de données sécurisé et efficace au sein et entre les unités commerciales de votre organisation, créez des unités de domaine au sein d'Amazon DataZone et permettez à des utilisateurs sélectionnés au sein de chaque unité commerciale de se connecter et de partager leurs actifs dans le catalogue. Les utilisateurs de n'importe quel endroit de l'entreprise peuvent facilement rechercher des actifs relevant de ces unités commerciales et demander l'accès à ces actifs.

Les unités de domaine peuvent également être utilisées pour permettre aux propriétaires de ressources, tels que les propriétaires de AWS comptes, de configurer des DataZone autorisations Amazon sur leurs ressources. Les unités de domaine fournissent une autorité déléguée des propriétaires de comptes aux propriétaires d'unités de domaine et elles peuvent configurer des autorisations d'autorisation sur les profils d'environnement (créés à l'aide de configurations de plan), au nom des propriétaires de comptes. Cela vous permet de limiter les personnes autorisées à créer et à utiliser quels profils d'environnement en fonction des unités commerciales auxquelles elles appartiennent. Les DataZone autorisations Amazon peuvent également être utilisées pour appliquer les normes relatives aux métadonnées et permettre uniquement à des projets sélectionnés de créer des formulaires de métadonnées et un glossaire. Cela peut aider à maintenir des métadonnées cohérentes et de qualité. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Au sein d'une unité de DataZone domaine Amazon, vous pouvez attribuer les politiques d'autorisation suivantes à vos utilisateurs et groupes afin de leur accorder des autorisations spécifiques :

- Politique de création d'unités de domaine
- Politique de création de projets
- Politique d'adhésion au projet
- Politique d'hypothèse de propriété des unités de domaine
- Politique d'hypothèse de propriété du projet

Pour de plus amples informations, veuillez consulter Attribuer des politiques d'autorisation aux utilisateurs et aux groupes au sein d'une unité DataZone de domaine Amazon.

Au sein d'une unité de DataZone domaine Amazon, vous pouvez attribuer les politiques d'autorisation suivantes à vos projets afin de leur accorder des autorisations spécifiques :

- Politique de création de glossaires
- Politique de création de formulaires de métadonnées
- Politique de création de types d'actifs personnalisés

Pour de plus amples informations, veuillez consulter Attribuer des politiques d'autorisation aux projets au sein d'une unité DataZone de domaine Amazon.

Une autre façon d'utiliser le mécanisme d'autorisation d'Amazon DataZone consiste à appliquer des politiques d'autorisation aux projets et aux propriétaires d'unités de domaine dans le cadre des configurations Amazon DataZone Blueprint.

Une configuration Amazon DataZone Blueprint est une entité qui encapsule les informations nécessaires à la création et à la configuration des ressources utilisées pour la publication et l'abonnement des flux de travail utilisateur. Ces informations incluent le numéro de AWS compte et la région, les CloudFormation modèles, les paramètres au niveau du compte tels que les sous-réseaux, VPCs et peuvent également contenir des informations de connexion à la base de données et des informations d'identification. Pour contrôler les coûts et améliorer la sécurité, les utilisateurs de plateformes de données doivent être en mesure de contrôler qui peut utiliser ces plans et créer des environnements.

Dans une configuration de plan spécifique, vous pouvez attribuer les politiques d'autorisation suivantes aux propriétaires de projets et d'unités de domaine :

- Créez des profils d'environnement à l'aide de ce plan. Cette politique peut être attribuée aux DataZone projets Amazon et les autorise à créer des profils d'environnement à l'aide de ce plan.
- Accordez des autorisations pour créer des profils d'environnement à l'aide de ce plan : cette politique peut être attribuée aux propriétaires d'unités de domaine et les autorise à accorder des autorisations aux projets pour créer des profils d'environnement à l'aide de ce plan.

Pour de plus amples informations, veuillez consulter Attribuez des politiques d'autorisation dans les configurations Amazon DataZone Blueprint.

Rubriques

- Création d'unités de domaine sur Amazon DataZone

- Modifier les unités de domaine dans Amazon DataZone
- Supprimer des unités de domaine sur Amazon DataZone
- Gérez les propriétaires d'unités de domaine sur Amazon DataZone
- Attribuer des politiques d'autorisation aux utilisateurs et aux groupes au sein d'une unité DataZone de domaine Amazon
- Attribuer des politiques d'autorisation aux projets au sein d'une unité DataZone de domaine Amazon
- Attribuez des politiques d'autorisation dans les configurations Amazon DataZone Blueprint

Création d'unités de domaine sur Amazon DataZone

Sur Amazon DataZone, les unités de domaine vous permettent d'organiser vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour créer une unité de domaine

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez créer des unités de domaine.
3. Sur la page des détails du domaine, accédez à l'onglet Unités de domaine.
4. Choisissez Créer une unité de domaine.
5. Spécifiez ce qui suit, puis choisissez Create domain unit :

    - Sous Détails de l'unité de domaine, dans Nom, spécifiez le nom de l'unité de domaine.
    - Sous Détails de l'unité de domaine, dans Description, spécifiez la description de l'unité de domaine.
    - Unité de domaine parent : choisissez l'unité de domaine parent sous laquelle vous souhaitez ajouter la nouvelle unité de domaine.
    - Propriétaires d'unités de domaine : spécifiez les propriétaires d'unités de domaine autorisés à modifier cette unité de domaine.

Modifier les unités de domaine dans Amazon DataZone

Sur Amazon DataZone, les unités de domaine vous permettent d'organiser vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour modifier une unité de domaine

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez modifier les unités de domaine.
3. Sur la page des détails du domaine, accédez à l'onglet Unités de domaine et choisissez l'unité de domaine que vous souhaitez modifier.
4. Développez les actions et choisissez Modifier l'unité de domaine.
5. Apportez vos modifications au nom et à la description de l'unité de domaine, puis choisissez Enregistrer les modifications.

Supprimer des unités de domaine sur Amazon DataZone

Sur Amazon DataZone, les unités de domaine vous permettent d'organiser vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour modifier une unité de domaine

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez supprimer des unités de domaine.

3. Sur la page des détails du domaine, accédez à l'onglet Unités de domaine et choisissez l'unité de domaine que vous souhaitez supprimer.
4. Développez les actions et choisissez Supprimer l'unité de domaine.
5. Dans la fenêtre contextuelle Supprimer l'unité de domaine, confirmez la suppression en choisissant Supprimer l'unité de domaine.

Gérez les propriétaires d'unités de domaine sur Amazon DataZone

Sur Amazon DataZone, les unités de domaine vous permettent d'organiser vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour ajouter des propriétaires à l'unité de domaine de premier niveau via la console DataZone de gestion Amazon, procédez comme suit.

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines et choisissez le DataZone domaine Amazon dans lequel vous souhaitez ajouter les propriétaires d'unités de domaine.
3. Sur la page des détails du domaine, accédez au tableau des propriétaires racines du domaine.
4. Choisissez Ajouter, puis spécifiez les utilisateurs que vous souhaitez désigner comme propriétaires d'unités de domaine. Choisissez Ajouter le (s) propriétaire (s) du domaine racine.

Pour ajouter des propriétaires d'unités de domaine via Amazon DataZone Data Portal, suivez la procédure suivante :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Choisissez Afficher les domaines, puis choisissez le domaine et l'unité de domaine auxquels vous souhaitez ajouter les propriétaires d'unités de domaine.
3. Sur la page des détails de l'unité de domaine, choisissez l'onglet Propriétaires, puis sélectionnez Ajouter des propriétaires.

4. Dans la fenêtre contextuelle Ajouter des propriétaires d'unités de domaine, spécifiez les utilisateurs que vous souhaitez désigner comme propriétaires d'unités de domaine, puis choisissez Ajouter des propriétaires.

Attribuer des politiques d'autorisation aux utilisateurs et aux groupes au sein d'une unité DataZone de domaine Amazon

Sur Amazon DataZone, les unités de domaine vous permettent d'organiser vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Dans une unité de DataZone domaine Amazon, vous pouvez attribuer les politiques d'autorisation suivantes à vos utilisateurs et groupes afin de leur accorder diverses autorisations d'autorisation au sein de cette unité de domaine :

- Politique de création d'unités de domaine
- Politique de création de projets
- Politique d'adhésion au projet
- Politique d'hypothèse de propriété des unités de domaine
- Politique d'hypothèse de propriété du projet

Pour attribuer des politiques d'autorisation aux utilisateurs et aux groupes au sein d'une unité de domaine, procédez comme suit :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Afficher les domaines, puis choisissez le domaine et l'unité de domaine auxquels vous souhaitez attribuer des politiques d'autorisation.
3. Sur la page des détails de l'unité de domaine, choisissez la politique d'autorisation à laquelle vous souhaitez attribuer, users/groups puis choisissez Ajouter des utilisateurs.
4. Dans la fenêtre contextuelle Ajouter des utilisateurs, effectuez l'une des opérations suivantes :

- Choisissez Utilisateurs et groupes sélectionnés, spécifiez les utilisateurs et les groupes auxquels vous souhaitez attribuer la politique d'autorisation sélectionnée, puis choisissez Ajouter des utilisateurs.
- Choisissez Tous les utilisateurs, puis sélectionnez Ajouter des utilisateurs.
- Choisissez Tous les groupes, puis sélectionnez Ajouter des utilisateurs.

5. Vous pouvez également activer ou désactiver les autorisations en cascade de la politique d'autorisation sélectionnée pour les utilisateurs sélectionnés. Pour ce faire, choisissez le ou les utilisateurs pour lesquels vous souhaitez activer les autorisations en cascade, puis développez Actions, puis choisissez Définir les autorisations en cascade sur true. Les utilisateurs sélectionnés bénéficieront des autorisations accordées par cette politique dans toutes les unités de domaine enfants de cette unité de domaine. Vous pouvez également choisir le ou les utilisateurs pour lesquels vous souhaitez désactiver les autorisations en cascade, puis développer le champ Actions et définir les autorisations en cascade sur False.

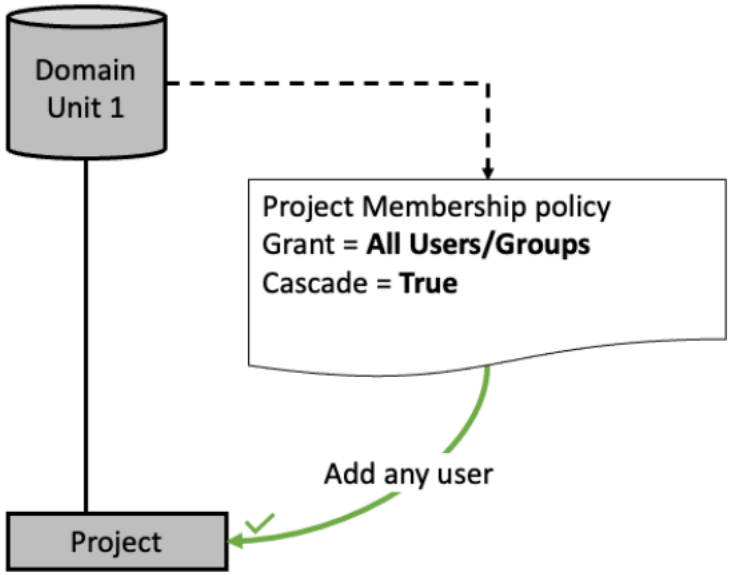
Politique d'adhésion au projet dans la hiérarchie des unités de domaine sur Amazon DataZone

La politique d'adhésion au projet définit les individus ou les groupes éligibles pour être ajoutés en tant que membres aux projets au sein d'une unité de domaine. Cette rubrique décrit les scénarios d'impact de la politique par rapport à une unité de domaine individuelle et à des unités de domaine dans une structure hiérarchique.

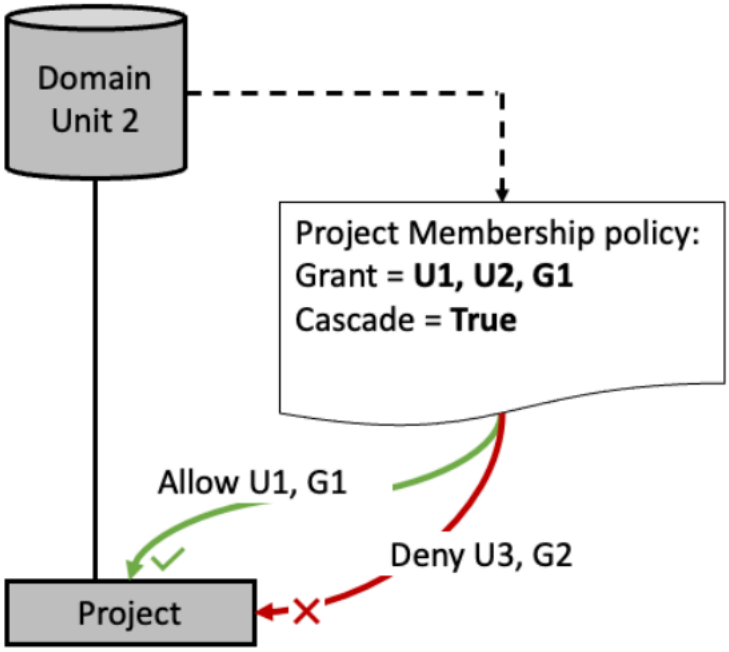
Il est important de noter plusieurs concepts utilisés dans cette rubrique :

- Pool de membres : les principaux (utilisateurs ou groupes) auxquels l'accès est accordé par le biais de la politique d'adhésion du projet sont considérés comme faisant partie du pool de membres du projet. Par exemple, si la politique d'unité de domaine DU1 est accordée aux utilisateurs U1 et U2, ainsi qu'au groupe d'authentification unique (SSO) G1, le pool de membres du projet pour DU1 serait composé de {U1, U2, G1}.
- Cascade : possibilité de transmettre l'allocation à toutes les unités de domaine enfants connectées via la hiérarchie des unités de domaine.
- Accorder : autorisation permettant à un utilisateur ou à un groupe d'effectuer une action.

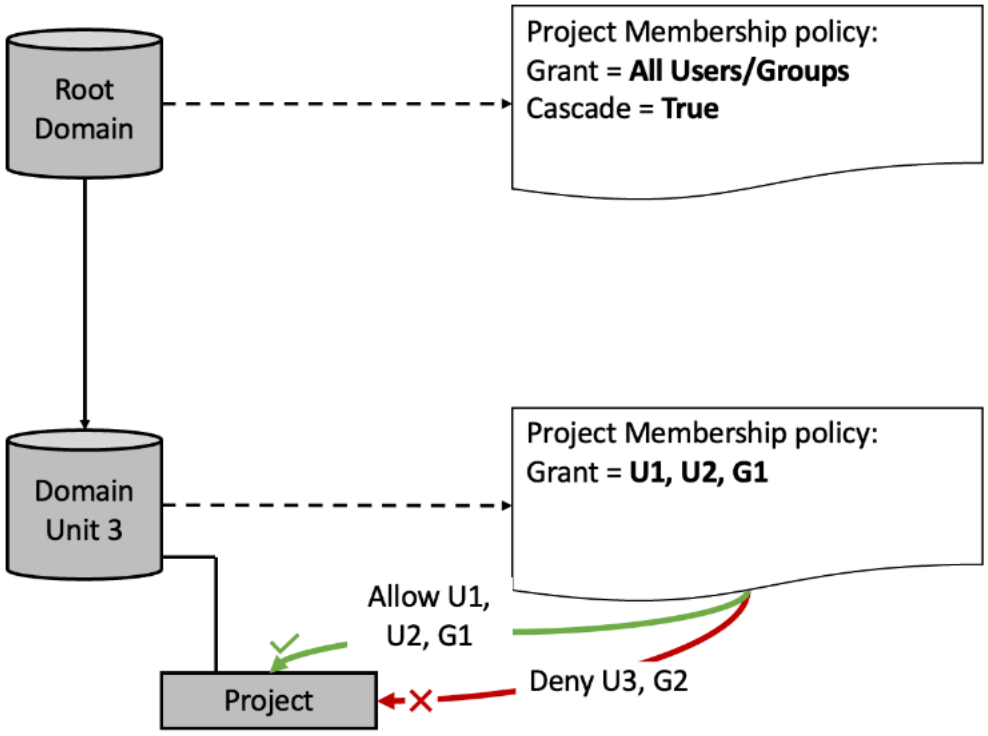
Scénario 1 : n'importe quel utilisateur ou groupe peut être ajouté au projet sous l'unité de domaine 1, car le pool de membres est composé de {Tous les utilisateurs/groupes}.



Scénario 2 - Les utilisateurs {U1, G1} peuvent être ajoutés au projet sous l'unité de domaine 2 puisqu'ils font partie du pool de membres sous l'unité de domaine 2. Les utilisateurs {U3, G2} ne peuvent être ajoutés à aucun projet car ils ne font pas partie du pool de membres.

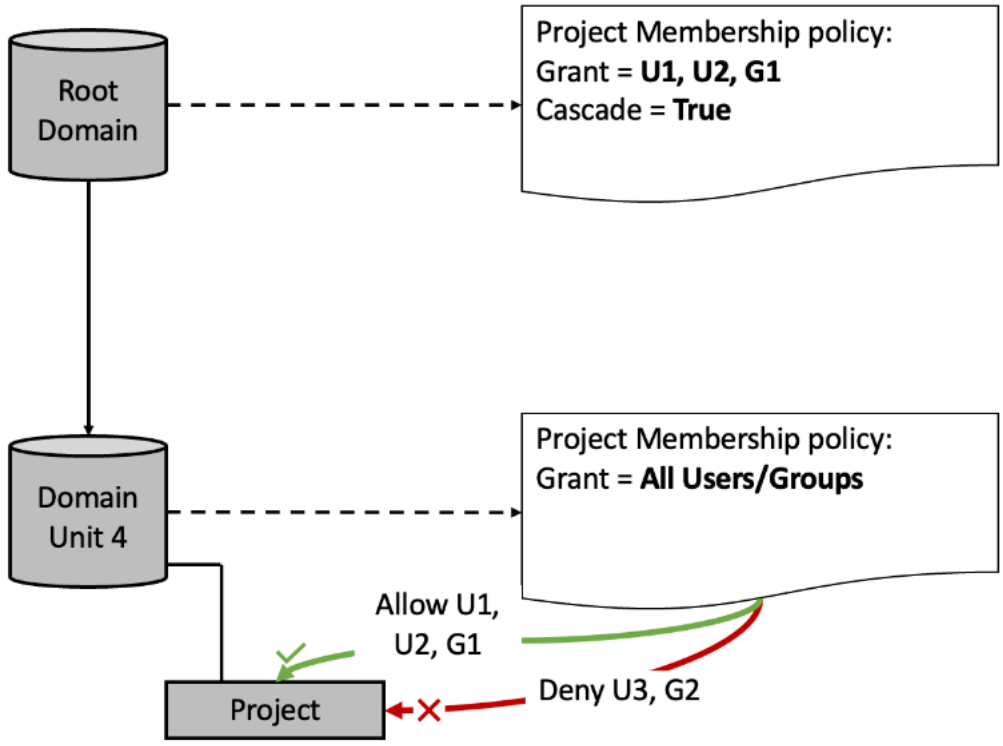


Scénario 3 - Intersection des groupes de membres : lorsqu'il existe des groupes de membres à différents niveaux hiérarchiques d'unités de domaine, seuls les utilisateurs et les groupes figurant dans tous les groupes de membres peuvent être ajoutés au projet.



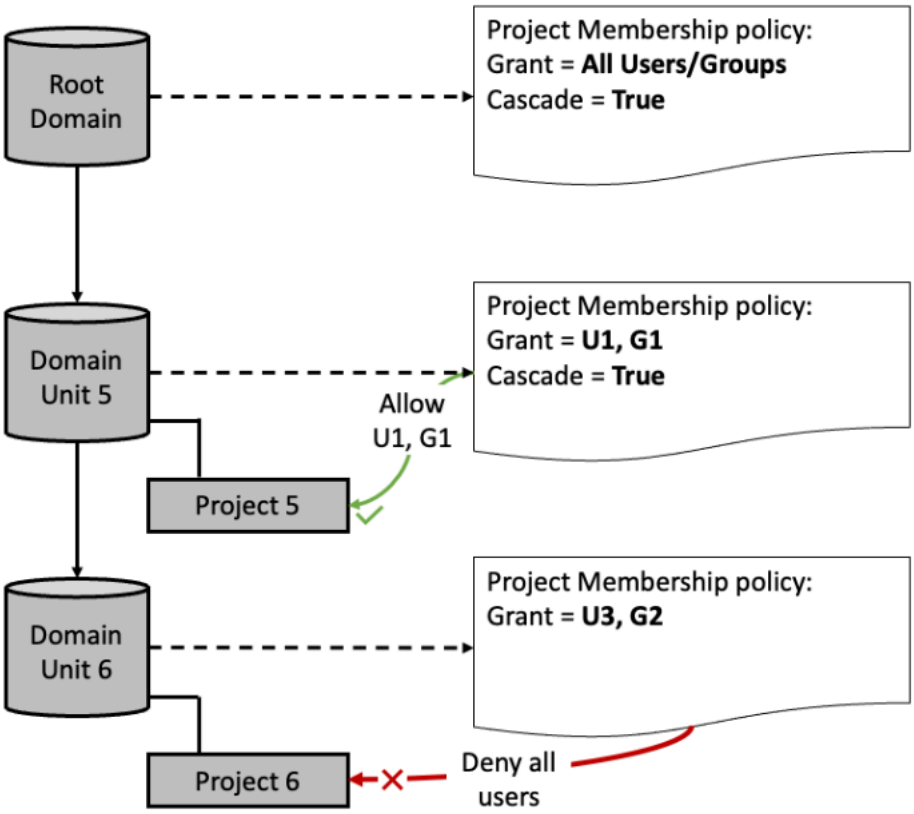
- L'intersection des utilisateurs des deux groupes de membres est {U1, U2, G1}.
- Les utilisateurs {U1, U2, G1} peuvent être ajoutés au projet sous l'unité de domaine 3.
- Les utilisateurs {U3, G2} ne peuvent pas être ajoutés au projet sous l'unité de domaine 3, même si tous les utilisateurs et tous les groupes figurent dans le pool de membres au niveau de l'unité de domaine racine.

Scénario 4 - Intersection des groupes de membres : lorsqu'il existe des groupes de membres à différents niveaux hiérarchiques d'unités de domaine, seuls les utilisateurs et les groupes figurant dans tous les groupes de membres peuvent être ajoutés au projet.

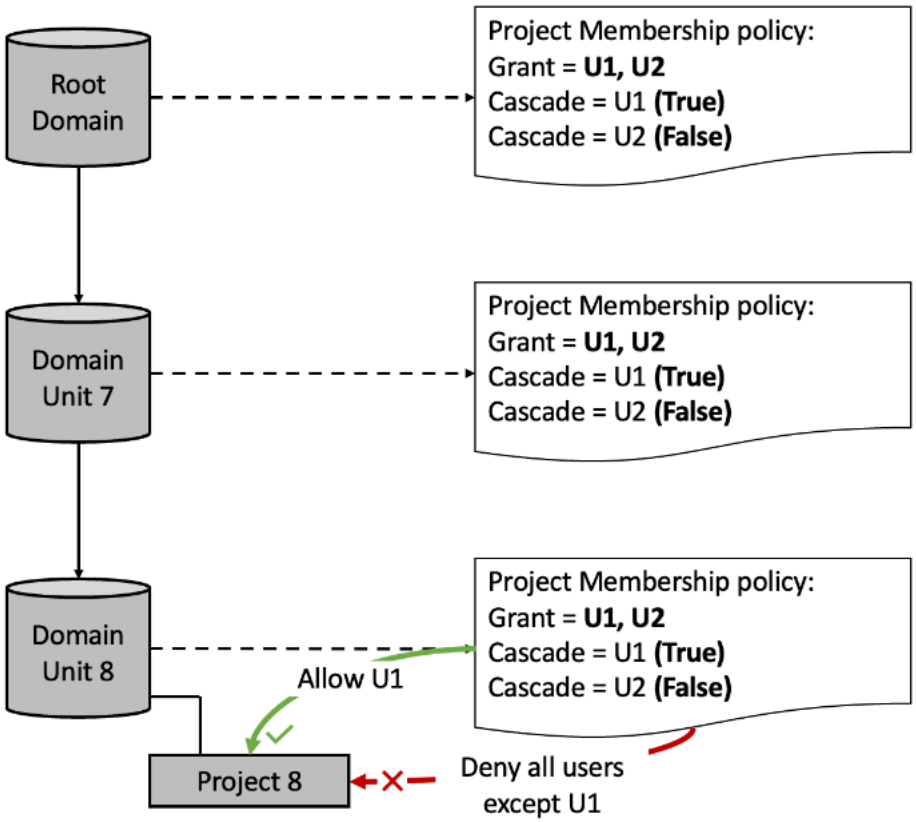


- L'intersection des utilisateurs des deux groupes de membres est {U1, U2, G1}.
- Le pool de membres de l'unité de domaine 4 est {Tous les utilisateurs/groupes} mais le pool de membres ne peut pas s'étendre au-delà du pool de membres du domaine racine {U1, U2, G1}.
- Les utilisateurs {U3, G2} ne peuvent pas être ajoutés au projet sous l'unité de domaine 4, même si tous les utilisateurs et tous les groupes figurent dans le pool de membres de l'unité de domaine 4.

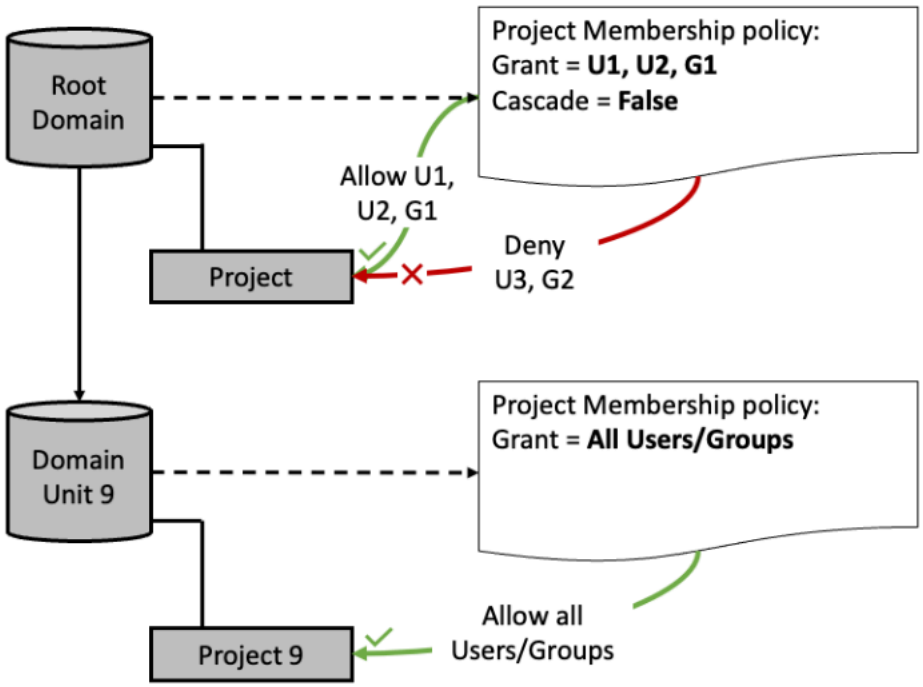
Scénario 5 - Les utilisateurs {U1, G1} peuvent être ajoutés au projet 5 car ils font partie de l'intersection des pools de membres entre le domaine racine et l'unité de domaine 5. Non user/group peut être ajouté à Project 6 car l'intersection des trois pools de membres est vide.



Scénario 6 - L'intersection entre les trois groupes de membres signifie que seul l'utilisateur {U1} peut être ajouté à Project 8. Les pools d'intersections de l'unité de domaine 8 sont {U1}, {U1}, {U1, U2}, seul {U1} étant commun aux trois.



Scénario 7 - Les utilisateurs {U1, U2, G1} peuvent être ajoutés au projet du domaine racine car ils font partie du pool de membres du domaine racine. Tout utilisateur ou groupe peut être ajouté au projet sous l'unité de domaine 9, car le pool de membres est composé de {Tous les utilisateurs/groupes}, car la cascade est définie sur false dans le domaine racine situé au-dessus.



Attribuer des politiques d'autorisation aux projets au sein d'une unité DataZone de domaine Amazon

Sur Amazon DataZone, les unités de domaine vous permettent d'organiser vos actifs et autres entités de domaine au sein d'unités commerciales et d'équipes spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Dans une unité de DataZone domaine Amazon, vous pouvez attribuer les politiques d'autorisation suivantes à vos projets afin d'accorder à ces entités diverses autorisations d'autorisation au sein de cette unité de domaine :

- Politique de création de glossaires
- Politique de création de formulaires de métadonnées
- Politique de création de types d'actifs personnalisés

Pour attribuer des politiques d'autorisation à des projets au sein d'une unité de domaine, procédez comme suit :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un

DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.

2. Choisissez Gérer le domaine, puis choisissez le domaine et l'unité de domaine auxquels vous souhaitez attribuer des politiques d'autorisation.
3. Sur la page des détails de l'unité de domaine, choisissez la politique d'autorisation que vous souhaitez attribuer et cliquez pour configurer.

Attribuez des politiques d'autorisation dans les configurations Amazon DataZone Blueprint

Une autre façon d'utiliser le mécanisme d'autorisation d'Amazon DataZone consiste à appliquer des politiques d'autorisation aux projets et aux propriétaires d'unités de domaine dans le cadre des configurations Amazon DataZone Blueprint.

Une configuration Amazon DataZone Blueprint est une entité qui encapsule les informations nécessaires à la création et à la configuration des ressources utilisées pour la publication et l'abonnement des flux de travail utilisateur. Ces informations incluent le numéro de AWS compte et la région, les modèles CFN, les paramètres au niveau du compte tels que les sous-réseaux VPCs et peuvent également contenir des informations de connexion à la base de données et des informations d'identification. Pour contrôler les coûts et améliorer la sécurité, les utilisateurs de plateformes de données doivent être en mesure de contrôler qui peut utiliser ces plans et créer des environnements.

Dans une configuration de plan spécifique, vous pouvez attribuer les politiques d'autorisation suivantes aux propriétaires de projets et d'unités de domaine :

- Créez des profils d'environnement à l'aide de ce plan. Cette politique peut être attribuée aux DataZone projets Amazon et les autorise à créer des profils d'environnement à l'aide de ce plan.
- Accordez des autorisations pour créer des profils d'environnement à l'aide de ce plan : cette politique peut être attribuée aux propriétaires d'unités de domaine et les autorise à accorder des autorisations aux projets pour créer des profils d'environnement à l'aide de ce plan.

Attribuez la politique d'autorisation de création de profils d'environnement à l'aide de ce plan aux projets à partir d'une configuration de plan via le portail de données Amazon DataZone

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de données, choisissez le domaine dans lequel le plan est activé avec lequel vous souhaitez travailler, puis accédez à l'onglet Configurations du plan.
3. Dans l'onglet Configurations du plan, choisissez le plan activé avec lequel vous souhaitez travailler, puis dans la page de détails de ce plan, accédez à l'onglet Politiques d'autorisation, puis choisissez la politique d'autorisation Créer des profils d'environnement à l'aide de ce plan.
4. Dans la page de détails de la politique d'autorisation Créer des profils d'environnement à l'aide de ce plan, développez Actions et choisissez Ajouter des projets.
5. Dans la fenêtre contextuelle Ajouter des projets, vous pouvez effectuer l'une des opérations suivantes :

   - Choisissez l'option Tous les projets d'une unité de domaine, puis recherchez et spécifiez les unités de domaine qui contiennent les projets que vous souhaitez autoriser à créer des profils d'environnement avec ce plan, puis choisissez Ajouter des projets.
   - Choisissez l'option Projets sélectionnés dans une unité de domaine, puis recherchez et spécifiez les unités de domaine qui contiennent les projets auxquels vous souhaitez attribuer cette politique, puis recherchez et choisissez les projets auxquels vous souhaitez attribuer cette politique, puis choisissez Ajouter des projets.

Attribuez les autorisations Grant pour créer des profils d'environnement à l'aide de cette politique d'autorisation de plan aux propriétaires d'unités de domaine à partir d'une configuration de plan via la console de gestion Amazon DataZone .

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Dans la DataZone console Amazon, choisissez le domaine dans lequel le plan est activé avec lequel vous souhaitez travailler, puis accédez à l'onglet Blueprints.
3. Dans l'onglet Blueprints, choisissez le plan activé avec lequel vous souhaitez travailler, puis dans la page de détails du plan, accédez à l'onglet Autorisations déléguées.

4. Dans l'onglet Autorisations déléguées, recherchez et choisissez les unités de domaine aux propriétaires desquelles vous souhaitez attribuer les autorisations d'octroi pour créer des profils d'environnement à l'aide de cette politique de modèle, puis choisissez Ajouter une autorisation déléguée.

Blueprints DataZone intégrés à Amazon

Un plan avec lequel un environnement est créé définit les outils et services que les membres du projet auquel appartient l'environnement peuvent utiliser lorsqu'ils travaillent avec les actifs du DataZone catalogue Amazon. Dans la version actuelle d'Amazon DataZone, il existe les plans intégrés suivants :

- Plan du lac de données
- Plan d'entrepôt de données
- SageMaker Plan Amazon

Vous pouvez suivre les étapes des procédures suivantes pour activer les plans par défaut dans Amazon DataZone :

- Activez les plans intégrés dans le AWS compte propriétaire du domaine Amazon DataZone
- Ajoutez Amazon SageMaker en tant que service de confiance dans le AWS compte propriétaire du DataZone domaine Amazon

Activez les plans intégrés dans le AWS compte propriétaire du domaine Amazon DataZone

Un plan avec lequel un environnement est créé définit les outils et services que les membres du projet auquel appartient l'environnement peuvent utiliser lorsqu'ils travaillent avec les actifs du DataZone catalogue Amazon.

Dans la version actuelle d'Amazon DataZone, plusieurs plans sont intégrés : le plan du lac de données, le plan de l'entrepôt de données et le plan Amazon. SageMaker

- Le plan du lac de données contient la définition du lancement et de la configuration d'un ensemble de services (AWS Glue, AWS Lake Formation, Amazon Athena) pour publier et utiliser les actifs des lacs de données dans le catalogue Amazon DataZone .
- Le plan d'entrepôt de données contient la définition du lancement et de la configuration d'un ensemble de services (Amazon Redshift) pour publier et utiliser les actifs Amazon Redshift dans le catalogue Amazon. DataZone

- Le SageMaker plan Amazon contient la définition du lancement et de la configuration d'un ensemble de services (Amazon SageMaker Studio) pour publier et utiliser les SageMaker ressources Amazon dans le DataZone catalogue Amazon.

Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Lors de la création d'un DataZone domaine Amazon, vous avez la possibilité de choisir la configuration rapide qui active automatiquement le lac de données par défaut et les plans intégrés de l'entrepôt de données par défaut dans le cadre du processus de création du domaine. La configuration rapide crée également des profils d'environnement par défaut et des environnements par défaut pour vous à l'aide de ces plans intégrés.

Si vous ne choisissez pas Configuration rapide lors de la création de votre DataZone domaine Amazon, vous pouvez utiliser la procédure ci-dessous pour activer les plans intégrés disponibles dans le AWS compte hébergeant ce DataZone domaine Amazon. Vous devez activer ces plans intégrés avant de pouvoir les utiliser pour créer des profils d'environnement et des environnements dans ce domaine.

Pour activer les plans intégrés dans un DataZone domaine Amazon via la console de DataZone gestion Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales.

Activer les plans intégrés dans un domaine Amazon DataZone

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez activer un ou plusieurs plans intégrés.
3. Sur la page des détails du domaine, accédez à l'onglet Blueprints.
4. Dans la liste des plans, choisissez le plan DefaultDataLakeou DefaultDataWarehousele SageMaker plan Amazon.
5. Sur la page de détails du plan choisi, choisissez Activer dans ce compte.
6. Sur la page Autorisations et ressources, spécifiez les éléments suivants :

    - Si vous activez le DefaultDataLakeplan, pour le rôle Glue Manage Access, spécifiez un rôle de service nouveau ou existant qui DataZone autorise Amazon à ingérer et à gérer l'accès aux tables dans AWS Glue and AWS Lake Formation.

- Si vous activez le DefaultDataWarehouseplan, pour le rôle Redshift Manage Access, spécifiez un rôle de service nouveau ou existant qui autorise DataZone Amazon à ingérer et à gérer l'accès aux partages de données, aux tables et aux vues dans Amazon Redshift.
- Si vous activez le SageMaker plan Amazon, pour le rôle de SageMaker gestion des accès, spécifiez un rôle de service nouveau ou existant qui accorde à Amazon l' DataZoneautorisation de publier les SageMaker données Amazon dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux ressources SageMaker publiées par Amazon dans le catalogue.

  > **Important**
  >
  > Lorsque vous activez le SageMaker plan Amazon, Amazon DataZone vérifie si les rôles IAM suivants pour Amazon DataZone existent dans le compte et la région actuels. Si ces rôles n'existent pas, Amazon les crée DataZone automatiquement.
  >
  > - AmazonDataZoneGlueAccess- <region>- <domainId>
  > - AmazonDataZoneRedshiftAccess- <region>- <domainId>

- Pour le rôle de provisionnement, spécifiez un rôle de service nouveau ou existant qui accorde à Amazon DataZone l'autorisation de créer et de configurer les ressources de l'environnement AWS CloudFormation à l'aide du compte et de la région d'environnement.
- Si vous activez le SageMaker plan Amazon, pour le compartiment Amazon S3 pour la source de données SageMaker -Glue, spécifiez un compartiment Amazon S3 qui doit être utilisé par tous les SageMaker environnements du AWS compte. Le préfixe de compartiment que vous spécifiez doit être l'un des suivants :
  - zone de données Amazon*
  - créateur de zones de données*
  - zone de données SageMaker*
  - DataZone- Sagemaker*
  - Sagemaker- * DataZone
  - DataZone-SageMaker*
  - SageMaker-DataZone*

7. Choisissez Activer le plan.

Une fois que vous avez activé le ou les plans choisis, vous pouvez contrôler quels projets peuvent utiliser les plans dans votre compte pour créer des profils d'environnement. Vous pouvez le faire en affectant la gestion des projets à la configuration du plan.

> **Important**
>
> Par défaut, aucun projet de gestion n'est spécifié pour les plans d'environnement, ce qui signifie que tout DataZone utilisateur d'Amazon peut créer des profils pour un plan d'environnement. Par conséquent, il est fortement recommandé de toujours spécifier la gestion des projets pour vos plans d'environnement afin de garantir une gouvernance plus solide.

Spécifiez la gestion des projets sur les plans activés

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines, puis choisissez le domaine dans lequel vous souhaitez ajouter le ou les projets de gestion pour le ou les plans choisis.
3. Choisissez l'onglet Blueprints, puis choisissez le plan avec lequel vous souhaitez travailler.
4. Par défaut, tous les projets du domaine peuvent utiliser les DefaultDataLake SageMaker plans ou Amazon du compte pour créer des profils d'environnement. DefaultDataWareshouse Toutefois, vous pouvez limiter cela en affectant la gestion des projets aux plans. Pour ajouter des projets de gestion, choisissez Sélectionner la gestion du projet, puis choisissez les projets que vous souhaitez ajouter en tant que gestion de projets dans le menu déroulant, puis sélectionnez Sélectionner la gestion de projets.

Une fois que vous avez activé le DefaultDataWarehouse plan dans votre AWS compte, vous pouvez ajouter des ensembles de paramètres à la configuration du plan. Un ensemble de paramètres est un groupe de clés et de valeurs, requis pour qu'Amazon DataZone établisse une connexion à votre cluster Amazon Redshift et est utilisé pour créer des environnements d'entrepôt de données. Ces paramètres incluent le nom de votre cluster Amazon Redshift, de votre base de données et le AWS secret contenant les informations d'identification du cluster.

Ajouter des ensembles de paramètres au DefaultDataWarehouse plan

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines, puis choisissez le domaine dans lequel vous souhaitez ajouter le jeu de paramètres.
3. Choisissez l'onglet Plans, puis choisissez le DefaultDataWareshouse plan pour ouvrir la page de détails du plan.
4. Dans l'onglet Ensembles de paramètres de la page de détails du plan, choisissez Créer un jeu de paramètres.

   - Entrez un nom pour le jeu de paramètres.
   - Fournissez éventuellement une description du jeu de paramètres.
   - Sélectionner une région
   - Sélectionnez le cluster Amazon Redshift ou Amazon Redshift Serverless.
   - Sélectionnez l'ARN AWS secret qui contient les informations d'identification du cluster Amazon Redshift sélectionné ou du groupe de travail Amazon Redshift Serverless. Le AWS secret doit être étiqueté avec le `AmazonDataZoneDomain : [Domain_ID]` tag afin de pouvoir être utilisé dans un ensemble de paramètres.
     - Si vous n'avez pas de AWS secret existant, vous pouvez également en créer un nouveau en choisissant Créer un nouveau AWS secret. Cela ouvre une boîte de dialogue dans laquelle vous pouvez fournir le nom du secret, le nom d'utilisateur et le mot de passe. Une fois que vous avez choisi Create New AWS Secret, Amazon DataZone crée un nouveau secret dans le service AWS Secrets Manager et s'assure que le secret est étiqueté avec le domaine dans lequel vous essayez de créer le jeu de paramètres.
   - Si vous avez choisi le cluster Amazon Redshift à l'étape ci-dessus, choisissez maintenant un cluster dans le menu déroulant. Si vous avez choisi le groupe de travail Amazon Redshift à l'étape ci-dessus, choisissez maintenant un groupe de travail dans le menu déroulant.
   - Entrez le nom de la base de données au sein du cluster Amazon Redshift ou du groupe de travail Amazon Redshift Serverless sélectionné.
   - Choisissez Créer un jeu de paramètres.

> **Note**
>
> Vous ne pouvez ajouter que 10 ensembles de paramètres au DefaultDataWarehouse plan.

Une fois que vous avez activé le SageMaker plan Amazon dans votre AWS compte, vous pouvez ajouter des ensembles de paramètres à la configuration du plan. Un ensemble de paramètres est un groupe de clés et de valeurs, requis pour DataZone qu'Amazon puisse établir une connexion avec votre Amazon SageMaker et utilisé pour créer des environnements Sagemaker.

Ajouter des ensembles de paramètres au SageMaker plan Amazon

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines, puis choisissez le domaine contenant le plan activé dans lequel vous souhaitez ajouter le jeu de paramètres.
3. Choisissez l'onglet Blueprints, puis choisissez le SageMaker plan Amazon pour ouvrir la page de détails du plan.
4. Sous l'onglet Ensembles de paramètres de la page de détails du plan, choisissez Créer un jeu de paramètres, puis spécifiez les éléments suivants :

   - Entrez un nom pour le jeu de paramètres.
   - Vous pouvez éventuellement fournir une description pour le jeu de paramètres.
   - Spécifiez le type d'authentification SageMaker du domaine Amazon. Vous pouvez choisir IAM ou IAM Identity Center (SSO).
   - Spécifiez une AWS région.
   - Spécifiez une clé AWS KMS pour le chiffrement des données. Vous pouvez choisir une clé existante ou en créer une nouvelle.
   - Sous Paramètres d'environnement, spécifiez les éléments suivants :
     - ID VPC : ID que vous utilisez pour le VPC de l'environnement Amazon. SageMaker Vous pouvez spécifier un VPC existant ou en créer un nouveau.
     - Sous-réseaux : un ou plusieurs IDs pour une plage d'adresses IP pour des ressources spécifiques au sein de votre VPC.
     - Accès au réseau : choisissez VPC uniquement ou Internet public uniquement.

- Groupe de sécurité : groupe de sécurité à utiliser lors de la configuration du VPC et des sous-réseaux.

- Sous Paramètres de la source de données, sélectionnez l'une des options suivantes :
  - AWS Glue uniquement
  - AWS Glue + Amazon Redshift sans serveur. Si vous choisissez cette option, spécifiez les éléments suivants :
    - Spécifiez l'ARN AWS secret qui contient les informations d'identification du cluster Amazon Redshift sélectionné. Le AWS secret doit être étiqueté avec le `AmazonDataZoneDomain : [Domain_ID]` tag afin de pouvoir être utilisé dans un ensemble de paramètres.

      Si vous n'avez pas de AWS secret existant, vous pouvez également en créer un nouveau en choisissant Créer un nouveau AWS secret. Cela ouvre une boîte de dialogue dans laquelle vous pouvez fournir le nom du secret, le nom d'utilisateur et le mot de passe. Une fois que vous avez choisi Create New AWS Secret, Amazon DataZone crée un nouveau secret dans le service AWS Secrets Manager et s'assure que le secret est étiqueté avec le domaine dans lequel vous essayez de créer le jeu de paramètres.
    - Spécifiez le groupe de travail Amazon Redshift que vous souhaitez utiliser lors de la création d'environnements.
    - Spécifiez le nom de la base de données (au sein du groupe de travail que vous avez choisi) que vous souhaitez utiliser lors de la création d'environnements.
  - AWS Glue uniquement + Amazon Redshift Cluster
    - Spécifiez l'ARN AWS secret qui contient les informations d'identification du cluster Amazon Redshift sélectionné. Le AWS secret doit être étiqueté avec le `AmazonDataZoneDomain : [Domain_ID]` tag afin de pouvoir être utilisé dans un ensemble de paramètres.

      Si vous n'avez pas de AWS secret existant, vous pouvez également en créer un nouveau en choisissant Créer un nouveau AWS secret. Cela ouvre une boîte de dialogue dans laquelle vous pouvez fournir le nom du secret, le nom d'utilisateur et le mot de passe. Une fois que vous avez choisi Create New AWS Secret, Amazon DataZone crée un nouveau secret dans le service AWS Secrets Manager et s'assure que le secret est étiqueté avec le domaine dans lequel vous essayez de créer le jeu de paramètres.
    - Spécifiez le cluster Amazon Redshift que vous souhaitez utiliser lors de la création d'environnements.

- Spécifiez le nom de la base de données (au sein du cluster que vous avez choisi) que vous souhaitez utiliser lors de la création d'environnements.

5. Choisissez Créer un jeu de paramètres.

Ajoutez Amazon SageMaker en tant que service de confiance dans le AWS compte propriétaire du DataZone domaine Amazon

Si vous avez activé le SageMaker plan Amazon, vous devez également l'ajouter SageMaker comme l'un des services fiables d'Amazon DataZone. Pour ce faire, suivez la procédure suivante :

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines, puis choisissez le domaine qui contient le SageMaker plan activé.
3. Choisissez les services fiables, puis Amazon SageMaker, puis sélectionnez Activer.

DataZone Schémas AWS de service personnalisés Amazon

Dans Amazon DataZone, les plans de AWS service personnalisés vous permettent d'optimiser l'utilisation des ressources et les coûts en configurant Amazon DataZone pour qu'il utilise vos propres rôles et AWS services AWS Identity and Access Management (IAM) existants que vous avez déjà configurés dans votre organisation.

Un plan avec lequel un DataZone environnement Amazon est créé définit les outils et services que les membres du projet auquel appartient l'environnement peuvent utiliser lorsqu'ils travaillent avec les actifs du DataZone catalogue Amazon. Dans la version actuelle d'Amazon DataZone, il existe les plans intégrés suivants :

- Plan du lac de données
- Plan d'entrepôt de données
- SageMaker Plan Amazon

Avec les plans AWS de services DataZone personnalisés Amazon, vous pouvez créer des environnements et des projets adaptés aux AWS services que vous utilisez actuellement dans votre organisation. Avec des plans personnalisés, vous pouvez inclure Amazon DataZone dans vos pipelines de données existants en le configurant pour utiliser vos rôles IAM existants afin d'améliorer la gouvernance lors de la configuration de l'infrastructure et de collaborer sur des initiatives commerciales.

> **Important**
>
> Avec Amazon DataZone Custom AWS Service Prints, vous pouvez migrer votre SageMaker domaine Amazon existant vers Amazon DataZone. Grâce à cette fonctionnalité, les administrateurs peuvent désormais configurer DataZone des projets Amazon en important leurs utilisateurs autorisés existants, leurs configurations de sécurité et leurs politiques depuis les SageMaker domaines Amazon. Pour plus d'informations, voir Configuration SageMaker des actifs (guide de l'administrateur).

Rubriques

- Activer un plan AWS de service personnalisé
- Création d'un environnement à l'aide d'un plan AWS de service personnalisé

- Créez des actions dans un environnement AWS de service personnalisé
- Ajouter des membres du projet à un environnement AWS de service personnalisé
- Configuration d'une source de données dans un environnement AWS de service
- Configuration d'un objectif d'abonnement dans un environnement AWS de service

Activer un plan AWS de service personnalisé

Suivez la procédure ci-dessous pour activer un plan AWS de service personnalisé dans votre domaine.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez activer un plan de AWS service personnalisé.
3. Choisissez l'onglet Blueprints, puis choisissez le plan du AWS service dans la liste des plans disponibles, puis sélectionnez Activer.

Création d'un environnement à l'aide d'un plan AWS de service personnalisé

Procédez comme suit pour créer un environnement à l'aide d'un plan AWS de service personnalisé.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel votre plan AWS de service personnalisé est activé.
3. Choisissez l'onglet Blueprints, puis choisissez le plan du AWS service activé, puis choisissez Create environment.
4. Sur la page Créer un environnement, spécifiez ce qui suit, puis choisissez Créer un environnement :
   - Nom : spécifiez le nom de l'environnement.
   - Description : spécifiez la description de l'environnement.

- Projet : spécifiez un projet propriétaire nouveau ou existant pour l'environnement. Les projets permettent à des groupes d'utilisateurs de découvrir, de publier, de s'abonner et de consommer des actifs sur Amazon DataZone. Cet environnement sera accessible à tous les membres du projet spécifié. Tous les environnements appartiennent à des projets dont les utilisateurs ont accès à l'environnement.
- Rôle dans l'environnement : spécifiez un rôle IAM existant qui permettra à Amazon DataZone d'accéder à vos AWS services et ressources existants, tels qu'Amazon S3 et AWS Glue, dans cet environnement.

> **Note**
>
> Amazon DataZone ne vous fournit pas ce rôle. Vous devez disposer d'un rôle IAM existant avec des autorisations pour accéder AWS aux services et ressources existants que vous souhaitez activer dans cet environnement.
> Assurez-vous que ce rôle IAM dispose des autorisations minimales requises, en d'autres termes, qu'il est limité pour fournir un accès uniquement aux AWS services et aux ressources que vous souhaitez activer dans cet environnement.
> Vous pouvez utiliser le générateur de AWS politiques pour créer une politique adaptée à vos besoins et l'associer au rôle IAM personnalisé que vous souhaitez utiliser.
> Assurez-vous que le rôle commence par `AmazonDataZone` suivre les conventions. Ce n'est pas obligatoire, mais recommandé. Si l'administrateur IAM utilise la `AmazonDataZoneFullAccess` politique, vous devez suivre cette convention car il existe une validation par vérification des rôles réussie.
> Lorsque vous créez votre rôle personnalisé, assurez-vous qu'il fait confiance `datazone.amazonaws.com` à sa politique de confiance :
>
> JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "datazone.amazonaws.com"
                ]
            },
```

```
            "Action": [
                "sts:AssumeRole",
                "sts:TagSession"
            ]
        }
    ]
}
```

- AWS région : spécifiez AWS la région dans laquelle vous souhaitez créer cet environnement.

Créez des actions dans un environnement AWS de service personnalisé

Procédez comme suit pour créer des actions dans un environnement AWS de service personnalisé. En créant des actions dans un environnement de AWS service personnalisé, vous ajoutez des liens profonds vers le portail de DataZone données Amazon vers l'outil d'analyse disponible dans cet environnement.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel votre plan AWS de service personnalisé est activé.
3. Choisissez l'onglet Blueprints, puis choisissez le plan de AWS service activé, puis choisissez l'environnement de AWS service dans lequel vous souhaitez ajouter des actions.
4. Sur la page des liens de AWS console, choisissez des liens (actions) dans les sections AWS Liens populaires ou AWS Liens personnalisés pour activer les liens profonds vers vos compartiments Amazon S3, vos groupes de travail Amazon Athena AWS , vos tâches Glue ou vers toute AWS autre ressource de console personnalisée depuis cet environnement via le portail de données DataZone Amazon.
5. Si vous accédez à cet environnement dans le portail de données en utilisant le lien Portail de données dans la section Résumé de cet environnement, vous pouvez voir les liens profonds que vous avez ajoutés dans la section Outils d'analyse.

Ajouter des membres du projet à un environnement AWS de service personnalisé

Suivez la procédure ci-dessous pour ajouter des membres du projet à un environnement AWS de service.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Cliquez sur l'onglet Projets, puis sélectionnez le projet dans un environnement de AWS service auquel vous souhaitez ajouter des membres.
3. Choisissez Ajouter, puis, sur la page Ajouter des membres, recherchez et ajoutez des membres issus d'utilisateurs IAM, d'utilisateurs SSO ou de groupes SSO. Spécifiez le rôle assigné dans le projet à un propriétaire, à un contributeur, à un consommateur, à un délégué ou à un spectateur. Lorsque vous avez terminé de rechercher et d'ajouter des membres, choisissez Ajouter des membres.

Configuration d'une source de données dans un environnement AWS de service

Procédez comme suit pour configurer une source de données dans un environnement AWS de service.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez l'onglet Blueprints, puis choisissez le plan de AWS service personnalisé.
3. Sous Environnements créés, choisissez l'environnement de AWS service dans lequel vous souhaitez configurer une source de données.
4. Choisissez l'onglet Sources de données, choisissez Ajouter, spécifiez ce qui suit, puis choisissez Ajouter.

    - Nom : nom de la source de données.
    - Ressource : choisissez AWS Glue ou Amazon Redshift.
        - Pour AWS Glue, spécifiez la base de données des ressources.

    - Pour Amazon Redshift, choisissez Cluster ou Serverless, puis spécifiez les informations d'identification Redshift, y compris un AWS secret nouveau ou existant, un cluster ou un groupe de travail sans serveur que vous souhaitez utiliser lors de la création d'environnements, la base de données que vous souhaitez utiliser lors de la création d'environnements et le schéma au sein de la base de données spécifiée.
- Autorisations : spécifiez un rôle de gestion des accès qui donnera DataZone à Amazon l'autorisation d'ingérer et de gérer l'accès aux tables dans AWS Lake Formation (pour AWS Glue) ou qui fournira à Amazon l'autorisation DataZone d'ingérer et de gérer l'accès aux tables dans Amazon Redshift.
- Utilisation à des fins de consommation de données : sur Amazon DataZone, les membres du projet peuvent consommer des données par le biais d'objectifs d'abonnement qu'Amazon DataZone utilise pour permettre l'accès aux données pour lesquelles vous vous êtes abonné dans le cadre de vos projets. Spécifiez si vous souhaitez également ajouter cette source de données en tant que cible d'abonnement.

Configuration d'un objectif d'abonnement dans un environnement AWS de service

Procédez comme suit pour configurer un objectif d'abonnement dans un environnement AWS de service.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez l'onglet Blueprints, puis choisissez le plan de AWS service.
3. Sous Environnements créés, choisissez l'environnement de AWS service dans lequel vous souhaitez configurer un objectif d'abonnement.
4. Choisissez l'onglet Objectifs d'abonnement, choisissez Ajouter, spécifiez ce qui suit, puis choisissez Ajouter.

    - Nom : nom de la cible de l'abonnement.
    - Ressource : choisissez AWS Glue ou Amazon Redshift.
        - Pour AWS Glue, spécifiez la base de données des ressources.
        - Pour Amazon Redshift, choisissez Cluster ou Serverless, puis spécifiez les informations d'identification Redshift, y compris un AWS secret nouveau ou existant, un cluster

ou un groupe de travail sans serveur que vous souhaitez utiliser lors de la création d'environnements, la base de données que vous souhaitez utiliser lors de la création d'environnements et le schéma au sein de la base de données spécifiée.

- Autorisations : spécifiez un rôle de gestion des accès qui donnera DataZone à Amazon l'autorisation d'ingérer et de gérer l'accès aux tables dans AWS Lake Formation (pour AWS Glue) ou qui fournira à Amazon l'autorisation DataZone d'ingérer et de gérer l'accès aux tables dans Amazon Redshift.
- Utilisation pour la consommation de données : dans Amazon DataZone, vous pouvez publier des données dans le catalogue de données via une source de données qui permet l'ingestion de métadonnées. Spécifiez s'il faut également ajouter cette cible d'abonnement en tant que source de données.

Comptes associés sur Amazon DataZone

L'association de vos AWS comptes à votre DataZone domaine Amazon permet aux utilisateurs du domaine de publier et de consommer les données de ces AWS comptes. La configuration d'une association de comptes se fait en trois étapes.

- Commencez par partager le domaine avec le AWS compte souhaité en demandant l'association. Amazon DataZone utilise AWS Resource Access Manager (RAM) si le AWS compte est différent du AWS compte du domaine. Une association de comptes ne peut être initiée que par le DataZone domaine Amazon.
- Ensuite, demandez au propriétaire du compte d'accepter la demande d'association.
- Troisièmement, demandez au propriétaire du compte d'activer les plans d'environnement souhaités. En activant un plan, le propriétaire du compte fournit aux utilisateurs du domaine les rôles IAM et les configurations de ressources nécessaires pour créer et accéder aux ressources de leur compte, telles que les bases de données AWS Glue et les clusters Amazon Redshift.

Procédez comme suit pour associer un compte à Amazon DataZone :

- Étape 1 - Demande d'association avec d'autres AWS comptes
- Étape 2 - Accepter une demande d'association de compte provenant d'un DataZone domaine Amazon et activer un plan d'environnement
- Étape 3 - Activer un plan d'environnement dans un compte associé AWS

Demande d'association avec d'autres AWS comptes

> **Note**
>
> En envoyant une demande d'association à un autre AWS compte, vous partagez votre domaine avec l'autre AWS compte avec AWS Resource Access Manager (RAM). Assurez-vous de vérifier l'exactitude de l'identifiant de compte que vous entrez.

Pour demander une association avec d'autres AWS comptes dans la DataZone console Amazon pour un DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de

DataZone gestion Amazonpour obtenir les autorisations minimales nécessaires pour demander une association de compte.

Procédez comme suit pour demander une association avec d'autres AWS comptes.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Faites défiler l'écran jusqu'à l'onglet Comptes associés et sélectionnez Demander une association.
4. Entrez IDs les comptes dont vous souhaitez demander l'association. Lorsque vous êtes satisfait de la liste des comptes IDs, choisissez Demander une association.
5. Sous Politique de RAM, spécifiez la politique de RAM pour l'association de comptes. Vous pouvez soit choisir `AWSRAMPermissionDataZonePortalReadWrite` celui qui permettra aux comptes associés d'exécuter Amazon DataZone APIs et d'accéder au portail de donnéesAWSRAMPermissionDataZoneDefault, soit celui qui permettra aux comptes associés d'exécuter uniquement Amazon DataZone APIs et ne fournira pas d'accès au portail de données. Amazon crée DataZone ensuite un partage de ressources dans le AWS Resource Access Manager au nom de votre compte, avec le ou les identifiants de compte saisis comme principaux.
6. Vous devez informer le propriétaire des autres AWS comptes pour qu'il accepte votre demande. Les invitations expirent au bout de sept (7) jours.

Fournissez un accès au compte à votre clé KMS gérée par le client

Les DataZone domaines Amazon et leurs métadonnées sont chiffrés, soit (par défaut) à l'aide d'une clé détenue par le client AWS, soit (éventuellement) d'une clé gérée par le client par le biais du AWS Key Management Service (KMS) que vous possédez et que vous fournissez lors de la création du domaine. Si votre domaine est chiffré à l'aide d'une clé gérée par le client, suivez la procédure ci-dessous pour autoriser le compte associé à utiliser la clé KMS.

1. Connectez-vous à la console de AWS gestion et ouvrez la console KMS à l'adresse https://console.aws.amazon.com/kms/.
2. Pour afficher les clés de votre compte que vous créez et gérez vous-même, dans le volet de navigation, choisissez Clés gérées par le client.

3. Pour afficher les clés de votre compte que vous créez et gérez vous-même, dans le volet de navigation, choisissez Clés gérées par le client.
4. Dans la liste des clés KMS, choisissez l'alias ou l'ID de clé de la clé KMS que vous souhaitez examiner.
5. Pour autoriser ou interdire aux AWS comptes externes d'utiliser la clé KMS, utilisez les commandes de la section Autres AWS comptes de la page. Les principaux IAM de ces comptes (dotés eux-mêmes des autorisations KMS appropriées) peuvent utiliser la clé KMS dans le cadre d'opérations cryptographiques, telles que le chiffrement, le déchiffrement, le rechiffrement et la génération de clés de données.

Accepter une demande d'association de compte provenant d'un DataZone domaine Amazon et activer un plan d'environnement

Pour accepter l'association dans la console DataZone de gestion Amazon avec un DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales.

Complétez ce qui suit pour accepter l'association avec un DataZone domaine Amazon.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les demandes et sélectionnez le domaine invitant dans la liste. L'état de l'invitation doit être demandé. Choisissez Demande de révision.
3. Choisissez d'activer les plans d'environnement d'entrepôt de données par défaut and/or des lacs de données en ne cochant aucune des cases, les deux ou l'une des cases. Tu pourras le faire plus tard.

   - Le plan d'environnement du lac de données permet aux utilisateurs du domaine de créer et de gérer les ressources AWS Glue, Amazon S3 et Amazon Athena à publier et à consommer à partir d'un lac de données.
   - Le plan d'environnement d'entrepôt de données permet aux utilisateurs du domaine de créer et de gérer des ressources Amazon Redshift à publier et à consommer à partir d'un entrepôt de données.
4. Si vous choisissez de sélectionner l'un des plans d'environnement par défaut ou les deux, configurez les autorisations et ressources suivantes.

- Le rôle IAM de gestion des accès fournit des autorisations à Amazon pour permettre DataZone aux utilisateurs du domaine d'ingérer et de gérer l'accès à des tables, telles que AWS Glue et Amazon Redshift. Vous pouvez choisir de demander à Amazon de DataZone créer et d'utiliser un nouveau rôle IAM, ou vous pouvez choisir parmi une liste de rôles IAM existants.
- Le rôle Provisioning IAM fournit des autorisations DataZone à Amazon pour permettre aux utilisateurs du domaine de créer et de configurer des ressources d'environnement, telles que les bases de données AWS Glue. Vous pouvez choisir de demander à Amazon de DataZone créer et d'utiliser un nouveau rôle IAM, ou vous pouvez choisir parmi une liste de rôles IAM existants.
- Le compartiment Amazon S3 pour Data Lake est le compartiment ou le chemin qu'Amazon utilisera lorsque les utilisateurs du domaine DataZone stockeront les données du lac de données. Vous pouvez utiliser le compartiment par défaut sélectionné par Amazon DataZone ou choisir votre propre chemin Amazon S3 existant en saisissant sa chaîne de chemin. Si vous sélectionnez votre propre chemin Amazon S3, vous devrez mettre à jour les politiques IAM pour autoriser Amazon DataZone à l'utiliser.

5. Lorsque vous êtes satisfait de vos configurations, choisissez Accepter et configurez l'association.

Activer un plan d'environnement dans un compte associé AWS

Pour activer un plan d'environnement dans la console de DataZone gestion Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales.

Procédez comme suit pour activer un plan dans un domaine associé.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Ouvrez le panneau de navigation de gauche et choisissez Domaines associés.
3. Choisissez le domaine pour lequel vous souhaitez activer un plan d'environnement.
4. Dans la liste des plans, choisissez soit le plan DefaultDataWarehouse, DefaultDataLakesoit le plan Amazon SageMaker, soit le plan du AWS service personnalisé.

> **Note**
>
> Si vous activez le plan de AWS service personnalisé, il n'est pas nécessaire de spécifier un rôle de gestion des accès. Les autorisations et le mécanisme d'autorisation pour le plan du AWS service personnalisé sont gérés lorsque vous créez des environnements à l'aide de ce plan. Pour de plus amples informations, veuillez consulter Création d'un environnement à l'aide d'un plan AWS de service personnalisé.

5. Sur la page de détails du plan choisi, choisissez Activer dans ce compte.
6. Sur la page Autorisations et ressources, spécifiez les éléments suivants :

- Si vous activez le DefaultDataLakeplan, pour le rôle Glue Manage Access, spécifiez un rôle de service nouveau ou existant qui DataZone autorise Amazon à ingérer et à gérer l'accès aux tables dans AWS Glue and AWS Lake Formation.
- Si vous activez le DefaultDataWarehouseplan, pour le rôle Redshift Manage Access, spécifiez un rôle de service nouveau ou existant qui autorise DataZone Amazon à ingérer et à gérer l'accès aux partages de données, aux tables et aux vues dans Amazon Redshift.
- Si vous activez le SageMaker plan Amazon, pour le rôle de SageMaker gestion des accès, spécifiez un rôle de service nouveau ou existant qui accorde à Amazon l' DataZone autorisation de publier les SageMaker données Amazon dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux ressources SageMaker publiées par Amazon dans le catalogue.

  > **Important**
  >
  > Lorsque vous activez le SageMaker plan Amazon, Amazon DataZone vérifie si les rôles IAM suivants pour Amazon DataZone existent dans le compte et la région actuels. Si ces rôles n'existent pas, Amazon les crée DataZone automatiquement.
  >
  > - AmazonDataZoneGlueAccess- <region>- <domainId>
  > - AmazonDataZoneRedshiftAccess- <region>- <domainId>

- Pour le rôle de provisionnement, spécifiez un rôle de service nouveau ou existant qui accorde à Amazon DataZone l'autorisation de créer et de configurer les ressources de l'environnement AWS CloudFormation à l'aide du compte et de la région d'environnement.
- Si vous activez le SageMaker plan Amazon, pour le compartiment Amazon S3 pour la source de données SageMaker -Glue, spécifiez un compartiment Amazon S3 qui doit être utilisé par

tous les SageMaker environnements du AWS compte. Le préfixe de compartiment que vous spécifiez doit être l'un des suivants :

- zone de données Amazon*
- créateur de zones de données*
- zone de données SageMaker*
- DataZone- Sagemaker*
- Sagemaker- * DataZone
- DataZone-SageMaker*
- SageMaker-DataZone*

7. Choisissez Activer le plan.

Une fois que vous avez activé le ou les plans choisis, vous pouvez contrôler quels projets peuvent utiliser les plans dans votre compte pour créer des profils d'environnement. Vous pouvez le faire en affectant la gestion des projets à la configuration du plan.

Spécifiez la gestion des projets sur Enabled DefaultDataLake ou DefaultDataWarehouse Blueprint

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous avec les informations d'identification de votre compte.
2. Ouvrez le panneau de navigation de gauche et sélectionnez Domaines associés, puis choisissez le domaine dans lequel vous souhaitez ajouter la gestion de projets.
3. Choisissez l'onglet Blueprints, puis choisissez DefaultDataLake ou DefaultDataWareshouse Blueprint.
4. Par défaut, tous les projets du domaine peuvent utiliser le DefaultDataWareshouse plan DefaultDataLake ou le plan du compte pour créer des profils d'environnement. Toutefois, vous pouvez limiter cela en affectant la gestion des projets au plan. Pour ajouter des projets de gestion, choisissez Sélectionner la gestion du projet, puis choisissez les projets que vous souhaitez ajouter en tant que gestion de projets dans le menu déroulant, puis sélectionnez Sélectionner la gestion de projets.

Une fois que vous avez activé le DefaultDataWarehouse plan dans votre AWS compte, vous pouvez ajouter des ensembles de paramètres à la configuration du plan. Un ensemble de paramètres est un groupe de clés et de valeurs, requis pour qu'Amazon DataZone établisse une connexion à votre cluster Amazon Redshift et est utilisé pour créer des environnements d'entrepôt de données. Ces

paramètres incluent le nom de votre cluster Amazon Redshift, de votre base de données et le AWS secret contenant les informations d'identification du cluster.

> **Important**
>
> Par défaut, aucun projet de gestion n'est spécifié pour les plans d'environnement, ce qui signifie que tout DataZone utilisateur d'Amazon peut créer des profils pour un plan d'environnement. Par conséquent, il est fortement recommandé de toujours spécifier la gestion des projets pour vos plans d'environnement afin de garantir une gouvernance plus solide.

Ajouter des ensembles de paramètres au DefaultDataWarehouse plan

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous avec les informations d'identification de votre compte.
2. Ouvrez le panneau de navigation de gauche et sélectionnez Domaines associés, puis choisissez le domaine dans lequel vous souhaitez ajouter des ensembles de paramètres.
3. Choisissez l'onglet Plans, puis choisissez le DefaultDataWareshouse plan pour ouvrir la page de détails du plan.
4. Dans l'onglet Ensembles de paramètres de la page de détails du plan, choisissez Créer un jeu de paramètres.
    - Entrez un nom pour le jeu de paramètres.
    - Fournissez éventuellement une description du jeu de paramètres.
    - Sélectionner une région
    - Sélectionnez le cluster Amazon Redshift ou Amazon Redshift Serverless.
    - Sélectionnez l'ARN AWS secret qui contient les informations d'identification du cluster Amazon Redshift sélectionné ou du groupe de travail Amazon Redshift Serverless. Le AWS secret doit être étiqueté avec le `AmazonDataZoneDomain : [Domain_ID]` tag afin de pouvoir être utilisé dans un ensemble de paramètres.
        - Si vous n'avez pas de AWS secret existant, vous pouvez également en créer un nouveau en choisissant Créer un nouveau AWS secret. Cela ouvre une boîte de dialogue dans laquelle vous pouvez fournir le nom du secret, le nom d'utilisateur et le mot de passe. Une fois que vous avez choisi Create New AWS Secret, Amazon DataZone crée un nouveau secret dans

le service AWS Secrets Manager et s'assure que le secret est étiqueté avec le domaine dans lequel vous essayez de créer le jeu de paramètres.

- Sélectionnez le cluster Amazon Redshift ou le groupe de travail Amazon Redshift Serverless.
- Entrez le nom de la base de données au sein du cluster Amazon Redshift ou du groupe de travail Amazon Redshift Serverless sélectionné.
- Choisissez Créer un jeu de paramètres.

> **Note**
>
> Vous ne pouvez ajouter que 10 ensembles de paramètres au DefaultDataWarehouse plan.

Une fois que vous avez activé le SageMaker plan Amazon dans votre AWS compte, vous pouvez ajouter des ensembles de paramètres à la configuration du plan. Un ensemble de paramètres est un groupe de clés et de valeurs, requis pour DataZone qu'Amazon puisse établir une connexion avec votre Amazon SageMaker et utilisé pour créer des environnements Sagemaker.

Ajouter des ensembles de paramètres au SageMaker plan Amazon

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous avec les informations d'identification de votre compte.
2. Choisissez Afficher les domaines, puis choisissez le domaine contenant le plan activé dans lequel vous souhaitez ajouter le jeu de paramètres.
3. Choisissez l'onglet Blueprints, puis choisissez le SageMaker plan Amazon pour ouvrir la page de détails du plan.
4. Sous l'onglet Ensembles de paramètres de la page de détails du plan, choisissez Créer un jeu de paramètres, puis spécifiez les éléments suivants :

   - Entrez un nom pour le jeu de paramètres.
   - Vous pouvez éventuellement fournir une description pour le jeu de paramètres.
   - Spécifiez le type d'authentification SageMaker du domaine Amazon. Vous pouvez choisir IAM ou IAM Identity Center (SSO).
   - Spécifiez une AWS région.
   - Spécifiez une clé AWS KMS pour le chiffrement des données. Vous pouvez choisir une clé existante ou en créer une nouvelle.

- Sous Paramètres d'environnement, spécifiez les éléments suivants :
  - ID VPC : ID que vous utilisez pour le VPC de l'environnement Amazon. SageMaker Vous pouvez spécifier un VPC existant ou en créer un nouveau.
  - Sous-réseaux : un ou plusieurs IDs pour une plage d'adresses IP pour des ressources spécifiques au sein de votre VPC.
  - Accès au réseau : choisissez VPC uniquement ou Internet public uniquement.
  - Groupe de sécurité : groupe de sécurité à utiliser lors de la configuration du VPC et des sous-réseaux.
- Sous Paramètres de la source de données, sélectionnez l'une des options suivantes :
  - AWS Glue uniquement
  - AWS Glue + Amazon Redshift sans serveur. Si vous choisissez cette option, spécifiez les éléments suivants :
    - Spécifiez l'ARN AWS secret qui contient les informations d'identification du cluster Amazon Redshift sélectionné. Le AWS secret doit être étiqueté avec le `AmazonDataZoneDomain : [Domain_ID]` tag afin de pouvoir être utilisé dans un ensemble de paramètres.

      Si vous n'avez pas de AWS secret existant, vous pouvez également en créer un nouveau en choisissant Créer un nouveau AWS secret. Cela ouvre une boîte de dialogue dans laquelle vous pouvez fournir le nom du secret, le nom d'utilisateur et le mot de passe. Une fois que vous avez choisi Create New AWS Secret, Amazon DataZone crée un nouveau secret dans le service AWS Secrets Manager et s'assure que le secret est étiqueté avec le domaine dans lequel vous essayez de créer le jeu de paramètres.
    - Spécifiez le groupe de travail Amazon Redshift que vous souhaitez utiliser lors de la création d'environnements.
    - Spécifiez le nom de la base de données (au sein du groupe de travail que vous avez choisi) que vous souhaitez utiliser lors de la création d'environnements.
  - AWS Glue uniquement + Amazon Redshift Cluster
    - Spécifiez l'ARN AWS secret qui contient les informations d'identification du cluster Amazon Redshift sélectionné. Le AWS secret doit être étiqueté avec le `AmazonDataZoneDomain : [Domain_ID]` tag afin de pouvoir être utilisé dans un ensemble de paramètres.

      Si vous n'avez pas de AWS secret existant, vous pouvez également en créer un nouveau en choisissant Créer un nouveau AWS secret. Cela ouvre une boîte de dialogue dans

laquelle vous pouvez fournir le nom du secret, le nom d'utilisateur et le mot de passe. Une fois que vous avez choisi Create New AWS Secret, Amazon DataZone crée un nouveau secret dans le service AWS Secrets Manager et s'assure que le secret est étiqueté avec le domaine dans lequel vous essayez de créer le jeu de paramètres.

- Spécifiez le cluster Amazon Redshift que vous souhaitez utiliser lors de la création d'environnements.
- Spécifiez le nom de la base de données (au sein du cluster que vous avez choisi) que vous souhaitez utiliser lors de la création d'environnements.

5. Choisissez Créer un jeu de paramètres.

Ajoutez Amazon SageMaker en tant que service de confiance dans le AWS compte associé

Si vous avez activé le SageMaker plan Amazon, vous devez également l'ajouter SageMaker comme l'un des services fiables d'Amazon DataZone. Pour ce faire, suivez la procédure suivante :

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous avec les informations d'identification de votre compte.
2. Choisissez Afficher les domaines, puis choisissez le domaine qui contient le SageMaker plan activé.
3. Choisissez les services fiables, puis Amazon SageMaker, puis sélectionnez Activer.

Rejeter une demande d'association de compte provenant d'un DataZone domaine Amazon

Pour rejeter une demande d'association dans la console de DataZone gestion Amazon depuis un DataZone domaine Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales.

Complétez ce qui suit pour rejeter une demande d'association provenant d'un DataZone domaine Amazon.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.

2. Choisissez Afficher les demandes et sélectionnez le domaine invitant dans la liste. L'état de l'invitation doit être demandé. Choisissez Refuser l'association. Confirmez votre choix en choisissant Refuser l'association.

Supprimer un compte associé sur Amazon DataZone

Pour supprimer un AWS compte associé dans la console de DataZone gestion Amazon, vous devez assumer un rôle IAM dans le compte avec des autorisations administratives. Configurer les autorisations IAM requises pour utiliser la console de DataZone gestion Amazonpour obtenir les autorisations minimales.

Suivez la procédure ci-dessous pour supprimer un compte associé de votre domaine.

1. Connectez-vous à la console de AWS gestion et ouvrez la console de DataZone gestion Amazon à l'adresse https://console.aws.amazon.com/datazone.
2. Choisissez Afficher les domaines et choisissez le nom du domaine dans la liste. Le nom est un hyperlien.
3. Faites défiler la page jusqu'à l'onglet Comptes associés. Choisissez l'identifiant du AWS compte que vous souhaitez supprimer.
4. Choisissez Dissocier. Confirmez votre choix en saisissant Dissocier dans le champ et en choisissant Dissocier.
5. Le compte est désormais supprimé de votre domaine et ne peut pas être utilisé par les utilisateurs du domaine pour publier et consommer des données.

Catalogue DataZone de données Amazon

Vous pouvez utiliser le catalogue de données Amazon DataZone Business pour cataloguer les données de votre organisation en fonction du contexte commercial et permettre ainsi à tous les membres de votre organisation de trouver et de comprendre rapidement les données.

Pour pouvoir utiliser Amazon DataZone pour cataloguer vos données, vous devez d'abord les importer (actifs) en tant qu'inventaire de votre projet sur Amazon DataZone. La création d'un inventaire pour un projet rend les actifs accessibles uniquement aux membres de ce projet. Les actifs de l'inventaire du projet ne sont pas accessibles à tous les utilisateurs du domaine search/browse sauf s'ils sont publiés explicitement.

Après avoir créé un inventaire de projet, les propriétaires de données peuvent organiser leurs actifs d'inventaire avec les métadonnées commerciales requises en ajoutant ou en mettant à jour les noms commerciaux (actif et schéma), les descriptions (actif et schéma), lisez-moi, les termes du glossaire (actif et schéma) et les formulaires de métadonnées.

L'étape suivante de l'utilisation d'Amazon DataZone pour cataloguer vos données consiste à rendre les actifs d'inventaire de votre projet accessibles aux utilisateurs du domaine. Vous pouvez le faire en publiant les actifs d'inventaire dans le DataZone catalogue Amazon. Seule la dernière version de la ressource d'inventaire peut être publiée dans le catalogue et seule la dernière version publiée est active dans le catalogue de découverte. Si un actif d'inventaire est mis à jour après sa publication dans le DataZone catalogue Amazon, vous devez le publier à nouveau de manière explicite pour que la dernière version figure dans le catalogue de découverte.

Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Rubriques

- Création d'un glossaire commercial sur Amazon DataZone
- Modifier un glossaire commercial sur Amazon DataZone
- Supprimer un glossaire professionnel sur Amazon DataZone
- Création d'un terme dans un glossaire sur Amazon DataZone
- Modifier un terme dans un glossaire sur Amazon DataZone
- Supprimer un terme dans un glossaire sur Amazon DataZone
- Création d'un formulaire de métadonnées dans Amazon DataZone

- Modifier un formulaire de métadonnées dans Amazon DataZone
- Supprimer un formulaire de métadonnées dans Amazon DataZone
- Création d'un champ dans un formulaire de métadonnées sur Amazon DataZone
- Modifier un champ dans un formulaire de métadonnées sur Amazon DataZone
- Supprimer un champ dans un formulaire de métadonnées sur Amazon DataZone

Création d'un glossaire commercial sur Amazon DataZone

Sur Amazon DataZone, un glossaire commercial est un ensemble de termes commerciaux (mots) qui peuvent être associés à des actifs (données). Il fournit des vocabulaires appropriés avec une liste de termes commerciaux et leurs définitions à l'intention des utilisateurs professionnels afin de garantir que les mêmes définitions sont utilisées dans l'ensemble de l'organisation lors de l'analyse des données. Les glossaires commerciaux sont créés dans le domaine du catalogue et peuvent être appliqués aux actifs et aux colonnes pour aider à comprendre les principales caractéristiques de cet actif ou de cette colonne. Un ou plusieurs termes du glossaire peuvent être appliqués. Un glossaire commercial peut être une liste plate de termes dans laquelle n'importe quel terme du glossaire commercial peut être associé à une sous-liste d'autres termes. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer un glossaire dans votre DataZone domaine Amazon, vous devez être membre du projet propriétaire disposant des autorisations appropriées pour ce domaine.

Pour créer un glossaire, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans Amazon DataZone Data Portal, choisissez Glossaires, puis Create glossary.
4. Spécifiez le nom, la description et le propriétaire du glossaire, puis choisissez Créer un glossaire.
5. Activez le nouveau glossaire en choisissant le bouton Activé.
6. Sur la page de détails du glossaire, vous pouvez choisir Create readme pour ajouter des informations supplémentaires sur ce glossaire.

Pour désactiver ou activer un glossaire professionnel, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans le portail Amazon DataZone Data, choisissez Glossaires et recherchez le glossaire professionnel que vous souhaitez désactiver/activer.
4. Sur la page de détails du glossaire, recherchez le bouton Activer/Désactiver et utilisez-le pour activer ou désactiver le glossaire sélectionné.

> **Note**
>
> La désactivation d'un glossaire désactive également tous les termes qu'il contient.

Modifier un glossaire commercial sur Amazon DataZone

Sur Amazon DataZone, un glossaire commercial est un ensemble de termes commerciaux (mots) qui peuvent être associés à des actifs (données). Il fournit des vocabulaires appropriés avec une liste de termes commerciaux et leurs définitions à l'intention des utilisateurs professionnels afin de garantir que les mêmes définitions sont utilisées dans l'ensemble de l'organisation lors de l'analyse des données. Les glossaires commerciaux sont créés dans le domaine du catalogue et peuvent être appliqués aux actifs et aux colonnes pour aider à comprendre les principales caractéristiques de cet actif ou de cette colonne. Un ou plusieurs termes du glossaire peuvent être appliqués. Un glossaire commercial peut être une liste plate de termes dans laquelle n'importe quel terme du glossaire commercial peut être associé à une sous-liste d'autres termes. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour modifier un glossaire dans votre DataZone domaine Amazon, vous devez être membre du projet propriétaire avec les autorisations appropriées pour ce domaine.

Pour modifier un glossaire professionnel, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes

DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.

2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Sur Amazon DataZone Data Portal, choisissez Glossaires et recherchez le glossaire commercial que vous souhaitez modifier.
4. Sur la page de détails du glossaire, développez Actions, puis choisissez Modifier pour modifier le glossaire.
5. Modifiez le nom et la description, puis choisissez Enregistrer.

Supprimer un glossaire professionnel sur Amazon DataZone

Sur Amazon DataZone, un glossaire commercial est un ensemble de termes commerciaux (mots) qui peuvent être associés à des actifs (données). Il fournit des vocabulaires appropriés avec une liste de termes commerciaux et leurs définitions à l'intention des utilisateurs professionnels afin de garantir que les mêmes définitions sont utilisées dans l'ensemble de l'organisation lors de l'analyse des données. Les glossaires commerciaux sont créés dans le domaine du catalogue et peuvent être appliqués aux actifs et aux colonnes pour aider à comprendre les principales caractéristiques de cet actif ou de cette colonne. Un ou plusieurs termes du glossaire peuvent être appliqués. Un glossaire commercial peut être une liste plate de termes dans laquelle n'importe quel terme du glossaire commercial peut être associé à une sous-liste d'autres termes. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour supprimer un glossaire de votre DataZone domaine Amazon, vous devez être membre du projet propriétaire avec les autorisations appropriées pour ce domaine.

Pour supprimer un glossaire professionnel, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Sur Amazon DataZone Data Portal, choisissez Glossaires et recherchez le glossaire professionnel que vous souhaitez supprimer.

4. Sur la page de détails du glossaire, développez Actions, puis choisissez Supprimer pour supprimer le glossaire.

> **Note**
>
> Vous devez supprimer tous les termes existants dans le glossaire avant de pouvoir supprimer le glossaire.

5. Confirmez la suppression du glossaire en choisissant Supprimer.

Création d'un terme dans un glossaire sur Amazon DataZone

Sur Amazon DataZone, un glossaire commercial est un ensemble de termes commerciaux qui peuvent être associés à des actifs (données). Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des termes dans un glossaire de votre DataZone domaine Amazon, vous devez être membre du projet propriétaire disposant des autorisations appropriées pour ce domaine.

Sur Amazon DataZone, les termes du glossaire commercial peuvent être décrits de manière détaillée. Pour définir le contexte d'un terme en particulier, vous pouvez définir les relations entre les termes. Lorsque vous définissez une relation pour un terme, celle-ci est automatiquement ajoutée à la définition du terme associé. Les relations terminologiques du glossaire disponibles sur Amazon DataZone sont les suivantes :

- Est un type de - indique que le terme actuel est un type du terme identifié. Indique que le terme identifié est le parent du terme actuel.
- Dispose de types : indique que le terme actuel est un terme générique désignant le ou les termes spécifiques indiqués. Cette relation peut désigner des termes secondaires pour le terme générique.

Pour créer un nouveau terme, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.

3. Dans Amazon DataZone Data Portal, sélectionnez Glossaires, puis choisissez le glossaire dans lequel vous souhaitez créer le nouveau terme.
4. Spécifiez le nom, la description et le propriétaire du terme, puis choisissez Créer un terme.
5. Activez le nouveau terme en choisissant le bouton Activé.
6. Pour ajouter un fichier Lisez-moi, accédez à la page de détails des termes, puis vous pouvez choisir Créer un fichier Lisez-moi pour ajouter des informations supplémentaires sur ce glossaire.
7. Pour ajouter des relations, accédez à la page de détails des termes, choisissez la section Relations terminologiques, puis choisissez Ajouter des termes au glossaire. Dans la boîte de dialogue, choisissez la relation et les termes que vous souhaitez associer, puis choisissez Fermer pour ajouter un terme au type de relation approprié. Cette relation est également ajoutée à tous les termes que vous avez associés.

Modifier un terme dans un glossaire sur Amazon DataZone

Sur Amazon DataZone, un glossaire commercial est un ensemble de termes commerciaux qui peuvent être associés à des actifs (données). Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des termes dans un glossaire de votre DataZone domaine Amazon, vous devez être membre du projet propriétaire disposant des autorisations appropriées pour ce domaine.

Sur Amazon DataZone, les termes du glossaire commercial peuvent être décrits de manière détaillée. Pour définir le contexte d'un terme en particulier, vous pouvez définir les relations entre les termes. Lorsque vous définissez une relation pour un terme, celle-ci est automatiquement ajoutée à la définition du terme associé. Les relations terminologiques du glossaire disponibles sur Amazon DataZone sont les suivantes :

- Est un type de - indique que le terme actuel est un type du terme identifié. Indique que le terme identifié est le parent du terme actuel.
- Dispose de types : indique que le terme actuel est un terme générique désignant le ou les termes spécifiques indiqués. Cette relation peut désigner des termes secondaires pour le terme générique.

Pour modifier un terme dans un glossaire, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant

à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.

2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Sur Amazon DataZone Data Portal, choisissez Glossaires, recherchez le glossaire contenant le terme que vous souhaitez modifier, puis choisissez-le.
4. Sur la page de détails du terme, développez Actions, puis choisissez Modifier pour modifier le terme.
5. Modifiez le nom et la description, puis choisissez Enregistrer.

Supprimer un terme dans un glossaire sur Amazon DataZone

Sur Amazon DataZone, un glossaire commercial est un ensemble de termes commerciaux qui peuvent être associés à des actifs (données). Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des termes dans un glossaire de votre DataZone domaine Amazon, vous devez être membre du projet propriétaire disposant des autorisations appropriées pour ce domaine.

Sur Amazon DataZone, les termes du glossaire commercial peuvent être décrits de manière détaillée. Pour définir le contexte d'un terme en particulier, vous pouvez définir les relations entre les termes. Lorsque vous définissez une relation pour un terme, celle-ci est automatiquement ajoutée à la définition du terme associé. Les relations terminologiques du glossaire disponibles sur Amazon DataZone sont les suivantes :

- Est un type de - indique que le terme actuel est un type du terme identifié. Indique que le terme identifié est le parent du terme actuel.
- Dispose de types : indique que le terme actuel est un terme générique désignant le ou les termes spécifiques indiqués. Cette relation peut désigner des termes secondaires pour le terme générique.

Pour supprimer un terme dans un glossaire, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.

3. Sur Amazon DataZone Data Portal, choisissez Glossaires, recherchez le glossaire contenant le terme que vous souhaitez supprimer, puis choisissez-le.
4. Sur la page de détails du glossaire, développez Actions, puis choisissez Supprimer pour supprimer le terme.
5. Confirmez la suppression du terme en choisissant Supprimer.

Création d'un formulaire de métadonnées dans Amazon DataZone

Sur Amazon DataZone, les formulaires de métadonnées sont des formulaires simples destinés à compléter le contexte commercial des métadonnées des actifs du catalogue. Il s'agit d'un mécanisme extensible permettant aux propriétaires de données d'enrichir la ressource avec des informations qui peuvent aider les utilisateurs des données lorsqu'ils recherchent et trouvent ces données. Les formulaires de métadonnées peuvent également servir de mécanisme pour assurer la cohérence de tous les actifs publiés dans le DataZone catalogue Amazon.

Une définition de formulaire de métadonnées est composée d'une ou de plusieurs définitions de champs, avec prise en charge des types de données booléens, datés, décimaux, entiers, chaînes et valeurs de champs du glossaire commercial. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des formulaires de métadonnées dans votre DataZone domaine Amazon, vous devez être membre du projet propriétaire et disposer des informations d'identification appropriées.

Pour créer un formulaire de métadonnées, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans Amazon DataZone Data Portal, sélectionnez Formulaires de métadonnées, puis sélectionnez Créer un formulaire.
4. Spécifiez le nom, la description et le propriétaire du formulaire de métadonnées, puis choisissez Créer un formulaire.

Modifier un formulaire de métadonnées dans Amazon DataZone

Sur Amazon DataZone, les formulaires de métadonnées sont des formulaires simples destinés à compléter le contexte commercial des métadonnées des actifs du catalogue. Il s'agit d'un mécanisme extensible permettant aux propriétaires de données d'enrichir la ressource avec des informations qui peuvent aider les utilisateurs des données lorsqu'ils recherchent et trouvent ces données. Les formulaires de métadonnées peuvent également servir de mécanisme pour assurer la cohérence de tous les actifs publiés dans le DataZone catalogue Amazon.

Une définition de formulaire de métadonnées est composée d'une ou de plusieurs définitions de champs, avec prise en charge des types de données booléens, datés, décimaux, entiers, chaînes et valeurs de champs du glossaire commercial. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des formulaires de métadonnées dans votre DataZone domaine Amazon, vous devez être membre du projet propriétaire et disposer des informations d'identification appropriées.

Pour modifier un formulaire de métadonnées, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans Amazon DataZone Data Portal, choisissez Formulaires de métadonnées, puis recherchez le formulaire de métadonnées que vous souhaitez modifier.
4. Sur la page de détails du formulaire de métadonnées, développez Actions, puis choisissez Modifier.
5. Mettez à jour les champs du nom, de la description et du propriétaire, puis choisissez Mettre à jour le formulaire.

Supprimer un formulaire de métadonnées dans Amazon DataZone

Sur Amazon DataZone, les formulaires de métadonnées sont des formulaires simples destinés à compléter le contexte commercial des métadonnées des actifs du catalogue. Il s'agit d'un mécanisme extensible permettant aux propriétaires de données d'enrichir la ressource avec des informations

qui peuvent aider les utilisateurs des données lorsqu'ils recherchent et trouvent ces données. Les formulaires de métadonnées peuvent également servir de mécanisme pour assurer la cohérence de tous les actifs publiés dans le DataZone catalogue Amazon.

Une définition de formulaire de métadonnées est composée d'une ou de plusieurs définitions de champs, avec prise en charge des types de données booléens, datés, décimaux, entiers, chaînes et valeurs de champs du glossaire commercial. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des formulaires de métadonnées dans votre DataZone domaine Amazon, vous devez être membre du projet propriétaire et disposer des informations d'identification appropriées.

Pour supprimer un formulaire de métadonnées, procédez comme suit :

> **Note**
>
> Avant de pouvoir supprimer un formulaire de métadonnées, vous devez le supprimer de tous les types de ressources ou de ressources auxquels il est appliqué.

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans Amazon DataZone Data Portal, choisissez Formulaires de métadonnées, puis recherchez le formulaire de métadonnées que vous souhaitez supprimer.
4. Si le formulaire de métadonnées que vous souhaitez supprimer est activé, désactivez-le en choisissant le bouton Activé.
5. Sur la page de détails du formulaire de métadonnées, développez Actions, puis choisissez Supprimer.
6. Confirmez la suppression en choisissant Supprimer.

Création d'un champ dans un formulaire de métadonnées sur Amazon DataZone

Sur Amazon DataZone, les formulaires de métadonnées sont des formulaires simples destinés à compléter le contexte commercial des métadonnées des actifs du catalogue. Il s'agit d'un mécanisme extensible permettant aux propriétaires de données d'enrichir la ressource avec des informations qui peuvent aider les utilisateurs des données lorsqu'ils recherchent et trouvent ces données. Les formulaires de métadonnées peuvent également servir de mécanisme pour assurer la cohérence de tous les actifs publiés dans le DataZone catalogue Amazon.

Une définition de formulaire de métadonnées est composée d'une ou de plusieurs définitions de champs, avec prise en charge des types de données booléens, datés, décimaux, entiers, chaînes et valeurs de champs du glossaire commercial. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des champs dans les formulaires de métadonnées de votre DataZone domaine Amazon, vous devez être membre du projet propriétaire et disposer des informations d'identification appropriées.

Pour créer un champ dans un formulaire de métadonnées, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans Amazon DataZone Data Portal, choisissez Formulaires de métadonnées, puis choisissez le formulaire de métadonnées dans lequel vous souhaitez créer des champs.
4. Sur la page de détails du formulaire, choisissez Créer un champ.
5. Spécifiez le nom du champ, sa description, son type et indiquez s'il s'agit d'un champ obligatoire, puis choisissez Créer un champ.

Modifier un champ dans un formulaire de métadonnées sur Amazon DataZone

Sur Amazon DataZone, les formulaires de métadonnées sont des formulaires simples destinés à compléter le contexte commercial des métadonnées des actifs du catalogue. Il s'agit d'un mécanisme

extensible permettant aux propriétaires de données d'enrichir la ressource avec des informations qui peuvent aider les utilisateurs des données lorsqu'ils recherchent et trouvent ces données. Les formulaires de métadonnées peuvent également servir de mécanisme pour assurer la cohérence de tous les actifs publiés dans le DataZone catalogue Amazon.

Une définition de formulaire de métadonnées est composée d'une ou de plusieurs définitions de champs, avec prise en charge des types de données booléens, datés, décimaux, entiers, chaînes et valeurs de champs du glossaire commercial. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des champs dans les formulaires de métadonnées de votre DataZone domaine Amazon, vous devez être membre du projet propriétaire et disposer des informations d'identification appropriées.

Pour modifier un champ dans un formulaire de métadonnées, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans Amazon DataZone Data Portal, choisissez Formulaires de métadonnées, puis choisissez le formulaire de métadonnées dans lequel vous souhaitez modifier les champs.
4. Sur la page de détails du formulaire, choisissez le champ que vous souhaitez modifier, puis développez Actions et choisissez Modifier.
5. Modifiez le nom du champ, sa description, son type et indiquez s'il s'agit d'un champ obligatoire, puis choisissez Mettre à jour le champ.

Supprimer un champ dans un formulaire de métadonnées sur Amazon DataZone

Sur Amazon DataZone, les formulaires de métadonnées sont des formulaires simples destinés à compléter le contexte commercial des métadonnées des actifs du catalogue. Il s'agit d'un mécanisme extensible permettant aux propriétaires de données d'enrichir la ressource avec des informations qui peuvent aider les utilisateurs des données lorsqu'ils recherchent et trouvent ces données. Les formulaires de métadonnées peuvent également servir de mécanisme pour assurer la cohérence de tous les actifs publiés dans le DataZone catalogue Amazon.

Une définition de formulaire de métadonnées est composée d'une ou de plusieurs définitions de champs, avec prise en charge des types de données booléens, datés, décimaux, entiers, chaînes et valeurs de champs du glossaire commercial. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer, modifier ou supprimer des champs dans les formulaires de métadonnées de votre DataZone domaine Amazon, vous devez être membre du projet propriétaire et disposer des informations d'identification appropriées.

Pour supprimer un champ dans un formulaire de métadonnées, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Accédez au menu Catalogue dans la barre de navigation supérieure à côté de Rechercher.
3. Dans Amazon DataZone Data Portal, choisissez Formulaires de métadonnées, puis choisissez le formulaire de métadonnées dans lequel vous souhaitez supprimer des champs.
4. Sur la page de détails du formulaire, choisissez le champ que vous souhaitez supprimer, puis développez Actions et choisissez Supprimer.
5. Confirmez la suppression en choisissant Supprimer.

DataZone Projets et environnements Amazon

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de ressources de données du DataZone catalogue Amazon. Chaque DataZone projet Amazon est soumis à un ensemble de contrôles d'accès afin que seuls les individus, groupes et rôles autorisés puissent accéder au projet et aux actifs de données auxquels ce projet est abonné, et ne puissent utiliser que les outils définis par les autorisations du projet. Les projets agissent en tant que principal d'identité qui reçoit des autorisations d'accès aux ressources sous-jacentes, ce qui permet DataZone à Amazon d'opérer au sein de l'infrastructure d'une organisation sans se fier aux informations d'identification des utilisateurs individuels.

Dans Amazon DataZone, un environnement est un ensemble de ressources configurées (par exemple, un compartiment Amazon S3, une AWS Glue base de données ou un groupe de travail Amazon Athena), avec un ensemble donné de principes IAM (avec des autorisations de contributeur attribuées) qui peuvent opérer sur ces ressources. Chaque environnement peut également avoir des utilisateurs principaux autorisés à accéder aux ressources et aux données par le biais d'un abonnement et d'un traitement des commandes. Les environnements sont conçus pour stocker des liens exploitables vers des AWS services, des applications externes IDEs et des consoles. Les membres du projet peuvent accéder à des services tels que la console Amazon Athena et bien d'autres via des liens profonds configurés dans un environnement. Les utilisateurs SSO et IAM du projet peuvent être limités à use/access des environnements spécifiques.

Dans Amazon DataZone, vous créez des environnements à l'aide de modèles appelés profils d'environnement. Les profils d'environnement, quant à eux, sont créés à l'aide de plans de AWS service intégrés et personnalisés. Avec les profils d'environnement, les administrateurs de domaine peuvent encapsuler des plans avec des paramètres préconfigurés, puis les travailleurs des données peuvent créer rapidement un certain nombre de nouveaux environnements en sélectionnant les profils d'environnement existants et en spécifiant les noms des nouveaux environnements. Cela permet aux travailleurs des données de gérer efficacement leurs projets et leurs environnements tout en s'assurant qu'ils respectent les politiques de gouvernance des données appliquées par leurs administrateurs de domaine.

Pour de plus amples informations, consultez DataZone Terminologie et concepts d'Amazon.

Rubriques

- Création d'un profil d'environnement

- Modifier un profil d'environnement
- Supprimer un profil d'environnement
- Créez un nouvel environnement
- Modifier un environnement
- Supprimer un environnement
- Création d'un nouveau projet.
- Modifier le projet
- Déplacer le projet vers une autre unité de domaine
- Supprimer le projet
- Quitter le projet
- Ajouter des membres à un projet
- Supprimer des membres d'un projet

Création d'un profil d'environnement

Sur Amazon DataZone, un profil d'environnement est un modèle que vous pouvez utiliser pour créer des environnements. L'objectif d'un profil d'environnement est de simplifier la création d'un environnement en intégrant des informations de placement telles que le AWS compte et la région dans les profils. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour créer des profils d'environnement dans un DataZone domaine Amazon, vous devez appartenir à un DataZone projet Amazon. Tous les profils d'environnement appartiennent aux projets et peuvent être utilisés par tous les utilisateurs autorisés, quel que soit le projet, pour créer de nouveaux environnements.

Pour créer un profil d'environnement

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Dans le portail de données, choisissez Parcourir les projets et sélectionnez le projet dans lequel vous souhaitez créer le profil d'environnement.

3. Accédez à l'onglet Environnements du projet, puis choisissez Créer un profil d'environnement.
4. Configurez les champs suivants :

   - Nom : nom de votre profil d'environnement.
   - Description — (Facultatif) Description de votre profil d'environnement.
   - Projet propriétaire - Le projet dans lequel le profil est créé est sélectionné par défaut dans ce champ.
   - Plan — Le plan pour lequel ce profil est créé. Vous pouvez choisir l'un des DataZone plans Amazon par défaut (Data Lake ou Data Warehouse).

     Si vous avez spécifié le plan de l'entrepôt de données, procédez comme suit :

     - Fournissez un ensemble de paramètres. Pour sélectionner un jeu de paramètres existant, choisissez l'option Choisir un jeu de paramètres. Si vous souhaitez entrer vos propres paramètres, choisissez Enter my own.
     - Si vous choisissez de sélectionner un paramètre existant, procédez comme suit :
       - Sélectionnez un AWS compte dans le menu déroulant.
       - Sélectionnez un ensemble de paramètres dans le menu déroulant.
     - Si vous choisissez de saisir vos propres paramètres, procédez comme suit :
       - Fournissez les AWS paramètres en sélectionnant le AWS compte et la région dans le menu déroulant.
       - Fournissez les paramètres du magasin de données Redshift :
         - Sélectionnez le cluster Amazon Redshift ou Amazon Redshift Serverless
         - Entrez l'ARN AWS secret qui contient les informations d'identification du cluster Amazon Redshift ou du groupe de travail Amazon Redshift Serverless sélectionné. Le AWS secret doit être marqué avec l'identifiant du domaine et l'identifiant du projet dans lesquels vous créez le profil d'environnement.
           - `AmazonDataZoneDomain: [Domain_ID]`
           - `AmazonDataZoneProject: [Project_ID]`
         - Entrez le nom du cluster Amazon Redshift ou du groupe de travail Amazon Redshift Serverless.
         - Entrez le nom de la base de données au sein du cluster Amazon Redshift ou du groupe de travail Amazon Redshift Serverless sélectionné.
       - Dans la section Projets autorisés, spécifiez les projets qui peuvent utiliser le profil d'environnement pour créer des environnements. Par défaut, tous les projets du domaine

peuvent utiliser les profils d'environnement du compte pour créer des environnements. Pour conserver ce paramètre par défaut, choisissez Tous les projets. Toutefois, vous pouvez restreindre cela en affectant des projets autorisés à l'environnement. Pour ce faire, sélectionnez Projets autorisés uniquement, puis spécifiez les projets qui peuvent utiliser ce profil de projet pour créer des environnements.

- Dans la section Publication, choisissez l'une des options suivantes :
  - Publier à partir de n'importe quel schéma : si vous choisissez cette option, les environnements créés à l'aide de ce profil d'environnement peuvent être utilisés pour publier à partir de n'importe quel schéma de base de données sélectionné dans les paramètres Redshift fournis ci-dessus. Les utilisateurs de l'environnement créé à l'aide de ces profils d'environnement peuvent également fournir leurs propres paramètres Amazon Redshift pour publier à partir de n'importe quel schéma du AWS compte et de la région sélectionnés dans le profil d'environnement.
  - Publier uniquement à partir du schéma d'environnement par défaut : si vous choisissez cette option, les environnements créés à l'aide de celle-ci peuvent être utilisés pour publier uniquement à partir du schéma par défaut créé par Amazon DataZone pour cet environnement. Les utilisateurs de l'environnement créé à l'aide de ces profils d'environnement ne peuvent pas fournir leurs propres paramètres Amazon Redshift.
  - Interdire la publication : si vous choisissez cette option, les environnements créés à l'aide de ce profil d'environnement ne peuvent être utilisés que pour l'abonnement et la consommation de données. Les environnements ne peuvent pas du tout être utilisés pour publier des données.

Si vous avez spécifié le plan Data Lake, procédez comme suit :

- Dans la section des paramètres du AWS compte, spécifiez le AWS numéro de AWS compte et la région du compte dans laquelle les environnements potentiels seront créés.
- Dans la section Projets autorisés, spécifiez les projets qui peuvent utiliser le profil d'environnement avec le profil d'environnement Data Lake intégré pour créer des environnements. Par défaut, tous les projets du domaine peuvent utiliser le plan du lac de données du compte pour créer des profils d'environnement. Pour conserver ce paramètre par défaut, choisissez Tous les projets. Toutefois, vous pouvez limiter cela en affectant des projets au plan. Pour ce faire, sélectionnez Projets autorisés uniquement, puis spécifiez les projets qui peuvent utiliser ce profil de projet pour créer des environnements.
- Dans la section Bases de données, choisissez N'importe quelle base de données pour activer la publication à partir de n'importe quelle base de données au sein du AWS compte

et de la région où l'environnement est créé ou choisissez Seule la base de données par défaut pour activer la publication uniquement à partir de la base de données de publication par défaut créée avec l'environnement.

5. Choisissez Créer un profil d'environnement.

Modifier un profil d'environnement

Sur Amazon DataZone, un profil d'environnement est un modèle que vous pouvez utiliser pour créer des environnements. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour modifier un profil d'environnement existant dans un DataZone domaine Amazon, vous devez appartenir à un DataZone projet Amazon.

Pour modifier un profil d'environnement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de données, choisissez Parcourir les projets et sélectionnez le projet dans lequel vous souhaitez modifier le profil d'environnement.
3. Accédez à l'onglet Environnements du projet, puis choisissez Profils d'environnement, puis choisissez le profil d'environnement que vous souhaitez modifier.

   Si vous modifiez un profil d'environnement d'entrepôt de données, vous ne pouvez modifier que le nom et la description d'un profil d'environnement existant.

   Si vous modifiez un profil d'environnement Data Lake, vous pouvez modifier le nom et la description du profil. Vous pouvez également modifier les projets autorisés à utiliser ce profil pour créer des environnements et vous pouvez modifier des bases de données. Pour modifier ces paramètres, procédez comme suit :

   - Dans la section Projets autorisés, spécifiez les projets qui peuvent utiliser le profil d'environnement avec le profil d'environnement Data Lake intégré pour créer des environnements. Par défaut, tous les projets du domaine peuvent utiliser le plan du lac de données du compte pour créer des profils d'environnement. Pour conserver ce paramètre par défaut, choisissez Tous les projets. Toutefois, vous pouvez limiter cela en affectant des projets

au plan. Pour ce faire, sélectionnez Projets autorisés uniquement, puis spécifiez les projets qui peuvent utiliser ce profil de projet pour créer des environnements.

- Dans la section Bases de données, choisissez N'importe quelle base de données pour activer la publication à partir de n'importe quelle base de données au sein du AWS compte et de la région où l'environnement est créé ou choisissez Seule la base de données par défaut pour activer la publication uniquement à partir de la base de données de publication par défaut créée avec l'environnement.

Lorsque vous avez terminé vos modifications, choisissez Modifier le profil d'environnement.

Supprimer un profil d'environnement

Sur Amazon DataZone, un profil d'environnement est un modèle que vous pouvez utiliser pour créer des environnements. L'objectif d'un profil d'environnement est de simplifier la création d'un environnement en intégrant des informations de placement telles que le AWS compte et la région dans les profils. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour supprimer des profils d'environnement dans un DataZone domaine Amazon, vous devez appartenir à un DataZone projet Amazon.

> **Note**
>
> Lorsque vous supprimez un profil d'environnement, vous ne pouvez plus créer d'autres environnements à l'aide de ce profil.

Pour supprimer un profil d'environnement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de données, choisissez Parcourir les projets et sélectionnez le projet dans lequel vous souhaitez supprimer le profil d'environnement.
3. Accédez à l'onglet Environnements du projet, puis choisissez Profils d'environnement, puis choisissez le profil d'environnement que vous souhaitez supprimer.

4. Sélectionnez le profil d'environnement que vous souhaitez supprimer, puis choisissez Actions, Supprimer et confirmez la suppression.

Créez un nouvel environnement

Dans les DataZone projets Amazon, les environnements sont des ensembles de ressources configurées (par exemple, un bucket Amazon S3, une base de données AWS Glue ou un groupe de travail Amazon Athena), avec un ensemble donné de principes IAM (rôles d'utilisateur de l'environnement) auxquels sont attribuées des autorisations de propriétaire ou de contributeur qui peuvent opérer sur ces ressources. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Tout DataZone utilisateur Amazon disposant des autorisations requises pour accéder au portail de données peut créer un DataZone environnement Amazon au sein d'un projet.

Pour créer un nouvel environnement, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Parcourir tous les projets et sélectionnez le projet dans lequel vous souhaitez créer un nouvel environnement.
3. Choisissez Créer un environnement, spécifiez des valeurs pour les champs suivants, puis choisissez Créer un environnement :

   - Nom : nom de l'environnement
   - Description : description de l'environnement
   - Profile d'environnement : choisissez un profil d'environnement existant ou créez-en un nouveau. Un profil d'environnement est un modèle que vous pouvez utiliser pour créer des environnements. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

     Une fois que vous avez sélectionné le profil d'environnement, dans la section Paramètres, spécifiez les valeurs des champs qui font partie de ce profil d'environnement.

Modifier un environnement

Dans les DataZone projets Amazon, les environnements sont des ensembles de ressources configurées (par exemple, un bucket Amazon S3, une base de données AWS Glue ou un groupe de travail Amazon Athena), avec un ensemble donné de principes IAM (avec des autorisations de contributeur attribuées) qui peuvent opérer sur ces ressources. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Tout DataZone utilisateur Amazon disposant des autorisations requises pour accéder au portail de données peut modifier un DataZone environnement Amazon au sein d'un projet.

Pour modifier un environnement existant, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Parcourir les projets dans le volet de navigation supérieur et sélectionnez le projet contenant l'environnement que vous souhaitez modifier.
3. Localisez et sélectionnez l'environnement pour ouvrir sa page de détails. Développez ensuite Actions et choisissez Modifier l'environnement.
4. Modifiez le nom et la description de l'environnement, puis choisissez Enregistrer les modifications.

Supprimer un environnement

Dans les DataZone projets Amazon, les environnements sont des ensembles de ressources configurées (par exemple, un bucket Amazon S3, une base de données AWS Glue ou un groupe de travail Amazon Athena), avec un ensemble donné de principes IAM (avec des autorisations de contributeur attribuées) qui peuvent opérer sur ces ressources. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Tout DataZone utilisateur Amazon disposant des autorisations requises pour accéder au portail de données peut supprimer un DataZone environnement Amazon au sein d'un projet.

Pour supprimer un environnement existant, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Parcourir le projet dans le volet de navigation supérieur et sélectionnez le projet contenant l'environnement que vous souhaitez supprimer.
3. Localisez et choisissez l'environnement pour ouvrir sa page de détails, puis développez Actions et choisissez Supprimer l'environnement.
4. Dans la fenêtre contextuelle Supprimer l'environnement, confirmez la suppression `Delete` en tapant dans le champ, puis choisissez Supprimer l'environnement.

   Vous ne pouvez supprimer correctement un environnement qu'une fois que toutes les entités dépendantes de cet environnement ont été supprimées. Pour supprimer un environnement, vous devez d'abord supprimer toutes ses sources de données et cibles d'abonnement associées.

Création d'un nouveau projet.

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de ressources de données du DataZone catalogue Amazon. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Tout DataZone utilisateur Amazon disposant des autorisations requises pour accéder au portail de données peut créer un DataZone projet Amazon.

Pour créer un nouveau projet, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de DataZone données Amazon, choisissez Create Project.
3. Spécifiez des valeurs pour les champs suivants, puis choisissez Créer un projet :

   - Nom — Le nom du projet.

- Description — Description du projet.
- Unité de domaine : unité de domaine sous laquelle vous souhaitez créer ce projet.

Modifier le projet

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de ressources de données du DataZone catalogue Amazon. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Pour modifier un DataZone projet Amazon, vous devez être le propriétaire de ce projet ou l'administrateur du domaine qui contient ce projet.

Pour modifier un projet existant, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Parcourir les projets.
3. Choisissez le projet que vous souhaitez modifier. Si vous ne le voyez pas facilement dans la liste des projets, vous pouvez le rechercher en spécifiant le nom du projet dans le champ Rechercher un projet.
4. Développez les actions et choisissez Modifier le projet.
5. Mettez à jour le nom et la description du projet, puis choisissez Enregistrer.

Déplacer le projet vers une autre unité de domaine

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de ressources de données du DataZone catalogue Amazon. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour déplacer un DataZone projet Amazon vers une autre unité de domaine, vous devez répondre aux exigences suivantes :

- Vous devez disposer d'une subvention politique pour la création de projets dans l'unité de domaine vers laquelle vous déplacez le projet.
- Tous les membres du projet doivent disposer des autorisations d'adhésion au projet dans l'unité de domaine vers laquelle vous déplacez le projet.
- Vous devez être propriétaire d'une unité de domaine dans l'unité de domaine vers laquelle vous déplacez le projet.
- Vous devez être le propriétaire du projet.

Pour déplacer un projet existant vers une autre unité de domaine, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Parcourir les projets.
3. Choisissez le projet que vous souhaitez déplacer. Si vous ne le voyez pas facilement dans la liste des projets, vous pouvez le rechercher en spécifiant le nom du projet dans le champ Rechercher un projet.
4. Développez les actions et choisissez Déplacer le projet.
5. Spécifiez l'unité de domaine sous laquelle vous souhaitez déplacer ce projet, puis choisissez Déplacer.

Supprimer le projet

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la and/or consommation de ressources de données du DataZone catalogue Amazon. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

L'acte de supprimer un projet est définitif. La suppression entraîne la suppression irrévocable du contenu du projet, notamment les sources de données, les environnements, les ressources, les glossaires et les formulaires de métadonnées. Amazon DataZone révoque les subventions DataZone qu'Amazon a accordées aux actifs gérés via Lake Formation et Amazon Redshift. La suppression d'un projet ne supprime pas les DataZone AWS ressources non Amazon qu'Amazon DataZone peut

vous avoir aidé à créer. Si vous n'avez plus besoin de ces AWS ressources, supprimez-les de leur AWS service et de leur compte respectifs.

Pour supprimer un DataZone projet Amazon, vous devez en être le propriétaire.

Pour supprimer un projet existant, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Un principal IAM peut accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et se connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Parcourir les projets dans le volet de navigation supérieur.
3. Choisissez le projet que vous souhaitez supprimer. S'il ne figure pas dans la liste des projets, vous pouvez le rechercher en spécifiant le nom du projet dans le champ Rechercher un projet.
4. Développez les actions et choisissez Supprimer le projet.

   Consultez les avertissements relatifs à l'impact potentiel de la suppression du projet.
5. Si vous acceptez les avertissements, saisissez le texte de confirmation et choisissez Supprimer.

> **Important**
>
> La suppression d'un projet est une action irrévocable qui ne peut être annulée ni par vous ni par vous. AWS

> **Note**
>
> Lorsque vous ou les utilisateurs de votre domaine créez un environnement dans un projet, Amazon DataZone crée AWS des ressources dans votre domaine ou dans les comptes associés afin de vous fournir, ainsi qu'aux utilisateurs de votre domaine, des fonctionnalités. Vous trouverez ci-dessous la liste des AWS ressources qu'Amazon DataZone peut créer pour un projet, ainsi que le nom par défaut. La suppression d'un projet ne supprime aucune de ces AWS ressources de vos AWS comptes.
>
> - <environmentId>Rôles IAM : datazone_usr_.

- <environmentName>Bases de données Glue : (1) <environmentName>_pub_db-*, (2) _sub_db-*. S'il existe déjà une base de données portant ce nom, Amazon DataZone ajoutera l'ID d'environnement.
- <environmentName>Groupes de travail Athena : -*. S'il existe déjà un groupe de travail portant ce nom, Amazon DataZone ajoutera l'ID d'environnement.
- CloudWatch groupe de journaux : datazone_ <environmentId>

Quitter le projet

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de ressources de données du DataZone catalogue Amazon. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour quitter un projet existant, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet.
3. Choisissez le projet que vous souhaitez quitter. Si vous ne le voyez pas facilement dans la liste des projets, vous pouvez le rechercher en spécifiant le nom du projet dans le champ Rechercher un projet.
4. Développez les actions et choisissez Quitter le projet.

Ajouter des membres à un projet

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de ressources de données du DataZone catalogue Amazon. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Vous devez être propriétaire ou contributeur du projet pour ajouter des membres à un projet. Vous pouvez ajouter des groupes SSO, des utilisateurs SSO ou des principaux IAM (rôles ou utilisateurs) en tant que membres du projet.

Pour ajouter des membres à un projet existant, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet.
3. Choisissez le projet auquel vous souhaitez ajouter des membres. Si vous ne le voyez pas facilement dans la liste des projets, vous pouvez le rechercher en spécifiant le nom du projet dans le champ Rechercher un projet.
4. Sur la page de détails du projet, sélectionnez l'onglet Membres, puis le nœud Choisir tous les membres.
5. Dans l'onglet Membres du projet, choisissez Ajouter des membres.
6. Dans la fenêtre contextuelle Ajouter des membres au projet, spécifiez le ou les utilisateurs que vous souhaitez ajouter et précisez leur rôle au sein du projet (propriétaire, contributeur, consommateur, délégué ou spectateur), puis choisissez Ajouter des membres.

Important

Vous ne pouvez ajouter en tant que membres du projet que les utilisateurs autorisés à être membres de ce projet conformément à la politique d'autorisation d'appartenance au projet configurée pour l'unité de domaine dans laquelle réside ce projet. Pour plus d'informations, consultez Attribuer des politiques d'autorisation aux utilisateurs et aux groupes au sein d'une unité DataZone de domaine Amazon.

Note

Vous pouvez ajouter un directeur IAM en tant que membre du projet s'il possède déjà un profil DataZone utilisateur Amazon dans le domaine. Amazon crée DataZone automatiquement un profil utilisateur pour un principal IAM lorsqu'il interagit avec succès avec

le domaine via le portail, l'API ou la CLI. Vous ne pouvez pas créer de profil utilisateur pour un directeur IAM. Pour ajouter les principaux IAM en tant que membres du projet dans le cas où le principal IAM ne possède pas de profil DataZone utilisateur Amazon existant dans le domaine, demandez à votre administrateur d'ajouter les deux autorisations IAM suivantes à celles de votre domaine AmazonDataZoneDomainExecutionRoledans la console IAM : et. `iam:GetUser` `iam:GetRole` Séparément, pour effectuer des actions dans le domaine, le principal IAM doit disposer des autorisations IAM correspondantes pour ces actions.

Supprimer des membres d'un projet

Dans Amazon DataZone, les projets permettent à un groupe d'utilisateurs de collaborer sur divers cas d'utilisation commerciale impliquant la publication, la découverte, l'abonnement et la consommation de ressources de données du DataZone catalogue Amazon. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. Vous devez être propriétaire d'un projet pour supprimer des membres d'un projet.

Pour supprimer des membres d'un projet existant, procédez comme suit.

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet.
3. Choisissez le projet dans lequel vous souhaitez supprimer des membres. Si vous ne le voyez pas facilement dans la liste des projets, vous pouvez le rechercher en spécifiant le nom du projet dans le champ Rechercher un projet.
4. Sur la page de détails du projet, sélectionnez l'onglet Membres, puis le nœud Choisir tous les membres.
5. Dans l'onglet Membres du projet, choisissez le ou les membres que vous souhaitez supprimer du projet, puis sélectionnez Supprimer.
6. Dans la fenêtre contextuelle Supprimer des membres, confirmez la suppression en choisissant Supprimer des membres.

Inventaire et publication des données sur Amazon DataZone

Cette section décrit les tâches et les procédures que vous souhaitez effectuer afin de créer un inventaire de vos données sur Amazon DataZone et de publier vos données sur Amazon DataZone.

Pour pouvoir utiliser Amazon DataZone pour cataloguer vos données, vous devez d'abord les importer (actifs) en tant qu'inventaire de votre projet sur Amazon DataZone. La création d'un inventaire pour un projet particulier rend les actifs accessibles uniquement aux membres de ce projet. Les ressources de l'inventaire du projet ne sont pas accessibles à tous les utilisateurs du domaine, search/browse sauf si elles sont publiées explicitement. Après avoir créé un inventaire de projet, les propriétaires de données peuvent organiser leurs actifs d'inventaire avec les métadonnées commerciales requises en ajoutant ou en mettant à jour les noms commerciaux (actif et schéma), les descriptions (actif et schéma), lisez-moi, les termes du glossaire (actif et schéma) et les formulaires de métadonnées.

L'étape suivante de l'utilisation d'Amazon DataZone pour cataloguer vos données consiste à rendre les actifs d'inventaire de votre projet accessibles aux utilisateurs du domaine. Vous pouvez le faire en publiant les actifs d'inventaire dans le DataZone catalogue Amazon. Seule la dernière version de la ressource d'inventaire peut être publiée dans le catalogue et seule la dernière version publiée est active dans le catalogue de découverte. Si un actif d'inventaire est mis à jour après sa publication dans le DataZone catalogue Amazon, vous devez le republier explicitement pour que la dernière version figure dans le catalogue de découverte.

Pour de plus amples informations, consultez DataZone Terminologie et concepts d'Amazon.

Rubriques

- Configurer les autorisations de Lake Formation pour Amazon DataZone
- Créez des types d'actifs personnalisés sur Amazon DataZone
- Créez et exécutez une source DataZone de données Amazon pour AWS Glue Data Catalog
- Création et gestion d'une source de DataZone données Amazon pour Amazon Redshift
- Modifier une source de données dans Amazon DataZone
- Supprimer une source de données dans Amazon DataZone
- Publier des actifs dans le DataZone catalogue Amazon à partir de l'inventaire du projet
- Gérez l'inventaire et organisez les actifs sur Amazon DataZone
- Création manuelle d'un actif sur Amazon DataZone

- Dépublier un actif du catalogue Amazon DataZone
- Supprimer un DataZone actif Amazon
- Lancer manuellement l'exécution d'une source de données dans Amazon DataZone
- Révisions des actifs sur Amazon DataZone
- Qualité des données sur Amazon DataZone
- Utilisation de l'apprentissage automatique et de l'IA générative sur Amazon DataZone
- Lignage des données sur Amazon DataZone
- Règles d'application des métadonnées pour la publication

Configurer les autorisations de Lake Formation pour Amazon DataZone

Lorsque vous créez un environnement à l'aide du plan de lac de données intégré (DefaultDataLake), une base de données AWS Glue est ajoutée à Amazon dans DataZone le cadre du processus de création de cet environnement. Si vous souhaitez publier des actifs à partir de cette base de données AWS Glue, aucune autorisation supplémentaire n'est requise.

Toutefois, si vous souhaitez publier des actifs et vous abonner à des actifs provenant d'une base de données AWS Glue qui existe en dehors de votre DataZone environnement Amazon, vous devez explicitement fournir à Amazon DataZone les autorisations d'accès aux tables de cette base de données AWS Glue externe. Pour ce faire, vous devez définir les paramètres suivants dans AWS Lake Formation et associer les autorisations Lake Formation nécessaires auAmazonDataZoneGlueAccess- <region>- <domainId>.

- Configurez l'emplacement Amazon S3 de votre lac de données dans AWS Lake Formation avec le mode d'autorisation Lake Formation ou le mode d'accès hybride. Pour plus d'informations, consultez le https://docs.aws.amazon.com/lake-formation/latest/dg/registerfichier -data-lake.html.
- Supprimez l'`IAMAllowedPrincipals`autorisation des tables Amazon Lake Formation pour lesquelles Amazon DataZone gère les autorisations. Pour plus d'informations, voir https://docs.aws.amazon.com/lake-formation/latest/dg/upgrade- glue-lake-formation-background .html.
- Associez les autorisations de AWS Lake Formation suivantes à AmazonDataZoneGlueAccess- <region>- <domainId> :
  - `Describe`et `Describe grantable` des autorisations sur la base de données où se trouvent les tables

- `Describe`,`Select`,`Describe Grantable`, `Select Grantable` autorisations sur toutes les tables de la base de données ci-dessus auxquelles vous DataZone souhaitez gérer l'accès en votre nom.

> **Note**
>
> Amazon DataZone prend en charge le mode hybride AWS Lake Formation. Le mode hybride Lake Formation vous permet de commencer à gérer les autorisations sur vos bases de données et tables AWS Glue via Lake Formation, tout en conservant les autorisations IAM existantes sur ces tables et bases de données. Pour de plus amples informations, consultez DataZone Intégration d'Amazon au mode hybride de AWS Lake Formation.

Pour de plus amples informations, veuillez consulter Résolution des problèmes liés aux autorisations de AWS Lake Formation pour Amazon DataZone.

DataZone Intégration d'Amazon au mode hybride de AWS Lake Formation

Amazon DataZone est intégré au mode hybride AWS Lake Formation. Cette intégration vous permet de publier et de partager facilement vos tables AWS Glue via Amazon DataZone sans avoir à les enregistrer au préalable dans AWS Lake Formation. Le mode hybride vous permet de commencer à gérer les autorisations sur vos tables AWS Glue via AWS Lake Formation tout en conservant les autorisations IAM existantes sur ces tables.

Pour commencer, vous pouvez activer le paramètre d'enregistrement de l'emplacement des données dans le DefaultDataLakeplan de la console de DataZone gestion Amazon.

Permettre l'intégration avec AWS le mode hybride Lake Formation

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez activer l'intégration avec le mode hybride AWS Lake Formation.
3. Sur la page des détails du domaine, accédez à l'onglet Blueprints.
4. Dans la liste des plans, choisissez le DefaultDataLakeplan.

5. Assurez-vous que le DefaultDataLake plan est activé. S'il n'est pas activé, suivez les étapes décrites Activez les plans intégrés dans le AWS compte propriétaire du domaine Amazon DataZone pour l'activer dans votre AWS compte.
6. Sur la page de DefaultDataLake détails, ouvrez l'onglet Provisioning et cliquez sur le bouton Modifier dans le coin supérieur droit de la page.
7. Sous Enregistrement de l'emplacement des données, cochez la case pour activer l'enregistrement de l'emplacement des données.
8. Pour le rôle de gestion de l'emplacement des données, vous pouvez créer un nouveau rôle IAM ou sélectionner un rôle IAM existant. Amazon DataZone utilise ce rôle pour gérer l'accès en lecture/écriture au ou aux compartiments Amazon S3 choisis pour Data Lake en utilisant le mode d'accès hybride AWS Lake Formation. Pour de plus amples informations, veuillez consulter AmazonDataZone<region>S3 Manage- - <domainId>.
9. Vous pouvez éventuellement choisir d'exclure certains sites Amazon S3 si vous ne souhaitez pas qu'Amazon DataZone les enregistre automatiquement en mode hybride. Pour cela, procédez comme suit :

    - Cliquez sur le bouton à bascule pour exclure les sites Amazon S3 spécifiés.
    - Indiquez l'URI du compartiment Amazon S3 que vous souhaitez exclure.
    - Pour ajouter des compartiments supplémentaires, choisissez Ajouter un emplacement S3.

    > **Note**
    >
    > Amazon autorise DataZone uniquement l'exclusion d'un emplacement S3 racine. Tous les emplacements S3 situés sur le chemin d'un emplacement S3 racine seront automatiquement exclus de l'enregistrement.

    - Sélectionnez Enregistrer les modifications.

Une fois que vous avez activé le paramètre d'enregistrement de la localisation des données dans votre AWS compte, lorsqu'un consommateur de données s'abonne à une table AWS Glue gérée via des autorisations IAM, Amazon enregistre d'abord les emplacements Amazon S3 de cette table en mode hybride, puis accorde l'accès au consommateur de données en gérant les autorisations sur la table via AWS Lake Formation. DataZone Cela garantit que les autorisations IAM disponibles continuent d'exister avec les autorisations AWS Lake Formation récemment accordées, sans perturber les flux de travail existants.

Comment gérer les emplacements Amazon S3 chiffrés lors de l'activation de l'intégration du mode hybride de AWS Lake Formation dans Amazon DataZone

Si vous utilisez un emplacement Amazon S3 chiffré à l'aide d'une clé KMS AWS gérée par le client ou gérée par le client, le rôle AmazonDataZoneS3Manage doit être autorisé à chiffrer et à déchiffrer les données avec la clé KMS, ou la politique de clé KMS doit accorder des autorisations sur la clé du rôle.

Si votre position Amazon S3 est chiffrée à l'aide d'une clé AWS gérée, ajoutez la politique en ligne suivante au AmazonDataZoneDataLocationManagementrôle :

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:Encrypt",
                "kms:Decrypt",
                "kms:ReEncrypt*",
                "kms:GenerateDataKey*",
                "kms:DescribeKey"
            ],
            "Resource": "arn:aws:kms:us-
east-1:111122223333:key/1234abcd-12ab-34cd-56ef-1234567890ab"
        }
    ]
}
```

Si votre position Amazon S3 est chiffrée à l'aide d'une clé gérée par le client, procédez comme suit :

1. Ouvrez la console AWS KMS à l'https://console.aws.amazon.comadresse /kms et connectez-vous en tant qu'utilisateur administratif d' AWS Identity and Access Management (IAM) ou en tant qu'utilisateur habilité à modifier la politique de clé KMS utilisée pour chiffrer l'emplacement.
2. Dans le volet de navigation, sélectionnez Clés gérées par le client, puis choisissez le nom de la clé KMS souhaitée.

3. Sur la page de détails des clés KMS, choisissez l'onglet Politique clé, puis effectuez l'une des opérations suivantes pour ajouter votre rôle personnalisé ou le rôle lié au service Lake Formation en tant qu'utilisateur clé KMS :

   - Si la vue par défaut s'affiche (avec les sections Administrateurs clés, Suppression des clés, Utilisateurs clés et Autres AWS comptes), dans la section Utilisateurs clés, ajoutez le AmazonDataZoneDataLocationManagementrôle.
   - Si la politique clé (JSON) s'affiche, modifiez la politique pour ajouter AmazonDataZoneDataLocationManagementun rôle à l'objet « Autoriser l'utilisation de la clé », comme indiqué dans l'exemple suivant

```
...
        {
            "Sid": "Allow use of the key",
            "Effect": "Allow",
            "Principal": {
                "AWS": [
                    "arn:aws:iam::111122223333:role/service-role/
AmazonDataZoneDataLocationManage-<region>-<domain-id>"
                ]
            },
            "Action": [
                "kms:Encrypt",
                "kms:Decrypt",
                "kms:ReEncrypt*",
                "kms:GenerateDataKey*",
                "kms:DescribeKey"
            ],
            "Resource": "*"
        },
        ...
```

> **Note**
>
> Si la clé KMS ou l'emplacement Amazon S3 ne se trouvent pas dans le même AWS compte que le catalogue de données, suivez les instructions de la section Enregistrement d'un emplacement Amazon S3 chiffré sur plusieurs AWS comptes.

Créez des types d'actifs personnalisés sur Amazon DataZone

Sur Amazon DataZone, les actifs représentent des types spécifiques de ressources de données, tels que des tables de base de données, des tableaux de bord ou des modèles d'apprentissage automatique. Pour garantir la cohérence et la standardisation lors de la description des actifs du catalogue, un DataZone domaine Amazon doit disposer d'un ensemble de types d'actifs qui définissent la manière dont les actifs sont représentés dans le catalogue. Un type d'actif définit le schéma d'un type d'actif spécifique. Un type de ressource comporte un ensemble de types de formulaires de métadonnées nominables obligatoires et facultatifs (par exemple, GovForm ou GovernanceFormType). Les types d'actifs sur Amazon DataZone sont versionnés. Lorsque des actifs sont créés, ils sont validés par rapport au schéma défini par leur type d'actif (généralement la dernière version), et si une structure non valide est spécifiée, la création des actifs échoue.

Types de ressources système : Amazon DataZone fournit des types de ressources système appartenant au service (y compris GlueTableAssetType, GlueViewAssetType, RedshiftTableAssetType RedshiftViewAssetType, et S3ObjectCollectionAssetType) et des types de formulaires système (y compris DataSourceReferenceFormType AssetCommonDetailsFormType, et SubscriptionTermsFormType). Les types de ressources du système ne peuvent pas être modifiés.

Types de ressources personnalisés : pour créer des types de ressources personnalisés, vous devez commencer par créer les types de formulaires de métadonnées et les glossaires requis à utiliser dans les types de formulaires. Vous pouvez ensuite créer des types de ressources personnalisés en spécifiant le nom, la description et les formulaires de métadonnées associés, qui peuvent être obligatoires ou facultatifs.

Pour les types d'actifs contenant des données structurées, pour représenter le schéma des colonnes dans le portail de données, vous pouvez utiliser le `RelationalTableFormType` pour ajouter les métadonnées techniques à vos colonnes (y compris les noms des colonnes, les descriptions et les types de données) et le `ColumnBusinessMetadataForm` pour ajouter les descriptions commerciales des colonnes, y compris les noms commerciaux, les termes du glossaire et les paires clé-valeur personnalisées.

Pour créer un type de ressource personnalisé via le portail de données, procédez comme suit :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.

2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet dans lequel vous souhaitez créer un type de ressource personnalisé.
3. Accédez à l'onglet Données du projet.
4. Choisissez Types de ressources dans le volet de navigation de gauche, puis choisissez Créer un type de ressource.
5. Spécifiez ce qui suit, puis choisissez Create.
   - Nom : nom du type de ressource personnalisé
   - Description : description du type de ressource personnalisé.
   - Choisissez Ajouter des formulaires de métadonnées pour ajouter des formulaires de métadonnées à ce type de ressource personnalisé.
6. Une fois le type d'actif personnalisé créé, vous pouvez l'utiliser pour créer des actifs.

Pour créer un type de ressource personnalisé via le APIs, procédez comme suit :

1. Créez un type de formulaire de métadonnées en appelant l'action `CreateFormType` API.

   Voici un SageMaker exemple d'Amazon :

```
m_model = "

structure SageMakerModelFormType {
   @required
   @amazon.datazone#searchable
   modelName: String

   @required
   modelArn: String

   @required
   creationTime: String
}
"

CreateFormType(
    domainIdentifier="my-dz-domain",
    owningProjectIdentifier="d4bywm0cja1dbb",
    name="SageMakerModelFormType",
```

```
    model=m_model
    status="ENABLED"
    )
```

2. Vous pouvez ensuite créer un type de ressource en appelant l'action `CreateAssetType` API. Vous pouvez créer des types de ressources uniquement via Amazon DataZone APIs en utilisant les types de formulaires système disponibles (`SubscriptionTermsFormType`dans l'exemple ci-dessous) ou vos types de formulaires personnalisés. Pour les types de formulaires système, le nom du type doit commencer par`amazon.datazone`.

```
CreateAssetType(
    domainIdentifier="my-dz-domain",
    owningProjectIdentifier="d4bywm0cja1dbb",
    name="SageMakerModelAssetType",
    formsInput={
        "SageMakerModelForm": {
            "typeIdentifier": "SageMakerModelFormType",
            "typeRevision": 7,
            "required": True,
        },
        "SubscriptionTerms": {
            "typeIdentifier": "amazon.datazone.SubscriptionTermsFormType",
            "typeRevision": 1,
            "required": False,
        },
    },
)
```

Voici un exemple de création d'un type de ressource pour les données structurées :

```
CreateAssetType(
    domainIdentifier="my-dz-domain",
    owningProjectIdentifier="d4bywm0cja1dbb",
    name="OnPremMySQLAssetType",
    formsInput={
        "OnpremMySQLForm": {
            "typeIdentifier": "OnpremMySQLFormType",
            "typeRevision": 5,
            "required": True,
```

```
        },
        "RelationalTableForm": {
            "typeIdentifier": "amazon.datazone.RelationalTableFormType",
            "typeRevision": 1,
            "required": True,
        },
        "ColumnBusinessMetadataForm": {
            "typeIdentifier": "amazon.datazone.ColumnBusinessMetadataFormType",
            "typeRevision": 1,
            "required": False,
        },
        "SubscriptionTerms": {
            "typeIdentifier": "amazon.datazone.SubscriptionTermsFormType",
            "typeRevision": 1,
            "required": False,
        },
    },
)
```

3. Vous pouvez désormais créer un actif à l'aide des types d'actifs personnalisés que vous avez créés dans les étapes ci-dessus.

```
CreateAsset(
   domainIdentifier="my-dz-domain",
   owningProjectIdentifier="d4bywm0cja1dbb",
   typeIdentifier="SageMakerModelAssetType",
   name="MyModelAsset",
   glossaryTerms="xxx",
   formsInput=[{
        "formName": "SageMakerModelForm",
        "typeIdentifier": "SageMakerModelFormType",
        "content": "{\n \"ModelName\" : \"sample-ModelName\",\n \"ModelArn\" :
 \"999999911111\",\n \"CreationTime\" : \"2025-01-01 18:00:00.000\"}"
        }
        ]
)
```

Et dans cet exemple, vous créez un actif de données structuré :

```
CreateAsset(
   domainIdentifier="my-dz-domain",
   owningProjectIdentifier="d4bywm0cja1dbb",
   typeIdentifier="OnPremMySQLAssetType",
   name="MyModelAsset",
   glossaryTerms="xxx",
   formsInput=[{
        "formName": "RelationalTableForm",
        "typeIdentifier": "amazon.datazone.RelationalTableFormType",
        "content": ".."
        },
        {
        "formName": "OnpremMySQLForm",
        "typeIdentifier": "OnpremMySQLFormType",
        "content": ".."
        },
        {
        "formName": "mySQLTableForm",
        "typeIdentifier": "MySQLTableFormType",
        "typeRevision": "1",
        "content": ".."
        },
        {
        "formName": "AssetCommonDetailsForm",
        "typeIdentifier": "amazon.datazone.AssetCommonDetailsFormType",
        "content": "..."
        },
        .....
        ]
)
```

Créez et exécutez une source DataZone de données Amazon pour AWS Glue Data Catalog

Dans Amazon DataZone, vous pouvez créer une source de AWS Glue Data Catalog données afin d'importer les métadonnées techniques des tables de base de données AWS Glue. Pour ajouter une source de données pour le AWS Glue Data Catalog, la base de données source doit déjà exister dans AWS Glue.

Lorsque vous créez et gérez une source de AWS Glue données, vous ajoutez des actifs de la AWS Glue base de données source à l'inventaire de votre DataZone projet Amazon. Vous pouvez exécuter vos sources de AWS Glue données selon un calendrier défini ou à la demande pour créer ou mettre à jour les métadonnées techniques de vos actifs. Pendant l'exécution des sources de données, vous pouvez éventuellement choisir de publier vos actifs dans le DataZone catalogue Amazon afin de les rendre accessibles à tous les utilisateurs du domaine. Vous pouvez également publier les actifs de l'inventaire de votre projet après avoir modifié leurs métadonnées commerciales. Les utilisateurs du domaine peuvent rechercher et découvrir vos actifs publiés, et demander des abonnements à ces actifs.

Pour ajouter une source AWS Glue de données

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel vous souhaitez ajouter la source de données.
3. Accédez à l'onglet Données du projet.
4. Choisissez Sources de données dans le volet de navigation de gauche, puis choisissez Créer une source de données.
5. Configurez les champs suivants :

   - Nom : nom de la source de données.
   - Description — Description de la source de données.
6. Sous Type de source de données, sélectionnez AWS Glue.
7. Sous Sélectionnez un environnement, spécifiez l'environnement dans lequel vous souhaitez publier les AWS Glue tables.
8. Sous Sélection des données, fournissez une AWS Glue base de données et entrez vos critères de sélection de table. Par exemple, si vous choisissez Inclure et entrez`*corporate`, la base de données inclura toutes les tables sources qui se terminent par le mot`corporate`.

   Vous pouvez choisir une AWS Glue base de données dans la liste déroulante ou saisir un nom de base de données. La liste déroulante inclut deux bases de données : la base de données de publication et la base de données d'abonnement de l'environnement. Si vous souhaitez importer

des actifs d'une base de données qui n'est pas créée par l'environnement, vous devez saisir le nom de la base de données au lieu de le sélectionner dans le menu déroulant.

Vous pouvez ajouter plusieurs règles d'inclusion et d'exclusion pour les tables d'une même base de données. Vous pouvez également ajouter plusieurs bases de données à l'aide du bouton Ajouter une autre base de données.

9. Sous Qualité des données, vous pouvez choisir d'activer la qualité des données pour cette source de données. Dans ce cas, Amazon DataZone importe vos résultats de qualité des données AWS Glue existants dans votre DataZone catalogue Amazon. Par défaut, Amazon DataZone importe les 100 derniers rapports de qualité existants sans date d'expiration depuis AWS Glue.

   Les indicateurs de qualité des données d'Amazon vous DataZone aident à comprendre l'exhaustivité et l'exactitude de vos sources de données. Amazon DataZone extrait ces indicateurs de qualité des données de AWS Glue afin de fournir du contexte à un moment donné, par exemple lors d'une recherche dans un catalogue de données commerciales. Les utilisateurs des données peuvent voir comment les indicateurs de qualité des données évoluent au fil du temps pour les actifs auxquels ils ont souscrit. Les producteurs de données peuvent ingérer les scores de qualité des données de AWS Glue selon un calendrier. Le catalogue de données Amazon DataZone Business peut également afficher des indicateurs de qualité des données provenant de systèmes tiers grâce à la qualité des données APIs. Pour de plus amples informations, consultez Qualité des données sur Amazon DataZone.
10. Choisissez Suivant.
11. Pour les paramètres de publication, choisissez si les actifs sont immédiatement détectables dans le catalogue de données commerciales. Si vous les ajoutez uniquement à l'inventaire, vous pourrez choisir les conditions d'abonnement ultérieurement et les publier dans le catalogue de données commerciales.
12. Pour la génération automatique de noms commerciaux, choisissez si vous souhaitez générer automatiquement des métadonnées pour les actifs lors de leur importation depuis la source.
13. (Facultatif) Pour les formulaires de métadonnées, ajoutez des formulaires pour définir les métadonnées collectées et enregistrées lorsque les actifs sont importés dans Amazon DataZone. Pour de plus amples informations, veuillez consulter the section called “Création d'un formulaire de métadonnées”.
14. Pour la préférence Exécuter, choisissez quand exécuter la source de données.

- Exécuter selon un calendrier : spécifiez les dates et heures d'exécution de la source de données.
- Exécuter à la demande : vous pouvez lancer manuellement des exécutions de sources de données.

15. Choisissez Suivant.
16. Vérifiez la configuration de votre source de données et choisissez Create.

> **Note**
>
> Lorsqu'une source de données AWS Glue est créée, Amazon DataZone crée les autorisations « lecture seule » de Lake Formation pour le rôle IAM de l'environnement utilisé pour créer la source de données afin d'accéder à toutes les tables des bases de données AWS Glue utilisées dans la source de données. Vous pouvez suivre l'état de ces subventions dans les sources de données sur la page de détails de votre environnement. Amazon DataZone ajoute les AWS balises suivantes à la base de données AWS Glue lorsqu'il accorde l'accès au rôle IAM de l'environnement de publication : `DataZoneDiscoverable_${domainId}: true`
> Pour les environnements créés avant la version actuelle d'Amazon DataZone, les membres du projet ne pourront pas voir les tables attribuées dans Amazon Athena.

Création et gestion d'une source de DataZone données Amazon pour Amazon Redshift

Dans Amazon DataZone, vous pouvez créer une source de données Amazon Redshift afin d'importer les métadonnées techniques des tables et des vues de base de données depuis l'entrepôt de données Amazon Redshift. Pour ajouter une source de DataZone données Amazon pour Amazon Redshift, l'entrepôt de données source doit déjà exister dans Amazon Redshift.

Lorsque vous créez et gérez une source de données Amazon Redshift, vous ajoutez des actifs de l'entrepôt de données Amazon Redshift source à l'inventaire de votre projet DataZone Amazon. Vous pouvez exécuter vos sources de données Amazon Redshift selon un calendrier défini ou à la demande pour créer ou mettre à jour les métadonnées techniques de vos actifs. Pendant l'exécution des sources de données, vous pouvez éventuellement choisir de publier les actifs de l'inventaire de votre projet dans le DataZone catalogue Amazon afin de les rendre accessibles à tous les

utilisateurs du domaine. Vous pouvez également publier vos actifs d'inventaire après avoir modifié leurs métadonnées commerciales. Les utilisateurs du domaine peuvent rechercher et découvrir vos actifs publiés et demander des abonnements à ces actifs.

Pour ajouter une source de données Amazon Redshift

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel vous souhaitez ajouter la source de données.
3. Accédez à l'onglet Données du projet.
4. Choisissez Sources de données dans le volet de navigation de gauche, puis choisissez Créer une source de données.
5. Configurez les champs suivants :
   - Nom : nom de la source de données.
   - Description — Description de la source de données.
6. Sous Type de source de données, sélectionnez Amazon Redshift.
7. Sous Sélectionnez un environnement, spécifiez un environnement dans lequel publier les tables Amazon Redshift.
8. En fonction de l'environnement que vous sélectionnez, Amazon DataZone appliquera automatiquement les informations d'identification Amazon Redshift et les autres paramètres directement depuis l'environnement ou vous donnera la possibilité de choisir les vôtres.
   - Si vous avez sélectionné un environnement qui autorise uniquement la publication à partir du schéma Amazon Redshift par défaut de l'environnement, Amazon DataZone appliquera automatiquement les informations d'identification Amazon Redshift et d'autres paramètres, notamment le nom du cluster ou du groupe de travail Amazon Redshift, le secret AWS , le nom de la base de données et le nom du schéma. Vous ne pouvez pas modifier ces paramètres renseignés automatiquement.
   - Si vous sélectionnez un environnement qui n'autorise pas la publication de données, vous ne pourrez pas procéder à la création de la source de données.

- Si vous sélectionnez un environnement qui permet de publier des données à partir de n'importe quel schéma, vous aurez la possibilité d'utiliser les informations d'identification et les autres paramètres Amazon Redshift de l'environnement ou de saisir vos propres identifiants/paramètres.

9. Si vous choisissez d'utiliser vos propres informations d'identification pour créer la source de données, fournissez les informations suivantes :

   - Sous Fournir les informations d'identification Amazon Redshift, choisissez d'utiliser un cluster Amazon Redshift provisionné ou un espace de travail sans serveur Amazon Redshift comme source de données.
   - En fonction de votre sélection à l'étape ci-dessus, choisissez votre cluster ou espace de travail Amazon Redshift dans le menu déroulant, puis choisissez le secret à utiliser pour l'authentification dans AWS Secrets Manager. Vous pouvez choisir un secret existant ou en créer un nouveau.
   - Pour que le secret existant apparaisse dans le menu déroulant, assurez-vous que votre secret dans AWS Secrets Manager inclut les balises suivantes (clé/valeur) :
     - AmazonDataZoneProject: <projectID>
     - AmazonDataZoneDomain: <domainID>

     Si vous choisissez de créer un nouveau secret, celui-ci est automatiquement étiqueté avec les balises référencées ci-dessus et aucune étape supplémentaire n'est nécessaire. Pour plus d'informations, voir Stockage des informations d'identification de base de données dans AWS Secrets Manager.

     Les utilisateurs d'Amazon Redshift dont le code AWS secret a été fourni pour créer la source de données doivent disposer d'`SELECT`autorisations sur les tables à publier. Si vous souhaitez qu'Amazon DataZone gère également les abonnements (accès) en votre nom, les utilisateurs de la base de données AWS secrète doivent également disposer des autorisations suivantes :

     - `CREATE DATASHARE`
     - `ALTER DATASHARE`
     - `DROP DATASHARE`

10. Sous Sélection des données, fournissez une base de données Amazon Redshift, un schéma et entrez les critères de sélection de votre table ou de votre vue. Par exemple, si vous choisissez Inclure et entrez`*corporate`, la ressource inclura toutes les tables sources qui se terminent par le mot`corporate`.

Vous pouvez ajouter plusieurs règles d'inclusion pour les tables d'une même base de données. Vous pouvez également ajouter plusieurs bases de données à l'aide du bouton Ajouter une autre base de données.

11. Choisissez Suivant.
12. Pour les paramètres de publication, choisissez si les actifs sont immédiatement détectables dans le catalogue de données. Si vous les ajoutez uniquement à l'inventaire, vous pourrez choisir les conditions d'abonnement ultérieurement et les publier dans le catalogue de données commerciales.
13. Pour la génération automatique de noms commerciaux, choisissez si vous souhaitez générer automatiquement des métadonnées pour les actifs au fur et à mesure de leur publication et de leur mise à jour à partir de la source.
14. (Facultatif) Pour les formulaires de métadonnées, ajoutez des formulaires pour définir les métadonnées collectées et enregistrées lorsque les actifs sont importés dans Amazon DataZone. Pour de plus amples informations, veuillez consulter the section called “Création d'un formulaire de métadonnées”.
15. Pour la préférence Exécuter, choisissez quand exécuter la source de données.
    - Exécuter selon un calendrier : spécifiez les dates et heures d'exécution de la source de données.
    - Exécuter à la demande : vous pouvez lancer manuellement des exécutions de sources de données.
16. Choisissez Suivant.
17. Vérifiez la configuration de votre source de données et choisissez Create.

> **Note**
>
> Lorsqu'une source de données Amazon Redshift est créée, Amazon DataZone accorde un accès en « lecture seule » à l'environnement utilisé pour créer la source de données afin d'accéder à toutes les tables des schémas Amazon Redshift utilisés dans la source de données. Vous pouvez suivre l'état de ces subventions dans les sources de données sur la page de détails de votre environnement.
> Lorsque vous utilisez un cluster Amazon Redshift ou un groupe de travail sans serveur différent de celui utilisé pour créer l'environnement, vous devez vous assurer que la AWS balise suivante est ajoutée au cluster ou au groupe de travail. Cela est nécessaire pour que

les utilisateurs de l'environnement puissent consulter la base de données autorisée dans Amazon Redshift Query Editor V2 : `DataZoneDiscoverable_${domainId}: true` Pour les environnements créés avant la version actuelle d'Amazon DataZone, les membres du projet ne pourront pas voir les tables attribuées dans Amazon Redshift.

Modifier une source de données dans Amazon DataZone

Après avoir créé une source de DataZone données Amazon, vous pouvez la modifier à tout moment pour modifier les détails de la source ou les critères de sélection des données. Lorsque vous n'avez plus besoin d'une source de données, vous pouvez la supprimer.

Pour effectuer ces étapes, la politique AmazonDataZoneFullAccess AWS gérée doit être jointe. Pour de plus amples informations, veuillez consulter the section called “AWS politiques gérées”.

Vous pouvez modifier une source de DataZone données Amazon pour modifier ses paramètres de sélection de données, notamment en ajoutant, en supprimant ou en modifiant les critères de sélection des tables. Vous pouvez également ajouter et supprimer des bases de données. Vous ne pouvez pas modifier le type de source de données ni l'environnement dans lequel une source de données est publiée.

Pour modifier une source de données

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel appartient la source de données.
3. Accédez à l'onglet Données du projet.
4. Choisissez Sources de données dans le volet de navigation de gauche, puis choisissez la source de données que vous souhaitez modifier.
5. Accédez à l'onglet Définition de la source de données et choisissez Modifier.
6. Apportez vos modifications à la définition de la source de données. Vous pouvez mettre à jour les détails de la source de données et modifier les critères de sélection des données.
7. Une fois les modifications terminées, choisissez Save (Enregistrer).

Supprimer une source de données dans Amazon DataZone

Après avoir créé une source de DataZone données Amazon, vous pouvez la modifier à tout moment pour modifier les détails de la source ou les critères de sélection des données.

Pour effectuer ces étapes, la politique AmazonDataZoneFullAccess AWS gérée doit être jointe. Pour de plus amples informations, veuillez consulter the section called “AWS politiques gérées”.

Lorsque vous n'avez plus besoin d'une source de DataZone données Amazon, vous pouvez la supprimer définitivement. Une fois que vous avez supprimé une source de données, toutes les ressources provenant de cette source de données sont toujours disponibles dans le catalogue et les utilisateurs peuvent toujours s'y abonner. Cependant, les ressources cesseront de recevoir des mises à jour de la source. Nous vous recommandons de déplacer d'abord les ressources dépendantes vers une autre source de données avant de les supprimer.

> **Note**
>
> Vous devez supprimer tous les envois de la source de données avant de pouvoir la supprimer. Pour de plus amples informations, veuillez consulter Découverte, abonnement et consommation de données.

Pour supprimer une source de données

1. Dans l'onglet Données du projet, sélectionnez Sources de données dans le volet de navigation de gauche.
2. Choisissez la source de données que vous souhaitez supprimer.
3. Choisissez Actions, Supprimer la source de données et confirmez la suppression.

Publier des actifs dans le DataZone catalogue Amazon à partir de l'inventaire du projet

Vous pouvez publier les DataZone actifs Amazon et leurs métadonnées à partir des inventaires de projets dans le DataZone catalogue Amazon. Vous ne pouvez publier que la version la plus récente d'une ressource dans le catalogue.

Tenez compte des points suivants lorsque vous publiez des actifs dans le catalogue :

- Pour publier une ressource dans le catalogue, vous devez être le propriétaire ou le contributeur de ce projet.
- Pour les actifs Amazon Redshift, assurez-vous que les clusters Amazon Redshift associés aux clusters d'éditeurs et d'abonnés répondent à toutes les exigences relatives au partage de données Amazon Redshift afin qu'Amazon puisse gérer l'accès aux tables et DataZone aux vues Redshift. Consultez la section Concepts de partage de données pour Amazon Redshift.
- Amazon prend DataZone uniquement en charge la gestion des accès pour les ressources publiées depuis Amazon AWS Glue Data Catalog Redshift et Amazon Redshift. Pour tous les autres actifs, tels que les objets Amazon S3, Amazon DataZone ne gère pas l'accès des abonnés approuvés. Si vous vous abonnez à ces actifs non gérés, le message suivant vous en informe :

  ```
  Subscription approval does not provide access to data. Subscription
  grants on this asset are not managed by Amazon DataZone. For more
  information or help, reach out to your administrator.
  ```

Publier une ressource sur Amazon DataZone

Si vous n'avez pas choisi de rendre les actifs immédiatement détectables dans le catalogue de données lorsque vous avez créé une source de données, effectuez les étapes suivantes pour les publier ultérieurement.

Pour publier un actif

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel appartient la ressource.
3. Accédez à l'onglet Données du projet.
4. Choisissez Données d'inventaire dans le volet de navigation de gauche, puis sélectionnez la ressource que vous souhaitez publier.

> **Note**
>
> Par défaut, toutes les ressources nécessitent une approbation d'abonnement, ce qui signifie que le propriétaire des données doit approuver toutes les demandes d'abonnement à la ressource. Si vous souhaitez modifier ce paramètre avant de publier la ressource, ouvrez les détails de la ressource et choisissez Modifier à côté de l'option Approbation de l'abonnement. Vous pouvez modifier ce paramètre ultérieurement en modifiant et en republiant la ressource.

5. Choisissez Publier la ressource. La ressource est directement publiée dans le catalogue.

   Si vous apportez des modifications à la ressource, par exemple en modifiant ses exigences d'approbation, vous pouvez choisir Republier pour publier les mises à jour du catalogue.

Gérez l'inventaire et organisez les actifs sur Amazon DataZone

Pour pouvoir utiliser Amazon DataZone pour cataloguer vos données, vous devez d'abord les importer (actifs) en tant qu'inventaire de votre projet sur Amazon DataZone. La création d'un inventaire pour un projet particulier rend les actifs accessibles uniquement aux membres de ce projet.

Une fois les actifs créés dans l'inventaire du projet, leurs métadonnées peuvent être conservées. Par exemple, vous pouvez modifier le nom et la description de la ressource ou me lire. Chaque modification de la ressource crée une nouvelle version de la ressource. Vous pouvez utiliser l'onglet Historique de la page de détails de la ressource pour afficher toutes les versions de la ressource.

Vous pouvez modifier la section Lisez-moi et ajouter des descriptions détaillées à la ressource. La section Read Me prend en charge le markdown, vous permettant ainsi de mettre en forme vos descriptions selon vos besoins et de décrire les informations clés sur un actif aux consommateurs.

Les termes du glossaire peuvent être ajoutés au niveau des actifs en remplissant les formulaires disponibles.

Pour modifier le schéma, vous pouvez consulter les colonnes, ajouter des noms commerciaux, des descriptions et ajouter des termes de glossaire au niveau des colonnes.

Si la génération automatique de métadonnées est activée lors de la création de la source de données, les noms commerciaux des actifs et des colonnes peuvent être examinés, acceptés ou rejetés individuellement ou en une seule fois.

Vous pouvez également modifier les conditions d'abonnement pour spécifier si l'approbation de la ressource est requise ou non.

Les formulaires de métadonnées d'Amazon vous DataZone permettent d'étendre le modèle de métadonnées d'un actif de données en ajoutant des attributs personnalisés (par exemple, région de vente, année de vente et trimestre de vente). Les formulaires de métadonnées attachés à un type de ressource sont appliqués à toutes les ressources créées à partir de ce type de ressource. Vous pouvez également ajouter des formulaires de métadonnées supplémentaires à des ressources individuelles dans le cadre de l'exécution de la source de données ou après sa création. Pour créer de nouveaux formulaires, voirthe section called “Création d'un formulaire de métadonnées”.

Pour mettre à jour les métadonnées d'une ressource, vous devez être le propriétaire ou le contributeur du projet auquel appartient la ressource.

Pour mettre à jour les métadonnées d'un actif

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet contenant la ressource dont vous souhaitez mettre à jour les métadonnées.
3. Accédez à l'onglet Données du projet.
4. Choisissez Données d'inventaire dans le volet de navigation de gauche, puis choisissez le nom de la ressource dont vous souhaitez mettre à jour les métadonnées.
5. Sur la page des détails de la ressource, sous Formulaires de métadonnées, choisissez Modifier et modifiez les formulaires existants selon vos besoins. Vous pouvez également joindre des formulaires de métadonnées supplémentaires à la ressource. Pour de plus amples informations, veuillez consulter the section called “Joindre des formulaires de métadonnées supplémentaires aux ressources”.
6. Lorsque vous avez terminé de faire des mises à jour, choisissez Enregistrer le formulaire.

   Lorsque vous enregistrez le formulaire, Amazon DataZone génère une nouvelle version d'inventaire de l'actif. Pour publier la version mise à jour dans le catalogue, sélectionnez Republier la ressource.

Joindre des formulaires de métadonnées supplémentaires aux ressources

Par défaut, les formulaires de métadonnées attachés à un domaine sont attachés à toutes les ressources publiées sur ce domaine. Les éditeurs de données peuvent associer des formulaires de métadonnées supplémentaires à des ressources individuelles afin de fournir un contexte supplémentaire.

Pour joindre des formulaires de métadonnées supplémentaires à une ressource

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet contenant la ressource dont vous souhaitez ajouter les métadonnées.
3. Accédez à l'onglet Données du projet.
4. Choisissez Données d'inventaire dans le volet de navigation de gauche, puis choisissez le nom de la ressource dont vous souhaitez ajouter les métadonnées.
5. Sur la page des détails de la ressource, sous Formulaires de métadonnées, sélectionnez Ajouter des formulaires.
6. Sélectionnez le ou les formulaires à ajouter à la ressource, puis choisissez Ajouter des formulaires.
7. Entrez des valeurs pour chacun des champs de métadonnées, puis choisissez Enregistrer le formulaire.

   Lorsque vous enregistrez le formulaire, Amazon DataZone génère une nouvelle version d'inventaire de l'actif. Pour publier la version mise à jour dans le catalogue, sélectionnez Republier la ressource.

Publier la ressource dans le catalogue après curation sur Amazon DataZone

Une fois satisfait de la curation des actifs, le propriétaire des données peut publier une version des actifs dans le DataZone catalogue Amazon et ainsi la rendre visible par tous les utilisateurs du domaine. La ressource indique la version d'inventaire et la version publiée. Dans le catalogue de

découverte, seule la dernière version publiée apparaît. Si les métadonnées sont mises à jour après la publication, une nouvelle version de l'inventaire sera disponible pour publication dans le catalogue.

Création manuelle d'un actif sur Amazon DataZone

Dans Amazon DataZone, un actif est une entité qui présente un seul objet de données physique (par exemple, un tableau, un tableau de bord, un fichier) ou un objet de données virtuel (par exemple, une vue). Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon. La publication manuelle d'une ressource est une opération ponctuelle. Vous ne spécifiez pas de calendrier d'exécution pour l'actif. Il n'est donc pas mis à jour automatiquement si sa source change.

Pour créer manuellement une ressource par le biais d'un projet, vous devez être le propriétaire ou le contributeur de ce projet.

Pour créer un actif manuellement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet pour lequel vous souhaitez créer la ressource.
3. Accédez à l'onglet Données du projet.
4. Choisissez Sources de données dans le volet de navigation de gauche, puis choisissez Créer une ressource de données.
5. Pour obtenir des informations détaillées sur les actifs, configurez les paramètres suivants :

   - Type d'actif : type d'actif.
   - Nom : nom de la ressource.
   - Description — Description de l'actif.
6. Pour l'emplacement S3, entrez le nom de ressource Amazon (ARN) du compartiment S3 source.

   Entrez éventuellement un point d'accès S3. Pour plus d'informations, consultez la rubrique Gestion de l'accès aux données à l'aide des points d'accès Amazon S3.

7. Pour les paramètres de publication, choisissez si les actifs sont immédiatement détectables dans le catalogue. Si vous les ajoutez uniquement à l'inventaire, vous pourrez choisir les conditions d'abonnement ultérieurement pour les publier dans le catalogue.
8. Choisissez Créer.

   Une fois la ressource créée, elle sera soit directement publiée en tant que ressource active dans le catalogue, soit stockée dans l'inventaire jusqu'à ce que vous décidiez de la publier.

Dépublier un actif du catalogue Amazon DataZone

Lorsque vous dépubliez une DataZone ressource Amazon du catalogue, elle n'apparaît plus dans les résultats de recherche globaux. Les nouveaux utilisateurs ne pourront pas trouver ou s'abonner à la liste des actifs dans le catalogue, mais tous les abonnements existants resteront les mêmes.

Pour dépublier une ressource, vous devez être le propriétaire ou le contributeur du projet auquel la ressource appartient :

Pour annuler la publication d'une ressource

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel appartient la ressource.
3. Accédez à l'onglet Données du projet.
4. Choisissez Données publiées dans le volet de navigation de gauche.
5. Recherchez la ressource dans la liste des ressources publiées, puis choisissez Dépublier.

   La ressource est supprimée du catalogue. Vous pouvez republier la ressource à tout moment en choisissant Publier.

Supprimer un DataZone actif Amazon

Lorsque vous n'avez plus besoin d'un actif sur Amazon DataZone, vous pouvez le supprimer définitivement. La suppression d'une ressource est différente de la dépublication d'une ressource

du catalogue. Vous pouvez supprimer une ressource et sa liste associée dans le catalogue afin qu'elle ne soit visible dans aucun résultat de recherche. Pour supprimer la liste des actifs, vous devez d'abord révoquer tous ses abonnements.

Pour supprimer un actif, vous devez être le propriétaire ou le contributeur du projet auquel appartient l'actif :

> **Note**
>
> Pour supprimer une liste d'actifs, vous devez d'abord révoquer tous les abonnements existants à l'actif. Vous ne pouvez pas supprimer une liste d'actifs qui compte déjà des abonnés.

Pour supprimer un actif

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet contenant la ressource que vous souhaitez supprimer.
3. Accédez à l'onglet Données du projet.
4. Choisissez Données publiées dans le volet de navigation de gauche, puis recherchez et choisissez la ressource que vous souhaitez supprimer. Cela ouvre la page de détails de l'actif.
5. Choisissez Actions, Supprimer et confirmez la suppression.

   Une fois la ressource supprimée, elle n'est plus disponible et les utilisateurs ne peuvent plus s'y abonner.

Lancer manuellement l'exécution d'une source de données dans Amazon DataZone

Lorsque vous exécutez une source de données, Amazon DataZone extrait toutes les métadonnées nouvelles ou modifiées de la source et met à jour les actifs associés dans l'inventaire. Lorsque vous ajoutez une source de données à Amazon DataZone, vous spécifiez la préférence d'exécution de

la source, qui définit si la source s'exécute selon un calendrier ou à la demande. Si votre source s'exécute à la demande, vous devez lancer une exécution manuelle de la source de données.

Même si votre source s'exécute selon un calendrier, vous pouvez toujours l'exécuter manuellement à tout moment. Après avoir ajouté des métadonnées commerciales aux actifs, vous pouvez sélectionner des actifs et les publier dans le DataZone catalogue Amazon afin que tous les utilisateurs du domaine puissent les découvrir. Seuls les actifs publiés sont consultables par les autres utilisateurs du domaine.

Pour exécuter une source de données manuellement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel appartient la source de données.
3. Accédez à l'onglet Données du projet.
4. Choisissez Sources de données dans le volet de navigation de gauche, puis recherchez et choisissez la source de données que vous souhaitez exécuter. Cela ouvre la page de détails de la source de données.
5. Choisissez Exécuter à la demande.

   Le statut de la source de données change au `Running` fur et à mesure qu'Amazon DataZone met à jour les métadonnées des actifs avec les données les plus récentes de la source. Vous pouvez surveiller l'état de l'exécution dans l'onglet Exécutions de la source de données.

Révisions des actifs sur Amazon DataZone

Amazon DataZone augmente la révision d'un actif lorsque vous modifiez ses métadonnées commerciales ou techniques. Ces modifications incluent la modification du nom de la ressource, de la description, des termes du glossaire, des noms de colonnes, des formulaires de métadonnées et des valeurs des champs de formulaires de métadonnées. Ces modifications peuvent résulter de modifications manuelles, de l'exécution de tâches de source de données ou d'opérations d'API. Amazon génère DataZone automatiquement une nouvelle révision d'actif chaque fois que vous apportez une modification à l'actif.

Une fois que vous avez mis à jour une ressource et qu'une nouvelle révision a été générée, vous devez publier la nouvelle révision dans le catalogue pour qu'elle soit mise à jour et disponible pour les abonnés. Pour de plus amples informations, veuillez consulter the section called "Publier des actifs dans le catalogue à partir de l'inventaire du projet". Vous ne pouvez publier que la version la plus récente d'une ressource dans le catalogue.

Pour consulter les révisions passées d'un actif

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet contenant la ressource.
3. Accédez à l'onglet Données du projet, puis localisez et sélectionnez la ressource. Cela ouvre la page de détails de l'actif.
4. Accédez à l'onglet Historique, qui affiche la liste des révisions passées de la ressource.

Qualité des données sur Amazon DataZone

Les indicateurs de qualité des données d'Amazon vous DataZone aident à comprendre les différents indicateurs de qualité tels que l'exhaustivité, l'actualité et l'exactitude de vos sources de données. Amazon DataZone s'intègre à AWS Glue Data Quality et propose APIs d'intégrer des indicateurs de qualité des données issus de solutions de qualité des données tierces. Les utilisateurs des données peuvent voir comment les indicateurs de qualité des données évoluent au fil du temps pour les actifs auxquels ils ont souscrit. Pour créer et appliquer les règles de qualité des données, vous pouvez utiliser l'outil de qualité des données de votre choix, tel que AWS Glue data quality. Grâce aux indicateurs de qualité des données d'Amazon DataZone, les consommateurs de données peuvent visualiser les scores de qualité des données pour les actifs et les colonnes, ce qui contribue à renforcer la confiance dans les données qu'ils utilisent pour prendre des décisions.

Conditions préalables et modifications des rôles IAM

Si vous utilisez les politiques AWS gérées DataZone d'Amazon, aucune étape de configuration supplémentaire n'est requise et ces politiques gérées sont automatiquement mises à jour pour garantir la qualité des données. Si vous utilisez vos propres politiques pour les rôles qui

accordent à Amazon les autorisations requises pour interagir avec DataZone les services pris en charge, vous devez mettre à jour les politiques associées à ces rôles afin de permettre la lecture des informations sur la qualité des données de AWS Glue dans le AWS politique gérée : AmazonDataZoneGlueManageAccessRolePolicy et de permettre la prise en charge des séries chronologiques APIs dans le AWS politique gérée : AmazonDataZoneDomainExecutionRolePolicy et leAWS politique gérée : AmazonDataZoneFullUserAccess.

Permettre la qualité des données pour les actifs AWS de Glue

Amazon DataZone extrait les indicateurs de qualité des données de AWS Glue afin de fournir du contexte à un moment donné, par exemple lors d'une recherche dans un catalogue de données commerciales. Les utilisateurs des données peuvent voir comment les indicateurs de qualité des données évoluent au fil du temps pour les actifs auxquels ils ont souscrit. Les producteurs de données peuvent ingérer les scores de qualité des données de AWS Glue selon un calendrier. Le catalogue de données Amazon DataZone Business peut également afficher des indicateurs de qualité des données provenant de systèmes tiers grâce à la qualité des données APIs. Pour plus d'informations, voir AWS Glue Data Quality et Getting started with AWS Glue Data Quality pour le catalogue de données.

Vous pouvez activer les mesures de qualité des données pour vos DataZone actifs Amazon de différentes manières :

- Utilisez le Data Portal ou Amazon DataZone APIs pour améliorer la qualité des données de votre source de données AWS Glue via le portail de données Amazon, soit lors de la création d'une nouvelle source de DataZone données Glue, soit lors de la modification d'une source de données AWS Glue existante.

  Pour plus d'informations sur l'activation de la qualité des données pour une source de données via le portail, consultezCréez et exécutez une source DataZone de données Amazon pour AWS Glue Data Catalog.

  > **Note**
  >
  > Vous pouvez utiliser le portail de données pour activer la qualité des données uniquement pour vos actifs d'inventaire AWS Glue. Dans cette version d'Amazon, l' DataZone activation de la qualité des données pour Amazon Redshift ou de types personnalisés de ressources via le portail de données n'est pas prise en charge.

Vous pouvez également utiliser le APIs pour améliorer la qualité des données pour vos sources de données nouvelles ou existantes. Vous pouvez le faire en invoquant le CreateDataSourceou UpdateDataSourceet en réglant le `autoImportDataQualityResult` paramètre sur « Vrai ».

Une fois la qualité des données activée, vous pouvez exécuter la source de données à la demande ou selon un calendrier. Chaque exécution peut générer jusqu'à 100 mesures par actif. Il n'est pas nécessaire de créer des formulaires ou d'ajouter des métriques manuellement lors de l'utilisation d'une source de données pour garantir la qualité des données. Lorsque l'actif est publié, les mises à jour apportées au formulaire de qualité des données (jusqu'à 30 points de données par règle d'historique) sont reflétées dans la liste destinée aux consommateurs. Par la suite, chaque nouvel ajout de métriques à l'actif est automatiquement ajouté à la liste. Il n'est pas nécessaire de republier la ressource pour mettre les derniers scores à la disposition des consommateurs.

Permettre la qualité des données pour les types d'actifs personnalisés

Vous pouvez utiliser Amazon DataZone APIs pour garantir la qualité des données pour tous vos types d'actifs personnalisés. Pour plus d'informations, consultez les ressources suivantes :

- PostTimeSeriesDataPoints
- ListTimeSeriesDataPoints
- GetTimeSeriesDataPoint
- DeleteTimeSeriesDataPoints

Les étapes suivantes fournissent un exemple d'utilisation de APIs notre CLI pour importer des métriques tierces pour vos actifs sur Amazon DataZone :

1. Appelez l'`PostTimeSeriesDataPoints`API comme suit :

```
aws datazone post-time-series-data-points  \
--cli-input-json file://createTimeSeriesPayload.json \
```

avec la charge utile suivante :

```
"domainId": "dzd_5oo7xzoqltu8mf",
    "entityId": "4wyh64k2n8czaf",
    "entityType": "ASSET",
    "form": {
        "content": "{\n  \"evaluations\" : [ {\n    \"types\" : [ \"MaximumLength
\" ],\n    \"description\" : \"ColumnLength \\\"ShippingCountry\\\" <= 6\",\n
   \"details\" : { },\n    \"applicableFields\" : [ \"ShippingCountry\" ],\n
   \"status\" : \"PASS\"\n  }, {\n    \"types\" : [ \"MaximumLength\" ],\n
 \"description\" : \"ColumnLength \\\"ShippingState\\\" <= 2\",\n    \"details
\" : { },\n    \"applicableFields\" : [ \"ShippingState\" ],\n    \"status\" :
 \"PASS\"\n  }, {\n    \"types\" : [ \"MaximumLength\" ],\n    \"description
\" : \"ColumnLength \\\"ShippingCity\\\" <= 8\",\n    \"details\" : { },\n
 \"applicableFields\" : [ \"ShippingCity\" ],\n    \"status\" : \"PASS\"\n  },
 {\n    \"types\" : [ \"Completeness\" ],\n    \"description\" : \"Completeness \
\\"ShippingStreet\\\" >= 0.59\",\n    \"details\" : { },\n    \"applicableFields
\" : [ \"ShippingStreet\" ],\n    \"status\" : \"PASS\"\n  }, {\n    \"types\" :
 [ \"MaximumLength\" ],\n    \"description\" : \"ColumnLength \\\"ShippingStreet\\
\" <= 101\",\n    \"details\" : { },\n    \"applicableFields\" : [ \"ShippingStreet
\" ],\n    \"status\" : \"PASS\"\n  }, {\n    \"types\" : [ \"MaximumLength\" ],\n
    \"description\" : \"ColumnLength \\\"BillingCountry\\\" <= 6\",\n    \"details
\" : { },\n    \"applicableFields\" : [ \"BillingCountry\" ],\n    \"status\" :
 \"PASS\"\n  }, {\n    \"types\" : [ \"Completeness\" ],\n    \"description\" :
 \"Completeness \\\"biLlingcountry\\\" >= 0.5\",\n    \"details\" : {\n
 \"EVALUATION_MESSAGE\" : \"Value: 0.26666666666666666 does not meet the constraint
 requirement!\"\n    },\n    \"applicableFields\" : [ \"biLlingcountry\" ],\n
   \"status\" : \"FAIL\"\n  }, {\n    \"types\" : [ \"Completeness\" ],\n
 \"description\" : \"Completeness \\\"Billingstreet\\\" >= 0.5\",\n    \"details
\" : { },\n    \"applicableFields\" : [ \"Billingstreet\" ],\n    \"status\" :
 \"PASS\"\n  } ],\n  \"passingPercentage\" : 88.0,\n  \"evaluationsCount\" : 8\n}",
        "formName": "shortschemaruleset",
        "id": "athp9dyw75gzhj",
        "timestamp": 1.71700477757E9,
        "typeIdentifier": "amazon.datazone.DataQualityResultFormType",
        "typeRevision": "8"
    },
    "formName": "shortschemaruleset"
}
```


Vous pouvez obtenir cette charge utile en invoquant l'`GetFormType`action suivante :

```
aws datazone get-form-type --domain-identifier <your_domain_id> --form-type-
identifier amazon.datazone.DataQualityResultFormType --region <domain_region> --
output text --query 'model.smithy'
```

2. Appelez l'`DeleteTimeSeriesDataPoints`API comme suit :

```
aws datazone delete-time-series-data-points\
--domain-identifier dzd_bqqlk3nz21zp2f \
--entity-identifier dzd_bqqlk3nz21zp2f \
--entity-type ASSET \
--form-name rulesET1 \
```

Utilisation de l'apprentissage automatique et de l'IA générative sur Amazon DataZone

> **Note**
>
> Propulsé par Amazon Bedrock : AWS implémente la détection automatique des abus. Comme les recommandations de l'IA pour les fonctionnalités de description dans Amazon DataZone sont basées sur Amazon Bedrock, les utilisateurs héritent des contrôles mis en œuvre dans Amazon Bedrock pour renforcer la sûreté, la sécurité et l'utilisation responsable de l'IA.

Dans la version actuelle d'Amazon DataZone, vous pouvez utiliser la fonctionnalité de recommandations de l'IA pour les noms et les descriptions afin d'automatiser la découverte et le catalogage des données. Support de l'IA générative dans Amazon : DataZone création de noms commerciaux et de descriptions pour les actifs et les colonnes. Vous pouvez utiliser ces noms et descriptions pour ajouter un contexte commercial à vos données et recommander des analyses pour les ensembles de données, ce qui peut contribuer à améliorer les résultats de découverte des données.

Basées sur les grands modèles linguistiques d'Amazon Bedrock, les recommandations de l'IA relatives aux noms et aux descriptions des actifs de données dans Amazon vous DataZone aident à garantir que vos données sont compréhensibles et faciles à découvrir. Les recommandations de

l'IA suggèrent également les applications analytiques les plus pertinentes pour les ensembles de données. En réduisant les tâches manuelles de documentation et en vous conseillant sur l'utilisation appropriée des données, les noms et descriptions générés automatiquement peuvent vous aider à améliorer la fiabilité de vos données et à minimiser le fait de négliger des données importantes afin d'accélérer la prise de décisions éclairées.

Régions prises en charge

Dans la DataZone version actuelle d'Amazon, la fonctionnalité de recommandations de l'IA pour les noms et les descriptions est prise en charge dans les régions suivantes :

- USA Est (Virginie du Nord)
- USA Ouest (Oregon)
- Asie-Pacifique (Tokyo)
- Europe (Francfort)
- Asie-Pacifique (Sydney)
- Canada (Centre)
- Europe (Londres)
- Amérique du Sud (Sao Paulo)
- Europe (Irlande)
- Asie-Pacifique (Singapour)
- USA Est (Ohio)
- Asie-Pacifique (Séoul)

Amazon DataZone prend en charge la génération de descriptions commerciales dans les régions suivantes.

- Asie-Pacifique (Mumbai)
- Europe (Paris)

Amazon DataZone prend en charge la génération de noms commerciaux dans les régions suivantes.

- Europe (Stockholm)

Inférence interrégionale de Bedrock

Amazon DataZone utilise le point de terminaison d'inférence interrégional d'Amazon Bedrock pour émettre des recommandations pour la région de l'est des États-Unis (Ohio). Toutes les autres régions utilisent un point de terminaison régional.

Étapes d'utilisation de GenAI

La procédure suivante explique comment générer des recommandations basées sur l'IA pour les noms et les descriptions dans Amazon DataZone :

- Accédez à l'URL du portail de DataZone données Amazon, puis connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous avec l' Compte AWS endroit où le domaine a été créé, puis choisissez Open data portal.
- Dans le volet de navigation supérieur, choisissez Sélectionner un projet, puis choisissez le projet contenant la ressource pour laquelle vous souhaitez générer des recommandations d'IA pour les descriptions.

Génération de descriptions et de résumés commerciaux

- Accédez à l'onglet Données du projet.
- Dans le volet de navigation de gauche, choisissez Données d'inventaire, puis choisissez le nom de la ressource pour laquelle vous souhaitez générer des recommandations d'IA pour les descriptions de la ressource.
- Sur la page de détails de la ressource, dans l'onglet Métadonnées commerciales, choisissez Générer des descriptions.

Génération de noms commerciaux

- Accédez à l'onglet Données du projet.
- Dans le volet de navigation de gauche, choisissez Sources de données, puis choisissez la source de données pour laquelle vous souhaitez activer la génération de noms commerciaux.

- Accédez à l'onglet Détails et activez la configuration AUTOMATED BUSINESS NAME GENERATION.
- BusinessNames peut également être généré par programmation lors de la création d'un actif en activant l' businessNameGeneration indicateur sous PredictionConfiguration dans la charge utile de l'API. CreateAsset

Accepter/rejeter les prédictions

- Une fois les descriptions générées, vous pouvez les modifier, les accepter ou les rejeter.
- Des icônes vertes sont affichées à côté de chaque description de métadonnées générée automatiquement pour la ressource de données. Dans l'onglet Métadonnées commerciales, vous pouvez choisir l'icône verte à côté du résumé généré automatiquement, puis choisir Modifier, Accepter ou Rejeter pour répondre à la description générée.
- Vous pouvez également choisir Accepter tout ou Rejeter toutes les options affichées en haut de la page lorsque l'onglet Métadonnées commerciales est sélectionné, et ainsi exécuter l'action sélectionnée sur toutes les descriptions générées automatiquement.
- Vous pouvez également choisir l'onglet Schéma, puis traiter individuellement les descriptions générées automatiquement en choisissant l'icône verte pour une description de colonne à la fois, puis en choisissant Accepter ou Rejeter.
- Dans l'onglet Schéma, vous pouvez également choisir Tout accepter ou Tout rejeter et ainsi exécuter l'action sélectionnée sur toutes les descriptions générées automatiquement.

Pour publier la ressource dans le catalogue avec les descriptions générées, choisissez Publier la ressource, puis confirmez cette action en choisissant à nouveau Publier la ressource dans la fenêtre contextuelle Publier la ressource.

> **Note**
>
> Si vous n'acceptez pas ou ne rejetez pas les descriptions générées pour une ressource, puis que vous publiez cette ressource, les métadonnées générées automatiquement et non révisées ne sont pas incluses dans la ressource de données publiée.

Support pour les types d'actifs relationnels personnalisés

Amazon DataZone prend en charge les fonctionnalités de GenAI pour les types d'actifs personnalisés. Auparavant, cette fonctionnalité n'était prise en charge que pour les types de ressources AWS Glue et AMazon Redshift gérés.

Pour activer cette fonctionnalité, créez votre propre définition de type de ressource et joignez-la `RelationalTableFormType` dans l'un des formulaires. Amazon détecte DataZone automatiquement la présence de tels formulaires et active les fonctionnalités GenAI pour ces actifs. L'expérience globale reste la même pour la génération de noms commerciaux (via PredictionConfiguration dans l' CreateAsset API) et de BusinessDescription (via le bouton Générer une description, cliquez sur la page de détails des actifs).

Pour plus d'informations sur la création de types de ressources personnalisés, voirCréez des types d'actifs personnalisés sur Amazon DataZone.

Quotas

Amazon DataZone prend en charge différents quotas pour la génération de noms commerciaux et de descriptions d'entreprise. Vous pouvez contacter l'équipe d' AWS assistance pour obtenir une augmentation de ces quotas.

- BusinessDescriptionGeneration: 10 000 invocations/mois
- BusinessNameGeneration: 50 000 invocations/mois

Lignage des données sur Amazon DataZone

Le lignage des données dans Amazon DataZone est une fonctionnalité OpenLineage compatible qui peut vous aider à capturer et à visualiser les événements de lignage, depuis ou via des systèmes OpenLineage compatibles, afin de retracer l'origine des données APIs, de suivre les transformations et de visualiser la consommation de données interorganisationnelle. Il vous fournit une vue d'ensemble de vos actifs de données afin de connaître leur origine et leur chaîne de connexions. Les données de lignage incluent des informations sur les activités du catalogue DataZone de données commerciales d'Amazon, notamment des informations sur les actifs catalogués, les abonnés à ces actifs et les activités qui se déroulent en dehors du catalogue de données commerciales capturé par programme à l'aide du. APIs

Rubriques

- Types de nœuds de lignage dans Amazon DataZone
- Attributs clés des nœuds de lignage
- Visualisation du lignage des données
- Autorisation du lignage des données sur Amazon DataZone
- Exemple d'expérience de lignage de données sur Amazon DataZone
- Activer le lignage des données dans la console de gestion
- Utilisation du lignage DataZone de données Amazon par programmation
- Automatisez le lignage pour le catalogue AWS Glue
- Automatisez le lignage depuis Amazon Redshift

Le lignage peut être configuré pour être automatiquement capturé à partir des bases de données AWS Glue et Amazon Redshift lorsqu'il est ajouté à Amazon. DataZone En outre, les tâches Spark ETL s' AWS exécutent dans la console Glue (v5.0 et versions ultérieures) ou les blocs-notes peuvent être configurés pour envoyer des événements de lignage aux domaines Amazon. DataZone

Dans Amazon DataZone, les administrateurs de domaine peuvent configurer le lignage tout en configurant des plans intégrés au lac de données et à l'entrepôt de données, ce qui garantit que toutes les séquences de source de données créées à partir de ces ressources sont activées pour la capture automatique du lignage.

Grâce à la OpenLineage compatibilité DataZone d'Amazon APIs, les administrateurs de domaines et les producteurs de données peuvent capturer et stocker des événements de lignage autres que ceux disponibles sur Amazon DataZone, y compris les transformations dans Amazon S3, AWS Glue et d'autres services. Cela fournit une vue complète aux consommateurs de données et les aide à avoir confiance dans l'origine de l'actif, tandis que les producteurs de données peuvent évaluer l'impact des modifications apportées à un actif en comprenant son utilisation. En outre, DataZone les versions d'Amazon font le suivi de chaque événement, ce qui permet aux utilisateurs de visualiser le lignage à tout moment ou de comparer les transformations liées à l'historique d'un actif ou d'une tâche. Cette généalogie historique permet de mieux comprendre l'évolution des données, ce qui est essentiel pour le dépannage, l'audit et la garantie de l'intégrité des actifs de données.

Grâce au lignage des données, vous pouvez accomplir les tâches suivantes sur Amazon DataZone :

- Comprendre la provenance des données : savoir d'où proviennent les données renforce la confiance dans les données en vous permettant de comprendre clairement leurs origines, leurs

dépendances et leurs transformations. Cette transparence permet de prendre des décisions fondées sur les données en toute confiance.

- Comprenez l'impact des modifications apportées aux pipelines de données : lorsque des modifications sont apportées aux pipelines de données, le lignage peut être utilisé pour identifier tous les consommateurs en aval qui seront concernés. Cela permet de garantir que les modifications sont apportées sans perturber les flux de données critiques.
- Identifiez la cause première des problèmes de qualité des données : si un problème de qualité des données est détecté dans un rapport en aval, le lignage, en particulier le lignage au niveau des colonnes, peut être utilisé pour retracer les données (au niveau des colonnes) afin d'identifier le problème jusqu'à sa source. Cela peut aider les ingénieurs de données à identifier et à résoudre le problème.
- Améliorez la gouvernance et la conformité des données : le lignage au niveau des colonnes peut être utilisé pour démontrer la conformité aux réglementations en matière de gouvernance et de confidentialité des données. Par exemple, le lignage au niveau des colonnes peut être utilisé pour indiquer où les données sensibles (telles que les informations personnelles) sont stockées et comment elles sont traitées dans les activités en aval.

Types de nœuds de lignage dans Amazon DataZone

dans Amazon DataZone, les informations relatives au lignage des données sont présentées dans des nœuds qui représentent des tables et des vues. Selon le contexte du projet, par exemple un projet sélectionné en haut à gauche dans le portail de données, les producteurs peuvent consulter à la fois l'inventaire et les actifs publiés, tandis que les consommateurs ne peuvent consulter que les actifs publiés. Lorsque vous ouvrez l'onglet Lignage pour la première fois sur la page des détails de l'actif, le nœud du jeu de données catalogué est le point de départ pour naviguer en amont ou en aval à travers les nœuds de lignage de votre graphe de lignage.

Les types de nœuds de lignage de données pris en charge par Amazon DataZone sont les suivants :

- Nœud de jeu de données : ce type de nœud inclut des informations de lignée de données relatives à un actif de données spécifique.
  - Les nœuds de jeu de données qui incluent des informations sur les actifs AWS Glue ou Amazon Redshift publiés dans le DataZone catalogue Amazon sont générés automatiquement et incluent une icône AWS Glue ou Amazon Redshift correspondante dans le nœud.
  - Les nœuds de jeu de données qui incluent des informations sur les actifs qui ne sont pas publiés dans le DataZone catalogue Amazon sont créés manuellement par les administrateurs de

domaine (producteurs) et sont représentés par une icône de ressource personnalisée par défaut dans le nœud.

- Nœud Job (run) : ce type de nœud affiche les détails de la tâche, y compris la dernière exécution d'une tâche donnée et les détails de l'exécution. Ce nœud capture également plusieurs exécutions de la tâche et peut être consulté dans l'onglet Historique des détails du nœud. Vous pouvez consulter les détails du nœud en choisissant l'icône du nœud.

Attributs clés des nœuds de lignage

L'`sourceIdentifier`attribut d'un nœud de lignée représente les événements qui se produisent sur un ensemble de données. Le nœud `sourceIdentifier` de lignée est l'identifiant de l'ensemble de données (table/vue, etc.). Il est utilisé pour l'application de l'unicité sur les nœuds de lignée. Par exemple, il ne peut pas y avoir deux nœuds de lignage identiques`sourceIdentifier`. Voici des exemples de `sourceIdentifier` valeurs pour différents types de nœuds :

- Pour le nœud de jeu de données avec le type de jeu de données correspondant :
  - Ressource : amazon.datazone.asset/ <assetId>
  - Annonce (ressource publiée) : amazon.datazone.listing/ <listingId>
  - AWS <region><account-id><database>Table à colle : arn:aws:glue : ::table//<table-name>
  - <redshift/redshift-serverless> <region><account-id><table-type (table/view etc) ><clusterIdentifier/workgroupName> <database><schema>Table/vue Amazon Redshift : arn:aws : : : :///<table-name>
  - Pour tout autre type de nœud de jeu de données importé à l'aide d'événements d'exécution open-lineage,<namespace>/<name>de l'ensemble de données d'entrée/sortie est utilisé en tant que nœud. `sourceIdentifier`
- Pour les emplois :
  - <jobs_namespace>Pour les nœuds de travail importés à l'aide d'événements d'exécution open-lineage,. <job_name>est utilisé comme SourceIdentifier.
- Pour les exécutions de tâches :
  - <jobs_namespace>Pour les nœuds d'exécution de tâches importés à l'aide d'événements d'exécution open-lineage,. <job_name>/<run_id>est utilisé comme SourceIdentifier.

Pour les actifs créés à l'aide de l'`createAsset`API, ils `sourceIdentifier` doivent être mis à jour à l'aide de l'`createAssetRevision`API pour permettre de mapper l'actif aux ressources en amont.

Visualisation du lignage des données

DataZoneLa page détaillée des actifs d'Amazon fournit une représentation graphique du lignage des données, ce qui permet de visualiser plus facilement les relations entre les données en amont ou en aval. La page de détails des actifs fournit les fonctionnalités suivantes pour naviguer dans le graphique :

- Lignage au niveau des colonnes : étendez le lignage au niveau des colonnes lorsqu'il est disponible dans les nœuds du jeu de données. Cela affiche automatiquement les relations avec les nœuds de jeu de données en amont ou en aval si les informations de la colonne source sont disponibles.
- Recherche par colonne : lorsque l'affichage par défaut du nombre de colonnes est 10. S'il y a plus de 10 colonnes, la pagination est activée pour accéder au reste des colonnes. Pour afficher rapidement une colonne en particulier, vous pouvez effectuer une recherche sur le nœud du jeu de données qui répertorie uniquement la colonne recherchée.
- Afficher uniquement les nœuds du jeu de données : si vous souhaitez afficher uniquement les nœuds de lignage des ensembles de données et filtrer les nœuds de travail, vous pouvez choisir l'icône de contrôle Ouvrir la vue en haut à gauche de la visionneuse de graphes et activer l'option Afficher uniquement les nœuds du jeu de données. Cela supprimera tous les nœuds de tâche du graphe et vous permettra de naviguer uniquement dans les nœuds du jeu de données. Notez que lorsque l'option Afficher uniquement les nœuds du jeu de données est activée, le graphe ne peut pas être développé en amont ou en aval.
- Volet de détails : les détails de chaque nœud de lignée sont capturés et affichés lorsqu'ils sont sélectionnés.
  - Le nœud du jeu de données possède un volet de détails qui affiche tous les détails capturés pour ce nœud pour un horodatage donné. Chaque nœud de jeu de données comporte 3 onglets, à savoir : informations sur le lignage, schéma et historique. L'onglet historique répertorie les différentes versions de l'événement de lignage capturé pour ce nœud. Tous les détails capturés à partir de l'API sont affichés à l'aide de formulaires de métadonnées ou d'un visualiseur JSON.
  - Le nœud Job dispose d'un volet détaillé pour afficher les détails de la tâche sous forme d'onglets, à savoir : Informations sur le travail et Historique. Le volet de détails capture également les requêtes ou expressions capturées dans le cadre de l'exécution du job. L'onglet historique répertorie les différentes versions de l'événement d'exécution d'une tâche capturée pour cette tâche. Tous les détails capturés à partir de l'API sont affichés à l'aide de formulaires de métadonnées ou d'un visualiseur JSON.

- Onglets de version : tous les nœuds de lignage d'Amazon DataZone Data Lineage sont dotés d'une fonction de version. Pour chaque nœud de jeu de données ou nœud de travail, les versions sont capturées sous forme d'historique, ce qui vous permet de naviguer entre les différentes versions pour identifier ce qui a changé au fil du temps. Chaque version ouvre un nouvel onglet sur la page de lignage pour faciliter la comparaison ou le contraste.

Autorisation du lignage des données sur Amazon DataZone

Autorisations d'écriture : pour publier des données de lignage sur Amazon DataZone, vous devez disposer d'un rôle IAM avec une politique d'autorisation incluant une `ALLOW` action sur l'`PostLineageEvent`API. Cette autorisation IAM se produit au niveau de la couche API Gateway.

Autorisations de lecture : `GetLineageNode` deux opérations sont incluses dans la politique `AmazonDataZoneDomainExecutionRolePolicy` gérée. Chaque utilisateur du DataZone domaine Amazon peut donc les invoquer pour parcourir le graphe de lignage des données. `ListLineageNodeHistory`

Exemple d'expérience de lignage de données sur Amazon DataZone

Vous pouvez utiliser l'expérience d'échantillonnage du lignage de données pour parcourir et comprendre le lignage des données sur Amazon DataZone, notamment en parcourant le tracé en amont ou en aval dans votre graphe de lignage de données, en explorant les versions et le lignage au niveau des colonnes.

Suivez la procédure suivante pour essayer l'exemple d'expérience de lignage de données sur Amazon : DataZone

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez n'importe quelle ressource de données disponible pour ouvrir la page de détails de la ressource.
3. Sur la page de détails de la ressource, cliquez sur l'onglet Lineage, puis passez la souris sur l'icône d'information, puis choisissez Try sample lineage.

4. Dans la fenêtre contextuelle de lignage des données, choisissez Démarrer la visite guidée du lignage des données.

   À ce stade, un onglet en plein écran qui fournit tout l'espace des informations sur le lignage s'affiche. L'exemple de graphe de lignage des données est initialement affiché avec un nœud de base d'une profondeur à chaque extrémité, en amont et en aval. Vous pouvez développer le graphique en amont ou en aval. Les informations des colonnes sont également disponibles pour que vous puissiez choisir et voir comment le lignage circule dans les nœuds.

Activer le lignage des données dans la console de gestion

Vous pouvez activer le lignage des données dans le cadre de la configuration de vos plans de lac de données par défaut et d'entrepôt de données par défaut.

Suivez la procédure suivante pour activer le lignage des données pour votre plan Data Lake par défaut.

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.
2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez activer le lignage des données pour votre DefaultDataLake plan.
3. Sur la page des détails du domaine, accédez à l'onglet Blueprints.
4. Sur la page de détails du DefaultDataLake plan, choisissez l'onglet Régions.
5. Vous pouvez activer le lignage des données dans le cadre de l'ajout d'une région à votre DefaultDataLake plan. Ainsi, si une région est déjà ajoutée mais que la fonctionnalité de lignage des données qu'elle contient n'est pas activée (la mention Non est affichée dans la colonne Importer le lignage des données), vous devez d'abord supprimer cette région. Pour activer le lignage des données, choisissez Ajouter une région, puis choisissez la région que vous souhaitez ajouter, et assurez-vous de cocher la case Activer l'importation du lignage des données dans la fenêtre contextuelle Ajouter une région.

Pour activer le lignage des données pour votre DefaultDataWarehouse plan, procédez comme suit.

1. Accédez à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et connectez-vous à l'aide des informations d'identification de votre compte.

2. Choisissez Afficher les domaines et choisissez le domaine dans lequel vous souhaitez activer le lignage des données pour votre DefaultDataWarehouse plan.
3. Sur la page des détails du domaine, accédez à l'onglet Blueprints.
4. Sur la page de détails du DefaultDataWarehouse plan, choisissez l'onglet Ensembles de paramètres.
5. Vous pouvez activer le lignage des données dans le cadre de l'ajout d'un ensemble de paramètres pour votre DefaultDataWarehouse plan. Pour ce faire, choisissez Créer un jeu de paramètres.
6. Sur la page Créer un jeu de paramètres, spécifiez ce qui suit, puis choisissez Créer un jeu de paramètres.

   - Nom du jeu de paramètres.
   - Description du jeu de paramètres.
   - AWS Région dans laquelle vous souhaitez créer des environnements.
   - Spécifiez si Amazon DataZone doit utiliser ces paramètres pour établir une connexion à votre cluster Amazon Redshift ou à votre groupe de travail sans serveur.
   - Spécifiez un AWS secret.
   - Spécifiez un cluster ou un groupe de travail sans serveur que vous souhaitez utiliser lors de la création d'environnements.
   - Spécifiez le nom de la base de données (au sein du cluster ou du groupe de travail que vous avez spécifié) que vous souhaitez utiliser lors de la création d'environnements.
   - Sous Importer le lignage des données, cochez la case Activer l'importation du lignage des données.

Utilisation du lignage DataZone de données Amazon par programmation

Pour utiliser la fonctionnalité de traçabilité des données dans Amazon DataZone, vous pouvez invoquer ce qui suit : APIs

- GetLineageNode
- ListLineageNodeHistory
- PostLineageEvent

Automatisez le lignage pour le catalogue AWS Glue

Au fur et à mesure que les bases de données et les tables AWS Glue sont ajoutées au DataZone catalogue Amazon, l'extraction du lignage est automatisée pour ces tables à l'aide d'analyses de sources de données. Le lignage est automatisé de plusieurs manières pour cette source :

- Configuration du plan : les administrateurs qui configurent des plans peuvent configurer des plans pour capturer automatiquement le lignage. Cela permet aux administrateurs de définir les sources de données importantes pour la capture du lignage plutôt que de se fier au catalogage des données par les producteurs de données. Pour de plus amples informations, veuillez consulter Activer le lignage des données dans la console de gestion.
- Configuration de la source de données : les producteurs de données, lorsqu'ils configurent les exécutions de sources de données pour les bases de données AWS Glue, reçoivent une vue ainsi que Data Quality pour les informer sur le lignage automatique des données pour cette source de données.
  - Le paramètre de lignage peut être consulté dans l'onglet Définition de la source de données. Cette valeur n'est pas modifiable par les producteurs de données.
  - La collection de lignage dans Data Source run extrait des informations à partir des métadonnées des tables pour créer le lignage. AWS Le crawler Glue prend en charge différents types de sources et les sources pour lesquelles le lignage est capturé dans le cadre de l'exécution de la source de données incluent Amazon S3, DynamoDB, Catalog, Delta Lake, les tables Iceberg et les tables Hudi stockées dans Amazon S3. JDBC et DocumentDB ou MongoDB ne sont actuellement pas pris en charge en tant que sources.
  - Limitation : si le nombre de tables est supérieur à 100, l'exécution du lignage échoue après 100 tables. Assurez-vous que le robot AWS Glue n'est pas configuré pour intégrer plus de 100 tables par série.
- AWS Configuration de Glue (v5.0) : lors de l'exécution des tâches AWS Glue dans AWS Glue Studio, le lignage des données peut être configuré pour que les tâches envoient des événements de lignage directement au domaine Amazon. DataZone

  1. Accédez à la console AWS Glue à l'adresse https://console.aws.amazon.com /gluestudio et connectez-vous à l'aide des informations d'identification de votre compte.
  2. Choisissez les tâches ETL et créez une nouvelle tâche ou cliquez sur l'une des tâches existantes.
  3. Accédez à l'onglet Détails de la tâche (y compris la tâche ETL Flows) et faites défiler la page vers le bas jusqu'à la section Générer des événements de lignage.

4. Cochez la case pour activer l'envoi d'événements de lignage. Elle s'agrandit pour afficher un champ de saisie permettant de saisir l'identifiant de DataZone domaine Amazon.

- AWS Configuration du bloc-notes Glue (V5.0) : dans un bloc-notes, vous pouvez automatiser la collecte des exécutions de Spark en ajoutant la magie de %%configure. Cette configuration enverra des événements au DataZone domaine Amazon.

```
%%configure --name project.spark -f
{
    "--
conf":"spark.extraListeners=io.openlineage.spark.agent.OpenLineageSparkListener
 --conf spark.openlineage.transport.type=amazon_datazone_api --
conf spark.openlineage.transport.domainId={DOMAIN_ID}  --conf
 spark.glue.accountId={ACCOUNT_ID} --conf
 spark.openlineage.facets.custom_environment_variables=[AWS_DEFAULT_REGION;GLUE_VERSION;GLUE_
 --conf spark.glue.JOB_NAME={JOB_NAME}"
}
```

Les détails des paramètres sont les suivants :

- `spark.extraListeners=io.openlineage.spark.agent.OpenLineageSparkListener`- OpenLineageSparkListener sera créé et enregistré avec le bus d'écoute de Spark
- `spark.openlineage.transport.type=amazon_datazone_api`- Il s'agit d'une OpenLineage spécification indiquant au OpenLineage plugin d'utiliser le transport DataZone API pour émettre des événements de lignage vers DataZone l' PostLineageEvent API. Pour plus d'informations, consultez https://openlineage. io/docs/integrations/spark/configuration/spark_conf
- `spark.openlineage.transport.domainId={DOMAIN_ID}`- Ce paramètre définit le domaine auquel le transport API soumettra les événements de lignage.
- `spark.openlineage.facets.custom_environment_variables [AWS_DEFAULT_REGION;GLUE_VERSION;GLUE_COMMAND_CRITERIA;GLUE_PYTHON_VERSION;]`- Les variables d'environnement suivantes (GLUE_VERSIONAWS_DEFAULT_REGION , GLUE_COMMAND_CRITERIA et GLUE_PYTHON_VERSION), renseignées par la session interactive Glue, seront ajoutées au LineageEvent
- `spark.glue.accountId=<ACCOUNT_ID>`- ID de compte du catalogue de données Glue où se trouvent les métadonnées. Cet identifiant de compte est utilisé pour créer Glue ARN lors d'un événement de lignée.

  - `spark.glue.JOB_NAME`- Nom du job de l'événement de lignage. Le nom de la tâche dans le bloc-notes peut être défini comme`spark.glue.JOB_NAME: ${projectId}.${pathToNotebook}`.
- Configurez les paramètres pour configurer la communication avec Amazon DataZone depuis AWS Glue

  Clé de paramètre : --conf

  Valeur du paramètre :

```
spark.extraListeners=io.openlineage.spark.agent.OpenLineageSparkListener
--conf spark.openlineage.transport.type=amazon_datazone_api
--conf spark.openlineage.transport.domainId=<DOMAIN_ID>
--conf
 spark.openlineage.facets.custom_environment_variables=[AWS_DEFAULT_REGION;GLUE_VERSION;GLUE_
--conf spark.glue.accountId=<ACCOUNT_ID> (replace <DOMAIN_ID> and <ACCOUNT_ID> with
 the right values)
```

  Pour les ordinateurs portables, ajoutez les paramètres supplémentaires suivants :

```
--conf spark.glue.JobName=<SessionId> --conf spark.glue.JobRunId=<SessionId or NONE?>
replace <SessionId> and <SessionId> with the right values
```

Automatisez le lignage depuis Amazon Redshift

Capture du lignage à partir du service Amazon Redshift Avec la configuration du plan d'entrepôt de données configuré par les administrateurs, le lignage est automatiquement capturé par Amazon. DataZone Le lignage exécute, capture les requêtes exécutées pour une base de données donnée et génère des événements de lignage à stocker sur Amazon DataZone afin d'être visualisés par les producteurs de données ou les consommateurs lorsqu'ils accèdent à un actif en particulier.

Le lignage peut être automatisé à l'aide des configurations suivantes :

- Configuration du plan : les administrateurs qui configurent des plans peuvent configurer des plans pour capturer automatiquement le lignage. Cela permet aux administrateurs de définir les

sources de données importantes pour la capture du lignage plutôt que de se fier au catalogage des données par les producteurs de données. Pour configurer, rendez-vous surActiver le lignage des données dans la console de gestion.

- Configuration de la source de données : lorsqu'ils configurent les exécutions de sources de données pour les bases de données Amazon Redshift, les producteurs de données se voient proposer un paramètre de lignage automatique pour cette source de données.

  Le paramètre de lignage peut être consulté dans l'onglet Définition de la source de données. Cette valeur n'est pas modifiable par les producteurs de données.

Règles d'application des métadonnées pour la publication

Les règles d'application des métadonnées pour la publication sur Amazon DataZone renforcent la gouvernance des données en permettant aux propriétaires d'unités de domaine d'établir des exigences claires en matière de métadonnées pour les producteurs de données, en rationalisant les demandes d'accès et en améliorant la gouvernance des données.

Cette fonctionnalité est prise en charge dans toutes les régions AWS commerciales dans lesquelles Amazon DataZone est actuellement disponible.

Les propriétaires d'unités de domaine peuvent suivre la procédure suivante pour configurer l'application des métadonnées dans Amazon DataZone :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Choisissez Domaines, accédez à l'onglet Unités de domaine et choisissez l'unité de domaine avec laquelle vous souhaitez travailler.
3. Cliquez sur l'onglet Règles, puis sur Ajouter.
4. Sur la page Créer une règle de formulaire de métadonnées requises, procédez comme suit, puis choisissez Ajouter une règle :

   - Donnez un nom à votre règle.
   - Sous Action, sélectionnez Publication de ressources de données et de produits.

- Sous Formulaires obligatoires, choisissez Ajouter un formulaire de métadonnées, choisissez un formulaire de métadonnées dans le domaine/l'unité de domaine que vous souhaitez ajouter à cette règle, puis choisissez Ajouter. Vous pouvez ajouter jusqu'à 5 formulaires de métadonnées par règle.
- Sous Étendue, spécifiez les entités de données auxquelles vous souhaitez associer ces formulaires. Vous pouvez choisir des produits de données et and/or des actifs de données.
- Sous Types de ressources de données, spécifiez si la règle s'applique à tous les types de ressources ou limitez-la aux types de ressources sélectionnés.
- Sous Projets, spécifiez si les formulaires requis seront associés aux and/or actifs de produits de données publiés par tous les projets ou uniquement aux projets sélectionnés dans cette unité de domaine. Vérifiez également la règle de cascade pour les unités de domaine enfant si vous souhaitez que les unités de domaine enfant héritent de cette exigence.

Produits DataZone de données Amazon

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. L'utilisation de produits de données cohérents et adaptés à l'entreprise améliore à la fois les processus de publication et d'abonnement. Les consommateurs de données peuvent facilement identifier les actifs de données interconnectés en les recherchant et en les trouvant sous forme d'unité unique. Cette approche réduit le temps et les efforts nécessaires pour trouver toutes les informations pertinentes et réduit le risque de manquer des données importantes. En outre, les produits de données simplifient l'accès aux données avec une seule demande en mettant en œuvre un modèle d'accès unifié. Cela élimine le besoin d'autorisations multiples, accélérant ainsi le lancement de l'analyse des données. En outre, en cataloguant les actifs sous forme de produits de données, les producteurs de données réduisent les frais administratifs en permettant la gestion des métadonnées et du contrôle d'accès au niveau du produit de données, plutôt qu'individuellement. De plus, la capacité à identifier ces actifs groupés spécialement conçus pour les utiliser rend la gouvernance des accès et l'utilisation des données plus efficaces, garantissant ainsi leur adéquation avec les objectifs commerciaux et leur accessibilité facile pour l'usage auquel ils sont destinés. Les équipes de gouvernance des données peuvent surveiller les taux de consommation de ces produits de données, fournissant ainsi des informations précieuses sur le degré de maturité de la data literacy. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Rubriques

- Créez de nouveaux produits de données sur Amazon DataZone
- Publiez des produits de données sur Amazon DataZone
- Modifier des produits de données sur Amazon DataZone
- Dépublier les produits de données sur Amazon DataZone
- Supprimer des produits de données sur Amazon DataZone
- S'abonner à un produit de données sur Amazon DataZone
- Examiner une demande d'abonnement et octroyer un abonnement à un produit de données sur Amazon DataZone
- Republiez les produits de données sur Amazon DataZone

Créez de nouveaux produits de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour créer des produits de données, vous devez être le propriétaire ou le contributeur du projet.

Pour créer un nouveau produit de données, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de DataZone données Amazon, choisissez le projet dans lequel vous souhaitez créer un produit de données.
3. Cliquez sur l'onglet Données, puis sur Données d'inventaire, puis sur Créer un nouveau produit de données.
4. Sur la page Créer un nouveau produit de données, spécifiez le nom et la description du produit de données, puis choisissez Sélectionner des actifs pour ajouter divers actifs à votre produit de données. Dans la fenêtre contextuelle Sélectionner les actifs, choisissez les actifs que vous souhaitez ajouter à ce produit de données, puis sélectionnez Sélectionner. Pour terminer la création du produit de données, choisissez Create.

Publiez des produits de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour publier des produits de données, vous devez être le propriétaire ou le contributeur du projet. Les règles d'application des métadonnées pour la publication peuvent être configurées pour établir des exigences claires en matière de métadonnées pour les producteurs de données, en limitant le moment où un produit de données peut être publié.

Pour publier un produit de données, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de DataZone données Amazon, choisissez le projet dans lequel réside le produit de données que vous souhaitez publier.
3. Cliquez sur l'onglet Données, puis sur Données d'inventaire, puis sur le filtre Produits de données. Cela affiche tous les produits de données existants non publiés.
4. Choisissez le produit de données que vous souhaitez publier, puis cliquez sur Publier. Confirmez la publication de ce produit de données en choisissant Publier le produit de données.

> **Note**
>
> Toutes les ressources de données non publiées figurant dans ce produit de données seront publiées, mais ne seront disponibles que par le biais de ce produit de données.

Modifier des produits de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour modifier un produit de données, vous devez être le propriétaire ou le contributeur du projet auquel le produit de données appartient.

Pour modifier un produit de données, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.

2. Dans le portail de DataZone données Amazon, choisissez le projet dans lequel réside le produit de données que vous souhaitez publier.
3. Choisissez l'onglet Données, puis choisissez Données d'inventaire ou Données publiées, puis choisissez le filtre Produits de données.
4. Choisissez le produit de données que vous souhaitez modifier. Dans le cadre de la modification d'un produit de données, vous pouvez effectuer les opérations suivantes :

   - Choisissez Créer un fichier readme pour ajouter un fichier readme afin d'aider les utilisateurs à mieux comprendre cette page.
   - Choisissez Ajouter des termes pour ajouter des termes au glossaire. Sélectionnez les termes du glossaire dans la fenêtre, puis choisissez Ajouter des termes.
   - Choisissez Ajouter un formulaire de métadonnées, puis sélectionnez votre formulaire dans la fenêtre Ajouter un formulaire de métadonnées et choisissez Ajouter.
   - Développez les actions, choisissez Modifier, modifiez le nom et la description du produit de données, puis choisissez Mettre à jour.

Dépublier les produits de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour dépublier un produit de données, vous devez être le propriétaire ou le contributeur du projet auquel appartient le produit de données.

Pour dépublier un produit de données, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de DataZone données Amazon, choisissez le projet dans lequel se trouve le produit de données que vous souhaitez dépublier.

3. Choisissez l'onglet Données, puis choisissez Données d'inventaire ou Données publiées, puis choisissez le filtre Produits de données. Cela affiche tous les produits de données existants.
4. Choisissez le produit de données que vous souhaitez dépublier, puis développez Actions et choisissez Dépublier. Confirmez l'annulation de la publication de ce produit de données en choisissant Dépublier.

> **Note**
>
> L'annulation de la publication d'un produit de données a les effets suivants :
>
> - Ce produit de données ne sera plus disponible à la consultation ou à l'abonnement.
> - Toutes les ressources de données uniquement disponibles via ce produit de données ne seront plus disponibles.
> - Tous les abonnements actifs à ce produit de données seront maintenus.
> - Les actifs de données publiés individuellement ne seront pas affectés.

Supprimer des produits de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour supprimer un produit de données, vous devez être le propriétaire ou le contributeur du projet auquel appartient le produit de données.

Pour supprimer un produit de données, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de DataZone données Amazon, choisissez le projet dans lequel réside le produit de données que vous souhaitez supprimer.

3. Choisissez l'onglet Données, puis choisissez Données d'inventaire ou Données publiées, puis choisissez le filtre Produits de données. Cela affiche tous les produits de données existants.
4. Choisissez le produit de données que vous souhaitez supprimer, puis développez Actions et choisissez Supprimer. Confirmez la suppression de ce produit de données `delete` en saisissant du texte dans le champ de texte, puis en choisissant Supprimer.

> **Note**
>
> La suppression d'un produit de données a les effets suivants :
>
> - Le produit de données ne sera plus disponible pour publication, consultation ou abonnement.
> - Les actifs de données uniquement disponibles via ce produit de données ne seront plus visibles dans le catalogue de données. Ils ne seront pas supprimés de vos actifs d'inventaire.

S'abonner à un produit de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Tout DataZone utilisateur Amazon disposant des autorisations requises pour accéder au portail de données peut s'abonner à un produit de DataZone données Amazon.

Pour vous abonner ou vous désabonner d'un produit de données, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Parcourir le catalogue pour trouver le produit de données auquel vous souhaitez vous abonner, puis choisissez ce produit de données.
3. Sur la page de détails du produit de données, choisissez S'abonner.

4. Spécifiez le projet et le motif de votre abonnement, puis choisissez S'abonner.

Examiner une demande d'abonnement et octroyer un abonnement à un produit de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Le projet propriétaire du produit de données peut examiner et accorder l'abonnement à un produit de DataZone données Amazon.

Pour examiner une demande d'abonnement et accorder un abonnement à un produit de données, procédez comme suit :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez le projet propriétaire du produit de données pour lequel vous souhaitez examiner une demande d'abonnement entrante.
3. Choisissez l'onglet Données, puis sélectionnez Demandes entrantes.
4. Choisissez la demande que vous souhaitez examiner, puis dans la fenêtre de demande d'abonnement, choisissez Approuver ou Rejeter, puis saisissez un commentaire de désignation.

Republiez les produits de données sur Amazon DataZone

Amazon DataZone permet aux producteurs de données de regrouper les actifs de données dans des packages autonomes bien définis, appelés produits de données, adaptés à des cas d'utilisation professionnels spécifiques. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Pour republier un produit de données, vous devez être le propriétaire ou le contributeur du projet. Les règles d'application des métadonnées pour la publication peuvent être configurées pour établir

des exigences claires en matière de métadonnées pour les producteurs de données, en limitant les circonstances dans lesquelles un produit de données peut être republié.

Pour republier un produit de données, procédez comme suit.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de DataZone données Amazon, choisissez le projet dans lequel réside le produit de données que vous souhaitez republier.
3. Cliquez sur l'onglet Données, puis sur Données publiées, puis sur le filtre Produits de données.
4. Choisissez le produit de données que vous souhaitez republier, puis cliquez sur l'onglet Ressources.
5. Dans l'onglet Ressources, effectuez l'une des opérations suivantes :

   - supprimez l'une des ressources existantes dans le produit de données en choisissant cette ressource, puis en développant l'icône d'action et en choisissant Supprimer la ressource. Confirmez la suppression de la ressource en choisissant Supprimer dans la fenêtre contextuelle Supprimer la ressource. Une fois que vous aurez republié, cette ressource sera supprimée de tous les abonnés à ce produit de données.
   - Ajoutez une nouvelle ressource au produit de données en cliquant sur le bouton Ajouter, puis en sélectionnant une ou plusieurs ressources à ajouter au produit de données.
6. Sur la page de détails du produit de données, sélectionnez Republier. Confirmez cette action en choisissant Republier dans la fenêtre contextuelle Republier le produit de données.

> **Note**
>
> La republication de ce produit de données mettra à jour les éléments suivants pour tous les abonnés :
>
> - Si des actifs ont été supprimés du produit de données, les abonnés n'auront plus accès à ces actifs.
> - Si des actifs ont été ajoutés au produit de données, les abonnés auront accès à ces actifs.

- Les nouvelles versions publiées des actifs de données seront disponibles.

Découverte, abonnement et consommation de DataZone données Amazon

Sur Amazon DataZone, une fois qu'un actif est publié sur un domaine, les abonnés peuvent le découvrir et demander un abonnement à cet actif. Le processus d'abonnement commence lorsque l'abonné recherche et parcourt le catalogue pour trouver l'actif qu'il souhaite. Depuis le DataZone portail Amazon, ils choisissent de s'abonner à l'actif en soumettant une demande d'abonnement incluant une justification et le motif de la demande. Le propriétaire de l'actif examine la demande. Ils peuvent approuver ou rejeter la demande.

Une fois l'abonnement accordé, un processus de traitement des commandes démarre afin de faciliter l'accès à l'actif pour l'abonné. Il existe deux principaux modes de contrôle d'accès et de traitement des actifs : ceux pour les actifs DataZone gérés par Amazon et ceux pour les actifs qui ne sont pas gérés par Amazon DataZone.

- Ressources gérées : Amazon DataZone peut gérer le traitement des commandes et les autorisations pour les actifs gérés, tels que AWS Glue les tables et les tables et vues Amazon Redshift.
- Actifs non gérés : Amazon DataZone publie des événements standard liés à vos actions (par exemple, l'approbation donnée à une demande d'abonnement) sur Amazon EventBridge. Vous pouvez utiliser ces événements standard pour intégrer d'autres AWS services ou des solutions tierces pour des intégrations personnalisées.

Rubriques

- Rechercher et consulter des actifs dans le DataZone catalogue Amazon
- Demandez un abonnement à des actifs sur Amazon DataZone
- Approuver ou rejeter une demande d'abonnement sur Amazon DataZone
- Révoquer un abonnement existant sur Amazon DataZone
- Annuler une demande d'abonnement sur Amazon DataZone
- Se désabonner d'un actif sur Amazon DataZone
- Utilisation des rôles IAM existants pour traiter les abonnements Amazon DataZone
- Accorder l'accès aux AWS Glue Data Catalog actifs gérés sur Amazon DataZone
- Accorder l'accès aux ressources Amazon Redshift gérées sur Amazon DataZone

- Accorder l'accès aux actifs non gérés sur Amazon pour les abonnements approuvés DataZone
- Interrogez des données dans Amazon Athena ou Amazon Redshift dans Amazon DataZone
- Règles d'application des métadonnées pour les demandes d'abonnement
- Analysez les données des DataZone abonnés Amazon avec des applications d'analyse externes via une connexion JDBC

Rechercher et consulter des actifs dans le DataZone catalogue Amazon

Amazon DataZone propose un moyen simplifié de rechercher des données. Tout DataZone utilisateur Amazon autorisé à accéder au portail de données peut rechercher des actifs dans le DataZone catalogue Amazon et consulter les noms des actifs ainsi que les métadonnées qui leur sont attribuées. Vous pouvez examiner un actif de plus près en consultant sa page de détails.

> **Note**
>
> Pour consulter les données réelles contenues dans une ressource, vous devez d'abord vous abonner à la ressource, faire approuver votre demande d'abonnement et obtenir l'accès accordé.

La recherche sur Amazon DataZone (dans les domaines nouveaux et existants) inclut des résultats basés sur des mots clés et des correspondances sémantiques. L'algorithme de recherche donne la priorité aux correspondances de mots clés, puis ajoute celles présentant des correspondances sémantiques.

La fonctionnalité de recherche sémantique permet aux utilisateurs, quels que soient leurs rôles et fonctions, de découvrir, d'accéder et d'exploiter plus efficacement les actifs de données de leur organisation, ce qui améliore la prise de décision, la collaboration et les capacités globales axées sur les données. Avec la recherche sémantique, les saisies de mots clés produisent des résultats de recherche basés sur des synonymes et basés sur le sens, en plus de simples résultats de correspondance de mots clés. Par exemple, avec la recherche sémantique, si vous saisissez « fleur » comme entrée de recherche, une ressource de données dont le nom contient le mot « rose » est renvoyée dans les résultats de recherche. Si vous saisissez « film » comme entrée de recherche, une ressource de données dont le nom contient le mot « film » s'affiche dans les résultats de recherche.

Si vous saisissez « football » comme entrée de recherche, une ressource de données dont le nom contient le mot « football » peut être renvoyée dans les résultats de recherche.

Avec la recherche par mot clé, vous pouvez saisir différents mots clés lors de la recherche des actifs auxquels vous êtes abonné. Par exemple, si une ressource est appeléeCatalog Sales Data, elle est renvoyée dans les résultats de recherche si vous saisissez l'un des mots clés suivants : catalog_salesCatalog Sales,CatalogSales, oucatalogsales.

Amazon améliore DataZone également l'expérience de recherche en activant des fonctionnalités de correspondance exacte et partielle précises pour les identifiants techniques tels que les noms de colonnes et de tables. Grâce à cette nouvelle fonctionnalité, vous pouvez effectuer des recherches en mettant vos mots clés entre guillemets (» «), afin de garantir que les résultats correspondent exactement ou partiellement aux noms techniques. Cette fonctionnalité s'appuie sur les fonctionnalités de recherche sémantique et par mots clés, qui vous permettent de découvrir des actifs par concepts et termes connexes. En ajoutant une couche de précision aux identifiants techniques, cette amélioration vous permet de gérer de grands catalogues de données avec des conventions de dénomination techniques complexes.

Lorsque vous parcourez vos données, vous devrez peut-être trouver des ressources techniques spécifiques pour soutenir vos cas d'utilisation. Grâce à la possibilité de rechercher des identifiants techniques, vous pouvez récupérer des actifs avec précision, gagner du temps et rationaliser le processus de découverte. Par exemple, une requête telle que « customer_id » renvoie des colonnes ou des tables avec l'identifiant exact, tandis qu'une requête partielle telle que « sales_ » peut identifier des actifs connexes tels que sales_summary et sales_data_2024. Cette amélioration permet aux consommateurs de données de trouver efficacement les actifs dont ils ont besoin, améliorant ainsi la productivité.

Pour rechercher des actifs dans le catalogue

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Vous pouvez saisir le nom de la ressource que vous recherchez dans la barre de recherche de la page d'accueil du portail de données.
3. Pour parcourir les espaces de noms, choisissez Catalogue en haut à droite de la page pour ouvrir le catalogue. Le catalogue propose une expérience de recherche à facettes qui

vous permet de trouver des actifs en effectuant une recherche selon des critères tels que le propriétaire des données et les termes du glossaire.

4. Entrez votre terme de recherche dans l'un des champs de recherche. Après avoir effectué une recherche, vous pouvez appliquer différents filtres pour affiner les résultats. Les filtres incluent le type d'actif, le compte source et le type d'actif Région AWS auquel appartient l'actif.
5. Pour afficher les détails d'une ressource spécifique, sélectionnez la ressource pour ouvrir sa page de détails. La page de détails contient les informations suivantes :

   - Le nom de la ressource, la source de données (AWS Glue Amazon Redshift ou Amazon S3), le type (table, vue ou objet S3), le nombre de colonnes et la taille.
   - Description de l'actif.
   - La version publiée en cours de la ressource, le propriétaire, si une approbation est requise pour les abonnements, le namepace et l'historique des mises à jour.
   - Un onglet Aperçu qui inclut les termes du glossaire et les formulaires de métadonnées.
   - Un onglet Schéma qui affiche le schéma de la ressource, y compris les noms des colonnes commerciales et techniques, les types de données et les descriptions commerciales des colonnes. L'onglet Schéma n'est visible que pour les tables et les vues (pas pour les objets Amazon S3).
   - Un onglet Abonnements qui inclut une liste des abonnés au domaine.
   - Un onglet Historique qui inclut une liste des révisions passées de la ressource.

Demandez un abonnement à des actifs sur Amazon DataZone

Amazon vous DataZone permet de trouver, d'accéder et de consommer les actifs du DataZone catalogue Amazon. Lorsque vous trouvez une ressource dans le catalogue à laquelle vous souhaitez accéder, vous devez vous y abonner, ce qui crée une demande d'abonnement. Un approbateur peut ensuite approuver ou demander votre demande.

Vous devez être membre d'un projet pour demander un abonnement à un actif de ce projet.

Pour souscrire à un actif

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à

l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.

2. Utilisez la barre de recherche pour rechercher et choisir la ressource à laquelle vous souhaitez vous abonner, puis choisissez S'abonner.
3. Dans la fenêtre contextuelle S'abonner, saisissez les informations suivantes :
   - Le projet auquel vous souhaitez souscrire à la ressource.
   - Brève justification de votre demande d'abonnement.
4. Choisissez Abonner.

   Vous recevez une notification sur le portail de données lorsque l'éditeur approuve votre demande.

Pour consulter le statut de la demande d'abonnement, recherchez et choisissez le projet avec lequel vous vous êtes abonné à la ressource. Accédez à l'onglet Données du projet, puis choisissez Données demandées dans le volet de navigation de gauche. Cette page répertorie les actifs auxquels le projet a demandé l'accès. Vous pouvez filtrer la liste en fonction du statut de la demande.

Approuver ou rejeter une demande d'abonnement sur Amazon DataZone

Amazon vous DataZone permet de trouver, d'accéder et de consommer les actifs du DataZone catalogue Amazon. Lorsque vous trouvez une ressource dans le catalogue à laquelle vous souhaitez accéder, vous devez vous y abonner, ce qui crée une demande d'abonnement. Un approbateur peut ensuite approuver ou rejeter votre demande.

Vous devez être membre du projet propriétaire (le projet qui a publié la ressource) pour approuver ou rejeter une demande d'abonnement.

Pour approuver ou rejeter une demande d'abonnement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.

2. Dans le portail de données, choisissez Parcourir la liste des projets et sélectionnez le projet contenant la ressource avec la demande d'abonnement.
3. Accédez à l'onglet Données, puis sélectionnez Demandes entrantes dans le volet de navigation de gauche.
4. Localisez la demande et choisissez Afficher la demande. Vous pouvez filtrer par En attente pour ne voir que les demandes encore ouvertes.
5. Passez en revue la demande d'abonnement et le motif de l'accès, puis décidez de l'approuver ou de la rejeter.
6. Pour approuver, sélectionnez l'une des deux options suivantes :

   - Accès complet : si vous choisissez d'approuver l'abonnement avec l'option d'accès complet, l'abonné aura accès à toutes les lignes et colonnes de votre ressource de données.
   - Approuver avec des filtres de ligne et de colonne : pour limiter l'accès à des lignes et colonnes de données spécifiques, vous pouvez choisir l'option d'approbation avec des filtres de ligne et de colonne. Pour de plus amples informations, veuillez consulter Contrôle d'accès précis aux données sur Amazon DataZone.
     - Sélectionnez Choisir des filtres, puis dans le menu déroulant, sélectionnez un ou plusieurs filtres disponibles que vous souhaitez appliquer à l'abonnement.
     - Pour créer un nouveau filtre, vous pouvez choisir l'option Créer un nouveau filtre, qui ouvre une nouvelle page pour créer un nouveau filtre de ligne ou de colonne. Pour plus d'informations, consultez Création de filtres de colonnes dans Amazon DataZone et Création de filtres de lignes dans Amazon DataZone.
7. (Facultatif) Entrez une réponse expliquant les raisons pour lesquelles vous avez accepté ou rejeté la demande.
8. Choisissez Approuver ou Rejeter.

En tant que propriétaire du projet, vous pouvez révoquer l'abonnement à tout moment. Pour de plus amples informations, veuillez consulter the section called "Révoquer un abonnement existant".

Pour consulter toutes les demandes d'abonnement, voirÉvénements et notifications.

> **Note**
>
> Amazon DataZone prend en charge le contrôle d'accès précis pour les tables AWS Glue, les tables Amazon Redshift et les vues Amazon Redshift.

Approbation automatique des demandes d'abonnement

Par défaut, les demandes d'abonnement à une ressource publiée nécessitent l'approbation manuelle du propriétaire des données. Amazon DataZone prend toutefois en charge deux scénarios dans lesquels les demandes d'abonnement peuvent être automatiquement approuvées :

- Approbation désactivée lors de la publication d'une ressource : lors de la publication d'une ressource de données, vous pouvez choisir de ne pas exiger l'approbation de l'abonnement. Dans ce cas, toutes les demandes d'abonnement entrantes pour cet actif sont automatiquement approuvées. Pour savoir comment désactiver l'approbation d'un actif, consultezPublier des actifs dans le DataZone catalogue Amazon à partir de l'inventaire du projet.
- Le demandeur est le propriétaire ou le contributeur du projet qui a publié la ressource. Une demande d'abonnement est également automatiquement approuvée si le demandeur est déjà autorisé à l'approuver manuellement. Plus précisément, s'ils sont membres à la fois du projet qui a publié la ressource et du projet demandant l'accès.

  Pour bénéficier de l'approbation automatique :

  - Le demandeur doit être répertorié en tant que propriétaire ou contributeur du projet dans lequel la ressource a été initialement publiée.
  - Le demandeur doit également être répertorié en tant que propriétaire ou contributeur du projet faisant la demande d'abonnement.

  Cela garantit que l'approbation automatique n'a lieu que lorsque le demandeur dispose d'une visibilité et d'autorisations dans les deux projets, celui qui partage la ressource et celui qui demande l'accès. Si le demandeur remplit les deux conditions, le système approuve automatiquement la demande.

Révoquer un abonnement existant sur Amazon DataZone

Amazon vous DataZone permet de trouver, d'accéder et de consommer les actifs du DataZone catalogue Amazon. Lorsque vous trouvez une ressource dans le catalogue à laquelle vous souhaitez accéder, vous devez vous y abonner, ce qui crée une demande d'abonnement. Un approbateur peut ensuite approuver ou demander votre demande. Il se peut que vous deviez révoquer un abonnement après l'avoir approuvé, soit parce que l'approbation était une erreur, soit parce que l'abonné n'a plus besoin d'accéder à l'actif.

Vous devez être membre du projet propriétaire (le projet qui a publié la ressource) pour révoquer un abonnement.

Pour révoquer un abonnement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet contenant l'abonnement que vous souhaitez révoquer.
3. Accédez à l'onglet Données, puis sélectionnez Demandes entrantes dans le volet de navigation de gauche.
4. Localisez l'abonnement que vous souhaitez révoquer et choisissez Afficher l'abonnement.
5. (Facultatif) Cochez la case pour permettre à l'abonné de conserver l'actif dans les objectifs d'abonnement du projet. Un objectif d'abonnement est une référence à un ensemble de ressources où les données souscrites peuvent être mises à disposition dans un environnement.

   Si vous souhaitez révoquer ultérieurement l'accès à la ressource depuis la cible d'abonnement, vous devez le faire dans AWS Lake Formation.
6. Choisissez Révoquer l'abonnement.

Vous ne pouvez pas réapprouver un abonnement après l'avoir révoqué. L'abonné doit s'abonner à nouveau à la ressource pour que vous puissiez l'approuver.

> **Note**
>
> La révocation d'un abonnement affecte uniquement l'accès de l'utilisateur concerné à l'actif, à savoir l'abonné dont vous révoquez l'abonnement. L'actif reste intact et l'utilisateur (abonné) le reste également. Cet utilisateur ne peut pas accéder à la ressource tant qu'il n'a pas soumis et approuvé une autre demande d'abonnement.

Annuler une demande d'abonnement sur Amazon DataZone

Amazon vous DataZone permet de trouver, d'accéder et de consommer les actifs du DataZone catalogue Amazon. Lorsque vous trouvez une ressource dans le catalogue à laquelle vous souhaitez accéder, vous devez vous y abonner, ce qui crée une demande d'abonnement. Un approbateur peut ensuite approuver ou demander votre demande. Vous devrez peut-être annuler une demande d'abonnement en attente, soit parce que vous l'avez soumise par erreur, soit parce que vous n'avez plus besoin d'un accès en lecture à la ressource.

Pour annuler une demande d'abonnement, vous devez être propriétaire du projet ou contributeur.

Pour annuler une demande d'abonnement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet contenant la demande d'abonnement.
3. Accédez à l'onglet Données du projet, puis choisissez Données demandées dans le volet de navigation de gauche. Cette page répertorie les actifs auxquels le projet a demandé l'accès.
4. Filtrez par Demandé pour ne voir que les demandes toujours en attente. Localisez la demande et choisissez Afficher la demande.
5. Passez en revue la demande d'abonnement et choisissez Annuler la demande.

Si vous souhaitez vous réabonner à la ressource (ou à une autre ressource), consultezthe section called “Demander un abonnement à des actifs”.

> **Note**
>
> Une demande d'abonnement en attente peut être annulée lorsqu'il n'est plus nécessaire d'accéder en « lecture » à la ressource. L'actif et l'utilisateur dont la demande d'abonnement en attente est annulée ne sont pas affectés par cette action.

Se désabonner d'un actif sur Amazon DataZone

Amazon vous DataZone permet de trouver, d'accéder et de consommer les actifs du DataZone catalogue Amazon. Lorsque vous trouvez une ressource dans le catalogue à laquelle vous souhaitez accéder, vous devez vous y abonner, ce qui crée une demande d'abonnement. Un approbateur peut ensuite approuver ou demander votre demande. Il se peut que vous deviez vous désinscrire d'un actif, soit parce que vous vous êtes inscrit par erreur et que vous avez été approuvé, soit parce que vous n'avez plus besoin d'un accès en lecture à l'actif.

Vous devez être membre d'un projet pour vous désinscrire de l'un de ses actifs.

Pour vous désabonner d'un actif

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet contenant la ressource dont vous souhaitez vous désabonner.
3. Accédez à l'onglet Données du projet, puis choisissez Données demandées dans le volet de navigation de gauche. Cette page répertorie les actifs auxquels le projet a demandé l'accès.
4. Filtrez par Approuvé pour ne voir que les demandes approuvées. Localisez la demande et choisissez Afficher l'abonnement.
5. Vérifiez l'abonnement et choisissez Se désabonner.

Si vous souhaitez vous réabonner à la ressource (ou à une autre ressource), consultezthe section called “Demander un abonnement à des actifs”.

> **Note**
>
> Lorsqu'un utilisateur n'a plus besoin d'accéder à une ressource, il peut choisir l'option de désinscription. La ressource reste intacte, aucune ressource n'est supprimée à la suite de cette action.

Utilisation des rôles IAM existants pour traiter les abonnements Amazon DataZone

Dans la version actuelle, Amazon vous DataZone aide à utiliser vos rôles IAM existants pour accéder aux données. Pour ce faire, vous pouvez créer un objectif d'abonnement dans l' DataZone environnement Amazon que vous utilisez pour exécuter votre abonnement. Pour créer un objectif d'abonnement pour un environnement dans l'un des AWS comptes associés, vous pouvez suivre les étapes suivantes :

Étape 1 : Assurez-vous que votre DataZone domaine Amazon utilise la version 2 ou supérieure de la politique de RAM

1. Accédez à la page Shared by me : Resource shares dans la console AWS RAM.
2. Étant donné que les partages de ressources AWS RAM existent dans des AWS régions spécifiques, choisissez la AWS région appropriée dans la liste déroulante située dans le coin supérieur droit de la console.
3. Sélectionnez le partage de ressources correspondant à votre DataZone domaine Amazon, puis choisissez Modifier. Vous pouvez identifier le partage de RAM pour le DataZone domaine Amazon à l'aide du nom ou de l'ID du domaine, car le partage de RAM est créé avec le nom :`DataZone-<domain-name>-<domain-id>`.
4. Choisissez Next pour passer à l'étape suivante où vous pouvez vérifier la version de la politique de RAM et la modifier.
5. Assurez-vous que la version de la politique de RAM est la version 2 ou supérieure. Si ce n'est pas le cas, utilisez le menu déroulant pour sélectionner la version 2 ou supérieure.
6. Choisissez Passer à l'étape 4 : Révision et mise à jour.
7. Choisissez Mettre à jour le partage de ressources.

Étape 2 : créer un objectif d'abonnement à partir d'un compte associé

- Dans la version actuelle, Amazon DataZone prend en charge la création d'objectifs d'abonnement en utilisant APIs uniquement. Vous trouverez ci-dessous quelques exemples de charge utile que vous pouvez utiliser pour créer un objectif d'abonnement afin de satisfaire les abonnements à vos tables AWS Glue et à vos tables ou vues Amazon Redshift. Pour de plus amples informations, veuillez consulter CreateSubscriptionTarget.

  Exemple d'objectif d'abonnement pour AWS Glue

```
{
        "domainIdentifier": "<DOMAIN_ID>",
        "environmentIdentifier": "<ENVIRONMENT_ID>",
        "name": "<SUBSCRIPTION_TARGET_NAME>",
        "type": "GlueSubscriptionTargetType",
        "authorizedPrincipals" : ["IAM_ROLE_ARN"],
        "subscriptionTargetConfig" : [{"content": "{\"databaseName\":
 \"<DATABASE_NAME>\"}", "formName": "GlueSubscriptionTargetConfigForm"}],
        "manageAccessRole": "<GLUE_DATA_ACCESS_ROLE_IN_ASSOCIATED_ACCOUNT_ARN>",
        "applicableAssetTypes" : ["GlueTableAssetType"],
        "provider": "Amazon DataZone"
}
```

Exemple d'objectif d'abonnement pour Amazon Redshift :

```
{
        "domainIdentifier": "<DOMAIN_ID>",
        "environmentIdentifier": "<ENVIRONMENT_ID>",
        "name": "<SUBSCRIPTION_TARGET_NAME>",
        "type": "RedshiftSubscriptionTargetType",
        "authorizedPrincipals" : ["REDSHIFT_DATABASE_ROLE_NAME"],
        "subscriptionTargetConfig" : [{"content": "{\"databaseName\":
 \"<DATABASE_NAME>\", \"secretManagerArn\": \"<SECRET_MANAGER_ARN>
\",\"clusterIdentifier\": \"<CLUSTER_IDENTIFIER>\"}", "formName":
 "RedshiftSubscriptionTargetConfigForm"}],
        "manageAccessRole":
 "<REDSHIFT_DATA_ACCESS_ROLE_IN_ASSOCIATED_ACCOUNT_ARN>",
        "applicableAssetTypes" : ["RedshiftViewAssetType",
 "RedshiftTableAssetType"],
        "provider": "Amazon DataZone"
}
```

⚠ Important

- L'EnvironmentIdentifier que vous utilisez dans l'appel d'API ci-dessus doit exister dans le même compte associé à partir duquel vous effectuez l'appel d'API. Dans le cas contraire, l'appel d'API échouera.
- L'ARN du rôle IAM que vous utilisez dans les « AuthorizedPrincipals » est le rôle auquel Amazon DataZone accordera l'accès après l'ajout d'un actif souscrit à la cible d'abonnement. Ces principaux autorisés doivent appartenir au même compte que l'environnement dans lequel l'objectif d'abonnement est créé.
- La valeur du champ fournisseur doit être « Amazon DataZone » pour qu'Amazon DataZone puisse terminer le traitement de l'abonnement.
- Le nom de base de données fourni dans subscriptionTargetConfig doit déjà exister dans le compte dans lequel la cible est créée. Amazon ne DataZone créera pas cette base de données. Assurez-vous également que le rôle de gestion des accès dispose de l'autorisation CREATE TABLE sur cette base de données.
- Assurez-vous également que les rôles (rôle IAM pour le AWS Glue et rôle de base de données pour Amazon Redshift) fournis en tant que principaux autorisés existent déjà dans le compte d'environnement. Pour les cibles d'abonnement Amazon Redshift, des mises à jour supplémentaires sont requises pour le rôle assumé lors de la connexion au cluster. Ce rôle doit être associé à une RedshiftDbRoles étiquette. La valeur de la balise peut être une liste séparée par des virgules. La valeur doit être le rôle de base de données fourni en tant que principal autorisé lors de la création de la cible d'abonnement.

Étape 3 : Abonnement à une nouvelle table et finalisation de l'abonnement à la nouvelle cible

- Une fois que vous avez créé l'objectif d'abonnement, vous pouvez vous abonner à un nouveau tableau et Amazon l' DataZone atteindra conformément à l'objectif ci-dessus.

Accorder l'accès aux AWS Glue Data Catalog actifs gérés sur Amazon DataZone

Sur Amazon DataZone, les demandes d'abonnement et les abonnements approuvés ou accordés pour l'accès en lecture aux actifs sont gérés par les propriétaires des actifs.

> **Note**
>
> La gestion de l'accès aux AWS Glue Data Catalog actifs à l'aide de la méthode AWS Lake Formation LF-TBAC n'est pas prise en charge.
> Support pour le partage d'actifs entre régions n' AWS Glue Data Catalog est pas pris en charge.

Une fois qu'une demande d'abonnement aux AWS Glue Data Catalog actifs gérés est approuvée, Amazon ajoute DataZone automatiquement ces actifs à tous les environnements de lac de données existants du projet. Amazon accorde et gère DataZone ensuite l'accès aux AWS Glue Data Catalog tables approuvées en votre nom par le biais de AWS Lake Formation. Pour le projet d'abonné, les actifs accordés apparaissent sous AWS Glue Data Catalog forme de ressources sur votre compte. Vous pouvez ensuite utiliser Amazon Athena pour interroger les tables.

> **Note**
>
> Si un nouvel environnement de lac de données est ajouté au projet après que les AWS Glue Data Catalog actifs souscrits ont été automatiquement ajoutés aux environnements de lac de données existants, vous devez ajouter manuellement ces AWS Glue Data Catalog actifs souscrits à ce nouvel environnement de lac de données. Vous pouvez le faire en choisissant l'option Ajouter une subvention dans l'onglet Données de la page de présentation du projet sur le portail de DataZone données Amazon.

Pour qu'Amazon DataZone puisse accorder l'accès aux tables du AWS Glue Data Catalog, les conditions suivantes doivent être remplies.

- La table AWS Glue doit être gérée par Lake Formation, car Amazon DataZone accorde l'accès en gérant les autorisations de Lake Formation.

- Le rôle Manage access pour l'environnement de lac de données utilisé pour publier la table AWS Glue Data Catalog doit disposer des autorisations Lake Formation suivantes :
  - `DESCRIBE`et `DESCRIBE GRANTABLE` des autorisations sur la base de données AWS Glue qui contient la table publiée.
  - `DESCRIBE`,`SELECT`,`DESCRIBE GRANTABLE`, `SELECT GRANTABLE` autorisations dans Lake Formation sur le tableau publié lui-même.

Pour plus d'informations, consultez la section Octroi et révocation d'autorisations sur les ressources du catalogue dans le Guide du AWS Lake Formation développeur.

Accorder l'accès aux ressources Amazon Redshift gérées sur Amazon DataZone

Lorsqu'un abonnement à une table ou à une vue Amazon Redshift est approuvé, Amazon DataZone peut automatiquement ajouter la ressource abonnée à tous les environnements d'entrepôt de données du projet, afin que les membres du projet puissent interroger les données à l'aide du lien de l'éditeur de requêtes Amazon Redshift dans leur environnement. Amazon DataZone crée en sous-main les subventions et les partages de données nécessaires entre la source et la cible de l'abonnement.

Le processus d'octroi de l'accès varie en fonction de l'emplacement de la base de données source (éditeur) et de la base de données cible (abonné).

- Même cluster, même base de données : si les données doivent être partagées au sein de la même base de données, Amazon DataZone accorde les autorisations directement sur la table source.
- Même cluster, base de données différente : si les données doivent être partagées entre deux bases de données au sein du même cluster, Amazon DataZone crée une vue dans la base de données cible et des autorisations sont accordées sur la vue créée.
- Même compte, cluster différent : Amazon DataZone crée un partage de données entre le cluster source et le cluster cible et crée une vue au-dessus de la table partagée. Les autorisations sont accordées sur la vue.
- Compte croisé : comme ci-dessus, mais une étape supplémentaire est requise pour autoriser le partage de données entre comptes du côté du cluster producteur et une autre étape pour associer le partage de données du côté du cluster de consommateurs.

> **Note**
>
> Si un nouvel environnement d'entrepôt de données est ajouté au projet après que les actifs Amazon Redshift souscrits ont été automatiquement ajoutés aux environnements d'entrepôt de données existants, vous devez ajouter manuellement ces actifs Amazon Redshift souscrits à ce nouvel environnement d'entrepôt de données. Vous pouvez le faire en choisissant l'option Ajouter une subvention dans l'onglet Données de la page de présentation du projet sur le portail de DataZone données Amazon.

Assurez-vous que les clusters Amazon Redshift que vous publiez et auxquels vous vous abonnez répondent à toutes les exigences relatives aux partages de données Amazon Redshift. Pour plus d'informations, consultez le guide du développeur Amazon Redshift.

> **Note**
>
> Amazon DataZone prend en charge l'attribution automatique d'abonnements aux actifs Amazon Redshift Cluster et Amazon Redshift Serverless.
> Le partage de données entre régions à l'aide d'Amazon Redshift n'est pas pris en charge.

Accorder l'accès aux actifs non gérés sur Amazon pour les abonnements approuvés DataZone

Sur Amazon DataZone, les demandes d'abonnement et les abonnements approuvés ou accordés pour l'accès en lecture aux actifs sont gérés par les propriétaires des actifs.

Amazon DataZone permet aux utilisateurs de publier n'importe quel type d'actif dans le catalogue de données commerciales. Pour certains de ces actifs, Amazon DataZone peut gérer automatiquement les autorisations d'accès. Ces actifs sont appelés actifs gérés et incluent les tables AWS Glue Data Catalog gérées par Lake Formation et les tables et vues Amazon Redshift. Tous les autres actifs auxquels Amazon ne DataZone peut pas automatiquement octroyer des abonnements sont considérés comme non gérés.

Amazon vous DataZone propose un moyen de gérer les autorisations d'accès pour vos actifs non gérés. Lorsqu'un abonnement à un actif du catalogue de données commerciales est approuvé par le propriétaire des données, Amazon DataZone publie un événement sur Amazon EventBridge dans votre compte avec toutes les informations nécessaires dans la charge utile qui vous permet de créer

les autorisations d'accès entre la source et la cible. Lorsque vous recevez cet événement, vous pouvez déclencher un gestionnaire personnalisé qui peut utiliser les informations de l'événement pour créer les autorisations ou les autorisations nécessaires. Une fois que vous avez accordé l'accès, vous pouvez signaler et mettre à jour le statut de l'abonnement sur Amazon DataZone afin que celui-ci puisse informer les utilisateurs abonnés à l'actif qu'ils peuvent commencer à consommer l'actif. Pour de plus amples informations, veuillez consulter DataZone Événements et notifications Amazon.

Interrogez des données dans Amazon Athena ou Amazon Redshift dans Amazon DataZone

Dans Amazon DataZone, une fois qu'un abonné a accès à un actif du catalogue, il peut l'utiliser (interroger et analyser) à l'aide d'Amazon Athena ou de l'éditeur de requêtes Amazon Redshift v2. Vous devez être propriétaire du projet ou contributeur pour effectuer cette tâche. En fonction des plans activés dans le projet, Amazon DataZone fournit des liens vers l'éditeur de requêtes Amazon and/or Athena Amazon Redshift v2 dans le volet droit de la page du projet dans le portail de données.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Sur le portail de DataZone données Amazon, choisissez Parcourir la liste des projets, puis recherchez et choisissez le projet pour lequel vous avez les données que vous souhaitez analyser.
3. Si le plan Data Lake est activé sur ce projet, un lien vers Amazon Athena s'affiche dans le panneau de droite de la page d'accueil du projet.

   Si le plan de l'entrepôt de données est activé sur ce projet, un lien vers l'éditeur de requêtes s'affiche dans le panneau de droite de la page d'accueil du projet.

   > **Note**
   >
   > Les plans sont définis dans le profil d'environnement avec lequel un projet est créé.

Rubriques

- Interrogez des données à l'aide d'Amazon Athena

- Interrogez des données à l'aide d'Amazon Redshift

Interrogez des données à l'aide d'Amazon Athena

Cliquez sur le lien Amazon Athena pour ouvrir l'éditeur de requêtes Amazon Athena dans un nouvel onglet du navigateur en utilisant les informations d'identification du projet pour l'authentification. Le DataZone projet Amazon sur lequel vous travaillez est automatiquement sélectionné comme groupe de travail actuel dans l'éditeur de requêtes.

Dans l'éditeur de requêtes Amazon Athena, rédigez et exécutez vos requêtes. Voici quelques tâches courantes :

- Interrogez et analysez vos actifs souscrits
- Création de nouvelles tables
- Création d'une table à partir des résultats de requête (CTAS) d'un compartiment S3 externe

Interrogez et analysez vos actifs souscrits

Si l'accès aux actifs auxquels votre projet est abonné n'est pas automatiquement accordé par Amazon DataZone, vous devez être autorisé à accéder aux données sous-jacentes. Pour plus d'informations sur la manière d'accorder l'accès à ces actifs, consultezAccorder l'accès aux actifs non gérés sur Amazon pour les abonnements approuvés DataZone.

Si l'accès aux ressources auxquelles votre projet est abonné est automatiquement accordé par Amazon DataZone, vous pouvez exécuter des requêtes SQL sur les tables et consulter les résultats dans Amazon Athena. Pour plus d'informations sur l'utilisation de SQL dans Amazon Athena, consultez la référence SQL pour Athena.

Lorsque vous accédez à l'éditeur de requêtes Amazon Athena après avoir choisi le lien Amazon Athena dans le panneau de droite de la page d'accueil du projet, une liste déroulante de projets s'affiche dans le coin supérieur droit de l'éditeur de requêtes Amazon Athena et le contexte de votre projet est automatiquement sélectionné.

Vous pouvez voir les bases de données suivantes dans le menu déroulant Base de données :

- Une base de données de publication (*{environmentname}*_pub_db). Le but de cette base de données est de vous fournir un environnement dans lequel vous pouvez produire de nouvelles données dans le contexte de votre projet, puis publier ces données dans le DataZone catalogue

Amazon. Les propriétaires de projets et les contributeurs ont un accès en lecture et en écriture à cette base de données. Les spectateurs du projet n'ont qu'un accès en lecture à cette base de données.

- Une base de données d'abonnement (*{environmentname}*_sub_db). Le but de cette base de données est de partager avec vous les données auxquelles vous vous êtes abonné en tant que membre du projet dans le DataZone catalogue Amazon, et de vous permettre d'interroger ces données.

Création de nouvelles tables

Si vous êtes connecté à un compartiment S3 externe, vous pouvez utiliser Amazon Athena pour interroger et analyser les actifs d'un compartiment Amazon S3 externe. Dans ce scénario, Amazon DataZone n'est pas autorisé à accorder l'accès directement aux données sous-jacentes dans le compartiment externe Amazon S3, et les données Amazon S3 externes créées en dehors du projet ne sont pas automatiquement gérées dans Lake Formation et ne peuvent pas être gérées par Amazon DataZone. Une autre solution consiste à copier les données du compartiment Amazon S3 externe vers une nouvelle table à l'intérieur du compartiment Amazon S3 du projet à l'aide d'une `CREATE TABLE` instruction dans Amazon Athena. Lorsque vous exécutez une `CREATE TABLE` requête dans Amazon Athena, vous enregistrez votre table auprès du. AWS Glue Data Catalog

Pour spécifier le chemin d'accès à vos données dans Simple Storage Service (Amazon S3), utilisez la propriété `LOCATION`, comme illustré dans l'exemple suivant :

```
CREATE EXTERNAL TABLE 'test_table'(
...
)
ROW FORMAT ...
STORED AS INPUTFORMAT ...
OUTPUTFORMAT ...
LOCATION 's3://bucketname/folder/'
```

Pour plus d'informations, consultez la section Emplacement des tables dans Amazon S3.

Création d'une table à partir des résultats de requête (CTAS) d'un compartiment S3 externe

Lorsque vous souscrivez à un actif, l'accès aux données sous-jacentes est en lecture seule. Vous pouvez utiliser Amazon Athena pour créer une copie du tableau. Dans Amazon Athena, `A CREATE TABLE AS SELECT (CTAS)` query crée une nouvelle table dans Amazon Athena à partir des résultats d'`SELECT`une instruction issue d'une autre requête. Pour plus d'informations sur la syntaxe CTAS, voir CREATE TABLE AS.

L'exemple suivant crée une table en copiant toutes les colonnes d'une table :

```
CREATE TABLE new_table AS
SELECT *
FROM old_table;
```

Dans la variation suivante du même exemple, votre instruction `SELECT` inclut également une clause `WHERE`. Dans ce cas, la requête sélectionne uniquement les lignes du tableau qui respectent la clause `WHERE` :

```
CREATE TABLE new_table AS
SELECT *
FROM old_table WHERE condition;
```

L'exemple suivant crée une nouvelle requête qui s'exécute sur un ensemble de colonnes à partir d'une autre table :

```
CREATE TABLE new_table AS
SELECT column_1, column_2, ... column_n
FROM old_table;
```

Cette variante du même exemple crée une nouvelle table à partir de colonnes spécifiques provenant de plusieurs tables :

```
CREATE TABLE new_table AS
SELECT column_1, column_2, ... column_n
FROM old_table_1, old_table_2, ... old_table_n;
```

Ces tables nouvellement créées font désormais partie de la AWS Glue base de données de vos projets et peuvent être rendues accessibles à d'autres personnes et partagées avec d'autres DataZone projets Amazon en publiant les données en tant que ressource dans le catalogue Amazon DataZone .

Interrogez des données à l'aide d'Amazon Redshift

Dans le portail de DataZone données Amazon, ouvrez un environnement qui utilise le plan de l'entrepôt de données. Choisissez le lien Amazon Redshift dans le panneau de droite de la page d'environnement. Cela ouvre une boîte de dialogue de confirmation contenant les informations nécessaires pour vous aider à établir une connexion au cluster Amazon Redshift ou au groupe de travail Amazon Redshift Serverless de votre environnement dans l'éditeur de requêtes Amazon Redshift v2.0. Une fois que vous avez identifié les informations nécessaires pour établir la connexion, cliquez sur le bouton Ouvrir Amazon Redshift. Cela ouvre l'éditeur de requêtes Amazon Redshift v2.0 dans un nouvel onglet du navigateur à l'aide des informations d'identification temporaires de l'environnement Amazon. DataZone

Dans l'éditeur de requêtes, suivez les étapes ci-dessous selon que votre environnement utilise un groupe de travail Amazon Redshift Serverless ou un cluster Amazon Redshift.

Pour un groupe de travail Amazon Redshift Serverless

1. Dans l'éditeur de requêtes, identifiez le groupe de travail Amazon Redshift Serverless de votre DataZone environnement Amazon, cliquez dessus avec le bouton droit de la souris et choisissez Create a connection.
2. Choisissez Federated User pour l'authentification.
3. Indiquez le nom de la base de données de DataZone l'environnement Amazon.
4. Choisissez Créer une connexion.

Pour un cluster Amazon Redshift :

1. Dans l'éditeur de requêtes, identifiez le cluster Amazon Redshift de votre DataZone environnement Amazon, cliquez dessus avec le bouton droit de la souris et choisissez Create a connection.
2. Sélectionnez Informations d'identification temporaires utilisant votre identité IAM pour l'authentification.
3. Si la méthode d'authentification ci-dessus n'est pas disponible, ouvrez les paramètres du compte en cliquant sur le bouton en forme de roue dentée dans le coin inférieur gauche, choisissez Authentifier avec les informations d'identification IAM et enregistrez. Il s'agit d'un one-time-only réglage.
4. Indiquez le nom de la base de données de DataZone l'environnement Amazon pour créer la connexion.
5. Choisissez Créer une connexion.

Vous pouvez maintenant commencer à interroger les tables et les vues du cluster Amazon Redshift ou du groupe de travail Amazon Redshift Serverless configuré pour votre environnement Amazon. DataZone

Toutes les tables ou vues Amazon Redshift auxquelles vous êtes abonné sont liées au cluster Amazon Redshift ou au groupe de travail Amazon Redshift Serverless configuré pour l'environnement. Vous pouvez vous abonner aux tables et aux vues ainsi que publier les nouvelles tables et vues que vous créez dans le cluster ou la base de données de votre environnement.

Prenons par exemple un scénario dans lequel un environnement est lié à un cluster Amazon Redshift appelé `redshift-cluster-1` et une base de données appelée `dev` dans ce cluster. À l'aide du portail de DataZone données Amazon, vous pouvez interroger les tables et les vues ajoutées à votre environnement. Dans la `Analytics tools` section du volet droit du portail de données, vous pouvez choisir le lien Amazon Redshift pour cet environnement, qui ouvre l'éditeur de requêtes. Vous pouvez ensuite cliquer avec le bouton droit sur le `redshift-cluster-1` cluster et créer une connexion à l'aide d'informations d'identification temporaires utilisant votre identité IAM. Une fois la connexion établie, vous pouvez voir toutes les tables et vues auxquelles votre environnement a accès dans la base de données de développement.

Règles d'application des métadonnées pour les demandes d'abonnement

La fonctionnalité relative aux règles d'application des métadonnées pour les demandes d'abonnement d'Amazon DataZone renforce la gouvernance des données en permettant aux propriétaires d'unités de domaine d'établir des exigences claires en matière de métadonnées pour les consommateurs de données, en rationalisant les demandes d'accès et en améliorant la gouvernance des données. Cette fonctionnalité permet aux organisations de s'aligner sur les normes de métadonnées de l'organisation, de mettre en œuvre des flux de travail personnalisés et de fournir une expérience d'accès aux données cohérente et gouvernée.

Cette fonctionnalité est prise en charge dans toutes les régions AWS commerciales dans lesquelles Amazon DataZone est actuellement disponible.

Les propriétaires d'unités de domaine peuvent suivre la procédure suivante pour configurer l'application des métadonnées dans Amazon DataZone :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Choisissez Domaines, accédez à l'onglet Unités de domaine et choisissez l'unité de domaine avec laquelle vous souhaitez travailler.
3. Cliquez sur l'onglet Règles, puis sur Ajouter.
4. Sur la page Créer une règle de formulaire de métadonnées requises, procédez comme suit, puis choisissez Ajouter une règle :

    - Donnez un nom à votre règle.
    - Sous Action, sélectionnez Demande d'abonnement.
    - Sous Formulaires obligatoires, choisissez Ajouter un formulaire de métadonnées, choisissez un formulaire de métadonnées dans le domaine/l'unité de domaine que vous souhaitez ajouter à cette règle, puis choisissez Ajouter. Vous pouvez ajouter jusqu'à 5 formulaires de métadonnées par règle.

- Sous Champ d'application, spécifiez les entités de données auxquelles vous souhaitez associer ces formulaires. Vous pouvez choisir des produits de données et and/or des actifs de données.
- Sous Types de ressources de données, spécifiez si la règle s'applique à tous les types de ressources ou limitez-la aux types de ressources sélectionnés.
- Sous Projets, spécifiez si les formulaires requis seront associés aux and/or actifs de produits de données publiés par tous les projets ou uniquement aux projets sélectionnés dans cette unité de domaine. Vérifiez également la règle de cascade pour les unités de domaine enfant si vous souhaitez que les unités de domaine enfant héritent de cette exigence.

Une fois l'application des métadonnées configurée, les consommateurs de données peuvent suivre la procédure suivante pour demander l'accès :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com /datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Utilisez la barre de recherche pour rechercher et choisir la ressource à laquelle vous souhaitez vous abonner, puis choisissez S'abonner.
3. Dans la fenêtre contextuelle S'abonner, saisissez les informations suivantes :
   - Le projet auquel vous souhaitez souscrire à la ressource.
   - Brève justification de votre demande d'abonnement.
   - Compléter les métadonnées requises : spécifiez les champs de métadonnées requis tels que spécifiés par l'unité de domaine. Si les champs obligatoires sont incomplets, ils sont surlignés et la soumission est désactivée jusqu'à ce qu'elle soit résolue. Une fois que tous les champs obligatoires sont remplis, sélectionnez Appliquer.
4. Sélectionnez Demande pour soumettre la demande d'abonnement. Après la soumission, un événement est généré dans EventBridge, qui peut être utilisé dans des flux de travail personnalisés en dehors d'Amazon DataZone selon les besoins. Vous recevez une notification sur le portail de données lorsque l'éditeur approuve votre demande.

Les producteurs de données peuvent suivre la procédure suivante pour approuver la demande d'abonnement :

Pour approuver ou rejeter une demande d'abonnement

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Dans le portail de données, choisissez Parcourir la liste des projets et sélectionnez le projet contenant la ressource avec la demande d'abonnement.
3. Accédez à l'onglet Données, puis sélectionnez Demandes entrantes dans le volet de navigation de gauche.
4. Localisez la demande et choisissez Afficher la demande. Vous pouvez filtrer par En attente pour ne voir que les demandes encore ouvertes.
5. Passez en revue la demande d'abonnement et le motif de l'accès, puis décidez de l'approuver ou de la rejeter.

   Les producteurs de données peuvent examiner les métadonnées fournies, y compris les liens vers les documents et le compte IDs, afin de déterminer si la demande répond aux exigences de conformité et de flux de travail avant d'accorder l'accès.
6. Pour approuver, sélectionnez l'une des deux options suivantes :

   - Accès complet : si vous choisissez d'approuver l'abonnement avec l'option d'accès complet, l'abonné aura accès à toutes les lignes et colonnes de votre ressource de données.
   - Approuver avec des filtres de ligne et de colonne : pour limiter l'accès à des lignes et colonnes de données spécifiques, vous pouvez choisir l'option d'approbation avec des filtres de ligne et de colonne. Pour de plus amples informations, veuillez consulter Contrôle d'accès précis aux données sur Amazon DataZone.
     - Sélectionnez Choisir des filtres, puis dans le menu déroulant, sélectionnez un ou plusieurs filtres disponibles que vous souhaitez appliquer à l'abonnement.
     - Pour créer un nouveau filtre, vous pouvez choisir l'option Créer un nouveau filtre, qui ouvre une nouvelle page pour créer un nouveau filtre de ligne ou de colonne. Pour plus d'informations, consultez Création de filtres de colonnes dans Amazon DataZone et Création de filtres de lignes dans Amazon DataZone.
7. (Facultatif) Entrez une réponse expliquant les raisons pour lesquelles vous avez accepté ou rejeté la demande.

8. Choisissez soit Approuver.

Analysez les données des DataZone abonnés Amazon avec des applications d'analyse externes via une connexion JDBC

Amazon DataZone permet aux consommateurs de données de localiser et de s'abonner facilement à des données provenant de plusieurs sources dans le cadre d'un même projet et d'analyser ces données à l'aide d'Amazon Athena, d'Amazon Redshift Query Editor et d'Amazon. SageMaker

Amazon prend DataZone également en charge l'authentification via le pilote Athena JDBC qui permet aux utilisateurs d'interroger les DataZone données Amazon auxquelles ils sont abonnés à l'aide d'outils SQL et d'analyse externes courants, tels que SQL Workbench, Tableau, Domino DBeaver, Power BI et bien d'autres. Les utilisateurs peuvent s'authentifier à l'aide de leurs informations d'identification professionnelles via SSO ou IAM et commencer à analyser les données qu'ils ont souscrites dans le cadre de leurs projets Amazon. DataZone

DataZoneLa prise en charge du pilote JDBC Athena par Amazon offre les avantages suivants :

- Un plus grand choix d'outils pour les requêtes et la visualisation : les consommateurs de données peuvent se connecter à Amazon à DataZone l'aide de leurs outils préférés parmi une large gamme d'outils d'analyse compatibles avec une connexion JDBC. Cela leur permet de continuer à utiliser les logiciels qu'ils connaissent bien sans avoir à apprendre de nouveaux outils de consommation de données.
- Accès programmatique : une connexion JDBC aux données régies par l'accès via des serveurs ou des applications personnalisées permet aux consommateurs de données d'effectuer des opérations de données automatisées et plus complexes.

Vous pouvez utiliser votre URL JDBC pour connecter vos outils d'analyse externes aux données que vous avez DataZone souscrites à Amazon. Pour obtenir votre URL JDBC, effectuez la procédure suivante :

> **Important**
>
> Dans la version actuelle, Amazon DataZone prend en charge l'authentification à l'aide du pilote Amazon Athena JDBC. Pour terminer cette procédure, assurez-vous d'avoir téléchargé et installé le dernier pilote Athena JDBC pour l'application d'analyse de votre choix.

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Sur le portail de DataZone données Amazon, choisissez Parcourir la liste des projets, puis recherchez et choisissez le projet pour lequel vous avez les données que vous souhaitez analyser.
3. Dans le panneau de droite de la page d'accueil du projet, choisissez Connect with JDBC.
4. Dans la fenêtre contextuelle des paramètres JDBC, choisissez votre méthode d'authentification (informations d'identification SSO ou informations d'identification IAM), puis copiez la chaîne ou les paramètres individuels de l'URL JDBC. Vous pouvez ensuite l'utiliser pour vous connecter à votre application d'analyse externe.

Lorsque vous connectez votre application d'analyse externe à Amazon à DataZone l'aide de votre requête ou de vos paramètres JBDC, vous appelez l'`RedeemAccessToken`API. L'`RedeemAccessToken`API échange un jeton d'accès à Identity Center contre les `AmazonDataZoneDomainExecutionRole` informations d'identification, qui sont utilisées pour appeler l'`GetEnvironmentCredentials`API.

Pour plus d'informations sur le mécanisme d'authentification qui utilise les informations d'identification IAM pour se connecter aux données DataZone gérées par Amazon dans Athena, consultez DataZone la section Fournisseur d'informations d'identification IAM. Pour plus d'informations sur le mécanisme d'authentification qui permet de se connecter aux données DataZone gérées par Amazon dans Athena à l'aide d'IAM Identity Center, DataZone consultez Idc Credentials Provider.

RedeemAccessToken Référence d'API

Syntaxe de demande

```
POST /sso/redeem-token HTTP/1.1
Content-type: application/json

{
   "domainId": "string",
   "accessToken": "string"
}
```

Paramètres de requête

La demande utilise les paramètres suivants.

DomainId

L'ID du DataZone domaine Amazon.

Modèle : ^dzd [-_] [A-zA-z0-9_-] {1,36} $

Obligatoire : oui

Jeton d'accès

Le jeton d'accès à Identity Center.

Type : chaîne

Obligatoire : oui

Syntaxe de réponse

```
HTTP/1.1 200
Content-type: application/json

{
   "credentials": AwsCredentials
}
```

Éléments de réponse

informations d'identification

Les `AmazonDataZoneDomainExecutionRole` informations d'identification utilisées pour appeler l'`GetEnvironmentCredentials`API.

Type : tableau d'`AwsCredentials`objets. Ce type de données inclut les propriétés suivantes :

- accessKeyId: AccessKeyId

- secretAccessKey: SecretAccessKey
- SessionToken : SessionToken
- expiration : Horodatage

Jeton d'accès

Le jeton d'accès à Identity Center.

Type : chaîne

Obligatoire : oui

Erreurs

AccessDeniedException

Vous ne disposez pas d’un accès suffisant pour effectuer cette action.

Code d’état HTTP : 403

ResourceNotFoundException

La ressource spécifiée est introuvable.

Code d’état HTTP :404

ValidationException

L'entrée ne satisfait pas les contraintes spécifiées par le AWS service.

Code d’état HTTP : 400

InternalServerException

La demande a échoué en raison d'une erreur, d'une exception ou d'un échec inconnu.

Code d’état HTTP : 500

Contrôle d'accès précis aux données sur Amazon DataZone

Dans la version actuelle d'Amazon DataZone, le contrôle d'accès précis de vos données est pris en charge, ce qui vous permet d'avoir un contrôle d'accès granulaire sur vos données sensibles. Vous pouvez contrôler quel projet peut accéder à des enregistrements de données spécifiques au sein de vos actifs de données publiés dans le catalogue de données Amazon DataZone Business. Amazon DataZone prend en charge les filtres de lignes et de colonnes pour mettre en œuvre un contrôle d'accès précis.

Les filtres de lignes vous permettent de restreindre l'accès à des lignes spécifiques en fonction des critères que vous définissez. Par exemple, si votre tableau contient des données pour deux régions (Amérique et Europe) et que vous souhaitez vous assurer que les employés en Europe ne peuvent accéder qu'aux données relatives à leur région, vous pouvez créer un filtre de lignes incluant les lignes où la région est l'Europe (par exemple, région = « Europe »). De cette façon, les employés en Europe n'auront pas accès aux données américaines.

Les filtres de colonnes vous permettent de limiter l'accès à des colonnes spécifiques de vos actifs de données. Par exemple, si votre tableau contient des informations sensibles telles que des informations personnelles identifiables (PII), vous pouvez créer un filtre de colonne pour exclure les colonnes PII. Cela garantit que les abonnés ne peuvent accéder qu'à des données non sensibles.

Pour utiliser un contrôle d'accès précis, vous pouvez créer des filtres de ligne et de colonne pour vos ressources AWS Glue et Amazon Redshift sur Amazon. DataZone Lorsqu'une demande d'abonnement pour accéder à vos actifs de données est reçue, vous pouvez l'approuver en appliquant les filtres de ligne et de colonne appropriés. Amazon DataZone veille à ce que l'abonné ne puisse accéder qu'aux lignes et aux colonnes autorisées par les filtres que vous avez appliqués au moment de l'approbation de l'abonnement.

Rubriques

- Création de filtres de lignes dans Amazon DataZone
- Création de filtres de colonnes dans Amazon DataZone
- Supprimer les filtres de ligne ou de colonne dans Amazon DataZone
- Modifier les filtres de ligne ou de colonne dans Amazon DataZone
- Accorder l'accès à l'aide de filtres sur Amazon DataZone

Création de filtres de lignes dans Amazon DataZone

Amazon vous DataZone permet de créer des filtres de lignes que vous pouvez utiliser lors de l'approbation des abonnements afin de vous assurer que l'abonné ne peut accéder qu'aux lignes de données définies dans les filtres de lignes. Pour créer un filtre de ligne, suivez les étapes ci-dessous :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel appartient la ressource.
3. Accédez à l'onglet Données du projet.
4. Choisissez Données publiées dans le volet de navigation de gauche, puis sélectionnez la ressource pour laquelle vous souhaitez créer le filtre de ligne. Vous pouvez ajouter des filtres de ligne si votre ressource de données dans Amazon DataZone est du type AWS Glue table, Amazon Redshift table ou Amazon Redshift view.
5. Sur la page détaillée des actifs, accédez à l'onglet Filtres des actifs, puis choisissez Ajouter un filtre des actifs.
6. Configurez les champs suivants :
   - Nom : nom du filtre
   - Description : description des filtres
7. Sous Type de filtre, choisissez Filtre de ligne.
8. Sous expression de filtre de ligne, fournissez une ou plusieurs expressions pour le filtre de ligne.
   - Choisissez une colonne dans la colonne de la liste déroulante.
   - Choisissez un opérateur dans la liste déroulante des opérateurs.
   - Entrez une valeur dans le champ Valeur.
9. Pour ajouter une autre condition à votre expression de filtre, choisissez Ajouter une condition.
10. Lorsque vous utilisez plusieurs conditions dans l'expression du filtre de ligne, choisissez Et ou Ou pour lier les conditions.
11. Choisissez Create filter (Créer un filtre).

Pour plus d'informations sur la façon d'appliquer des filtres de ligne à un abonnement, consultezApprouver ou rejeter une demande d'abonnement sur Amazon DataZone.

Création de filtres de colonnes dans Amazon DataZone

Amazon vous DataZone permet de créer des filtres de colonne que vous pouvez utiliser lors de l'approbation des abonnements afin de vous assurer que l'abonné ne peut accéder qu'aux colonnes de données définies dans les filtres de colonnes. Pour créer un filtre de colonne, suivez les étapes ci-dessous :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Choisissez Sélectionner un projet dans le volet de navigation supérieur et sélectionnez le projet auquel appartient la ressource.
3. Accédez à l'onglet Données du projet.
4. Choisissez Données publiées dans le volet de navigation de gauche, puis sélectionnez la ressource pour laquelle vous souhaitez créer le filtre de colonne. Vous pouvez ajouter des filtres de colonne si votre ressource de données dans Amazon DataZone est du type AWS Glue table, Amazon Redshift table ou Amazon Redshift view.
5. Sur la page détaillée des actifs, accédez à l'onglet Filtres des actifs, puis choisissez Ajouter un filtre des actifs.
6. Configurez les champs suivants :
   - Nom : nom du filtre
   - Description : description des filtres
7. Sous Type de filtre, choisissez Filtre par colonne.
8. Sélectionnez les colonnes que vous souhaitez inclure dans les filtres en cochant à nouveau les colonnes de la ressource de données.
9. Choisissez Créer un filtre

Pour plus d'informations sur la façon d'appliquer des filtres de colonne à un abonnement, consultezApprouver ou rejeter une demande d'abonnement sur Amazon DataZone.

Supprimer les filtres de ligne ou de colonne dans Amazon DataZone

Pour supprimer un filtre de ligne ou de colonne, suivez les étapes ci-dessous :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Accédez à l'onglet Données du projet.
3. Choisissez Données publiées ou Données d'inventaire dans le volet de navigation de gauche, puis sélectionnez la ressource pour laquelle vous souhaitez supprimer un filtre de ligne ou de colonne.
4. Sur la page des détails des actifs, accédez à l'onglet Filtres des actifs, puis ouvrez le filtre que vous souhaitez supprimer.
5. Choisissez Actions, Supprimer, puis confirmez la suppression.

> **Note**
>
> Vous ne pouvez supprimer un filtre que s'il n'est pas utilisé dans les abonnements actifs.

Modifier les filtres de ligne ou de colonne dans Amazon DataZone

Pour modifier un filtre de ligne ou de colonne, suivez les étapes ci-dessous :

1. Accédez à l'URL du portail de DataZone données Amazon et connectez-vous à l'aide de l'authentification unique (SSO) ou de vos AWS informations d'identification. Si vous êtes un DataZone administrateur Amazon, vous pouvez accéder à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone et vous connecter avec l' Compte AWS endroit où le domaine a été créé, puis choisir Open data portal.
2. Accédez à l'onglet Données du projet.

3. Choisissez Données publiées ou Données d'inventaire dans le volet de navigation de gauche, puis sélectionnez la ressource pour laquelle vous souhaitez modifier un filtre de ligne ou de colonne.
4. Sur la page détaillée des actifs, accédez à l'onglet Filtres des actifs, puis ouvrez le filtre que vous souhaitez modifier.
5. Vous pouvez modifier le champ suivant :

   - Description : description des filtres
6. Si vous modifiez un filtre de ligne, vous pouvez mettre à jour l'expression du filtre de ligne.
7. Si vous modifiez un filtre de colonne, vous pouvez ajouter ou supprimer les colonnes sélectionnées dans le filtre.
8. Une fois que vous avez apporté les modifications, choisissez Modifier le filtre de ressources.

> **Note**
>
> Si vous modifiez un filtre utilisé dans les abonnements actifs, Amazon DataZone mettra automatiquement à jour les autorisations accordées aux projets des abonnés. Cela signifie que les abonnés ne pourront accéder qu'aux lignes ou aux colonnes définies dans le filtre mis à jour, ce qui garantit que vos politiques d'accès aux données sont appliquées de manière cohérente.

Accorder l'accès à l'aide de filtres sur Amazon DataZone

Amazon DataZone permet un contrôle d'accès précis en traduisant les filtres de ligne et de colonne définis en autorisations appropriées pour AWS Lake Formation et Amazon Redshift. Vous trouverez ci-dessous une explication de la manière dont Amazon DataZone matérialise ces filtres pour les tables AWS Glue et Amazon Redshift.

AWS Tables à colle

Lorsqu'un abonnement à une table AWS Glue avec des filtres par and/or colonnes de lignes est approuvé, Amazon DataZone matérialise l'abonnement en créant des subventions dans AWS Lake Formation with Data Cell Filters, garantissant ainsi que les membres du projet d'abonné ne peuvent accéder qu'aux lignes et colonnes auxquels ils sont autorisés à accéder en fonction des filtres appliqués à l'abonnement.

Amazon traduit DataZone d'abord les filtres de lignes et de colonnes appliqués dans Amazon DataZone en filtres de cellules de données AWS Lake Formation. Si plusieurs filtres de lignes et de colonnes sont utilisés, Amazon DataZone réunit toutes les colonnes et toutes les conditions de filtre de ligne pour calculer les autorisations effectives au niveau des lignes et des colonnes. Amazon crée DataZone ensuite un filtre de cellule de données unique sur la AWS Lake Formation en utilisant des autorisations de ligne et de colonne efficaces.

Une fois le filtre de cellule de données créé, Amazon DataZone partage la table des abonnés avec le projet d'abonné en créant des autorisations en lecture seule (SELECT) dans AWS Lake Formation à l'aide de ce filtre de cellule de données.

Amazon Redshift

Lorsqu'un abonnement à un Amazon Redshift table/view avec filtres de and/or colonnes de lignes est approuvé, Amazon DataZone matérialise l'abonnement en créant des vues de retard limitées dans Amazon Redshift, garantissant ainsi que les membres du projet d'abonné ne peuvent accéder qu'aux lignes et colonnes auxquels ils sont autorisés à accéder en fonction des filtres de lignes et de colonnes appliqués à l'abonnement.

Amazon traduit DataZone d'abord les filtres de lignes et de colonnes appliqués à un abonnement sur Amazon en une DataZone vue de reliure tardive sur Amazon Redshift. Si plusieurs filtres de lignes et de colonnes sont utilisés, Amazon DataZone réunit toutes les colonnes et toutes les conditions de filtrage de lignes pour calculer les autorisations effectives au niveau des lignes et des colonnes. Amazon crée DataZone ensuite la vue de liaison tardive en utilisant les autorisations de ligne et de colonne effectives.

Une fois la vue exécutoire tardive créée, Amazon DataZone partage cette vue avec les membres du projet d'abonnés en créant des autorisations de lecture seule (SELECT) dans Amazon Redshift.

DataZone Événements et notifications Amazon

Amazon vous DataZone tient informé des activités importantes au sein de votre portail de données, telles que les demandes d'abonnement, les mises à jour, les commentaires et les événements du système. Amazon vous DataZone fournit ces informations en envoyant des messages dans la boîte de réception dédiée du portail de données ou via le bus EventBridge par défaut d'Amazon.

Événements via la boîte de réception dédiée du portail de DataZone données Amazon

Amazon DataZone fournit une boîte de réception dédiée dans le portail de données, dans laquelle vous pouvez consulter vos messages et y donner suite. Les messages récents apparaissent également sur la page d'accueil, la page du projet et la page du catalogue. Par exemple, si un utilisateur demande l'accès à une ressource de données, les propriétaires du projet de publication et les contributeurs de cette ressource voient la demande dans le portail de données et une fois qu'une action est entreprise, les membres du projet abonné lié à cette demande voient la notification dans le portail de données. Il existe deux types de messages :

- Tâches : ces messages informent le destinataire qu'une action est nécessaire quelque part. Ils ont un champ de statut facultatif que vous pouvez utiliser pour le suivi.
- Événements : ces messages sont informatifs et n'ont aucun statut attribué. Les événements fournissent une piste d'audit des mises à jour récentes.

Dans Amazon DataZone, des messages sont générés pour les types d'événements suivants :

Catégorie d'événement	Nom de l’événement	Description de l'événement	Type d’événement
Abonnement	Demande d'abonnement créée	L'événement est généré lorsqu'une demande d'abonnement est créée	Sous-tâche
Abonnement	Demande d'abonnement acceptée	L'événement est généré lorsqu'une	Événement

Catégorie d'événement	Nom de l'événement	Description de l'événement	Type d'événement
		demande d'abonnement est acceptée	
Abonnement	Demande d'abonnement rejetée	Un événement est généré lorsqu'une demande d'abonnement est rejetée	Événement
Abonnement	Demande d'abonnement supprimée	L'événement est généré lorsqu'une demande d'abonnement est supprimée	Événement
Project	Création du projet réussie	L'événement est généré lorsque la création du projet réussit	Événement
Adhésion au projet	L'ajout d'un membre au projet a réussi	L'événement est généré lorsqu'un nouveau membre est ajouté à un projet	Événement
Adhésion au projet	Suppression d'un membre du projet réussie	L'événement est généré lorsqu'un membre est supprimé d'un projet	Événement
Adhésion au projet	Le changement de rôle du membre du projet a réussi	Un événement est généré, le rôle d'un membre dans le projet est modifié	Événement

Catégorie d'événement	Nom de l'événement	Description de l'événement	Type d'événement
Environnement	Déploiement de l'environnement lancé	L'événement est généré lorsqu'un déploiement d'environnement est lancé	Événement
Environnement	Déploiement de l'environnement terminé	Un événement est généré lorsqu'un déploiement d'environnement est terminé avec succès	Événement
Environnement	Le déploiement de l'environnement a échoué	Un événement est généré en cas d'échec du déploiement d'un environnement	Événement
Environnement	Flux de travail personnalisé de déploiement de l'environnement lancé	L'événement est généré lorsqu'un environnement avec un flux de travail personnalisé est lancé	Événement
Actif de données	Actif ajouté à l'inventaire	L'événement est généré lorsqu'une nouvelle ressource de données est ajoutée à l'inventaire, c'est-à-dire ajoutée au catalogue à l'état de brouillon	Événement

Catégorie d'événement	Nom de l'événement	Description de l'événement	Type d'événement
Actif de données	Ressource publiée	L'événement est généré lorsqu'une nouvelle ressource de données est publiée, c'est-à-dire disponible pour abonnement	Événement
Actif de données	Schéma des actifs modifié	Un événement est généré lorsqu'un schéma d'actif a changé depuis une tâche d'ingestion précédente	Événement
Abonnement en cours	Abonnement créé	L'événement est généré lorsqu'une personne demande à s'abonner à un actif de données	Sous-tâche
Abonnement en cours	Abonnement approuvé	L'événement est généré lorsqu'un abonnement est approuvé par le propriétaire ou le contributeur du projet de publication	Événement

Catégorie d'événement	Nom de l'événement	Description de l'événement	Type d'événement
Abonnement en cours	Abonnement refusé	L'événement est généré lorsqu'un abonnement est refusé par le propriétaire ou le contributeur du projet de publication	Événement
Abonnement en cours	Abonnement supprimé	L'événement est généré lorsqu'un abonnement est annulé par l'abonné	Événement
Abonnement en cours	Subvention d'abonnement demandée	L'événement est généré lorsqu'une personne demande l'accès à un actif	Événement
Abonnement en cours	Subvention d'abonnement terminée	L'événement est généré lorsqu'un abonnement est autorisé à accéder à la ressource par le propriétaire ou le contributeur du projet de publication	Événement
Abonnement en cours	Échec de la subvention d'abonnement	Un événement est généré lorsqu'une subvention d'abonnement échoue	Événement

Catégorie d'événement	Nom de l'événement	Description de l'événement	Type d'événement
Abonnement en cours	Demande de révocation de la subvention d'abonnement	Un événement est généré lorsqu'une subvention d'abonnement révoquée est initiée par le propriétaire ou le contributeur du projet de publication	Événement
Abonnement en cours	Révocation de la subvention d'abonnement terminée	L'événement est généré lorsqu'une révocation d'une subvention d'abonnement est terminée	Événement
Abonnement en cours	La révocation de l'autorisation d'abonnement a échoué	Un événement est généré en cas d'échec de la révocation d'une autorisation d'abonnement	Événement
Génération automatique de noms commerciaux	Nom commercial généré avec succès	Un événement est généré lorsque la tâche générée automatiquement par le nom de l'entreprise s'achève avec succès	Événement

Catégorie d'événement	Nom de l’événement	Description de l'événement	Type d’événement
Génération automatique de noms commerciaux	Le nom commercial généré a échoué	L'événement est généré lorsque la tâche générée automatiquement par le nom de l'entreprise échoue	Événement
Exécution de la source de données	Source de données créée	L'événement est généré lorsqu'une nouvelle source de données est créée	Événement
Exécution de la source de données	Source de données mise à jour	L'événement est généré lorsqu'une source de données existante est mise à jour	Événement
Exécution de la source de données	Exécution de la source de données déclenchée	L'événement est généré lorsqu'une exécution de source de données est lancée	Événement
Exécution de la source de données	L'exécution de la source de données a réussi	L'événement est généré lorsqu'une exécution de source de données réussit	Événement
Exécution de la source de données	L'exécution de la source de données a échoué	Un événement est généré en cas d'échec de l'exécution d'une source de données	Événement

Pour afficher les tâches dans la boîte de réception de votre portail de données, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le DataZone domaine Amazon a été créé.
2. Dans le portail de données, pour afficher une fenêtre contextuelle présentant les tâches récentes, sélectionnez l'icône en forme de cloche à côté de la barre de recherche.
3. Sélectionnez Afficher tout pour afficher toutes les tâches. Vous pouvez changer de vue et voir tous les événements en sélectionnant l'onglet Événements.
4. Vous pouvez filtrer la recherche en fonction du sujet de l'événement, du statut actif ou inactif ou de la plage de dates.
5. Choisissez une tâche individuelle pour accéder à l'emplacement où vous pouvez répondre à la tâche.

Pour consulter les événements dans la boîte de réception de votre portail de données, procédez comme suit :

1. Accédez au portail de DataZone données Amazon à l'aide de l'URL du portail de données et connectez-vous à l'aide de votre SSO ou de vos AWS informations d'identification. Si vous êtes DataZone administrateur Amazon, vous pouvez obtenir l'URL du portail de données en accédant à la DataZone console Amazon à l'adresse https://console.aws.amazon.com/datazone dans le AWS compte sur lequel le domaine DataZone racine Amazon a été créé.
2. Dans le portail de données, pour afficher la fenêtre contextuelle des derniers événements, sélectionnez l'icône en forme de cloche à côté de la barre de recherche.
3. Sélectionnez Afficher tout pour afficher tous les événements. Vous pouvez modifier les vues et voir toutes les tâches en sélectionnant l'onglet Tâches.
4. Filtrez la recherche par sujet de l'événement ou par plage de dates.
5. Choisissez un événement individuel pour accéder à l'emplacement où vous pouvez consulter les détails de cet événement.

Événements via le bus EventBridge par défaut d'Amazon

En plus d'envoyer des messages à votre boîte de réception dédiée dans le portail de données, il envoie DataZone également ces messages à votre bus d'événements EventBridge par défaut Amazon sur le même AWS compte où est hébergé votre domaine DataZone racine Amazon. Cela permet une automatisation axée sur les événements, telle que le traitement des abonnements ou des intégrations personnalisées avec d'autres outils. Vous pouvez créer des règles qui correspondent aux EventBridge événements Amazon entrants et les envoyer aux EventBridge cibles Amazon pour traitement. Une seule règle peut envoyer un événement à plusieurs cibles, qui peuvent ensuite s'exécuter en parallèle.

Voici un exemple d'événement :


```
{
  "version": "0",
  "id": "bd3d6239-2877-f464-0572-b1d76760e085",
  "detail-type": "Subscription Request Created",
  "source": "aws.datazone",
  "account": "111111111111",
  "time": "2023-11-13T17:57:00Z",
  "region": "us-east-1",
  "resources": [],
  "detail": {
    "version": "655",
    "metadata": {
      "domain": "dzd_bc8e1ez8r2a6xz",
      "user": "44f864b8-50a1-70cc-736f-c1f763934ab7",
      "id": "5jbc0lie0sr99j",
      "version": "1",
      "typeName": "SubscriptionRequestEntityType",
      "owningProjectId": "6oy92hwk937pgn",
      "awsAccountId": "111111111111",
      "clientToken": "e781b7b5-78c5-4608-961e-3792a6c3ff0d"
    },
    "data": {
      "autoApproved": true,
      "requesterId": "44f864b8-50a1-70cc-736f-c1f763934ab7",
      "status": "PENDING",
      "subscribedListings": [
        {
          "id": "ayzstznnx4dxyf",
```

```
          "ownerProjectId": "5a3se66qm88947",
          "version": "12"
        }
      ],
      "subscribedPrincipals": [
        {
          "id": "6oy92hwk937pgn",
          "type": "PROJECT"
        }
      ]
    }
  }
}
```

La liste complète des types de détails pris en charge par Amazon DataZone inclut :

- Demande d'abonnement créée
- Demande d'abonnement acceptée
- Demande d'abonnement rejetée
- Demande d'abonnement supprimée
- Subvention d'abonnement demandée
- Subvention d'abonnement terminée
- Échec de la subvention d'abonnement
- Demande de révocation de la subvention d'abonnement
- Révocation de l'autorisation d'abonnement terminée
- Échec de la révocation de l'autorisation d'abonnement
- Actif ajouté à l'inventaire
- Ressource ajoutée au catalogue
- Schéma des actifs modifié
- Modification du statut de la source de données
- Source de données créée
- Source de données mise à jour
- Exécution de la source de données déclenchée
- Exécution de la source de données réussie

- L'exécution de la source de données a échoué
- Création de domaine réussie
- Echec de la création du domaine
- Suppression du domaine réussie
- Echec de la suppression du domaine
- Déploiement de l'environnement commencé
- Déploiement de l'environnement terminé
- Échec du déploiement de l'environnement
- La suppression de l'environnement a commencé
- Suppression de l'environnement terminée
- La suppression de l'environnement a échoué
- Création du projet réussie
- L'ajout d'un membre au projet a réussi
- Suppression d'un membre du projet réussie
- Changement de rôle du membre du projet réussi
- Déploiement de l'environnement : flux de travail client
- Génération de noms commerciaux réussie
- Échec de la génération du nom commercial

Pour plus d'informations, consultez Amazon EventBridge.

Sécurité sur Amazon DataZone

La sécurité du cloud AWS est la priorité absolue. En tant que AWS client, vous bénéficiez de centres de données et d'architectures réseau conçus pour répondre aux exigences des entreprises les plus sensibles en matière de sécurité.

La sécurité est une responsabilité partagée entre vous AWS et vous. Le modèle de responsabilité partagée décrit ceci comme la sécurité du cloud et la sécurité dans le cloud :

- Sécurité du cloud : AWS est chargée de protéger l'infrastructure qui exécute les AWS services dans le AWS Cloud. AWS vous fournit également des services que vous pouvez utiliser en toute sécurité. Des auditeurs tiers testent et vérifient régulièrement l'efficacité de notre sécurité dans le cadre des programmes de AWS conformité Programmes de de conformité. Pour en savoir plus sur les programmes de conformité applicables à Amazon DataZone, consultez la section AWS Services concernés par programme de conformitéAWS .
- Sécurité dans le cloud — Votre responsabilité est déterminée par le AWS service que vous utilisez. Vous êtes également responsable d'autres facteurs, y compris de la sensibilité de vos données, des exigences de votre entreprise, ainsi que de la législation et de la réglementation applicables.

Cette documentation vous aide à comprendre comment appliquer le modèle de responsabilité partagée lors de l'utilisation d'Amazon DataZone. Les rubriques suivantes expliquent comment configurer Amazon pour répondre DataZone à vos objectifs de sécurité et de conformité. Vous apprendrez également à utiliser d'autres AWS services qui vous aident à surveiller et à sécuriser vos DataZone ressources Amazon.

Rubriques

- Protection des données sur Amazon DataZone
- Autorisation sur Amazon DataZone
- Contrôle de l'accès aux DataZone ressources Amazon à l'aide d'IAM
- Validation de conformité pour Amazon DataZone
- Bonnes pratiques en matière de sécurité pour Amazon DataZone
- Résilience chez Amazon DataZone
- Sécurité de l'infrastructure sur Amazon DataZone
- Prévention interservices confuse des adjoints sur Amazon DataZone

- Analyse de configuration et de vulnérabilité pour Amazon DataZone
- Domaines à ajouter à votre liste d'autorisations

Protection des données sur Amazon DataZone

Le modèle de responsabilité AWS partagée de s'applique à la protection des données sur Amazon DataZone. Comme décrit dans ce modèle, AWS est chargé de protéger l'infrastructure mondiale qui gère tous les AWS Cloud. La gestion du contrôle de votre contenu hébergé sur cette infrastructure relève de votre responsabilité. Vous êtes également responsable des tâches de configuration et de gestion de la sécurité des Services AWS que vous utilisez. Pour plus d'informations sur la confidentialité des données, consultez Questions fréquentes (FAQ) sur la confidentialité des données. Pour en savoir plus sur la protection des données en Europe, consultez le billet de blog Modèle de responsabilité partagée d'AWS et RGPD (Règlement général sur la protection des données) sur le Blog de sécuritéAWS .

À des fins de protection des données, nous vous recommandons de protéger les Compte AWS informations d'identification et de configurer les utilisateurs individuels avec AWS IAM Identity Center ou Gestion des identités et des accès AWS (IAM). Ainsi, chaque utilisateur se voit attribuer uniquement les autorisations nécessaires pour exécuter ses tâches. Nous vous recommandons également de sécuriser vos données comme indiqué ci-dessous :

- Utilisez l'authentification multifactorielle (MFA) avec chaque compte.
- SSL/TLS À utiliser pour communiquer avec AWS les ressources. Nous exigeons TLS 1.2 et recommandons TLS 1.3.
- Configurez l'API et la journalisation de l'activité des utilisateurs avec AWS CloudTrail. Pour plus d'informations sur l'utilisation des CloudTrail sentiers pour capturer AWS des activités, consultez la section Utilisation des CloudTrail sentiers dans le guide de AWS CloudTrail l'utilisateur.
- Utilisez des solutions de AWS chiffrement, ainsi que tous les contrôles de sécurité par défaut qu'ils contiennent Services AWS.
- Utilisez des services de sécurité gérés avancés tels qu'Amazon Macie, qui contribuent à la découverte et à la sécurisation des données sensibles stockées dans Amazon S3.
- Si vous avez besoin de modules cryptographiques validés par la norme FIPS 140-3 pour accéder AWS via une interface de ligne de commande ou une API, utilisez un point de terminaison FIPS. Pour plus d'informations sur les points de terminaison FIPS disponibles, consultez Norme FIPS (Federal Information Processing Standard) 140-3.

Nous vous recommandons fortement de ne jamais placer d'informations confidentielles ou sensibles, telles que les adresses e-mail de vos clients, dans des balises ou des champs de texte libre tels que le champ Nom. Cela inclut lorsque vous travaillez avec Amazon DataZone ou une autre entreprise Services AWS à l'aide de la console AWS CLI, de l'API ou AWS SDKs. Toutes les données que vous entrez dans des balises ou des champs de texte de forme libre utilisés pour les noms peuvent être utilisées à des fins de facturation ou dans les journaux de diagnostic. Si vous fournissez une adresse URL à un serveur externe, nous vous recommandons fortement de ne pas inclure d'informations d'identification dans l'adresse URL permettant de valider votre demande adressée à ce serveur.

Chiffrement des données

Lorsque vous accordez des autorisations, vous décidez qui obtient quelles autorisations pour quelles DataZone ressources Amazon. Vous activez des actions spécifiques que vous souhaitez autoriser sur ces ressources. Par conséquent, vous devez accorder uniquement les autorisations qui sont requises pour exécuter une tâche. L'implémentation d'un accès sur la base du moindre privilège est fondamentale pour réduire les risques de sécurité et l'impact que pourraient avoir des d'erreurs ou des actes de malveillance.

Chiffrement au repos

Amazon DataZone chiffre toutes vos données par défaut à l'aide d'une AWS clé de service de gestion des clés (AWS KMS) qui vous AWS appartient et gère pour vous. Vous pouvez également chiffrer les données stockées dans le DataZone catalogue Amazon à l'aide de clés que vous gérez avec AWS KMS.

Lorsque vous créez un domaine sur Amazon DataZone, vous pouvez fournir des paramètres de chiffrement en cochant la case à côté de Personnaliser les paramètres de chiffrement (avancés) sous Chiffrement des données, et en fournissant une clé KMS.

Chiffrement en transit

Amazon DataZone utilise le protocole TLS (Transport Layer Security) et le chiffrement côté client pour le chiffrement en transit. La communication avec Amazon DataZone se fait toujours via HTTPS, de sorte que vos données sont toujours cryptées pendant le transport.

Confidentialité du trafic inter-réseaux

Pour sécuriser les connexions entre les comptes, Amazon DataZone utilise des rôles de service et des rôles IAM pour se connecter en toute sécurité aux comptes clients et exécuter des opérations pour le compte du client.

Rubriques

- Le chiffrement des données est au repos pour Amazon DataZone
- Utilisation des points de terminaison VPC d'interface pour Amazon DataZone

Le chiffrement des données est au repos pour Amazon DataZone

Le chiffrement des données au repos par défaut permet de réduire les frais opérationnels et la complexité liés à la protection des données sensibles. Dans le même temps, il vous permet de créer des applications sécurisées qui répondent aux exigences réglementaires et de conformité strictes en matière de chiffrement.

Amazon DataZone utilise des clés AWS détenues par défaut pour chiffrer automatiquement vos données au repos. Vous ne pouvez pas consulter, gérer ou auditer l'utilisation des clés que vous AWS possédez. Pour plus d'informations, consultez Clés détenues par AWS.

Bien que vous ne puissiez pas désactiver cette couche de chiffrement ou sélectionner un autre type de chiffrement, vous pouvez choisir une clé gérée par le client lorsque vous créez vos domaines Amazon DataZone . Amazon DataZone prend en charge l'utilisation de clés symétriques gérées par le client que vous pouvez créer, posséder et gérer. Comme vous avez le contrôle total du chiffrement, vous pouvez effectuer les tâches suivantes :

- Établissement et gestion des stratégies de clé
- Établissement et gestion des politiques IAM et des octrois
- Activation et désactivation des stratégies de clé
- Rotation des matériaux de chiffrement de clé
- Ajout de balises
- Création d'alias de clé
- Planification de clés pour suppression

Pour utiliser votre propre clé, choisissez une clé gérée par le client lorsque vous créez votre DataZone domaine Amazon.

Pour plus d'informations, consultez Clés gérées par le client.

> **Note**
>
> Amazon active DataZone automatiquement le chiffrement au repos à l'aide de clés AWS détenues pour protéger gratuitement les données des clients.
> AWS Des frais KMS s'appliquent pour l'utilisation de clés gérées par le client. Pour plus d'informations sur la tarification, consultez Tarification du service de gestion des clés AWS.

Comment Amazon DataZone utilise les subventions dans AWS KMS

Amazon a DataZone besoin de deux autorisations pour utiliser votre clé gérée par le client. Lorsque vous créez un DataZone domaine Amazon chiffré à l'aide d'une clé gérée par le client, Amazon DataZone crée des subventions en votre nom en envoyant des CreateGrantdemandes à AWS KMS. Les subventions dans AWS KMS sont utilisées pour donner à Amazon DataZone l'accès à une clé KMS dans votre compte. Amazon DataZone crée les autorisations suivantes pour utiliser votre clé gérée par le client pour les opérations internes suivantes :

Une autorisation pour le chiffrement de vos données au repos pour les opérations suivantes :

- Envoyez DescribeKeydes demandes à AWS KMS pour vérifier que l'ID de clé KMS symétrique géré par le client saisi lors de la création d'un DataZone domaine Amazon est valide.
- Envoyez GenerateDataKeyà AWS KMS pour générer des clés de données chiffrées par votre clé gérée par le client.
- La demande Send Decrypt permet DataZone à Amazon de déchiffrer les données stockées.
- RetireGrantpour annuler la subvention lorsque le domaine est supprimé.

Une subvention pour la recherche, la découverte et l'exportation de vos données :

- DescribeKey- fournit au client les informations clés gérées par le client qui permettent DataZone à Amazon de valider la clé.
- Déchiffrer : permet DataZone à Amazon de déchiffrer les données stockées.

Vous pouvez révoquer à tout moment l'accès à l'attribution de la clé gérée par le client. Si vous le faites, Amazon DataZone ne pourra accéder à aucune des données chiffrées par la clé gérée par le client, ce qui affectera les opérations qui dépendent de ces données.

Création d'une clé gérée par le client

Vous pouvez créer une clé symétrique gérée par le client à l'aide de la console AWS de gestion ou du AWS KMS APIs.

Pour créer une clé symétrique gérée par le client, suivez les étapes de création d'une clé symétrique gérée par le client dans le guide du développeur du service de gestion des AWS clés.

Stratégie de clé : les stratégies de clé contrôlent l'accès à votre clé gérée par le client. Chaque clé gérée par le client doit avoir exactement une stratégie de clé, qui contient des instructions qui déterminent les personnes pouvant utiliser la clé et comment elles peuvent l'utiliser. Lorsque vous créez votre clé gérée par le client, vous pouvez spécifier une stratégie de clé. Pour plus d'informations, consultez la section Gestion de l'accès aux clés gérées par le client dans le Guide du développeur du service de gestion des AWS clés.

Pour utiliser votre clé gérée par le client avec vos DataZone ressources Amazon, les opérations d'API suivantes doivent être autorisées dans la politique relative aux clés :

- kms : CreateGrant — ajoute une autorisation à une clé gérée par le client. Accorde un accès de contrôle à une clé KMS spécifiée, ce qui permet d'accéder aux opérations d'octroi DataZone requises par Amazon. Pour plus d'informations sur l'utilisation des subventions, consultez le guide du développeur du service de gestion des AWS clés.
- kms : DescribeKey — fournit les informations relatives aux clés gérées par le client pour permettre DataZone à Amazon de valider la clé.
- kms : GenerateDataKey — renvoie une clé de données symétrique unique à utiliser en dehors de AWS KMS.
- KMS:Decrypt : déchiffre le texte chiffré par une clé KMS.

Voici des exemples de déclarations de politique que vous pouvez ajouter pour Amazon DataZone :

```
"Statement": [
    {
      "Sid": "Enable IAM User Permissions for DescribeKey",
      "Effect": "Allow",
```

```
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
      "Action": "kms:DescribeKey",
      "Resource": "arn:aws:kms:region:111122223333:key/key_ID"
    },
    {
      "Sid": "Allow access to principals authorized to manage Amazon DataZone",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
      ],
      "Resource": "arn:aws:kms:region:111122223333:key/key_ID",
      "Condition": {
        "ForAnyValue:StringEquals": {
          "kms:EncryptionContextKeys": "aws:datazone:domainId"
        }
      }
    },
    {
      "Sid": "Allow creating grants when creating an Amazon DataZone for all principals
 in the account that are authorized to manage Amazon DataZone",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:root"
      },
      "Action": "kms:CreateGrant",
      "Resource": "arn:aws:kms:region:111122223333:key/key_ID",
      "Condition": {
        "StringLike": {
          "kms:CallerAccount": "111122223333",
          "kms:ViaService": "datazone.region.amazonaws.com"
        },
        "Bool": {
          "kms:GrantIsForAWSResource": "true"
        },
        "ForAnyValue:StringEquals": {
          "kms:EncryptionContextKeys": "aws:datazone:domainId"
        }
      }
```

```
    }
]
```

> **Note**
>
> Le portail de DataZone données Amazon a accès à votre clé gérée par le client via le principal du rôle d'exécution du domaine.

Pour plus d'informations sur la spécification des autorisations dans une politique, consultez le Guide du développeur du service de gestion des AWS clés.

Pour plus d'informations sur la résolution des problèmes d'accès par clé, consultez le Guide du développeur du service de gestion des AWS clés.

Spécifier une clé gérée par le client pour Amazon DataZone

Vous pouvez spécifier une clé gérée par le client comme deuxième couche de chiffrement lors de la création du domaine.

Contexte DataZone de chiffrement Amazon

Un contexte de chiffrement est un ensemble facultatif de paires clé-valeur qui contient des informations contextuelles supplémentaires sur les données.

AWS KMS utilise le contexte de chiffrement comme données authentifiées supplémentaires pour prendre en charge le chiffrement authentifié. Lorsque vous incluez un contexte de chiffrement dans une demande de chiffrement de données, AWS KMS lie le contexte de chiffrement aux données chiffrées. Pour déchiffrer les données, vous devez inclure le même contexte de chiffrement dans la demande.

Amazon DataZone utilise le contexte de chiffrement suivant :

```
"encryptionContextSubset": {
    "aws:datazone:domainId": "{dzd_samleid}"
}
```

Utilisation du contexte de chiffrement à des fins de surveillance : lorsque vous utilisez une clé symétrique gérée par le client pour chiffrer Amazon DataZone, vous pouvez également utiliser le contexte de chiffrement dans les dossiers d'audit et les journaux pour identifier la manière dont la clé gérée par le client est utilisée. Le contexte de chiffrement apparaît également dans les journaux générés par AWS CloudTrail ou Amazon CloudWatch Logs.

Utilisation du contexte de chiffrement pour contrôler l'accès à votre clé gérée par le client : vous pouvez utiliser le contexte de chiffrement dans les politiques clés et les politiques IAM comme conditions pour contrôler l'accès à votre clé symétrique gérée par le client. Vous pouvez également utiliser des contraintes de contexte de chiffrement dans un octroi.

Amazon DataZone utilise une contrainte de contexte de chiffrement dans les autorisations afin de contrôler l'accès à la clé gérée par le client dans votre compte ou votre région. La contrainte d'octroi exige que les opérations autorisées par l'octroi utilisent le contexte de chiffrement spécifié.

Vous trouverez ci-dessous des exemples d'instructions de stratégie de clé permettant d'accorder l'accès à une clé gérée par le client dans un contexte de chiffrement spécifique.

```
{
     "Sid": "Enable DescribeKey",
     "Effect": "Allow",
     "Principal": {
       "AWS": "arn:aws:iam::111122223333:role/ExampleRole"
     },
     "Action": "kms:DescribeKey",
     "Resource": "arn:aws:kms:region:111122223333:key/key_ID"
   },
   {
     "Sid": "Allow access to principal to manage an Amazon DataZone domain with the
given domain id",
     "Effect": "Allow",
     "Principal": {
       "AWS": "arn:aws:iam::111122223333:role/ExampleRole"
     },
     "Action": [
       "kms:Decrypt",
       "kms:GenerateDataKey"
     ],
     "Resource": "arn:aws:kms:region:111122223333:key/key_ID",
     "Condition": {
       "StringEquals": {
```

```
          "kms:EncryptionContext:aws:datazone:domainId": "dzd_sampleid"
        }
      }
    },
    {
      "Sid": "Allow creating grants when creating an Amazon DataZone domain to
principal",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::111122223333:role/ExampleRole"
      },
      "Action": "kms:CreateGrant",
      "Resource": "arn:aws:kms:region:111122223333:key/key_ID",
      "Condition": {
        "StringLike": {
          "kms:CallerAccount": "111122223333",
          "kms:ViaService": "datazone.region.amazonaws.com"
        },
        "Bool": {
          "kms:GrantIsForAWSResource": "true"
        },
        "ForAnyValue:StringEquals": {
          "kms:EncryptionContextKeys": "aws:datazone:domainId"
        }
      }
    }
```

Surveillance de vos clés de chiffrement pour Amazon DataZone

Lorsque vous utilisez une clé gérée par le client AWS KMS avec vos DataZone ressources Amazon, vous pouvez l'utiliser AWS CloudTrailpour suivre les demandes qu'Amazon DataZone envoie à AWS KMS. Les exemples suivants sont AWS CloudTrail des événements pour`CreateGrant`, `GenerateDataKeyDecrypt`, et `RetireGrant` pour surveiller les opérations KMS appelées par Amazon DataZone pour accéder aux données chiffrées par votre clé gérée par le client.

CreateGrant

Lorsque vous utilisez une clé gérée par le client AWS KMS pour chiffrer votre DataZone domaine Amazon, Amazon DataZone envoie une `CreateGrant` demande en votre nom pour accéder à la clé KMS de votre AWS compte. Les subventions créées par Amazon DataZone sont spécifiques à la ressource associée à la clé gérée par le client AWS KMS. En outre, Amazon DataZone

utilise cette `RetireGrant` opération pour supprimer une autorisation lorsque vous supprimez un domaine.

L'exemple d'événement suivant enregistre l'opération `CreateGrant` :


```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "AROAIGDTESTANDEXAMPLE:Sampleuser01",
        "arn": "arn:aws:sts::111122223333:assumed-role/Example/Sampleuser01",
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AROAIGDTESTANDEXAMPLE",
                "arn": "arn:aws:iam::111122223333:role/Example",
                "accountId": "111122223333",
                "userName": "Example"
            },
            "attributes": {
                "creationDate": "2024-04-22T17:02:00Z",
                "mfaAuthenticated": "false"
            }
        },
        "invokedBy": "datazone.amazonaws.com"
    },
    "eventTime": "2024-04-22T17:02:00Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "CreateGrant",
    "awsRegion": "us-east-2",
    "sourceIPAddress": "datazone.amazonaws.com",
    "userAgent": "datazone.amazonaws.com",
    "requestParameters": {
        "retiringPrincipal": "datazone.us-east-2.amazonaws.com",
        "operations": [
            "GenerateDataKey",
            "RetireGrant",
            "DescribeKey",
            "Decrypt"
        ],
        "granteePrincipal": "datazone.us-east-2.amazonaws.com",
```

```
        "constraints": {
            "encryptionContextSubset": {
                "aws:datazone:domainId": "dzd_sampleid"
            }
        },
        "keyId": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
    },
    "responseElements": {
        "grantId":
 "0ab0ac0d0b000f00ea00cc0a0e00fc00bce000c000f0000000c0bc0a0000aaafSAMPLE",
        "keyId": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
    },
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": false,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management",
    "sessionCredentialFromConsole": "true"
}
```

```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "AROAIGDTESTANDEXAMPLE:Sampleuser01",
        "arn": "arn:aws:sts::111122223333:assumed-role/Example/Sampleuser01",
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
        "sessionContext": {
```

```
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AROAIGDTESTANDEXAMPLE",
                "arn": "arn:aws:iam::111122223333:role/Example",
                "accountId": "111122223333",
                "userName": "Example"
            },
            "attributes": {
                "creationDate": "2024-04-22T17:10:00Z",
                "mfaAuthenticated": "false"
            }
        },
        "invokedBy": "datazone.amazonaws.com"
    },
    "eventTime": "2024-04-22T17:49:00Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "CreateGrant",
    "awsRegion": "us-east-2",
    "sourceIPAddress": "datazone.amazonaws.com",
    "userAgent": "datazone.amazonaws.com",
    "requestParameters": {
        "retiringPrincipal": "datazone.us-east-2.amazonaws.com",
        "operations": [
            "DescribeKey",
            "Decrypt"
        ],
        "granteePrincipal": "datazone.us-east-2.amazonaws.com",
        "constraints": {
            "encryptionContextSubset": {
                "aws:datazone:domainId": "dzd_sampleid"
            }
        },
        "keyId": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
    },
    "responseElements": {
        "grantId":
 "0ab0ac0d0b000f00ea00cc0a0e00fc00bce000c000f0000000c0bc0a0000aaafSAMPLE",
        "keyId": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
    },
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": false,
```

```
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management",
    "sessionCredentialFromConsole": "true"
}
```

GenerateDataKey

Lorsque vous activez une clé gérée par le client AWS KMS pour votre DataZone domaine Amazon, Amazon DataZone génère des clés de données. Il envoie une `GenerateDataKey` demande à AWS KMS qui spécifie la clé gérée par le client AWS KMS pour le domaine.

L'exemple d'événement suivant enregistre l' GenerateDataKey opération :

```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "AROAIGDTESTANDEXAMPLE:AmazonSageMakerDomainExecution",
        "arn": "arn:aws:sts::111122223333:assumed-role/
AmazonSageMakerDomainExecution/AmazonSageMakerDomainExecution",
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AROAIGDTESTANDEXAMPLE",
                "arn": "arn:aws:iam::111122223333:role/service-role/
AmazonSageMakerDomainExecution",
                "accountId": "111122223333",
                "userName": "AmazonSageMakerDomainExecution"
            },
```

```
            "attributes": {
                "creationDate": "2024-04-22T19:50:39Z",
                "mfaAuthenticated": "false"
            }
        },
        "invokedBy": "datazone.amazonaws.com"
    },
    "eventTime": "2024-04-22T19:50:40Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "GenerateDataKey",
    "awsRegion": "us-east-2",
    "sourceIPAddress": "datazone.amazonaws.com",
    "userAgent": "datazone.amazonaws.com",
    "requestParameters": {
        "keySpec": "AES_256",
        "encryptionContext": {
            "aws:datazone:domainId": "dzd_sampleid",
            "V": "2024-04-22T17:49:12.98177136Z|cacf3df7-7b99-49f6-ae14-sample",
            "version": "0",
            "N": "dzd_sampleid|arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE",
            "*aws-kms-table*": "awsdatazoneroaring-data-store-datakeys-prod-us-
east-2"
        },
        "keyId": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management"
}
```

```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AWSService",
        "invokedBy": "AWS Internal"
    },
    "eventTime": "2024-04-22T19:50:40Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "GenerateDataKey",
    "awsRegion": "us-east-2",
    "sourceIPAddress": "AWS Internal",
    "userAgent": "AWS Internal",
    "requestParameters": {
        "encryptionContext": {
            "aws:datazone:domainId": "dzd_sampleid",
            "aws:s3:arn": "arn:aws:s3:::amazon-datazone-us-
east-2-422ceee9465430bdb354d1c9efsample"
        },
        "keyId": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE",
        "keySpec": "AES_256"
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "sharedEventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventCategory": "Management"
}
```

Decrypt

Lorsque vous accédez à un DataZone domaine Amazon crypté, Amazon DataZone appelle l'`Decrypt`opération pour utiliser la clé de données cryptée stockée afin d'accéder aux données cryptées.

L'exemple d'événement suivant enregistre l'opération `Decrypt` :


```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "AROAIGDTESTANDEXAMPLE:AmazonSageMakerDomainExecution",
        "arn": "arn:aws:sts::111122223333:assumed-role/
AmazonSageMakerDomainExecution/AmazonSageMakerDomainExecution",
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE3",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AROAIGDTESTANDEXAMPLE",
                "arn": "arn:aws:iam::111122223333:role/service-role/
AmazonSageMakerDomainExecution",
                "accountId": "111122223333",
                "userName": "AmazonSageMakerDomainExecution"
            },
            "attributes": {
                "creationDate": "2024-04-22T19:50:39Z",
                "mfaAuthenticated": "false"
            }
        },
        "invokedBy": "datazone.amazonaws.com"
    },
    "eventTime": "2024-04-22T19:51:54Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "Decrypt",
    "awsRegion": "us-east-2",
    "sourceIPAddress": "datazone.amazonaws.com",
    "userAgent": "datazone.amazonaws.com",
    "requestParameters": {
```

```
        "encryptionAlgorithm": "SYMMETRIC_DEFAULT",
        "encryptionContext": {
            "aws:datazone:domainId": "dzd_sampleid",
            "V": "2024-04-22T17:49:12.98177136Z|cacf3df7-7b99-49f6-ae14-sample",
            "version": "0",
            "N": "dzd_sampleid|arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE",
            "*aws-kms-table*": "awsdatazoneroaring-data-store-datakeys-prod-us-
east-2"
        }
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management"
}
```



```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AWSService",
        "invokedBy": "datazone.amazonaws.com"
    },
    "eventTime": "2024-04-22T19:51:54Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "Decrypt",
    "awsRegion": "us-east-2",
    "sourceIPAddress": "datazone.amazonaws.com",
    "userAgent": "datazone.amazonaws.com",
```

```
    "requestParameters": {
        "encryptionContext": {
            "aws:datazone:domainId": "dzd_sampleid",
            "V": "2024-04-22T17:49:12.98177136Z|cacf3df7-7b99-49f6-ae14-sample",
            "version": "0",
            "N": "dzd_sampleid|arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE",
            "*aws-kms-table*": "awsdatazoneroaring-data-store-datakeys-prod-us-
east-2"
        },
        "encryptionAlgorithm": "SYMMETRIC_DEFAULT"
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "sharedEventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventCategory": "Management"
}
```



```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AWSService",
        "invokedBy": "AWS Internal"
    },
    "eventTime": "2024-04-22T19:51:54Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "Decrypt",
    "awsRegion": "us-east-2",
```

```
    "sourceIPAddress": "AWS Internal",
    "userAgent": "AWS Internal",
    "requestParameters": {
        "encryptionAlgorithm": "SYMMETRIC_DEFAULT",
        "encryptionContext": {
            "aws:datazone:domainId": "dzd_sampleid",
            "aws:s3:arn": "arn:aws:s3:::amazon-datazone-us-
east-2-422ceee9465430bdb354d1c9efsample"
        }
    },
    "responseElements": null,
    "requestID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "readOnly": true,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "sharedEventID": "ff000af-00eb-00ce-0e00-ea000fb0fba0SAMPLE",
    "eventCategory": "Management"
}
```

RetireGrant

L'exemple d'événement suivant enregistre l'opération `RetireGrant` :

```
{
    "eventVersion": "1.11",
    "userIdentity": {
        "type": "AWSService",
        "invokedBy": "datazone.amazonaws.com"
    },
    "eventTime": "2025-04-29T22:18:50Z",
    "eventSource": "kms.amazonaws.com",
    "eventName": "RetireGrant",
```

```
    "awsRegion": "us-east-2",
    "sourceIPAddress": "datazone.amazonaws.com",
    "userAgent": "datazone.amazonaws.com",
    "requestParameters": null,
    "responseElements": {
        "keyId": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
    },
    "additionalEventData": {
        "grantId":
 "0ab0ac0d0b000f00ea00cc0a0e00fc00bce000c000f0000000c0bc0a0000aaafSAMPLE"
    },
    "requestID": "294308c0-7617-4727-b5c9-34eaf75aa8e3",
    "eventID": "273708f7-5fbb-3a90-b04d-2b3138bf0ec9",
    "readOnly": false,
    "resources": [
        {
            "accountId": "111122223333",
            "type": "AWS::KMS::Key",
            "ARN": "arn:aws:kms:us-
east-2:111122223333:key/1234abcd-12ab-34cd-56ef-123456SAMPLE"
        }
    ],
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "sharedEventID": "b46377d7-b3c3-4bfd-a257-722bd3f3411d",
    "eventCategory": "Management"
}
```


Création d'environnements Data Lake impliquant des catalogues AWS Glue cryptés

Dans les cas d'utilisation avancés, lorsque vous travaillez avec un catalogue AWS Glue chiffré, vous devez autoriser l'accès au DataZone service Amazon pour utiliser votre clé KMS gérée par le client. Vous pouvez le faire en mettant à jour votre politique KMS personnalisée et en ajoutant une balise à la clé. Pour autoriser l'accès au DataZone service Amazon afin de travailler avec les données d'un catalogue AWS Glue crypté, procédez comme suit :

- Ajoutez la politique suivante à votre clé KMS personnalisée. Consultez Modification d'une stratégie de clé pour de plus amples informations.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "Allow datazone environment roles to decrypt using the key",
            "Effect": "Allow",
            "Principal": {
                "AWS": "*"
            },
            "Action": [
                "kms:Decrypt"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "kms:EncryptionContext:glue_catalog_id":
 "<GLUE_CATALOG_ID>"
                },
                "ArnLike": {
                    "aws:PrincipalArn": [
                        "arn:aws:iam::111122223333:role/*datazone_usr*",
                        "arn:aws:iam::444455556666:role/*datazone_usr*"
                    ]
                }
            }
        },
        {
            "Sid": "Allow datazone environment roles to describe the key",
            "Effect": "Allow",
            "Principal": {
                "AWS": "*"
            },
            "Action": [
                "kms:DescribeKey"
            ],
            "Resource": "*",
            "Condition": {
                "ArnLike": {
                    "aws:PrincipalArn": [
                    "arn:aws:iam::111122223333:role/*datazone_usr*",
                    "arn:aws:iam::444455556666:role/*datazone_usr*"
```

```
                    ]
                }
            }
        }
    ]
}
```

> **Important**
>
> - Vous devez modifier le `"aws:PrincipalArn"` ARNs contenu de la politique à l'aide du compte IDs dans lequel vous souhaitez créer les environnements. Chaque compte dans lequel vous souhaitez créer un environnement doit être répertorié dans la politique en tant que`"aws:PrincipalArn"`.
> - Vous devez également le <GLUE_CATALOG_ID>remplacer par l'identifiant de AWS compte valide dans lequel se trouve votre catalogue AWS Glue.
> - Notez que cette politique autorise l'utilisation de la clé pour tous les rôles d'utilisateur de l' DataZoneenvironnement Amazon dans le ou les comptes spécifiés. Si vous souhaitez autoriser uniquement des rôles d'utilisateur de l'environnement spécifiques à utiliser la clé, vous devez spécifier le nom complet du rôle utilisateur de l'environnement (par exemple, `arn:aws:iam::<ENVIRONMENT_ACCOUNT_ID>:role/datazone_usr_<ENVIRONMENT_ID>` (où <ENVIRONMENT_ID>est l'ID de l'environnement) plutôt que le format générique.

- Ajoutez la balise suivante à votre clé KMS personnalisée. Pour plus d'informations, consultez la section Utilisation de balises pour contrôler l'accès aux clés KMS.

```
key: AmazonDataZoneEnvironment
value: all
```

Utilisation des points de terminaison VPC d'interface pour Amazon DataZone

Si vous utilisez Amazon Virtual Private Cloud (Amazon VPC) pour héberger vos AWS ressources, vous pouvez établir une connexion entre votre Amazon VPC et Amazon. DataZone Vous pouvez utiliser cette connexion avec Amazon DataZone sans passer par l'Internet public.

Amazon VPC vous permet de lancer AWS des ressources dans un réseau virtuel personnalisé. Vous pouvez utiliser un VPC pour contrôler vos paramètres réseau, tels que la plage d'adresses IP, les sous-réseaux, les tables de routage et les passerelles réseau. Pour plus d'informations VPCs, consultez le guide de l'utilisateur Amazon VPC.

Pour connecter votre Amazon VPC à Amazon DataZone, vous devez d'abord définir un point de terminaison VPC d'interface, qui vous permet de connecter votre VPC à d'autres services. AWS Le point de terminaison assure une connectivité évolutive et fiable, sans qu'une passerelle Internet, une instance NAT (Network Address Translation) ou une connexion VPN ne soit nécessaire. Pour plus d'informations et des étapes détaillées sur la création d'un point de terminaison VPC, consultez Interface VPC Endpoints () dans le guide de AWS PrivateLink l'utilisateur Amazon VPC.

> **Important**
>
> Dans un VPC, une politique de point de terminaison est une politique basée sur les ressources que vous pouvez associer à un point de terminaison VPC pour contrôler quels AWS principaux peuvent utiliser le point de terminaison pour accéder à un service. AWS La version actuelle d'Amazon DataZone prend en charge l'utilisation de politiques relatives aux terminaux pour établir et utiliser des connexions entre votre Amazon VPC et les points de terminaison Amazon DataZone .

Autorisation sur Amazon DataZone

L'interface DataZone d'Amazon se compose d'une console de gestion intégrée à la console AWS et d'une application Web hors console (portail de données).

La console DataZone de gestion Amazon peut être utilisée par AWS les administrateurs top-level-resource APIs, notamment pour créer et gérer des domaines, des associations de AWS comptes pour ces domaines et des sources de données pour lesquelles vous souhaitez déléguer la gestion des accès à Amazon DataZone. Vous pouvez utiliser la console de DataZone gestion Amazon pour gérer tous les rôles IAM et toutes les configurations nécessaires pour déléguer le contrôle de gestion des accès au DataZone service Amazon pour leurs AWS comptes configurés de manière explicite. Le portail de DataZone données Amazon est une application de centre d' AWS identité propriétaire destinée aux utilisateurs du SSO. Si elle est activée, la console peut également être utilisée par les principaux IAM autorisés pour se fédérer dans le portail de données au lieu d'utiliser une identité SSO.

Le portail DataZone de données d'Amazon est principalement conçu pour être utilisé par les utilisateurs authentifiés par l' AWS IAM Identity Center afin de gérer l'accès aux données et d'effectuer des tâches de publication, de découverte, d'abonnement et d'analyse des données.

Autorisation dans la DataZone console Amazon

Le modèle d'autorisation de DataZone la console Amazon utilise l'autorisation IAM. La console est principalement utilisée par les administrateurs pour la configuration. Amazon DataZone utilise le concept d'un AWS compte d'administrateur de domaine et de AWS comptes de membres, et la console est utilisée à partir de tous ces comptes pour établir des relations de confiance tout en respectant les limites de AWS l'organisation.

Autorisation sur le DataZone portail Amazon

Le modèle d'autorisation du portail de DataZone données Amazon est une ACL hiérarchique avec des archétypes de rôles statiques (profils) qui incluent les administrateurs et les utilisateurs. Par exemple, les utilisateurs peuvent avoir un profil d'administrateur ou d'utilisateur. Au niveau d'un domaine, ils peuvent avoir une désignation d'utilisateur du domaine en tant que propriétaire des données. Au niveau d'un projet, un utilisateur peut être propriétaire ou contributeur. Ces profils peuvent être configurés selon l'un des deux types suivants : utilisateurs et groupes. Ces profils sont ensuite associés à des domaines et à des projets, et l'état de ces autorisations est stocké dans une table d'association.

Dans le cadre de ce modèle d'autorisation, Amazon DataZone permet aux utilisateurs de gérer les autorisations des utilisateurs et des groupes. Les utilisateurs gèrent l'adhésion aux projets, demandent l'adhésion aux projets et approuvent les adhésions. Les utilisateurs publient des données, s'abonnent aux données et approuvent les abonnements.

Les utilisateurs effectuent des analyses de données dans le cadre de projets spécifiques lorsque le client de leur portail de données demande des informations d'identification de session IAM qu'Amazon DataZone génère en fonction du profil effectif de l'utilisateur dans le contexte du projet spécifique. Cette session est limitée à la fois aux autorisations de l'utilisateur et aux ressources spécifiques du projet. Les utilisateurs se rendent ensuite sur Athena ou Redshift pour rechercher les données pertinentes, et tout le travail IAM sous-jacent est complètement abstrait.

DataZone Profils et rôles Amazon

Une fois qu'un utilisateur est authentifié, le contexte authentifié correspond à un ID de profil utilisateur. Ce profil utilisateur peut comporter plusieurs associations différentes (propriétaire du

projet, administrateur de domaine, etc.) utilisées pour autoriser les utilisateurs. Chaque association (par exemple, propriétaire du projet, administrateur de domaine, etc.) dispose d'autorisations pour certaines activités en fonction du contexte. Par exemple, un utilisateur qui possède une association d'administrateurs de domaine peut créer des domaines supplémentaires, affecter d'autres administrateurs de domaine au domaine et créer des modèles de projet au sein de son domaine. Un propriétaire de projet peut ajouter ou supprimer des membres pour son projet et publier des actifs dans un domaine.

Contrôle de l'accès aux DataZone ressources Amazon à l'aide d'IAM

Vous devez Gestion des identités et des accès AWS (IAM) effectuer les tâches liées à la sécurité suivantes :

- Créez des utilisateurs et des groupes sous votre Compte AWS.
- Attribuez des informations de sécurité uniques à chaque utilisateur situé sous votre Compte AWS.
- Contrôlez les autorisations de chaque utilisateur pour effectuer des tâches avec AWS des ressources.
- Autorisez les utilisateurs d'un autre Compte AWS pays à partager vos AWS ressources.
- Créez des rôles pour vous Compte AWS et définissez les utilisateurs ou les services qui peuvent les assumer.
- Utilisez les identités existantes pour que votre entreprise accorde des autorisations pour effectuer des tâches à l'aide de AWS ressources

Pour plus d'informations sur IAM, consultez les ressources suivantes :

- Gestion des identités et des accès AWS (JE SUIS)
- Commencer à utiliser IAM
- Guide de l'utilisateur IAM

Les sections suivantes décrivent les politiques et les autorisations requises pour configurer Amazon DataZone et ses composants, tels que les domaines (y compris le domaine), les comptes associés, les projets et les sources de données. Pour de plus amples informations, veuillez consulter DataZone Terminologie et concepts d'Amazon.

Table des matières

- AWS politiques gérées pour Amazon DataZone
- Rôles IAM pour Amazon DataZone
- Informations d'identification temporaires
- Autorisations de principal

AWS politiques gérées pour Amazon DataZone

Une politique AWS gérée est une politique autonome créée et administrée par AWS. AWS les politiques gérées sont conçues pour fournir des autorisations pour de nombreux cas d'utilisation courants afin que vous puissiez commencer à attribuer des autorisations aux utilisateurs, aux groupes et aux rôles.

N'oubliez pas que les politiques AWS gérées peuvent ne pas accorder d'autorisations de moindre privilège pour vos cas d'utilisation spécifiques, car elles sont accessibles à tous les AWS clients. Nous vous recommandons de réduire encore les autorisations en définissant des politiques gérées par le client qui sont propres à vos cas d'utilisation.

Vous ne pouvez pas modifier les autorisations définies dans les politiques AWS gérées. Si les autorisations définies dans une politique AWS gérée sont AWS mises à jour, la mise à jour affecte toutes les identités principales (utilisateurs, groupes et rôles) auxquelles la politique est attachée. AWS est le plus susceptible de mettre à jour une politique AWS gérée lorsqu'une nouvelle politique Service AWS est lancée ou lorsque de nouvelles opérations d'API sont disponibles pour les services existants.

Pour plus d'informations, consultez Politiques gérées par AWS dans le Guide de l'utilisateur IAM.

Table des matières

- AWS politique gérée : AmazonDataZoneFullAccess
- AWS politique gérée : AmazonDataZoneFullUserAccess
- AWS politique gérée : AmazonDataZoneEnvironmentRolePermissionsBoundary
- AWS politique gérée : AmazonDataZoneRedshiftGlueProvisioningPolicy
- AWS politique gérée : AmazonDataZoneGlueManageAccessRolePolicy
- AWS politique gérée : AmazonDataZoneRedshiftManageAccessRolePolicy
- AWS politique gérée : AmazonDataZoneDomainExecutionRolePolicy

- AWS politique gérée : AmazonDataZoneSageMakerProvisioningRolePolicy
- AWS politique gérée : AmazonDataZoneSageMakerManageAccessRolePolicy
- AWS politique gérée : AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary
- Amazon DataZone met à jour AWS ses politiques gérées

AWS politique gérée : AmazonDataZoneFullAccess

Vous pouvez associer la politique `AmazonDataZoneFullAccess` à vos identités IAM.

Cette politique fournit un accès complet à Amazon DataZone via le AWS Management Console. Cette politique donne également des autorisations à AWS KMS pour les paramètres SSM chiffrés. La clé KMS doit être étiquetée EnableKeyForAmazonDataZone pour permettre le déchiffrement des paramètres SSM.

Détails de l'autorisation

Cette politique inclut les autorisations suivantes :

- `datazone`— accorde aux principaux un accès complet à Amazon DataZone via le AWS Management Console.
- `kms`— Permet aux principaux de répertorier les alias, de décrire les clés et de déchiffrer les clés.
- `s3`— Permet aux principaux de choisir des compartiments S3 existants ou d'en créer de nouveaux pour stocker les données Amazon DataZone .
- `ram`— Permet aux principaux de partager des DataZone domaines Comptes AWS Amazon entre eux.
- `iam`— Permet aux directeurs de répertorier et de transmettre des rôles et d'obtenir des politiques.
- `sso`— Permet aux principaux d'obtenir les régions où cette option AWS IAM Identity Center est activée.
- `secretsmanager`— Permet aux principaux de créer, de baliser et de répertorier des secrets avec un préfixe spécifique.
- `aoss`— Permet aux principaux de créer et de récupérer des informations pour les politiques de sécurité OpenSearch sans serveur.
- `bedrock`— Permet aux principaux de créer, de répertorier et de récupérer des informations pour les profils d'inférence et les modèles de base.

- `codeconnections`— Permet aux principaux de supprimer, de récupérer des informations, de répertorier les connexions et de gérer les balises associées aux connexions.
- `codewhisperer`— Permet aux principaux de répertorier les CodeWhisperer profils.
- `ssm`— Permet aux principaux de saisir, de supprimer et de récupérer des informations pour les paramètres.
- `redshift`— Permet aux principaux de décrire les clusters et de répertorier les groupes de travail sans serveur
- `glue`— Permet aux principaux d'obtenir des bases de données.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneFullAccess dans le AWS Guide de référence des stratégies gérées par.

Considérations et limites relatives aux politiques

Certaines fonctionnalités ne sont pas couvertes par la `AmazonDataZoneFullAccess` politique.

- Si vous créez un DataZone domaine Amazon avec votre propre AWS KMS clé, vous devez disposer des autorisations nécessaires `kms:CreateGrant` pour que la création du domaine réussisse, et `kms:Decrypt` pour que cette clé puisse invoquer d'autres Amazon`kms:GenerateDataKey`, DataZone APIs tels que `listDataSources` et`createDataSource`. Et vous devez également disposer des autorisations nécessaires pour `kms:CreateGrant`kms:Decrypt,`kms:GenerateDataKey`, et `kms:DescribeKey` dans la politique de ressources de cette clé.

  Si vous utilisez la clé KMS appartenant au service par défaut, cela n'est pas obligatoire.

  Pour de plus amples informations, veuillez consulter AWS Key Management Service.
- Si vous souhaitez utiliser les fonctionnalités de création et de mise à jour de rôles dans la DataZone console Amazon, vous devez disposer des privilèges d'administrateur ou des autorisations IAM requises pour créer des rôles IAM et créer/mettre à jour des politiques. Les autorisations requises incluent`iam:CreateRole`,`iam:CreatePolicy`, `iam:CreatePolicyVersion`iam:DeletePolicyVersion, et `iam:AttachRolePolicy` les autorisations.
- Si vous créez un nouveau domaine sur Amazon DataZone avec l'identifiant AWS IAM Identity Center utilisateur activé, ou si vous l'activez pour un domaine existant sur Amazon DataZone, vous devez disposer des autorisations suivantes :

- organisations : DescribeOrganization
- organisations : ListDelegatedAdministrators
- SSO : CreateInstance
- SSO : ListInstances
- SSO : GetSharedSsoConfiguration
- SSO : PutApplicationGrant
- SSO : PutApplicationAssignmentConfiguration
- SSO : PutApplicationAuthenticationMethod
- SSO : PutApplicationAccessScope
- SSO : CreateApplication
- SSO : DeleteApplication
- SSO : CreateApplicationAssignment
- SSO : DeleteApplicationAssignment
- répertoire SSO : CreateUser
- répertoire SSO : SearchUsers
- SSO : ListApplications

- Pour accepter une demande d'association de AWS compte sur Amazon DataZone, vous devez en avoir l'`ram:AcceptResourceShareInvitation`autorisation.
- Si vous souhaitez créer la ressource requise pour la configuration réseau d' SageMaker Unified Studio, vous devez disposer des autorisations suivantes et associer une AmazonVpcFullAccess politique :
  - iam : PassRole
  - formation des nuages : CreateStack

AWS politique gérée : AmazonDataZoneFullUserAccess

Cette politique accorde un accès complet à Amazon DataZone, mais elle n'autorise pas la gestion des domaines, des utilisateurs ou des comptes associés.

Détails de l'autorisation

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneFullUserAccess dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneEnvironmentRolePermissionsBoundary

> **Note**
>
> Cette politique est une limite des autorisations. Une limite des autorisations définit le nombre maximum d'autorisations qu'une politique basée sur l'identité peut accorder à une entité IAM. Vous ne devez pas utiliser et joindre vous-même les politiques relatives aux limites des DataZone autorisations d'Amazon. Les politiques DataZone relatives aux limites des autorisations Amazon ne doivent être associées qu'aux rôles DataZone gérés par Amazon. Pour plus d'informations sur les limites d'autorisations, consultez la section Limites d'autorisations pour les entités IAM dans le guide de l'utilisateur IAM.

Lorsque vous créez un environnement via le portail de DataZone données Amazon, Amazon DataZone applique cette limite d'autorisations aux rôles IAM produits lors de la création de l'environnement. La limite des autorisations limite l'étendue des rôles créés par Amazon DataZone et de tous les rôles que vous ajoutez.

Amazon DataZone utilise la politique `AmazonDataZoneEnvironmentRolePermissionsBoundary` gérée pour limiter le principal IAM provisionné auquel elle est attachée. Les principes peuvent prendre la forme des rôles d'utilisateur qu'Amazon DataZone peut assumer pour le compte des utilisateurs d'entreprise interactifs ou des services d'analyse (par exemple)AWS Glue, puis effectuer des actions pour traiter des données telles que la lecture et l'écriture depuis Amazon S3 ou l'exécution. AWS Glue crawler

La `AmazonDataZoneEnvironmentRolePermissionsBoundary` politique accorde DataZone à Amazon un accès en lecture et en écriture à des services tels qu' AWS Glue Amazon S3 AWS Lake Formation, Amazon Redshift et Amazon Athena. La politique accorde également des autorisations de lecture et d'écriture à certaines ressources d'infrastructure requises pour utiliser ces services, telles que les interfaces réseau et AWS KMS les clés.

Amazon DataZone applique la politique `AmazonDataZoneEnvironmentRolePermissionsBoundary` AWS gérée en tant que limite d'autorisations pour tous les rôles de DataZone l'environnement Amazon (propriétaire et contributeur). Cette limite d'autorisations restreint ces rôles afin de n'autoriser l'accès qu'aux ressources requises et aux actions nécessaires à un environnement.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneEnvironmentRolePermissionsBoundary dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneRedshiftGlueProvisioningPolicy

La AmazonDataZoneRedshiftGlueProvisioningPolicy politique accorde à Amazon DataZone les autorisations nécessaires pour interagir avec AWS Glue et Amazon Redshift.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneRedshiftGlueProvisioningPolicy dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneGlueManageAccessRolePolicy

Cette politique autorise Amazon DataZone à publier les données AWS Glue dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux ressources publiées par AWS Glue dans le catalogue.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneGlueManageAccessRolePolicy dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneRedshiftManageAccessRolePolicy

Cette politique autorise Amazon DataZone à publier les données Amazon Redshift dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux ressources publiées par Amazon Redshift ou Amazon Redshift Serverless dans le catalogue.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneRedshiftManageAccessRolePolicy dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneDomainExecutionRolePolicy

Il s'agit de la politique par défaut pour le rôle DataZone `DomainExecutionRole` de service Amazon. Ce rôle est utilisé par Amazon DataZone pour cataloguer, découvrir, gérer, partager et analyser les données du DataZone domaine Amazon. Ce rôle permet d'accéder à tous les Amazon DataZone APIs nécessaires à l'utilisation du portail de données, ainsi que des autorisations de RAM pour prendre en charge l'utilisation des comptes associés dans un DataZone domaine Amazon.

Vous pouvez joindre la AmazonDataZoneDomainExecutionRolePolicy politique à votre`AmazonDataZoneDomainExecutionRole`.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneDomainExecutionRolePolicy dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneSageMakerProvisioningRolePolicy

La AmazonDataZoneSageMakerProvisioningRolePolicy politique accorde à Amazon DataZone les autorisations nécessaires pour interagir avec Amazon SageMaker.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneSageMakerProvisioningRolePolicy dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneSageMakerManageAccessRolePolicy

Cette politique autorise Amazon DataZone à publier des SageMaker actifs Amazon dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux actifs SageMaker publiés par Amazon dans le catalogue.

Cette politique inclut les autorisations pour effectuer les opérations suivantes :

- cloudtrail — récupère des informations sur les CloudTrail sentiers.
- cloudwatch — récupère les CloudWatch alarmes en cours.
- logs — récupère les filtres métriques pour les CloudWatch journaux.
- sns — récupère la liste des abonnements à une rubrique SNS.
- config — récupère des informations sur les enregistreurs de configuration, les ressources et les règles de AWS configuration. Permet également au rôle lié au service de créer et de supprimer des règles de AWS configuration, et d'exécuter des évaluations par rapport à ces règles.
- iam — Obtenir et générer des rapports d'identification pour les comptes.
- organisations : récupérer les informations relatives au compte et à l'unité organisationnelle (UO) d'une organisation.
- securityhub : récupère des informations sur la manière dont le service, les normes et les contrôles Security Hub sont configurés.
- tag — récupère des informations sur les balises de ressources.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneSageMakerManageAccessRolePolicy dans le AWS Guide de référence des stratégies gérées par.

AWS politique gérée : AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary

> **Note**
>
> Cette politique est une limite des autorisations. Une limite des autorisations définit le nombre maximum d'autorisations qu'une politique basée sur l'identité peut accorder à une entité IAM. Vous ne devez pas utiliser et joindre vous-même les politiques relatives aux limites des DataZone autorisations d'Amazon. Les politiques DataZone relatives aux limites des autorisations Amazon ne doivent être associées qu'aux rôles DataZone gérés par Amazon. Pour plus d'informations sur les limites d'autorisations, consultez la section Limites d'autorisations pour les entités IAM dans le guide de l'utilisateur IAM.

Lorsque vous créez un SageMaker environnement Amazon via le portail de DataZone données Amazon, Amazon DataZone applique cette limite d'autorisations aux rôles IAM produits lors de la création de l'environnement. La limite des autorisations limite l'étendue des rôles créés par Amazon DataZone et de tous les rôles que vous ajoutez.

Amazon DataZone utilise la politique `AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary` gérée pour limiter le principal IAM provisionné auquel elle est attachée. Les principes peuvent prendre la forme des rôles d'utilisateur qu'Amazon DataZone peut assumer pour le compte des utilisateurs d'entreprise interactifs ou des services d'analyse (par exemple)AWS SageMaker, puis effectuer des actions pour traiter des données telles que la lecture et l'écriture depuis Amazon S3 ou Amazon Redshift ou l'exécution AWS du robot d'exploration Glue.

Cette `AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary` politique accorde DataZone à Amazon un accès en lecture et en écriture à des services tels qu'Amazon SageMaker, AWS Glue, Amazon S3, AWS Lake Formation, Amazon Redshift et Amazon Athena. La politique accorde également des autorisations de lecture et d'écriture à certaines ressources d'infrastructure requises pour utiliser ces services, telles que les interfaces réseau, les référentiels Amazon ECR et les clés AWS KMS. Il donne également accès à des SageMaker applications Amazon comme Amazon SageMaker Canvas.

Amazon DataZone applique la politique `AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary` gérée en tant que limite d'autorisations pour tous les rôles de DataZone l'environnement Amazon (propriétaire et contributeur). Cette limite d'autorisations restreint ces rôles afin de n'autoriser l'accès qu'aux ressources requises et aux actions nécessaires à un environnement.

Pour voir les autorisations de cette stratégie, consultez AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary dans le AWS Guide de référence des stratégies gérées par.

Amazon DataZone met à jour AWS ses politiques gérées

Consultez les informations relatives aux mises à jour des politiques AWS gérées pour Amazon DataZone depuis que ce service a commencé à suivre ces modifications. Pour recevoir des alertes automatiques concernant les modifications apportées à cette page, abonnez-vous au flux RSS sur la page d'historique des DataZone documents Amazon.

Modifier	Description	Date
AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary - mises à jour des politiques	Mises à jour des politiques du AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary. Ajout d'une instruction Deny pour l'`sagemaker:UpdateNotebookInstanceLifecycleConfig` action visant à restreindre cette opération à privilèges élevés.	11 mars 2026
AmazonDataZoneDomainExecutionRolePolicy - mises à jour des politiques	Mises à jour des politiques relatives à l'AmazonDataZoneDomainExecutionRolePolicyajout d'autorisations pour l'`QueryGraph` action afin de prendre en	25 février 2026

Modifier	Description	Date
	charge les fonctionnalités de recherche d'entités basées sur des graphes.	
AmazonDataZoneGlue ManageAccessRolePolicy - mises à jour des politiques	Mises à jour de la AmazonDataZoneGlueManageAccessRolePolicypolitique concernant l'ajout d'autorisations à l'`GetConnection` action afin de prendre en charge la capture du lignage des données pour les sources de données de AWS Glue basées sur la connexion.	30 juillet 2025
AmazonDataZoneFullAccess - mises à jour des politiques	Mises à jour des AmazonDataZoneFullAccesspolitiques relatives à la généralisation de la portée SecretsManager `create` et `tag` des autorisations pour les nouveaux domaines qui auront le format de `dzd-` au lieu de`dzd_...`	23 juillet 2025
AmazonDataZoneFullAccess - mises à jour des politiques	Mises à jour des AmazonDataZoneFullAccesspolitiques visant à permettre à la console d'associer ou de mettre à jour les autorisations AWS gérées dans les partages de ressources AWS RAM.	22 mai 2025

Modifier	Description	Date
AmazonDataZoneGlue ManageAccessRolePolicy - mises à jour des politiques	Mises à jour de la politique AmazonDat aZoneGlueManageAcc essRolePolicy: le rôle d'utilisa teur DataZone du projet Amazon est utilisé comme rôle de transfert de données pour les tables fédérées. Cette mise `datazone_ usr_role*` à jour complète la `iam:PassR ole` déclaration, permettant d'utiliser le rôle d'utilisateur du projet à cette fin.	21 mai 2025
AmazonDataZoneSage MakerProvisioningRolePolicy - mises à jour des politiques	Mises à jour de la politique AmazonDataZoneSage MakerProvisioningRolePolicy - ajout d'un support à l'`glue:GetConnection` action.	2 janvier 2025
AmazonDataZoneSage MakerEnvironmentRo lePermissionsBoundary - mises à jour des politiques	Mises à jour de la politique AmazonDat aZoneSageMakerEnvi ronmentRolePermiss ionsBoundary: cette modificat ion ajoute la limite `sagemaker :AddTags` d'autorisation afin de permettre DataZone à Amazon d'appeler `CreateUse rProfile` avec succès avec les balises nécessaires.	3 décembre 2024

Modifier	Description	Date
AmazonDataZoneSageMakerAccess, et AmazonDataZoneGlueManageAccessRolePolicy - mises à jour des politiques	Mises à jour des politiques relatives à AmazonDataZoneFullAccessAmazonDataZoneSageMakerAccess, et AmazonDataZoneGlueManageAccessRolePolicy- pour permettre la prise en charge de l'expérience Amazon SageMaker Unified Studio.	3 décembre 2024
AmazonDataZoneDomainExecutionRolePolicy et AmazonDataZoneFullUserAccess - mises à jour des politiques	Mises à jour des politiques relatives à AmazonDataZoneDomainExecutionRolePolicyet AmazonDataZoneFullUserAccess- pour permettre la prise en charge des règles d'application des métadonnées pour les demandes d'abonnement.	19 novembre 2024

Modifier	Description	Date
AmazonDataZoneRedshiftGlueProvisioningPolicy - mises à jour des politiques	Mises à jour des politiques du AmazonDataZoneRedshiftGlueProvisioningPolicy- `iam:DeletePolicyVersion` vers Ajouter pour permettre aux utilisateurs de supprimer les versions des politiques créées avec`datazone*` . Cela permet de débloquer les utilisateurs qui doivent mettre à jour leur politique de rôle d'utilisateur dans leur environnement.	22 octobre 2024
AmazonDataZoneDomainExecutionRolePolicy et AmazonDataZoneFullUserAccess - mises à jour des politiques	Mises à jour des politiques relatives aux AmazonDataZoneDomainExecutionRolePolicyet AmazonDataZoneFullUserAccess- pour permettre la prise en charge APIs des nouveaux outils utilisés pour créer et gérer des unités de DataZone domaine et des produits de données Amazon.	31 juillet 2024

Modifier	Description	Date
AmazonDataZoneGlue ManageAccessRolePolicy - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneGlue ManageAccessRolePolicy- Amazon DataZone ajoute des autorisations IAM qui sont utilisées pour des fonctionn alités de contrôle d'accès précises afin de limiter l'octroi d'autorisations à Lake Formation.	2 juillet 2024
AmazonDataZoneExec utionRolePolicy et AmazonDat aZoneFullUserAccess - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneExec utionRolePolicy et AmazonDat aZoneFullUserAccesspour permettre la prise en charge du lignage des données et du contrôle d'accès détaillé. APIs	27 juin 2024

Modifier	Description	Date
AmazonDataZoneGlue ManageAccessRolePolicy - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneGlue ManageAccessRolePolicyqui ajoute les autorisations IAM requises pour la fonctionn alité d'auto-abonnement d'Amazon DataZone afin de limiter les autorisations accordées lors de la formation des lacs. Grâce à la fonctionn alité d'auto-abonnement, les autorisations de formation de lacs ne peuvent être accordées qu'aux ressources étiquetées.	14 juin 2024
AmazonDataZoneDoma inExecutionRolePolicy - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneDoma inExecutionRolePolicyqui ajoute de nouvelles APIs fonctionnalités à Amazon DataZone qui permettent aux utilisateurs de configurer des actions pour leurs DataZone environnements Amazon.	14 juin 2024

Modifier	Description	Date
AmazonDataZoneFullAccess - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneFull Accessqui permet à la console DataZone de gestion Amazon de créer des secrets au nom de l'utilisateur avec des balises de domaine et de projet. Inclut également l'`ram:ListResourceSharePermissions` action permettant aux administrateurs du compte du propriétaire du domaine de consulter le statut d'association des comptes associés.	14 juin 2024
AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary - nouvelle limite d'autorisations	Nouvelle limite d'autorisations appelée AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary. Lorsque vous créez un SageMaker environnement Amazon via le portail de DataZone données Amazon, Amazon DataZone applique cette limite d'autorisations aux rôles IAM produits lors de la création de l'environnement. La limite des autorisations limite l'étendue des rôles créés par Amazon DataZone et de tous les rôles que vous ajoutez.	30 avril 2024

Modifier	Description	Date
AmazonDataZoneSageMakerAccess - nouvelle politique	La nouvelle politique appelée AmazonDataZoneSageMakerAccessdonne à Amazon l' DataZone autorisation de publier SageMaker les actifs Amazon dans le catalogue . Cela donne également à Amazon l' DataZoneautorisation d'accorder ou de révoquer l'accès aux actifs SageMaker publiés par Amazon dans le catalogue.	30 avril 2024
AmazonDataZoneFullAccess - mise à jour de la politique	Une mise à jour de la AmazonDataZoneFullAccesspolitique qui ajoute un accès à l'`DescribeSecurityGroups` action pour améliorer la facilité d'utilisation pour les administrateurs de compte qui configurent des plans dans la console et des `GetPolicy` actions pour aider à récupérer des informations sur la politique gérée spécifiée.	30 avril 2024

Modifier	Description	Date
AmazonDataZoneSage MakerProvisioningRolePolicy - nouvelle politique	La nouvelle politique appelée AmazonDataZoneSage MakerProvisioningRolePolicy accorde à Amazon DataZone les autorisations nécessaires pour interagir avec Amazon SageMaker.	30 avril 2024
AmazonDataZoneS3Ma nage- <region>- <domainId>- nouveau rôle	Nouveau rôle appelé AmazonDataZoneS3Ma nage-<region>, <domainId >utilisé lorsqu'Amazon DataZone appelle AWS Lake Formation pour enregistrer un emplacement Amazon Simple Storage Service (Amazon S3). AWS Lake Formation assume ce rôle lors de l'accès aux données à cet emplacement.	1er avril 2024
AmazonDataZoneGlue ManageAccessRolePolicy - Mise à jour de la politique	Mise à jour du AmazonDat aZoneGlueManageAcc essRolePolicypour permettre la prise en charge des autorisations permettant DataZone à Amazon d'activer les autorisations de publication et d'accès aux données.	1er avril 2024

Modifier	Description	Date
AmazonDataZoneDomainExecutionRolePolicy et AmazonDataZoneFullUserAccess - Mise à jour de la politique	Mise à jour du AmazonDataZoneDomainExecutionRolePolicyet AmazonDataZoneFullUserAccesspour permettre la prise en charge de l'`CancelMetadataGenerationRun` API.	29 mars 2024
AmazonDataZoneFullAccess - Mise à jour de la politique	Mise à jour du `AmazonDataZoneFullAccess` pour permettre aux utilisateurs de choisir leurs secrets, leurs clusters, leurs VPC et leurs sous-réseaux dans la console de DataZone gestion Amazon plutôt que de les saisir dans une zone de texte.	13 mars 2024
AmazonDataZoneDomainExecutionRolePolicy - Mise à jour de la politique	Mise à jour du AmazonDataZoneDomainExecutionRolePolicypour permettre la prise en charge de l'`ListEnvironmentBlueprintConfigurationSummaries` API requise pour créer des profils d'environnement en identifiant quels plans sont activés dans quel compte et dans quelle région.	01 février 2024

Modifier	Description	Date
AmazonDataZoneGlue ManageAccessRolePolicy - Mise à jour de la politique	Mise à jour du AmazonDat aZoneGlueManageAcc essRolePolicypour permettre la prise en charge du mode hybride AWS Lake Formation.	14 décembre 2023
AmazonDataZoneFull UserAccess et AmazonDat aZoneDomainExecuti onRolePolicy - Mises à jour des politiques	Mise à jour des politique s AmazonDataZoneFull UserAccesset des AmazonDat aZoneDomainExecuti onRolePolicypolitiques afin de prendre en charge la fonctionnalité de descripti on des données basée sur l'IA générative sur Amazon. DataZone	28 novembre 2023
AmazonDataZoneEnvi ronmentRolePermiss ionsBoundary - Mise à jour de la politique	Amazon DataZone a mis à jour la politique AmazonDat aZoneEnvironmentRo lePermissionsBound arygérée qui consiste en une `athena:GetQueryRes ultsStream` autorisation supplémentaire limitée à la `ResourceTag` condition.	17 novembre 2023

Modifier	Description	Date
AmazonDataZoneRedshiftManageAccessRolePolicy - Mise à jour de la politique	Amazon a DataZone mis à jour le AmazonDataZoneRedshiftManageAccessRolePolicyen supprimant la vérification de l'identifiant de l'organisation pour l'`redshift:AssociateDataShareConsumer` action. Cela vous permet de partager les ressources entre les AWS organisations.	16 novembre 2023
AmazonDataZoneFullUserAccess - Mise à jour de la politique	Amazon a DataZone mis à jour la AmazonDataZoneFullUserAccesspolitique qui accorde un accès complet à Amazon DataZone, mais elle n'autorise pas la gestion des domaines, des utilisateurs ou des comptes associés.	2 octobre 2023
AmazonDataZonePortalFullAccessPolicy - politique déconseillée	Amazon DataZone a déconseillé le. AmazonDataZonePortalFullAccessPolicy	29 septembre 2023
AmazonDataZonePreviewConsoleFullAccess - politique déconseillée	Amazon DataZone a déconseillé le. AmazonDataZonePreviewConsoleFullAccess	29 septembre 2023

Modifier	Description	Date
AmazonDataZoneDomainExecutionRolePolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneDomainExecutionRolePolicy. Il s'agit de la politique par défaut pour le rôle DataZone `AmazonDataZoneDomainExecutionRole` de service Amazon. Ce rôle est utilisé par Amazon DataZone pour cataloguer, découvrir, gérer, partager et analyser les données du DataZone domaine Amazon. Vous pouvez joindre la `AmazonDataZoneDomainExecutionRolePolicy` politique à votre`AmazonDataZoneDomainExecutionRole` .	25 septembre 2023
AmazonDataZoneCrossAccountAdmin - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneCrossAccountAdminqui permet aux utilisateurs de travailler avec Amazon DataZone et ses comptes associés.	19 septembre 2023

Modifier	Description	Date
AmazonDataZoneFull UserAccess - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneFull UserAccessqui accorde un accès complet à Amazon DataZone, mais elle n'autorise pas la gestion des domaines, des utilisateurs ou des comptes associés.	12 septembre 2023
AmazonDataZoneReds hiftManageAccessRolePolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneReds hiftManageAccessRo lePolicyqui accorde des autorisations permettant DataZone à Amazon d'activer les autorisations de publication et d'accès aux données.	12 septembre 2023

Modifier	Description	Date
AmazonDataZoneGlue ManageAccessRolePolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique AmazonDataZoneGlue ManageAccessRolePo licyqui accorde à Amazon l' DataZone autorisation de publier les données AWS Glue dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux ressources publiées par AWS Glue dans le catalogue.	12 septembre 2023
AmazonDataZoneReds hiftGlueProvisioningPolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneReds hiftGlueProvisioningPolicyqui accorde à Amazon DataZone les autorisations nécessaires pour interagir avec les sources de données prises en charge.	12 septembre 2023
AmazonDataZoneEnvi ronmentRolePermiss ionsBoundary - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneEnvi ronmentRolePermiss ionsBoundaryqui limite le principal IAM provisionné auquel il est attaché.	12 septembre 2023

Modifier	Description	Date
AmazonDataZoneFullAccess - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneFull Accessqui fournit un accès complet à Amazon DataZone via la console AWS de gestion.	12 septembre 2023
Mise à jour de politique gérée	Mises à jour de la politique AmazonDataZonePrev iewConsoleFullAcce ssgérée qui consiste en des `iam:GetPolicy` autorisat ions supplémentaires.	13 juin 2023
Amazon DataZone a commencé à suivre les modifications	Amazon DataZone a commencé à suivre les modifications apportées AWS à ses politiques gérées.	20 mars 2023

Rôles IAM pour Amazon DataZone

Rubriques

- AmazonDataZoneProvisioningRole-<domainAccountId>
- AmazonDataZoneDomainExecutionRole
- AmazonDataZoneGlueAccess- <region>- <domainId>
- AmazonDataZoneRedshiftAccess- <region>- <domainId>
- AmazonDataZone<region>S3 Manage- - <domainId>
- AmazonDataZoneSageMakerManageAccessRole- <region>- <domainId>
- AmazonDataZoneSageMakerProvisioningRolePolicyRole-<domainAccountId>

AmazonDataZoneProvisioningRole-<domainAccountId>

`AmazonDataZoneProvisioningRole-<domainAccountId>`Il a la `AmazonDataZoneRedshiftGlueProvisioningPolicy` pièce jointe. Ce rôle accorde à Amazon DataZone les autorisations nécessaires pour interagir avec AWS Glue et Amazon Redshift.

La politique de confiance suivante `AmazonDataZoneProvisioningRole-<domainAccountId>` est attachée à la valeur par défaut :

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "datazone.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
            "aws:SourceAccount": "{{domain_account}}"
        }
      }
    }
  ]
}
```

AmazonDataZoneDomainExecutionRole

AmazonDataZoneDomainExecutionRoleLa politique AWS gérée est AmazonDataZoneDomainExecutionRolePolicyjointe. Amazon DataZone crée ce rôle pour vous en votre nom. Pour certaines actions sur le portail de données, Amazon DataZone assume ce rôle dans le compte dans lequel le rôle est créé et vérifie que ce rôle est autorisé à effectuer l'action.

Le AmazonDataZoneDomainExecutionRolerôle est obligatoire dans celui Compte AWS qui héberge votre DataZone domaine Amazon. Ce rôle est automatiquement créé pour vous lorsque vous créez votre DataZone domaine Amazon.

Le AmazonDataZoneDomainExecutionRolerôle par défaut est soumis à la politique de confiance suivante.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "datazone.amazonaws.com"
            },
            "Action": [
                "sts:AssumeRole",
                "sts:TagSession"
            ],
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "{{source_account_id}}"
                },
                "ForAllValues:StringLike": {
                    "aws:TagKeys": [
                        "datazone*"
                    ]
                }
            }
        }
    ]
}
```

AmazonDataZoneGlueAccess- <region>- <domainId>

Le `AmazonDataZoneGlueAccess-<region>-<domainId>` rôle comporte ce qui est `AmazonDataZoneGlueManageAccessRolePolicy` joint. Ce rôle autorise Amazon DataZone à

publier les données AWS Glue dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux ressources publiées par AWS Glue dans le catalogue.

Le `AmazonDataZoneGlueAccess-<region>-<domainId>` rôle par défaut est associé à la politique de confiance suivante :

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "datazone.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
            "StringEquals": {
            "aws:SourceAccount": "111122223333"
            },
            "ArnEquals": {
            "aws:SourceArn": "arn:aws:datazone:us-east-1:111122223333:domain/
dzd-12345"
            }
        }
    }
  ]
}
```

AmazonDataZoneRedshiftAccess- <region>- <domainId>

Le `AmazonDataZoneRedshiftAccess-<region>-<domainId>` rôle comporte ce qui est `AmazonDataZoneRedshiftManageAccessRolePolicy` joint. Ce rôle autorise Amazon DataZone à publier les données Amazon Redshift dans le catalogue. Cela donne également à Amazon l' DataZone autorisation d'accorder ou de révoquer l'accès aux ressources publiées par Amazon Redshift ou Amazon Redshift Serverless dans le catalogue.

Le AmazonDataZoneRedshiftAccess-<region>-<domainId> rôle par défaut est associé à la politique d'autorisation intégrée suivante :

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Sid": "RedshiftSecretStatement",
         "Effect":"Allow",
         "Action":"secretsmanager:GetSecretValue",
         "Resource":"*",
         "Condition":{
            "StringEquals":{
               "secretsmanager:ResourceTag/AmazonDataZoneDomain":"{{domainId}}"
            }
         }
      }
   ]
}
```

La politique de confiance suivante AmazonDataZoneRedshiftManageAccessRole<timestamp> est attachée à la valeur par défaut :

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "datazone.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
```

```
      "Condition": {
            "StringEquals": {
            "aws:SourceAccount": "111122223333"
            },
            "ArnEquals": {
            "aws:SourceArn": "arn:aws:datazone:us-east-1:111122223333:domain/
dzd-12345"
            }
        }
    }
  ]
}
```

AmazonDataZone<region>S3 Manage- - <domainId>

Le AmazonDataZone S3Manage- <region>- <domainId>est utilisé lorsqu'Amazon DataZone appelle AWS Lake Formation pour enregistrer un site Amazon Simple Storage Service (Amazon S3). AWS Lake Formation assume ce rôle lors de l'accès aux données à cet emplacement. Pour plus d'informations, consultez la section Exigences relatives aux rôles utilisés pour enregistrer des sites.

La politique d'autorisation intégrée suivante est jointe à ce rôle.

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "LakeFormationDataAccessPermissionsForS3",
            "Effect": "Allow",
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
                "s3:DeleteObject"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceAccount": "{{accountId}}"
```

```
                }
            }
        },
        {
            "Sid": "LakeFormationDataAccessPermissionsForS3ListBucket",
            "Effect": "Allow",
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceAccount": "{{accountId}}"
                }
            }
        },
        {
            "Sid": "LakeFormationDataAccessPermissionsForS3ListAllMyBuckets",
            "Effect": "Allow",
            "Action": [
                "s3:ListAllMyBuckets"
            ],
            "Resource": "arn:aws:s3:::*",
            "Condition": {
                "StringEquals": {
                    "aws:ResourceAccount": "{{accountId}}"
                }
            }
        },
        {
            "Sid": "LakeFormationExplicitDenyPermissionsForS3",
            "Effect": "Deny",
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
                "s3:DeleteObject"
            ],
            "Resource": [
                "arn:aws:s3:::[[BucketNames]]/*"
            ],
            "Condition": {
                "StringEquals": {
                    "aws:ResourceAccount": "{{accountId}}"
                }
```

```
            }
        },
        {
            "Sid": "LakeFormationExplicitDenyPermissionsForS3ListBucket",
            "Effect": "Deny",
            "Action": [
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::[[BucketNames]]"
            ],
            "Condition": {
                "StringEquals": {
                    "aws:ResourceAccount": "{{accountId}}"
                }
            }
        }
    ]
}
```

Le AmazonDataZone S3Manage- <region>- <domainId>est associé à la politique de confiance suivante :

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "TrustLakeFormationForDataLocationRegistration",
            "Effect": "Allow",
            "Principal": {
                "Service": "lakeformation.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "{{source_account_id}}"
                }
```

```
            }
        }
    ]
}
```

AmazonDataZoneSageMakerManageAccessRole- <region>- <domainId>

Le `AmazonDataZoneSageMakerManageAccessRole` rôle a le`AmazonDataZoneSageMakerAccess`, le`AmazonDataZoneRedshiftManageAccessRolePolicy`, et ce qui `AmazonDataZoneGlueManageAccessRolePolicy` y est attaché. Ce rôle autorise Amazon à DataZone publier et à gérer les abonnements pour les actifs Data Lake, Data Warehouse et Amazon Sagemaker.

Le `AmazonDataZoneSageMakerManageAccessRole` rôle est associé à la politique en ligne suivante :

JSON

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Sid": "RedshiftSecretStatement",
         "Effect":"Allow",
         "Action":"secretsmanager:GetSecretValue",
         "Resource":"*",
         "Condition":{
            "StringEquals":{
               "secretsmanager:ResourceTag/AmazonDataZoneDomain":"{{domainId}}"
            }
         }
      }
   ]
}
```

Le `AmazonDataZoneSageMakerManageAccessRole` rôle est associé à la politique de confiance suivante :

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Sid": "DatazoneTrustPolicyStatement",
      "Effect": "Allow",
      "Principal": {
        "Service":  ["datazone.amazonaws.com",
                    "sagemaker.amazonaws.com"]
      },
      "Action": "sts:AssumeRole",
      "Condition": {
            "StringEquals": {
            "aws:SourceAccount": "111122223333"
            },
            "ArnEquals": {
            "aws:SourceArn": "arn:aws:datazone:us-east-1:111122223333:domain/
dzd-12345"
            }
        }
    }
]
}
```

AmazonDataZoneSageMakerProvisioningRolePolicyRole-<domainAccountId>

Le `AmazonDataZoneSageMakerProvisioningRolePolicyRole` rôle a `AmazonDataZoneSageMakerProvisioningRolePolicy` et ce qui `AmazonDataZoneRedshiftGlueProvisioningPolicy` y est attaché. Ce rôle accorde à Amazon DataZone les autorisations nécessaires pour interagir avec AWS Glue, Amazon Redshift et Amazon Sagemaker.

Le AmazonDataZoneSageMakerProvisioningRolePolicyRole rôle est associé à la politique en ligne suivante :

JSON

```
{
    "Version":"2012-10-17",
    "Statement": [
        {
            "Sid": "SageMakerStudioTagOnCreate",
            "Effect": "Allow",
            "Action": [
                "sagemaker:AddTags"
            ],
            "Resource": "arn:aws:sagemaker:*:111122223333:*/*",
            "Condition": {
                "Null": {
                    "sagemaker:TaggingAction": "false"
                }
            }
        }
    ]
}
```

Le AmazonDataZoneSageMakerProvisioningRolePolicyRole rôle est associé à la politique de confiance suivante :

JSON

```
{
  "Version":"2012-10-17",
  "Statement": [
    {
      "Sid": "DataZoneTrustPolicyStatement",
      "Effect": "Allow",
      "Principal": {
        "Service": "datazone.amazonaws.com"
```

```
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
            "aws:SourceAccount": "{{domain_account}}"
        }
      }
    }
  ]
}
```

Informations d'identification temporaires

Certains AWS services ne fonctionnent pas lorsque vous vous connectez à l'aide d'informations d'identification temporaires. Pour plus d'informations, notamment sur les AWS services qui fonctionnent avec des informations d'identification temporaires, consultez la section AWS Services compatibles avec IAM dans le Guide de l'utilisateur IAM.

Vous utilisez des informations d'identification temporaires si vous vous connectez à l' AWS Management Console aide d'une méthode autre qu'un nom d'utilisateur et un mot de passe. Par exemple, lorsque vous accédez à AWS l'aide du lien d'authentification unique (SSO) de votre entreprise, ce processus crée automatiquement des informations d'identification temporaires. Vous créez également automatiquement des informations d'identification temporaires lorsque vous vous connectez à la console en tant qu'utilisateur, puis changez de rôle. Pour plus d'informations sur le changement de rôle, consultez Changement de rôle (console) dans le Guide de l'utilisateur IAM.

Vous pouvez créer manuellement des informations d'identification temporaires à l'aide de l' AWS API AWS CLI or. Vous pouvez ensuite utiliser ces informations d'identification temporaires pour y accéder AWS. AWS recommande de générer dynamiquement des informations d'identification temporaires au lieu d'utiliser des clés d'accès à long terme. Pour plus d'informations, consultez Informations d'identification de sécurité temporaires dans IAM.

Informations d'identification temporaires DataZone du portail Amazon

Lorsque vous vous connectez au DataZone portail Amazon, vous recevez des informations d'identification temporaires pour AmazonDataZoneDomainExecutionRole. Lorsque vous utilisez leAmazonDataZoneDomainExecutionRole, ces informations d'identification sont automatiquement

actualisées lorsqu'elles sont utilisées. Lorsqu'ils ne sont pas utilisés pendant un certain temps, ils expirent automatiquement.

Autorisations de principal

Lorsque vous utilisez un utilisateur ou un rôle IAM pour effectuer des actions AWS, vous êtes considéré comme un mandant. Les politiques accordent des autorisations au principal. Lorsque vous utilisez certains services, vous pouvez effectuer une action qui déclenche une autre action dans un autre service. Dans ce cas, vous devez disposer des autorisations nécessaires pour effectuer les deux actions. Pour savoir si une action nécessite des actions dépendantes supplémentaires dans une politique, consultez la section Actions, ressources et clés de condition pour AWS Documentation Essentials dans la référence d'autorisation de service.

Validation de conformité pour Amazon DataZone

Pour savoir si un programme Services AWS de conformité Service AWS s'inscrit dans le champ d'application de programmes de conformité spécifiques, consultez Services AWS la section de conformité et sélectionnez le programme de conformité qui vous intéresse. Pour des informations générales, voir Programmes de AWS conformité Programmes AWS de .

Vous pouvez télécharger des rapports d'audit tiers à l'aide de AWS Artifact. Pour plus d'informations, voir Téléchargement de rapports dans AWS Artifact .

Votre responsabilité en matière de conformité lors de l'utilisation Services AWS est déterminée par la sensibilité de vos données, les objectifs de conformité de votre entreprise et les lois et réglementations applicables. Pour plus d'informations sur votre responsabilité en matière de conformité lors de l'utilisation Services AWS, consultez AWS la documentation de sécurité.

Bonnes pratiques en matière de sécurité pour Amazon DataZone

Amazon DataZone propose un certain nombre de fonctionnalités de sécurité à prendre en compte lors de l'élaboration et de la mise en œuvre de vos propres politiques de sécurité. Les bonnes pratiques suivantes doivent être considérées comme des instructions générales et ne représentent pas une solution de sécurité complète. Étant donné que ces bonnes pratiques peuvent ne pas être appropriées ou suffisantes pour votre environnement, considérez-les comme des remarques utiles plutôt que comme des recommandations.

Implémentation d'un accès sur la base du moindre privilège

Lorsque vous accordez des autorisations, vous décidez qui obtient quelles autorisations pour quelles DataZone ressources Amazon. Vous activez des actions spécifiques que vous souhaitez autoriser sur ces ressources. Par conséquent, vous devez accorder uniquement les autorisations qui sont requises pour exécuter une tâche. L'implémentation d'un accès sur la base du moindre privilège est fondamentale pour réduire les risques en matière de sécurité et l'impact que pourraient avoir des d'erreurs ou des actes de malveillance.

Pour plus d'informations, consultez AWS politiques gérées pour Amazon DataZone la section Politiques de contrôle des services (SCPs).

Utilisation des rôles IAM

Les applications productrices et clientes doivent disposer d'informations d'identification valides pour accéder aux DataZone ressources Amazon. Vous ne devez pas stocker les AWS informations d'identification directement dans une application cliente ou dans un compartiment Amazon S3. Il s'agit d'autorisations à long terme qui ne font pas automatiquement l'objet d'une rotation et qui pourraient avoir un impact commercial important si elles étaient compromises.

Vous devez plutôt utiliser un rôle IAM pour gérer les informations d'identification temporaires permettant à vos applications productrices et clientes d'accéder aux DataZone ressources Amazon. Lorsque vous utilisez un rôle, vous n'avez pas à utiliser d'informations d'identification à long terme (par exemple, un nom d'utilisateur et un mot de passe ou des clés d'accès) pour accéder à d'autres ressources.

Pour plus d'informations, consultez les rubriques suivantes dans le Guide de l'utilisateur IAM :

- Rôles IAM
- Scénarios courants pour les rôles : utilisateurs, applications et services.

Implémentation d'un chiffrement côté serveur dans des ressources dépendantes

Les données au repos et les données en transit peuvent être cryptées sur Amazon DataZone.

CloudTrail À utiliser pour surveiller les appels d'API

Amazon DataZone est intégré à AWS CloudTrail un service qui fournit un enregistrement des actions entreprises par un utilisateur, un rôle ou un AWS service sur Amazon DataZone.

À l'aide des informations collectées par CloudTrail, vous pouvez déterminer la demande qui a été faite à Amazon DataZone, l'adresse IP à partir de laquelle la demande a été faite, l'auteur de la demande, la date à laquelle elle a été faite et des informations supplémentaires.

Utilisation de la RAM sur Amazon DataZone

L'association de vos AWS comptes aux DataZone domaines Amazon permet aux utilisateurs du domaine de publier et de consommer les données de ces AWS comptes. Amazon DataZone utilise AWS Resource Access Manager (RAM) pour gérer l'accès entre comptes. Pour plus d'informations, reportez-vous à Comptes associés sur Amazon DataZone la section Sécurité de la AWS RAM.

Résilience chez Amazon DataZone

L'infrastructure AWS mondiale est construite autour Régions AWS de zones de disponibilité. Régions AWS fournissent plusieurs zones de disponibilité physiquement séparées et isolées, connectées par un réseau à faible latence, à haut débit et hautement redondant. Avec les zones de disponibilité, vous pouvez concevoir et exploiter des applications et des bases de données qui basculent automatiquement d'une zone à l'autre sans interruption. Les zones de disponibilité sont davantage disponibles, tolérantes aux pannes et ont une plus grande capacité de mise à l'échelle que les infrastructures traditionnelles à un ou plusieurs centres de données.

Pour plus d'informations sur les zones de disponibilité Régions AWS et les zones de disponibilité, consultez la section Infrastructure AWS globale.

Outre l'infrastructure AWS mondiale, Amazon DataZone propose plusieurs fonctionnalités pour répondre à vos besoins en matière de résilience et de sauvegarde des données.

Rubriques

- Résilience des sources de données
- Résilience des actifs
- Résilience des types d'actifs et des formulaires de métadonnées
- Glossaire : résilience
- Résilience des recherches mondiales

- Résilience des abonnements
- Résilience environnementale
- Résilience du plan environnemental
- Résilience des projets
- résilience de la RAM
- Résilience de la gestion des profils utilisateur
- Résilience du domaine

Résilience des sources de données

Lors d'un événement de DataZone disponibilité d'Amazon, les `DataSource` offres d'emploi seront régulièrement réessayées pendant 24 heures au maximum. Si une tâche échoue en raison d'une mauvaise configuration, un `DataSourceRunFailed` événement sera émis. Si le DataZone domaine Amazon est configuré avec une clé KMS et qu' AmazonDataZoneDomainExecutionRole il perd l'accès à cette clé pendant l'exécution d'une tâche, l'exécution se terminera dans son `INACCESSIBLE` état actuel. Une fois l'accès KMS rétabli, la tâche doit être mise à jour manuellement pour déclencher le retour à un état utilisable.

Résilience des actifs

Dans Amazon DataZone, les actifs sont versionnés. Si une version d'un actif doit être annulée, vous pouvez créer une nouvelle version en utilisant le contenu de la dernière version stable. Une version d'un actif peut être publiée. Une version publiée d'une ressource ne peut pas être modifiée, sauf en publiant une nouvelle version. Il est possible de s'abonner à un actif publié (alias une liste). Pour empêcher de nouveaux abonnements à un actif, celui-ci peut être dépublié. L'annulation de la publication d'un actif n'a aucun effet sur les abonnements existants. La suppression d'une ressource entraîne la suppression de toutes les versions non publiées de cette ressource. Les versions publiées de la ressource doivent être supprimées séparément. Une version publiée d'un actif ne peut être supprimée que s'il n'y a aucun abonnement.

Résilience des types d'actifs et des formulaires de métadonnées

Dans Amazon DataZone, les types de ressources et les types de formulaires de métadonnées sont versionnés. Un type de ressource ne peut pas être supprimé s'il est utilisé par une ressource. Un type de formulaire de métadonnées ne peut pas être supprimé s'il est utilisé par un type de ressource ou une ressource. Si vous ne souhaitez pas que des metadata-form-type informations spécifiques soient

utilisées pour la curation, vous pouvez les désactiver, sans affecter celles auxquelles elles sont déjà associées.

Glossaire : résilience

Sur Amazon DataZone, les glossaires et les termes du glossaire ne peuvent pas être supprimés s'ils sont utilisés. Si vous ne souhaitez pas qu'un glossaire ou un terme de glossaire spécifique soit utilisé pour la curation, vous pouvez les désactiver sans affecter ceux auxquels ils sont déjà attachés.

Résilience des recherches mondiales

Sur Amazon DataZone, les actifs publiés (alias listes) peuvent être découverts par le biais d'une recherche globale. La publication d'une ressource peut être annulée en annulant la publication de la ressource. L'annulation de la publication d'un actif n'affecte pas les abonnements existants. Une ressource publiée peut être ramenée à une version spécifique de la ressource en republiant cette version. Cela n'affectera pas les abonnements existants.

Résilience des abonnements

Sur Amazon DataZone, SubscriptionGrant Fulfillment tentera deux tentatives de retrait avant d'échouer. En cas d'échec, il doit être supprimé manuellement pour réessayer.
Si Amazon DataZone ne parvient pas à révoquer les autorisations relatives à un abonnement, la suppression de l'abonnement risque d'échouer. L'erreur sous-jacente doit être corrigée, ou l'`retainPermissions`indicateur peut être utilisé dans l'opération d'`DeleteSubscriptionGrant`API pour forcer la suppression de la subvention auprès d'Amazon DataZone sans révoquer les autorisations.

Si le DataZone domaine Amazon est configuré avec une clé KMS et qu'`AmazonDataZoneDomainExecutionRole`il perd l'accès à cette clé pendant le `SubscriptionGrant` flux de travail, l'autorisation est marquée`INACCESSIBLE`. Une fois l'accès KMS rétabli, les `INACCESSIBLE` autorisations doivent être supprimées et recréées.

Résilience environnementale

Si le DataZone domaine Amazon est configuré avec une clé KMS et qu'`AmazonDataZoneDomainExecutionRole`il perd l'accès à cette clé pendant le flux de travail environnemental, l'environnement sera marqué`INACCESSIBLE`. Une fois l'accès KMS rétabli, l'`INACCESSIBLE`environnement doit être supprimé et recréé. La création de l'environnement tentera deux fois de se retirer avant d'échouer. En cas d'échec, il doit être supprimé manuellement pour

réessayer. Si le flux de travail environnemental échoue, l'environnement passe à l'état d'échec. À ce stade, il ne peut être que supprimé et recréé.

Résilience du plan environnemental

Dans Amazon DataZone, un plan d'environnement ne peut pas être supprimé s'il existe des profils d'environnement sous-jacents.

Résilience des projets

Dans Amazon DataZone, un projet ne peut pas être supprimé s'il existe des environnements confinés.

résilience de la RAM

Pour obtenir des informations sur la résilience de la RAM, voir https://docs.aws.amazon.com/ram/latest/userguide/security- disaster-recovery-resiliency .html.

Résilience de la gestion des profils utilisateur

Pour obtenir des informations sur la résilience des profils utilisateur, consultez AWS Identity Center.

Résilience du domaine

Sur Amazon DataZone, un domaine ne peut pas être supprimé s'il contient des projets ou des sources de données.

Sécurité de l'infrastructure sur Amazon DataZone

En tant que service géré, Amazon DataZone est protégé par la sécurité du réseau AWS mondial. Pour plus d'informations sur les services AWS de sécurité et sur la manière dont AWS l'infrastructure est protégée, consultez la section Sécurité du AWS cloud. Pour concevoir votre AWS environnement en utilisant les meilleures pratiques en matière de sécurité de l'infrastructure, consultez la section Protection de l'infrastructure dans le cadre AWS bien architecturé du pilier de sécurité.

Vous utilisez des appels d'API AWS publiés pour accéder à Amazon DataZone via le réseau. Les clients doivent prendre en charge les éléments suivants :

- Protocole TLS (Transport Layer Security). Nous exigeons TLS 1.2 et recommandons TLS 1.3.

- Ses suites de chiffrement PFS (Perfect Forward Secrecy) comme DHE (Ephemeral Diffie-Hellman) ou ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). La plupart des systèmes modernes tels que Java 7 et les versions ultérieures prennent en charge ces modes.

Prévention interservices confuse des adjoints sur Amazon DataZone

Le problème de député confus est un problème de sécurité dans lequel une entité qui n'est pas autorisée à effectuer une action peut contraindre une entité plus privilégiée à le faire. En AWS, l'usurpation d'identité interservices peut entraîner la confusion des adjoints. L'usurpation d'identité entre services peut se produire lorsqu'un service (le service appelant) appelle un autre service (le service appelé). Le service appelant peut être manipulé et ses autorisations utilisées pour agir sur les ressources d'un autre client auxquelles on ne serait pas autorisé à accéder autrement. Pour éviter cela, AWS fournit des outils qui vous aident à protéger vos données pour tous les services auprès des principaux fournisseurs de services qui ont obtenu l'accès aux ressources de votre compte.

Nous recommandons d'utiliser la clé de contexte aws : SourceAccount global condition dans les politiques relatives aux ressources afin de limiter les autorisations qu'Amazon DataZone accorde à un autre service sur la ressource. Utilisez aws : SourceAccount si vous souhaitez autoriser l'association de toute ressource de ce compte à l'utilisation interservices.

Analyse de configuration et de vulnérabilité pour Amazon DataZone

AWS gère les tâches de sécurité de base telles que l'application de correctifs au système d'exploitation client (OS) et aux bases de données, la configuration du pare-feu et la reprise après sinistre. Ces procédures ont été vérifiées et certifiées par les tiers appropriés. Pour plus d'informations, consultez le modèle de responsabilité AWS partagée.

Domaines à ajouter à votre liste d'autorisations

Pour que le portail de DataZone données Amazon puisse accéder au DataZone service Amazon, vous devez ajouter les domaines suivants à la liste d'autorisation sur le réseau à partir duquel le portail de données tente d'accéder au service.

- `*.api.aws`
- `*.on.aws`

Surveillance d'Amazon DataZone

La surveillance joue un rôle important dans le maintien de la fiabilité, de la disponibilité et des performances d'Amazon DataZone et de vos autres AWS solutions. AWS fournit les outils de surveillance suivants pour surveiller Amazon DataZone, signaler un problème et prendre des mesures automatiques le cas échéant :

- Amazon CloudWatch surveille vos AWS ressources et les applications que vous utilisez AWS en temps réel. Vous pouvez collecter et suivre les métriques, créer des tableaux de bord personnalisés, et définir des alarmes qui vous informent ou prennent des mesures lorsqu'une métrique spécifique atteint un seuil que vous spécifiez. Par exemple, vous pouvez CloudWatch suivre l'utilisation du processeur ou d'autres indicateurs de vos EC2 instances Amazon et lancer automatiquement de nouvelles instances en cas de besoin. Pour plus d'informations, consultez le guide de CloudWatch l'utilisateur Amazon.
- Amazon CloudWatch Logs vous permet de surveiller, de stocker et d'accéder à vos fichiers journaux à partir d' EC2 instances Amazon et d'autres sources. CloudTrail CloudWatch Les journaux peuvent surveiller les informations contenues dans les fichiers journaux et vous avertir lorsque certains seuils sont atteints. Vous pouvez également archiver vos données de journaux dans une solution de stockage hautement durable. Pour plus d'informations, consultez le guide de l'utilisateur Amazon CloudWatch Logs.
- Amazon EventBridge peut être utilisé pour automatiser vos AWS services et répondre automatiquement aux événements du système, tels que les problèmes de disponibilité des applications ou les modifications des ressources. Les événements AWS liés aux services sont diffusés EventBridge en temps quasi réel. Vous pouvez écrire des règles simples pour préciser les événements qui vous intéressent et les actions automatisées à effectuer quand un événement correspond à une règle. Pour plus d'informations, consultez le guide de EventBridge l'utilisateur Amazon.
- AWS CloudTrailcapture les appels d'API et les événements associés effectués par ou pour le compte de votre AWS compte et envoie les fichiers journaux dans un compartiment Amazon S3 que vous spécifiez. Vous pouvez identifier les utilisateurs et les comptes appelés AWS, l'adresse IP source à partir de laquelle les appels ont été effectués et la date des appels. Pour plus d'informations, consultez le Guide de l'utilisateur AWS CloudTrail.

Surveillance des DataZone événements Amazon sur Amazon EventBridge

Vous pouvez surveiller les DataZone événements Amazon dans EventBridge, qui fournit un flux de données en temps réel à partir de vos propres applications, applications software-as-a-service (SaaS) et AWS services. EventBridgeachemine ces données vers des cibles telles qu' AWS Lambda Amazon Simple Notification Service. Ces événements sont les mêmes que ceux qui apparaissent dans Amazon CloudWatch Events, qui fournit un flux quasi en temps réel d'événements système décrivant les modifications apportées aux AWS ressources.

Pour de plus amples informations, veuillez consulter Événements via le bus EventBridge par défaut d'Amazon.

Journalisation des appels d' DataZone API Amazon à l'aide de AWS CloudTrail

Amazon DataZone est intégré à AWS CloudTrail un service qui fournit un enregistrement des actions entreprises par un utilisateur, un rôle ou un AWS service sur Amazon DataZone. CloudTrail capture tous les appels d'API pour Amazon DataZone sous forme d'événements. Les appels capturés incluent des appels provenant de la DataZone console Amazon et des appels de code vers les opérations de DataZone l'API Amazon. Si vous créez un suivi, vous pouvez activer la diffusion continue d' CloudTrailévénements vers un compartiment Amazon S3, y compris des événements pour Amazon DataZone. Si vous ne configurez pas de suivi, vous pouvez toujours consulter les événements les plus récents dans la CloudTrail console dans Historique des événements. À l'aide des informations collectées par CloudTrail, vous pouvez déterminer la demande qui a été faite à Amazon DataZone, l'adresse IP à partir de laquelle la demande a été faite, l'auteur de la demande, la date à laquelle elle a été faite et des informations supplémentaires.

Pour en savoir plus CloudTrail, consultez le guide de AWS CloudTrail l'utilisateur.

DataZone Informations Amazon dans CloudTrail

CloudTrail est activé sur votre compte Compte AWS lorsque vous créez le compte. Lorsqu'une activité se produit dans la console DataZone de gestion Amazon, cette activité est enregistrée dans un CloudTrail événement avec d'autres événements de AWS service dans l'historique des événements. Vous pouvez consulter, rechercher et télécharger les événements récents dans votre

Compte AWS. Pour plus d'informations, consultez la section Affichage des événements à l'aide de l'historique des CloudTrail événements.

Pour un enregistrement continu des événements de votre entreprise Compte AWS, y compris des événements pour Amazon DataZone, créez un parcours. Un suivi permet CloudTrail de fournir des fichiers journaux à un compartiment Amazon S3. Par défaut, lorsque vous créez un journal d'activité dans la console, il s'applique à toutes les régions Régions AWS. Le journal enregistre les événements de toutes les régions de la AWS partition et transmet les fichiers journaux au compartiment Amazon S3 que vous spécifiez. En outre, vous pouvez configurer d'autres AWS services pour analyser plus en détail les données d'événements collectées dans les CloudTrail journaux et agir en conséquence. Pour plus d'informations, consultez les ressources suivantes :

- Présentation de la création d'un journal de suivi
- CloudTrail services et intégrations pris en charge
- Configuration des notifications Amazon SNS pour CloudTrail
- Réception de fichiers CloudTrail journaux de plusieurs régions et réception de fichiers CloudTrail journaux de plusieurs comptes

Toutes les DataZone actions d'Amazon sont enregistrées par CloudTrail.

Résolution des problèmes liés à Amazon DataZone

Si vous rencontrez des problèmes de refus d'accès ou des difficultés similaires lorsque vous travaillez avec Amazon, DataZone consultez les rubriques de cette section.

Résolution des problèmes liés aux autorisations de AWS Lake Formation pour Amazon DataZone

Cette section contient des instructions de résolution des problèmes que vous pourriez rencontrer lorsque vousConfigurer les autorisations de Lake Formation pour Amazon DataZone.

Message d'erreur dans le portail de données	Résolution
Impossible d'assumer le rôle d'accès aux données.	Cette erreur s'affiche lorsqu'Amazon n' DataZone est pas en mesure de supposer AmazonDataZoneGlueDataAccessRoleque vous avez utilisé pour activer le DefaultDa taLakeBlueprintdans votre compte. Pour résoudre le problème, accédez à la console AWS IAM du compte sur lequel se trouve votre ressource de données et assurez-vous qu'il existe AmazonDataZoneGlueDataAcces sRoleune relation de confiance appropriée avec le responsable du DataZone service Amazon. Pour de plus amples informations, consultez AmazonDataZoneGlueAccess- <region>- <domainId>.
Le rôle d'accès aux données ne dispose pas des autorisations nécessaires pour lire les métadonnées de la ressource à laquelle vous essayez de vous abonner.	Cette erreur s'affiche lorsqu'Amazon assume le AmazonDataZoneGlueDataAccessRolerôle DataZone avec succès, mais que celui-ci ne dispose pas des autorisations nécessair es. Pour résoudre le problème, accédez à la console AWS IAM du compte auquel se trouve votre ressource de données et assurez-vous que le rôle est AmazonDataZoneGlue

Message d'erreur dans le portail de données	Résolution
	ManageAccessRolePolicyassocié à ce dernier. Pour de plus amples informations, veuillez consulter AmazonDataZoneGlueAccess-<region>- <domainId>.
L'actif est un lien vers une ressource. Amazon DataZone ne prend pas en charge les abonnements à des liens vers des ressources.	Cette erreur s'affiche lorsque la ressource que vous essayez de publier sur Amazon DataZone est un lien de ressource vers une table AWS Glue.

Message d'erreur dans le portail de données	Résolution
L'actif n'est pas géré par AWS Lake Formation.	Cette erreur indique que les autorisations de AWS Lake Formation ne sont pas appliquées à la ressource que vous souhaitez publier. Cela peut se produire dans les cas suivants. • L'emplacement Amazon S3 de la ressource n'est pas enregistré dans AWS Lake Formation. Pour résoudre le problème, connectez-vous à votre console AWS Lake Formation sur le compte sur lequel se trouve la table et enregistrez l'emplacement Amazon S3 en mode AWS Lake Formation ou en mode hybride. Pour plus d'informations, consultez la rubrique Enregistrement d'un emplacement Amazon S3. Plusieurs scénarios nécessitent des modifications supplémentaires. Il s'agit notamment de compartiments Amazon S3 cryptés ou d'un compartiment S3 multi-comptes et d'une configuration Glue AWS Catalog. Dans de tels cas, il peut être nécessaire de modifier les paramètres de KMS and/or S3. Pour plus d'informations, consultez la rubrique Enregistrement d'un emplacement Amazon S3 chiffré. • L'emplacement Amazon S3 est enregistré en mode AWS Lake Formation mais IAMAllowedPrincipal est ajouté aux autorisations de la table. Pour résoudre le problème, vous pouvez soit supprimer le IAMAllowedprincipal des autorisations de la table, soit enregistrer l'emplacement S3 en mode hybride. Pour plus d'informations, voir À propos de la mise à niveau vers le modèle d'autorisations de Lake Formation. Si votre position S3 est

Message d'erreur dans le portail de données	Résolution
	cryptée ou si elle se trouve sur un compte différent de celui de votre table AWS Glue, suivez les instructions de la section Enregistrer une position Amazon S3 cryptée.
Le rôle Data Access ne dispose pas des autorisations Lake Formation nécessaires pour accorder l'accès à cette ressource.	Cette erreur indique que le AmazonDataZoneGlueDataAccessRolefichier que vous utilisez pour activer le DefaultDataLakeBlueprintdans votre compte ne dispose pas des autorisations nécessaires pour qu'Amazon puisse DataZone gérer les autorisations sur l'actif publié. Vous pouvez résoudre le problème soit en l'ajoutant en AmazonDataZoneGlueDataAccessRoletant qu'administrateur de AWS Lake Formation, soit en accordant les autorisations suivantes AmazonDataZoneGlueDataAccessRoleà la ressource que vous souhaitez publier. • Décrire et décrire les autorisations pouvant être accordées sur la base de données où se trouve l'actif • Décrivez, sélectionnez, décrivez les autorisations pouvant être accordées, sélectionnez les autorisations pouvant être accordées sur tous les actifs de la base de données dont vous souhaitez qu'Amazon gère l'accès en votre nom. DataZone

Résolution des problèmes liés à la liaison des actifs Amazon DataZone Lineage avec des ensembles de données en amont

Cette section contient des instructions de résolution des problèmes que vous pourriez rencontrer avec Amazon DataZone Lineage. Pour certains événements d'exécution de AWS Glue lignage

ouvert liés à Amazon Redshift, vous constaterez peut-être que le lignage des actifs n'est pas lié à un ensemble de données en amont. Cette rubrique explique les scénarios et quelques approches pour atténuer les problèmes. Pour plus d'informations sur le lignage, voirLignage des données sur Amazon DataZone.

SourceIdentifier sur un nœud de lignée

L'`sourceIdentifier`attribut d'un nœud de lignée représente les événements qui se produisent sur un ensemble de données. Pour plus d'informations, consultez la section Attributs clés dans les nœuds de lignage.

Le nœud de lignage représente tous les événements qui se produisent sur le jeu de données ou le travail correspondant. Le nœud de lignée contient un attribut « SourceIdentifier » qui contient l'identifiant du jeu de données/tâches correspondant. Comme nous prenons en charge les événements de lignage ouvert, la `sourceIdentifier` valeur est renseignée par défaut sous la forme de la combinaison de « namespace » et de « nom » pour un ensemble de données, une tâche et des exécutions de tâches.

Pour AWS des ressources telles qu' AWS Glue Amazon Redshift, il s'`sourceIdentifier`agirait de l' AWS Glue ARN de la table et de la table Redshift à partir ARNs desquelles Amazon DataZone créera l'événement d'exécution et les autres informations suivantes :

> **Note**
>
> Dans AWS, l'ARN contient des informations telles que l'AccountID, la région, la base de données et la table pour chaque ressource.

- OpenLineage l'événement pour ces ensembles de données contient le nom de la base de données et de la table.
- La région est capturée dans la facette « propriétés de l'environnement » d'une course. Si ce n'est pas le cas, le système utilise la région indiquée dans les informations d'identification de l'appelant.
- AccountId est extrait des informations d'identification de l'appelant.

SourceIdentifier sur les actifs qu'il contient DataZone

`AssetCommonDetailForm`possède un attribut appelé « SourceIdentifier » qui représente l'identifiant de l'ensemble de données représenté par la ressource. Pour que les nœuds de lignage des

actifs soient liés à un ensemble de données en amont, l'attribut doit être renseigné avec la valeur correspondant à celle du nœud du jeu de `sourceIdentifier` données. Si les actifs sont importés par source de données, le flux de travail est automatiquement renseigné `sourceIdentifier` sous forme de tableau ARN/Redshift ARN, tandis que les autres actifs (y compris les actifs personnalisés) créés via `CreateAsset` l'API doivent avoir cette valeur renseignée par l'appelant. AWS Glue

Comment Amazon crée-t-il le DataZone SourceIdentifier à partir de l' OpenLineage événement ?

Pour les actifs Redshift AWS Glue et Redshift, ils `sourceIdentifier` sont conçus à partir de Glue et Redshift. ARNs Voici comment Amazon le DataZone construit :

AWS Glue ARN

L'objectif est de créer un OpenLineage événement dont le nœud de lignée en sortie `sourceIdentifier` est :

```
arn:aws:glue:us-east-1:123456789012:table/testlfdb/testlftb-1
```

Pour déterminer si une course utilise des données provenant de AWS Glue, recherchez la présence de certains mots clés dans la `environment-properties` facette. Plus précisément, si l'un de ces champs désignés est présent, le système suppose qu'il `RunEvent` provient de AWS Glue.

- GLUE_VERSION
- CRITÈRES_COMMANDE_COLLE
- GLUE_PYTHON_VERSION

```
"run": {
   "runId":"4e3da9e8-6228-4679-b0a2-fa916119fthr",
   "facets":{
      "environment-properties":{
         "_producer":"https://github.com/OpenLineage/OpenLineage/tree/1.9.1/
integration/spark",
         "_schemaURL":"https://openlineage.io/spec/2-0-2/OpenLineage.json#/$defs/
RunFacet",
         "environment-properties":{
            "GLUE_VERSION":"3.0",
            "GLUE_COMMAND_CRITERIA":"glueetl",
            "GLUE_PYTHON_VERSION":"3"
```

```
         }
      }
   }
```

Pour une AWS Glue exécution, vous pouvez utiliser le nom de la `symlinks` facette pour obtenir le nom de la base de données et de la table, qui peuvent être utilisés pour créer l'ARN.

Vous devez vous assurer que le nom est `databaseName.tableName` :

```
"symlinks": {
   "_producer":"https://github.com/OpenLineage/OpenLineage/tree/1.9.1/integration/
spark",
   "_schemaURL":"https://openlineage.io/spec/facets/1-0-0/SymlinksDatasetFacet.json#/
$defs/SymlinksDatasetFacet",
   "identifiers":[
      {
         "namespace":"s3://object-path",
         "name":"testlfdb.testlftb-1",
         "type":"TABLE"
      }
   ]
}
```

Exemple d'événement COMPLET :

```
{
   "eventTime":"2024-07-01T12:00:00.000000Z",
   "producer":"https://github.com/OpenLineage/OpenLineage/tree/1.9.1/integration/glue",
   "schemaURL":"https://openlineage.io/spec/2-0-2/OpenLineage.json#/$defs/RunEvent",
   "eventType":"COMPLETE",
   "run": {
      "runId":"4e3da9e8-6228-4679-b0a2-fa916119fthr",
      "facets":{
         "environment-properties":{
            "_producer":"https://github.com/OpenLineage/OpenLineage/tree/1.9.1/
integration/spark",
            "_schemaURL":"https://openlineage.io/spec/2-0-2/OpenLineage.json#/$defs/
RunFacet",
            "environment-properties":{
               "GLUE_VERSION":"3.0",
               "GLUE_COMMAND_CRITERIA":"glueetl",
               "GLUE_PYTHON_VERSION":"3"
            }
```

```
         }
      }
   },
   "job":{
      "namespace":"namespace",
      "name":"job_name",
      "facets":{
         "jobType":{
            "_producer":"https://github.com/OpenLineage/OpenLineage/tree/1.9.1/
integration/glue",
            "_schemaURL":"https://openlineage.io/spec/facets/2-0-2/
JobTypeJobFacet.json#/$defs/JobTypeJobFacet",
            "processingType":"BATCH",
            "integration":"glue",
            "jobType":"JOB"
         }
      }
   },
   "inputs":[
      {
         "namespace":"namespace",
         "name":"input_name"
      }
   ],
   "outputs":[
      {
         "namespace":"namespace.output",
         "name":"output_name",
         "facets":{
            "symlinks":{
               "_producer":"https://github.com/OpenLineage/OpenLineage/tree/1.9.1/
integration/spark",
               "_schemaURL":"https://openlineage.io/spec/facets/1-0-0/
SymlinksDatasetFacet.json#/$defs/SymlinksDatasetFacet",
               "identifiers":[
                  {
                     "namespace":"s3://object-path",
                     "name":"testlfdb.testlftb-1",
                     "type":"TABLE"
                  }
               ]
            }
         }
      }
```

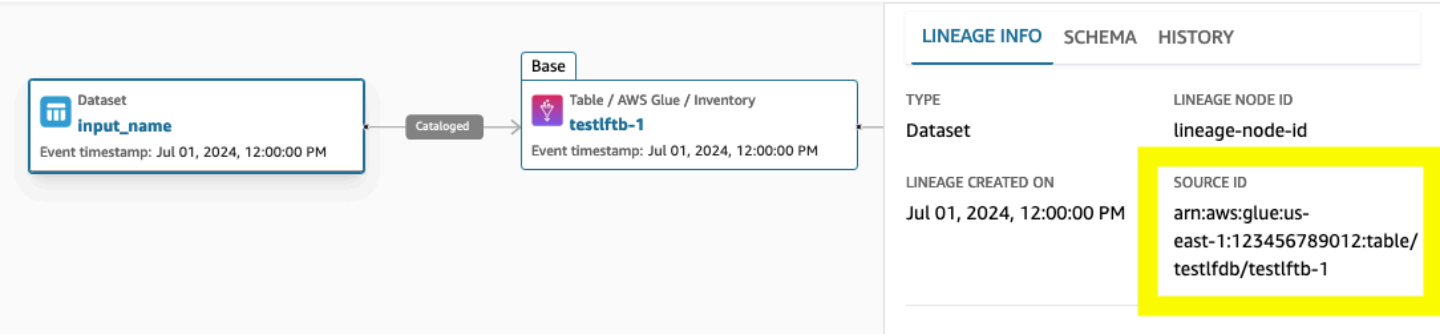
```
    ]
}
```

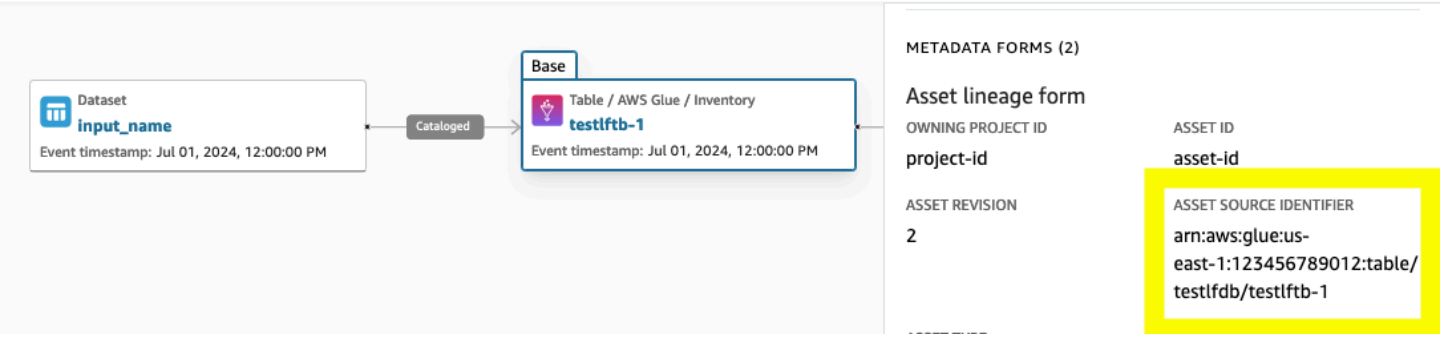
Sur la base de l'`OpenLineage`événement soumis, le nœud `sourceIdentifier` de lignée de sortie sera :

```
arn:aws:glue:us-east-1:123456789012:table/testlfdb/testlftb-1
```

Le nœud de lignage en sortie sera connecté au nœud de lignage d'un actif où l'actif est `sourceIdentifier` :

```
arn:aws:glue:us-east-1:123456789012:table/testlfdb/testlftb-1
```





Amazon Redshift (ARN)

L'objectif est de créer un OpenLineage événement dont le nœud de lignée en sortie `sourceIdentifier` est :

```
arn:aws:redshift:us-east-1:123456789012:table/workgroup-20240715/tpcds_data/public/
dws_tpcds_7
```

Le système détermine si une entrée ou une sortie est stockée dans Redshift en fonction de l'espace de noms. Plus précisément, si l'espace de noms commence par redshift ://ou contient les chaînes

`redshift-serverless.amazonaws.com` ou `redshift.amazonaws.com` s'il s'agit d'une ressource Redshift.

```
"outputs": [
    {
        "namespace":"redshift://workgroup-20240715.123456789012.us-
east-1.redshift.amazonaws.com:5439",
        "name":"tpcds_data.public.dws_tpcds_7"
    }
]
```

Notez que l'espace de noms doit être au format suivant :

```
provider://{cluster_identifier}.{region_name}:{port}
```

Pour `redshift-serverless` :

```
"outputs": [
    {
        "namespace":"redshift://workgroup-20240715.123456789012.us-east-1.redshift-
serverless.amazonaws.com:5439",
        "name":"tpcds_data.public.dws_tpcds_7"
    }
]
```

Les résultats sont les suivants `sourceIdentifier`

```
arn:aws:redshift-serverless:us-east-1:123456789012:table/workgroup-20240715/tpcds_data/
public/dws_tpcds_7
```

Sur la base de l' OpenLineage événement soumis, le nœud de lignage `sourceIdentifier` à mapper vers un nœud de lignage en aval (c'est-à-dire une sortie de l'événement) est le suivant :

```
arn:aws:redshift-serverless:us-e:us-east-1:123456789012:table/workgroup-20240715/
tpcds_data/public/dws_tpcds_7
```

Il s'agit du mappage qui vous permet de visualiser le lignage d'un actif dans le catalogue.

Approche alternative

Lorsqu'aucune des conditions ci-dessus n'est remplie, le système utilise l'espace de noms/name pour construire : `sourceIdentifier`

```
"inputs": [
  {
     "namespace":"arn:aws:redshift:us-east-1:123456789012:table",
     "name":"workgroup-20240715/tpcds_data/public/dws_tpcds_7"
  }
],
"outputs": [
  {
     "namespace":"arn:aws:glue:us-east-1:123456789012:table",
     "name":"testlfdb/testlftb-1"
  }
]
```

Résolution d'un manque d'amont pour le nœud de traçabilité des actifs

Si vous ne voyez pas l'amont du nœud de lignage des actifs, vous pouvez effectuer les opérations suivantes pour déterminer pourquoi il n'est pas lié au jeu de données :

1. Invoquez `GetAsset` tout en fournissant le `domainId` et `assetId` :

   ```
   aws datazone get-asset --domain-identifier <domain-id> --identifier <asset-id>
   ```

   La réponse apparaît comme suit :

   ```
   {
       .....
       "formsOutput": [
           .....
           {
               "content": "{\"sourceIdentifier\":\"arn:aws:glue:eu-
   west-1:123456789012:table/testlfdb/testlftb-1\"}",
               "formName": "AssetCommonDetailsForm",
               "typeName": "amazon.datazone.AssetCommonDetailsFormType",
               "typeRevision": "6"
           },
           .....
   ```

```
    ],
    "id": "<asset-id>",
    ....
}
```

2. Appelez `GetLineageNode` pour obtenir le nœud `sourceIdentifier` de lignée de l'ensemble de données. Comme il n'existe aucun moyen d'obtenir directement le nœud de lignage pour le nœud de jeu de données correspondant, vous pouvez commencer `GetLineageNode` par :

```
aws datazone get-lineage-node --domain-identifier <domain-id> --identifier
 <job_namespace>.<job_name>/<run_id>

if you are using the getting started scripts, job name and run ID are printed in the
 console
and namespace is "default". Otherwise you can get these values from run event
 content.
```

L'exemple de réponse se présente comme suit :

```
{
    .....
    "downstreamNodes": [
        {
            "eventTimestamp": "2024-07-24T18:08:55+08:00",
            "id": "afymge5k4v0euf"
        }
    ],
    "formsOutput": [
        <some forms corresponding to run and job>
    ],
    "id": "<system generated node-id for run>",
    "sourceIdentifier": "default.redshift.create/2f41298b-1ee7-3302-
a14b-09addffa7580",
    "typeName": "amazon.datazone.JobRunLineageNodeType",
    ....
    "upstreamNodes": [
        {
            "eventTimestamp": "2024-07-24T18:08:55+08:00",
            "id": "6wf2z27c8hghev"
        },
        {
            "eventTimestamp": "2024-07-24T18:08:55+08:00",
```

```
            "id": "4tjbcsnre6banb"
        }
    ]
}
```

3. Appelez `GetLineageNode` à nouveau en transmettant l'identifiant du downstream/upstream nœud (qui, selon vous, devrait être lié au nœud de ressource) car celui-ci correspond à l'ensemble de données :

   Exemple de commande utilisant l'exemple de réponse ci-dessus :

```
aws datazone get-lineage-node --domain-identifier <domain-id> --identifier
 afymge5k4v0euf
```

   Cela renvoie les détails du nœud de lignée correspondant à l'ensemble de données : afymge5k4v0euf

```
{
    .....
    "domainId": "dzd_cklzc5s2jcr7on",
    "downstreamNodes": [],
    "eventTimestamp": "2024-07-24T18:08:55+08:00",
    "formsOutput": [
        .....
    ],
    "id": "afymge5k4v0euf",
    "sourceIdentifier": "arn:aws:redshift:us-east-1:123456789012:table/
workgroup-20240715/tpcds_data/public/dws_tpcds_7",
    "typeName": "amazon.datazone.DatasetLineageNodeType",
    "typeRevision": "1",
    ....
    "upstreamNodes": [
        ...
    ]
}
```

4. Comparez le nœud `sourceIdentifier` de ce jeu de données et la réponse de`GetAsset`. S'ils ne sont pas liés, ils ne correspondront pas et ne seront donc pas visibles dans l'interface utilisateur du lignage.

Scénarios et mesures d'atténuation non correspondants

Voici les scénarios les plus courants dans lesquels ils ne correspondent pas, ainsi que les mesures d'atténuation possibles :

Cause première : les tables sont présentes dans un compte différent de celui du compte de DataZone domaine Amazon.

Atténuation : vous pouvez invoquer l'`PostLineageEvent`opération depuis un compte associé. Comme l'ARN est sélectionné `accountId` à partir des informations d'identification de l'appelant, vous pouvez assumer le rôle depuis le compte contenant les tables lorsque vous exécutez le script de démarrage ou que vous l'`PostLineageEvent`invoquez. Cela aidera à construire ARNs correctement les nœuds d'actifs et à établir des liens avec eux.

Cause première : L'ARN de Redshift table/views contient Redshift/Redshift-Serverless en fonction de l'espace de noms et des attributs de nom des informations de jeu de données correspondantes lors de l'événement d'exécution. OpenLineage

Atténuation : comme il n'existe aucun moyen déterministe de savoir si le nom donné appartient à un cluster ou à un groupe de travail, nous utilisons l'heuristique suivante :

- Si le « nom » correspondant à l'ensemble de données contient `redshift-serverless.amazonaws.com` « », nous utilisons redshift-serverless dans l'ARN, sinon la valeur par défaut est « redshift ».
- Ce qui précède signifie que les alias sur les noms de groupes de travail ne fonctionneront pas.

Cause première : les ensembles de données en amont ne sont pas correctement liés pour les actifs personnalisés.

Atténuation : assurez-vous de renseigner le `sourceIdentifier` champ sur la ressource en invoquant`CreateAsset`/`CreateAssetRevision`qui correspond au nœud `sourceIdentifier` du jeu de données (qui serait<namespace>/<name>pour les nœuds personnalisés).

Quotas pour Amazon DataZone

Votre AWS compte dispose de quotas par défaut, anciennement appelés limites, pour chaque AWS service. Sauf indication contraire, chaque quota est spécifique à une région.

Amazon DataZone applique les quotas et limites suivants.

DataZone Quotas Amazon

Ressource	Description	Valeur
Types de ressources de données	Nombre maximal de types de ressources de données pouvant être créés dans un DataZone domaine	1 000
Actifs de données	Le nombre maximum de ressources de données pouvant être créées dans un DataZone domaine Amazon	1 million
Glossaires	Le nombre maximum de glossaires commerciaux que vous pouvez créer dans un domaine	1 000
Termes du glossaire commercial	Le nombre maximum de termes du glossaire commercial que vous pouvez créer dans un domaine	10 000
Environnements d'un domaine	Le nombre maximum d'environnements dans un DataZone domaine Amazon	500

Ressource	Description	Valeur
Nombre de filtres d'actifs par actif	Le nombre maximum de filtres d'actifs par DataZone actif Amazon	100
Nombre de filtres par abonnement	Le nombre maximum de filtres par DataZone abonnement Amazon	5
Unités de domaine dans un domaine	Le nombre maximum d'unités de domaine dans un DataZone domaine Amazon	500
Niveaux hiérarchiques dans une unité de domaine	Le nombre maximal de niveaux hiérarchiques pour une unité de domaine	5
Subventions par politique et par unité de domaine	Le nombre maximum de subventions par politique et par unité de domaine	20
Produits de données	Le nombre maximum de produits de données pouvant être créés dans un DataZone domaine	500 000
La source de données s'exécute	Le nombre maximal d'exécutions de source de données par source de données par jour	25

Limites de débit des DataZone API Amazon

Le tableau suivant décrit les limites de débit pour Amazon DataZone APIs. Ces limites s'appliquent par AWS compte et par région.

Limites de débit des DataZone API Amazon

« Hello, World! »	Limites de débit d'API
CreateGlossary	5 transactions par seconde (TPS)
UpdateGlossary	20 TPS
GetGlossary	20 TPS
DeleteGlossary	20 TPS
UpdateGlossaryTerm	20 TPS
DeleteGlossaryTerm	20 TPS
CreateAsset	20 TPS
ListAssetRevisions	20 TPS
CreateAssetRevision	20 TPS
DeleteAsset	20 TPS
CreateDataProduct	20 TPS
ListDataProductRevisions	20 TPS
CreateDataProductRevision	20 TPS
DeleteDataProduct	20 TPS
CreateAssetType	20 TPS
DeleteAssetType	20 TPS
CreateFormType	20 TPS
DeleteFormType	20 TPS
Rechercher	20 TPS
SearchTypes	20 TPS

« Hello, World! »	Limites de débit d'API
AcceptPredictions	20 TPS
RejectPredictions	20 TPS
AcceptSubscriptionRequest	3 TPS
CancelSubscription	3 TPS
CreateSubscriptionGrant	3 TPS
CreateSubscriptionRequest	3 TPS
GetSubscriptionEligibility	30 TPS
DeleteSubscriptionGrant	3 TPS
DeleteSubscriptionRequest	3 TPS
DeleteSubscriptionTarget	3 TPS
GetSubscription	8 CUILLÈRES À CAFÉ
GetSubscriptionGrant	8 CUILLÈRES À CAFÉ
GetSubscriptionRequestDetails	8 CUILLÈRES À CAFÉ
ListSubscriptionGrants	8 CUILLÈRES À CAFÉ
ListSubscriptionRequests	8 CUILLÈRES À CAFÉ
ListSubscriptions	8 CUILLÈRES À CAFÉ
ListSubscriptionTargets	8 CUILLÈRES À CAFÉ
RejectSubscriptionRequest	3 TPS
RevokeSubscription	3 TPS
UpdateSubscriptionRequest	3 TPS

« Hello, World! »	Limites de débit d'API
UpdateSubscriptionTarget	3 TPS
CreateProjectProfile	3 TPS
UpdateProjectProfile	3 TPS
CreateDomain	8 CUILLÈRES À CAFÉ
UpdateDomain	8 CUILLÈRES À CAFÉ
CreateProject	3 TPS
UpdateProject	3 TPS
DeleteProject	3 TPS
ListProjects	8 CUILLÈRES À CAFÉ
CreateProjectMembership	3 TPS
ListProjectMemberships	8 CUILLÈRES À CAFÉ
DeleteProjectMembership	3 TPS
CreateEnvironment	3 TPS
DeleteEnvironment	3 TPS
UpdateEnvironment	3 TPS
ListEnvironments	8 CUILLÈRES À CAFÉ
GetEnvironment	8 CUILLÈRES À CAFÉ
GetEnvironmentCredentials	8 CUILLÈRES À CAFÉ
CreateEnvironmentProfile	8 CUILLÈRES À CAFÉ
ListEnvironmentProfiles	8 CUILLÈRES À CAFÉ

« Hello, World! »	Limites de débit d'API
ListEnvironmentBlueprints	8 CUILLÈRES À CAFÉ
PutEnvironmentBlueprintConfiguration	10 TPS
StartMetadataGenerationRun	10 TPS
CancelMetadataGenerationRun	20 TPS
CreateDomainUnit	20 TPS
AddPolicyGrant	20 TPS
AddEntityOwner	20 TPS
CreateRule	20 TPS
UpdateRule	20 TPS
CreateDataSource	20 TPS
UpdateDataSource	20 TPS
DeleteDataSource	20 TPS
ListDataSources	20 TPS
SearchListings	16 CUILLÈRES À CAFÉ
StartDataSourceRun	20 TPS
UpdateDataSourceRunActivities	20 TPS
PostLineageEvent	20 TPS
CreateConnection	20 TPS
UpdateConnection	20 TPS
GetConnection	20 TPS

« Hello, World! »	Limites de débit d'API
ListConnections	20 TPS
DeleteConnection	20 TPS
CreateListingChangeSet	20 TPS

Historique du document pour le guide de DataZone l'utilisateur Amazon

Le tableau suivant décrit les versions de documentation pour Amazon DataZone.

Modification	Description	Date
AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary - mises à jour des politiques	Mises à jour des politiques du AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary. Ajout d'une instruction Deny pour l'`sagemaker:UpdateNotebookInstanceLifecycleConfig` action visant à restreindre cette opération à privilèges élevés. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	11 mars 2026
AmazonDataZoneDomainExecutionRolePolicy - mises à jour des politiques	Mises à jour des politiques relatives à l'AmazonDataZoneDomainExecutionRolePolicyajout d'autorisations à l'`QueryGraph` action afin de prendre en charge les fonctionnalités de recherche d'entités basées sur des graphes. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	25 février 2026

AmazonDataZoneGlueManageAccessRolePolicy - mises à jour des politiques	Mises à jour de la AmazonDataZoneGlueManageAccessRolePolicypolitique concernant l'ajout d'autorisations à l'`GetConnection` action afin de prendre en charge la capture du lignage des données pour les sources de données basées sur la connexion de AWS Glue. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	30 juillet 2025
AmazonDataZoneFullAccess - mises à jour des politiques	Mises à jour des politiques relatives à la AmazonDataZoneFullAccessgénéralisation de la portée SecretsManager `create` et `tag` des autorisations pour les nouveaux domaines qui auront le format de `dzd-` au lieu de. `dzd_..` Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	23 juillet 2025

AmazonDataZoneFullAccess - mises à jour des politiques	Mises à jour des AmazonDataZoneFullAccesspolitiques visant à permettre à la console d'associer ou de mettre à jour les autorisations AWS gérées dans les partages de ressources AWS RAM. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	22 mai 2025
AmazonDataZoneGlueManageAccessRolePolicy - mises à jour des politiques	Mises à jour de la politique AmazonDataZoneGlueManageAccessRolePolicy: le rôle d'utilisateur DataZone du projet Amazon est utilisé comme rôle de transfert de données pour les tables fédérées. Cette mise à jour `datazone_usr_role*` complète la `iam:PassRole` déclaration, permettant d'utiliser le rôle d'utilisateur du projet à cette fin. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	21 mai 2025

AmazonDataZoneSageMakerProvisioningRolePolicy - mises à jour des politiques	Mises à jour de la politique AmazonDataZoneSageMakerProvisioningRolePolicy - ajout d'un support à l'`glue:GetConnection` action. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	2 janvier 2025
AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary - mises à jour des politiques	Mises à jour de la politique AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary: cette modification ajoute la limite `sagemaker:AddTags` d'autorisation afin de permettre DataZone à Amazon d'appeler `CreateUserProfile` avec succès avec les balises nécessaires. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	3 décembre 2024

AmazonDataZoneSageMakerAccess, et AmazonDataZoneGlueManageAccessRolePolicy - mises à jour des politiques	Mises à jour des politiques relatives à AmazonDataZoneFullAccessAmazonDataZoneSageMakerAccess, et AmazonDataZoneGlueManageAccessRolePolicy- pour permettre la prise en charge de l'expérience Amazon SageMaker Unified Studio. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	3 décembre 2024
AmazonDataZoneDomainExecutionRolePolicy et AmazonDataZoneFullUserAccess - mises à jour des politiques	Mises à jour des politiques pour permettre la prise en charge des règles d'application des métadonnées pour les demandes d'abonnement. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	20 novembre 2024

Amazon DataZone lance des règles d'application des métadonnées pour les demandes d'abonnement	Les nouvelles règles d'application des métadonnées pour les demandes d'abonnement sur Amazon DataZone renforcent la gouvernance des données en permettant aux propriétaires d'unités de domaine d'établir des exigences claires en matière de métadonnées pour les consommateurs de données, en rationalisant les demandes d'accès et en améliorant la gouvernance des données. Cette fonctionnalité permet aux organisations de s'aligner sur les normes de métadonnées de l'organisation, de mettre en œuvre des flux de travail personnalisés et de fournir une expérience d'accès aux données cohérente et gouvernée. Pour plus d'informations, consultez la section Règles d'application des métadonnées pour les demandes d'abonnement.	20 novembre 2024

AmazonDataZoneRedshiftGlueProvisioningPolicy - mises à jour des politiques	Ajout `iam:DeletePolicyVersion` pour permettre aux utilisateurs de supprimer les versions des politiques créées avec`datazone*` . Cela permet de débloquer les utilisateurs qui doivent mettre à jour leur politique de rôle d'utilisateur dans leur environnement. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	22 octobre 2024

AWS CloudFormation support pour un plan AWS de service personnalisé	Amazon DataZone a ajouté la AWS CloudFormation prise en charge du plan AWS de service personnalisé. Cette nouvelle fonctionnalité vous permet d' AWS CloudFormation automatiser la création d'environnements dans Amazon DataZone. Grâce à des plans personnalisés, les administrateurs peuvent désormais intégrer facilement Amazon à leurs pipelines de données existants DataZone en utilisant les rôles IAM existants pour publier les actifs de données dans le DataZone catalogue Amazon, en facilitant le partage régi de ces actifs et en améliorant la gouvernance dans l'ensemble de l'infrastructure. Pour plus d'informations, consultez la référence DataZone des types de ressources Amazon.	12 septembre 2024

Unités de domaine	Amazon DataZone introduit un ensemble de nouvelles fonctionnalités de gouvernance des données appelées unités de domaine et politiques d'autorisation qui permettent aux clients de créer une organisation unit/team au niveau de l'entreprise et de gérer des politiques en fonction de leurs besoins commerciaux. Avec l'ajout d'unités de domaine, les utilisateurs peuvent organiser, créer, rechercher et trouver des actifs de données et des projets associés à des unités commerciales ou à des équipes. Grâce aux politiques d'autorisation, ces utilisateurs d'unités de domaine peuvent définir des politiques d'accès pour créer des projets, des glossaires et utiliser des ressources informatiques au sein d'Amazon DataZone.	5 août 2024

Produits de données	Amazon DataZone lance des produits de données, qui permettent de regrouper les actifs de données dans des packages autonomes bien définis, adaptés à des cas d'utilisation professionnels spécifiques. Par exemple, un produit de données d'analyse marketing peut regrouper divers actifs de données, tels que les données de campagnes marketing, les données de pipeline et les données clients. Grâce aux produits de données, les clients peuvent simplifier les processus de découverte et d'abonnement, en les alignant sur les objectifs commerciaux et en réduisant la redondanc e dans la gestion des actifs individuels.	5 août 2024

AmazonDataZoneDomainExecutionRolePolicy et AmazonDataZoneFullUserAccess - mises à jour des politiques	Mises à jour de la politique AmazonDataZoneDomainExecutionRolePolicyet AmazonDataZoneFullUserAccesspour permettre la prise en charge des nouvelles APIs règles utilisées pour créer et gérer les unités de DataZone domaine et les produits de données Amazon. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	5 août 2024

Contrôle précis des accès	Amazon DataZone a introduit un contrôle d'accès précis, qui vous permet de contrôler de manière précise vos actifs de données dans le catalogue de données commerciales d'Amazon, à travers les lacs DataZone de données et les entrepôts de données. Grâce à cette nouvelle fonctionnalité, les propriétaires de données peuvent désormais restreind re l'accès à des enregistr ements de données spécifiqu es au niveau des lignes et des colonnes, au lieu d'accorder l'accès à l'ensemble des actifs de données. Par exemple, si vos données contienne nt des colonnes contenant des informations sensibles telles que des informations personnelles identifiables (PII), vous pouvez limiter l'accès aux seules colonnes nécessair es, afin de garantir la protectio n des informations sensibles tout en autorisant l'accès aux données non sensibles. De même, vous pouvez contrôler l'accès au niveau des lignes, en permettant aux utilisateurs de ne voir que les enregistr ements correspondant à leur rôle ou à leur tâche.	2 juillet 2024

AmazonDataZoneGlue ManageAccessRolePolicy - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneGlue ManageAccessRolePolicy- Amazon DataZone ajoute des autorisations IAM qui sont utilisées pour des fonctionn alités de contrôle d'accès précises afin de limiter l'octroi d'autorisations à Lake Formation. Pour plus d'informa tions, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	2 juillet 2024

Lignage des données	Amazon DataZone lance le lignage des données en version préliminaire, aidant les clients à visualiser les événements de lignage à partir de systèmes OpenLineage compatibles ou via une API et à suivre le mouvement des données de la source à la consommation. Grâce à la OpenLineage compatibilité DataZone d'Amazon APIs, les administrateurs de domaines et les producteurs de données peuvent capturer et stocker des événements de lignage autres que ceux disponibl es sur Amazon DataZone, y compris les transformations dans Amazon S3, AWS Glue et d'autres services. En outre, DataZone les versions d'Amazon font le suivi de chaque événement, ce qui permet aux utilisateurs de visualiser le lignage à tout moment ou de comparer les transformations liées à l'histori que d'un actif ou d'une tâche. Cette généalogie historique permet de mieux comprendr e l'évolution des données, ce qui est essentiel pour le dépannage, l'audit et la validation de l'intégrité des actifs de données.	27 juin 2024

AmazonDataZoneExecutionRolePolicy et AmazonDataZoneFullUserAccess - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneExecutionRolePolicyet AmazonDataZoneFullUserAccesspour permettre la prise en charge du lignage des données et du contrôle d'accès détaillé. APIs Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	27 juin 2024

Plan AWS de service personnalisé	Avec les plans de AWS service personnalisés, si vous disposez de AWS ressource s existantes, notamment des rôles IAM, des lacs de données, des maillages de données, des compartiments Amazon S3 et des clusters Amazon Redshift, vous pouvez désormais spécifier des autorisations pour ces ressources existantes à l'aide de votre propre rôle IAM personnalisé, afin que vos DataZone utilisateurs Amazon puissent tirer parti de la publication et de l'abonnem ent pour partager et gérer ces ressources. Avec des plans AWS de service personnal isés, DataZone les administr ateurs Amazon peuvent configurer des environne ments AWS de service en utilisant leurs propres rôles personnalisés. Ils peuvent configurer des liens d'actions pour ces environnements de AWS service et fournir ainsi un accès fédéré à toutes leurs AWS ressource s existantes. Ils peuvent également configurer des cibles d'abonnement et des sources de données dans ces environnements AWS	17 juin 2024

	de service personnalisés. Les administrateurs peuvent configurer des environnements de AWS service dans leur propre compte de DataZone domaine Amazon ou dans tout compte associé à partir duquel ils souhaitent publier, s'abonner, découvrir ou gérer des données.	
AmazonDataZoneGlue ManageAccessRolePolicy - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneGlue ManageAccessRolePolicy qui ajoute les autorisations IAM requises pour la fonctionn alité d'auto-abonnement d'Amazon DataZone afin de limiter les autorisations accordées lors de la formation des lacs. Grâce à la fonctionn alité d'auto-abonnement, les autorisations de formation de lacs ne peuvent être accordées qu'aux ressources étiquetées. Pour plus d'informa tions, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	14 juin 2024

AmazonDataZoneFullAccess - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneFullAccess qui permet à la console DataZone de gestion Amazon de créer des secrets au nom de l'utilisateur avec des balises de domaine et de projet. Inclut également l'`ram:ListResourceSharePermissions` action permettant aux administrateurs du compte du propriétaire du domaine de consulter le statut d'association des comptes associés. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	14 juin 2024
AmazonDataZoneDomainExecutionRolePolicy - mise à jour de la politique	Mise à jour de la politique AmazonDataZoneDomainExecutionRolePolicy qui ajoute de nouvelles APIs fonctionnalités à Amazon DataZone qui permettent aux utilisateurs de configurer des actions pour leurs DataZone environnements Amazon. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	14 juin 2024

Améliorations apportées à la création de sources de données	Amazon DataZone a apporté des améliorations au flux de création des sources de données afin de simplifier la gestion des accès pour les producteurs de données. Avec ces mises à jour, lorsqu'un producteur de données crée une source de données pour publier ses actifs AWS Glue et Amazon Redshift, Amazon DataZone accorde des autorisations de lecture seule aux membres du projet. Lors de la création d'une source de données AWS Glue, Amazon accorde DataZone automatiquement des autorisations « en lecture seule » au rôle IAM de l'environnement utilisé pour créer la source de données, permettant ainsi l'accès à toutes les tables des bases de données Glue AWS associées. De même, pour les sources de données Amazon Redshift, Amazon DataZone accorde un accès « en lecture seule » à toutes les tables des schémas Amazon Redshift utilisés dans la source de données.	10 juin 2024

Intégration avec Amazon SageMaker	Amazon DataZone lance l'intégration avec Amazon SageMaker pour aider les producteurs de données et les consommateurs à passer facilement SageMaker à Amazon pour collaborer sur des projets d'apprentissage automatique (ML) tout en renforçant la gouvernan ce de l'accès aux données et aux actifs de machine learning. Grâce à la nouvelle intégration intégrée entre Amazon DataZone et Amazon SageMaker, les consommat eurs et les producteurs de données peuvent rationaliser la gouvernance du machine learning lors de la configura tion de l'infrastructure, collaborer sur des initiativ es commerciales et gérer facilement les données et les actifs de machine learning.	6 mai 2024

AmazonDataZoneSageMakerProvisioningRolePolicy - nouvelle politique	La nouvelle politique appelée AmazonDataZoneSageMakerProvisioningRolePolicy accorde à Amazon DataZone les autorisations nécessaires pour interagir avec Amazon SageMaker. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	30 avril 2024
AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary - nouvelle limite d'autorisations	Nouvelle limite d'autorisations appelée AmazonDataZoneSageMakerEnvironmentRolePermissionsBoundary . Lorsque vous créez un SageMaker environnement Amazon via le portail de DataZone données Amazon, Amazon DataZone applique cette limite d'autorisations aux rôles IAM produits lors de la création de l'environnement. La limite des autorisations limite l'étendue des rôles créés par Amazon DataZone et de tous les rôles que vous ajoutez. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	30 avril 2024

AmazonDataZoneSageMakerAccess - nouvelle politique	Une nouvelle politique appelée AmazonDataZoneSageMakerAccess accorde à Amazon DataZone les autorisations requises pour accorder aux utilisateurs l'accès à diverses ressources de l' SageMaker environnement Amazon. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	30 avril 2024
AmazonDataZoneFullAccess - mise à jour de la politique	Une mise à jour de la AmazonDataZoneFullAccesspolitique qui ajoute un accès à l'`DescribeSecurityGroups` action pour améliorer la facilité d'utilisation pour les administrateurs de compte qui configurent des plans dans la console et des `GetPolicy` actions pour aider à récupérer des informations sur la politique gérée spécifiée. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	30 avril 2024

Mode d'accès hybride à Lake Formation	Amazon DataZone a introduit une intégration avec AWS le mode d'accès hybride Lake Formation. Cette intégrati on vous permet de publier et de partager facilemen t vos tables AWS Glue via Amazon DataZone, sans avoir à les enregistrer au préalable dans AWS Lake Formation. Pour commencer , les administrateurs activent le paramètre d'enregistrement de la localisation des données dans le `DefaultDataLake` plan de la DataZone console Amazon. Ensuite, lorsqu'un consommateur de données s'abonne à une table AWS Glue gérée via des autorisat ions IAM, Amazon enregistr e d' DataZoneabord les emplacements Amazon S3 de cette table en mode hybride, puis accorde l'accès au consommateur de données en gérant les autorisations sur la table via AWS Lake Formation. Cela garantit que les autorisations IAM disponibl es continuent d'exister avec les autorisations AWS Lake Formation récemment accordées, sans perturber les flux de travail existants. Pour plus d'informations, consultez	3 avril 2024

	la section DataZone Intégration d'Amazon au mode hybride de AWS Lake Formation.	
Qualité des données	Amazon DataZone lance l'intégration avec AWS Glue Data Quality et propose APIs d'intégrer des indicateurs de qualité des données issus de solutions de qualité des données tierces. La nouvelle intégration vous permet de publier automatiquement les scores de AWS Glue Data Quality dans le catalogue de données Amazon DataZone Business. Amazon DataZone APIs peut être utilisé pour ingérer des indicateurs de qualité provenant de sources tierces. Une fois publiées, les consommateurs de données peuvent facilement recherche r des actifs de données, consulter des indicateurs de qualité détaillés et identifie r les contrôles et les règles défaillants, ce qui permet aux entreprises de prendre des décisions éclairées. Pour plus d'informations, consultez la section Qualité des données sur Amazon DataZone.	3 avril 2024

AmazonDataZoneS3Manage- - nouveau rôle <region>< domainId>	Nouveau rôle appelé AmazonDataZoneS3Ma nage-<region>, <domainId >utilisé lorsqu'Amazon DataZone appelle AWS Lake Formation pour enregistrer un emplacement Amazon Simple Storage Service (Amazon S3). AWS Lake Formation assume ce rôle lors de l'accès aux données à cet emplaceme nt. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	1er avril 2024
AmazonDataZoneGlue ManageAccessRolePolicy - Mise à jour de la politique	Mise à jour du AmazonDat aZoneGlueManageAcc essRolePolicypour permettre la prise en charge des autorisations permettant DataZone à Amazon d'activer les autorisations de publication et d'accès aux données. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	1er avril 2024

AmazonDataZoneDomainExecutionRolePolicy et AmazonDataZoneFullUserAccess - Mise à jour de la politique	Mise à jour du AmazonDataZoneDomainExecutionRolePolicyet AmazonDataZoneFullUserAccesspour permettre la prise en charge de l'`CancelMetadataGenerationRun` API. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	29 mars 2024
AmazonDataZoneFullAccess - Mise à jour de la politique	Amazon DataZone a annoncé la mise à disposition générale de la nouvelle fonctionnalité générative basée sur l'IA afin d'améliorer la découverte des données, leur compréhension et leur utilisation en enrichissant le catalogue de données commerciales. En un seul clic, les producteurs de données peuvent générer des descriptions et un contexte complets des données commerciales, mettre en évidence les colonnes pertinentes et inclure des recommandations sur des cas d'utilisation analytiques. Le lancement ajoute un support APIs que les producteurs de données peuvent utiliser pour générer par programmation des descriptions d'actifs.	27 mars 2024

AmazonDataZoneFullAccess - Mise à jour de la politique	Amazon DataZone a apporté plusieurs améliorations à son intégration avec Amazon Redshift, simplifiant le processus de publication et d'abonnement aux tables et aux vues Amazon Redshift. Ces mises à jour simplifient l'expérience des producteurs de données et des consommateurs, en leur permettant de créer rapidement des environnements d'entrepôt de données à l'aide d'informations d'identification et de paramètres de connexion préconfigurés fournis par leurs DataZone administrateurs Amazon. En outre, ces améliorations permettent aux administrateurs de mieux contrôler qui peut utiliser les ressources de leurs AWS comptes et des clusters Amazon Redshift, et dans quel but.	21 mars 2024

AmazonDataZoneFullAccess - Mise à jour de la politique	Mise à jour du `AmazonDataZoneFullAccess` pour permettre aux utilisateurs de choisir leurs secrets, leurs clusters, leurs VPC et leurs sous-réseaux dans la console de DataZone gestion Amazon plutôt que de les saisir dans une zone de texte. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	13 mars 2024
AmazonDataZoneDomainExecutionRolePolicy - Mise à jour de la politique	Mise à jour du AmazonDataZoneDomainExecutionRolePolicypour permettre la prise en charge de l' ListEnvironmentBlueprintConfigurationSummaries API requise pour créer des profils d'environnement en identifiant quels plans sont activés dans quel compte et dans quelle région. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	1er février 2024

Améliorations apportées à l'utilisation de Cloud Formation	Les utilisateurs d'Amazon DataZone peuvent désormais en tirer parti AWS CloudFormation pour modéliser et gérer efficacement une suite de DataZone ressources Amazon. Cette approche facilite le provisionnement cohérent des ressources, tout en permettant la gestion du cycle de vie par le biais de pratiques liées à l'infrastructure en tant que code. Grâce aux modèles personnalisés, vous pouvez définir avec précision les ressources dont vous avez besoin et leurs interdépendances. Pour plus d'informations, consultez la référence DataZone des types de ressources Amazon.	18 janvier 2024

Actifs personnalisés	La prise en charge des actifs personnalisés permet DataZone à Amazon de cataloguer les actifs via le portail de données pour les données non structurées, notamment les tableaux de bord, les requêtes et les modèles, ce qui vous permet d'ajouter plus facilement des actifs personnalisés directeme nt dans le portail de données, en plus du support d'API précédemment disponible. La possibilité de créer, de mettre à jour et de publier des ressources personnal isées sur Amazon DataZone vous permet de partager, de rechercher, de vous abonner à tout type d'actif et de créer un flux de travail qui assure la gouvernance de ces actifs. Pour plus d'informations, consultez la section Création de types de ressources personnalisés.	5 janvier 2024

Ajouter les principaux IAM en tant que membres du projet	Vous pouvez désormais ajouter des responsables IAM en tant que membres du projet, même s'ils ne se sont pas encore connectés à Amazon DataZone (exigence précédente). Une fois qu'un administrateur de domaine ou un administrateur informati que a ajouté `iam:GetUs er` et `iam:GetRole` au rôle d'exécution du domaine, les propriétaires de projet peuvent ajouter les principaux IAM en tant que membres en fournissant simplement le nom de ressource Amazon (ARN) du rôle IAM ou de l'utilisateur IAM. Le principal IAM doit toujours disposer des autorisat ions IAM requises pour accéder à Amazon DataZone et celles-ci peuvent être configurées dans la console IAM. Pour plus d'informations, voir Ajouter des membres à un projet.	5 janvier 2024

Supprimer le domaine	Supprimer le domaine est une fonctionnalité qui vous permet de supprimer plus facilemen t vos domaines. Vous pouvez désormais procéder à la suppression du domaine même s'il n'est pas vide (car il contient des projets, des environnements, des actifs, des sources de données, etc.). Pour plus d'informa tions, consultez Supprimer des DataZone domaines Amazon.	27 décembre 2023

Mode hybride de Lake Formation	Amazon DataZone a ajouté la prise en charge du mode hybride AWS Lake Formation . Grâce à ce support, si vous publiez une table AWS Glue sur Amazon DataZone dont l'emplacement AWS S3 est enregistré dans Lake Formation en mode hybride, Amazon DataZone traite cette table comme un actif géré et peut gérer les subventions d'abonnement à cette table. Avant la sortie de cette fonctionnalité, Amazon DataZone traitait ce tableau comme un actif non géré, c'est-à-dire DataZone qu'Amazon ne serait pas en mesure d'accorder des abonnements à ce tableau. Pour plus d'informations, consultez Configurer les autorisations de Lake Formation pour Amazon DataZone.	22 décembre 2023

Conformité à la loi HIPAA	Amazon DataZone est désormais conforme à la loi américaine HIPAA (Health Insurance Portability and Accountability Act) de 1996. Pour consulter la liste des AWS services conformes à la loi HIPAA, voir https://aws.amazon.com/compliance/hipaa-eligible-services-reference/.	14 décembre 2023
AmazonDataZoneGlueManageAccessRolePolicy - Mise à jour de la politique	Mise à jour du AmazonDataZoneGlueManageAccessRolePolicypour permettre la prise en charge du mode hybride AWS Lake Formation . Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	14 décembre 2023
AmazonDataZoneFullUserAccess et AmazonDataZoneDomainExecutionRolePolicy - Mises à jour des politiques	Amazon a DataZone mis à jour les politiques AmazonDataZoneFullUserAccesset les AmazonDataZoneDomainExecutionRolePolicypolitiques afin de prendre en charge la fonctionnalité de description des données basée sur l'IA générative d'Amazon. DataZone Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	28 novembre 2023

Recommandations de l'IA	AWS annonce la prévisualisation d'une nouvelle fonctionnalité générative basée sur l'IA dans Amazon DataZone afin d'améliorer la découverte des données, leur compréhension et leur utilisation en enrichissant le catalogue de données commerciales. En un seul clic, les producteurs de données peuvent générer des descriptions et un contexte complets des données commerciales, mettre en évidence les colonnes pertinentes et inclure des recommandations sur des cas d'utilisation analytiques. Grâce aux recommandations de l'IA pour les descriptions sur Amazon DataZone, les consommateurs de données peuvent identifier les tables de données et les colonnes nécessaires à l'analyse, ce qui améliore la découvrabilité des données et réduit les back-and-forth communications avec les producteurs de données. La version préliminaire est disponible dans DataZone les domaines Amazon fournis dans les AWS régions suivantes : USA Est (Virginie du Nord), USA Ouest (Oregon). Pour plus	28 novembre 2023

	d'informations, consultez la section Utilisation de l'apprentissage automatique et de l'IA générative.	
DefaultDataLake plan	Amazon DataZone a ajouté une amélioration au DefaultDataLake plan qui vous permet de mieux contrôler qui peut publier quelles données depuis votre AWS compte. Deux modifications majeures ont été introduites avec le lancement de cette fonctionnalité.	20 novembre 2023
AmazonDataZoneEnvironmentRolePermissionsBoundary - Mise à jour de la politique	Amazon DataZone a mis à jour la politique AmazonDataZoneEnvironmentRolePermissionsBoundarygérée qui consiste en une `athena:GetQueryResultsStream` autorisation supplémentaire limitée à la `ResourceTag` condition . Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	17 novembre 2023

AmazonDataZoneRedshiftManageAccessRolePolicy - Mise à jour de la politique	Amazon a DataZone mis à jour la AmazonDataZoneRedshiftManageAccessRolePolicypolitique en supprimant la vérification de l'identifiant de l'organisation pour `redshift:AssociateDataShareConsumer` cette action. Cela vous permet de partager les ressources entre les AWS organisations. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	16 novembre 2023
Publication générale du guide de l'utilisateur	Version de disponibilité générale (GA) du guide de DataZone l'utilisateur Amazon.	15 octobre 2023
AmazonDataZoneFullUserAccess - Mise à jour de la politique	Amazon a DataZone mis à jour la AmazonDataZoneFullUserAccesspolitique qui accorde un accès complet à Amazon DataZone, mais elle n'autorise pas la gestion des domaines, des utilisateurs ou des comptes associés. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	2 octobre 2023

AmazonDataZonePreviewConsoleFullAccess - politique déconseillée	Amazon DataZone a déconseillé le AmazonDataZonePreviewConsoleFullAccess. Pour plus d'informations, consultez les DataZone mises à jour des politiques gérées par Amazon. AWS	29 septembre 2023
AmazonDataZonePortalFullAccessPolicy - politique déconseillée	Amazon DataZone a déconseillé le AmazonDataZonePortalFullAccessPolicy. Pour plus d'informations, consultez les DataZone mises à jour des politiques gérées par Amazon. AWS	29 septembre 2023

AmazonDataZoneDomainExecutionRolePolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneDomainExecutionRolePolicy. Il s'agit de la politique par défaut pour le rôle DataZone de service Amazon `AmazonDataZoneDomainExecutionRole`. Ce rôle est utilisé par Amazon DataZone pour cataloguer, découvrir, gérer, partager et analyser les données du DataZone domaine Amazon. Vous pouvez joindre la `AmazonDataZoneDomainExecutionRolePolicy` politique à votre`AmazonDataZoneDomainExecutionRole`. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	25 septembre 2023
AmazonDataZoneCrossAccountAdmin - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneCrossAccountAdminqui permet aux utilisateurs de travailler avec Amazon DataZone et ses comptes associés. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	19 septembre 2023

AmazonDataZoneRedshiftManageAccessRolePolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneRedshiftManageAccessRolePolicyqui accorde des autorisations permettant DataZone à Amazon d'activer les autorisations de publication et d'accès aux données. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	12 septembre 2023
AmazonDataZoneRedshiftGlueProvisioningPolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneRedshiftGlueProvisioningPolicyqui accorde à Amazon DataZone les autorisations nécessaires pour interagir avec les sources de données prises en charge. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	12 septembre 2023

AmazonDataZoneGlueManageAccessRolePolicy - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique intitulée « AmazonDataZoneGlueManageAccessRolePolicyaccorde à Amazon l' DataZoneautorisation de publier les données AWS Glue dans le catalogue ». Cela donne également à Amazon l' DataZoneautorisation d'accorder ou de révoquer l'accès aux ressources publiées par AWS Glue dans le catalogue. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	12 septembre 2023
AmazonDataZoneFullUserAccess - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneFullUserAccessqui accorde un accès complet à Amazon DataZone via le portail de données. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	12 septembre 2023

AmazonDataZoneFullAccess - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneFull Accessqui fournit un accès complet à Amazon DataZone via la console AWS de gestion. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	12 septembre 2023
AmazonDataZoneEnvironmentRolePermissionsBoundary - Nouvelle politique	Amazon DataZone a ajouté une nouvelle politique appelée AmazonDataZoneEnvironmentRolePermissionsBoundaryqui limite le principal IAM provisionné auquel il est attaché. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	12 septembre 2023
Mise à jour de politique gérée	Mises à jour de la politique AmazonDataZonePreviewConsoleFullAccess gérée. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	13 juin 2023

Mise à jour de politique gérée	Mises à jour de la politique AmazonDataZoneProjectDeploymentPermissionsBoundary gérée. Pour plus d'informations, consultez les DataZone mises à jour des politiques AWS gérées par Amazon.	3 avril 2023
???	Version initiale du guide de l'utilisateur Amazon DataZone (version préliminaire).	29 mars 2023

Les traductions sont fournies par des outils de traduction automatique. En cas de conflit entre le contenu d'une traduction et celui de la version originale en anglais, la version anglaise prévaudra.