Amazon Elastic Compute Cloud

Guide de l'utilisateur pour les instances Linux

Amazon Elastic Compute Cloud: Guide de l'utilisateur pour les instances Linux

Copyright © 2020 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Présentation d'Amazon EC2 1
Fonctionnalités d'Amazon EC2 1
Comment démarrer avec Amazon EC2 1
Services connexes 2
Accès à Amazon EC2 3
Tarification pour Amazon EC2 4
Conformité DSS PCI 4
Instances et AMI 4
Instances 5
AMI 7
Régions et zones 7
Régions 8
Zones de disponibilité 13
Zones locales 15
Zones de longueur d'onde 18
AWS Outposts 20
volume du périphérique racine 21
Concepts du stockage de périphérique racine 21
Choix d'une AMI par type de périphérique racine 23
Détermination du type de périphérique racine de votre instance 24
Modification du volume racine pour qu'il persiste 25
Configuration de 28
Inscrivez-vous à AWS 28
Création d'une paire de clés 28
Création d'un groupe de sécurité 30
Didacticiel de démarrage 32
Présentation 32
Prérequis 33
Étape 1 : Lancement d'une instance 33
Étape 2 : Connexion à l'instance 34
Étape 3 : Nettoyage de votre instance 35
Étapes suivantes 35
Bonnes pratiques 36
Didacticiels 38
Installation d'un serveur LAMP (Amazon Linux 2) 38
Étape 1 : Préparer le serveur LAMP 39
Étape 2 : Tester votre serveur LAMP 42
Étape 3 : Sécuriser le serveur de base de données 43
Étape 4 : (Facultatif) Installer phpMyAdmin 44
Dépannage 47
Voir aussi 47
Installation d'un serveur LAMP (Amazon Linux AMI) 48
Étape 1 : Préparer le serveur LAMP 48
Étape 2 : Tester votre serveur LAMP 52
Étape 3 : Sécuriser le serveur de base de données 53
Étape 4 : (Facultatif) Installer phpMyAdmin 55
Dépannage 58
Voir aussi 59
Didacticiel : Hébergement d'un blog WordPress 59
Prérequis 60
Installer WordPress 60
Étapes suivantes 67
Aide! Mon nom DNS public a changé et mon blog ne fonctionne plus 68
Didacticiel : Configurer SSL/TLS sur Amazon Linux 2 69

Prérequis ... 69
Étape 1 : Activer TLS sur le serveur ... 70
Étape 2 : Obtenir un certificat signé par une autorité de certification (CA) ... 72
Étape 3 : Tester et renforcer la configuration de sécurité ... 77
Dépannage ... 80
Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux 2 ... 81
Didacticiel : Configurer SSL/TLS sur Amazon Linux ... 84
Prérequis ... 85
Étape 1 : Activer TLS sur le serveur ... 85
Étape 2 : Obtenir un certificat signé par une autorité de certification (CA) ... 87
Étape 3 : Tester et renforcer la configuration de sécurité ... 92
Dépannage ... 94
Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux ... 95
Didacticiel : Augmenter la disponibilité des applications ... 98
Prérequis ... 99
Dimensionner votre application et en répartir la charge ... 100
Test de votre équilibreur de charge ... 101
Amazon Machine Images ... 103
Utilisation d'une AMI ... 103
Création de votre propre AMI ... 104
Achat, partage et vente d'AMI ... 104
Annulation de l'inscription de votre AMI ... 105
Amazon Linux 2 et Amazon Linux AMI ... 105
Types d'AMI ... 105
Autorisations de lancement ... 105
Stockage pour le périphérique racine ... 106
Types de virtualisation ... 108
Rechercher une AMI Linux ... 110
Rechercher une AMI Linux à l'aide de la console Amazon EC2 ... 111
Rechercher une AMI à l'aide de l'AWS CLI ... 112
Rechercher l'AMI Amazon Linux la plus récente avec Systems Manager ... 112
Utiliser un paramètre Systems Manager pour rechercher une AMI ... 113
Rechercher une AMI Quick Start ... 116
AMI partagées ... 117
Rechercher des AMI partagées ... 117
Rendre une AMI publique ... 119
Partager une AMI avec des comptes AWS spécifiques ... 121
Utiliser des signets ... 123
Consignes pour les AMI Linux partagées ... 123
AMI payantes ... 128
Vendre votre AMI ... 129
Rechercher une AMI payante ... 129
Acheter une AMI payante ... 130
Obtenir le code produit pour votre instance ... 130
Utiliser le support payant ... 131
Factures pour les AMI payantes et supportées ... 131
Gérer vos abonnements AWS Marketplace ... 131
Création d'une AMI Linux basée sur Amazon EBS ... 132
Présentation de la création d'AMIs basées sur des volumes Amazon EBS ... 132
Création d'une AMI Linux à partir d'une instance ... 133
Création d'une AMI Linux à partir d'un instantané ... 135
Lancement d'une instance à partir d'une AMI que vous avez créée ... 136
Création d'une AMI Linux basée sur le stockage d'instance ... 136
Présentation du processus de création des AMI basées sur le stockage d'instance ... 136
Prérequis ... 137
Installation des outils AMI ... 138
Création d'une AMI à partir d'une instance basée sur le stockage d'instance ... 141

Conversion en AMI basée sur Amazon EBS 148
Référence des outils AMI 151
Utiliser le chiffrement avec des AMI basées sur EBS 168
Scénarios de lancement d'instances 168
Scénarios de copie d'images 171
Copie une AMI 173
Autorisations nécessaires pour copier une AMI basée sur le stockage d'instance 174
Copie entre régions 175
Copie entre comptes 176
Chiffrement et copie 176
Copie d'une AMI 177
Arrêt d'une opération de copie d'une AMI en attente 179
Obtenir des informations de facturation 179
Champs d'informations de facturation d'AMI 179
Détails de la plateforme et valeurs des opérations d'utilisation 180
Affichage des détails de la plateforme et des valeurs des opérations d'utilisation 180
Confirmation des informations de facturation sur votre facture 182
Annulation de l'inscription de votre AMI Linux 182
Nettoyage de vos AMI basées sur des volumes Amazon EBS 183
Nettoyage de votre AMI basée sur le stockage d'instance 184
Amazon Linux 186
Disponibilité Amazon Linux 186
Connexion à une instance Amazon Linux 186
Identification des images Amazon Linux 187
Outils de ligne de commande AWS 188
Référentiel de package 189
Bibliothèque Extras (Amazon Linux 2) 191
Accès aux packages source à des fins de référence 192
cloud-init 192
Abonnement aux notifications Amazon Linux 194
Exécution d'Amazon Linux 2 sur site 196
Kernel Live Patching 200
Noyaux fournis par l'utilisateur 205
AMIs HVM (GRUB) 205
AMIs paravirtuelles (PV-GRUB) 206
Utilisation de l'environnement de bureau MATE 211
Instances 213
Types d'instance 213
Types d'instance disponibles 214
Spécifications matérielles 217
Types de virtualisation AMI 218
Instances reposant sur le système Nitro 218
Fonctions de mise en réseau et de stockage 219
Limites d'instance 222
Usage général 222
Calcul optimisé 270
Mémoire optimisée 278
Stockage optimisé 289
Calcul accéléré 296
Recherche d'un type d'instance 312
Modification du type d'instance 313
Obtention de recommandations 319
Options d'achat d'instance 323
Détermination du cycle de vie de l'instance 324
Instances à la demande 325
Instances réservées 329
Instances planifiées 369

Instances Spot 373
Hôtes dédiés 465
Instances dédiées 498
Réservations de capacité à la demande 504
Cycle de vie d'une instance 524
Lancement d'une instance 525
Arrêt et démarrage d'une instance (instances basées sur les volumes Amazon EBS uniquement) . 526
Mise en veille prolongée d'une instance (instances basées sur Amazon EBS uniquement) 526
Redémarrage d'instance 527
Mise hors service d'instance 527
Terminaison d'instance 527
Différences entre redémarrage, arrêt, mise en veille prolongée et résiliation 528
Lancer 529
Connexion 598
Arrêt et démarrage 625
Mise en veille prolongée 628
Redémarrer 641
Mise hors service 642
Terminer 645
Récupération 652
Configurer les instances 653
Scénarios de configuration courants 653
Gestion de logiciels 654
Gestion des utilisateurs 660
Contrôle des états du processeur 662
Réglage de l'heure 668
Optimisations des options d'UC 673
Changement du nom d'hôte 690
Configuration d'un DNS dynamique 692
Exécution de commandes lors du lancement 694
Métadonnées d'instance et données utilisateur 701
Elastic Inference 736
Identification d'instances 736
Inspection du Documents d'identité d'instance 737
Inspection de l'UUID du système 737
Surveillance 739
Surveillance automatique et surveillance manuelle 740
Outils de surveillance automatique 740
Outils de surveillance manuelle 741
Bonnes pratiques de surveillance 742
Surveillance du statut de vos instances 742
Contrôles de statut des instances 742
Événements planifiés 749
Surveillance de vos instances avec CloudWatch 761
Activer la surveillance détaillée 762
Répertorier les métriques disponibles 764
Obtenir les statistiques des métriques 775
Graphique de métriques 783
Créer une alarme 783
Créer des alarmes qui arrêtent, finissent, redémarrent ou récupèrent une instance 785
Automatisation d'Amazon EC2 avec EventBridge 795
Mémoire de surveillance et métriques de disques 795
Collecte de métriques à l'aide de l'agent CloudWatch 796
Obsolète : Collecte de métriques à l'aide des scripts de surveillance CloudWatch 796
Journalisation des appels d'API avec AWS CloudTrail 804
Informations sur Amazon EC2 et Amazon EBS dans CloudTrail 804
Compréhension des entrées du fichier journal Amazon EC2 et Amazon EBS 805

Audit des utilisateurs qui se connectent via EC2 Instance Connect 806
Mise en réseau 808
Adressage IP des instances 808
Adresses IPv4 privées et noms d'hôte DNS internes 808
Adresses IPv4 publiques et noms d'hôte DNS externes 809
Adresses IP Elastic (IPv4) 810
Serveur Amazon DNS 810
Adresses IPv6 810
Utilisation des adresses IPv4 pour vos instances 811
Utilisation des adresses IPv6 pour vos instances 815
Plusieurs adresses IP 817
Fourniture de vos propres adresses IP 825
Exigences 826
Préparation à l'importation de votre plage d'adresses dans votre compte AWS 826
Mise en service de la plage d'adresses pour une utilisation avec AWS 828
Publication de la plage d'adresses via AWS 829
Utiliser votre plage d'adresses 830
Mise hors service de la plage d'adresses 831
Adresses IP Elastic 831
Principes de base d'une adresse IP Elastic 831
Utilisation d'adresses IP Elastic 832
Utilisation d'enregistrements DNS inverse pour les applications de messagerie 839
Limite appliquée aux adresses IP Elastic 839
Interfaces réseau 840
Notions fondamentales concernant l'interface réseau 840
Adresses IP par interface réseau et par type d'instance 841
Utilisation des interfaces réseau 853
Scénarios pour les interfaces réseau 859
Meilleures pratiques pour la configuration des interfaces réseau 861
Interfaces réseau gérées par demandeur 863
Mise en réseau améliorée 864
Prise en charge de la mise en réseau améliorée 864
Activation de la mise en réseau améliorée sur une instance 865
Mise en réseau améliorée : ENA 865
Mise en réseau améliorée : Intel 82599 VF 878
Dépannage de l'adaptateur ENA 884
Elastic Fabric Adapter (EFA) 891
Principes de base EFA 891
Interfaces et bibliothèques prises en charge 893
Types d'instance pris en charge 893
AMIs prises en charge 893
Restrictions liées à EFA 894
Démarrer avec EFA et MPI 894
Démarrer avec EFA et NCCL 902
Utilisation d’EFA 920
Surveillance d'un EFA 922
Vérification du programme d'installation EFA à l'aide d'un total de contrôle 923
Groupes de placement 924
Groupes de placement du cluster 925
Groupes de placement par partition 926
Groupes de placement par répartition 927
Règles et restrictions des groupes de placement 927
Création d'un groupe de placement 929
Balisage d'un groupe de placement 930
Lancement d'instances dans un groupe de placement 932
Description des instances dans un groupe de placement 933
Modification du groupe de placement d'une instance 935

Suppression d'un groupe de placement 936
MTU réseau 937
Trames jumbo (MTU de 9001) 937
Détection de la MTU du chemin 938
Vérifier la MTU du chemin entre deux hôtes 938
Vérification et définition de la MTU sur votre instance Linux 939
Dépannage 940
Clouds privés virtuels 940
Documentation Amazon VPC 940
EC2-Classic 941
Suppression des plateformes prises en charge 941
Types d'instances disponibles dans EC2-Classic 942
Différences entre les instances d'EC2-Classic et d'un VPC 943
Partager et accéder à des ressources entre EC2-Classic et un VPC 948
ClassicLink 949
Migration d'EC2-Classic vers un VPC 961
Sécurité 971
Sécurité de l'infrastructure 971
Isolement de réseau 972
Isolation sur les hôtes physiques 972
Contrôle du trafic réseau 972
Points de terminaison de VPC d'interface 973
Création d'un point de terminaison de VPC d'interface 973
Création d'une stratégie de point de terminaison de VPC d'interface 973
Résilience 974
Protection des données 975
Chiffrement au repos 976
Chiffrement en transit 976
Identity and Access Management 976
Accès réseau à votre instance 977
Attributs d'autorisation Amazon EC2 977
IAM et Amazon EC2 977
Stratégies IAM 979
Rôles IAM 1034
Accès réseau 1044
Paires de clés 1047
Création ou importation d'une paire de clés 1047
Balisage d'une paire de clés 1051
Récupération de la clé publique pour votre paire de clés sous Linux 1053
Récupération de la clé publique de votre paire de clés via les métadonnées d'instance 1053
Localisation de la clé publique sur une instance 1054
Identification de la paire de clés spécifiée au lancement 1055
Vérification de l'empreinte de votre paire de clés 1055
Ajout ou remplacement d'une paire de clés pour votre instance 1055
Connexion à votre instance Linux en cas de perte de votre clé privée 1056
Suppression de votre paire de clés 1061
Groupes de sécurité 1062
Règles des groupes de sécurité 1063
Groupes de sécurité par défaut 1066
Custom security groups 1066
Utilisation des groupes de sécurité 1067
Référence des règles de groupe de sécurité 1075
Gestion des mises à jour 1081
Validation de la conformité 1082
Stockage 1083
Amazon EBS 1084
Fonctions d'Amazon EBS 1085

Volumes EBS 1086
Instantanés EBS 1129
Services de données EBS 1181
Volumes EBS et NVMe 1209
Optimisation EBS 1213
Performances EBS 1230
Métriques CloudWatch EBS 1247
Événements CloudWatch EBS 1253
Quotas EBS 1264
Stockage d'instance 1264
Durée de vie d'un stockage d'instance 1265
Volumes de stockage d'instance 1266
Ajouter des volumes de stockage d'instance 1272
Volumes de stockage d'instance SSD 1276
Volumes d'échange de stockage d'instance 1278
Optimisation des performances disque 1280
Stockage de fichiers 1281
Amazon S3 1281
Amazon EFS 1283
Limites de volume d'instance 1286
Limites de volume du système Nitro 1286
Limites de volume spécifiques à Linux 1287
Bande passante et capacité 1287
Noms d'appareil 1287
Noms d'appareil disponibles 1288
Considérations sur les noms d'appareil 1289
Mappage de périphérique de stockage en mode bloc 1289
Concepts de mappage de périphérique de stockage en mode bloc 1290
Mappage de périphérique de stockage en mode bloc d'une AMI 1293
Mappage de périphérique de stockage en mode bloc d'une instance 1295
Ressources et balises 1301
Emplacements des ressources 1301
ID de ressource 1302
Liste et filtrage de vos ressources 1303
Création d'une liste et filtrage des ressources à l'aide de la console 1303
Affichage et filtrage à l'aide de l'interface de commande et des API 1307
Balisage de vos ressources 1309
Principes de base des balises 1309
Balisage de vos ressources 1310
Restrictions liées aux balises 1313
Balisage de vos ressources pour facturation 1314
Gestion des balises à l'aide de la console 1314
Utilisation des balises à l'aide de la ligne de commande 1318
Ajout de balises à une ressource à l'aide de CloudFormation 1321
Quotas de service 1322
Affichages de vos limites actuelles 1322
Demande d'augmentation 1324
Limites relatives à des e-mails envoyés à l'aide du port 25 1324
Rapports d'utilisation 1324
Dépannage 1325
Résolution des problèmes de lancement 1325
Dépassement de la limite d'instance 1325
Capacité d'instance insuffisante 1326
Mise hors service immédiate de l'instance 1326
Connexion à votre instance 1327
Causes courantes des problèmes de connexion 1328
Erreur de connexion à votre instance : connexion expirée 1329

Erreur : unable to load key… Expecting: ANY PRIVATE KEY 1331
Erreur : clé de l'utilisateur non reconnue par le serveur 1331
Erreur : autorisation refusée ou connexion fermée par [instance] port 22 1332
Erreur : fichier de clé privée non protégé 1334
Erreur : La clé privée doit commencer par « -----BEGIN RSA PRIVATE KEY----- » et se terminer par « -----END RSA PRIVATE KEY----- » 1334
Erreur : le serveur a refusé notre clé ou Aucune méthode d'authentification prise en charge disponible 1334
Impossible d'envoyer une commande ping à l'instance 1335
Erreur : le serveur a fermé la connexion réseau de manière inopinée 1335
Arrêt de votre instance 1335
Création d'une instance de remplacement 1336
Mise hors service de votre instance 1337
Mise à fin d'instance retardée 1337
Instance terminée toujours affichée 1337
Instances lancées ou terminées automatiquement 1338
Contrôles de statut échoués 1338
Examen des informations de contrôle de statut 1339
Récupération des journaux système 1339
Dépannage des erreurs du journal du système pour les instances basées sur Linux 1340
Mémoire insuffisante : processus d'arrêt 1341
ERROR: mmu_update failed (la mise à jour de la gestion de la mémoire a échoué) 1341
Erreur d'E/S (échec du périphérique de stockage en mode bloc) 1342
I/O ERROR: neither local nor remote disk (le périphérique de stockage en mode bloc distribué ne fonctionne plus) 1343
request_module: runaway loop modprobe (modprobe en boucle sur le noyau hérité sur des versions Linux plus anciennes) 1344
« FATAL: kernel too old » et « fsck: No such file or directory while trying to open /dev » (décalage entre le noyau et l'AMI) 1345
« FATAL: Could not load /lib/modules » ou « BusyBox » (modules noyau manquants) 1346
ERROR Invalid kernel (noyau incompatible EC2) 1347
fsck: No such file or directory while trying to open... (système de fichiers non trouvé) 1348
General error mounting filesystems (Montage en échec) 1350
VFS: Unable to mount root fs on unknown-block (le système de fichiers racine ne correspond pas) 1351
Erreur : Unable to determine major/minor number of root device... (décalage du système de fichiers/périphérique racine) 1352
XENBUS : Device with no driver... 1353
... days without being checked, check forced (Contrôle du système de fichiers nécessaire) 1354
fsck died with exit status... (périphérique manquant) 1355
Invite GRUB (grubdom>) 1356
Bringing up interface eth0: Device eth0 has different MAC address than expected, ignoring. (Adresse MAC codée de manière irréversible) 1358
Unable to load SELinux Policy. Machine is in enforcing mode. Halting now. (Erreur de configuration SELinux) 1359
XENBUS: Timeout connecting to devices (délai d'attente Xenbus) 1360
Résolution d'un problème d'instance inaccessible 1360
Redémarrage d'instance 1361
Sortie de la console de l'instance 1361
Création d'une capture d'écran d'une instance inaccessible 1362
Récupération d'instance en cas de plantage de l'ordinateur hôte 1363
Démarrage à partir du mauvais volume 1363
EC2Rescue pour Linux 1364
Installation d'EC2Rescue pour Linux 1365
(Facultatif) Vérification de la signature de EC2Rescue pour Linux 1366
Utilisation d'EC2Rescue pour Linux 1368
Développement des modules EC2Rescue 1370

Envoi d'une interruption de diagnostic 1375
Types d'instance pris en charge 1375
Prérequis 1375
Envoi d'une interruption de diagnostic 1378
Historique du document 1379
Historique des années précédentes 1384

Présentation d'Amazon EC2

Amazon Elastic Compute Cloud (Amazon EC2) offre une capacité de calcul évolutive dans le cloud Amazon Web Services (AWS). L'utilisation d'Amazon EC2 vous dispense d'investir à l'avance dans du matériel et, par conséquent, vous pouvez développer et déployer les applications plus rapidement. Vous pouvez utiliser Amazon EC2 pour lancer autant de serveurs virtuels que nécessaire, configurer la sécurité et la mise en réseau, et gérer le stockage. Amazon EC2 permet également de monter ou descendre en puissance rapidement afin de gérer l'évolution des exigences ou des pics de popularité, ce qui permet de réduire la nécessité de prévoir le trafic du serveur.

Pour plus d'informations sur le cloud computing, consultez What is Cloud Computing?

Fonctionnalités d'Amazon EC2

Amazon EC2 offre les fonctionnalités suivantes :

- Environnements de calcul virtuels, appelés instances
- Modèles préconfigurés pour vos instances, appelés Amazon Machine Images (AMI), qui combinent les composants dont vous avez besoin pour votre serveur (système d'exploitation et logiciels supplémentaires inclus)
- Diverses configurations de capacité d'UC de mémoire, de stockage et de mise en réseau pour vos instances, appelées types d'instance
- Sécuriser les informations de connexion de vos instances à l'aide de paires de clés (AWS stocke la clé publique, tandis que vous stockez la clé privée dans un endroit sécurisé)
- Volumes de stockage pour les données temporaires qui sont supprimées lorsque vous arrêtez, mettez en veille prolongée ou résiliez votre instance, appelés volumes de stockage d'instance
- Volumes de stockage permanents pour vos données à l'aide d'Amazon Elastic Block Store (Amazon EBS), appelés volumes Amazon EBS
- Plusieurs emplacements physiques pour vos ressources, tels que les instances et les volumes Amazon EBS, appelés Régions et Zones de disponibilité
- Pare-feu permettant de spécifier les protocoles, ports et plages d'adresses IP source qui peuvent atteindre vos instances à l'aide des groupes de sécurité
- Adresses IPv4 statiques pour le cloud computing dynamique, appelées adresses IP Elastic
- Métadonnées, appelées balises, que vous pouvez créer et affecter à vos ressources Amazon EC2
- Réseaux virtuels que vous pouvez créer et qui sont logiquement isolés du reste du cloud AWS, et que, le cas échéant, vous pouvez connecter à votre propre réseau, appelés clouds privés virtuels (VPC)

Pour plus d'informations sur les fonctions de Amazon EC2, consultez la page produit Amazon EC2.

Pour plus d'informations sur l'exécution de votre site web sur AWS, consultez Hébergement web.

Comment démarrer avec Amazon EC2

Tout d'abord, vous devez vous préparer à utiliser Amazon EC2. Une fois la configuration terminée, vous êtes prêt à compléter le didacticiel Démarrez pour Amazon EC2. Chaque fois que vous avez besoin

d'informations supplémentaires sur une fonction Amazon EC2, vous pouvez lire la documentation technique.

Etre opérationnel

- Configuration avec Amazon EC2 (p. 28)
- Didacticiel : Premiers pas avec les instances Amazon EC2 Linux (p. 32)

Les bases

- Instances et AMI (p. 4)
- Régions et zones (p. 7)
- Types d'instance (p. 213)
- Balises (p. 1309)

Mise en réseau et sécurité

- Paires de clés (p. 1047)
- Groupes de sécurité (p. 1062)
- Adresses IP Elastic (p. 831)
- Clouds privés virtuels (p. 940)

Stockage

- Amazon EBS (p. 1084)
- Stockage d'instance (p. 1264)

Utilisation des instances Linux

- Run Command AWS Systems Manager dans le Guide de l'utilisateur AWS Systems Manager
- Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2 (p. 38)
- Didacticiel : Configurer SSL/TLS sur Amazon Linux 2 (p. 69)

Si vous souhaitez savoir si AWS vous convient, contactez le service commercial AWS. Pour toute question technique sur Amazon EC2, utilisez le Amazon EC2 forum.

Services connexes

Vous pouvez allouer des ressources Amazon EC2, telles que des instances et des volumes, en utilisant directement Amazon EC2. Vous pouvez aussi allouer des ressources Amazon EC2 à l'aide d'autres services dans AWS. Pour plus d'informations, consultez la documentation suivante :

- Amazon EC2 Auto Scaling Guide de l'utilisateur
- AWS CloudFormation Guide de l'utilisateur
- AWS Elastic Beanstalk Manuel du développeur
- AWS OpsWorks User Guide

Pour répartir automatiquement le trafic applicatif entrant sur plusieurs instances, utilisez Elastic Load Balancing. Pour plus d'informations, consultez le Elastic Load Balancing Guide de l'utilisateur.

Pour obtenir une base de données relationnelle gérée dans le cloud, utilisez Amazon Relational Database Service (Amazon RDS) pour lancer une instance de base de données. Même si vous pouvez configurer une base de données sur une instance EC2, Amazon RDS présente l'avantage de gérer vos tâches de gestion de base de données, telles que l'application de correctifs logiciels, la sauvegarde et le stockage de sauvegardes. Pour plus d'informations, consultez le Amazon Relational Database Service Guide du développeur.

Pour faciliter la gestion des conteneurs Docker sur un cluster d'instances EC2, utilisez Amazon Elastic Container Service (Amazon ECS). Pour de plus amples informations, veuillez consulter le Amazon Elastic Container Service Developer Guide ou le Guide de l'utilisateur Amazon Elastic Container Service pour AWS Fargate.

Pour superviser les statistiques de base de vos instances et les volumes Amazon EBS, utilisez Amazon CloudWatch. Pour plus d'informations, consultez Guide de l'utilisateur Amazon CloudWatch. Pour détecter une utilisation potentiellement autorisée ou malveillante de vos instances EC2, utilisez Amazon GuardDuty. Pour de plus amples informations, veuillez consulter le Guide de l'utilisateur Amazon GuardDuty.

Accès à Amazon EC2

Amazon EC2 fournit une interface utilisateur basée sur le Web, la console Amazon EC2. Si vous êtes inscrit à un compte AWS, vous pouvez accéder à la console Amazon EC2 en vous connectant à la console de gestion AWS Management Console et en sélectionnant EC2 depuis la page d'accueil de la console.

Si vous préférez utiliser une interface ligne de commande, vous disposez des options suivantes :

Interface de ligne de commande (CLI) AWS

Fournit des commandes pour une large gamme de produits AWS et est prise en charge par Windows, Mac et Linux/UNIX. Consultez le AWS Command Line Interface Guide de l'utilisateur pour démarrer. Pour plus d'informations sur les commandes pour Amazon EC2, consultez ec2 dans le manuel AWS CLI Command Reference.

Outils AWS pour Windows PowerShell

Fournit des commandes pour une large gamme de produits AWS à ceux qui écrivent des scripts dans l'environnement PowerShell. Consultez le Outils AWS pour Windows PowerShell Guide de l'utilisateur pour démarrer. Pour plus d'informations sur les applets de commande pour Amazon EC2, consultez le Outils AWS pour PowerShell Cmdlet Reference.

Amazon EC2 prend en charge la création de ressources avec AWS CloudFormation. Vous créez un modèle, dans JSON ou YAML, qui décrit vos ressources AWS, et AWS CloudFormation alloue et configure ces ressources pour vous. Vous pouvez réutiliser vos modèles CloudFormation pour allouer les mêmes ressources plusieurs fois, que ce soit dans la même région et le même compte ou dans des régions et comptes différents. Pour de plus amples informations sur les types de ressources et les propriétés pour Amazon EC2, veuillez consulter EC2 Référence de type de ressource dans le AWS CloudFormation Guide de l'utilisateur.

Amazon EC2 fournit une API de requête. Ces requêtes sont des requêtes HTTP ou HTTPS qui utilisent les verbes HTTP GET ou POST et un paramètre de requête nommé `Action`. Pour plus d'informations sur les actions d'API pour Amazon EC2, consultez Actions dans le Amazon EC2 API Reference.

Si vous préférez développer des applications utilisant des API propres au langage au lieu d'envoyer une demande via HTTP ou HTTPS, AWS fournit des bibliothèques, des exemples de code, des didacticiels et d'autres ressources aux développeurs de logiciels. Ces bibliothèques offrent des fonctions de base qui automatisent les tâches telles que la signature cryptographique des demandes, les nouvelles tentatives de demande et la gestion des réponses d'erreur. Vous pouvez ainsi démarrer plus facilement. Pour plus d'informations, consultez Kits SDK et outils AWS.

Tarification pour Amazon EC2

Lorsque vous vous inscrivez à AWS, vous pouvez démarrer gratuitement avec Amazon EC2 grâce à l'offre gratuite AWS.

Amazon EC2 propose les options d'achat suivantes pour les instances :

Instances à la demande

Payez les instances que vous utilisez à la seconde, sans engagement à long terme ou paiement initial.

Plans d'économies

Vous pouvez réduire les coûts de Amazon EC2 en vous engageant pour une utilisation continue, en USD par heure, pour une durée de 1 à 3 ans.

Instances réservées

Vous pouvez réduire les coûts de Amazon EC2 en vous engageant pour une configuration d'instance spécifique, incluant le type et la région, pour une durée de 1 à 3 ans.

Instances Spot

Demande d'instances EC2 inutilisées, ce qui peut réduire vos coûts de Amazon EC2 de façon considérable.

Pour obtenir la liste complète des tarifs de Amazon EC2, veuillez consulter Tarification Amazon EC2.

Pour calculer le coût d'un exemple d'environnement alloué, consultez le Centre d'optimisation des coûts du cloud.

Pour consulter votre facture, dirigez-vous vers le Tableau de bord de gestion des coûts et de la facturation dans la console AWS Billing and Cost Management. Votre facture contient des liens vers les rapports d'utilisation qui fournissent des détails sur votre facture. Pour en savoir plus sur la facturation du compte AWS, consultez Facturation du compte AWS.

Pour toute question relative à la facturation, aux comptes et aux événements AWS, contactez AWS Support.

Pour une vue d'ensemble de Trusted Advisor, service qui vous aide à optimiser les coûts, la sécurité et les performances de votre environnement AWS, consultez AWS Trusted Advisor.

Conformité DSS PCI

Amazon EC2 prend en charge le traitement, le stockage et la transmission des données de cartes bancaires par un commerçant ou un fournisseur de services, et a été validé comme étant conforme à la norme PCI (Payment Card Industry) DSS (Data Security Standard). Pour plus d'information sur PCI DSS, et notamment sur la manière de demander une copie de l'AWS PCI Compliance Package, consultez PCI DSS, niveau 1.

Instances et AMI

Une image machine Amazon (AMI) est un modèle qui contient une configuration logicielle (par exemple, un système d'exploitation, un serveur d'applications et des applications). À partir d'une AMI, vous lancez

une instance qui est une copie de l'AMI s'exécutant en tant que serveur virtuel dans le cloud. Vous pouvez lancer plusieurs instances d'une AMI, comme illustré sur la figure suivante.

Vos instances continuent de s'exécuter jusqu'à ce que vous les arrêtiez, les mettiez en veille prolongée ou les résiliiez, ou jusqu'à ce qu'elles connaissent une défaillance. En cas de défaillance d'une instance, vous pouvez en lancer une nouvelle à partir de l'AMI.

Instances

Une instance est un serveur virtuel figurant dans le cloud. Sa configuration au moment du lancement est une copie de l'AMI que vous avez spécifiée quand vous avez lancé l'instance.

Vous pouvez lancer différents types d'instance à partir d'une seule AMI. Un type d'instance détermine essentiellement les capacités matérielles de l'ordinateur hôte utilisé pour votre instance. Chaque type d'instance offre des capacités de calcul et de mémoire différentes. Sélectionnez un type d'instance en fonction de la quantité de mémoire et de la puissance de calcul dont vous avez besoin pour l'application ou le logiciel que vous prévoyez d'exécuter sur l'instance. Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.

Une fois une instance lancée, celle-ci ressemble à un hôte classique et vous pouvez interagir avec elle comme avec n'importe quel ordinateur. Vous exercez un contrôle total sur vos instances. Vous pouvez utiliser sudo pour exécuter des commandes nécessitant des privilèges racine.

Votre compte AWS est associé à une limite quant au nombre d'instances pouvant être en cours d'exécution. Pour de plus amples informations sur cette limite et savoir comment demander que celle-ci soit augmentée, veuillez consulter Combien d'instances puis-je exécuter dans Amazon EC2 dans la FAQ générale sur Amazon EC2.

Stockage pour votre instance

Le périphérique racine pour votre instance contient l'image utilisée pour démarrer l'instance. Pour plus d'informations, consultez Volume du périphérique racine Amazon EC2 (p. 21).

Votre instance peut inclure des volumes de stockage local, appelés volumes de stockage d'instance, que vous pouvez configurer au moment du lancement avec la fonctionnalité de mappage de périphérique

de stockage en mode bloc. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289). Une fois que ces volumes ont été ajoutés et mappés à votre instance, ils sont à votre disposition pour être montés et utilisés. Si votre instance est défaillante, ou si elle est arrêtée ou terminée, les données sur ces volumes sont perdues. Ces volumes sont donc mieux adaptés aux données temporaires. Pour maintenir en sécurité des données importantes, vous devez utiliser une stratégie de réplication sur plusieurs instances ou stocker vos données persistantes sur des volumes Amazon S3 ou Amazon EBS. Pour plus d'informations, consultez Stockage (p. 1083).

Bonnes pratiques de sécurité

- Utilisez AWS Identity and Access Management (IAM) pour contrôler l'accès à vos ressources AWS, y compris vos instances. Vous pouvez créer des utilisateurs et des groupes IAM sous votre compte AWS, affecter des informations d'identification (credentials) de sécurité à chacun d'entre eux et contrôler l'accès aux ressources et services dans AWS. Pour de plus amples informations, veuillez consulter Identity and Access Management pour Amazon EC2 (p. 976).
- Limitez l'accès en permettant uniquement aux hôtes et réseaux approuvés d'accéder à des ports sur votre instance. Par exemple, vous pouvez restreindre l'accès SSH en limitant le trafic entrant sur le port 22. Pour plus d'informations, consultez Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062).
- Vérifiez régulièrement les règles de vos groupes de sécurité et veillez à appliquer le principe du moindre privilège —en donnant accès uniquement aux permissions dont vous avez besoin. Vous pouvez également créer différents groupes de sécurité pour gérer les instances ayant des exigences de sécurité différentes. Envisagez de créer un groupe de sécurité bastion qui autorise les connexions externes, et conservez le reste de vos instances dans un groupe n'autorisant pas les connexions externes.
- Désactivez les connexions basées sur mot de passe pour les instances lancées à partir de votre AMI. Les mots de passe peuvent être trouvés ou craqués, et représentent donc un risque pour la sécurité. Pour plus d'informations, consultez Désactivation des connexions à distance basées sur mot de passe à la racine (p. 124). Pour plus d'informations sur comment partager des AMI en toute sécurité, consultez AMI partagées (p. 117).

Arrêt et fin des instances

Vous pouvez arrêter ou mettre fin à une instance en cours d'exécution à tout moment.

Arrêt d'une instance

Lorsqu'une instance est arrêtée, celle-ci exécute une fermeture normale, puis passe à un état `stopped`. Tous ses volumes Amazon EBS restent attachés et vous pouvez redémarrer l'instance ultérieurement.

L'utilisation d'instance supplémentaire ne vous est pas facturée pendant que l'instance est à un état arrêté. Une une minute au minimum sera facturée pour chaque passage d'un état arrêté à un état d'exécution. Si le type d'une instance a été modifié pendant que l'instance était arrêtée, le taux du nouveau type d'instance vous sera facturé après le démarrage de l'instance. Toute l'utilisation Amazon EBS associée de votre instance, y compris l'utilisation du périphérique racine, est facturée selon des tarifs Amazon EBS classiques.

Lorsqu'une instance est à l'état arrêté, vous pouvez attacher ou détacher des volumes Amazon EBS. Vous pouvez également créer une AMI à partir de l'instance, et vous pouvez modifier le noyau, le disque RAM et le type d'instance.

Mise hors service d'une instance

Lorsque vous mettez fin à une instance, celle-ci procède à une fermeture normale. Le volume du périphérique racine est supprimé par défaut, mais tous les volumes Amazon EBS attachés sont conservés par défaut, selon le paramètre de l'attribut `deleteOnTermination` de chaque volume. L'instance proprement dite est également supprimée et vous ne pourrez pas la redémarrer ultérieurement.

Pour éviter toute fin accidentelle, vous pouvez désactiver la possibilité de mettre fin à une instance. Dans ce cas, assurez-vous que l'attribut `disableApiTermination` est défini sur `true` pour l'instance. Pour contrôler le comportement de la fermeture d'une instance, par exemple, `shutdown -h` sous Linux ou `shutdown` sous Windows, définissez l'attribut de l'instance `instanceInitiatedShutdownBehavior` sur `stop` ou `terminate` si nécessaire. Les instances avec des volumes Amazon EBS pour le périphérique racine ont par défaut la valeur `stop`, et les instances avec des périphériques racine de stockage d'instance sont toujours terminés suite à une fermeture d'instance.

Pour plus d'informations, consultez Cycle de vie d'une instance (p. 524).

AMI

Amazon Web Services (AWS) publie de nombreuses AMI (Amazon Machine Images) contenant des configurations logicielles courantes pour un usage public. En outre, des membres de la communauté de développeurs AWS ont publié leurs propres AMI personnalisées. Vous pouvez également créer votre ou vos propres AMI personnalisées. Cela vous permet de démarrer rapidement et facilement de nouvelles instances disposant de tout ce dont vous avez besoin. Par exemple, si votre application est un site web ou un service web, votre AMI peut inclure un serveur web, le contenu statique associé et le code pour les pages dynamiques. Ainsi, quand vous lancez une instance à partir de cette AMI, votre serveur web démarre et votre application est prête à accepter des demandes.

Toutes les AMI sont classées comme étant basées sur Amazon EBS, ce qui signifie que le périphérique racine pour une instance lancée à partir de l'AMI est un volume Amazon EBS, ou basée sur un stockage d'instance, ce qui signifie que le périphérique racine pour une instance lancée à partir de l'AMI est un volume de stockage d'instance créé à partir d'un template stocké dans Amazon S3.

La description d'une AMI indique le type de périphérique racine (`ebs` ou `instance store`). Ceci est important, car ce que vous pouvez faire avec chaque type d'AMI présente des différences significatives. Pour plus d'informations sur ces différences, consultez Stockage pour le périphérique racine (p. 106).

Vous pouvez annuler l'inscription de votre AMI lorsque vous avez terminé de l'utiliser. Après cette opération, vous ne pouvez pas utiliser l'AMI pour lancer de nouvelles instances. Les instances existantes lancées à partir de l'AMI ne sont pas affectées. Par conséquent, si vous avez également terminé avec les instances lancées à partir de ces AMI, vous devez les arrêter.

Régions et zones

Amazon EC2 est hébergé à plusieurs endroits dans le monde. Ces emplacements sont composés de régions, de zones de disponibilité, de Zones locales et de zones de longueur d'onde. Chaque Région constitue une zone géographique séparée.

- Les zones de disponibilité sont des emplacements multiples isolés dans chaque région.
- Les Zones locales vous permettent de placer des ressources, telles que le calcul et le stockage, dans plusieurs emplacements plus proches de vos utilisateurs finaux.
- AWS Outposts offre les services, l'infrastructure et les modèles d'exploitation AWS natifs à la quasi-totalité de centres de données, d'espaces de colocalisation d'infrastructures ou d'installations sur site.
- Les zones Wavelength permettent aux développeurs de créer des applications qui offrent des latences ultra-faibles aux appareils 5G et aux utilisateurs finaux. Wavelength déploie des services de calcul et de stockage AWS standard à la périphérie des réseaux 5G des opérateurs de télécommunications.

AWS gère des centres de données à la pointe de la technologie et hautement disponibles. Bien qu'elles soient rares, des pannes touchant la disponibilité des instances se trouvant au même emplacement peuvent se produire. Si vous hébergez toutes vos instances dans un seul emplacement touché par une panne, aucune de vos instances ne sera disponible.

Pour vous aider à déterminer le déploiement qui vous convient le mieux, veuillez consulter les Questions fréquentes (FAQ) AWS Wavelength.

Sommaire

- Régions (p. 8)
- Zones de disponibilité (p. 13)
- Zones locales (p. 15)
- Zones de longueur d'onde (p. 18)
- AWS Outposts (p. 20)

Régions

Chaque région Amazon EC2 est conçue pour être complètement isolée des autres régions Amazon EC2. Cela permet d'atteindre la plus grande tolérance aux pannes possible et une stabilité optimale.

Le diagramme suivant illustre plusieurs régions AWS.

Lorsque vous consultez vos ressources, vous voyez uniquement celles liées à la région que vous avez spécifiée. Cela est dû au fait que les régions sont éloignées les unes des autres et que nous ne répliquons pas automatiquement les ressources entre régions.

Lorsque vous lancez une instance, vous devez sélectionner une AMI se trouvant dans la même région. Si l'AMI est dans une autre région, vous pouvez copier l'AMI dans la région que vous utilisez. Pour plus d'informations, consultez Copie une AMI (p. 173).

Notez qu'il n'y a pas de frais pour le transfert de données entre régions. Pour plus d'informations, consultez Tarification Amazon EC2 - Transfert de données.

Sommaire

- Régions disponibles (p. 8)
- Régions et points de terminaison (p. 10)
- Décrire vos régions (p. 10)
- Obtenir le nom de la région (p. 12)
- Spécification de la région pour une ressource (p. 12)

Régions disponibles

Votre compte détermine les régions qui vous sont disponibles. Exemples :

- Un compte AWS fournit plusieurs régions afin que vous puissiez lancer des instances Amazon EC2 dans des emplacements qui satisfont vos exigences. Par exemple, vous pouvez souhaiter lancer des instances en Europe afin d'être plus proche de vos clients européens ou pour satisfaire à des exigences légales.
- Un compte AWS GovCloud (USA Ouest) fournit un accès uniquement à la région AWS GovCloud (USA Ouest) et à la région AWS GovCloud (US-East). Pour de plus amples informations, veuillez consulter AWS GovCloud (États-Unis).

- Un compte Amazon AWS (Chine) fournit un accès aux régions Pékin et Ningxia uniquement. Pour plus d'informations, consultez la section relative à AWS en Chine.

Le tableau suivant répertorie les régions fournies par un compte AWS. Vous ne pouvez pas décrire des régions supplémentaires ou y accéder depuis un compte AWS, par exemple Région AWS GovCloud (US) ou les régions chinoises. Pour utiliser une région introduite après le 20 mars 2019, vous devez l'activer. Pour de plus amples informations, veuillez consulter Gestion des régions AWS dans le manuel AWS General Reference.

Pour de plus amples informations sur les zones de longueur d'onde disponibles, veuillez consulter Zones de longueur d'onde disponibles dans le Manuel du développeur sur les zones de longueur d'onde AWS.

Code	Nom	Statut d'inscription	Zone locale
`us-east-2`	USA Est (Ohio)	Facultatif	Non disponible
`us-east-1`	USA Est (Virginie du Nord)	Facultatif	Non disponible
`us-west-1`	USA Ouest (Californie du Nord)	Facultatif	Non disponible
`us-west-2`	USA Ouest (Oregon)	Facultatif	`us-west-2-lax-1a` `us-west-2-lax-1b`
`af-south-1`	Afrique (Le Cap)	Obligatoire	Non disponible
`ap-east-1`	Asie-Pacifique (Hong Kong)	Obligatoire	Non disponible
`ap-south-1`	Asie-Pacifique (Mumbai)	Facultatif	Non disponible
`ap-northeast-3`	Asie-Pacifique (Osaka-Local)	Facultatif	Non disponible
`ap-northeast-2`	Asie-Pacifique (Séoul)	Facultatif	Non disponible
`ap-southeast-1`	Asie-Pacifique (Singapour)	Facultatif	Non disponible
`ap-southeast-2`	Asie-Pacifique (Sydney)	Facultatif	Non disponible
`ap-northeast-1`	Asie-Pacifique (Tokyo)	Facultatif	Non disponible
`ca-central-1`	Canada (Centre)	Facultatif	Non disponible
`eu-central-1`	Europe (Francfort)	Facultatif	Non disponible
`eu-west-1`	Europe (Irlande)	Facultatif	Non disponible
`eu-west-2`	Europe (Londres)	Facultatif	Non disponible
`eu-south-1`	UE (Milan)	Obligatoire	Non disponible
`eu-west-3`	Europe (Paris)	Facultatif	Non disponible
`eu-north-1`	Europe (Stockholm)	Facultatif	Non disponible
`me-south-1`	Moyen-Orient (Bahreïn)	Obligatoire	Non disponible

Code	Nom	Statut d'inscription	Zone locale
`sa-east-1`	Amérique du Sud (São Paulo)	Facultatif	Non disponible

Pour de plus amples informations, veuillez consulter Infrastructure globale AWS.

Le nombre et le mappage des zones de disponibilité par région peuvent varier d'un compte AWS à l'autre. Pour obtenir la liste des zones de disponibilité qui sont disponibles pour votre compte, vous pouvez utiliser la console Amazon EC2 ou l'interface ligne de commande. Pour plus d'informations, consultez Décrire vos régions (p. 10).

Régions et points de terminaison

Lorsque vous utilisez une instance à l'aide de l'interface de ligne de commande ou des actions d'API, vous devez spécifier son point de terminaison régional. Pour de plus amples informations sur les régions et points de terminaison disponibles pour Amazon EC2, veuillez consulter Points de terminaison et quotas Amazon EC2 dans le Référence générale d'Amazon Web Services.

Pour plus d'informations sur les points de terminaison et les protocoles dans AWS GovCloud (USA Ouest), consultez Points de terminaison AWS GovCloud (USA Ouest) dans le AWS GovCloud (US) User Guide.

Décrire vos régions

Vous pouvez utiliser la console Amazon EC2 ou l'interface de ligne de commande pour déterminer quelles régions et zones de disponibilité sont disponibles pour votre compte. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

Pour rechercher vos régions à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Affichez les options dans le sélecteur de région à partir de la barre de navigation.

3. Vos ressources EC2 pour cette région sont affichées dans le tableau de bord EC2 dans la section Ressources.

Pour rechercher vos régions à l'aide de AWS CLI

- Utilisez la commande describe-regions comme suit pour décrire les régions qui sont activées pour votre compte.

```
aws ec2 describe-regions
```

Pour décrire toutes les régions, y compris celles qui sont désactivées pour votre compte, ajoutez l'option `--all-regions` comme suit.

```
aws ec2 describe-regions --all-regions
```

Pour rechercher vos régions à l'aide de Outils AWS pour Windows PowerShell

- Utilisez la commande Get-EC2Region comme suit pour décrire les régions de votre compte.

```
PS C:\> Get-EC2Region
```

Obtenir le nom de la région

Vous pouvez utiliser l'API Amazon Lightsail pour afficher le nom d'une région.

Pour afficher le nom de la région à l'aide de la commande AWS CLI

- Utilisez la commande get-regions comme suit pour décrire le nom de la région spécifiée.

```
aws lightsail get-regions --query "regions[?name=='region-name'].displayName" --output
 text
```

L'exemple suivant renvoie le nom de la région `us-east-2`.

```
aws lightsail get-regions --query "regions[?name=='us-east-2'].displayName" --output
 text
```

En voici la sortie :

```
Ohio
```

Spécification de la région pour une ressource

Vous pouvez spécifier la région pour la ressource à chaque fois que vous créez une ressource Amazon EC2. Vous pouvez spécifier la région pour une ressource à l'aide de la AWS Management Console ou de la ligne de commande.

Considérations

Il est possible que certaines ressources AWS ne soient pas disponibles dans toutes les régions. Assurez-vous de pouvoir créer les ressources dont vous avez besoin dans les régions souhaitées avant de lancer une instance.

Pour spécifier la région pour une ressource avec la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Utilisez le sélecteur de région dans la barre de navigation.

Pour spécifier la région par défaut à l'aide de la ligne de commande

Vous pouvez définir la valeur d'une variable d'environnement sur le point de terminaison régional souhaité (par exemple, `https://ec2.us-east-2.amazonaws.com`) :

- `AWS_DEFAULT_REGION` (AWS CLI)
- `Set-AWSDefaultRegion` (Outils AWS pour Windows PowerShell)

L'autre solution consiste à utiliser l'option ligne de commande `--region` (AWS CLI) ou `-Region` (Outils AWS pour Windows PowerShell) avec chaque commande individuelle. Par exemple, `--region us-east-2`.

Pour plus d'informations sur les régions et les points de terminaison pour Amazon EC2, consultez Points de terminaison Amazon Elastic Compute Cloud.

Zones de disponibilité

Chaque région se compose de plusieurs emplacements isolés appelés zones de disponibilité. Lorsque vous lancez une instance, vous pouvez choisir une zone de disponibilité, ou nous laisser en choisir une pour vous. Si vous distribuez vos instances dans plusieurs zones de disponibilité et si une instance connaît une défaillance, vous pouvez concevoir votre application afin qu'une instance dans une autre zone de disponibilité puisse gérer les requêtes.

Le diagramme suivant illustre plusieurs zones de disponibilité dans une région AWS.

Vous pouvez également utiliser les adresses IP Elastic pour masquer la défaillance d'une instance dans une zone de disponibilité en remappant rapidement l'adresse à une instance dans une autre zone de disponibilité. Pour plus d'informations, consultez Adresses IP Elastic (p. 831).

Une zone de disponibilité est représentée par un code de région suivi d'un identifiant à lettre ; par exemple, `us-east-1a`. Pour garantir que les ressources sont réparties entre les zones de disponibilité d'une région, nous mappons indépendamment les zones de disponibilité aux noms de chaque compte AWS. Par exemple, la zone de disponibilité `us-east-1a` de votre compte AWS peut se trouver à un emplacement différent de la zone de disponibilité `us-east-1a` d'un autre compte AWS.

Pour coordonner les zones de disponibilité entre les comptes, vous devez utiliser un ID de zone de disponibilité, qui représente l'identifiant unique et cohérent d'une zone de disponibilité. Par exemple, `use1-az1` est l'ID de zone de disponibilité de la région `us-east-1` dont l'emplacement est identique dans chaque compte AWS.

Vous pouvez afficher les ID des zones de disponibilité afin de déterminer l'emplacement des ressources d'un compte par rapport aux ressources d'un autre compte. Par exemple, si vous partagez avec un autre compte un sous-réseau dans la zone de disponibilité portant l'ID `use-az2`, ce sous-réseau est accessible par cet autre compte dans la zone de disponibilité portant également l'ID `use-az2`. L'ID de zone de disponibilité de chaque VPC et de chaque sous-réseau s'affiche dans la console Amazon VPC. Pour de

plus amples informations, veuillez consulter Utilisation de VPC partagés dans le Amazon VPC Guide de l'utilisateur.

Alors que les zones de disponibilité augmentent avec le temps, notre capacité à les développer peut devenir limitée. Dans ce cas, nous pouvons vous empêcher de lancer une instance dans une zone de disponibilité limitée, à moins que vous n'ayez déjà une instance dans cette zone de disponibilité. Finalement, nous pouvons également retirer la zone de disponibilité limitée de la liste des zones de disponibilité pour les nouveaux comptes. Par conséquent, votre compte peut avoir un nombre différent de zones de disponibilité disponibles dans une région qu'un autre compte.

Sommaire

- Décrire vos zones de disponibilité (p. 14)
- Lancer des instances dans une zone de disponibilité (p. 14)
- Migrer une instance vers une autre zone de disponibilité (p. 15)

Décrire vos zones de disponibilité

Vous pouvez utiliser la console Amazon EC2 ou l'interface de ligne de commande pour déterminer quelles zones de disponibilité sont disponibles pour votre compte. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

Pour rechercher vos zones de disponibilité à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Affichez les options dans le sélecteur de région à partir de la barre de navigation.
3. Dans le panneau de navigation, choisissez Tableau de bord EC2.
4. Les zones de disponibilité apparaissent sous État du service, dans Statut de la zone.

Pour rechercher vos zones de disponibilité à l'aide de AWS CLI

1. Utilisez la commande describe-availability-zones comme suit pour décrire les zones de disponibilité dans la région spécifiée.

```
aws ec2 describe-availability-zones --region region-name
```

2. Utilisez la commande describe-availability-zones comme suit pour décrire les zones de disponibilité, quel que soit leur statut d'inscription.

```
aws ec2 describe-availability-zones --all-availability-zones
```

Pour rechercher vos zones de disponibilité à l'aide de Outils AWS pour Windows PowerShell

Utilisez la commande Get-EC2AvailabilityZone comme suit pour décrire les zones de disponibilité dans la région spécifiée.

```
PS C:\> Get-EC2AvailabilityZone -Region region-name
```

Lancer des instances dans une zone de disponibilité

Lorsque vous lancez une instance, sélectionnez une région qui rapproche vos instances de clients spécifiques, ou qui satisfait à vos exigences légales ou autres. En lançant vos instances dans des zones de disponibilité distinctes, vous pouvez protéger vos applications de la défaillance d'un emplacement unique.

Lorsque vous lancez une instance, vous pouvez éventuellement spécifier une zone de disponibilité dans la région que vous utilisez. Si vous ne spécifiez pas de zone de disponibilité, nous sélectionnons une zone de disponibilité pour vous. Lorsque vous lancez vos instances initiales, nous vous recommandons d'accepter la zone de disponibilité par défaut, car cela nous permet de sélectionner la meilleure zone de disponibilité pour vous, en fonction de l'état de santé du système et de la capacité disponible. Si vous lancez des instances additionnelles, ne spécifiez une zone que si vos nouvelles instances doivent être proches ou séparées de vos instances en cours d'exécution.

Migrer une instance vers une autre zone de disponibilité

Si nécessaire, vous pouvez migrer une instance d'une zone de disponibilité à une autre. Par exemple, supposons que vous essayez de modifier le type d'instance de votre instance et que nous ne pouvons pas lancer une instance du nouveau type dans la zone de disponibilité actuelle. Dans ce cas, vous pouvez migrer l'instance vers une zone de disponibilité où nous pouvons lancer une instance de ce type.

Le processus de migration comporte les étapes suivantes :

- Création d'une AMI à partir de l'instance d'origine
- Lancement d'une instance dans la nouvelle zone de disponibilité
- Mise à jour de la configuration de la nouvelle instance, comme indiqué dans la procédure suivante

Pour migrer une instance vers une autre zone de disponibilité

1. Créez une AMI à partir de l'instance. La procédure varie selon votre système d'exploitation et le type de volume du périphérique racine pour l'instance. Pour plus d'informations, consultez la documentation qui correspond à votre système d'exploitation et à votre volume du périphérique racine :
   - Création d'une AMI Linux basée sur Amazon EBS (p. 132)
   - Création d'une AMI Linux basée sur le stockage d'instance (p. 136)
   - Création d'une AMI Windows Amazon basée sur des volumes EBS
2. Si vous avez besoin de préserver l'adresse IPv4 privée d'une instance, vous devez supprimer le sous-réseau dans la zone de disponibilité actuelle, puis créer un sous-réseau dans la nouvelle zone de disponibilité avec la même plage d'adresses IPv4 que le sous-réseau d'origine. Veuillez noter que vous devez mettre fin à toutes les instances dans un sous-réseau avant de pouvoir le supprimer. Par conséquent, vous devez créer des AMI à partir de toutes les instances de votre sous-réseau de façon à pouvoir déplacer toutes les instances du sous-réseau actuel vers le nouveau sous-réseau.
3. Lancez une instance depuis une AMI que vous avez créée, en spécifiant la nouvelle zone de disponibilité ou le nouveau sous-réseau. Vous pouvez utiliser le même type d'instance que pour l'instance d'origine ou sélectionner un nouveau type d'instance. Pour plus d'informations, consultez Lancer des instances dans une zone de disponibilité (p. 14).
4. Si l'instance d'origine a une adresse IP Elastic associée, associez-la à la nouvelle instance. Pour plus d'informations, consultez Dissociation d'une adresse IP Elastic (p. 837).
5. Si l'instance d'origine est une Instance réservée, changez la zone de disponibilité pour votre réservation. (Si vous avez également changé le type d'instance, vous pouvez aussi modifier le type d'instance de votre réservation). Pour plus d'informations, consultez Envoi de demandes de modification (p. 363).
6. (Facultatif) Mettez fin à l'instance d'origine. Pour plus d'informations, consultez Mise hors service d'une instance (p. 647).

Zones locales

Un Zone locale est une extension d'une région AWS située à proximité géographique de vos utilisateurs. Les Zones locales ont leurs propres connexions à Internet et au support AWS Direct Connect, de sorte que

les ressources créées dans un Zone locale peuvent servir les utilisateurs locaux avec des communications à faible latence. Pour de plus amples informations, veuillez consulter les zones locales AWS.

Une Zone locale est représentée par un code de région suivi d'un identifiant qui indique l'emplacement ; par exemple, `us-west-2-lax-1a`.

Pour utiliser une Zone locale, vous devez d'abord l'activer. Pour de plus amples informations, veuillez consulter the section called “Activer Zones locales” (p. 17). Ensuite, créez un sous-réseau dans la Zone locale. Enfin, lancez l'une des ressources suivantes dans le sous-réseau de la Zone locale, afin que vos applications soient proches de vos utilisateurs finaux :

- Instances Amazon EC2
- Volumes Amazon EBS
- Serveurs de fichiers Amazon FSx
- Application Load Balancers
- Hôtes dédiés

Pour de plus amples informations sur les zones locales disponibles, veuillez consulter the section called “Régions disponibles” (p. 8).

Sommaire

- Décrire votre Zones locales (p. 16)
- Activer Zones locales (p. 17)
- Lancement d'instances dans un Zone locale (p. 17)

Décrire votre Zones locales

Vous pouvez utiliser la console Amazon EC2 ou l'interface de ligne de commande pour déterminer quels Zones locales sont disponibles pour votre compte. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

Pour rechercher votre Zones locales à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Affichez les options dans le sélecteur de région à partir de la barre de navigation.
3. Dans le panneau de navigation, choisissez Tableau de bord EC2.
4. Les Zones locales sont répertoriés sous Intégrité du service, État de la zone.

Pour rechercher votre Zones locales à l'aide de AWS CLI

1. Utilisez la commande describe-availability-zones comme suit pour décrire les Zones locales dans la région spécifiée.

```
aws ec2 describe-availability-zones --region region-name
```

2. Utilisez la commande describe-availability-zones comme suit pour décrire les Zones locales, qu'elles soient activées ou non.

```
aws ec2 describe-availability-zones --all-availability-zones
```

Pour rechercher votre Zones locales à l'aide de Outils AWS pour Windows PowerShell

Utilisez la commande Get-EC2AvailabilityZone comme suit pour décrire les Zones locales dans la région spécifiée.

```
PS C:\> Get-EC2AvailabilityZone -Region region-name
```

Activer Zones locales

Avant de spécifier une Zone locale pour une ressource ou un service, vous devez activer Zones locales.

Considération

Il est possible que certaines ressources AWS ne soient pas disponibles dans toutes les régions. Assurez-vous que vous pouvez créer les ressources dont vous avez besoin dans les régions ou Zones locales souhaitées avant de lancer une instance dans une Zone locale spécifique.

Pour activer les Zones locales à l'aide de la console :

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le coin supérieur gauche de la page, sélectionnez Nouvelle expérience EC2. Vous ne pouvez pas exécuter cette tâche à l'aide de l'ancienne console.
3. Dans le sélecteur de région de la barre de navigation, sélectionnez la région de la Zone locale.
4. Dans le panneau de navigation, choisissez Tableau de bord EC2.
5. En haut à droite de la page, choisissez Attributs du compte, Zones.
6. Choisissez Gérer.
7. Pour Groupe Zone, choisissez Activé.
8. Choisissez Mettre à jour le groupe de zones.

Pour activer les Zones locales à l'aide de l'AWS CLI

- Utilisez la commande modify-availability-zone-group.

Lancement d'instances dans un Zone locale

Lorsque vous lancez une instance, vous pouvez spécifier un sous-réseau qui se trouve dans un Zone locale. Vous allouez également une adresse IP à partir d'un groupe de frontières réseau, qui est un ensemble unique de zones de disponibilité, de Zones locales ou de zones de longueur d'onde à partir desquelles AWS annonce des adresses IP, par exemple `us-west-2-lax-1a`.

Vous pouvez allouer les adresses IP suivantes à partir d'un groupe de frontières réseau :

- Adresses IPv4 Elastic fournies par Amazon
- Adresses IPv6 de VPC fournies par Amazon

Pour lancer une instance dans un Zone locale :

1. Activez Zones locales. Pour de plus amples informations, veuillez consulter Activer Zones locales (p. 17).
2. Créez un VPC dans une région qui prend en charge le Zone locale. Pour de plus amples informations, veuillez consulter Création d'un VPC dans le Amazon VPC Guide de l'utilisateur.
3. Créez un sous-réseau. Sélectionnez le Zone locale lorsque vous créez le sous-réseau. Pour de plus amples informations, veuillez consulter Création d'un sous-réseau dans votre VPC dans le Amazon VPC Guide de l'utilisateur.
4. Lancez une instance et sélectionnez le sous-réseau que vous avez créé dans le Zone locale. Pour de plus amples informations, veuillez consulter Lancer votre instance (p. 529).

Zones de longueur d'onde

AWS Wavelength permet aux développeurs de créer des applications qui offrent des latences ultra-faibles aux appareils mobiles et aux utilisateurs finaux. Wavelength déploie des services de calcul et de stockage AWS standard à la périphérie des réseaux 5G des opérateurs de télécommunications. Les développeurs peuvent étendre un VPC à une ou plusieurs zones de longueur d'onde, puis utiliser des ressources AWS comme les instances Amazon EC2 pour exécuter des applications nécessitant une latence ultra-faible et une connexion aux services AWS dans la région.

Une zone de longueur d'onde est une zone isolée située à l'emplacement du transporteur où l'infrastructure de longueur d'onde est déployée. Les zones de longueur d'onde sont liées à une région. Une zone de longueur d'onde est une extension logique d'une région et est gérée par le plan de contrôle de la région.

Une zone de longueur d'onde est représentée par un code de région suivi d'un identificateur qui indique la zone de longueur d'onde, par exemple `us-east-1-wl1-bos-wlz-1`.

Pour utiliser une zone de longueur d'onde, vous devez d'abord vous inscrire à la zone. Pour plus d'informations, consultez the section called “Activer les zones de longueur d'onde” (p. 19). Ensuite, créez un sous-réseau dans la zone de longueur d'onde. Enfin, lancez vos ressources dans le sous-réseau de zones de longueur d'onde, afin que vos applications soient plus proches de vos utilisateurs finaux.

Les zones de longueur d'onde ne sont pas disponibles dans toutes les régions. Pour de plus amples informations sur les régions qui prennent en charge les zones de longueur d'onde, veuillez consulter Zones de longueur d'onde dans le Manuel du développeur AWS Wavelength.

Sommaire

- Décrire vos zones de longueur d'onde (p. 19)
- Activer les zones de longueur d'onde (p. 19)
- Lancement d'instances dans une zone de longueur d'onde (p. 20)

Décrire vos zones de longueur d'onde

Vous pouvez utiliser la console Amazon EC2 ou l'interface de ligne de commande pour déterminer quelles zones de longueur d'onde sont disponibles pour votre compte. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

Pour rechercher vos zones de longueur d'onde à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Affichez les options dans le sélecteur de région à partir de la barre de navigation.
3. Dans le panneau de navigation, choisissez Tableau de bord EC2.
4. Les zones de longueur d'onde sont répertoriées sous Intégrité du service, État de zone.

Pour rechercher vos zones de longueur d'onde à l'aide de AWS CLI

1. Utilisez la commande describe-availability-zones comme suit pour décrire les zones de longueur d'onde dans la région spécifiée.

```
aws ec2 describe-availability-zones --region region-name
```

2. Utilisez la commande describe-availability-zones comme suit pour décrire les zones de longueur d'onde, quel que soit leur statut d'inscription.

```
aws ec2 describe-availability-zones --all-availability-zones
```

Pour rechercher votre zone de longueur d'onde à l'aide de Outils AWS pour Windows PowerShell

Utilisez la commande Get-EC2AvailabilityZone comme suit pour décrire les zones de longueur d'onde dans la région spécifiée.

```
PS C:\> Get-EC2AvailabilityZone -Region region-name
```

Activer les zones de longueur d'onde

Avant de spécifier une zone de longueur d'onde pour une ressource ou un service, vous devez activer les zones de longueur d'onde.

Considérations

- Vous devez demander l'accès pour utiliser les zones de longueur d'onde. Pour de plus amples informations sur la demande d'accès à la zone de longueur d'onde, veuillez consulter AWS Wavelength.
- Certaines ressources AWS ne sont pas disponibles dans toutes les régions. Assurez-vous de pouvoir créer les ressources dont vous avez besoin dans la région ou la zone de longueur d'onde souhaitée avant de lancer une instance dans une zone de longueur d'onde spécifique.

Pour vous inscrire aux zones de longueur d'onde à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le coin supérieur gauche de la page, sélectionnez Nouvelle expérience EC2. Vous ne pouvez pas exécuter cette tâche à l'aide de l'ancienne console.
3. Dans le sélecteur de région de la barre de navigation, sélectionnez la région pour la zone de longueur d'onde.
4. Dans le panneau de navigation, choisissez Tableau de bord EC2.
5. En haut à droite de la page, choisissez Attributs du compte, Zones.
6. Sous Zones de longueur d'onde, activez chaque zone de longueur d'onde.
7. Activez la zone de longueur d'onde.

Pour activer les zones de longueur d'onde à l'aide de AWS CLI

Utilisez la commande modify-availability-zone-group.

Lancement d'instances dans une zone de longueur d'onde

Lorsque vous lancez une instance, vous pouvez spécifier un sous-réseau qui se trouve dans une zone de longueur d'onde. Vous allouez également une adresse IP de transporteur à partir d'un groupe de frontières réseau, qui est un ensemble unique de zones de disponibilité, de zones locales ou de zones de longueur d'onde à partir desquelles AWS annonce des adresses IP, par exemple `us-east-1-wl1-bos-wlz-1`.

Pour de plus amples informations sur le lancement d'une instance dans une zone de longueur d'onde, veuillez consulter Démarrer avec AWS Wavelength dans le Manuel du développeur AWS Wavelength.

AWS Outposts

AWS Outposts est un service entièrement géré qui étend l'infrastructure, les services, les API et les outils AWS aux sites du client. En fournissant un accès local à l'infrastructure gérée par AWS, AWS Outposts permet aux clients de créer et d'exécuter des applications sur site à l'aide des mêmes interfaces de programmation que dans les régions AWS, tout en utilisant les ressources de calcul et de stockage locales pour réduire la latence et les besoins de traitement des données locales.

Un outpost est un pool de capacités de calcul et de stockage AWS déployées sur un site client. AWS exploite, surveille et gère cette capacité dans le cadre d'une région AWS. Vous pouvez créer des sous-réseaux sur votre outpost et les spécifier lorsque vous créez des ressources AWS telles que des instances EC2, des volumes EBS, des clusters ECS et des instances RDS. Les instances se trouvant dans des sous-réseaux outpost communiquent avec d'autres instances de la région AWS à l'aide d'adresses IP privées se trouvant toutes dans le même VPC.

Pour commencer à utiliser AWS Outposts, vous devez créer un outpost et commander une capacité outpost. Pour de plus amples informations sur les configurations outposts, veuillez consulter notre catalogue. Une fois votre équipement Ooutpost installé, la capacité de calcul et de stockage est disponible pour vous lorsque vous lancez des instances Amazon EC2 et que vous créez des volumes Amazon EBS sur votre outpost.

Lancement d'instances sur un outpost

Vous pouvez lancer des instances EC2 dans le sous-réseau outpost que vous avez créé. Les groupes de sécurité contrôlent le trafic entrant et sortant pour les instances d'un sous-réseau outpost, comme ils le font pour les instances d'un sous-réseau de zone de disponibilité. Pour vous connecter à une instance EC2 dans un sous-réseau outpost, vous pouvez spécifier une paire de clés lorsque vous lancez l'instance, comme vous le faites pour les instances d'un sous-réseau de zone de disponibilité.

Le volume racine doit être de 30 Go ou moins. Vous pouvez spécifier des volumes de données dans le mappage de périphériques de blocs de l'AMI ou de l'instance pour fournir un stockage supplémentaire.

Pour couper les blocs inutilisés du volume de démarrage, veuillez consulter How to Build Sparse EBS Volumes (Comment créer des volumes EBS fragmentés) dans le blog AWS Partner Network.

Nous vous recommandons d'augmenter le délai d'attente de NVMe pour le volume racine. Pour plus d'informations, consultez Expiration de l'intégration des E/S (p. 1212).

Pour de plus amples informations sur la création d'un outpost, veuillez consulter Démarrage avec AWS Outposts dans le Guide de l'utilisateur AWS Outposts.

Création d'un volume sur un outpost

Vous pouvez créer des volumes EBS dans le sous-réseau outpost que vous avez créé. Lorsque vous créez le volume, spécifiez l'ARN (Amazon Resource Name) de l'outpost.

La commande create-volume suivante crée un volume vide de 50 Go sur l'outpost spécifié.

```
aws ec2 create-volume --availability-zone us-east-2a --outpost-arn arn:aws:outposts:us-
east-2:123456789012:outpost/op-03e6fecad652a6138 --size 50
```

Vous devez détacher un volume Outpost avant de pouvoir le modifier. Pour plus d'informations sur le détachement des volumes, consultez la section Détacher un volume Amazon EBS d'une instance Linux (p. 1127).

Volume du périphérique racine Amazon EC2

Lorsque vous lancez une instance, le volume du périphérique racine contient l'image utilisée pour démarrer l'instance. Quand nous avons introduit Amazon EC2, toutes les AMI étaient basées sur le stockage d'instance Amazon EC2, ce qui signifie que le périphérique racine d'une instance lancée à partir de l'AMI est un volume de stockage d'instance créé à partir d'un modèle stocké dans Amazon S3. Après Amazon EBS, nous avons introduit les AMI basées sur les volumes Amazon EBS. Cela signifie que le périphérique racine d'une instance lancée à partir de l'AMI est un volume Amazon EBS créé à partir d'un instantané Amazon EBS.

Vous avez le choix entre les AMI basées sur le stockage d'instance Amazon EC2 et les AMI basées sur les volumes Amazon EBS. Il est recommandé d'utiliser les AMI basées sur les volumes Amazon EBS, car ils se lancent plus rapidement et utilisent le stockage permanent.

> **Important**
>
> Seuls les types d'instance suivants prennent en charge un volume de stockage d'instance en tant que périphérique racine : C3, D2, G2, I2, M3 et R3.

Pour plus d'informations sur l'utilisation des noms d'unité Amazon EC2 pour vos volumes racines, consultez Noms d'appareil pour les instances Linux (p. 1287).

Sommaire

- Concepts du stockage de périphérique racine (p. 21)
- Choix d'une AMI par type de périphérique racine (p. 23)
- Détermination du type de périphérique racine de votre instance (p. 24)
- Modification du volume racine pour qu'il persiste (p. 25)

Concepts du stockage de périphérique racine

Vous pouvez lancer une instance à partir d'une AMI basée sur le stockage d'instance ou d'une AMI basée sur Amazon EBS. La description d'une AMI inclut son type. Vous remarquerez qu'en certains endroits, il

est fait référence au périphérique racine comme `ebs` (pour les AMI basées sur les volumes Amazon EBS) ou comme `instance store` (pour les AMI basées sur le stockage d'instance). Ce point est important, car il existe des différences importantes entre ce que vous pouvez faire avec chaque type d'AMI. Pour plus d'informations sur ces différences, consultez Stockage pour le périphérique racine (p. 106).

Instances basées sur le stockage d'instance

Les instances qui utilisent les stockages d'instance pour le périphérique racine ont automatiquement un ou plusieurs volumes de stockage d'instance disponibles, l'un faisant office de volume du périphérique racine. Quand une instance est lancée, l'image utilisée pour démarrer l'instance est copiée sur le volume racine. Notez que vous pouvez utiliser le cas échéant des volumes de stockage d'instance supplémentaires, suivant le type d'instance.

Les données présentes sur les volumes de stockage d'instance demeurent aussi longtemps que l'instance s'exécute, mais ces données sont supprimées quand il est procédé à la terminaison de l'instance (les instances basées sur le stockage d'instance ne prennent pas en charge l'action Stop) ou en cas de défaillance de l'instance (problèmes rencontrés par un lecteur sous-jacent, par exemple).

Après qu'une instance basée sur le stockage d'instances a échoué ou s'est terminée, elle ne peut pas être restaurée. Si vous prévoyez d'utiliser les instances basées sur le stockage d'instance Amazon EC2, il est vivement recommandé de répartir les données de vos stockages d'instance entre plusieurs zones de disponibilité. Vous devez aussi sauvegarder régulièrement les données critiques de vos volumes de stockage d'instance sur un stockage permanent.

Pour plus d'informations, consultez Stockage d'instance Amazon EC2 (p. 1264).

Instances basées sur les volumes Amazon EBS

Les instances qui ont recours à Amazon EBS pour le périphérique racine sont automatiquement associées à un volume Amazon EBS. Lorsque vous lancez une instance basée sur les volumes Amazon EBS, nous créons un volume Amazon EBS pour chaque instantané Amazon EBS référencé par l'AMI que vous utilisez. Vous pouvez aussi utiliser d'autres volumes Amazon EBS ou des volumes de stockage d'instance, suivant le type d'instance.

Une instance basée sur Amazon EBS peut être arrêtée et redémarrée ultérieurement sans affecter les données stockées dans les volumes attachés. Il existe diverses tâches liées aux instances et aux volumes que vous pouvez effectuer quand une instance basée sur Amazon EBS est dans un état arrêté. Par exemple, vous pouvez modifier les propriétés de l'instance, changer sa taille ou mettre à jour le noyau qu'elle utilise, ou vous pouvez aussi attacher votre volume racine à une autre instance en cours d'exécution à des fins de débogage ou autre.

Si une instance basée sur Amazon EBS échoue, vous pouvez restaurer votre session en suivant l'une de ces méthodes :

- Arrêtez l'instance et redémarrez-la (essayez cette méthode en premier).
- Prenez automatiquement un instantané de tous les volumes appropriés et créez un nouvel AMI. Pour plus d'informations, consultez Création d'une AMI Linux basée sur Amazon EBS (p. 132).
- Attachez le volume à la nouvelle instance à l'aide des étapes suivantes :
  1. Créez un instantané du volume racine.
  2. Inscrivez un nouvel AMI à l'aide de l'instantané.
  3. Lancez une nouvelle instance à partir du nouvel AMI.
  4. Détachez les volumes Amazon EBS restants de l'ancienne instance.
  5. Rattachez les volumes Amazon EBS à la nouvelle instance.

Pour plus d'informations, consultez Volumes Amazon EBS (p. 1086).

Choix d'une AMI par type de périphérique racine

L'AMI que vous spécifiez au lancement de votre instance détermine le type de volume du périphérique racine de votre instance.

Pour choisir une AMI basée sur des volumes Amazon EBS avec la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, sélectionnez AMI.
3. Dans les listes de filtres, sélectionnez le type d'image (par exemple, Public images). Dans la barre de recherche, choisissez Platform pour sélectionner le système d'exploitation (comme Amazon Linux) et Root Device Type pour sélectionner les EBS images.
4. (Facultatif) Pour obtenir des informations supplémentaires afin de vous aider à choisir, choisissez l'icône Show/Hide Columns, mettez à jour les colonnes à afficher et choisissez Close.

5. Choisissez une AMI et notez l'ID d'AMI.

Pour choisir une AMI basée sur le stockage d'instance avec la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, sélectionnez AMI.
3. Dans les listes de filtres, sélectionnez le type d'image (par exemple, Public images). Dans la barre de recherche, choisissez Platform pour sélectionner le système d'exploitation (comme Amazon Linux) et Root Device Type pour sélectionner les Instance store.
4. (Facultatif) Pour obtenir des informations supplémentaires afin de vous aider à choisir, choisissez l'icône Show/Hide Columns, mettez à jour les colonnes à afficher et choisissez Close.
5. Choisissez une AMI et notez l'ID d'AMI.

Pour vérifier le volume du périphérique racine d'une AMI à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-images (AWS CLI)
- Get-EC2Image (Outils AWS pour Windows PowerShell)

Détermination du type de périphérique racine de votre instance

Nouvelle console

Pour déterminer le type de périphérique racine d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez l'instance.
3. Sous l'onglet Stockage, sous Détails de l'appareil racine, vérifiez la valeur de Type d'appareil racine comme suit :
   - Si la valeur est `EBS`, il s'agit d'une instance basée sur Amazon EBS.
   - Si la valeur est `INSTANCE-STORE`, il s'agit d'une instance basée sur le stockage d'instance.

Ancienne console

Pour déterminer le type de périphérique racine d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez l'instance.
3. Vérifiez la valeur de Type d'appareil racine sous l'onglet Description comme suit :
   - Si la valeur est `ebs`, il s'agit d'une instance basée sur Amazon EBS.
   - Si la valeur est `instance store`, il s'agit d'une instance basée sur le stockage d'instance.

Pour déterminer le type de périphérique racine d'une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-instances (AWS CLI)
- Get-EC2Instance

Modification du volume racine pour qu'il persiste

Par défaut, le volume racine d'une AMI basée sur Amazon EBS est supprimé quand l'instance se termine. Vous pouvez modifier le comportement par défaut pour vous assurer que le volume persiste après la fin de l'instance. Pour modifier le comportement par défaut, définissez l'attribut `DeleteOnTermination` avec la valeur `false` à l'aide d'un mappage de périphérique de stockage en mode bloc.

Rubriques

- Configuration du volume racine pour qu'il persiste pendant le lancement de l'instance (p. 25)
- Configuration du volume racine pour qu'il persiste pour une instance existante (p. 26)
- Confirmation qu'un volume racine est configuré pour persister (p. 26)

Configuration du volume racine pour qu'il persiste pendant le lancement de l'instance

Vous pouvez configurer le volume racine pour qu'il persiste lorsque vous lancez une instance à l'aide de la console Amazon EC2 ou des outils de ligne de commande.

Configurer le volume racine pour qu'il persiste lorsque vous lancez une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, puis Lancer une instance.
3. Sur la page Choose an Amazon Machine Image (AMI), sélectionnez l'AMI à utiliser, puis choisissez Select.
4. Suivez l'Assistant pour compléter les pages Choisir un type d'instance et Configurer les détails de l'instance.
5. Sur la page Add Storage, désélectionnez Delete On Termination pour le volume racine.
6. Complétez les pages restantes de l'Assistant, puis sélectionnez Launch.

Configurer le volume racine pour qu'il persiste lorsque vous lancez une instance à l'aide de l'AWS CLI

Utilisez la commande run-instances et incluez un mappage de périphérique en mode bloc qui définit l'attribut `DeleteOnTermination` avec la valeur `false`.

```
$ aws ec2 run-instances --block-device-mappings file://mapping.json ...other parameters...
```

Spécifiez les éléments suivants dans `mapping.json`.

```
[
    {
        "DeviceName": "/dev/sda1",
        "Ebs": {
            "DeleteOnTermination": false
        }
    }
]
```

Configurer le volume racine pour qu'il persiste lorsque vous lancez une instance à l'aide de Outils pour Windows PowerShell

Utilisez la commande New-EC2Instance et incluez un mappage de périphérique en mode bloc qui définit l'attribut `DeleteOnTermination` avec la valeur `false`.

```
C:\> $ebs = New-Object Amazon.EC2.Model.EbsBlockDevice
C:\> $ebs.DeleteOnTermination = $false
C:\> $bdm = New-Object Amazon.EC2.Model.BlockDeviceMapping
C:\> $bdm.DeviceName = "dev/xvda"
C:\> $bdm.Ebs = $ebs
C:\> New-EC2Instance -ImageId ami-0abcdef1234567890 -BlockDeviceMapping $bdm ...other
 parameters...
```

Configuration du volume racine pour qu'il persiste pour une instance existante

Vous pouvez configurer le volume racine pour qu'il persiste pour une instance en cours d'exécution à l'aide des outils de ligne de commande uniquement.

Configurer le volume racine pour qu'il persiste pour une instance existante à l'aide de AWS CLI

Utilisez la commande modify-instance-attribute et incluez un mappage de périphérique en mode bloc qui définit l'attribut `DeleteOnTermination` sur `false`.

```
aws ec2 modify-instance-attribute --instance-id i-1234567890abcdef0 --block-device-mappings
 "[{\"DeviceName\": \"/dev/xvda\",\"Ebs\":{\"DeleteOnTermination\":false}}]"
```

Pour configurer le volume racine pour qu'il persiste pour une instance en cours d'exécution à l'aide de Outils AWS pour Windows PowerShell

Utilisez la commande Edit-EC2InstanceAttribute et incluez un mappage de périphérique en mode bloc qui définit l'attribut `DeleteOnTermination` sur `false`.

```
C:\> $ebs = New-Object Amazon.EC2.Model.EbsInstanceBlockDeviceSpecification
C:\> $ebs.DeleteOnTermination = $false
C:\> $bdm = New-Object Amazon.EC2.Model.InstanceBlockDeviceMappingSpecification
C:\> $bdm.DeviceName = "/dev/xvda"
C:\> $bdm.Ebs = $ebs
C:\> Edit-EC2InstanceAttribute -InstanceId i-1234567890abcdef0 -BlockDeviceMapping $bdm
```

Confirmation qu'un volume racine est configuré pour persister

Vous pouvez confirmer qu'un volume racine est configuré pour persister à l'aide de la console Amazon EC2 ou des outils de ligne de commande.

Nouvelle console

Confirmer qu'un volume racine est configuré pour persister à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances, puis sélectionnez l'instance.
3. Dans l'onglet Stockage, sous Bloquer les appareil, recherchez l'entrée du volume racine. Si la valeur Supprimer lors de la résiliation est définie avec la valeur `No`, le volume est configuré pour persister.

Ancienne console

Confirmer qu'un volume racine est configuré pour persister à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances, puis sélectionnez l'instance.
3. Dans l'onglet Description choisissez l'entrée pour le Root device (Périphérique racine). Si la valeur Supprimer lors de la résiliation est définie avec la valeur `False`, le volume est configuré pour persister.

Confirmer qu'un volume racine est configuré pour persister à l'aide de l'AWS CLI

Utilisez la commande describe-instances et vérifiez que l'attribut `DeleteOnTermination` de l'élément de réponse `BlockDeviceMappings` est défini avec la valeur `false`.

```
$ aws ec2 describe-instances --instance-id i-1234567890abcdef0
```

```
...
    "BlockDeviceMappings": [
    {
        "DeviceName": "/dev/sda1",
        "Ebs": {
            "Status": "attached",
            "DeleteOnTermination": false,
            "VolumeId": "vol-1234567890abcdef0",
            "AttachTime": "2013-07-19T02:42:39.000Z"
        }
    }
...
```

Confirmer qu'un volume racine est configuré pour persister à l'aide de Outils AWS pour Windows PowerShell

Utilisez Get-EC2Instance et vérifiez que l'attribut `DeleteOnTermination` de l'élément de réponse `BlockDeviceMappings` est défini avec la valeur `false`.

```
C:\> (Get-EC2Instance -InstanceId i-i-1234567890abcdef0).Instances.BlockDeviceMappings.Ebs
```

Configuration avec Amazon EC2

Effectuez les tâches décrites dans cette section pour configurer le lancement d'une instance Amazon EC2 pour la première fois :

1. Inscrivez-vous à AWS (p. 28)
2. Création d'une paire de clés (p. 28)
3. Création d'un groupe de sécurité (p. 30)

Lorsque vous avez terminé, vous serez prêt pour le tutoriel Démarrez avec Amazon EC2 (p. 32).

Inscrivez-vous à AWS

Lorsque vous vous inscrivez à Amazon Web Services (AWS), votre compte AWS est automatiquement inscrit à tous les services d'AWS, y compris Amazon EC2. Seuls les services que vous utilisez vous sont facturés.

Avec Amazon EC2, vous ne payez que ce que vous utilisez. Si vous êtes un nouveau client AWS, vous pouvez démarrer avec Amazon EC2 gratuitement. Pour plus d'informations, consultez la page sur l'offre gratuite AWS.

Si vous possédez déjà un compte AWS, passez à la prochaine étape. Si tel n'est pas le cas, observez la procédure suivante pour en créer un.

Pour créer un compte AWS

1. Ouvrez https://portal.aws.amazon.com/billing/signup.
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.

Création d'une paire de clés

AWS utilise le chiffrement de clé publique pour sécuriser les informations de connexion de votre instance. Une instance Linux n'ayant pas de mot de passe, vous utilisez une paire de clés pour vous connecter à votre instance en toute sécurité. Vous indiquez le nom de la paire de clés au lancement de votre instance, puis fournissez la clé privée lorsque vous vous connectez avec SSH.

Si vous n'avez pas encore créé de paire de clés, vous pouvez le faire à l'aide de la console Amazon EC2. Notez que si vous prévoyez de lancer des instances dans plusieurs régions, vous devez créer une paire de clés dans chaque région. Pour plus d'informations sur les régions, consultez Régions et zones (p. 7).

Vous pouvez créer une paire de clés à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour créer votre paire de clés

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le volet de navigation, cliquez sur Key Pairs.
3. Choisissez Créer une paire de clés.
4. Dans Nom, entrez un nom descriptif pour la paire de clés. Amazon EC2 associe la clé publique au nom que vous spécifiez comme nom de clé. Le nom peut inclure jusqu'à 255 caractères ASCII. Il ne peut pas inclure d'espaces de début ou de fin.
5. Dans File format (Format de fichier), choisissez le format dans lequel enregistrer la clé privée. Pour enregistrer la clé privée dans un format qui peut être utilisé avec OpenSSH, choisissez pem. Pour enregistrer la clé privée dans un format qui peut être utilisé avec PuTTY, choisissez ppk.
6. Choisissez Créer une paire de clés.
7. Le fichier de clé privée est automatiquement téléchargé dans votre navigateur. Le nom de fichier de base est le nom que vous avez spécifié pour votre paire de clés, et l'extension de nom de fichier est déterminée par le format de fichier que vous avez choisi. Enregistrez le fichier de clé privée en lieu sûr.

   Important

   C'est votre seule occasion d'enregistrer le fichier de clé privée.
8. Si vous envisagez d'utiliser un client SSH sur un ordinateur macOS ou Linux pour vous connecter à votre instance Linux, utilisez la commande suivante pour définir les autorisations de votre fichier de clé privée afin d'être la seule personne autorisée à le lire.

```
chmod 400 my-key-pair.pem
```

   Si vous ne définissez pas ces autorisations, vous ne pouvez pas vous connecter à votre instance à l'aide de cette paire de clés. Pour plus d'informations, consultez Erreur : fichier de clé privée non protégé (p. 1334).

Ancienne console

Pour créer votre paire de clés

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous NETWORK & SECURITY, choisissez Key Pairs.

   Note

   Le panneau de navigation se trouve sur la gauche de la console Amazon EC2. Si le volet n'est pas visible, il se peut qu'il soit réduit ; cliquez sur la flèche pour le développer.
3. Choisissez Create Key Pair.
4. Pour Nom de la paire de clés, entrez un nom pour la nouvelle paire de clés, puis choisissez Créer. Le nom peut inclure jusqu'à 255 caractères ASCII. Il ne peut pas inclure d'espaces de début ou de fin.
5. Le fichier de clé privée est automatiquement téléchargé dans votre navigateur. Le nom de fichier de base est le nom que vous avez spécifié comme nom de votre paire de clés et l'extension du nom de fichier est `.pem`. Enregistrez le fichier de clé privée en lieu sûr.

   Important

   C'est votre seule occasion d'enregistrer le fichier de clé privée.
6. Si vous envisagez d'utiliser un client SSH sur un ordinateur macOS ou Linux pour vous connecter à votre instance Linux, utilisez la commande suivante pour définir les autorisations de votre fichier de clé privée afin d'être la seule personne autorisée à le lire.

```
chmod 400 my-key-pair.pem
```

Si vous ne définissez pas ces autorisations, vous ne pouvez pas vous connecter à votre instance à l'aide de cette paire de clés. Pour plus d'informations, consultez Erreur : fichier de clé privée non protégé (p. 1334).

Pour plus d'informations, consultez Paires de clés Amazon EC2 et instances Linux (p. 1047).

Création d'un groupe de sécurité

Les groupes de sécurité font office de pare-feu pour les instances associées, en contrôlant le trafic entrant et le trafic sortant au niveau de l'instance. Vous devez ajouter des règles à un groupe de sécurité qui vous permettent de vous connecter à votre adresse IP avec SSH. Vous pouvez aussi ajouter des règles qui permettent les accès HTTP et HTTPS entrants et sortants depuis n'importe quel emplacement.

Notez que si vous prévoyez de lancer des instances dans plusieurs régions, vous devez créer un groupe de sécurité dans chaque région. Pour plus d'informations sur les régions, consultez Régions et zones (p. 7).

Conditions préalables requises

Vous aurez besoin de l'adresse IPv4 publique de votre ordinateur local. L'éditeur de groupe de sécurité de la console Amazon EC2 peut détecter automatiquement l'adresse IPv4 publique pour vous. Sinon, vous pouvez utiliser l'expression de recherche « quelle est mon adresse IP ? » dans un navigateur Internet, ou utiliser le service suivant : Check IP. Si votre connexion s'effectue via un fournisseur de services Internet (ISP) ou derrière un pare-feu sans adresse IP statique, vous devez déterminer la plage d'adresses IP utilisée par les ordinateurs clients.

Vous pouvez créer un groupe de sécurité personnalisé à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour créer un groupe de sécurité avec le principe du moindre privilège

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez une région pour le groupe de sécurité. Comme les groupes de sécurité sont propres à une région, vous devez sélectionner la même région que celle que vous avez créée avec votre paire de clés.
3. Dans le volet de navigation, choisissez Groupes de sécurité.
4. Sélectionnez Créer un groupe de sécurité.
5. Dans la section Basic details (Détails de base) procédez comme suit.
   a. Entrez un nom et une description pour le nouveau groupe de sécurité. Choisissez un nom facile à retenir, tel que votre nom d'utilisateur suivi de _SG_ et du nom de la région. Par exemple, moi-SG-uswest2.
   b. Dans la liste VPC sélectionnez votre VPC par défaut pour la région.
6. Dans la section Inbound rules (Règles entrantes), créez les règles suivantes (choisissez Add rule (Ajouter une règle) pour chaque nouvelle règle) :
   - Choisissez HTTP dans la liste Type et vérifiez que Source a bien la valeur N'importe où (`0.0.0.0/0`).
   - Choisissez HTTPS dans la liste Type et vérifiez que Source a bien la valeur N'importe où (`0.0.0.0/0`).
   - Choisissez SSH dans la liste Type. Dans la zone Source, choisissez Mon IP pour remplir automatiquement le champ avec l'adresse IPv4 publique de votre ordinateur local. Sinon,

choisissez Personnalisé et spécifiez l'adresse IPv4 publique de votre ordinateur ou réseau en notation CIDR. Pour spécifier une adresse IP individuelle en notation CIDR, ajoutez le suffixe de routage `/32` : `203.0.113.25/32`, par exemple. Si votre entreprise alloue des adresses à partir d'une plage, spécifiez la plage complète, telle que `203.0.113.0/24`.

> **Warning**
>
> Pour des raisons de sécurité, il est déconseillé d'autoriser l'accès de SSH à votre instance à partir de toutes les adresses IPv4 (`0.0.0.0/0`), si ce n'est à titre de test et pour une très brève durée.

7. Sélectionnez Créer un groupe de sécurité.

Ancienne console

Pour créer un groupe de sécurité avec le principe du moindre privilège

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez Create Security Group.
4. Entrez un nom et une description pour le nouveau groupe de sécurité. Choisissez un nom facile à retenir, tel que votre nom d'utilisateur suivi de _SG_ et du nom de la région. Par exemple, moi-SG-uswest2.
5. Dans la liste VPC sélectionnez votre VPC par défaut pour la région.
6. Sous l'onglet Inbound (Entrant), créez les règles suivantes (choisissez Add rule (Ajouter une règle) pour chaque nouvelle règle) :
   - Choisissez HTTP dans la liste Type et vérifiez que Source a bien la valeur N'importe où (`0.0.0.0/0`).
   - Choisissez HTTPS dans la liste Type et vérifiez que Source a bien la valeur N'importe où (`0.0.0.0/0`).
   - Choisissez SSH dans la liste Type. Dans la zone Source, choisissez Mon IP pour remplir automatiquement le champ avec l'adresse IPv4 publique de votre ordinateur local. Sinon, choisissez Personnalisé et spécifiez l'adresse IPv4 publique de votre ordinateur ou réseau en notation CIDR. Pour spécifier une adresse IP individuelle en notation CIDR, ajoutez le suffixe de routage `/32` : `203.0.113.25/32`, par exemple. Si votre entreprise alloue des adresses à partir d'une plage, spécifiez la plage complète, telle que `203.0.113.0/24`.

     > **Warning**
     >
     > Pour des raisons de sécurité, il est déconseillé d'autoriser l'accès de SSH à votre instance à partir de toutes les adresses IPv4 (`0.0.0.0/0`), si ce n'est à titre de test et pour une très brève durée.
7. Sélectionnez Créer.

Command line

Pour créer un groupe de sécurité avec le principe du moindre privilège

Utilisez l'une des commandes suivantes :

- create-security-group (AWS CLI)
- New-EC2SecurityGroup (Outils AWS pour Windows PowerShell)

Pour plus d'informations, consultez Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062).

Didacticiel : Premiers pas avec les instances Amazon EC2 Linux

Utilisez ce didacticiel pour commencer avec Amazon Elastic Compute Cloud (Amazon EC2). Vous apprendrez à lancer une instance Linux, à vous y connecter et à l'utiliser. Une instance est un serveur virtuel dans le cloud AWS. Avec Amazon EC2, vous pouvez installer et configurer le système d'exploitation et les applications qui s'exécutent sur votre instance.

Pour démarrer avec une instance Windows, veuillez consulter Démarrez avec les instances Windows Amazon EC2.

Lorsque vous vous inscrivez à AWS, vous pouvez commencer à utiliser Amazon EC2 grâce à l'offre gratuite d'AWS. Si vous avez créé votre compte AWS il y a moins de 12 mois et que vous n'avez pas encore dépassé les avantages de l'offre gratuite pour Amazon EC2, cela ne vous coûtera rien de suivre ce didacticiel, parce que nous vous aidons à sélectionner des options faisant partie des avantages de l'offre gratuite. Sinon, vous devrez payer les frais d'utilisation Amazon EC2 standard à partir du moment où vous lancez l'instance et jusqu'à ce que vous mettiez fin à celle-ci (ce qui constitue la tâche finale de ce didacticiel), même si elle reste inactive.

Table des matières

- Présentation (p. 32)
- Prérequis (p. 33)
- Étape 1 : Lancement d'une instance (p. 33)
- Étape 2 : Connexion à l'instance (p. 34)
- Étape 3 : Nettoyage de votre instance (p. 35)
- Étapes suivantes (p. 35)

Présentation

L'instance est une instance basée sur Amazon EBS (ce qui signifie que le volume racine est un volume EBS). Vous pouvez spécifier la zone de disponibilité dans laquelle s'exécute votre instance ou laisser Amazon EC2 la sélectionner pour vous. Lorsque vous lancez votre instance, vous la sécurisez en spécifiant une paire de clés (key pair) et un groupe de sécurité. Lorsque vous vous connectez à votre instance, vous devez indiquer la clé privée de la paire de clés que vous avez spécifiée au lancement de l'instance.

Tâches

Pour suivre ce didacticiel, effectuez les tâches suivantes :

1. Lancez une instance (p. 33)
2. Connectez-vous à votre instance (p. 34)
3. Nettoyez votre instance (p. 35)

Didacticiels connexes

- Si vous préférez lancer une instance Windows, consultez ce didacticiel dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows : Premiers pas avec les instances Windows Amazon EC2.
- Si vous préférez utiliser la ligne de commande, consultez ce didacticiel dans le manuel AWS Command Line Interface Guide de l'utilisateur : Utilisation de Amazon EC2 via l'interface de ligne de commande AWS.

Prérequis

Avant de commencer, assurez-vous d'avoir terminé les étapes de Configuration avec Amazon EC2 (p. 28).

Étape 1 : Lancement d'une instance

Vous pouvez lancer une instance Linux à l'aide de l'AWS Management Console comme décrit dans la procédure suivante. Ce didacticiel a pour but de vous aider à lancer rapidement votre première instance. Il ne couvrira donc pas toutes les options possibles. Pour de plus amples informations sur les options avancées, veuillez consulter Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Pour lancer une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sur le tableau de bord de la console, sélectionnez Launch Instance.

3. La page Choose an Amazon Machine Image (AMI) affiche une liste de configurations de base nommées Amazon Machine Images (AMI) qui servent de templates pour votre instance. Sélectionnez une version HVM d'Amazon Linux 2. Notez que ces AMI sont indiquées comme « Éligible à l'offre gratuite ».
4. Sur la page Choisir un type d'instance, vous pouvez sélectionner la configuration matérielle de votre instance. Sélectionnez le type d'instance `t2.micro` qui est sélectionné par défaut. Le type d'instance `t2.micro` est éligible pour l'offre gratuite. Dans les régions où `t2.micro` n'est pas disponible, vous pouvez utiliser une instance `t3.micro` avec l'offre gratuite. Pour plus d'informations, consultez la page sur l'offre gratuite AWS.
5. Sélectionnez Vérifier et lancer afin de laisser l'assistant compléter les autres paramètres de configuration pour vous.
6. Sur la page Examiner le lancement de l'instance, sous Groupes de sécurité, vous verrez que l'assistant a créé et sélectionné un groupe de sécurité pour vous. Vous pouvez utiliser ce groupe de sécurité ou sélectionner un groupe de sécurité que vous avez créé lors de la configuration à l'aide des étapes suivantes :

   a. Sélectionnez Edit security groups.
   b. Sur la page Configure Security Group, vérifiez que Select an existing security group est sélectionné.
   c. Choisissez votre groupe de sécurité dans la liste des groupes de sécurité existants, puis sélectionnez Vérifier et lancer.
7. Sur la page Review Instance Launch, sélectionnez Launch.
8. Lorsque vous êtes invité à choisir une paire de clés, sélectionnez Choisir une paire de clés existante, puis sélectionnez la paire de clés que vous avez créée lors de la configuration.

   > **Warning**
   >
   > Ne sélectionnez pas Poursuivre sans paire de clés. Si vous lancez votre instance sans une paire de clés, vous ne pourrez pas vous y connecter.

   Lorsque vous êtes prêt, cochez la case de confirmation, puis sélectionnez Launch Instances.
9. Une page de confirmation indique que l'instance est en cours de lancement. Sélectionnez View Instances pour fermer la page de confirmation et revenir à la console.
10. Sur l'écran Instances, vous pouvez afficher le statut du lancement. Il suffit de peu de temps pour lancer une instance. Lorsque vous lancez une instance, son état initial est `pending`. Une fois que l'instance a démarré, son état devient `running` et elle reçoit un nom DNS public. (Si la colonne DNS public (IPv4) est masquée, sélectionnez l'icône Afficher / Masquer les colonnes (icône en forme d'engrenage) dans le coin supérieur droit de la page, puis sélectionnez DNS public (IPv4).)
11. Cela peut prendre quelques minutes avant que l'instance soit prête pour que vous puissiez vous y connecter. Vérifiez que votre instance a réussi ses contrôles de statut ; vous pouvez voir cette information dans la colonne Status Checks.

Étape 2 : Connexion à l'instance

Vous pouvez vous connecter à une instance Linux de différentes façons. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).

> **Important**
>
> Vous ne pouvez pas vous connecter à votre instance si vous ne l'avez pas lancée avec une paire de clés pour laquelle vous disposez du fichier `.pem` et avec un groupe de sécurité autorisant l'accès SSH depuis votre ordinateur. Si vous ne pouvez pas vous connecter à votre instance, consultez Résolution de la connexion à votre instance (p. 1327) pour obtenir de l'aide.

Étape 3 : Nettoyage de votre instance

Une fois que vous avez fini avec l'instance que vous avez créée pour ce didacticiel, vous devez effectuer un nettoyage en mettant fin à l'instance. Si vous souhaitez exécuter d'autres opérations avec cette instance avant le nettoyage, consultez Étapes suivantes (p. 35).

> **Important**
>
> Mettre fin à une instance la supprime ; vous ne pouvez pas vous reconnecter à une instance une fois que vous y avez mis fin.

Si vous avez lancé une instance qui ne fait pas partie de l'offre gratuite AWS, cette instance ne vous est plus facturée dès que son statut passe à `shutting down` ou `terminated`. Si vous souhaitez conserver une instance pour l'utiliser ultérieurement, mais ne pas payer de frais, vous pouvez arrêter l'instance maintenant et la redémarrer plus tard. Pour plus d'informations, veuillez consulter Arrêt des instances.

Pour mettre fin à une instance

1. Dans le panneau de navigation, choisissez Instances. Sélectionnez l'instance dans la liste des instances.
2. Choisissez État de l'instance, Résilier l'instance.
3. Choisissez Résilier lorsque vous êtes invité à confirmer.

   Amazon EC2 arrête et met fin à votre instance. Après que votre instance a pris fin, elle reste visible sur la console pendant un court instant, puis l'entrée est supprimée automatiquement. Vous ne pouvez pas supprimer vous-même l'instance résiliée de l'affichage de la console.

Étapes suivantes

Après avoir démarré votre instance, vous souhaiterez peut-être essayer quelques-uns des exercices suivants :

- Découvrez comment gérer à distance votre instance EC2 à l'aide de la fonctionnalité Exécuter la commande. Pour de plus amples informations, veuillez consulter AWS Systems Manager Run Command (Exécuter la commande) dans le Guide de l'utilisateur AWS Systems Manager.
- Configurez une alarme CloudWatch pour vous avertir si votre utilisation dépasse le niveau d'offre gratuite. Pour plus d'informations, consultez Création d'une alarme de facturation dans le AWS Billing and Cost Management Guide de l'utilisateur.
- Ajoutez un volume EBS. Pour plus d'informations, consultez Création d'un volume Amazon EBS (p. 1107) et Attacher un volume Amazon EBS à une instance (p. 1109).
- Installez la stack LAMP. Pour plus d'informations, consultez Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2 (p. 38).

Bonnes pratiques relatives à Amazon EC2.

Cette liste de bonnes pratiques vous aidera à tirer le meilleur profit et la plus grande satisfaction d'Amazon EC2.

Sécurité

- Gérez l'accès aux ressources AWS et aux API à l'aide de la fédération d'identité, des utilisateurs IAM et des rôles IAM. Définissez les stratégies et les procédures de gestion des informations d'identification pour créer, distribuer, faire tourner et révoquer les informations d'identification de l'accès à AWS. Pour plus d'informations, consultez Bonnes pratiques IAM dans le IAM Guide de l'utilisateur.
- Implémentez les règles les moins permissives pour votre groupe de sécurité. Pour plus d'informations, consultez Règles des groupes de sécurité (p. 1063).
- Corrigez, mettez à jour et sécurisez régulièrement le système d'exploitation et les applications de votre instance. Pour plus d'informations sur la mise à jour d'Amazon Linux 2 ou de l'Amazon Linux AMI, consultez Gestion de logiciels sur votre instance Linux dans le Amazon EC2 Guide de l'utilisateur pour les instances Linux.

Stockage

- Maîtrisez les implications du type de périphérique racine pour la persistance, la sauvegarde et la récupération des données. Pour plus d'informations, consultez Stockage pour le périphérique racine (p. 106).
- Utilisez des volumes Amazon EBS distincts pour le système d'exploitation et vos données. Assurez-vous que le volume avec vos données persiste après la fin de l'instance. Pour plus d'informations, consultez Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649).
- Utilisez le stockage d'instance disponible pour que votre instance stocke les données temporaires. Souvenez-vous que les données stockées dans un stockage d'instance sont supprimées quand vous arrêtez, mettez en veille prolongée ou résiliez votre instance. Si vous utilisez le stockage d'instance pour le stockage de base de données, assurez-vous d'avoir un cluster avec un facteur de réplication qui garantit la tolérance aux pannes.
- Chiffrez les volumes EBS et les instantanés. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Gestion des ressources

- Utilisez les métadonnées d'instances et les balises de ressource personnalisées pour suivre et identifier vos ressources AWS. Pour plus d'informations, consultez Métadonnées d'instance et données utilisateur (p. 701) et Balisage de vos ressources Amazon EC2 (p. 1309).
- Affichez vos limites actuelles pour Amazon EC2. Prévoyez de demander les augmentations de limite avant le moment où vous en aurez besoin. Pour plus d'informations, consultez Quotas de service Amazon EC2 (p. 1322).

Sauvegarde et restauration

- Sauvegardez régulièrement vos volumes EBS à l'aide des instantanés Amazon EBS (p. 1129) et créez une Amazon Machine Image (AMI) (p. 103) à partir de votre instance afin d'enregistrer la configuration en tant que modèle pour lancer les futures instances.

- Déployez les composants critiques de votre application à travers plusieurs zones de disponibilité et répliquez vos données de manière appropriée.
- Concevez vos applications pour gérer l'adressage IP dynamique au redémarrage de votre instance. Pour plus d'informations, consultez Adressage IP des instances Amazon EC2 (p. 808).
- Surveillez les événements et répondez-y. Pour plus d'informations, consultez Surveillance d'Amazon EC2 (p. 739).
- Vérifiez bien que vous êtes prêt à gérer le failover (basculement). Pour une solution de base, vous pouvez attacher manuellement une interface réseau ou une adresse IP Elastic à une instance de remplacement. Pour plus d'informations, consultez Interfaces réseau Elastic (p. 840). Pour une solution automatisée, vous pouvez utiliser Amazon EC2 Auto Scaling. Pour plus d'informations, consultez le manuel Amazon EC2 Auto Scaling Guide de l'utilisateur.
- Testez régulièrement le processus de récupération de vos instances et volumes Amazon EBS en cas de défaillance.

Mise en réseau

- Définissez la valeur de durée de vie (TTL) pour vos applications sur 255, pour IPv4 et IPv6. Si vous utilisez une valeur inférieure, la durée de vie risque d'expirer pendant le transit du trafic de l'application, ce qui entraînerait des problèmes d'accessibilité pour vos instances.

Didacticiels pour les instances Amazon EC2 exécutant des systèmes d'exploitation Linux

Les didacticiels suivants vous montrent comment effectuer des tâches courantes à l'aide des instances EC2 exécutant des systèmes d'exploitation Linux. Pour voir les vidéos, consultez Vidéos et ateliers didactiques AWS.

Didacticiels

- Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2 (p. 38)
- Didacticiel : Installation d'un serveur web LAMP avec l'Amazon Linux AMI (p. 48)
- Didacticiel : Hébergement d'un blog WordPress avec Amazon Linux (p. 59)
- Didacticiel : Configurer SSL/TLS sur Amazon Linux 2 (p. 69)
- Didacticiel : Configurer SSL/TLS sur Amazon Linux (p. 84)
- Didacticiel : Augmenter la disponibilité des applications sur Amazon EC2 (p. 98)

Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2

Les procédures suivantes vous aident à installer un serveur web Apache avec PHP et le support MariaDB (une fourche développée par la communauté de MySQL) sur votre instance Amazon Linux 2 (parfois appelé serveur web LAMP ou pile LAMP). Vous pouvez utiliser ce serveur pour héberger un site web statique ou déployer une application PHP dynamique qui lit et écrit des informations sur une base de données.

> **Important**
>
> Pour configurer un serveur web LAMP sur Amazon Linux AMI, consultez Didacticiel : Installation d'un serveur web LAMP avec l'Amazon Linux AMI (p. 48).
> Si vous essayez de configurer un serveur web LAMP sur une instance Ubuntu ou Red Hat Enterprise Linux, ce didacticiel ne fonctionnera pas pour vous. Pour obtenir plus d'informations sur d'autres distributions, consultez leur documentation spécifique. Pour plus d'informations sur les serveurs web LAMP sur Ubuntu, consultez la rubrique ApacheMySQLPHP de la documentation proposée par la communauté Ubuntu.

Option : Effectuer ce didacticiel en utilisant Automation

Pour effectuer ce didacticiel en utilisant AWS Systems Manager Automation au lieu des tâches suivantes, exécutez le document Automation AWSDocs-InstallALAMPServer-AL2.

Tâches

- Étape 1 : Préparer le serveur LAMP (p. 39)
- Étape 2 : Tester votre serveur LAMP (p. 42)
- Étape 3 : Sécuriser le serveur de base de données (p. 43)
- Étape 4 : (Facultatif) Installer phpMyAdmin (p. 44)
- Dépannage (p. 47)
- Voir aussi (p. 47)

Étape 1 : Préparer le serveur LAMP

Prérequis

- Ce didacticiel suppose que vous avez déjà lancé une nouvelle instance à l'aide d'Amazon Linux 2 avec un nom DNS public que l'on peut atteindre à partir d'Internet. Pour plus d'informations, consultez Étape 1 : Lancement d'une instance (p. 33). Vous devez aussi avoir configuré votre groupe de sécurité pour permettre les connexions SSH (port 22), HTTP (port 80) et HTTPS (port 443). Pour obtenir plus d'informations sur ces conditions préalables, consultez le didacticiel Autorisation du trafic entrant pour vos instances Linux (p. 1044).
- La procédure suivante installe la dernière version PHP disponible sur Amazon Linux 2, actuellement PHP 7.2. Si vous prévoyez d'utiliser d'autres applications PHP que celles décrites dans ce didacticiel, vous pouvez vérifier qu'elles sont compatibles avec PHP 7.2.

Pour préparer le serveur LAMP

1. Connectez-vous à votre instance (p. 34).
2. Pour vous assurer que tous vos packages logiciels sont mis à jour, effectuez une mise à jour logicielle rapide sur votre instance. Ce processus peut prendre quelques minutes, mais il est important pour vous assurer que vous disposez des dernières mises à jour de sécurité et des nouveaux correctifs de bogues.

   L'option `-y` installe les mises à jour sans demander de confirmation. Si vous souhaitez examiner les mises à jour avant l'installation, vous pouvez omettre cette option.

   ```
   [ec2-user ~]$ sudo yum update -y
   ```

3. Installez les référentiels supplémentaires Amazon Linux `lamp-mariadb10.2-php7.2` et `php7.2` pour obtenir les dernières versions des packages LAMP MariaDB et PHP pour Amazon Linux 2.

   ```
   [ec2-user ~]$ sudo amazon-linux-extras install -y lamp-mariadb10.2-php7.2 php7.2
   ```

   Si vous recevez une erreur `sudo: amazon-linux-extras: command not found`, cela signifie que votre instance n'a pas été lancée avec une AMI Amazon Linux 2 (vous utilisez peut-être l'Amazon Linux AMI). Vous pouvez afficher votre version d'Amazon Linux avec la commande suivante

   ```
   cat /etc/system-release
   ```

   Pour configurer un serveur web LAMP sur Amazon Linux AMI, consultez Didacticiel : Installation d'un serveur web LAMP avec l'Amazon Linux AMI (p. 48).
4. Maintenant que votre instance est à jour, vous pouvez installer le serveur web MariaDB et les packages logiciels PHP.

   Utilisez la commande yum install pour installer plusieurs packages logiciels et toutes les dépendances associées au même moment.

   ```
   [ec2-user ~]$ sudo yum install -y httpd mariadb-server
   ```

   Vous pouvez afficher les versions actuelles de ces packages avec la commande suivante :

   ```
   yum info package_name
   ```

5. Démarrez le serveur web Apache.

```
[ec2-user ~]$ sudo systemctl start httpd
```

6. Utilisez la commande systemctl pour configurer le serveur Web Apache afin qu'il soit lancé à chaque démarrage système.

```
[ec2-user ~]$ sudo systemctl enable httpd
```

Vous pouvez vérifier que httpd est activé en exécutant la commande suivante :

```
[ec2-user ~]$ sudo systemctl is-enabled httpd
```

7. Ajoutez une règle de sécurité pour autoriser les connexions HTTP entrantes (port 80) à votre instance si vous ne l'avez pas déjà fait. Par défaut, un groupe de sécurité launch-wizard-*N* a été configuré pour votre instance lors de l'initialisation. Ce groupe contient une règle unique pour autoriser les connexions SSH.

   a. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

   b. Choisissez Instances et sélectionnez votre instance.

   c. Sous l'onglet Sécurité, affichez les règles entrantes. Vous devriez voir la règle suivante :

```
Port range   Protocol     Source
22           tcp          0.0.0.0/0
```

   d. Choisissez le lien pour le groupe de sécurité. En utilisant les procédures de Ajout de règles à un groupe de sécurité (p. 1069), ajoutez une nouvelle règle de sécurité entrante avec les valeurs suivantes :

   - Type : HTTP
   - Protocole : TCP
   - Plage de ports : 80
   - Source : Personnalisé

8. Testez votre serveur web. Dans un navigateur web, saisissez l'adresse DNS publique (ou l'adresse IP publique) de votre instance. S'il n'existe aucun contenu dans `/var/www/html`, vous devriez voir la page test Apache. Vous pouvez obtenir le DNS public de votre instance en utilisant la console Amazon EC2 (vérifiez la colonne DNS public ; si cette colonne est masquée, choisissez l'icône Afficher/Masquer les colonnes (icône en forme d'engrenage) et sélectionnez DNS public).

   Si vous ne pouvez pas voir la page test Apache, vérifiez que le groupe de sécurité que vous utilisez contient une règle pour autoriser le trafic HTTP (port 80). Pour plus d'informations sur l'ajout d'une règle HTTP à votre groupe de sécurité, consultez Ajout de règles à un groupe de sécurité (p. 1069).

   Important

   Si vous n'utilisez pas Amazon Linux, il se peut que vous deviez aussi configurer le pare-feu sur votre instance pour autoriser ces connexions. Pour obtenir plus d'informations sur la configuration du pare-feu, consultez la documentation de votre distribution spécifique.

La commande httpd traite les fichiers qui sont conservés dans un répertoire appelé racine du document Apache. La racine du document Apache d'Amazon Linux est `/var/www/html` qui est détenu par défaut par la racine.

Pour autoriser le compte `ec2-user` à manipuler les fichiers de ce répertoire, vous devez modifier la propriété et les autorisations du répertoire. Il existe plusieurs façons d'accomplir cette tâche. Dans ce didacticiel, vous ajoutez l'utilisateur `ec2-user` au groupe `apache` pour donner au groupe `apache` la propriété du répertoire `/var/www` et attribuer les autorisations d'écriture au groupe.

Pour définir les autorisations sur les fichiers

1. Ajoutez votre utilisateur (dans ce cas, `ec2-user`) au groupe `apache`.

```
[ec2-user ~]$ sudo usermod -a -G apache ec2-user
```

2. Déconnectez-vous, puis reconnectez-vous pour sélectionner le nouveau groupe, puis vérifiez votre adhésion.

   a. Déconnectez-vous (utilisez la commande exit ou fermez la fenêtre de terminal) :

```
[ec2-user ~]$ exit
```

   b. Pour vérifier votre adhésion au groupe `apache`, reconnectez-vous à votre instance, puis exécutez la commande suivante :

```
[ec2-user ~]$ groups
ec2-user adm wheel apache systemd-journal
```

3. Remplacez la propriété de groupe de `/var/www` et son contenu par le groupe `apache`.

```
[ec2-user ~]$ sudo chown -R ec2-user:apache /var/www
```

4. Pour ajouter des autorisations d'écriture de groupe et définir l'ID de groupe pour les futurs sous-répertoires, modifiez les autorisations sur les répertoires de `/var/www` et ses sous-répertoires.

```
[ec2-user ~]$ sudo chmod 2775 /var/www && find /var/www -type d -exec sudo chmod 2775
 {} \;
```

5. Pour ajouter des autorisations d'écriture de groupe, modifiez de façon récursive les autorisations sur les fichiers de `/var/www` et ses sous-répertoires :

```
[ec2-user ~]$ find /var/www -type f -exec sudo chmod 0664 {} \;
```

Maintenant, `ec2-user` (et tous les futurs membres du groupe `apache`) peut ajouter, supprimer et modifier les fichiers à la racine du document Apache. Vous pouvez ainsi ajouter du contenu, tel qu'un site Web statique ou une application PHP.

Pour sécuriser votre serveur web (facultatif)

Un serveur web exécutant le protocole HTTP ne fournit aucune sécurité de transport pour les données qu'il envoie ou reçoit. Lorsque vous vous connectez à un serveur HTTP via un navigateur Web, les URL que vous visitez, le contenu des pages web que vous recevez et le contenu (y compris les mots de passe) de tous les formulaires HTML que vous envoyez peuvent être vus par des personnes malveillantes sur le chemin d'accès réseau. Les bonnes pratiques en matière de sécurisation de votre serveur web consistent à installer la prise en charge HTTPS (HTTP Secure), qui protège vos données grâce au chiffrement SSL/TLS.

Pour plus d'informations sur l'activation de HTTPS sur votre serveur, consultez Didacticiel : Configurer SSL/TLS sur Amazon Linux 2 (p. 69).

Étape 2 : Tester votre serveur LAMP

Si votre serveur est installé et en cours d'exécution, et que vos autorisations sur les fichiers sont correctement définies, votre compte `ec2-user` doit pouvoir créer un fichier PHP simple dans le répertoire `/var/www/html` qui est disponible à partir d'Internet.

Pour tester votre serveur LAMP

1. Créez un fichier PHP à la racine du document Apache.

```
[ec2-user ~]$ echo "<?php phpinfo(); ?>" > /var/www/html/phpinfo.php
```

   Si l'erreur « Permission denied » s'affiche lorsque vous essayez d'exécuter cette commande, essayez de vous déconnecter et de vous reconnecter pour récupérer les autorisations d'un groupe que vous avez configurées dans Pour définir les autorisations sur les fichiers (p. 41).

2. Dans un navigateur web, saisissez l'URL du fichier que vous venez de créer. Cette URL est l'adresse DNS publique de votre instance suivie par une barre oblique et le nom du fichier. Par exemple :

```
http://my.public.dns.amazonaws.com/phpinfo.php
```

   Vous devriez voir la page d'informations PHP:

PHP Version 7.2.0

System	Linux ip-172-31-22-15.us-west-2.compute.internal 4.9.62-10.57.amzn2.x86_ 2017 x86_64
Build Date	Dec 13 2017 03:34:37
Server API	Apache 2.0 Handler
Virtual Directory Support	disabled
Configuration File (php.ini) Path	/etc
Loaded Configuration File	/etc/php.ini
Scan this dir for additional .ini files	/etc/php.d
Additional .ini files parsed	/etc/php.d/20-bz2.ini, /etc/php.d/20-calendar.ini, /etc/php.d/20-ctype.ini, /etc/ /etc/php.d/20-fileinfo.ini, /etc/php.d/20-ftp.ini, /etc/php.d/20-gettext.ini, /etc/ph /etc/php.d/20-mysqlnd.ini, /etc/php.d/20-pdo.ini, /etc/php.d/20-phar.ini, /etc/p sqlite3.ini, /etc/php.d/20-tokenizer.ini, /etc/php.d/30-mysqli.ini, /etc/php.d/30- pdo_sqlite.ini
PHP API	20170718
PHP Extension	20170718
Zend Extension	320170718
Zend Extension Build	API320170718,NTS
PHP Extension Build	API20170718,NTS

Si vous ne voyez pas cette page, vérifiez que le fichier `/var/www/html/phpinfo.php` a été créé correctement à l'étape précédente. Vous pouvez également vérifier que les packages requis ont été installés avec la commande suivante.

```
[ec2-user ~]$ sudo yum list installed httpd mariadb-server php-mysqlnd
```

Si l'un des packages requis n'est pas présent dans votre sortie, installez-les avec la commande sudo yum install *package*. Vérifiez également que les référentiels supplémentaires `php7.2` et `lamp-mariadb10.2-php7.2` sont activés dans la sortie de la commande amazon-linux-extras.

3. Supprimez le fichier `phpinfo.php`. Même si ces informations peuvent vous être utiles, elles ne doivent pas être diffusées sur Internet pour des raisons de sécurité.

```
[ec2-user ~]$ rm /var/www/html/phpinfo.php
```

Vous devriez maintenant avoir un serveur web LAMP entièrement fonctionnel. Si vous ajoutez un contenu à la racine du document Apache à l'emplacement `/var/www/html`, vous devez pouvoir voir ce contenu à l'adresse du DNS public de votre instance.

Étape 3 : Sécuriser le serveur de base de données

L'installation par défaut du serveur MariaDB possède plusieurs fonctions qui sont parfaites pour les tests et le développement, mais elles devraient être désactivées ou supprimées des serveurs de production. La commande mysql_secure_installation vous guide à travers le processus de paramétrage d'un mot de passe racine et de suppression des fonctions non sécurisées de votre installation. Même si vous ne comptez pas utiliser le serveur MariaDB, nous vous recommandons de suivre cette procédure.

Pour sécuriser le serveur MariaDB

1. Démarrez le serveur MariaDB.

```
[ec2-user ~]$ sudo systemctl start mariadb
```

2. Exécutez mysql_secure_installation.

```
[ec2-user ~]$ sudo mysql_secure_installation
```

   a. A l'invite, saisissez un mot de passe pour le compte racine.

      i. Saisissez le mot de passe racine actuel. Par défaut, le compte racine n'a pas de mot de passe défini. Appuyez sur Entrée.

      ii. Tapez **Y** pour définir un mot de passe et saisissez deux fois un mot de passe sécurisé. Pour plus d'informations sur la création d'un mot de passe fiable, consultez https://identitysafe.norton.com/password-generator/. Assurez-vous de stocker ce mot de passe en lieu sûr.

      La mesure la plus simple pour sécuriser votre base de données consiste à définir un mot de passe racine pour MariaDB. Lorsque vous concevez ou installez une application reposant sur une base de données, vous devez généralement créer un utilisateur de services de base de données pour cette application et éviter d'utiliser le compte racine, sauf pour administrer la base de données.

   b. Tapez **Y** pour supprimer les comptes d'utilisateur anonymes.

   c. Tapez **Y** pour désactiver la connexion racine à distance.

   d. Tapez **Y** pour supprimer la base de données de test.

   e. Tapez **Y** pour recharger les tableaux de privilèges et enregistrer vos changements.

3. (Facultatif) Si vous ne comptez pas utiliser le serveur MariaDB tout de suite, arrêtez-le. Vous pouvez le redémarrer lorsque vous en avez de nouveau besoin.

```
[ec2-user ~]$ sudo systemctl stop mariadb
```

4. (Facultatif) Si vous voulez que le serveur MariaDB soit lancé à chaque démarrage, saisissez la commande suivante.

```
[ec2-user ~]$ sudo systemctl enable mariadb
```

Étape 4 : (Facultatif) Installer phpMyAdmin

phpMyAdmin est un outil de gestion de bases de données basé sur le Web que vous pouvez utiliser pour visualiser et modifier les bases de données MySQL sur votre instance EC2. Suivez les étapes ci-dessous pour installer et configurer `phpMyAdmin` sur votre instance Amazon Linux.

Important

Nous ne vous recommandons pas d'utiliser `phpMyAdmin` pour accéder à un serveur LAMP, sauf si vous avez activé SSL/TLS dans Apache. Sinon, votre mot de passe administrateur de base de données et d'autres données sont transmises de façon non sécurisée sur Internet. Pour accéder à des recommandations de sécurité des développeurs, consultez Securing your phpMyAdmin installation. Pour obtenir des informations générales sur la sécurisation d'un serveur Web sur une instance EC2, consultez Didacticiel : Configurer SSL/TLS sur Amazon Linux 2 (p. 69).

Pour installer phpMyAdmin

1. Installez les dépendances obligatoires.

```
[ec2-user ~]$ sudo yum install php-mbstring -y
```

2. Redémarrez Apache.

```
[ec2-user ~]$ sudo systemctl restart httpd
```

3. Redémarrez `php-fpm`.

```
[ec2-user ~]$ sudo systemctl restart php-fpm
```

4. Accédez à la racine du document Apache sur `/var/www/html`.

```
[ec2-user ~]$ cd /var/www/html
```

5. Sélectionnez un package source pour la dernière version de phpMyAdmin sur https://www.phpmyadmin.net/downloads. Pour télécharger le fichier directement sur votre instance, copiez le lien et collez-le dans une commande wget, comme dans cet exemple :

```
[ec2-user html]$ wget https://www.phpmyadmin.net/downloads/phpMyAdmin-latest-all-languages.tar.gz
```

6. Créez un dossier `phpMyAdmin` et extrayez le package dans celui-ci avec la commande suivante.

```
[ec2-user html]$ mkdir phpMyAdmin && tar -xvzf phpMyAdmin-latest-all-languages.tar.gz -C phpMyAdmin --strip-components 1
```

7. Supprimez le tarball *`phpMyAdmin-latest-all-languages.tar.gz`*.

```
[ec2-user html]$ rm phpMyAdmin-latest-all-languages.tar.gz
```

8. (Facultatif) Si le serveur MySQL n'est pas en cours d'exécution, démarrez-le maintenant.

```
[ec2-user ~]$ sudo systemctl start mariadb
```

9. Dans un navigateur web, saisissez l'URL de votre installation phpMyAdmin. Cette URL est l'adresse DNS publique (ou l'adresse IP publique) de votre instance suivie par une barre oblique et le nom du fichier de votre répertoire d'installation. Par exemple :

```
http://my.public.dns.amazonaws.com/phpMyAdmin
```

La page de connexion phpMyAdmin devrait s'afficher :

10. Connectez-vous à votre installation phpMyAdmin avec le nom d'utilisateur `root` et le mot de passe racine MySQL que vous avez créés précédemment.

    Votre installation doit être configurée avant que vous la mettiez en service. Pour configurer phpMyAdmin, vous pouvez créer manuellement un fichier de configuration, utiliser la console de configuration ou combiner ces deux approches.

    Pour plus d'informations sur l'utilisation de phpMyAdmin, consultez le Guide de l'utilisateur phpMyAdmin.

Dépannage

Cette section propose des suggestions pour résoudre les problèmes courants que vous pouvez rencontrer lors de la configuration d'un nouveau serveur LAMP.

Je ne peux pas me connecter à mon serveur à l'aide d'un navigateur Internet.

Effectuez les vérifications suivantes pour voir si votre serveur web Apache est en cours d'exécution et accessible.

- Le serveur web est-il en cours d'exécution ?

  Vous pouvez vérifier que httpd est activé en exécutant la commande suivante :

  ```
  [ec2-user ~]$ sudo systemctl is-enabled httpd
  ```

  Si le processus httpd n'est pas en cours d'exécution, répétez les étapes décrites dans Pour préparer le serveur LAMP (p. 39).
- Le pare-feu est-il configuré correctement ?

  Si vous ne pouvez pas voir la page test Apache, vérifiez que le groupe de sécurité que vous utilisez contient une règle pour autoriser le trafic HTTP (port 80). Pour plus d'informations sur l'ajout d'une règle HTTP à votre groupe de sécurité, consultez Ajout de règles à un groupe de sécurité (p. 1069).

Voir aussi

Pour plus d'informations sur le transfert des fichiers vers votre instance ou l'installation d'un blog WordPress sur votre serveur web, consultez la documentation suivante :

- Transfert de fichiers vers votre instance Linux à l'aide de WinSCP (p. 618)
- Transfert de fichiers des instances Linux à partir de Linux à l'aide de SCP (p. 603)
- Didacticiel : Hébergement d'un blog WordPress avec Amazon Linux (p. 59)

Pour plus d'informations sur les commandes et le logiciel utilisés dans ce didacticiel, consultez les pages web suivantes :

- Serveur Web Apache : http://httpd.apache.org/
- Serveur de base de données MariaDB : https://mariadb.org/
- Langage de programmation PHP : http://php.net/
- La commande `chmod` : https://en.wikipedia.org/wiki/Chmod
- La commande `chown` : https://en.wikipedia.org/wiki/Chown

Pour plus d'informations sur l'enregistrement d'un nom de domaine pour votre serveur web ou le transfert d'un nom de domaine existant vers cet hôte, consultez Création et migration de domaines et de sous-domaines vers Amazon Route 53 dans le Amazon Route 53 Manuel du développeur.

Didacticiel : Installation d'un serveur web LAMP avec l'Amazon Linux AMI

Les procédures suivantes vous aident à installer un serveur Web Apache avec PHP et le support MySQL sur votre instance Amazon Linux (parfois appelé serveur Web LAMP ou pile LAMP). Vous pouvez utiliser ce serveur pour héberger un site web statique ou déployer une application PHP dynamique qui lit et écrit des informations sur une base de données.

> **Important**
>
> Pour configurer un serveur web LAMP sur Amazon Linux 2, consultez Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2 (p. 38).
> Si vous essayez de configurer un serveur web LAMP sur une instance Ubuntu ou Red Hat Enterprise Linux, ce didacticiel ne fonctionnera pas pour vous. Pour obtenir plus d'informations sur d'autres distributions, consultez leur documentation spécifique. Pour plus d'informations sur les serveurs web LAMP sur Ubuntu, consultez la rubrique ApacheMySQLPHP de la documentation proposée par la communauté Ubuntu.

Option : Effectuer ce didacticiel en utilisant Automation

Pour effectuer ce didacticiel en utilisant AWS Systems Manager Automation au lieu des tâches suivantes, exécutez le document Automation AWSDocs-InstallALAMPServer-AL.

Tâches

- Étape 1 : Préparer le serveur LAMP (p. 48)
- Étape 2 : Tester votre serveur LAMP (p. 52)
- Étape 3 : Sécuriser le serveur de base de données (p. 53)
- Étape 4 : (Facultatif) Installer phpMyAdmin (p. 55)
- Dépannage (p. 58)
- Voir aussi (p. 59)

Étape 1 : Préparer le serveur LAMP

Prérequis

Ce didacticiel suppose que vous avez déjà lancé une nouvelle instance à l'aide de l'Amazon Linux AMI avec un nom DNS public que l'on peut atteindre à partir d'Internet. Pour plus d'informations, consultez Étape 1 : Lancement d'une instance (p. 33). Vous devez aussi avoir configuré votre groupe de sécurité pour permettre les connexions SSH (port 22), HTTP (port 80) et HTTPS (port 443). Pour obtenir plus d'informations sur ces conditions préalables, consultez le didacticiel Autorisation du trafic entrant pour vos instances Linux (p. 1044).

Pour installer et démarrer le serveur web LAMP avec l'Amazon Linux AMI

1. Connectez-vous à votre instance (p. 34).
2. Pour vous assurer que tous vos packages logiciels sont mis à jour, effectuez une mise à jour logicielle rapide sur votre instance. Ce processus peut prendre quelques minutes, mais il est important pour vous assurer que vous disposez des dernières mises à jour de sécurité et des nouveaux correctifs de bogues.

   L'option `-y` installe les mises à jour sans demander de confirmation. Si vous souhaitez examiner les mises à jour avant l'installation, vous pouvez omettre cette option.

   ```
   [ec2-user ~]$ sudo yum update -y
   ```

3. Maintenant que votre instance est à jour, vous pouvez installer le serveur web Apache, MySQL et les packages logiciels PHP.

   Important

   Il est possible que certaines applications ne soient pas compatibles avec l'environnement logiciel recommandé suivant. Avant d'installer ces packages, vérifiez si vos applications LAMP sont compatibles avec eux. En cas de problème, vous devrez peut-être installer un autre environnement. Pour de plus amples informations, veuillez consulter Le logiciel d'application que je veux exécuter sur mon serveur est incompatible avec la version PHP installée ou d'autres logiciels (p. 58)

   Utilisez la commande yum install pour installer plusieurs packages logiciels et toutes les dépendances associées au même moment.

```
[ec2-user ~]$ sudo yum install -y httpd24 php72 mysql57-server php72-mysqlnd
```

   Si vous recevez l'erreur `No package package-name available`, cela signifie que votre instance n'a pas été lancée avec l'Amazon Linux AMI (vous utilisez peut-être Amazon Linux 2). Vous pouvez afficher votre version d'Amazon Linux avec la commande suivante

```
cat /etc/system-release
```

4. Démarrez le serveur web Apache.

```
[ec2-user ~]$ sudo service httpd start
Starting httpd:                                            [  OK  ]
```

5. Utilisez la commande chkconfig pour configurer le serveur Web Apache afin qu'il soit lancé à chaque démarrage système.

```
[ec2-user ~]$ sudo chkconfig httpd on
```

   La commande chkconfig n'indique aucun message de confirmation lorsque vous parvenez à activer un service avec elle.

   Vous pouvez vérifier que httpd est activé en exécutant la commande suivante :

```
[ec2-user ~]$ chkconfig --list httpd
httpd           0:off   1:off   2:on    3:on    4:on    5:on    6:off
```

   Ici, httpd a la valeur `on` sur les niveaux d'exécution 2, 3, 4 et 5 (ce qui correspond à ce que vous voulez voir).

6. Ajoutez une règle de sécurité pour autoriser les connexions HTTP entrantes (port 80) à votre instance si vous ne l'avez pas déjà fait. Par défaut, un groupe de sécurité launch-wizard-*N* a été configuré pour votre instance lors de l'initialisation. Ce groupe contient une règle unique pour autoriser les connexions SSH.

   a. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
   b. Choisissez Instances et sélectionnez votre instance.
   c. Sous l'onglet Sécurité, affichez les règles entrantes. Vous devriez voir la règle suivante :

```
Port range   Protocol     Source
22           tcp          0.0.0.0/0
```

d. Choisissez le lien pour le groupe de sécurité. En utilisant les procédures de Ajout de règles à un groupe de sécurité (p. 1069), ajoutez une nouvelle règle de sécurité entrante avec les valeurs suivantes :

- Type : HTTP
- Protocole : TCP
- Plage de ports : 80
- Source : Personnalisé

7. Testez votre serveur web. Dans un navigateur web, saisissez l'adresse DNS publique (ou l'adresse IP publique) de votre instance. Vous pouvez obtenir l'adresse DNS publique de votre instance à l'aide de la console Amazon EC2. S'il n'existe aucun contenu dans `/var/www/html`, vous devriez voir la page test Apache. Lorsque vous ajoutez du contenu à la racine du document, votre contenu apparaît à l'adresse DNS publique de votre instance au lieu de cette page test.

   Si vous ne pouvez pas voir la page test Apache, vérifiez que le groupe de sécurité que vous utilisez contient une règle pour autoriser le trafic HTTP (port 80). Pour plus d'informations sur l'ajout d'une règle HTTP à votre groupe de sécurité, consultez Ajout de règles à un groupe de sécurité (p. 1069).

   Si vous n'utilisez pas Amazon Linux, il se peut que vous deviez aussi configurer le pare-feu sur votre instance pour autoriser ces connexions. Pour obtenir plus d'informations sur la configuration du pare-feu, consultez la documentation de votre distribution spécifique.

La commande httpd traite les fichiers qui sont conservés dans un répertoire appelé racine du document Apache. La racine du document Apache d'Amazon Linux est `/var/www/html` qui est détenu par défaut par la racine.

```
[ec2-user ~]$ ls -l /var/www
total 16
drwxr-xr-x 2 root root 4096 Jul 12 01:00 cgi-bin
drwxr-xr-x 3 root root 4096 Aug  7 00:02 error
drwxr-xr-x 2 root root 4096 Jan  6  2012 html
drwxr-xr-x 3 root root 4096 Aug  7 00:02 icons
drwxr-xr-x 2 root root 4096 Aug  7 21:17 noindex
```

Pour autoriser le compte `ec2-user` à manipuler les fichiers de ce répertoire, vous devez modifier la propriété et les autorisations du répertoire. Il existe plusieurs façons d'accomplir cette tâche. Dans ce didacticiel, vous ajoutez l'utilisateur `ec2-user` au groupe `apache` pour donner au groupe `apache` la propriété du répertoire `/var/www` et attribuer les autorisations d'écriture au groupe.

Pour définir les autorisations sur les fichiers

1. Ajoutez votre utilisateur (dans ce cas, `ec2-user`) au groupe `apache`.

   ```
   [ec2-user ~]$ sudo usermod -a -G apache ec2-user
   ```

2. Déconnectez-vous, puis reconnectez-vous pour sélectionner le nouveau groupe, puis vérifiez votre adhésion.

   a. Déconnectez-vous (utilisez la commande exit ou fermez la fenêtre de terminal) :

      ```
      [ec2-user ~]$ exit
      ```

   b. Pour vérifier votre adhésion au groupe `apache`, reconnectez-vous à votre instance, puis exécutez la commande suivante :

      ```
      [ec2-user ~]$ groups
      ec2-user wheel apache
      ```

3. Remplacez la propriété de groupe de `/var/www` et son contenu par le groupe `apache`.

   ```
   [ec2-user ~]$ sudo chown -R ec2-user:apache /var/www
   ```

4. Pour ajouter des autorisations d'écriture de groupe et définir l'ID de groupe pour les futurs sous-répertoires, modifiez les autorisations sur les répertoires de `/var/www` et ses sous-répertoires.

   ```
   [ec2-user ~]$ sudo chmod 2775 /var/www
   [ec2-user ~]$ find /var/www -type d -exec sudo chmod 2775 {} \;
   ```

5. Pour ajouter des autorisations d'écriture de groupe, modifiez de façon récursive les autorisations sur les fichiers de `/var/www` et ses sous-répertoires :

   ```
   [ec2-user ~]$ find /var/www -type f -exec sudo chmod 0664 {} \;
   ```

Maintenant, `ec2-user` (et tous les futurs membres du groupe `apache`) peut ajouter, supprimer et modifier les fichiers à la racine du document Apache. Vous pouvez ainsi ajouter du contenu, tel qu'un site Web statique ou une application PHP.

(Facultatif) Sécuriser votre serveur web

Un serveur web exécutant le protocole HTTP ne fournit aucune sécurité de transport pour les données qu'il envoie ou reçoit. Lorsque vous vous connectez à un serveur HTTP via un navigateur Web, les URL que vous visitez, le contenu des pages web que vous recevez et le contenu (y compris les mots de passe) de tous les formulaires HTML que vous envoyez peuvent être vus par des personnes malveillantes sur le chemin d'accès réseau. Les bonnes pratiques en matière de sécurisation de votre serveur web consistent à installer la prise en charge HTTPS (HTTP Secure), qui protège vos données grâce au chiffrement SSL/TLS.

Pour plus d'informations sur l'activation de HTTPS sur votre serveur, consultez Didacticiel : Configurer SSL/TLS sur Amazon Linux (p. 84).

Étape 2 : Tester votre serveur LAMP

Si votre serveur est installé et en cours d'exécution, et que vos autorisations sur les fichiers sont correctement définies, votre compte `ec2-user` doit pouvoir créer un fichier PHP simple dans le répertoire `/var/www/html` qui est disponible à partir d'Internet.

Pour tester votre serveur web LAMP

1. Créez un fichier PHP à la racine du document Apache.

   ```
   [ec2-user ~]$ echo "<?php phpinfo(); ?>" > /var/www/html/phpinfo.php
   ```

   Si l'erreur « Permission denied » s'affiche lorsque vous essayez d'exécuter cette commande, essayez de vous déconnecter et de vous reconnecter pour récupérer les autorisations d'un groupe que vous avez configurées dans Étape 1 : Préparer le serveur LAMP (p. 48).

2. Dans un navigateur web, saisissez l'URL du fichier que vous venez de créer. Cette URL est l'adresse DNS publique de votre instance suivie par une barre oblique et le nom du fichier. Par exemple :

   ```
   http://my.public.dns.amazonaws.com/phpinfo.php
   ```

   Vous devriez voir la page d'informations PHP:

PHP Version 7.2.0

System	Linux ip-172-31-22-15.us-west-2.compute.internal 4.9.62-10.57.amzn2.x86_ 2017 x86_64
Build Date	Dec 13 2017 03:34:37
Server API	Apache 2.0 Handler
Virtual Directory Support	disabled
Configuration File (php.ini) Path	/etc
Loaded Configuration File	/etc/php.ini
Scan this dir for additional .ini files	/etc/php.d
Additional .ini files parsed	/etc/php.d/20-bz2.ini, /etc/php.d/20-calendar.ini, /etc/php.d/20-ctype.ini, /etc/ /etc/php.d/20-fileinfo.ini, /etc/php.d/20-ftp.ini, /etc/php.d/20-gettext.ini, /etc/ph /etc/php.d/20-mysqlnd.ini, /etc/php.d/20-pdo.ini, /etc/php.d/20-phar.ini, /etc/p sqlite3.ini, /etc/php.d/20-tokenizer.ini, /etc/php.d/30-mysqli.ini, /etc/php.d/30- pdo_sqlite.ini
PHP API	20170718
PHP Extension	20170718
Zend Extension	320170718
Zend Extension Build	API320170718,NTS
PHP Extension Build	API20170718,NTS

Si vous ne voyez pas cette page, vérifiez que le fichier `/var/www/html/phpinfo.php` a été créé correctement à l'étape précédente. Vous pouvez également vérifier que les packages requis ont été installés avec la commande suivante. Les versions de package dans la deuxième colonne n'ont pas besoin de correspondre à cet exemple de sortie.

```
[ec2-user ~]$ sudo yum list installed httpd24 php72 mysql57-server php72-mysqlnd
Loaded plugins: priorities, update-motd, upgrade-helper
Installed Packages
httpd24.x86_64                          2.4.25-1.68.amzn1                   @amzn-
updates
mysql56-server.x86_64                   5.6.35-1.23.amzn1                   @amzn-
updates
php70.x86_64                            7.0.14-1.20.amzn1                   @amzn-
updates
php70-mysqlnd.x86_64                    7.0.14-1.20.amzn1                   @amzn-
updates
```

Si l'un des packages requis n'est pas présent dans votre sortie, installez-le avec la commande sudo yum install ***package***.

3. Supprimez le fichier `phpinfo.php`. Même si ces informations peuvent vous être utiles, elles ne doivent pas être diffusées sur Internet pour des raisons de sécurité.

```
[ec2-user ~]$ rm /var/www/html/phpinfo.php
```

Étape 3 : Sécuriser le serveur de base de données

L'installation par défaut du serveur MySQL possède plusieurs fonctions qui sont excellentes pour les tests et les développements, mais elles devraient être désactivées ou supprimées des serveurs de production.

La commande mysql_secure_installation vous guide à travers le processus de paramétrage d'un mot de passe racine et de suppression des fonctions non sécurisées de votre installation. Même si vous ne comptez pas utiliser le serveur MySQL, nous vous recommandons de suivre cette procédure.

Pour sécuriser le serveur de base de données

1. Démarrez le serveur MySQL.

```
[ec2-user ~]$ sudo service mysqld start
Initializing MySQL database:
...

PLEASE REMEMBER TO SET A PASSWORD FOR THE MySQL root USER !
...

Starting mysqld:                                           [  OK  ]
```

2. Exécutez mysql_secure_installation.

```
[ec2-user ~]$ sudo mysql_secure_installation
```

   a. A l'invite, saisissez un mot de passe pour le compte racine.

      i. Saisissez le mot de passe racine actuel. Par défaut, le compte racine n'a pas de mot de passe défini. Appuyez sur Entrée.

      ii. Tapez **Y** pour définir un mot de passe et saisissez deux fois un mot de passe sécurisé. Pour plus d'informations sur la création d'un mot de passe fiable, consultez https://identitysafe.norton.com/password-generator/. Assurez-vous de stocker ce mot de passe en lieu sûr.

      La mesure la plus simple pour sécuriser votre base de données consiste à définir un mot de passe racine pour MySQL. Lorsque vous concevez ou installez une application reposant sur une base de données, vous devez généralement créer un utilisateur de services de base de données pour cette application et éviter d'utiliser le compte racine, sauf pour administrer la base de données.

   b. Tapez **Y** pour supprimer les comptes d'utilisateur anonymes.

   c. Tapez **Y** pour désactiver la connexion racine à distance.

   d. Tapez **Y** pour supprimer la base de données de test.

   e. Tapez **Y** pour recharger les tableaux de privilèges et enregistrer vos changements.

3. (Facultatif) Si vous ne comptez pas utiliser le serveur MySQL tout de suite, arrêtez-le. Vous pouvez le redémarrer lorsque vous en avez de nouveau besoin.

```
[ec2-user ~]$ sudo service mysqld stop
Stopping mysqld:                                           [  OK  ]
```

4. (Facultatif) Si vous voulez que le serveur MySQL soit lancé à chaque démarrage, saisissez la commande suivante.

```
[ec2-user ~]$ sudo chkconfig mysqld on
```

Vous devriez maintenant avoir un serveur web LAMP entièrement fonctionnel. Si vous ajoutez un contenu à la racine du document Apache à l'emplacement `/var/www/html`, vous devez pouvoir voir ce contenu à l'adresse du DNS public de votre instance.

Étape 4 : (Facultatif) Installer phpMyAdmin

Pour installer phpMyAdmin

phpMyAdmin est un outil de gestion de bases de données basé sur le Web que vous pouvez utiliser pour visualiser et modifier les bases de données MySQL sur votre instance EC2. Suivez les étapes ci-dessous pour installer et configurer phpMyAdmin sur votre instance Amazon Linux.

> **Important**
>
> Nous ne vous recommandons pas d'utiliser phpMyAdmin pour accéder à un serveur LAMP, sauf si vous avez activé SSL/TLS dans Apache. Sinon, votre mot de passe administrateur de base de données et d'autres données sont transmises de façon non sécurisée sur Internet. Pour accéder à des recommandations de sécurité des développeurs, consultez Securing your phpMyAdmin installation.

> **Note**
>
> Le système de gestion des packages Amazon Linux ne prend pas en charge actuellement l'installation automatique de phpMyAdmin dans un environnement PHP 7. Ce didacticiel explique comment installer phpMyAdmin manuellement.

1. Connectez-vous à votre instance EC2 à l'aide de SSH.
2. Installez les dépendances obligatoires.

```
[ec2-user ~]$ sudo yum install php72-mbstring.x86_64 -y
```

3. Redémarrez Apache.

```
[ec2-user ~]$ sudo service httpd restart
Stopping httpd:                                            [  OK  ]
Starting httpd:                                            [  OK  ]
```

4. Accédez à la racine du document Apache sur `/var/www/html`.

```
[ec2-user ~]$ cd /var/www/html
[ec2-user html]$
```

5. Sélectionnez un package source pour la dernière version de phpMyAdmin sur https://www.phpmyadmin.net/downloads. Pour télécharger le fichier directement sur votre instance, copiez le lien et collez-le dans une commande wget, comme dans cet exemple :

```
[ec2-user html]$ wget https://www.phpmyadmin.net/downloads/phpMyAdmin-latest-all-
languages.tar.gz
```

6. Créez un dossier phpMyAdmin et extrayez le package dans celui-ci à l'aide de la commande suivante.

```
[ec2-user html]$ mkdir phpMyAdmin && tar -xvzf phpMyAdmin-latest-all-languages.tar.gz -
C phpMyAdmin --strip-components 1
```

7. Supprimez le tarball *`phpMyAdmin-latest-all-languages.tar.gz`*.

```
[ec2-user html]$ rm phpMyAdmin-latest-all-languages.tar.gz
```

8. (Facultatif) Si le serveur MySQL n'est pas en cours d'exécution, démarrez-le maintenant.

```
[ec2-user ~]$ sudo service mysqld start
```

```
Starting mysqld:                                           [  OK  ]
```

9. Dans un navigateur web, saisissez l'URL de votre installation phpMyAdmin. Cette URL est l'adresse DNS publique (ou l'adresse IP publique) de votre instance suivie par une barre oblique et le nom du fichier de votre répertoire d'installation. Par exemple :

```
http://my.public.dns.amazonaws.com/phpMyAdmin
```

La page de connexion phpMyAdmin devrait s'afficher :

10. Connectez-vous à votre installation phpMyAdmin avec le nom d'utilisateur `root` et le mot de passe racine MySQL que vous avez créés précédemment.

    Votre installation doit être configurée avant que vous la mettiez en service. Pour configurer phpMyAdmin, vous pouvez créer manuellement un fichier de configuration, utiliser la console de configuration ou combiner ces deux approches.

    Pour plus d'informations sur l'utilisation de phpMyAdmin, consultez le Guide de l'utilisateur phpMyAdmin.

Dépannage

Cette section propose des suggestions pour résoudre les problèmes courants que vous pouvez rencontrer lors de la configuration d'un nouveau serveur LAMP.

Je ne peux pas me connecter à mon serveur à l'aide d'un navigateur Internet.

Effectuez les vérifications suivantes pour voir si votre serveur web Apache est en cours d'exécution et accessible.

- Le serveur web est-il en cours d'exécution ?

  Vous pouvez vérifier que httpd est activé en exécutant la commande suivante :

  ```
  [ec2-user ~]$ chkconfig --list httpd
  httpd           0:off   1:off   2:on    3:on    4:on    5:on    6:off
  ```

  Ici, httpd a la valeur `on` sur les niveaux d'exécution 2, 3, 4 et 5 (ce qui correspond à ce que vous voulez voir).

  Si le processus httpd n'est pas en cours d'exécution, répétez les étapes décrites dans Étape 1 : Préparer le serveur LAMP (p. 48).

- Le pare-feu est-il configuré correctement ?

  Si vous ne pouvez pas voir la page test Apache, vérifiez que le groupe de sécurité que vous utilisez contient une règle pour autoriser le trafic HTTP (port 80). Pour plus d'informations sur l'ajout d'une règle HTTP à votre groupe de sécurité, consultez Ajout de règles à un groupe de sécurité (p. 1069).

Le logiciel d'application que je veux exécuter sur mon serveur est incompatible avec la version PHP installée ou d'autres logiciels

Ce didacticiel recommande d'installer les versions les plus récentes d'Apache HTTP Server, de PHP et de MySQL. Avant d'installer une application LAMP supplémentaire, vérifiez sa configuration requise pour confirmer qu'elle est compatible avec votre environnement installé. Si la dernière version de PHP n'est pas prise en charge, il est possible (en toute sécurité) de revenir à une configuration antérieure prise en charge. Vous pouvez également installer plusieurs versions de PHP en parallèle, ce qui résout certains problèmes de compatibilité avec un minimum d'effort. Pour plus d'informations sur la configuration d'une préférence parmi plusieurs versions PHP installées, consultez Notes de mise à jour d'Amazon Linux AMI 2016.09.

Comment revenir à une version plus ancienne

La version précédente bien testée de ce didacticiel appelait les packages LAMP de base suivants :

- `httpd24`
- `php56`
- `mysql55-server`
- `php56-mysqlnd`

Si vous avez déjà installé les derniers packages comme cela est recommandé au début de ce didacticiel, vous devez commencer par les désinstaller, ainsi que toutes les autres dépendances, comme suit :

```
[ec2-user ~]$ sudo yum remove -y httpd24 php72 mysql57-server php72-mysqlnd perl-DBD-
MySQL57
```

Installez ensuite l'environnement de remplacement :

```
[ec2-user ~]$ sudo yum install -y  httpd24 php56 mysql55-server php56-mysqlnd
```

Si vous décidez ensuite d'effectuer une mise à niveau vers l'environnement recommandé, vous devez d'abord supprimer les packages et dépendances personnalisés :

```
[ec2-user ~]$ sudo yum remove -y  httpd24 php56 mysql55-server php56-mysqlnd perl-DBD-
MySQL56
```

Vous pouvez maintenant installer les derniers packages comme indiqué précédemment.

Voir aussi

Pour plus d'informations sur le transfert des fichiers vers votre instance ou l'installation d'un blog WordPress sur votre serveur web, consultez la documentation suivante :

- Transfert de fichiers vers votre instance Linux à l'aide de WinSCP (p. 618)
- Transfert de fichiers des instances Linux à partir de Linux à l'aide de SCP (p. 603)
- Didacticiel : Hébergement d'un blog WordPress avec Amazon Linux (p. 59)

Pour plus d'informations sur les commandes et le logiciel utilisés dans ce didacticiel, consultez les pages web suivantes :

- Serveur Web Apache : http://httpd.apache.org/
- Serveur de base de données MySQL : http://www.mysql.com/
- Langage de programmation PHP : http://php.net/
- La commande `chmod` : https://en.wikipedia.org/wiki/Chmod
- La commande `chown` : https://en.wikipedia.org/wiki/Chown

Pour plus d'informations sur l'enregistrement d'un nom de domaine pour votre serveur web ou le transfert d'un nom de domaine existant vers cet hôte, consultez Création et migration de domaines et de sous-domaines vers Amazon Route 53 dans le Amazon Route 53 Manuel du développeur.

Didacticiel : Hébergement d'un blog WordPress avec Amazon Linux

Les procédures suivantes vous aideront à installer, configurer et sécuriser un blog WordPress sur votre instance Amazon Linux. Ce didacticiel est une bonne introduction sur l'utilisation d'Amazon EC2, dans la mesure où vous contrôlez entièrement un serveur web qui héberge votre blog WordPress, ce qui n'est pas classique avec un service d'hébergement traditionnel.

Vous êtes responsable de la mise à jour des packages logiciels et de la gestion des correctifs de sécurité pour votre serveur. Pour une installation de WordPress plus automatisée qui ne nécessite aucune interaction directe avec la configuration du serveur web, le service AWS CloudFormation propose un modèle WordPress qui peut vous permettre de débuter rapidement. Pour de plus amples informations,

veuillez consulter Mise en route dans le AWS CloudFormation Guide de l'utilisateur. Si vous préférez héberger votre blog WordPress sur une instance Windows, veuillez consulter Déploiement d'un blog WordPress sur votre instance Windows Amazon EC2 dans le manuel Amazon EC2 Guide de l'utilisateur pour les instances Windows. Si vous avez besoin d'une solution hautement disponible avec une base de données découplée, veuillez consulter Déploiement d'un site web WordPress haute disponibilité dans le AWS Elastic Beanstalk Manuel du développeur.

> **Important**
>
> Ces procédures sont destinées à une utilisation avec Amazon Linux. Pour obtenir plus d'informations sur d'autres distributions, consultez leur documentation spécifique. Plusieurs étapes de ce didacticiel ne fonctionnent pas sur des instances Ubuntu. Pour obtenir de l'aide sur l'installation de WordPress sur une instance Ubuntu, consultez WordPress dans la documentation Ubuntu.

Option : Effectuer ce didacticiel en utilisant Automation

Pour effectuer ce didacticiel en utilisant AWS Systems Manager Automation au lieu des tâches suivantes, exécutez l'un des documents Automation suivants : AWSDocs-HostingAWordPressBlog-AL (Amazon Linux) ou AWSDocs-HostingAWordPressBlog-AL2 (Amazon Linux 2).

Rubriques

- Prérequis (p. 60)
- Installer WordPress (p. 60)
- Étapes suivantes (p. 67)
- Aide! Mon nom DNS public a changé et mon blog ne fonctionne plus (p. 68)

Prérequis

Ce didacticiel suppose que vous avez lancé une instance Amazon Linux avec un serveur Web opérationnel à l'aide de la prise en charge PHP et de base de données (MySQL ou MariaDB) en suivant toutes les étapes dans Didacticiel : Installation d'un serveur web LAMP avec l'Amazon Linux AMI (p. 48) pour l'AMI Amazon Linux ou Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2 (p. 38) pour Amazon Linux 2. Ce didacticiel propose aussi des étapes pour la configuration d'un groupe de sécurité afin de permettre le trafic `HTTP` et `HTTPS` ainsi que plusieurs étapes afin de vous assurer que les autorisations sur les fichiers sont définies correctement pour votre serveur web. Pour plus d'informations sur l'ajout de règles à votre groupe de sécurité, consultez le didacticiel Ajout de règles à un groupe de sécurité (p. 1069).

Nous vous recommandons vivement d'associer une adresse IP Elastic à l'instance que vous utilisez pour héberger un blog WordPress. Cela évite à l'adresse DNS publique de votre instance de changer et de détériorer votre installation. Si vous possédez un nom de domaine et que vous voulez l'utiliser pour votre blog, vous pouvez mettre à jour l'enregistrement DNS pour que le nom de domaine pointe vers votre EIP (afin d'obtenir de l'aide à ce sujet, veuillez contacter votre serveur d'inscriptions des noms de domaine). Vous pouvez avoir une EIP associée à une instance en cours d'exécution sans coût aucun. Pour plus d'informations, consultez Adresses IP Elastic (p. 831).

Si vous n'avez pas encore de nom de domaine pour votre blog, vous pouvez enregistrer un nom de domaine avec Route 53 et associer l'EIP de votre instance à votre nom de domaine. Pour de plus amples informations, veuillez consulter Inscription de noms de domaines à l'aide d'Amazon Route 53 dans le manuel Amazon Route 53 Manuel du développeur.

Installer WordPress

Connectez-vous à votre instance et téléchargez le package d'installation WordPress.

Pour télécharger et décompresser le package d'installation WordPress

1. Téléchargez le dernier package d'installation WordPress avec la commande wget. La commande suivante devrait toujours télécharger la dernière version.

```
[ec2-user ~]$ wget https://wordpress.org/latest.tar.gz
```

2. Décompressez et désarchivez le package d'installation. Le dossier d'installation est décompressé dans un dossier appelé `wordpress`.

```
[ec2-user ~]$ tar -xzf latest.tar.gz
```

Pour créer une base de données et un utilisateur pour votre installation WordPress

Votre installation WordPress doit stocker des informations dans une base de données comme des billets de blog et des commentaires des utilisateurs. Cette procédure vous aide à créer la base de données de votre blog et un utilisateur qui est autorisé à lire et à enregistrer des informations dans cette dernière.

1. Démarrez le serveur de base de données.
   - Amazon Linux 2

     ```
     [ec2-user ~]$ sudo systemctl start mariadb
     ```

   - AMI Amazon Linux

     ```
     [ec2-user ~]$ sudo service mysqld start
     ```

2. Connectez-vous au serveur de base de données en tant qu'utilisateur `root`. Saisissez votre mot de passe `root` de base de données lorsque vous y êtes invité. Il peut être différent du mot de passe `root` de votre système ou il peut même être inexistant si vous n'avez pas sécurisé votre serveur de base de données.

   Si vous n'avez pas encore sécurisé votre serveur de base de données, il est important de le faire. Pour plus d'informations, consultez Pour sécuriser le serveur MariaDB (p. 44) (Amazon Linux 2) ou Pour sécuriser le serveur de base de données (p. 54) (AMI Amazon Linux).

```
[ec2-user ~]$ mysql -u root -p
```

3. Créez un utilisateur et un mot de passe pour votre base de données MySQL. Votre installation WordPress utilise ces valeurs pour communiquer avec votre base de données MySQL. Saisissez la commande suivante en remplaçant les informations par un nom utilisateur et un mot de passe uniques.

```
CREATE USER 'wordpress-user'@'localhost' IDENTIFIED BY 'your_strong_password';
```

   Assurez-vous de créer un mot de passe fiable pour votre utilisateur. N'utilisez pas l'apostrophe (') dans votre mot de passe, car elle détériorera la commande précédente. Pour plus d'informations sur la création d'un mot de passe fiable, consultez http://www.pctools.com/guides/password/. Ne réutilisez pas un mot de passe existant et assurez-vous de stocker ce mot de passe dans un endroit sûr.

4. Créez votre base de données. Donnez à votre base de données un nom descriptif pertinent comme `wordpress-db`.

   Note

   Les signes de ponctuation autour du nom de la base de données dans la commande ci-dessous sont appelés « accents graves ». La touche « accent grave » (`) est généralement située au-dessus de la touche `Tab` d'un clavier QWERTY standard. Les « accents graves »

ne sont pas toujours nécessaires, mais ils vous permettent d'utiliser des caractères qui sont normalement interdits dans les noms de base de données, comme les traits d'union.

```
CREATE DATABASE `wordpress-db`;
```

5. Accordez l'ensemble des privilèges de votre base de données à l'utilisateur WordPress que vous avez créé précédemment.

```
GRANT ALL PRIVILEGES ON `wordpress-db`.* TO "wordpress-user"@"localhost";
```

6. Annulez les privilèges de base de données pour récupérer tous vos changements.

```
FLUSH PRIVILEGES;
```

7. Quittez le client `mysql`.

```
exit
```

Pour créer et modifier le fichier wp-config.php

Le dossier d'installation WordPress contient un modèle de fichier de configuration appelé `wp-config-sample.php`. Dans cette procédure, vous copiez ce fichier avant de le modifier pour respecter votre configuration spécifique.

1. Copiez le fichier `wp-config-sample.php` sur un fichier appelé `wp-config.php`. Cela crée un nouveau fichier de configuration et garde le modèle de fichier original intact comme sauvegarde.

```
[ec2-user ~]$ cp wordpress/wp-config-sample.php wordpress/wp-config.php
```

2. Modifiez le fichier `wp-config.php` avec votre éditeur de texte préféré (comme nano ou vim) et saisissez les valeurs pour votre installation. Si vous n'avez pas d'éditeur de texte préféré, `nano` convient aux débutants.

```
[ec2-user ~]$ nano wordpress/wp-config.php
```

   a. Trouvez la ligne qui définit `DB_NAME` et remplacez `database_name_here` par le nom de la base de données que vous avez créée à l'Step 4 (p. 61) de la procédure Pour créer une base de données et un utilisateur pour votre installation WordPress (p. 61).

```
define('DB_NAME', 'wordpress-db');
```

   b. Trouvez la ligne qui définit `DB_USER` et remplacez `username_here` par l'utilisateur de base de données que vous avez créé à l'Step 3 (p. 61) de la procédure Pour créer une base de données et un utilisateur pour votre installation WordPress (p. 61).

```
define('DB_USER', 'wordpress-user');
```

   c. Trouvez la ligne qui définit `DB_PASSWORD` et remplacez `password_here` par le mot de passe fiable que vous avez créé à l'Step 3 (p. 61) de la procédure Pour créer une base de données et un utilisateur pour votre installation WordPress (p. 61).

```
define('DB_PASSWORD', 'your_strong_password');
```

   d. Trouvez la section appelée `Authentication Unique Keys and Salts`. Ces valeurs `KEY` et `SALT` offrent une layer de chiffrement aux cookies du navigateur que les utilisateurs de WordPress

stockent sur leurs machines locales. En gros, l'ajout de valeurs longues aléatoires à cet endroit rend votre site plus sécurisé. Consultez https://api.wordpress.org/secret-key/1.1/salt/ pour générer de façon aléatoire un ensemble de valeurs clés que vous pouvez copier et coller dans votre fichier `wp-config.php`. Pour coller du texte dans un terminal PuTTY, placez le curseur où vous voulez coller le texte et faites un clic droit avec votre souris dans le terminal PuTTY.

Pour plus d'informations sur les clés de sécurité, consultez https://wordpress.org/support/article/editing-wp-config-php/#security-keys.

> **Note**
>
> Les valeurs ci-dessous sont proposées à titre d'exemple seulement. N'utilisez pas ces valeurs pour votre installation.

```
define('AUTH_KEY',         ' #U$$+[RXN8:b^-L 0(WU_+ c+WFkI~c]o]-bHw+)/
Aj[wTwSiZ<Qb[mghEXcRh-');
define('SECURE_AUTH_KEY',  'Zsz._P=l/|y.Lq)XjlkwS1y5NJ76E6EJ.AV0pCKZZB,*~*r ?6OP
$eJT@;+(ndLg');
define('LOGGED_IN_KEY',    'ju}qwre3V*+8f_zOWf?{LlGsQ]Ye@2Jh^,8x>)Y |;(^[Iw]Pi
+LG#A4R?7N`YB3');
define('NONCE_KEY',        'P(g62HeZxEes|LnI^i=H,[XwK9I&[2s|:?0N}VJM%?;v2v]v+;
+^9eXUahg@::Cj');
define('AUTH_SALT',        'C$DpB4Hj[JK:?{ql`sRVa:{:7yShy(9A@5wg+`JJVb1fk%_-
Bx*M4(qc[Qg%JT!h');
define('SECURE_AUTH_SALT', 'd!uRu#}+q#{f$Z?Z9uFPG.${+S{n~1M&%@~gL>U>NV<zpD-@2-
Es7Q1O-bp28EKv');
define('LOGGED_IN_SALT',   ';j{00P*owZf)kVD+FVLn-~ >.|Y%Ug4#I^*LVd9QeZ^&XmK|e(76miC
+&W&+^0P/');
define('NONCE_SALT',       '-97r*V/cgxLmp?Zy4zUU4r99QQ_rGs2LTd%P;|
_e1tS)8_B/,.6[=UK<J_y9?JWG');
```

e. Enregistrez le fichier et quittez votre éditeur de texte.

Pour placer vos fichiers WordPress à la racine du document Apache

- Maintenant que vous avez décompressé le dossier d'installation, créé une base de données et un utilisateur MySQL et personnalisé le fichier de configuration WordPress, vous êtes prêt à copier vos fichiers d'installation à la racine du document du serveur web. Ainsi, vous pouvez exécuter le script d'installation pour terminer votre installation. L'emplacement de ces fichiers varie selon que vous préfériez avoir votre blog WordPress disponible à la racine réelle de votre serveur web (par exemple, *`my.public.dns.amazonaws.com`*) ou dans un sous-répertoire ou dossier à la racine (par exemple, *`my.public.dns.amazonaws.com/blog`*).
  - Si vous souhaitez que WordPress s'exécute à la racine de votre document, copiez le contenu du répertoire d'installation wordpress (et non le répertoire lui-même), comme suit :

    ```
    [ec2-user ~]$ cp -r wordpress/* /var/www/html/
    ```

  - Si vous souhaitez que WordPress s'exécute dans un autre répertoire à la racine du document, commencez par créer ce répertoire, puis copiez-y les fichiers. Dans cet exemple, WordPress s'exécutera à partir du répertoire `blog` :

    ```
    [ec2-user ~]$ mkdir /var/www/html/blog
    [ec2-user ~]$ cp -r wordpress/* /var/www/html/blog/
    ```

> **Important**
>
> A des fins de sécurité, si vous ne passez pas à la prochaine procédure immédiatement, arrêtez le serveur web Apache (`httpd`) dès maintenant. Une fois que vous avez déplacé votre installation

à la racine du document Apache, le script d'installation WordPress n'est pas protégé, et un pirate informatique est susceptible d'accéder à votre blog si le serveur Web Apache est en cours d'exécution. Pour arrêter le serveur Web Apache, saisissez la commande sudo service httpd stop. Si vous ne passez pas à la prochaine procédure, vous n'avez pas à arrêter le serveur web Apache.

Pour autoriser WordPress à utiliser les permalinks

Les permalinks WordPress doivent utiliser les fichiers `.htaccess` Apache pour fonctionner correctement, mais ce n'est pas activé par défaut sur Amazon Linux. Utilisez cette procédure pour permettre tous les remplacements à la racine du document Apache.

1. Ouvrez le fichier `httpd.conf` avec votre éditeur de texte préféré (comme nano ou vim). Si vous n'avez pas d'éditeur de texte préféré, `nano` convient aux débutants.

```
[ec2-user ~]$ sudo vim /etc/httpd/conf/httpd.conf
```

2. Trouvez la section qui commence par `<Directory "/var/www/html">`.

```
<Directory "/var/www/html">
    #
    # Possible values for the Options directive are "None", "All",
    # or any combination of:
    #   Indexes Includes FollowSymLinks SymLinksifOwnerMatch ExecCGI MultiViews
    #
    # Note that "MultiViews" must be named *explicitly* --- "Options All"
    # doesn't give it to you.
    #
    # The Options directive is both complicated and important.  Please see
    # http://httpd.apache.org/docs/2.4/mod/core.html#options
    # for more information.
    #
    Options Indexes FollowSymLinks

    #
    # AllowOverride controls what directives may be placed in .htaccess files.
    # It can be "All", "None", or any combination of the keywords:
    #   Options FileInfo AuthConfig Limit
    #
    AllowOverride None

    #
    # Controls who can get stuff from this server.
    #
    Require all granted
</Directory>
```

3. Modifiez la ligne `AllowOverride None` dans la section ci-dessus par `AllowOverride All`.

   Note

   Il existe plusieurs lignes `AllowOverride` dans ce fichier. Assurez-vous de modifier la ligne dans la section `<Directory "/var/www/html">`.

```
AllowOverride All
```

4. Enregistrez le fichier et quittez votre éditeur de texte.

Pour installer la bibliothèque de dessins graphiques PHP sur Amazon Linux 2

La bibliothèque GD pour PHP vous permet de modifier des images. Installez cette bibliothèque si vous devez recadrer l'image d'en-tête pour votre blog. La version de phpMyAdmin que vous installez peut nécessiter une version minimale spécifique de cette bibliothèque (par exemple, la version 7.2).

Utilisez la commande suivante pour installer la bibliothèque de dessin graphique PHP sur Amazon Linux 2. Par exemple, si vous avez installé php7.2 depuis amazon-linux-extras dans le cadre de l'installation de la pile LAMP, cette commande installe la version 7.2 de la bibliothèque de dessins graphiques PHP.

```
[ec2-user ~]$ sudo yum install php-gd
```

Pour vérifier la version installée, utilisez la commande suivante :

```
[ec2-user ~]$ sudo yum list installed | grep php-gd
```

Voici un exemple de sortie :

```
php-gd.x86_64                    7.2.30-1.amzn2            @amzn2extra-php7.2
```

Pour installer la bibliothèque de dessins graphiques PHP sur Amazon Linux AMI

La bibliothèque GD pour PHP vous permet de modifier des images. Installez cette bibliothèque si vous devez recadrer l'image d'en-tête pour votre blog. La version de phpMyAdmin que vous installez peut nécessiter une version minimale spécifique de cette bibliothèque (par exemple, la version 7.2).

Pour vérifier quelles versions sont disponibles, utilisez la commande suivante :

```
[ec2-user ~]$ yum list | grep php-gd
```

Voici un exemple de ligne de sortie de la bibliothèque de dessin graphique PHP (version 7.2) :

```
php72-gd.x86_64                  7.2.30-1.22.amzn1         amzn-updates
```

Utilisez la commande suivante pour installer une version spécifique de la bibliothèque de dessin graphique PHP (par exemple, la version 7.2) sur Amazon Linux AMI :

```
[ec2-user ~]$ sudo yum install php72-gd
```

Pour corriger les autorisations sur les fichiers pour le serveur web Apache

Certaines fonctions disponibles dans WordPress nécessitent un accès en écriture à la racine du document Apache (comme le chargement de supports via des écrans d'administration). Si vous ne l'avez pas déjà fait, appliquez les autorisations et appartenances aux groupes suivantes (comme décrit plus en détail dans le didacticiel sur le serveur web LAMP (p. 48)).

1. Accordez la propriété du fichier `/var/www` et de son contenu à l'utilisateur `apache`.

   ```
   [ec2-user ~]$ sudo chown -R apache /var/www
   ```

2. Accordez la propriété de groupe de `/var/www` et de son contenu au groupe `apache`.

   ```
   [ec2-user ~]$ sudo chgrp -R apache /var/www
   ```

3. Modifiez les autorisations sur les répertoires de `/var/www` et ses sous-répertoires pour ajouter des autorisations d'écriture de groupe et définir l'ID de groupe pour les futurs sous-répertoires.

   ```
   [ec2-user ~]$ sudo chmod 2775 /var/www
   ```

```
[ec2-user ~]$ find /var/www -type d -exec sudo chmod 2775 {} \;
```

4. Modifiez de façon récursive les autorisations sur les fichiers de `/var/www` et ses sous-répertoires pour ajouter des autorisations d'écriture de groupe.

```
[ec2-user ~]$ find /var/www -type f -exec sudo chmod 0664 {} \;
```

5. Redémarrez le serveur web Apache pour récupérer les nouveaux groupe et autorisations.
   - Amazon Linux 2

     ```
     [ec2-user ~]$ sudo systemctl restart httpd
     ```

   - AMI Amazon Linux

     ```
     [ec2-user ~]$ sudo service httpd restart
     ```

Pour exécuter le script d'installation WordPress avec Amazon Linux 2

Vous êtes maintenant prêt à installer WordPress. Les commandes que vous utilisez dépendent du système d'exploitation. Les commandes dans cette procédure sont destinées à une utilisation avec Amazon Linux 2. Exécutez la procédure suivant celle-ci avec AMI Amazon Linux.

1. Utilisez la commande systemctl pour vous assurer que les services `httpd` et de base de données commencent à chaque démarrage système.

   ```
   [ec2-user ~]$ sudo systemctl enable httpd && sudo systemctl enable mariadb
   ```

2. Vérifiez que le serveur de base de données est en cours d'exécution.

   ```
   [ec2-user ~]$ sudo systemctl status mariadb
   ```

   Si le service de base de données n'est pas en cours d'exécution, démarrez-le.

   ```
   [ec2-user ~]$ sudo systemctl start mariadb
   ```

3. Vérifiez que votre serveur web Apache (`httpd`) est en cours d'exécution.

   ```
   [ec2-user ~]$ sudo systemctl status httpd
   ```

   Si le service `httpd` n'est pas en cours d'exécution, démarrez-le.

   ```
   [ec2-user ~]$ sudo systemctl start httpd
   ```

4. Dans un navigateur Web, saisissez l'URL de votre blog WordPress (soit l'adresse DNS publique pour votre instance soit cette adresse suivie par le dossier `blog`). Vous devriez voir le script d'installation WordPress. Fournissez les informations requises lors de l'installation de WordPress. Choisissez Installer WordPress pour terminer l'installation. Pour de plus amples informations, veuillez consulter la section relative à l'étape 5 : Exécuter le script d'installation sur le site web de WordPress.

Exécuter le script d'installation WordPress avec AMI Amazon Linux

1. Utilisez la commande chkconfig pour vous assurer que les services `httpd` et de base de données commencent à chaque démarrage système.

```
[ec2-user ~]$ sudo chkconfig httpd on && sudo chkconfig mysqld on
```

2. Vérifiez que le serveur de base de données est en cours d'exécution.

```
[ec2-user ~]$ sudo service mysqld status
```

Si le service de base de données n'est pas en cours d'exécution, démarrez-le.

```
[ec2-user ~]$ sudo service mysqld start
```

3. Vérifiez que votre serveur web Apache (`httpd`) est en cours d'exécution.

```
[ec2-user ~]$ sudo service httpd status
```

Si le service `httpd` n'est pas en cours d'exécution, démarrez-le.

```
[ec2-user ~]$ sudo service httpd start
```

4. Dans un navigateur Web, saisissez l'URL de votre blog WordPress (soit l'adresse DNS publique pour votre instance soit cette adresse suivie par le dossier `blog`). Vous devriez voir le script d'installation WordPress. Fournissez les informations requises lors de l'installation de WordPress. Choisissez Installer WordPress pour terminer l'installation. Pour de plus amples informations, veuillez consulter la section relative à l'étape 5 : Exécuter le script d'installation sur le site web de WordPress.

Étapes suivantes

Après avoir testé votre blog WordPress, pensez à mettre à jour sa configuration.

Utiliser un nom de domaine personnalisé

Si vous avez un nom de domaine associé à l'EIP de votre instance EC2, vous pouvez configurer votre blog pour utiliser ce nom au lieu de l'adresse DNS publique EC2. Pour de plus amples informations, veuillez consulter la section relative à la modification de l'URL du site sur le site WordPress.

Configurer votre blog

Vous pouvez configurer votre blog pour utiliser différents thèmes et plugins afin de proposer une expérience plus personnalisée à vos lecteurs. Cependant, il peut arriver que le processus d'installation échoue ce qui entraînera la perte de tout votre blog. Nous vous recommandons vivement de créer une sauvegarde Amazon Machine Image (AMI) de votre instance avant d'essayer d'installer des thèmes ou des plugins. Ainsi, vous pouvez restaurer votre blog en cas de problème pendant l'installation. Pour plus d'informations, consultez Création de votre propre AMI (p. 104).

Augmenter la capacité

Si votre blog WordPress devient populaire et que vous avez besoin de plus de puissance de calcul ou de stockage, pensez aux étapes suivantes :

- Développez l'espace de stockage sur votre instance. Pour plus d'informations, consultez Amazon EBS Elastic Volumes (p. 1182).
- Déplacez votre base de données MySQL vers Amazon RDS pour profiter de la capacité du service à se mettre à l'échelle facilement.
- Migrez vers un type d'instance plus important. Pour plus d'informations, consultez Modification du type d'instance (p. 313).

- Ajoutez d'autres instances. Pour plus d'informations, consultez Didacticiel : Augmenter la disponibilité des applications sur Amazon EC2 (p. 98).

En savoir plus sur WordPress

Pour obtenir des informations sur WordPress, consultez la documentation d'aide au sujet de WordPress Codex sur http://codex.wordpress.org/. Pour de plus amples informations sur la résolution des problèmes de votre installation, rendez-vous sur https://wordpress.org/support/article/how-to-install-wordpress/#common-installation-problems. Pour obtenir des informations sur la sécurisation de votre blog WordPress, veuillez consultez https://wordpress.org/support/article/hardening-wordpress/. Pour obtenir des informations sur la mise à jour de votre blog WordPress, veuillez consulter https://wordpress.org/support/article/updating-wordpress/.

Aide! Mon nom DNS public a changé et mon blog ne fonctionne plus

Votre installation WordPress est automatiquement configurée à l'aide de l'adresse DNS publique pour votre instance EC2. Si vous arrêtez et redémarrez l'instance, l'adresse DNS publique change (à moins qu'elle soit associée à une adresse IP Elastic) et votre blog ne fonctionne plus, car il fait référence à des ressources à une adresse qui n'existe plus (ou qui est assignée à une autre instance EC2). Une description plus détaillée du problème et plusieurs solutions possibles sont présentées sur https://wordpress.org/support/article/changing-the-site-url/.

Si cela est arrivé à votre installation WordPress, vous pouvez récupérer votre blog avec la procédure ci-dessous qui utilise l'interface de ligne de commande wp-cli pour WordPress.

Pour changer l'URL de votre site WordPress avec la commande wp-cli

1. Connectez-vous à votre instance EC2 avec SSH.
2. Notez l'ancienne URL de site et la nouvelle URL de site pour votre instance. L'ancienne URL de site ressemble au nom DNS public pour votre instance EC2 lorsque vous avez installé WordPress. La nouvelle URL de site correspond au nom DNS public actuel pour votre instance EC2. Si vous n'êtes pas certain de votre ancienne URL de site, vous pouvez utiliser curl pour la trouver avec la commande suivante.

   ```
   [ec2-user ~]$ curl localhost | grep wp-content
   ```

   Vous devriez voir des références à votre ancien nom DNS public dans les données de sortie qui ressembleront à cela (ancienne URL de site en rouge) :

   ```
   <script type='text/javascript' src='http://ec2-52-8-139-223.us-
   west-1.compute.amazonaws.com/wp-content/themes/twentyfifteen/js/functions.js?
   ver=20150330'></script>
   ```

3. Téléchargez le kit wp-cli avec la commande suivante.

   ```
   [ec2-user ~]$ curl -O https://raw.githubusercontent.com/wp-cli/builds/gh-pages/phar/wp-
   cli.phar
   ```

4. Recherchez et remplacez l'ancienne URL de site dans votre installation WordPress avec la commande suivante. Remplacez l'ancienne et la nouvelle URL de site pour votre instance EC2 ainsi que le chemin vers votre installation WordPress (généralement `/var/www/html` ou `/var/www/html/blog`).

   ```
   [ec2-user ~]$ php wp-cli.phar search-replace 'old_site_url' 'new_site_url' --path=/
   path/to/wordpress/installation --skip-columns=guid
   ```

5. Dans un navigateur web, saisissez la nouvelle URL de site de votre blog WordPress pour vérifier que le site fonctionne correctement à nouveau. Si ce n'est pas le cas, veuillez consulter https://wordpress.org/support/article/changing-the-site-url/ et https://wordpress.org/support/article/how-to-install-wordpress/#common-installation-problems pour de plus amples informations.

Didacticiel : Configurer SSL/TLS sur Amazon Linux 2

SSL/TLS (Secure Sockets Layer/Transport Layer Security) crée un canal chiffré entre un serveur web et un client web qui empêche les données en transit d'être écoutées. Ce didacticiel explique comment ajouter manuellement la prise en charge de SSL/TLS sur une instance EC2 avec Amazon Linux 2 et le serveur Web Apache. Ce didacticiel suppose que vous n'utilisez pas d'équilibreur de charge. Si vous utilisez un équilibreur de charge, vous pouvez choisir d'utiliser AWS Certificate Manager à la place.

Pour des raisons historiques, le chiffrement web est communément appelé SSL. Alors que les navigateurs web prennent toujours en charge SSL, son protocole successeur TLS est moins vulnérable en cas d'attaque. Amazon Linux 2 désactive la prise en charge côté serveur pour toutes les versions SSL par défaut. Les organismes de normes de sécurité considèrent TLS 1.0 comme peu sûr, et TLS 1.0 et TLS 1.1 sont sur le point d'être déclarés formellement obsolètes par l'IETF. Ce didacticiel contient des conseils pour l'activation de TLS 1.2 exclusivement. (Un protocole TLS 1.3 plus récent existe, mais il n'est pas installé par défaut sur Amazon Linux 2). Pour plus d'informations sur les normes de chiffrement mises à jour, consultez RFC 7568 et RFC 8446.

Ce didacticiel fait référence au chiffrement Web moderne simplement comme TLS.

> **Important**
>
> Ces procédures sont destinées à une utilisation avec Amazon Linux 2. Nous supposons également que vous commencez avec une nouvelle instance Amazon EC2. Si vous essayez de configurer un serveur web LAMP sur une instance avec une autre distribution ou si vous réaffectez une instance existante plus ancienne, certaines procédures de ce didacticiel peuvent ne pas fonctionner pour vous. Pour plus d'informations sur les serveurs web LAMP sur Ubuntu, consultez la rubrique ApacheMySQLPHP de la documentation proposée par la communauté Ubuntu. Pour plus d'informations sur Red Hat Enterprise Linux, consultez la rubrique Web Servers du portail client.

Table des matières

- Prérequis (p. 69)
- Étape 1 : Activer TLS sur le serveur (p. 70)
- Étape 2 : Obtenir un certificat signé par une autorité de certification (CA) (p. 72)
- Étape 3 : Tester et renforcer la configuration de sécurité (p. 77)
- Dépannage (p. 80)
- Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux 2 (p. 81)

Prérequis

Avant de commencer ce didacticiel, suivez les étapes suivantes :

- Lancez une instance Amazon Linux 2 basée sur les volumes EBS. Pour plus d'informations, consultez Étape 1 : Lancement d'une instance (p. 33).

- Configurez vos groupes de sécurité afin que votre instance puisse accepter des connexions sur les ports TCP suivants :
  - SSH (port 22)
  - HTTP (port 80)
  - HTTPS (port 443)

  Pour plus d'informations, consultez Autorisation du trafic entrant pour vos instances Linux (p. 1044).
- Installez le serveur Web Apache. Pour des instructions pas à pas, consultez le Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2 (p. 38). Seuls le package httpd et ses dépendances sont nécessaires. Par conséquent, vous pouvez ignorer les instructions impliquant PHP et MariaDB.
- Pour identifier et authentifier les sites web, l'infrastructure à clés publiques (PKI) TLS repose sur le système de noms de domaine (DNS). Pour utiliser votre instance EC2 pour héberger un site web public, vous devez enregistrer un nom de domaine pour votre serveur web ou transférer un nom de domaine existant vers votre hôte Amazon EC2. Plusieurs services d'enregistrement de domaines tiers et d'hébergement DNS sont disponibles pour cela, ou vous pouvez utiliser Amazon Route 53.

Étape 1 : Activer TLS sur le serveur

Cette procédure vous décrit le processus de paramétrage de TLS sur Amazon Linux 2 avec un certificat auto-signé numérique.

> **Note**
>
> Un certificat auto-signé est acceptable dans un environnement de test, mais pas pour les environnements de production. Si vous exposez votre certificat auto-signé sur Internet, les visiteurs de votre site verront s'afficher des messages d'avertissement de sécurité.

Pour activer TLS sur un serveur

1. Connectez-vous à votre instance (p. 34) et confirmez qu'Apache est en cours d'exécution.

   ```
   [ec2-user ~]$ sudo systemctl is-enabled httpd
   ```

   Si la valeur renvoyée n'est pas « activé », démarrez Apache et configurez-le pour qu'il démarre à chaque amorçage du système.

   ```
   [ec2-user ~]$ sudo systemctl start httpd && sudo systemctl enable httpd
   ```

2. Pour vous assurer que tous vos packages logiciels sont mis à jour, effectuez une mise à jour logicielle rapide sur votre instance. Ce processus peut prendre quelques minutes, mais il est important pour vous assurer que vous disposez des dernières mises à jour de sécurité et des nouveaux correctifs de bogues.

   > **Note**
   >
   > L'option `-y` installe les mises à jour sans demander de confirmation. Si vous souhaitez examiner les mises à jour avant l'installation, vous pouvez omettre cette option.

   ```
   [ec2-user ~]$ sudo yum update -y
   ```

3. Maintenant que votre instance est à jour, ajoutez la prise en charge de TLS en installant le module Apache `mod_ssl`.

   ```
   [ec2-user ~]$ sudo yum install -y mod_ssl
   ```

Votre instance dispose désormais des fichiers suivants que vous utilisez pour configurer votre serveur sécurisé et créer un certificat pour les tests :

- `/etc/httpd/conf.d/ssl.conf`

  Le fichier de configuration de mod_ssl. Il contient des directives indiquant à Apache où trouver les clés et les certificats de chiffrement, les versions de protocoles TLS à autoriser et les algorithmes de chiffrement à accepter.
- `/etc/pki/tls/certs/make-dummy-cert`

  Script pour générer un certificat X.509 auto-signé et une clé privée pour votre hôte serveur. Ce certificat est utile pour vérifier qu'Apache est correctement paramétré pour utiliser TLS. Comme il n'offre aucune preuve d'identité, il ne doit pas être utilisé en production. S'il est utilisé en production, il déclenche des avertissements dans les navigateurs web.

4. Exécutez le script pour générer un certificat factice auto-signé et une clé pour les tests.

```
[ec2-user ~]$ cd /etc/pki/tls/certs
sudo ./make-dummy-cert localhost.crt
```

Cela génère un nouveau fichier `localhost.crt` dans le répertoire `/etc/pki/tls/certs/`. Le nom de fichier spécifié correspond au nom par défaut attribué dans la directive SSLCertificateFile dans `/etc/httpd/conf.d/ssl.conf`.

Ce fichier contient un certificat auto-signé et la clé privée du certificat. Apache exige que le certificat et la clé soient au format PEM qui est constitué de caractères ASCII codés en Base64, encadrés par des lignes « BEGIN » et « END », comme dans l'exemple abrégé ci-après.


```
-----BEGIN PRIVATE KEY-----
MIIEvgIBADANBgkqhkiG9w0BAQEFAASCBKgwggSkAgEAAoIBAQD2KKx/8Zk94m1q
3gQMZF9ZN66Ls19+3tHAgQ5Fpo9KJDhzLjOOCI8u1PTcGmAah5kEitCEc0wzmNeo
BCl0wYR6G0rGaKtK9Dn7CuIjvubtUysVyQoMVPQ97ldeakHWeRMiEJFXg6kZZ0vr
GvwnKoMh3DlK44D9dX7IDua2PlYx5+eroA+1Lqf32ZSaAO0bBIMIYTHigwbHMZoT
...
56tE7THvH7vOEf4/iUOsIrEzaMaJ0mqkmY1A70qQGQKBgBF3H1qNRNHuyMcPODFs
27hDzPDinrquSEvoZIggkDMlh2irTiipJ/GhkvTpoQlv0fK/VXw8vSgeaBuhwJvS
LXU9HvYq0U6O4FgD3nAyB9hI0BE13r1HjUvbjT7moH+RhnNz6eqqdscCS09VtRAo
4QQvAqOa8UheYeoXLdWcHaLP
-----END PRIVATE KEY-----

-----BEGIN CERTIFICATE-----
MIIEazCCA1OgAwIBAgICWxQwDQYJKoZIhvcNAQELBQAwgbExCzAJBgNVBAYTAi0t
MRIwEAYDVQQIDAlTb21lU3RhdGUxETAPBgNVBAcMCFNvbWVDaXR5MRkwFwYDVQQK
DBBTb21lT3JnYW5pemF0aW9uMR8wHQYDVQQLDBZTb21lT3JnYW5pemF0aW9uYWxV
bml0MRkwFwYDVQQDDBBpcC0xNzItMzEtMjAtMjM2MSQwIgYJKoZIhvcNAQkBFhVy
...
z5rRUE/XzxRLBZOoWZpNWTXJkQ3uFYH6s/sBwtHpKKZMzOvDedREjNKAvk4ws6F0
CuIjvubtUysVyQoMVPQ97ldeakHWeRMiEJFXg6kZZ0vrGvwnKoMh3DlK44D9dlU3
WanXWehT6FiSZvB4sTEXXJN2jdw8g+sHGnZ8zCOsclknYhHrCVD2vnBlZJKSZvak
3ZazhBxtQSukFMOnWPP2a0DMMFGYUHOd0BQE8sBJxg==
-----END CERTIFICATE-----
```


Les noms et extensions de fichiers sont fournis à titre indicatif et n'ont aucun effet sur la fonction. Par exemple, vous pouvez appeler un certificat `cert.crt`, `cert.pem`, ou tout autre nom de fichier dans la mesure où la directive associée dans le fichier `ssl.conf` utilise le même nom.

> **Note**
>
> Lorsque vous remplacez les fichiers TLS par défaut par vos propres fichiers personnalisés, veillez à ce qu'ils soient au format PEM.

5. Ouvrez le fichier `/etc/httpd/conf.d/ssl.conf` et supprimez le commentaire sur la ligne suivante, car le certificat fictif signé automatiquement contient aussi la clé. Si vous ne le faites pas avant d'exécuter l'étape suivante, le service Apache ne peut pas démarrer.

```
SSLCertificateKeyFile /etc/pki/tls/private/localhost.key
```

6. Redémarrez Apache.

```
[ec2-user ~]$ sudo systemctl restart httpd
```

   Note

   Assurez-vous que le port TCP 443 est accessible sur votre instance EC2, tel que décrit précédemment.

7. Votre serveur web Apache devrait maintenant prendre en charge HTTPS (HTTP sécurisé) sur le port 443. Testez-le en saisissant l'adresse IP ou le nom de domaine complet de votre instance EC2 dans une barre URL du navigateur avec le préfixe **`https://`**.

   Étant donné que vous vous connectez à un site avec un certificat d'hôte auto-signé non approuvé, il se peut que votre navigateur affiche une série d'avertissements de sécurité. Ignorez-les et poursuivez sur le site.

   Si la page de test Apache par défaut s'ouvre, cela signifie que vous avez configuré correctement TLS sur votre serveur. Toutes les données transmises entre le navigateur et le serveur sont maintenant chiffrées.

   Note

   Pour éviter aux visiteurs du site d'avoir des avertissements, vous devez obtenir un certificat signé par une CA qui chiffre mais vous authentifie aussi publiquement comme le propriétaire du site.

Étape 2 : Obtenir un certificat signé par une autorité de certification (CA)

Vous pouvez utiliser le processus suivant pour obtenir un certificat signé par une CA :

- Générez une demande de signature de certificat (CSR) à partir d'une clé privée
- Envoyez la demande de signature de certificat (CSR) à une autorité de certification (CA)
- Obtenez un certificat d'hôte signé
- Configurez Apache pour utiliser le certificat

Le chiffrement d'un certificat X.509 TLS auto-signé est identique à celui d'un certificat signé par une autorité de certification. La différence est sociale, pas mathématique. Une autorité de certification promet, au minimum, de valider la propriété d'un domaine avant de générer un certificat pour un demandeur. Chaque navigateur web contient une liste d'autorités de certification approuvées par le fournisseur de navigateur pour faire cela. Un certificat X.509 se compose surtout d'une clé publique qui correspond à votre clé de serveur privée et d'une signature de l'autorité de certification qui est cryptographiquement reliée à la clé publique. Lorsqu'un navigateur se connecte à un serveur web sur HTTPS, le serveur présente un certificat que le navigateur doit vérifier par rapport à sa liste d'autorités de certification approuvées. Si le signataire est sur la liste ou s'il est accessible via une chaîne de confiance composée d'autres utilisateurs de confiance, le navigateur négocie un canal de données chiffrées rapide avec le serveur et charge la page.

> **Important**
>
> Les certificats coûtent généralement de l'argent à cause travail impliqué dans la validation des requêtes, donc il est intéressant de comparer les prix. Une liste d'autorités de certification connues est proposée sur dmoztools.net. Quelques autorités de certification offrent des certificats basiques gratuits. La plus importante autorité de certification est le projet Let's Encrypt, qui prend également en charge l'automatisation du processus de création et de renouvellement des certificats. Pour de plus amples informations sur l'utilisation de Let's Encrypt comme autorité de certification, consultez Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux 2 (p. 81).

La clé est l'élément sous-jacent du certificat d'hôte. Depuis 2019, des groupes gouvernementaux et industriels recommandent l'utilisation d'une taille de clé minimale (module) de 2048 bits pour les clés RSA conçues pour protéger des documents, jusqu'en 2030. La taille de module par défaut générée par OpenSSL dans Amazon Linux 2 est de 2048 bits qui convient à l'utilisation d'un certificat signé par une CA. Dans la procédure suivante, étape facultative pour ceux qui souhaitent une clé personnalisée, par exemple, une clé avec un module plus important ou utilisant un algorithme de chiffrement différent.

Ces instructions pour l'acquisition de certificats d'hôte signés par l'autorité de certification (CA) ne fonctionnent pas à moins que vous possédiez un domaine DNS enregistré et hébergé.

Pour obtenir un certificat signé par une CA

1. Connectez-vous à votre instance (p. 34) et accédez à /etc/pki/tls/private/. Il s'agit du répertoire où vous stockez la clé privée du serveur pour TLS. Si vous préférez utiliser une clé d'hôte existante pour générez la CSR, passez à l'étape 3.
2. (Facultatif) Générez une nouvelle clé privée. Voici quelques exemples de configurations de clés. Toutes les clés obtenues fonctionnent avec votre serveur web, mais elles diffèrent dans le degré et le type de sécurité qu'elles mettent en œuvre.

   - Exemple 1 : création d'une clé d'hôte RSA par défaut. Le fichier obtenu, **`custom.key`**, est une clé privée RSA 2048 bits.

     ```
     [ec2-user ~]$ sudo openssl genrsa -out custom.key
     ```

   - Exemple 2 : création d'une clé RSA plus forte avec un modulus plus grand. Le fichier obtenu, **`custom.key`**, est une clé privée RSA 4096 bits.

     ```
     [ec2-user ~]$ sudo openssl genrsa -out custom.key 4096
     ```

   - Exemple 3 : création d'une clé RSA chiffrée 4096 bits avec protection par mot de passe. Le fichier obtenu, **`custom.key`**, est une clé privée RSA 4096 bits chiffrée avec le chiffrement AES-128.

     > **Important**
     >
     > Le chiffrement de la clé offre une plus grande sécurité, mais comme une clé chiffrée nécessite un mot de passe, les services qui en dépendent ne peuvent pas démarrer automatiquement. A chaque fois que vous utilisez cette clé, vous devez fournir le mot de passe (« abcde12345 » dans l'exemple précédent) sur une connexion SSH.

     ```
     [ec2-user ~]$ sudo openssl genrsa -aes128 -passout pass:abcde12345 -out custom.key
      4096
     ```

   - Exemple 4 : création d'une clé avec un chiffrement non RSA. La cryptographie RSA peut être relativement lente en raison de la taille de ses clés publiques, lesquelles sont basées sur le produit de deux grands nombres premiers. Cependant, il est possible de créer des clés pour TLS qui utilisent des chiffrements non RSA. Les clés basées sur les mathématiques des courbes elliptiques sont plus petites et plus rapides en termes de calcul, tout en offrant un niveau de sécurité équivalent.

```
[ec2-user ~]$ sudo openssl ecparam -name prime256v1 -out custom.key -genkey
```

Le résultat est une clé privée 256 bits à courbes elliptiques utilisant prime256v1, une « courbe nommée » que OpenSSL prend en charge. Sa qualité cryptographique est légèrement plus importante qu'une clé RSA 2048 bits, selon NIST.

> **Note**
>
> Les autorités de certification ne fournissent pas toutes le même niveau de support pour les clés basées sur les courbes elliptiques que pour les clés RSA.

Assurez-vous que la nouvelle clé privée possède un critère de propriété et d'autorisations très restrictif (propriétaire=racine, groupe=racine, lecture/écriture pour propriétaire uniquement). Les commandes seraient similaires à celles illustrées dans l'exemple suivant.

```
[ec2-user ~]$ sudo chown root:root custom.key
[ec2-user ~]$ sudo chmod 600 custom.key
[ec2-user ~]$ ls -al custom.key
```

Les commandes ci-avant produisent le résultat suivant.

```
-rw------- root root custom.key
```

Une fois que vous avez créé et configuré une clé satisfaisante, vous pouvez créer une CSR.

3. Créez une CSR à l'aide de votre clé préférée. L'exemple suivant utilise **`custom.key`**.

```
[ec2-user ~]$ sudo openssl req -new -key custom.key -out csr.pem
```

OpenSSL ouvre une boîte de dialogue et vous invite à compléter les informations affichées dans le tableau ci-dessous. Tous les champs à l'exception de Common Name sont facultatifs pour un certificat d'hôte basique avec validation de domaine.

Nom	Description	Exemple
Nom du pays	Abréviation ISO de deux lettres de votre pays.	US (=Etats-Unis)
Nom de l'état ou de la province	Nom de l'état ou de la province où votre organisation se situe. Ce nom ne peut pas être abrégé.	Washington
Nom de la localité	L'emplacement de votre organisation, comme une ville.	Seattle
Nom de l'organisation	Nom légal complet de votre organisation. N'abrégez pas le nom de votre organisation.	Exemple d'entreprise
Nom de l'unité d'organisation	Informations supplémentaires sur l'organisation, s'il y en a.	Exemple de service
Nom commun	Cette valeur doit correspondre exactement à l'adresse web que les utilisateurs saisiront dans un navigateur, selon vous. Il s'agit généralement d'un nom de domaine avec un nom d'hôte ou un alias préfixé sous la forme **`www.example.com`**.	www.exemple.com

Nom	Description	Exemple
	Dans les essais avec un certificat auto-signé et aucune résolution DNS, le nom commun peut se composer uniquement du nom d'hôte. Les autorités de certification proposent aussi des certificats onéreux qui acceptent les noms inconnus comme **`*.example.com`**.	
Adresse e-mail	L'adresse e-mail de l'administrateur du serveur.	quelquun@exemple.com

Au final, OpenSSL vous invite à donner un mot de passe de stimulation facultatif. Ce mot de passe s'applique uniquement à la CSR et aux transactions entre vous et votre autorité de certification, donc suivez les recommandations de l'autorité de certification sur cela, l'autre champ facultatif et le nom de l'entreprise facultatif. Le mot de passe de stimulation de la CSR n'a aucun effet sur le fonctionnement du serveur.

Le fichier obtenu **`csr.pem`** contient votre clé publique, la signature numérique de votre clé publique et les métadonnées que vous avez saisies.

4. Envoyez la CSR à une autorité de certification. Elle consiste généralement en l'ouverture de votre fichier CSR dans un éditeur de texte et la reproduction du contenu dans un formulaire web. A ce moment-là, il se peut que l'on vous demande de fournir un SAN (subject alternate name) ou plus à placer sur le certificat. Si **`www.example.com`** est le nom commun, **`example.com`** serait un bon SAN, et vice versa. Un visiteur de votre site qui saisit l'un de ces noms devrait bénéficier d'une connexion sans erreur. Si le formulaire web de votre autorité de certification le permet, incluez le nom commun dans la liste des SAN. Certaines autorités de certification l'incluent automatiquement.

   Une fois que votre demande a été approuvée, vous recevrez un nouveau certificat d'hôte signé par l'autorité de certification. Il se peut que l'on vous demande également de télécharger un fichier de certificat intermédiaire qui contient des certificats supplémentaires nécessaires pour compléter la chaîne de confiance de l'autorité de certification.

   > **Note**
   >
   > Votre autorité de certification peut vous envoyer des fichiers sous différents formats en fonction des finalités recherchées. Dans ce didacticiel, vous devez utiliser uniquement un fichier de certificat au format PEM, qui comporte habituellement (mais pas toujours) une extension de fichier `.pem` ou `.crt`. Si vous ne savez pas quel fichier utiliser, ouvrez les fichiers dans un éditeur de texte et recherchez celui qui contient un ou plusieurs blocs commençant par la ligne suivante.
   >
   > ```
   > - - - - -BEGIN CERTIFICATE - - - - -
   > ```
   >
   > Le fichier doit également se terminer par la ligne suivante.
   >
   > ```
   > - - - -END CERTIFICATE - - - - -
   > ```
   >
   > Vous pouvez également tester le fichier dans la ligne de commande, comme suit.
   >
   > ```
   > [ec2-user certs]$ openssl x509 -in certificate.crt -text
   > ```
   >
   > Vérifiez que ces lignes apparaissent dans le fichier. N'utilisez pas de fichiers se terminant par `.p7b`, `.p7c` ou autres extensions similaires.

5. Placez le nouveau certificat signé par une CA et les certificats intermédiaires dans le répertoire `/etc/pki/tls/certs`.

Note

Il existe plusieurs méthodes pour charger votre nouveau certificat dans votre instance EC2, mais le moyen le plus simple et informatif consiste à ouvrir un éditeur de texte (vi, nano, Bloc-notes, etc.) sur votre ordinateur local et votre instance, puis à copier et coller le contenu du fichier. Pour effectuer ces opérations sur l'instance EC2, vous devez disposer de privilèges racine [sudo]. Vous voyez ainsi immédiatement s'il existe des problèmes d'autorisation ou de chemin d'accès. Veillez toutefois à ne pas d'ajouter des lignes supplémentaires lors de la copie du contenu, ou à les modifier de quelque façon.

À partir du répertoire `/etc/pki/tls/certs`, vérifiez que les paramètres de propriété du fichier, de groupe et d'autorisation correspondent aux valeurs par défaut Amazon Linux 2 très restrictives (propriétaire=racine, groupe=racine, lecture/écriture pour propriétaire uniquement). L'exemple suivant illustre les commandes à utiliser.

```
[ec2-user certs]$ sudo chown root:root custom.crt
[ec2-user certs]$ sudo chmod 600 custom.crt
[ec2-user certs]$ ls -al custom.crt
```

Ces commandes devraient générer le résultat suivant.

```
-rw------- root root custom.crt
```

Les autorisations pour le fichier de certificat intermédiaire sont moins contraignantes (propriétaire=racine, groupe=racine, le propriétaire peut écrire, le groupe peut lire, tout le monde peut lire). L'exemple suivant illustre les commandes à utiliser.

```
[ec2-user certs]$ sudo chown root:root intermediate.crt
[ec2-user certs]$ sudo chmod 644 intermediate.crt
[ec2-user certs]$ ls -al intermediate.crt
```

Ces commandes devraient générer le résultat suivant.

```
-rw-r--r-- root root intermediate.crt
```

6. Placez la clé privée que vous avez utilisée pour créer la CSR dans le répertoire `/etc/pki/tls/private/`.

   Note

   Il existe plusieurs méthodes pour charger votre clé personnalisée dans votre instance EC2, mais le moyen le plus simple et informatif consiste à ouvrir un éditeur de texte (vi, nano, Bloc-notes, etc.) sur votre ordinateur local et votre instance, puis à copier et coller le contenu du fichier. Pour effectuer ces opérations sur l'instance EC2, vous devez disposer de privilèges racine [sudo]. Vous voyez ainsi immédiatement s'il existe des problèmes d'autorisation ou de chemin d'accès. Veillez toutefois à ne pas d'ajouter des lignes supplémentaires lors de la copie du contenu, ou à les modifier de quelque façon.

   À partir du répertoire `/etc/pki/tls/private`, utilisez les commandes suivantes pour vérifier que les paramètres de propriété du fichier, de groupe et d'autorisation correspondent aux valeurs par défaut Amazon Linux 2 très restrictives (propriétaire=racine, groupe=racine, lecture/écriture pour propriétaire uniquement).

```
[ec2-user private]$ sudo chown root:root custom.key
[ec2-user private]$ sudo chmod 600 custom.key
[ec2-user private]$ ls -al custom.key
```

Ces commandes devraient générer le résultat suivant.

```
-rw------- root root custom.key
```

7. Modifiez `/etc/httpd/conf.d/ssl.conf` pour refléter les nouveaux fichiers de certificat et de clé.

   a. Fournissez le chemin et le nom de fichier du certificat d'hôte signé par une CA dans la directive `SSLCertificateFile` d'Apache :

   ```
   SSLCertificateFile /etc/pki/tls/certs/custom.crt
   ```

   b. Si vous avez reçu un fichier de certificat intermédiaire (`intermediate.crt` dans cet exemple), indiquez son nom correct de chemin et de fichier à l'aide de la directive `SSLCACertificateFile` d'Apache :

   ```
   SSLCACertificateFile /etc/pki/tls/certs/intermediate.crt
   ```

   > **Note**
   >
   > Certaines autorités de certification combinent le certificat d'hôte et les certificats intermédiaires dans un seul fichier ; la directive `SSLCACertificateFile` devient alors inutile. Consultez les instructions fournies par votre autorité de certification.

   c. Fournissez le chemin et le nom de fichier de la clé privée (`custom.key` dans cet exemple) dans la directive `SSLCertificateKeyFile` d'Apache :

   ```
   SSLCertificateKeyFile /etc/pki/tls/private/custom.key
   ```

8. Enregistrez `/etc/httpd/conf.d/ssl.conf` et redémarrez Apache.

   ```
   [ec2-user ~]$ sudo systemctl restart httpd
   ```

9. Testez votre serveur en saisissant votre nom de domaine dans la barre d'URL de navigateur avec le préfixe `https://`. Votre navigateur doit charger la page de test via HTTPS sans générer d'erreurs.

Étape 3 : Tester et renforcer la configuration de sécurité

Une fois que votre TLS est opérationnel et exposé au public, vous devriez tester son niveau de sécurité. Il est facile de le faire avec des services en ligne comme Qualys SSL Labs qui effectue une analyse gratuite et complète de votre configuration de sécurité. En fonction des résultats, vous pouvez décider de renforcer la configuration de sécurité par défaut en contrôlant les protocoles que vous acceptez, les chiffrements que vous préférez et que vous excluez. Pour plus d'informations, consultez comment Qualys formule ses scores.

> **Important**
>
> Le test concret est essentiel pour la sécurité de votre serveur. Les petites erreurs de configuration peuvent entraîner des failles de sécurité et des pertes de données. Comme les pratiques de sécurité recommandées changent constamment en réponse à la recherche et aux menaces émergentes, des audits de sécurité périodiques sont essentiels pour la bonne administration du serveur.

Sur le site Qualys SSL Labs, saisissez le nom de domaine complet de votre serveur dans le formulaire **`www.example.com`**. Après environ deux minutes, vous recevrez une note (de A à F) pour votre site et

une analyse détaillée des résultats. Le tableau suivant présente le rapport pour un domaine avec des paramètres identiques à la configuration Apache par défaut sur Amazon Linux 2 et avec un certificat Certbot par défaut.

Score général	B
Certificat	100 %
Support du protocole	95 %
Échange de clés	70 %
Force du chiffrement	90 %

Même si l'aperçu montre que la configuration est principalement sûre, le rapport détaillé indique plusieurs problèmes potentiels, répertoriés ici dans l'ordre de gravité :

✗ Le chiffrement RC4 est pris en charge pour être utilisé par certains navigateurs plus anciens. Un chiffrement est le noyau mathématique d'un algorithme de chiffrement. RC4, un chiffrement rapide utilisé pour chiffrer les flux de données TLS, est connu pour avoir plusieurs failles importantes. À moins que vous ayez une très bonne raison de prendre en charge des navigateurs existants, vous devez désactiver cette option.

✗ Les anciennes versions de TLS sont prises en charge. La configuration prend en charge TLS 1.0 (déjà obsolète) et TLS 1.1 (bientôt obsolète). Seul TLS 1.2 est recommandé depuis 2018.

✗ La confidentialité persistante n'est pas entièrement prise en charge. La confidentialité persistante est une fonction des algorithmes qui chiffrent à l'aide de clés de session temporaires (éphémères) issues de la clé privée. Ceci signifie en pratique que les pirates informatiques ne peuvent pas déchiffrer les données HTTPS même s'ils possèdent la clé privée à long terme d'un serveur web.

Pour corriger et prévenir les erreurs de configuration TLS

1. Ouvrez le fichier de configuration `/etc/httpd/conf.d/ssl.conf` dans un éditeur de texte et mettez en commentaire la ligne suivante en saisissant « # » au début de la ligne.

   ```
   #SSLProtocol all -SSLv3
   ```

2. Ajoutez la directive suivante :

   ```
   #SSLProtocol all -SSLv3
   SSLProtocol -SSLv2 -SSLv3 -TLSv1 -TLSv1.1 +TLSv1.2
   ```

   Cette directive désactive explicitement les versions SSL 2 et 3, ainsi que les versions TLS 1.0 et 1.1. Le serveur refuse désormais d'accepter les connexions chiffrées avec des clients utilisant tout sauf TLS 1.2. La formulation des commentaires dans la directive indique plus clairement, à un lecteur humain, ce pour quoi le serveur est configuré.

   Note

   La désactivation des versions TLS 1.0 et 1.1 de cette manière empêche un faible pourcentage de navigateurs web obsolètes d'accéder à votre site.

Pour modifier la liste des chiffrements autorisés

1. Dans le fichier de configuration `/etc/httpd/conf.d/ssl.conf`, recherchez la section avec la directive **`SSLCipherSuite`** et mettez en commentaire la ligne existante en saisissant « # » au début de la ligne.

```
#SSLCipherSuite HIGH:MEDIUM:!aNULL:!MD5
```

2. Spécifiez explicitement des suites de chiffrement et un ordre de chiffrement qui donnent la priorité à la confidentialité persistante et évitent les chiffrements peu sûrs. La directive `SSLCipherSuite` utilisée ici est basée sur la sortie du Mozilla SSL Configuration Generator, qui adapte une configuration TLS au logiciel spécifique s'exécutant sur votre serveur. (Pour plus d'informations, consultez la ressource utile de Mozilla Security/Server Side TLS.) Déterminez d'abord vos versions d'Apache et OpenSSL en utilisant la sortie des commandes suivantes.

```
[ec2-user ~]$ yum list installed | grep httpd

[ec2-user ~]$ yum list installed | grep openssl
```

Par exemple, si l'information renvoyée est Apache 2.4.34 et OpenSSL 1.0.2, nous saisissons cela dans le générateur. Si vous choisissez le modèle de compatibilité « moderne », il crée une directive `SSLCipherSuite` qui applique la sécurité de façon stricte, tout en étant compatible avec la plupart des navigateurs. Si votre logiciel ne prend pas en charge la configuration moderne, vous pouvez mettre à jour le logiciel ou choisir la configuration « intermédiaire ».

```
SSLCipherSuite ECDHE-ECDSA-AES256-GCM-SHA384:ECDHE-RSA-AES256-GCM-SHA384:ECDHE-ECDSA-
CHACHA20-POLY1305:
ECDHE-RSA-CHACHA20-POLY1305:ECDHE-ECDSA-AES128-GCM-SHA256:ECDHE-RSA-AES128-GCM-SHA256:
ECDHE-ECDSA-AES256-SHA384:ECDHE-RSA-AES256-SHA384:ECDHE-ECDSA-AES128-SHA256:ECDHE-RSA-
AES128-SHA256
```

Les chiffrements sélectionnés contiennent ECDHE (une abréviation pour Elliptic Curve Diffie-Hellman Ephemeral) dans leur nom. Le terme ephemeral indique la confidentialité persistante. Comme corollaire, ces chiffrements ne prennent pas en charge RC4.

Nous vous recommandons d'utiliser une liste explicite de chiffrements au lieu de compter sur les valeurs par défaut ou les directives succinctes dont le contenu n'est pas visible.

Copiez la directive générée dans `/etc/httpd/conf.d/ssl.conf`.

> **Note**
>
> Même si la directive est affichée ici sur plusieurs lignes afin d'être plus lisible, elle doit être sur une seule ligne lorsqu'elle est copiée dans `/etc/httpd/conf.d/ssl.conf` avec un point (pas d'espace) entre les noms des chiffrements.

3. En dernier lieu, supprimez la mise en commentaire de la ligne suivante en retirant le « # » au début de la ligne.

```
#SSLHonorCipherOrder on
```

Cette directive force le serveur à préférer les chiffrements de niveau élevé notamment (dans ce cas) ceux qui prennent en charge la confidentialité persistante. Avec cette directive activée, le serveur essaie d'établir une connexion très sécurisée avant d'avoir recours aux chiffrements autorisés dotés d'une sécurité moindre.

Après avoir terminé ces deux procédures, enregistrez les modifications dans `/etc/httpd/conf.d/ssl.conf` et redémarrez Apache.

Si vous testez de nouveau le domaine sur Qualys SSL Labs, vous devriez voir que la vulnérabilité RC4 et les autres avertissements ont été supprimés et que le résumé ressemble à ce qui suit.

Score général	A
Certificat	100 %
Support du protocole	100 %
Échange de clés	90 %
Force du chiffrement	90 %

> **Important**
>
> Chaque mise à jour d'OpenSSL présente de nouveaux chiffrements et supprime le support des anciens. Gardez à jour votre instance EC2 Amazon Linux 2, surveillez les annonces de sécurité d'OpenSSL et restez attentif aux rapports de nouvelles attaques de la sécurité dans la presse technique. Pour plus d'informations, consultez la section Predefined SSL Security Policies for Elastic Load Balancing du manuel Guide de l'utilisateur pour les Equilibreurs de charge classiques.

Dépannage

- Mon serveur web Apache ne démarre pas si je ne fournis pas un mot de passe

  Il s'agit du comportement attendu si vous avez installé une clé de serveur privée chiffrée et protégée par mot de passe.

  Vous pouvez supprimer l'obligation de chiffrement et de mot de passe de la clé. En supposant que vous disposez d'une clé RSA privée chiffrée nommée `custom.key` dans le répertoire par défaut et que le mot de passe de celle-ci est **`abcde12345`**, exécutez les commandes suivantes sur votre instance EC2 pour générer une version non chiffrée de la clé.

  ```
  [ec2-user ~]$ cd /etc/pki/tls/private/
  [ec2-user private]$ sudo cp custom.key custom.key.bak
  [ec2-user private]$ sudo openssl rsa -in custom.key -passin pass:abcde12345 -out
   custom.key.nocrypt
  [ec2-user private]$ sudo mv custom.key.nocrypt custom.key
  [ec2-user private]$ sudo chown root:root custom.key
  [ec2-user private]$ sudo chmod 600 custom.key
  [ec2-user private]$ sudo systemctl restart httpd
  ```

  Apache devrait maintenant démarrer sans vous demander de fournir un mot de passe.
- J'obtiens des erreurs lors de l'exécution **`sudo yum install -y mod_ssl`**.

  Lorsque vous installez les packages requis pour SSL, vous pouvez voir des erreurs similaires à ce qui suit.

  ```
  Error: httpd24-tools conflicts with httpd-tools-2.2.34-1.16.amzn1.x86_64
  Error: httpd24 conflicts with httpd-2.2.34-1.16.amzn1.x86_64
  ```

  Cela signifie généralement que votre instance EC2 n'exécute pas Amazon Linux 2. Ce didacticiel prend uniquement en charge les instances récemment créées à partir d'une AMI Amazon Linux 2 officielle.

Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux 2

L'autorité de certification Let's Encrypt est la pièce maîtresse de l'effort réalisé par l'organisation Electronic Frontier Foundation (EFF) pour chiffrer Internet à 100 %. Conformément à cet objectif, les certificats d'hôte Let's Encrypt sont conçus pour être créés, validés, installés et tenus à jour avec une intervention humaine limitée. Les aspects automatisés de la gestion de certificats sont effectués par un agent logiciel exécuté sur votre serveur web. Après avoir installé et configuré l'agent, il communique de façon sécurisée avec Let's Encrypt et effectue des tâches administratives sur Apache et sur le système de gestion des clés. Ce didacticiel utilise un agent Certbot gratuit, car il vous permet soit de fournir une clé de chiffrement personnalisée comme base pour vos certificats, soit d'autoriser l'agent lui-même à créer une clé basée sur ses valeurs par défaut. Vous pouvez également configurer Certbot pour renouveler vos certificats de façon régulière sans intervention humaine, comme décrit dans Pour automatiser Certbot (p. 84). Pour plus d'informations, consultez le guide de l'utilisateur Certbot et la documentation Man.

Même si Certbot n'est pas officiellement pris en charge sur Amazon Linux 2, il est disponible en téléchargement et fonctionne correctement une fois installé. Nous vous recommandons d'effectuer les sauvegardes suivantes pour protéger vos données et éviter tout incident :

- Avant de commencer, prenez un instantané de votre volume racine Amazon EBS. Vous pouvez ainsi restaurer l'état d'origine de votre instance EC2. Pour plus d'informations sur la création d'instantanés EBS, consultez Création d'instantanés Amazon EBS (p. 1132).
- La procédure ci-dessous nécessite la modification du fichier `httpd.conf`, qui contrôle le fonctionnement d'Apache. Certbot lui apporte automatiquement ses propres modifications, ainsi qu'à d'autres fichiers de configuration. Effectuez une copie de sauvegarde de l'ensemble de votre répertoire `/etc/httpd` au cas où vous auriez besoin de le restaurer.

Préparation de l'installation

Exécutez les procédures suivantes avant d'installer Certbot.

1. Téléchargez les packages de référentiels EPEL (Extra Packages for Enterprise Linux) 7. Ils sont requis pour fournir les dépendances requises par Certbot.

   a. Accédez à votre répertoire de base (`/home/ec2-user`). Téléchargez le kit EPEL avec la commande suivante.

   ```
   [ec2-user ~]$ sudo wget -r --no-parent -A 'epel-release-*.rpm' http://
   dl.fedoraproject.org/pub/epel/7/x86_64/Packages/e/
   ```

   b. Installez les packages de référentiel comme illustré dans la commande suivante.

   ```
   [ec2-user ~]$ sudo rpm -Uvh dl.fedoraproject.org/pub/epel/7/x86_64/Packages/e/epel-
   release-*.rpm
   ```

   c. Activez EPEL comme dans la commande suivante.

   ```
   [ec2-user ~]$ sudo yum-config-manager --enable epel*
   ```

   Vous pouvez vérifier qu'EPEL est activé avec la commande suivante. Les informations renvoyées devraient être semblables à ce qui suit.

   ```
   [ec2-user ~]$ sudo yum repolist all
   ```

```
...
epel/x86_64                          Extra Packages for Enterprise Linux 7 - x86_64
                                 enabled: 12949+175
epel-debuginfo/x86_64                Extra Packages for Enterprise Linux 7 - x86_64
 - Debug                         enabled:      2890
epel-source/x86_64                   Extra Packages for Enterprise Linux 7 - x86_64
 - Source                        enabled:         0
epel-testing/x86_64                  Extra Packages for Enterprise Linux 7 -
 Testing - x86_64                     enabled:    778+12
epel-testing-debuginfo/x86_64        Extra Packages for Enterprise Linux 7 -
 Testing - x86_64 - Debug             enabled:       107
epel-testing-source/x86_64           Extra Packages for Enterprise Linux 7 -
 Testing - x86_64 - Source            enabled:         0
...
```

2. Modifiez le fichier de configuration principal Apache, `/etc/httpd/conf/httpd.conf`. Localisez la directive « `Listen 80` » et ajoutez les lignes suivantes après cette dernière, en remplaçant les exemples de noms de domaine par le nom commun et le nom SAN (Subject Alternative Name) réels.

```
<VirtualHost *:80>
    DocumentRoot "/var/www/html"
    ServerName "example.com"
    ServerAlias "www.example.com"
</VirtualHost>
```

Enregistrez le fichier et redémarrez Apache.

```
[ec2-user ~]$ sudo systemctl restart httpd
```

Installation et exécution de Certbot

Cette procédure est basée sur la documentation d'EFF pour installer Certbot sur Fedora et RHEL 7. Elle décrit l'utilisation par défaut de Certbot, générant un certificat basé sur une clé RSA 2 048 bits.

1. Installez les dépendances et les packages Certbot à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo yum install -y certbot python2-certbot-apache
```

2. Exécutez Certbot.

```
[ec2-user ~]$ sudo certbot
```

3. A l'invite « Enter email address (used for urgent renewal and security notices) », saisissez l'adresse d'un contact et appuyez sur Entrée.
4. A l'invite, acceptez les conditions de service Let's Encrypt. Saisissez « A » et appuyez sur Entrée pour poursuivre.

```
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Please read the Terms of Service at
https://letsencrypt.org/documents/LE-SA-v1.2-November-15-2017.pdf. You must
agree in order to register with the ACME server at
https://acme-v02.api.letsencrypt.org/directory
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
(A)gree/(C)ancel: A
```

5. À l'autorisation EFF vous permettant de vous ajouter à la liste de diffusion, saisissez « Y » ou « N », puis appuyez sur Entrée.

6. Certbot affiche le nom commun et le nom SAN (Subject Alternative Name) que vous avez fournis dans le bloc VirtualHost.

```
Which names would you like to activate HTTPS for?
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
1: example.com
2: www.example.com
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Select the appropriate numbers separated by commas and/or spaces, or leave input
blank to select all options shown (Enter 'c' to cancel):
```

Laissez le champ vide et appuyez sur Entrée.

7. Certbot affiche le résultat suivant lorsqu'il crée des certificats et configure Apache. Il vous invite ensuite à réacheminer les requêtes HTTP vers HTTPS.

```
Obtaining a new certificate
Performing the following challenges:
http-01 challenge for example.com
http-01 challenge for www.example.com
Waiting for verification...
Cleaning up challenges
Created an SSL vhost at /etc/httpd/conf/httpd-le-ssl.conf
Deploying Certificate for example.com to VirtualHost /etc/httpd/conf/httpd-le-ssl.conf
Enabling site /etc/httpd/conf/httpd-le-ssl.conf by adding Include to root configuration
Deploying Certificate for www.example.com to VirtualHost /etc/httpd/conf/httpd-le-
ssl.conf

Please choose whether or not to redirect HTTP traffic to HTTPS, removing HTTP access.
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
1: No redirect - Make no further changes to the webserver configuration.
2: Redirect - Make all requests redirect to secure HTTPS access. Choose this for
new sites, or if you're confident your site works on HTTPS. You can undo this
change by editing your web server's configuration.
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -
Select the appropriate number [1-2] then [enter] (press 'c' to cancel):
```

Si vous souhaitez autoriser les visiteurs à se connecter à votre serveur via un protocole HTTP non chiffré, saisissez « 1 ». Si vous souhaitez accepter uniquement les connexions chiffrées via HTTPS, saisissez « 2 ». Appuyez sur Entrée pour soumettre votre choix.

8. Certbot termine la configuration d'Apache et signale un succès et d'autres informations.

```
Congratulations! You have successfully enabled https://example.com and
https://www.example.com

You should test your configuration at:
https://www.ssllabs.com/ssltest/analyze.html?d=example.com
https://www.ssllabs.com/ssltest/analyze.html?d=www.example.com
- - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - - -

IMPORTANT NOTES:
 - Congratulations! Your certificate and chain have been saved at:
   /etc/letsencrypt/live/certbot.oneeyedman.net/fullchain.pem
   Your key file has been saved at:
   /etc/letsencrypt/live/certbot.oneeyedman.net/privkey.pem
   Your cert will expire on 2019-08-01. To obtain a new or tweaked
   version of this certificate in the future, simply run certbot again
   with the "certonly" option. To non-interactively renew *all* of
   your certificates, run "certbot renew"
 - Your account credentials have been saved in your Certbot
   configuration directory at /etc/letsencrypt. You should make a
   secure backup of this folder now. This configuration directory will
```

```
  also contain certificates and private keys obtained by Certbot so
  making regular backups of this folder is ideal.
```

9. Une fois que vous avez terminé l'installation, testez et optimisez la sécurité de votre serveur comme décrit dans Étape 3 : Tester et renforcer la configuration de sécurité (p. 77).

Configuration du renouvellement automatique de certificats

Certbot est conçu pour devenir une partie de votre système de serveur invisible et sans erreur. Par défaut, il génère des certificats d'hôte avec un court délai d'expiration de 90 jours. Si vous n'avez pas configuré votre système pour appeler la commande automatiquement, vous devez réexécuter la commande certbot manuellement avant l'expiration. Cette procédure indique comment automatiser Certbot en configurant une tâche cron.

Pour automatiser Certbot

1. Ouvrez le fichier `/etc/crontab` avec un éditeur de texte tel que vim ou nano, avec sudo. Sinon, utilisez sudo crontab -e.
2. Ajoutez une ligne similaire à la suivante et enregistrez le fichier.

```
39      1,13    *       *       *       root    certbot renew --no-self-upgrade
```

Voici un descriptif de chaque composant :

`39 1,13 * * *`

Planifie l'exécution d'une commande à 1 h 39 et 13 h 39 chaque jour. Les valeurs sélectionnées sont arbitraires, mais les développeurs Certbot suggèrent d'exécuter la commande au moins deux fois par jour. Cela garantit la révocation et le remplacement rapides de tous les certificats compromis.

`root`

Des autorisations racine sont nécessaires pour exécuter la commande.

`certbot renew --no-self-upgrade`

La commande à exécuter. La sous-commande renew provoque la vérification des certificats obtenus et le renouvellement de ceux qui s'approchent de la date d'expiration par Certbot. L'indicateur `--no-self-upgrade` empêche la mise à niveau de Certbot sans votre intervention.

3. Redémarrez le processus cron.

```
[ec2-user ~]$ sudo systemctl restart crond
```

Didacticiel : Configurer SSL/TLS sur Amazon Linux

SSL/TLS (Secure Sockets Layer/Transport Layer Security) crée un canal chiffré entre un serveur web et un client web qui empêche les données en transit d'être écoutées. Ce didacticiel explique comment ajouter manuellement la prise en charge de SSL/TLS sur une instance EC2 avec l'AMI Amazon Linux et le serveur Web Apache. Si vous prévoyez de proposer des services de niveau commercial, AWS Certificate Manager, non présenté ici, est une bonne option.

Pour des raisons historiques, le chiffrement web est communément appelé SSL. Alors que les navigateurs web prennent toujours en charge SSL, son protocole successeur TLS est moins vulnérable en cas d'attaque. L'AMI Amazon Linux désactive la prise en charge, côté serveur, de toutes les versions de SSL

par défaut. Les organismes de normes de sécurité considèrent TLS 1.0 comme peu sûr, et TLS 1.0 et TLS 1.1 sont sur le point d'être déclarés formellement obsolètes par l'IETF. Ce didacticiel contient des conseils pour l'activation de TLS 1.2 exclusivement. (Un protocole TLS 1.3 plus récent existe sous forme d'ébauche, mais il n'est pas pris en charge sur Amazon Linux). Pour plus d'informations sur les normes de chiffrement mises à jour, consultez RFC 7568 et RFC 8446.

Ce didacticiel fait référence au chiffrement Web moderne simplement comme TLS.

> **Important**
>
> Ces procédures sont destinées à une utilisation avec l'AMI Amazon Linux. Si vous essayez de configurer un serveur web LAMP sur une instance d'une distribution différente, certaines procédures de ce didacticiel ne fonctionneront peut-être pas pour vous. Pour plus d'informations sur les serveurs web LAMP sur Ubuntu, consultez la rubrique ApacheMySQLPHP de la documentation proposée par la communauté Ubuntu. Pour plus d'informations sur Red Hat Enterprise Linux, consultez la rubrique Web Servers du portail client.

Table des matières

- Prérequis (p. 85)
- Étape 1 : Activer TLS sur le serveur (p. 85)
- Étape 2 : Obtenir un certificat signé par une autorité de certification (CA) (p. 87)
- Étape 3 : Tester et renforcer la configuration de sécurité (p. 92)
- Dépannage (p. 94)
- Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux (p. 95)

Prérequis

Avant de commencer ce didacticiel, suivez les étapes suivantes :

- Lancez une instance basée sur EBS avec l'AMI Amazon Linux. Pour plus d'informations, consultez Étape 1 : Lancement d'une instance (p. 33).
- Configurez votre groupe de sécurité afin que votre instance puisse accepter des connexions sur les ports TCP suivants :
  - SSH (port 22)
  - HTTP (port 80)
  - HTTPS (port 443)

  Pour plus d'informations, consultez Autorisation du trafic entrant pour vos instances Linux (p. 1044).
- Installez le serveur web Apache. Pour des instructions pas à pas, consultez le Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux. (p. 48) Seuls le package http24 et ses dépendances sont nécessaires. Vous pouvez ignorer les instructions impliquant PHP et MySQL.
- Pour identifier et authentifier les sites web, l'infrastructure à clés publiques (PKI) TLS repose sur le système de noms de domaine (DNS). Pour utiliser votre instance EC2 pour héberger un site web public, vous devez enregistrer un nom de domaine pour votre serveur web ou transférer un nom de domaine existant vers votre hôte Amazon EC2. Plusieurs services d'enregistrement de domaines tiers et d'hébergement DNS sont disponibles pour cela, ou vous pouvez utiliser Amazon Route 53.

Étape 1 : Activer TLS sur le serveur

Cette procédure vous décrit le processus de paramétrage de TLS sur Amazon Linux avec un certificat auto-signé numérique.

Note

Un certificat auto-signé est acceptable dans un environnement de test, mais pas pour les environnements de production. Si vous exposez votre certificat auto-signé sur Internet, les visiteurs de votre site verront s'afficher des messages d'avertissement de sécurité.

Pour activer TLS sur un serveur

1. Connectez-vous à votre instance (p. 34) et confirmez qu'Apache est en cours d'exécution.

```
[ec2-user ~]$ sudo service httpd status
```

Si nécessaire, démarrez Apache.

```
[ec2-user ~]$ sudo service httpd start
```

2. Pour vous assurer que tous vos packages logiciels sont mis à jour, effectuez une mise à jour logicielle rapide sur votre instance. Ce processus peut prendre quelques minutes, mais il est important pour vous assurer que vous disposez des dernières mises à jour de sécurité et des nouveaux correctifs de bogues.

   Note

   L'option `-y` installe les mises à jour sans demander de confirmation. Si vous souhaitez examiner les mises à jour avant l'installation, vous pouvez omettre cette option.

```
[ec2-user ~]$ sudo yum update -y
```

3. Maintenant que votre instance est à jour, ajoutez la prise en charge de TLS en installant le module Apache `mod_ssl` :

```
[ec2-user ~]$ sudo yum install -y mod24_ssl
```

Votre instance dispose désormais des fichiers suivants que vous utilisez pour configurer votre serveur sécurisé et créer un certificat pour les tests :

`/etc/httpd/conf.d/ssl.conf`

Le fichier de configuration de mod_ssl. Il contient des « directives » indiquant à Apache où trouver les clés et les certificats de chiffrement, les versions de protocoles TLS à autoriser et les algorithmes de chiffrement à accepter.

`/etc/pki/tls/private/localhost.key`

Une clé privée RSA 2048 bits générée automatiquement pour votre hôte Amazon EC2. Pendant l'installation, OpenSSL a utilisé cette clé pour générer un certificat d'hôte auto-signé ; vous pouvez également utiliser cette clé pour générer une demande de signature de certificat (CSR) à envoyer à une autorité de certification (CA).

`/etc/pki/tls/certs/localhost.crt`

Un certificat X.509 auto-signé généré automatiquement pour votre serveur hôte. Ce certificat est utile pour vérifier qu'Apache est correctement paramétré pour utiliser TLS.

Les fichiers `.key` et `.crt` sont au format PEM qui est constitué de caractères ASCII codés en Base64, encadrés par des lignes « BEGIN » et « END », comme dans cet exemple abrégé d'un certificat :

```
-----BEGIN CERTIFICATE-----
```

```
MIIEazCCA1OgAwIBAgICWxQwDQYJKoZIhvcNAQELBQAwgbExCzAJBgNVBAYTAi0t
MRIwEAYDVQQIDAlTb211lU3RhdGUxETAPBgNVBAcMCFNvbWVDaXR5MRkwFwYDVQQK
DBBTb211lT3JnYW5pemF0aW9uMR8wHQYDVQQLDBZTb211lT3JnYW5pemF0aW9uYWxV
bml0MRkwFwYDVQQDDBBpcC0xNzItMzEtMjAtMjM2MSQwIgYJKoZIhvcNAQkBFhVy
...
z5rRUE/XzxRLBZOoWZpNWTXJkQ3uFYH6s/sBwtHpKKZMzOvDedREjNKAvk4ws6F0
WanXWehT6FiSZvB4sTEXXJN2jdw8g+sHGnZ8zCOsclknYhHrCVD2vnBlZJKSZvak
3ZazhBxtQSukFMOnWPP2a0DMMFGYUHOd0BQE8sBJxg==
-----END CERTIFICATE-----
```

Les noms et extensions de fichier sont fournis à des fins de commodité et n'ont aucun effet sur la fonction ; vous pouvez appeler un certificat `cert.crt`, `cert.pem` ou tout autre nom de fichier dans la mesure où la directive associée dans le fichier `ssl.conf` utilise le même nom.

> **Note**
>
> Lorsque vous remplacez les fichiers TLS par défaut par vos propres fichiers personnalisés, veillez à ce qu'ils soient au format PEM.

4. Redémarrez Apache.

```
[ec2-user ~]$ sudo service httpd restart
```

5. Votre serveur web Apache devrait maintenant prendre en charge HTTPS (HTTP sécurisé) sur le port 443. Testez-le en tapant l'adresse IP ou le nom de domaine complet de votre instance EC2 dans une barre URL du navigateur avec le préfixe **`https://`**. Étant donné que vous vous connectez à un site avec un certificat d'hôte auto-signé non approuvé, il se peut que votre navigateur affiche une série d'avertissements de sécurité.

   Ignorez-les et poursuivez sur le site. Si la page de test Apache par défaut s'ouvre, cela signifie que vous avez configuré correctement TLS sur votre serveur. Toutes les données transmises entre le navigateur et le serveur sont maintenant chiffrées en toute sécurité.

   Pour éviter aux visiteurs du site d'avoir des avertissements, vous devez obtenir un certificat qui chiffre mais vous authentifie aussi publiquement comme le propriétaire du site.

Étape 2 : Obtenir un certificat signé par une autorité de certification (CA)

Vous pouvez utiliser le processus suivant pour obtenir un certificat signé par une CA :

- Générez une demande de signature de certificat (CSR) à partir d'une clé privée
- Envoyez la demande de signature de certificat (CSR) à une autorité de certification (CA)
- Obtenez un certificat d'hôte signé
- Configurez Apache pour utiliser le certificat

Le chiffrement d'un certificat X.509 TLS auto-signé est identique à celui d'un certificat signé par une autorité de certification. La différence est sociale, pas mathématique. Une autorité de certification promet de valider au minimum la propriété d'un domaine avant de générer un certificat pour un demandeur. Chaque navigateur web contient une liste d'autorités de certification approuvées par le fournisseur de navigateur pour faire cela. Un certificat X.509 se compose surtout d'une clé publique qui correspond à votre clé de serveur privée et d'une signature de l'autorité de certification qui est cryptographiquement reliée à la clé publique. Lorsqu'un navigateur se connecte à un serveur web sur HTTPS, le serveur présente un certificat que le navigateur doit vérifier par rapport à sa liste d'autorités de certification approuvées. Si le signataire est sur la liste ou s'il est accessible via une chaîne de confiance composée d'autres utilisateurs

de confiance, le navigateur négocie un canal de données chiffrées rapide avec le serveur et charge la page.

Les certificats coûtent généralement de l'argent à cause travail impliqué dans la validation des requêtes, donc il est intéressant de comparer les prix. Une liste d'autorités de certification connues est proposée sur dmoztools.net. Quelques autorités de certification offrent des certificats basiques gratuits. La plus importante autorité de certification est le projet Let's Encrypt, qui prend également en charge l'automatisation du processus de création et de renouvellement des certificats. Pour de plus amples informations sur l'utilisation de Let's Encrypt comme autorité de certification, consultez Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux (p. 95).

La clé est l'élément sous-jacent du certificat d'hôte. Depuis 2017, des groupes gouvernementaux et industriels recommandent l'utilisation d'une taille de clé minimale (module) de 2048 bits pour les clés RSA conçues pour protéger des documents jusqu'en 2030. La taille de module par défaut générée par OpenSSL dans Amazon Linux est de 2048 bits ce qui signifie que la clé existante générée automatiquement convient à l'utilisation d'un certificat signé par une CA. Une autre procédure est décrite ci-dessous pour ceux qui désirent une clé personnalisée, par exemple, une avec un module plus important ou utilisant un algorithme de chiffrement différent.

Ces instructions pour l'acquisition de certificats d'hôte signés par l'autorité de certification (CA) ne fonctionnent pas à moins que vous possédiez un domaine DNS enregistré et hébergé.

Pour obtenir un certificat signé par une CA

1. Connectez-vous à votre instance (p. 34) et accédez à /etc/pki/tls/private/. Il s'agit du répertoire où la clé privée du serveur de SSL/TLS est stockée. Si vous préférez utiliser votre clé d'hôte existante pour générez la CSR, passez à l'étape 3.
2. (Facultatif) Générez une nouvelle clé privée. Voici quelques exemples de configurations de clés. Toutes les clés obtenues fonctionnent avec votre serveur web, mais elles diffèrent dans leur façon de mettre en œuvre la sécurité (et en ce qui concerne le niveau de sécurité assuré).

   - Exemple 1 : création d'une clé d'hôte RSA par défaut. Le fichier obtenu, **`custom.key`**, est une clé privée RSA 2048 bits.

     ```
     [ec2-user ~]$ sudo openssl genrsa -out custom.key
     ```

   - Exemple 2 : création d'une clé RSA plus forte avec un modulus plus grand. Le fichier obtenu, **`custom.key`**, est une clé privée RSA 4096 bits.

     ```
     [ec2-user ~]$ sudo openssl genrsa -out custom.key 4096
     ```

   - Exemple 3 : création d'une clé RSA chiffrée 4096 bits avec protection par mot de passe. Le fichier obtenu, **`custom.key`**, est une clé privée RSA 4096 bits chiffrée avec le chiffrement AES-128.

     > **Important**
     >
     > Le chiffrement de la clé offre une plus grande sécurité, mais comme une clé chiffrée nécessite un mot de passe, les services qui en dépendent ne peuvent pas démarrer automatiquement. A chaque fois que vous utilisez cette clé, vous devez fournir le mot de passe (« abcde12345 » dans l'exemple précédent) sur une connexion SSH.

     ```
     [ec2-user ~]$ sudo openssl genrsa -aes128 -passout pass:abcde12345 -out custom.key
      4096
     ```

   - Exemple 4 : création d'une clé avec un chiffrement non RSA. La cryptographie RSA peut être relativement lente en raison de la taille de ses clés publiques, lesquelles sont basées sur le produit de deux grands nombres premiers. Cependant, il est possible de créer des clés pour TLS qui utilisent des chiffrements non RSA. Les clés basées sur les mathématiques des courbes elliptiques sont plus petites et plus rapides en termes de calcul, tout en offrant un niveau de sécurité équivalent.

```
[ec2-user ~]$ sudo openssl ecparam -name prime256v1 -out custom.key -genkey
```

Le résultat est une clé privée 256 bits à courbes elliptiques utilisant prime256v1, une « courbe nommée » que OpenSSL prend en charge. Sa qualité cryptographique est légèrement plus importante qu'une clé RSA 2048 bits, selon NIST.

> **Note**
>
> Les autorités de certification ne fournissent pas toutes le même niveau de support pour les clés basées sur les courbes elliptiques que pour les clés RSA.

Assurez-vous que la nouvelle clé privée possède un critère de propriété et d'autorisations très restrictif (propriétaire=racine, groupe=racine, lecture/écriture pour propriétaire uniquement). Les commandes seraient les suivantes :

```
[ec2-user ~]$ sudo chown root.root custom.key
[ec2-user ~]$ sudo chmod 600 custom.key
[ec2-user ~]$ ls -al custom.key
```

Les commandes ci-dessus devraient générer le résultat suivant :

```
-rw------- root root custom.key
```

Une fois que vous avez créé et configuré une clé satisfaisante, vous pouvez créer une CSR.

3. Créez une CSR à l'aide de votre clé préférée. L'exemple ci-dessous utilise **`custom.key`** :

```
[ec2-user ~]$ sudo openssl req -new -key custom.key -out csr.pem
```

OpenSSL ouvre une boîte de dialogue et vous invite à compléter les informations affichées dans le tableau ci-dessous. Tous les champs à l'exception de Common Name sont facultatifs pour un certificat d'hôte basique avec validation de domaine.

Nom	Description	Exemple
Nom du pays	Abréviation ISO de deux lettres de votre pays.	US (=Etats-Unis)
Nom de l'état ou de la province	Nom de l'état ou de la province où votre organisation se situe. Ce nom ne peut pas être abrégé.	Washington
Nom de la localité	L'emplacement de votre organisation, comme une ville.	Seattle
Nom de l'organisation	Nom légal complet de votre organisation. N'abrégez pas le nom de votre organisation.	Exemple d'entreprise
Nom de l'unité d'organisation	Informations supplémentaires sur l'organisation, s'il y en a.	Exemple de service
Nom commun	Cette valeur doit exactement correspondre à l'adresse web que les utilisateurs taperont dans un navigateur, selon vous. Il s'agit généralement d'un nom de domaine avec un nom d'hôte ou un alias préfixé sous la forme **`www.example.com`**.	www.exemple.com

Nom	Description	Exemple
	Dans les essais avec un certificat auto-signé et aucune résolution DNS, le nom commun peut se composer uniquement du nom d'hôte. Les autorités de certification proposent aussi des certificats onéreux qui acceptent les noms inconnus comme **`*.example.com`**.	
Adresse e-mail	L'adresse e-mail de l'administrateur du serveur.	quelquun@exemple.com

Au final, OpenSSL vous invite à donner un mot de passe de stimulation facultatif. Ce mot de passe s'applique uniquement à la CSR et aux transactions entre vous et votre autorité de certification, donc suivez les recommandations de l'autorité de certification sur cela, l'autre champ facultatif et le nom de l'entreprise facultatif. Le mot de passe de stimulation de la CSR n'a aucun effet sur le fonctionnement du serveur.

Le fichier obtenu **`csr.pem`** contient votre clé publique, la signature numérique de votre clé publique et les métadonnées que vous avez saisies.

4. Envoyez la CSR à une autorité de certification. Elle consiste généralement en l'ouverture de votre fichier CSR dans un éditeur de texte et la reproduction du contenu dans un formulaire web. A ce moment-là, il se peut que l'on vous demande de fournir un SAN (subject alternate name) ou plus à placer sur le certificat. Si **`www.example.com`** est le nom commun, **`example.com`** serait un bon SAN, et vice versa. Un visiteur de votre site qui tape l'un de ces noms devrait bénéficier d'une connexion sans erreur. Si le formulaire web de votre autorité de certification le permet, incluez le nom commun dans la liste des SAN. Certaines autorités de certification l'incluent automatiquement.

   Une fois que votre demande a été approuvée, vous recevrez un nouveau certificat d'hôte signé par l'autorité de certification. Il se peut que l'on vous demande également de télécharger un fichier de certificat intermédiaire qui contient des certificats supplémentaires nécessaires pour compléter la chaîne de confiance de l'autorité de certification.

   > **Note**
   >
   > Votre autorité de certification peut vous envoyer des fichiers sous différents formats en fonction des finalités recherchées. Dans ce didacticiel, vous devez utiliser uniquement un fichier de certificat au format PEM qui comporte habituellement (mais pas toujours) une extension `.pem` ou `.crt`. Si vous ne savez pas quel fichier utiliser, ouvrez les fichiers dans un éditeur de texte et recherchez celui qui contient un ou plusieurs blocs commençant par :
   >
   > ```
   > - - - - -BEGIN CERTIFICATE - - - - -
   > ```
   >
   > Le fichier doit également se terminer par :
   >
   > ```
   > - - - -END CERTIFICATE - - - - -
   > ```
   >
   > Vous pouvez également tester un fichier dans la ligne de commande, comme suit :
   >
   > ```
   > [ec2-user certs]$ openssl x509 -in certificate.crt -text
   > ```
   >
   > Vérifiez que ces lignes apparaissent dans le fichier. N'utilisez pas de fichiers se terminant par `.p7b`, `.p7c` ou autres extensions similaires.

5. Placez le nouveau certificat signé par une CA et les certificats intermédiaires dans le répertoire `/etc/pki/tls/certs`.

Note

Il existe plusieurs méthodes pour charger votre clé personnalisée dans votre instance EC2, mais le moyen le plus simple et informatif consiste à ouvrir un éditeur de texte (vi, nano, Bloc-notes, etc.) sur votre ordinateur local et votre instance, puis à copier et coller le contenu du fichier. Pour effectuer ces opérations sur l'instance EC2, vous devez disposer de privilèges racine [sudo]. Vous voyez ainsi immédiatement s'il existe des problèmes d'autorisation ou de chemin d'accès. Veillez toutefois à ne pas d'ajouter des lignes supplémentaires lors de la copie du contenu, ou à les modifier de quelque façon.

À partir du répertoire `/etc/pki/tls/certs`, utilisez les commandes suivantes pour vérifier que les paramètres de propriété du fichier, de groupe et d'autorisation correspondent aux valeurs par défaut Amazon Linux très restrictives (propriétaire=racine, groupe=racine, lecture/écriture pour propriétaire uniquement).

```
[ec2-user certs]$ sudo chown root.root custom.crt
[ec2-user certs]$ sudo chmod 600 custom.crt
[ec2-user certs]$ ls -al custom.crt
```

Les commandes ci-dessus devraient générer le résultat suivant :

```
-rw------- root root custom.crt
```

Les autorisations pour le fichier de certificat intermédiaire sont moins contraignantes (propriétaire=racine, groupe=racine, le propriétaire peut écrire, le groupe peut lire, tout le monde peut lire). Les commandes seraient :

```
[ec2-user certs]$ sudo chown root.root intermediate.crt
[ec2-user certs]$ sudo chmod 644 intermediate.crt
[ec2-user certs]$ ls -al intermediate.crt
```

Les commandes ci-dessus devraient générer le résultat suivant :

```
-rw-r--r-- root root intermediate.crt
```

6. Si vous avez utilisé une clé personnalisée pour créer votre CSR et le certificat d'hôte correspondant, supprimez ou renommez l'ancienne clé du répertoire `/etc/pki/tls/private/`, puis ajoutez-lui la nouvelle clé.

   Note

   Il existe plusieurs méthodes pour charger votre clé personnalisée dans votre instance EC2, mais le moyen le plus simple et informatif consiste à ouvrir un éditeur de texte (vi, nano, Bloc-notes, etc.) sur votre ordinateur local et votre instance, puis à copier et coller le contenu du fichier. Pour effectuer ces opérations sur l'instance EC2, vous devez disposer de privilèges racine [sudo]. Vous voyez ainsi immédiatement s'il existe des problèmes d'autorisation ou de chemin d'accès. Veillez toutefois à ne pas d'ajouter des lignes supplémentaires lors de la copie du contenu, ou à les modifier de quelque façon.

   À partir du répertoire `/etc/pki/tls/private`, vérifiez que les paramètres de propriété du fichier, de groupe et d'autorisation correspondent aux valeurs par défaut Amazon Linux très restrictives (propriétaire=racine, groupe=racine, lecture/écriture pour propriétaire uniquement). Les commandes seraient les suivantes :

```
[ec2-user private]$ sudo chown root.root custom.key
[ec2-user private]$ sudo chmod 600 custom.key
```

```
[ec2-user private]$ ls -al custom.key
```

Les commandes ci-dessus devraient générer le résultat suivant :

```
-rw------- root root custom.key
```

7. Modifiez `/etc/httpd/conf.d/ssl.conf` pour refléter les nouveaux fichiers de certificat et de clé.

   a. Fournissez le chemin et le nom de fichier du certificat d'hôte signé par une CA dans la directive `SSLCertificateFile` d'Apache :

   ```
   SSLCertificateFile /etc/pki/tls/certs/custom.crt
   ```

   b. Si vous avez reçu un fichier de certificat intermédiaire (`intermediate.crt` dans cet exemple), indiquez son nom correct de chemin et de fichier à l'aide de la directive `SSLCACertificateFile` d'Apache :

   ```
   SSLCACertificateFile /etc/pki/tls/certs/intermediate.crt
   ```

   > **Note**
   >
   > Certaines autorités de certification combinent le certificat d'hôte et les certificats intermédiaires dans un seul fichier ; cette directive devient alors inutile. Consultez les instructions fournies par votre autorité de certification.

   c. Fournissez le chemin et le nom de fichier de la clé privée dans la directive `SSLCertificateKeyFile` d'Apache :

   ```
   SSLCertificateKeyFile /etc/pki/tls/private/custom.key
   ```

8. Enregistrez `/etc/httpd/conf.d/ssl.conf` et redémarrez Apache.

```
[ec2-user ~]$ sudo service httpd restart
```

9. Testez votre serveur en saisissant votre nom de domaine dans la barre d'URL de navigateur avec le préfixe `https://`. Votre navigateur doit charger la page de test via HTTPS sans générer d'erreurs.

Étape 3 : Tester et renforcer la configuration de sécurité

Une fois que votre TLS est opérationnel et exposé au public, vous devriez tester son niveau de sécurité. Il est facile de le faire avec des services en ligne comme Qualys SSL Labs qui effectue une analyse gratuite et complète de votre configuration de sécurité. En fonction des résultats, vous pouvez décider de renforcer la configuration de sécurité par défaut en contrôlant les protocoles que vous acceptez, les chiffrements que vous préférez et que vous excluez. Pour plus d'informations, consultez comment Qualys formule ses scores.

> **Important**
>
> Le test concret est essentiel pour la sécurité de votre serveur. Les petites erreurs de configuration peuvent entraîner des failles de sécurité et des pertes de données. Comme les pratiques de sécurité recommandées changent constamment en réponse à la recherche et aux menaces émergentes, des audits de sécurité périodiques sont essentiels pour la bonne administration du serveur.

Sur le site Qualys SSL Labs, tapez le nom de domaine complet de votre serveur dans le formulaire **`www.example.com`**. Après environ deux minutes, vous recevrez une note (de A à F) pour votre site et une analyse détaillée des résultats. Même si l'aperçu montre que la configuration est principalement sûre, le rapport détaillé indique plusieurs problèmes potentiels. Exemples :

✗ Le chiffrement RC4 est pris en charge pour être utilisé par certains navigateurs plus anciens. Un chiffrement est le noyau mathématique d'un algorithme de chiffrement. RC4, un chiffrement rapide utilisé pour chiffrer les flux de données TLS, est connu pour avoir plusieurs failles importantes. À moins que vous ayez une très bonne raison de prendre en charge des navigateurs existants, vous devez désactiver cette option.

✗ Les anciennes versions de TLS sont prises en charge. La configuration prend en charge TLS 1.0 (déjà obsolète) et TLS 1.1 (bientôt obsolète). Seul TLS 1.2 est recommandé depuis 2018.

Pour corriger la configuration de TLS

1. Ouvrez le fichier de configuration `/etc/httpd/conf.d/ssl.conf` dans un éditeur de texte et mettez en commentaire les lignes qui suivent en saisissant « # » au début de chacune d'entre elles :

```
#SSLProtocol all -SSLv3
#SSLProxyProtocol all -SSLv3
```

2. Ajoutez les directives suivantes :

```
SSLProtocol -SSLv2 -SSLv3 -TLSv1 -TLSv1.1 +TLSv1.2
SSLProxyProtocol -SSLv2 -SSLv3 -TLSv1 -TLSv1.1 +TLSv1.2
```

   Ces directives désactivent explicitement les versions SSL 2 et 3, ainsi que les versions TLS 1.0 et 1.1. Le serveur refuse désormais d'accepter les connexions chiffrées avec des clients utilisant tout sauf TLS 1.2. La formulation des commentaires dans la directive communique plus clairement, à un lecteur humain, ce pour quoi le serveur est configuré.

   > **Note**
   >
   > La désactivation des versions TLS 1.0 et 1.1 de cette manière empêche un faible pourcentage de navigateurs web obsolètes d'accéder à votre site.

Pour modifier la liste des chiffrements autorisés

1. Ouvrez le fichier de configuration `/etc/httpd/conf.d/ssl.conf` et trouvez la section avec les exemples mis en commentaire pour la configuration de **`SSLCipherSuite`** et **`SSLProxyCipherSuite`**.

```
#SSLCipherSuite HIGH:MEDIUM:!aNULL:!MD5
#SSLProxyCipherSuite HIGH:MEDIUM:!aNULL:!MD5
```

   Laissez-les tels quels et ajoutez en dessous les directives suivants :

   > **Note**
   >
   > Même si elle est affichée ici sur plusieurs lignes pour plus de lisibilité, chacune de ces deux directives doit être sur une seule ligne sans espaces entre les noms des chiffrements.

```
SSLCipherSuite ECDHE-ECDSA-AES256-GCM-SHA384:ECDHE-RSA-AES256-GCM-SHA384:ECDHE-ECDSA-
CHACHA20-POLY1305:
ECDHE-RSA-CHACHA20-POLY1305:ECDHE-ECDSA-AES128-GCM-SHA256:ECDHE-RSA-AES128-GCM-
SHA256:ECDHE-ECDSA-AES256-SHA384:
```

```
ECDHE-RSA-AES256-SHA384:ECDHE-ECDSA-AES128-SHA256:ECDHE-RSA-AES128-SHA256:AES:!aNULL:!
eNULL:!EXPORT:!DES:
!RC4:!MD5:!PSK:!aECDH:!EDH-DSS-DES-CBC3-SHA:!EDH-RSA-DES-CBC3-SHA:!KRB5-DES-CBC3-SHA

SSLProxyCipherSuite ECDHE-ECDSA-AES256-GCM-SHA384:ECDHE-RSA-AES256-GCM-SHA384:ECDHE-
ECDSA-CHACHA20-POLY1305:
ECDHE-RSA-CHACHA20-POLY1305:ECDHE-ECDSA-AES128-GCM-SHA256:ECDHE-RSA-AES128-GCM-
SHA256:ECDHE-ECDSA-AES256-SHA384:
ECDHE-RSA-AES256-SHA384:ECDHE-ECDSA-AES128-SHA256:ECDHE-RSA-AES128-SHA256:AES:!aNULL:!
eNULL:!EXPORT:!DES:
!RC4:!MD5:!PSK:!aECDH:!EDH-DSS-DES-CBC3-SHA:!EDH-RSA-DES-CBC3-SHA:!KRB5-DES-CBC3-SHA
```

Ces chiffrements constituent un sous-ensemble de la liste beaucoup plus longue des chiffrements pris en charge dans OpenSSL. Ils sont été sélectionnés et classés selon les critères suivants :

- Prise en charge de la confidentialité persistante
- Force
- Rapidité
- Chiffrements spécifiques avant les familles de chiffrements
- Chiffrements autorisés avant les chiffrements refusés

Notez que les chiffrements de niveau élevé possèdent ECDHE dans leur nom, pour Elliptic Curve Diffie-Hellman Ephemeral ; le terme éphémère indique la confidentialité persistante. Par ailleurs, RC4 compte maintenant parmi les chiffrements interdits vers la fin.

Nous vous recommandons d'utiliser une liste explicite de chiffrements au lieu de compter sur les valeurs par défaut ou les directives succinctes dont le contenu n'est pas visible.

> **Important**
>
> La liste des chiffrements affichée compte parmi les nombreuses listes possibles. Par exemple, il se peut que vous vouliez optimiser une liste pour la rapidité plutôt que la confidentialité persistante.
> Si vous anticipez un besoin de prise en charge des clients plus anciens, vous pouvez autoriser la suite de chiffrement DES-CBC3-SHA.
> Enfin, chaque mise à jour de OpenSSL présente de nouveaux chiffrements et dévalorise les anciens. Gardez à jour votre instance EC2 Amazon Linux, surveillez les annonces de sécurité d'OpenSSL et restez attentif aux rapports de nouvelles attaques de la sécurité dans la presse technique. Pour plus d'informations, consultez la section Predefined SSL Security Policies for Elastic Load Balancing du manuel Guide de l'utilisateur pour les Equilibreurs de charge classiques.

2. Supprimez la mise en commentaire de la ligne suivante en retirant le « # » :

```
#SSLHonorCipherOrder on
```

   Cette commande force le serveur à préférer les chiffrements de niveau élevé notamment (dans ce cas) ceux qui prennent en charge la confidentialité persistante. Avec cette directive activée, le serveur essaie d'établir une connexion très sécurisée avant d'avoir recours aux chiffrements autorisés dotés d'une sécurité moindre.
3. Redémarrez Apache. Si vous testez de nouveau le domaine sur Qualys SSL Labs, vous devriez voir que la vulnérabilité RC4 a été supprimée.

Dépannage

- Mon serveur web Apache ne démarre pas si je ne fournis pas un mot de passe

Il s'agit du comportement attendu si vous avez installé une clé de serveur privée chiffrée et protégée par mot de passe.

Vous pouvez supprimer l'obligation de chiffrement et de mot de passe de la clé. En supposant que vous disposez d'une clé RSA privée chiffrée nommée `custom.key` dans le répertoire par défaut et que le mot de passe de celle-ci est **`abcde12345`**, exécutez les commandes suivantes sur votre instance EC2 pour générer une version non chiffrée de la clé.

```
[ec2-user ~]$ cd /etc/pki/tls/private/
[ec2-user private]$ sudo cp custom.key custom.key.bak
[ec2-user private]$ sudo openssl rsa -in custom.key -passin pass:abcde12345 -out
 custom.key.nocrypt
[ec2-user private]$ sudo mv custom.key.nocrypt custom.key
[ec2-user private]$ sudo chown root.root custom.key
[ec2-user private]$ sudo chmod 600 custom.key
[ec2-user private]$ sudo service httpd restart
```

Apache devrait maintenant démarrer sans vous demander de fournir un mot de passe.

Automatisation de certificat : Utilisation de Let's Encrypt avec Certbot sur Amazon Linux

L'autorité de certification Let's Encrypt est la pièce maîtresse de l'effort de l'organisation Electronic Frontier Foundation (EFF) pour chiffrer Internet à 100 %. Conformément à cet objectif, les certificats d'hôte Let's Encrypt sont conçus pour être créés, validés, installés et tenus à jour avec une intervention humaine limitée. Les aspects automatisés de la gestion de certificats sont effectués par un agent sur le serveur web. Après avoir installé et configuré l'agent, il communique de façon sécurisée avec Let's Encrypt et effectue des tâches administratives sur Apache et sur le système de gestion des clés. Ce didacticiel utilise un agent Certbot gratuit, car il vous permet soit de fournir une clé de chiffrement personnalisée comme base pour vos certificats, soit d'autoriser l'agent lui-même à créer une clé basée sur ses valeurs par défaut. Vous pouvez également configurer Certbot pour renouveler vos certificats de façon régulière sans intervention humaine, comme décrit dans Pour automatiser Certbot (p. 84). Pour plus d'informations, consultez le guide de l'utilisateur Certbot ou la page relative à Man.

Même si Certbot n'est pas officiellement pris en charge sur l'AMI Amazon Linux, il est disponible en téléchargement et fonctionne correctement une fois installé. Nous vous recommandons d'effectuer les sauvegardes suivantes pour protéger vos données et éviter tout incident :

- Avant de commencer, prenez un instantané de votre volume racine Amazon EBS. Vous pouvez ainsi restaurer l'état d'origine de votre instance EC2. Pour plus d'informations sur la création d'instantanés EBS, consultez Création d'instantanés Amazon EBS (p. 1132).
- La procédure ci-dessous nécessite la modification du fichier `httpd.conf`, qui contrôle le fonctionnement d'Apache. Certbot lui apporte automatiquement ses propres modifications, ainsi qu'à d'autres fichiers de configuration. Effectuez une copie de sauvegarde de l'ensemble de votre répertoire `/etc/httpd` au cas où vous auriez besoin de le restaurer.

Pour installer et exécuter Certbot

1. Activez le référentiel EPEL (Extra Packages for Enterprise Linux) à partir du projet Fedora sur votre instance. Les référentiels EPEL doivent être utilisés comme des dépendances lorsque vous exécutez le script d'installation Certbot.

```
[ec2-user ~]$ sudo yum-config-manager --enable epel
```

2. Téléchargez la dernière version de Certbot à partir de EFF sur votre instance EC2 à l'aide de la commande suivante.

```
[ec2-user ~]$ wget https://dl.eff.org/certbot-auto
```

3. Transformez le fichier téléchargé en fichier exécutable.

```
[ec2-user ~]$ chmod a+x certbot-auto
```

4. Exécutez le fichier avec les autorisations racine et l'indicateur `--debug`.

```
[ec2-user ~]$ sudo ./certbot-auto --debug
```

5. A partir de l'invite de commande « Is this ok [y/d/N] », tapez « y » et appuyez sur Entrée.
6. A l'invite « Enter email address (used for urgent renewal and security notices) », entrez l'adresse d'un contact et appuyez sur Entrée.
7. A l'invite, acceptez les conditions de service Let's Encrypt. Saisissez « A » et appuyez sur Entrée pour poursuivre :

```
-------------------------------------------------------------------------------
Please read the Terms of Service at
https://letsencrypt.org/documents/LE-SA-v1.1.1-August-1-2016.pdf. You must agree
in order to register with the ACME server at
https://acme-v01.api.letsencrypt.org/directory
-------------------------------------------------------------------------------
(A)gree/(C)ancel: A
```

8. Le lien d'autorisation EFF vous permet de vous ajouter à la liste de diffusion en indiquant « Y » ou « N », puis en appuyant sur Entrée.
9. A l'invite indiqué ci-dessous, entrez votre Common Name (le nom de votre domaine tel que décrit précédemment) et votre nom SAN (Subject Alternative Name), en séparant les deux noms par un espace ou une virgule, et appuyez sur Entrée. Dans cet exemple, les noms ont été fournis :

```
No names were found in your configuration files. Please enter in your domain
name(s) (comma and/or space separated)  (Enter 'c' to cancel):example.com
 www.example.com
```

10. Dans un système Amazon Linux avec une configuration Apache par défaut, comme dans l'exemple ci-dessous, un message vous demande d'indiquer le premier nom que vous avez fourni. Entrez « 1 » et appuyez sur Entrée.

```
Obtaining a new certificate
Performing the following challenges:
tls-sni-01 challenge for example.com
tls-sni-01 challenge for www.example.com

We were unable to find a vhost with a ServerName or Address of example.com.
Which virtual host would you like to choose?
(note: conf files with multiple vhosts are not yet supported)
-------------------------------------------------------------------------------
1: ssl.conf                       |                       | HTTPS | Enabled
-------------------------------------------------------------------------------
Press 1 [enter] to confirm the selection (press 'c' to cancel): 1
```

11. Certbot vous demande ensuite d'indiquer le deuxième nom. Entrez « 1 » et appuyez sur Entrée.

```
We were unable to find a vhost with a ServerName or Address of www.example.com.
Which virtual host would you like to choose?
(note: conf files with multiple vhosts are not yet supported)
```

```
-------------------------------------------------------------------------------
1: ssl.conf                       |                       | HTTPS | Enabled
-------------------------------------------------------------------------------
Press 1 [enter] to confirm the selection (press 'c' to cancel): 1
```

A ce stade-là, Certbot crée votre clé et une CSR :

```
Waiting for verification...
Cleaning up challenges
Generating key (2048 bits): /etc/letsencrypt/keys/0000_key-certbot.pem
Creating CSR: /etc/letsencrypt/csr/0000_csr-certbot.pem
```

12. Autorisez Certbot à créer votre clé et tous les certificats d'hôte nécessaires. Lorsque vous êtes invité à entrer chaque nom, tapez « 1 » et appuyez sur Entrée comme indiqué dans l'exemple :

```
We were unable to find a vhost with a ServerName or Address of example.com.
Which virtual host would you like to choose?
(note: conf files with multiple vhosts are not yet supported)
-------------------------------------------------------------------------------
1: ssl.conf                       |                       | HTTPS | Enabled
-------------------------------------------------------------------------------
Press 1 [enter] to confirm the selection (press 'c' to cancel): 1
Deploying Certificate for example.com to VirtualHost /etc/httpd/conf.d/ssl.conf

We were unable to find a vhost with a ServerName or Address of www.example.com.
Which virtual host would you like to choose?
(note: conf files with multiple vhosts are not yet supported)
-------------------------------------------------------------------------------
1: ssl.conf                       | example.com        | HTTPS | Enabled
-------------------------------------------------------------------------------
Press 1 [enter] to confirm the selection (press 'c' to cancel): 1
Deploying Certificate for www.example.com to VirtualHost /etc/httpd/conf.d/ssl.conf
```

13. Choisissez d'autoriser les connexions non sécurisées sur votre serveur web. Si vous choisissez l'option 2 (comme dans l'exemple), toutes les connexions vers votre serveur seront soit chiffrées, soit rejetées.

```
Please choose whether HTTPS access is required or optional.
-------------------------------------------------------------------------------
1: Easy - Allow both HTTP and HTTPS access to these sites
2: Secure - Make all requests redirect to secure HTTPS access
-------------------------------------------------------------------------------
Select the appropriate number [1-2] then [enter] (press 'c' to cancel): 2
```

Certbot termine la configuration d'Apache et signale un succès et d'autres informations :

```
Congratulations! You have successfully enabled https://example.com and
https://www.example.com

You should test your configuration at:
https://www.ssllabs.com/ssltest/analyze.html?d=example.com
https://www.ssllabs.com/ssltest/analyze.html?d=www.example.com
-------------------------------------------------------------------------------

IMPORTANT NOTES:
 - Congratulations! Your certificate and chain have been saved at
   /etc/letsencrypt/live/example.com/fullchain.pem. Your cert will
   expire on 2017-07-19. To obtain a new or tweaked version of this
   certificate in the future, simply run certbot-auto again with the
   "certonly" option. To non-interactively renew *all* of your
   certificates, run "certbot-auto renew"
```

```
....
```

14. Une fois que vous avez terminé l'installation, testez et optimisez la sécurité de votre serveur comme décrit dans Étape 3 : Tester et renforcer la configuration de sécurité (p. 77).

Certbot est conçu pour devenir une partie de votre système de serveur invisible et sans erreur. Par défaut, il génère des certificats d'hôte avec un court délai d'expiration de 90 jours. Si vous n'avez pas configuré votre système pour appeler la commande automatiquement, vous devez ré-exécuter la commande certbot manuellement. Cette procédure indique comment automatiser Certbot en configurant une tâche cron.

Pour configurer le renouvellement automatique de certificats

1. Ouvrez le fichier `/etc/crontab` avec un éditeur de texte (tel que vim ou nano) à l'aide de sudo. Sinon, utilisez sudo crontab -e.
2. Ajoutez une ligne similaire à la suivante et enregistrez le fichier.

```
39      1,13    *       *       *       root    certbot renew --no-self-upgrade
```

   Voici un descriptif de chaque composant :

   `39 1,13 * * *`

   Planifie l'exécution d'une commande à 1 h 39 et 13 h 39 chaque jour. Les valeurs sélectionnées sont arbitraires, mais les développeurs Certbot suggèrent d'exécuter la commande au moins deux fois par jour. Cela garantit la révocation et le remplacement rapides de tous les certificats compromis.

   `root`

   Des privilèges racine sont nécessaires pour exécuter la commande.

   `certbot renew --no-self-upgrade`

   La commande à exécuter. La sous-commande renew provoque la vérification des certificats obtenus et le renouvellement de ceux qui s'approchent de la date d'expiration par Certbot. L'indicateur `--no-self-upgrade` empêche la mise à niveau de Certbot sans votre intervention.
3. Redémarrez le processus cron :

```
[ec2-user ~]$ sudo service crond restart
```

Didacticiel : Augmenter la disponibilité des applications sur Amazon EC2

Imaginons que vous commenciez l'exécution de votre application ou site web sur une seule instance EC2 et qu'au fil du temps, le trafic augmente au point que plusieurs instances deviennent nécessaires pour satisfaire la demande. Vous pouvez lancer plusieurs instances EC2 depuis votre AMI, puis utiliser Elastic Load Balancing pour répartir le trafic entrant de votre application entre ces instances EC2. La disponibilité de votre application s'en trouve accrue. Le placement de vos instances dans plusieurs zones de disponibilité améliore aussi la tolérance aux pannes de votre application. Si une zone de disponibilité est confrontée à une panne, le trafic est acheminé vers l'autre zone de disponibilité.

Vous pouvez utiliser Amazon EC2 Auto Scaling pour maintenir en permanence un nombre minimal d'instances pour votre application. Amazon EC2 Auto Scaling peut détecter lorsque votre instance ou application est défectueuse, et la remplacer automatiquement pour préserver sa disponibilité. Vous

pouvez aussi utiliser Amazon EC2 Auto Scaling pour agrandir ou réduire votre capacité Amazon EC2 automatiquement en fonction de la demande, à l'aide de critères que vous spécifiez.

Dans ce didacticiel, nous utilisons Amazon EC2 Auto Scaling avec Elastic Load Balancing pour garantir que vous conserviez un nombre spécifié d'instances EC2 saines derrière votre équilibreur de charge. Notez que ces instances n'ont pas besoin d'adresses IP publiques, parce que le trafic est acheminé d'abord vers l'équilibreur de charge, puis vers les instances. Pour plus d'informations, consultez Amazon EC2 Auto Scaling et Elastic Load Balancing.

Table des matières

- Prérequis (p. 99)
- Dimensionner votre application et en répartir la charge (p. 100)
- Test de votre équilibreur de charge (p. 101)

Prérequis

Le didacticiel présume que vous avez déjà effectué les étapes suivantes :

1. Créez un VPC (Virtual Private Cloud) avec un sous-réseau public dans deux zones de disponibilité ou plus.
2. Lancez une instance dans le VPC.
3. Connectez-vous à l'instance et personnalisez-la. Par exemple, installation des logiciels et des applications, copie des données et attachement des volumes EBS supplémentaires. Pour plus d'informations sur la configuration d'un serveur web sur votre instance, consultez Didacticiel : Installation d'un serveur web LAMP avec l'Amazon Linux AMI (p. 48).
4. Testez votre application sur votre instance pour vous assurer qu'elle est correctement configurée.
5. Créez un Amazon Machine Image (AMI) personnalisé à partir de votre instance. Pour plus d'informations, consultez Création d'une AMI Linux basée sur Amazon EBS (p. 132) ou Création d'une AMI Linux basée sur le stockage d'instance (p. 136).
6. (Facultatif) Mettez fin à l'instance si vous n'en n'avez plus besoin.
7. Créez un rôle IAM qui accorde à votre application l'accès à AWS dont elle a besoin. Pour plus d'informations, consultez Pour créer un rôle IAM avec la console IAM (p. 1038).

Dimensionner votre application et en répartir la charge

Utilisez la procédure suivante pour créer un équilibreur de charge, créer une configuration de lancement pour vos instances, créer un groupe Auto Scaling composé de deux instances ou plus et associer l'équilibreur de charge au groupe Auto Scaling.

Pour dimensionner votre application et en répartir la charge

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous LOAD BALANCING, choisissez Load Balancers.
3. Sélectionnez Create Load Balancer.
4. Pour Application Load Balancer, choisissez Create.
5. Sur la page Configure Load Balancer, procédez de la façon suivante :
   a. Dans Nom, entrez un nom pour votre équilibreur de charge. Par exemple, **`my-lb`**.
   b. Pour Scheme, conservez la valeur par défaut internet-facing.
   c. Pour Listeners, conservez la valeur par défaut, soit un écouteur qui accepte le trafic HTTP sur le port 80.
   d. Pour Zones de disponibilité, sélectionnez le VPC que vous avez utilisé pour vos instances. Sélectionnez une zone de disponibilité puis le sous-réseau public de cette zone de disponibilité. Répétez cette opération pour une seconde zone de disponibilité.
   e. Choisissez Next: Configure Security Settings.
6. Pour ce didacticiel, vous n'utilisez pas un écouteur sécurisé. Choisissez Next: Configure Security Groups.
7. Sur la page Configure Security Groups, procédez de la façon suivante :
   a. Choisissez Create a new security group.
   b. Tapez un nom et une description pour le groupe de sécurité, ou conservez le nom et la description par défaut. Le nouveau groupe de sécurité contient une règle qui autorise le trafic en direction du port configuré pour l'écouteur.
   c. Choisissez Next: Configure Routing.
8. Sur la page Configure Routing, procédez de la façon suivante :
   a. Pour Target group, conservez la valeur par défaut New target group.
   b. Dans Nom, entrez un nom pour le groupe cible.
   c. Conservez HTTP pour Protocol, 80 pour Port et instance pour Target type.
   d. Pour Health checks, conservez le protocole par défaut et le chemin d'accès.
   e. Choisissez Next: Register Targets.
9. Sur la page Register Targets, choisissez Next: Review pour accéder à la page suivante, car nous utiliserons Amazon EC2 Auto Scaling pour ajouter des instances EC2 au groupe cible.
10. Sur la page Review, choisissez Create. Une fois que l'équilibreur de charge est créé, cliquez sur Close.
11. Dans le panneau de navigation, sous AUTO SCALING, sélectionnez Launch Configurations.
    - Si vous utilisez Amazon EC2 Auto Scaling pour la première fois, une page de bienvenue s'affiche. Sélectionnez Create Auto Scaling group pour démarrer l'Assistant Create Auto Scaling Group, puis cliquez sur Create launch configuration.
    - Sinon, choisissez Create launch configuration.
12. Dans Nom de la configuration de lancement, entrez un nom pour votre configuration de lancement (par exemple, **`my-launch-config`**).
13. Pour Image machine Amazon (AMI), sous Mes AMI, choisissez l'AMI que vous avez créée dans Prérequis (p. 99).

14. Pour Type d'instance, utilisez Choisir le type d'instance pour sélectionner un type d'instance.
15. (Facultatif) Pour Configuration supplémentaire, procédez comme suit :
    a. Choisissez Demander des instances ponctuelles. Sinon, les instances sont des instances à la demande.
    b. Pour Profil d'instance IAM, sélectionnez le rôle IAM que vous avez créé dans Prérequis (p. 99).
    c. Choisissez Activer la surveillance détaillée des instances EC2 dans CloudWatch. Sinon, la surveillance de base est activée.
    d. Pour configurer les métadonnées d'instance, développez Détails avancés, activez ou désactivez les métadonnées d'instance et configurez la version et la limite de saut selon les besoins.
    e. Pour exécuter un script de démarrage, développez Détails avancés et entrez le script dans Données utilisateur.
    f. Pour attribuer des adresses IP publiques à vos instances, développez Détails avancés et définissez le Type d'adresse IP selon vos besoins.
16. Pour Stockage (volumes), vous ajoutez des volumes selon vos besoins. Vous pouvez créer des volumes EBS vides ou créer des volumes EBS à partir d'instantanés EBS.
17. Pour Groupes de sécurité, vous pouvez sélectionner un groupe de sécurité existant ou en créer un nouveau. Le groupe de sécurité doit autoriser le trafic HTTP et les vérifications de l'état depuis l'équilibreur de charge. Si vous avez attribué des adresses IP publiques à vos instances, vous pouvez éventuellement autoriser le trafic SSH afin de pouvoir vous y connecter.
18. Pour Paire de clés (connexion), choisissez une paire de clés existante, créez une nouvelle paire de clés ou continuez sans paire de clés. Cochez la case de confirmation.
19. Choisissez Créer une configuration de lancement.
20. Une fois la configuration de lancement créée, vous devez créer un groupe Auto Scaling.
    - Si vous découvrez Amazon EC2 Auto Scaling et que vous utilisez l'Assistant Créer le groupe Auto Scaling, vous passez automatiquement à l'étape suivante.
    - Sinon, sélectionnez le groupe Auto Scaling et choisissez Actions, Créer un groupe Auto Scaling.
21. Dans la page Choisir un modèle de lancement ou une configuration, entrez un nom pour le groupe Auto Scaling. Par exemple, **`my-asg`**. Choisissez Suivant.
22. Sur la page Configurer les paramètres, choisissez votre VPC et vos deux sous-réseaux publics. Choisissez Suivant.
23. Dans la page Configurer les options avancées, sélectionnez Activer l'équilibrage de charge et choisissez votre groupe cible. Sélectionnez le type de vérification de l'état ELB et choisissez Suivant.
24. Pour Group size, saisissez le nombre d'instances (par exemple, **`2`**). Notez qu'il est recommandé de conserver approximativement le même nombre d'instances dans chaque zone de disponibilité.
25. Sur la page Configurer la taille du groupe et les stratégies de mise à l'échelle, vous pouvez configurer la taille du groupe ou configurer ce dernier de façon à ce qu'il évolue dynamiquement en fonction de la demande. Choisissez Suivant.
26. (Facultatif) Choisissez Ajouter des notifications afin de configurer les notifications SNS pour les activités de mise à l'échelle. Choisissez Suivant.
27. (Facultatif) Choisissez Ajouter une balise pour ajouter des balises. Choisissez Suivant.
28. Sur la page Révision, modifiez les détails selon vos besoins, puis choisissez Créer un groupe Auto Scaling.

Test de votre équilibreur de charge

Lorsqu'un client envoie une requête à votre équilibreur de charge, ce dernier achemine la requête vers l'une des instances enregistrées.

Pour tester l'équilibreur de charge

1. Vérifiez que vos instances sont prêtes. Sur la page Auto Scaling Groups, sélectionnez votre groupe Auto Scaling, puis choisissez l'onglet Instances. Initialement, vos instances sont à l'état `Pending`. Lorsque le cycle de vie est `InService`, vos instances sont prêtes à l'emploi.
2. Vérifiez que vos instances sont enregistrées auprès de l'équilibreur de charge. Dans la page Groupes cibles choisissez le nom du groupe cible pour ouvrir sa page de détails, puis choisissez Cibles. Si l'état de vos instances est `initial`, il est possible qu'elles soient toujours en cours d'inscription. Lorsque l'état de vos instances indique `healthy`, elles sont prêtes à être utilisées. Une fois vos instances prêtes, vous pouvez passer à l'étape suivante.
3. Sur la page Load Balancers, sélectionnez votre équilibreur de charge.
4. Sous l'onglet Description, recherchez le nom DNS. Ce nom présente la forme suivante :

```
my-lb-xxxxxxxxxx.us-west-2.elb.amazonaws.com
```

5. Dans un navigateur web, collez le nom DNS de l'équilibreur de charge dans la barre d'adresse et appuyez sur Entrée. Votre site web s'affiche.

Amazon Machine Images (AMI)

Une AMI (Amazon Machine Image) fournit les informations requises pour lancer une instance. Vous devez spécifier une AMI lorsque vous lancez une instance. Lorsque vous avez besoin de plusieurs instances configurées de manière identique, il est possible de lancer plusieurs instances à partir d'une même AMI. Lorsque vous avez besoin d'instances configurées de manière différente, vous pouvez utiliser différentes AMI pour lancer ces instances.

Une AMI comprend les éléments suivants :

- Un ou plusieurs instantanés EBS ou, dans le cas des AMI basées sur le stockage d'instance, un modèle pour le volume racine de l'instance (par exemple, un système d'exploitation, un serveur d'applications et des applications).
- Les autorisations de lancement qui contrôlent les comptes AWS qui peuvent utiliser l'AMI pour lancer les instances.
- Un mappage de périphérique de stockage en mode bloc qui spécifie les volumes à attacher à l'instance lorsqu'elle est lancée.

Rubriques

- Utilisation d'une AMI (p. 103)
- Création de votre propre AMI (p. 104)
- Achat, partage et vente d'AMI (p. 104)
- Annulation de l'inscription de votre AMI (p. 105)
- Amazon Linux 2 et Amazon Linux AMI (p. 105)
- Types d'AMI (p. 105)
- Types de virtualisation AMI Linux (p. 108)
- Rechercher une AMI Linux (p. 110)
- AMI partagées (p. 117)
- AMI payantes (p. 128)
- Création d'une AMI Linux basée sur Amazon EBS (p. 132)
- Création d'une AMI Linux basée sur le stockage d'instance (p. 136)
- Utiliser le chiffrement avec des AMI basées sur EBS (p. 168)
- Copie une AMI (p. 173)
- Obtenir des informations de facturation (p. 179)
- Annulation de l'inscription de votre AMI Linux (p. 182)
- Amazon Linux (p. 186)
- Noyaux fournis par l'utilisateur (p. 205)
- Utilisation de l'environnement de bureau MATE fourni avec Amazon Linux 2 (p. 211)

Utilisation d'une AMI

Le diagramme suivant résume le cycle de vie de l'AMI. Après avoir créé et enregistré une AMI, vous pouvez l'utiliser pour lancer de nouvelles instances. Vous pouvez également lancer des instances depuis

une AMI si son propriétaire vous octroie des autorisations de lancement. Vous pouvez copier une AMI dans la même région ou dans des régions différentes. Lorsque vous n'avez plus besoin d'une AMI, vous pouvez annuler son inscription.

Vous pouvez rechercher une AMI répondant aux critères de votre instance. Vous pouvez rechercher des AMI fournies par AWS ou des AMI fournies par la communauté. Pour de plus amples informations, veuillez consulter Types d'AMI (p. 105) et Recherche d'une AMI.

Une fois que vous avez lancé une instance à partir d'une AMI, vous pouvez vous y connecter. Lorsque vous êtes connecté à une instance, vous pouvez l'utiliser comme vous le feriez avec n'importe quel autre serveur. Pour plus d'informations sur le lancement, la connexion et l'utilisation de votre instance, consultez Instances Amazon EC2 (p. 213).

Création de votre propre AMI

Vous pouvez lancer une instance à partir d'une AMI existante, personnaliser cette instance, puis enregistrer la configuration mise à jour comme AMI personnalisée. Les instances lancées à partir de cette nouvelle AMI incluront les personnalisations apportée lors de sa création.

Le périphérique de stockage racine de l'instance détermine le processus à suivre pour créer une AMI. Le volume racine d'une instance est soit un volume Amazon EBS, soit un volume de stockage d'instance. Pour plus d'informations sur les volumes du périphérique racine, consultez Volume du périphérique racine Amazon EC2 (p. 21).

- Pour créer une AMI basée sur Amazon EBS, consultez Création d'une AMI Linux basée sur Amazon EBS (p. 132).
- Pour créer une AMI basée sur le stockage d'instance, consultez Création d'une AMI Linux basée sur le stockage d'instance (p. 136).

Pour faciliter le classement et la gestion de vos AMI, vous pouvez leur attribuer des balises personnalisées. Pour plus d'informations, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Achat, partage et vente d'AMI

Une fois que vous avez créé une AMI, vous pouvez conserver son statut privé afin d'être la seule personne à pouvoir l'utiliser ou vous pouvez la partager avec une liste spécifiée de comptes AWS. Vous pouvez également rendre votre AMI personnalisée publique afin que la communauté puisse l'utiliser. La création d'une AMI sûre, sécurisée et utilisable à des fins d'utilisation publique est un processus relativement simple,

à condition que vous respectiez quelques consignes simples. Pour plus d'informations sur la création et l'utilisation d'AMI partagées, consultez AMI partagées (p. 117).

Vous pouvez acheter des AMI auprès d'un tiers, notamment si elle est fournie avec des contrats de service proposés par des organisations telles que Red Hat. Vous pouvez également créer une AMI et la vendre à d'autres utilisateurs Amazon EC2. Pour plus d'informations sur la vente ou l'achat d'AMI, consultez AMI payantes (p. 128).

Annulation de l'inscription de votre AMI

Vous pouvez annuler l'inscription de votre AMI lorsque vous avez terminé de l'utiliser. Après cette opération, l'AMI ne peut plus être utilisée pour lancer de nouvelles instances. Les instances existantes lancées à partir de l'AMI ne sont pas affectées. Pour plus d'informations, consultez Annulation de l'inscription de votre AMI Linux (p. 182).

Amazon Linux 2 et Amazon Linux AMI

Amazon Linux 2 et l'Amazon Linux AMI sont des images Linux prises en charge et gérées, fournies par AWS. Voici quelques caractéristiques d'Amazon Linux 2 et de l'Amazon Linux AMI :

- Un environnement d'exécution stable, sécurisé et très performant pour des applications exécutées sur Amazon EC2.
- Fournis sans frais supplémentaires aux utilisateurs d'Amazon EC2.
- Accès du référentiel à plusieurs versions de MySQL, PostgreSQL, Python, Ruby, Tomcat et bien d'autres packages courants.
- Mises à jour régulières incluant les derniers composants et rendues disponibles dans les référentiels yum en vue de leur installation sur les instances en cours d'exécution.
- Inclut des packages qui facilitent l'intégration aux services AWS, par exemple les outils AWS CLI, Amazon EC2 API et AMI, la bibliothèque Boto pour Python et les outils Elastic Load Balancing.

Pour plus d'informations, consultez Amazon Linux (p. 186).

Types d'AMI

Vous pouvez sélectionner une AMI en fonction des caractéristiques suivantes :

- Région (consultez Régions et zones (p. 7))
- Système d'exploitation
- Architecture (32 bits ou 64 bits)
- Autorisations de lancement (p. 105)
- Stockage pour le périphérique racine (p. 106)

Autorisations de lancement

Le propriétaire d'une AMI détermine sa disponibilité en spécifiant les autorisations de lancement. Les autorisations de lancement sont réparties en plusieurs catégories.

Autorisation de lancement	Description
public	Le propriétaire octroie des autorisations de lancement à tous les comptes AWS.
explicite	Le propriétaire octroie des autorisations de lancement à des comptes AWS spécifiques.
implicite	Le propriétaire a des autorisations de lancement implicites pour une AMI.

Amazon et la communauté Amazon EC2 proposent un large éventail d'AMI publiques. Pour plus d'informations, consultez AMI partagées (p. 117). Les développeurs peuvent faire payer leurs AMI. Pour plus d'informations, consultez AMI payantes (p. 128).

Stockage pour le périphérique racine

Toutes les AMI sont réparties en deux catégories : basées sur Amazon EBS ou basées sur le stockage d'instance. La première catégorie signifie que le périphérique racine d'une instance lancée à partir de l'AMI est un volume Amazon EBS créé à partir d'un instantané Amazon EBS. La deuxième catégorie signifie que le périphérique racine d'une instance lancée à partir de l'AMI est un volume de stockage d'instance créé à partir d'un modèle stocké dans Amazon S3. Pour plus d'informations, consultez Volume du périphérique racine Amazon EC2 (p. 21).

Le tableau suivant résume les différences importantes lors de l'utilisation des deux types d'AMI.

Caractéristiques	AMI basée sur des volumes Amazon EBS	AMI basée sur le stockage d'instance Amazon
Temps de démarrage pour une instance	Généralement inférieur à 1 minute	Généralement inférieur à 5 minutes
Limite de taille d'un périphérique racine	16 TiO	10 GiO
volume du périphérique racine	Volume Amazon EBS	Volume de stockage d'instance
Persistance des données	Par défaut, le volume racine est supprimé lorsque l'instance est arrêtée.* Par défaut, les données des autres volumes Amazon EBS sont conservées après la mise hors service de l'instance.	Les données des volumes de stockage d'instance sont conservées uniquement pendant la durée de vie de l'instance.
Modifications	Le type d'instance, le noyau, le disque RAM et les données utilisateur peuvent être modifiés pendant que l'instance est arrêtée.	Les attributs de l'instance restent les mêmes pendant la durée de vie de l'instance.
Frais	Les éléments suivants vous sont facturés : utilisation de l'instance, utilisation du volume Amazon EBS et stockage de votre AMI sous forme d'instantané Amazon EBS.	L'utilisation de l'instance et le stockage de l'AMI dans Amazon S3 vous sont facturés.
Création d'AMI/bundle	Utilise une seule commande/un seul appel	Requiert l'installation et l'utilisation des outils AMI

Caractéristiques	AMI basée sur des volumes Amazon EBS	AMI basée sur le stockage d'instance Amazon
État d'arrêt	Peut être à l'état arrêté. Même lorsque l'instance est arrêtée et ne s'exécute pas, le volume racine est conservé dans Amazon EBS	État d'arrêt impossible, les instances sont en cours d'exécution ou hors service

* Par défaut, les volumes racine d'instances basées sur Amazon EBS ont l'indicateur `DeleteOnTermination` défini sur `true`. Pour plus d'informations sur la modification de cet indicateur afin que le volume soit conservé après la mise hors service, consultez Modification du volume racine pour qu'il persiste (p. 25).

Sélection du type de périphérique racine de votre AMI

Pour déterminer le type de périphérique racine d'une AMI à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, cliquez sur AMIs, puis sélectionnez l'AMI.
3. Vérifiez la valeur de Root Device Type sous l'onglet Details comme suit :
   - Si la valeur est `ebs`, il s'agit d'une AMI basée sur Amazon EBS.
   - Si la valeur est `instance store`, il s'agit d'une AMI basée sur le stockage d'instance.

Pour déterminer le type de périphérique racine d'une AMI à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-images (AWS CLI)
- Get-EC2Image (Outils AWS pour Windows PowerShell)

État d'arrêt

Vous pouvez arrêter une instance basée sur Amazon EBS, mais pas une instance basée sur le stockage d'instance Amazon EC2. Lorsque vous choisissez d'arrêter une instance, vous interrompez son exécution (son statut passe de `running` à `stopping`, puis à `stopped`). Une instance arrêtée est conservée sur Amazon EBS, ce qui lui permet d'être redémarrée. L'arrêt est différent de la mise hors service dans la mesure où vous ne pouvez pas redémarrer une instance qui a été mise hors service. Etant donné que les instances basées sur le stockage d'instance Amazon EC2 ne peuvent pas être arrêtées, elle sont soit en cours d'exécution, soit hors service. Pour plus d'informations sur ce qui se produit et ce que vous pouvez faire lors de l'arrêt d'une instance, consultez Arrêt et démarrage de votre instance (p. 625).

Persistance et stockage de données par défaut

Les instances qui utilisent un volume de stockage d'instance pour le périphérique racine ont un stockage d'instance disponible automatiquement (le volume racine contient la partition racine et vous pouvez stocker des données supplémentaires). Vous pouvez ajouter un stockage persistant à votre instance en attachant un ou plusieurs volumes Amazon EBS. Toute donnée présente sur un volume de stockage d'instance est effacée lorsque l'instance échoue ou qu'elle est mise hors service. Pour plus d'informations, consultez Durée de vie d'un stockage d'instance (p. 1265).

Les instances qui ont recours à Amazon EBS pour le périphérique racine sont automatiquement associées à un volume Amazon EBS. Le volume apparaît dans votre liste de volumes comme tous les autres. Avec

la plupart des types d'instance, les instances basées sur des volumes Amazon EBS n'ont aucun volume de stockage d'instance par défaut. Vous pouvez ajouter des volumes de stockage d'instance ou des volumes Amazon EBS supplémentaires à l'aide d'un mappage de périphérique de stockage en mode bloc. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289).

Temps de démarrage

Les instances lancées depuis une AMI basée sur Amazon EBS se lancent plus rapidement que celles lancées depuis une AMI basée sur le stockage d'instance. Lorsque vous lancez une instance depuis une AMI basée sur le stockage d'instance, toutes les parties doivent être extraites d'Amazon S3 avant que l'instance soit disponible. Dans le cadre d'une AMI basée sur Amazon EBS, seules les parties nécessaires au démarrage de l'instance doivent être extraites de l'instantané avant que l'instance soit disponible. Toutefois, les performances d'une instance qui utilise un volume Amazon EBS pour son périphérique racine sont plus lentes pendant un bref moment, tandis que les parties restantes sont extraites de l'instantané et chargées dans le volume. Lorsque vous arrêtez et redémarrez l'instance, elle est lancée rapidement dans la mesure où l'état est stocké dans un volume Amazon EBS.

Création d'AMI

Pour créer des AMI Linux basées sur le stockage d'instance, vous devez créer une AMI à partir de votre instance sur l'instance elle-même à l'aide des outils AMI Amazon EC2.

La création d'AMI est nettement plus simple pour les AMI basées sur Amazon EBS. L'action d'API `CreateImage` crée votre AMI basée sur Amazon EBS et l'inscrit. Un bouton sur l'AWS Management Console vous permet de créer une AMI à partir d'une instance en cours d'exécution. Pour plus d'informations, consultez Création d'une AMI Linux basée sur Amazon EBS (p. 132).

Facturation

Pour les AMI basées sur le stockage d'instance, le stockage de l'AMI dans Amazon S3 et l'utilisation de l'instance vous sont facturés. Avec les AMI basées sur Amazon EBS, l'utilisation de l'instance, le stockage de volume Amazon EBS et son utilisation, et le stockage de votre AMI comme instantané Amazon EBS vous sont facturés.

Avec les AMI basées sur le stockage d'instance Amazon EC2, chaque fois que vous personnalisez une AMI et que vous en créez une nouvelle, toutes les parties sont stockées dans Amazon S3 pour chaque AMI. Par conséquent, c'est la taille complète de chaque AMI personnalisée qui est stockée. Avec les AMI basées sur Amazon EBS, chaque fois que vous personnalisez une AMI et que vous en créez une nouvelle, seules les modifications sont stockées. L'espace de stockage nécessaire pour les AMI personnalisées qui suivent la première est donc bien plus réduit, ce qui entraîne des frais de stockage des AMI moins élevés.

Lorsqu'une instance basée sur Amazon EBS est arrêtée, l'utilisation de l'instance n'est pas facturée, mais le stockage du volume l'est. Dès que vous démarrez votre instance, nous facturons au minimum une minute pour l'utilisation. Après une minute, seules les secondes utilisées vous sont facturées. Si, par exemple, vous exécutez une instance pendant 20 secondes, puis que vous l'arrêtez, nous vous facturons une minute complète. Si vous exécutez une instance pendant 3 minutes et 40 secondes, nous vous facturons exactement 3 minutes et 40 secondes d'utilisation. Chaque seconde, avec un minimum d'une minute, pendant laquelle votre instance s'exécute vous est facturée, même si l'instance demeure inactive et que vous ne vous y connectez pas.

Types de virtualisation AMI Linux

Les Amazon Machine Images Linux utilisent l'un des deux types de virtualisation : virtualisation paravirtuelle ou virtualisation HVM. Les principales différences entre les AMI de virtualisation paravirtuelle ou virtualisation HVM résident dans leur façon de démarrer et leur capacité à tirer parti des extensions matérielles spéciales (UC, réseau et stockage) pour obtenir une meilleure performance.

Pour obtenir les meilleures performances, nous vous recommandons d'utiliser les types d'instance de la génération actuelle et les AMI HVM quand vous lancez vos instances. Pour plus d'informations sur les types d'instance de la génération actuelle, consultez Types d'instance Amazon EC2. Si vous utilisez des types d'instance de la génération précédente et souhaitez effectuer une mise à niveau, consultez Chemins de mise à niveau.

Le tableau suivant compare les AMI HVM et PV.

	HVM	Virtualisation paravirtuelle
Description	Les AMIs HVM sont présentées avec un ensemble entièrement virtualisé de matériel et démarrent en exécutant l'enregistrement d'amorçage maître du périphérique de stockage en mode bloc racine de votre image. Ce type de virtualisation permet d'exécuter un système d'exploitation directement par-dessus une machine virtuelle sans aucune modification, comme si elle était exécutée sur le matériel bare-metal. Le système hôte Amazon EC2 émule une partie ou tout le matériel sous-jacent qui est présenté à l'invité.	Les AMIs de virtualisation paravirtuelle démarrent avec un chargeur de démarrage spécial appelé PV-GRUB qui lance le cycle de démarrage, puis charge en chaîne le noyau spécifié dans le fichier `menu.lst` sur votre image. Les invités de virtualisation paravirtuelle peuvent s'exécuter sur du matériel hôte qui ne prend pas explicitement en charge la virtualisation. Traditionnellement, les invités de virtualisation paravirtuelle avaient de meilleures performances que les invités HVM. A cause des améliorations de la virtualisation HVM et de la disponibilité des pilotes de virtualisation paravirtuelle pour les AMI HVM, ce n'est plus le cas. Pour plus d'informations sur PV-GRUB et son utilisation sur Amazon EC2, consultez Activation de vos propres noyaux Linux (p. 205).
Prise en charge des extensions matérielles	Oui. Contrairement aux invités PV, les invités HVM peuvent profiter des extensions matérielles qui offrent un accès rapide au matériel sous-jacent sur le système hôte. Pour obtenir plus d'informations sur les extensions de virtualisation au niveau de l'UC disponibles dans Amazon EC2, consultez Intel Virtualization Technology sur le site Web Intel. Les AMI HVM sont obligatoires pour tirer parti de la mise en réseau améliorée et du traitement GPU. Afin de passer les instructions sur le réseau spécialisé et les appareils GPU, le système d'exploitation doit pouvoir avoir accès à la plate-forme matérielle initiale.	Non, ils ne peuvent pas tirer parti des extensions matérielles spéciales telles que la mise en réseau améliorée ou le traitement GPU.

	HVM	Virtualisation paravirtuelle
	La virtualisation HVM donne cet accès. Pour plus d'informations, consultez Mise en réseau améliorée sur Linux (p. 864) et Instances à calcul accéléré Linux (p. 296).	
Types d'instance pris en charge	Tous les types d'instance de la génération actuelle prennent en charge les AMI HVM.	Les types d'instance de la génération précédente prennent en charge les AMI PV suivantes :C1, C3, HS1, M1, M3, M2 et T1. Les types d'instance de la génération actuelle ne prennent en charge les AMI de virtualisation paravirtuelle.
Régions prises en charge	Toutes les régions prennent en charge les instances HVM.	Asie-Pacifique (Tokyo), Asie-Pacifique (Singapour), Asie-Pacifique (Sydney), Europe (Francfort), Europe (Irlande), Amérique du Sud (São Paulo), USA Est (Virginie du Nord), USA Ouest (Californie du Nord) et USA Ouest (Oregon)
Comment trouver	Vérifiez que le type de virtualisation de l'AMI est défini sur `hvm` à l'aide de la console ou de la commande describe-images.	Vérifiez que le type de virtualisation de l'AMI est défini sur `paravirtual` à l'aide de la console ou de la commande describe-images.

Virtualisation paravirtuelle sur HVM

Les invités paravirtuels avaient traditionnellement de meilleures performances en ce qui concerne les opérations de stockage et les opérations réseau que les invités HVM, car ils pouvaient tirer parti de pilotes spéciaux pour les E/S qui évitaient la surcharge du réseau et du disque matériel en émulation. Les invités HVM devaient quant à eux appliquer ces instructions à du matériel émulé. Maintenant, les pilotes de virtualisation paravirtuelle sont disponibles pour les invités HVM, donc les systèmes d'exploitation qui ne peuvent pas être utilisés dans un environnement paravirtualisé peuvent encore connaître des avantages en termes de performance en ce qui concerne le stockage et l'E/S du réseau en les utilisant. Avec ces pilotes de virtualisation paravirtuelle sur HVM, les invités HVM peuvent obtenir une performance similaire, ou meilleure, que les invités paravirtuels.

Rechercher une AMI Linux

Avant de pouvoir lancer une instance, vous devez sélectionner une AMI à utiliser. Lorsque vous sélectionnez une AMI, prenez en compte les exigences que vous pourriez avoir pour les instances que vous lancerez :

- La région
- Système d'exploitation
- L'architecture : 32 bits (`i386`), 64 bits (`x86_64`) ou ARM 64 bits (arm64)
- Le type de périphérique racine : Amazon EBS ou stockage d'instance

- Le fournisseur (par exemple, Amazon Web Services)
- Les logiciels supplémentaires (par exemple, SQL Server)

Si vous devez trouver une AMI Windows, consultez la section Recherche d'une AMI Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Sommaire

- Rechercher une AMI Linux à l'aide de la console Amazon EC2 (p. 111)
- Rechercher une AMI à l'aide de l'AWS CLI (p. 112)
- Rechercher l'AMI Amazon Linux la plus récente avec Systems Manager (p. 112)
- Utiliser un paramètre Systems Manager pour rechercher une AMI (p. 113)
- Rechercher une AMI Quick Start (p. 116)

Rechercher une AMI Linux à l'aide de la console Amazon EC2

Vous pouvez rechercher une AMI Linux à l'aide de la console Amazon EC2. Vous pouvez sélectionner dans la liste des AMI lorsque vous utilisez l'assistant de lancement pour lancer une instance, ou rechercher toutes les AMI disponibles à l'aide de la page Images. Chaque région utilise ses propres ID d'AMI.

Pour rechercher une AMI Linux à l'aide de l'assistant de lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région dans laquelle lancer vos instances. Vous pouvez sélectionner n'importe quelle région disponible, quel que soit votre emplacement.
3. Sur le tableau de bord de la console, sélectionnez Launch instance (Lancer une instance).
4. Dans l'onglet Lancement rapide, choisissez l'une des AMI couramment utilisées dans la liste. Si vous ne voyez pas l'AMI dont vous avez besoin, sélectionnez l'onglet My AMIs (Mes AMI), AWS Marketplace ou Community AMIs (AMI de la communauté) pour trouver des AMI supplémentaires. Pour de plus amples informations, veuillez consulter Étape 1 : Sélection d'une Amazon Machine Image (AMI) (p. 531).

Pour rechercher une AMI Linux à l'aide de la page Images

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région dans laquelle lancer vos instances. Vous pouvez sélectionner n'importe quelle région disponible, quel que soit votre emplacement.
3. Dans le panneau de navigation, sélectionnez AMI.
4. (Facultatif) Utilisez les options de filtre pour restreindre la liste des AMI affichées aux seules AMI qui vous intéressent. Par exemple, pour répertorier toutes les AMI Linux fournies par AWS, sélectionnez Images publiques. Cliquez dans la barre de recherche et sélectionnez Propriétaire dans le menu, puis Images Amazon. Cliquez à nouveau dans la barre de recherche pour sélectionner Plateforme, puis le système d'exploitation dans la liste fournie.
5. (Facultatif) Cliquez sur l'icône Afficher / Masquer les colonnes pour sélectionner les attributs d'image à afficher, comme le type de périphérique racine. Vous pouvez également sélectionner une AMI dans la liste et afficher ses propriétés sous l'onglet Détails.
6. Avant de sélectionner une AMI, il est important de vérifier si celle-ci est basée sur le stockage d'instance ou sur Amazon EBS et d'être conscient des effets de cette différence. Pour plus d'informations, consultez Stockage pour le périphérique racine (p. 106).

7. Pour lancer une instance à partir de cette AMI, sélectionnez-la et cliquez sur Lancer. Pour de plus amples informations sur le lancement d'une instance à l'aide de la console, veuillez consulter Lancement d'une instance depuis une AMI (p. 532). Si vous n'êtes pas prêt à lancer l'instance maintenant, notez l'ID de l'AMI pour plus tard.

Rechercher une AMI à l'aide de l'AWS CLI

Vous pouvez utiliser les commandes de AWS CLI spécifiques à Amazon EC2 pour obtenir uniquement la liste des AMI Linux qui vous intéressent. Une fois que vous avez trouvé une AMI répondant à vos besoins, notez son ID afin de pouvoir l'utiliser pour lancer des instances. Pour plus d'informations, consultez Lancement d'une instance à l'aide de la AWS CLI dans le AWS Command Line Interface Guide de l'utilisateur.

La commande describe-images prend en charge les paramètres de filtrage. Par exemple, utilisez le paramètre `--owners` pour afficher les AMI publiques détenues par Amazon.

```
aws ec2 describe-images --owners self amazon
```

Vous pouvez ajouter le filtre suivant à la commande précédente pour afficher uniquement les AMI basées sur Amazon EBS :

```
--filters "Name=root-device-type,Values=ebs"
```

> **Important**
>
> Si l'indicateur `--owners` n'est pas spécifié dans la commande describe-images, les images renvoyées sont celles pour lesquelles vous avez des autorisations de lancement, quel que soit le propriétaire.

Rechercher l'AMI Amazon Linux la plus récente avec Systems Manager

Amazon EC2 fournit des paramètres AWS Systems Manager publics pour les AMI publiques maintenues par AWS que vous pouvez utiliser lors du lancement d'instances. Par exemple, le paramètre fourni par EC2 `/aws/service/ami-amazon-linux-latest/amzn2-ami-hvm-x86_64-gp2` est disponible dans toutes les régions et pointe toujours vers la dernière version de l'AMI Amazon Linux 2 dans une région donnée.

Les paramètres publics de l'AMI Amazon EC2 sont disponibles à partir des chemins suivants :

- `/aws/service/ami-amazon-linux-latest`
- `/aws/service/ami-windows-latest`

Vous pouvez afficher une liste de toutes les AMI Linux dans la région AWS actuelle à l'aide de la commande suivante de l'interface de ligne de commande AWS.

```
aws ssm get-parameters-by-path --path /aws/service/ami-amazon-linux-latest --query
 Parameters[].Name
```

Pour lancer une instance à l'aide d'un paramètre public

L'exemple suivant utilise le paramètre public fourni par EC2 pour lancer une instance `m5.xlarge` à l'aide de la dernière AMI Amazon Linux 2.

Pour spécifier le paramètre dans la commande, utilisez la syntaxe suivante : `resolve:ssm:public-parameter`, où `resolve:ssm` est le préfixe standard et `public-parameter` le chemin et le nom du paramètre public.

Dans l'exemple, les paramètres `--count` et `--security-group` ne sont pas inclus. Pour `--count`, la valeur par défaut est 1. Si vous avez un VPC par défaut et un groupe de sécurité par défaut, ils sont utilisés.

```
aws ec2 run-instances
    --image-id resolve:ssm:/aws/service/ami-amazon-linux-latest/amzn2-ami-hvm-x86_64-gp2
    --instance-type m5.xlarge
    --key-name MyKeyPair
```

Pour de plus amples informations, veuillez consulter Utilisation de paramètres publics dans le Guide de l'utilisateur AWS Systems Manager et Requête pour les derniers ID AMI Amazon Linux à l'aide du stockage de paramètres AWS Systems Manager .

Utiliser un paramètre Systems Manager pour rechercher une AMI

Lorsque vous lancez une instance à l'aide de l'assistant de lancement EC2 dans la console, vous pouvez sélectionner une AMI dans la liste ou sélectionner un paramètre AWS Systems Manager pointant vers un ID AMI. Si vous utilisez le code d'automatisation pour lancer vos instances, vous pouvez spécifier le paramètre Systems Manager au lieu de l'ID AMI.

Un paramètre Systems Manager est une paire clé-valeur définie par le client que vous pouvez créer dans le stockage de paramètres Systems Manager. Le stockage de paramètres fournit un magasin central pour externaliser les valeurs de configuration de vos applications. Pour de plus amples informations, veuillez consulter Stockage de paramètres AWS Systems Manager dans le Guide de l'utilisateur AWS Systems Manager.

Lorsque vous créez un paramètre qui pointe vers un ID AMI, assurez-vous que vous spécifiez le type de données comme `aws:ec2:image`. Ce type de données garantit que lorsque le paramètre est créé ou modifié, la valeur du paramètre est validée en tant qu'ID AMI. Pour de plus amples informations, veuillez consulter Prise en charge des paramètres natifs pour les ID AMI dans le Guide de l'utilisateur AWS Systems Manager.

Sommaire

- Cas d'utilisation (p. 113)
- Lancer une instance à l'aide d'un paramètre Systems Manager (p. 114)
- Autorisations (p. 115)
- Limites (p. 115)

Cas d'utilisation

En utilisant les paramètres Systems Managerpour pointer vers les ID AMI, vous pouvez faciliter la sélection de l'AMI correcte pour vos utilisateurs lors du lancement d'instances, et vous pouvez simplifier la maintenance du code d'automatisation.

Plus facile pour les utilisateurs

Si vous devez lancer des instances à l'aide d'une AMI spécifique et si cette AMI est mise à jour régulièrement, nous vous recommandons de demander à vos utilisateurs de sélectionner un paramètre

Systems Manager pour trouver l'AMI. En demandant à vos utilisateurs de sélectionner un paramètre Systems Manager, vous pouvez vous assurer que la dernière AMI est utilisée pour lancer des instances.

Par exemple, chaque mois dans votre organisation, vous pouvez créer une nouvelle version de votre AMI dotée des derniers correctifs du système d'exploitation et des applications. Vous devez également demander à vos utilisateurs de lancer des instances à l'aide de la dernière version de votre AMI. Pour vous assurer que vos utilisateurs utilisent la dernière version, vous pouvez créer un paramètre Systems Manager (par exemple, `golden-ami`) qui pointe vers l'ID AMI correct. Chaque fois qu'une nouvelle version de l'AMI est créée, vous mettez à jour la valeur de l'ID AMI dans le paramètre afin qu'elle pointe toujours vers la dernière AMI. Vos utilisateurs n'ont pas besoin de connaître les mises à jour périodiques de l'AMI, car ils continuent à sélectionner le même paramètre Systems Manager à chaque fois. En demandant aux utilisateurs de sélectionner un paramètre Systems Manager, il leur est plus facile de sélectionner l'AMI appropriée pour le lancement d'une instance.

Simplifier la maintenance du code d'automatisation

Si vous utilisez le code d'automatisation pour lancer vos instances, vous pouvez spécifier le paramètre Systems Manager au lieu de l'ID AMI. Si une nouvelle version de l'AMI est créée, vous modifiez la valeur de l'ID AMI dans le paramètre afin qu'elle pointe vers la dernière AMI. Le code d'automatisation qui fait référence au paramètre n'a pas besoin d'être modifié chaque fois qu'une nouvelle version de l'AMI est créée. Cela simplifie grandement la maintenance de l'automatisation et réduit les coûts de déploiement.

> Note
>
> Les instances en cours d'exécution ne sont pas affectées lorsque vous modifiez l'ID AMI vers lequel le paramètre Systems Manager pointe.

Lancer une instance à l'aide d'un paramètre Systems Manager

Vous pouvez lancer une instance à l'aide de la console ou de l'AWS CLI. Au lieu de spécifier un ID AMI, vous pouvez spécifier un paramètre AWS Systems Manager qui pointe vers un ID AMI.

Pour rechercher une AMI Linux à l'aide d'un paramètre Systems Manager (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région dans laquelle lancer vos instances. Vous pouvez sélectionner n'importe quelle région disponible, quel que soit votre emplacement.
3. Sur le tableau de bord de la console, sélectionnez Launch instance (Lancer une instance).
4. Choisissez Rechercher par paramètre Systems Manager (en haut à droite).
5. Pour Paramètre Systems Manager, sélectionnez un paramètre. L'ID AMI correspondant apparaît à côté de Currently resolves to (Se résout en).
6. Choisissez Search (Rechercher). Les AMI correspondant à l'ID AMI apparaissent dans la liste.
7. Sélectionnez l'AMI dans la liste, puis choisissez Select (Sélectionner).

Pour de plus amples informations sur le lancement d'une instance à partir d'une AMI à l'aide de l'assistant de lancement, veuillez consulter Étape 1 : Sélection d'une Amazon Machine Image (AMI) (p. 531).

Pour lancer une instance à l'aide d'un paramètre AWS Systems Manager au lieu d'un ID AMI (AWS CLI)

L'exemple suivant utilise le paramètre Systems Manager `golden-ami` pour lancer une instance `m5.xlarge`. Le paramètre pointe vers un ID AMI.

Pour spécifier le paramètre dans la commande, utilisez la syntaxe suivante : `resolve:ssm:/`*`parameter-name`*, où `resolve:ssm` est le préfixe standard et `parameter-name` est le nom du paramètre unique. Notez que le nom du paramètre est sensible à la casse. Les barres obliques inverses pour le nom du paramètre ne sont nécessaires que si le paramètre fait partie d'une hiérarchie,

par exemple `/amis/production/golden-ami`. Vous pouvez omettre la barre oblique inverse si le paramètre ne fait pas partie d'une hiérarchie.

Dans l'exemple, les paramètres `--count` et `--security-group` ne sont pas inclus. Pour `--count`, la valeur par défaut est 1. Si vous avez un VPC par défaut et un groupe de sécurité par défaut, ils sont utilisés.

```
aws ec2 run-instances
    --image-id resolve:ssm:/golden-ami
    --instance-type m5.xlarge
    ...
```

Pour lancer une instance à l'aide d'une version spécifique d'un paramètre AWS Systems Manager (AWS CLI)

Les paramètres Systems Manager ont la prise en charge de la version. Chaque itération d'un paramètre se voit attribuer un numéro de version unique. Vous pouvez référencer la version du paramètre comme suit : `resolve:ssm:parameter-name:version`, où `version` est le numéro de version unique. Par défaut, la dernière version du paramètre est utilisée lorsqu'aucune version n'est spécifiée.

L'exemple suivant utilise la version 2 du paramètre.

Dans l'exemple, les paramètres `--count` et `--security-group` ne sont pas inclus. Pour `--count`, la valeur par défaut est 1. Si vous avez un VPC par défaut et un groupe de sécurité par défaut, ils sont utilisés.

```
aws ec2 run-instances
    --image-id resolve:ssm:/golden-ami:2
    --instance-type m5.xlarge
    ...
```

Pour lancer une instance à l'aide d'un paramètre public fourni par AWS

Amazon EC2 propose des paramètres Systems Manager publics pour les AMI publiques fournies par AWS. Par exemple, le paramètre public /aws/service/ami-amazon-linux-latest/amzn2-ami-hvm-x86_64-gp2 est disponible dans toutes les régions et pointe toujours vers la dernière version de l'AMI Amazon Linux 2 de la région.

```
aws ec2 run-instances
    --image-id resolve:ssm:/aws/service/ami-amazon-linux-latest/amzn2-ami-hvm-x86_64-gp2
    --instance-type m5.xlarge
    ...
```

Autorisations

Si vous utilisez des paramètres Systems Manager qui pointent vers des ID AMI dans l'assistant de lancement d'instance, vous devez ajouter `ssm:DescribeParameters` et `ssm:GetParameters` à votre stratégie IAM. `ssm:DescribeParameters` accorde à vos utilisateurs IAM l'autorisation d'afficher et de sélectionner des paramètres Systems Manager. `ssm:GetParameters` accorde à vos utilisateurs IAM l'autorisation d'obtenir les valeurs des paramètres Systems Manager. Vous pouvez également restreindre l'accès à des paramètres Systems Manager spécifiques. Pour de plus amples informations, veuillez consulter Utilisation de l'assistant de lancement d'EC2 (p. 1027).

Limites

Les AMI et les paramètres Systems Manager sont spécifiques à la région. Pour utiliser le même nom de paramètre Systems Manager dans les régions, créez un paramètre Systems Manager dans chaque région

avec le même nom (par exemple, `golden-ami`). Dans chaque région, pointez le paramètre Systems Manager sur une AMI de cette région.

Rechercher une AMI Quick Start

Lorsque vous lancez une instance à l'aide de la console Amazon EC2, la page Sélection d'une Amazon Machine Image (AMI) inclut la liste des AMI populaires sous l'onglet Lancement rapide. Si vous souhaitez automatiser le lancement d'une instance à l'aide d'une de ces images machine Amazon de lancement rapide, vous devez localiser par programme l'ID de la version actuelle de l'IMA.

Pour localiser la version actuelle d'une AMI de lancement rapide, vous pouvez énumérer toutes les AMI avec son nom d'AMI, puis trouver celle avec la date de création la plus récente.

Example Exemple : Rechercher l'IMA Amazon Linux 2 actuelle

```
aws ec2 describe-images \
    --owners amazon \
    --filters 'Name=name,Values=amzn2-ami-hvm-2.0.????????.?-x86_64-gp2'
 'Name=state,Values=available' \
    --query 'reverse(sort_by(Images, &CreationDate))[:1].ImageId' \
    --output text
```

Example Exemple : Rechercher l'IMA Amazon Linux actuelle

```
aws ec2 describe-images \
    --owners amazon \
    --filters 'Name=name,Values=amzn-ami-hvm-????.??.?.????????-x86_64-gp2'
 'Name=state,Values=available' \
    --query 'reverse(sort_by(Images, &CreationDate))[:1].ImageId' \
    --output text
```

Example Exemple : Rechercher l'IMA Ubuntu Server 16.04 LTS actuelle

```
aws ec2 describe-images \
    --owners 099720109477 \
    --filters 'Name=name,Values=ubuntu/images/hvm-ssd/ubuntu-xenial-16.04-amd64-
server-????????' 'Name=state,Values=available' \
    --query 'reverse(sort_by(Images, &CreationDate))[:1].ImageId' \
    --output text
```

Example Exemple : Rechercher l'IMA Red Hat Enterprise Linux 7.5 actuelle

```
aws ec2 describe-images \
    --owners 309956199498 \
    --filters 'Name=name,Values=RHEL-7.5_HVM_GA*' 'Name=state,Values=available' \
    --query 'reverse(sort_by(Images, &CreationDate))[:1].ImageId' \
    --output text
```

Example Exemple : Rechercher l'IMA SUSE Linux Enterprise Server 15 actuelle

```
aws ec2 describe-images \
    --owners amazon \
    --filters 'Name=name,Values=suse-sles-15-v????????-hvm-ssd-x86_64'
 'Name=state,Values=available' \
    --query 'reverse(sort_by(Images, &CreationDate))[:1].ImageId' \
```

```
    --output text
```

AMI partagées

Une AMI partagée est une AMI qu'un développeur a créée et mise à disposition pour qu'elle puisse être utilisée par d'autres développeurs. L'une des façons les plus simples de se lancer avec Amazon EC2 est d'utiliser une AMI partagée qui possède les composants dont vous avez besoin, puis d'y ajouter du contenu personnalisé. Vous pouvez également créer vos propres AMI et les partager avec d'autres.

Vous utilisez une AMI partagée à vos propres risques. Amazon ne peut se porter garant de l'intégrité ou de la sécurité des AMI partagées par d'autres utilisateurs Amazon EC2. Par conséquent, vous devriez traiter les AMI partagées de la même façon que vous traiteriez un code étranger que vous envisageriez de déployer dans votre propre centre de données, et prendre toutes les précautions nécessaires. Nous vous recommandons d'utiliser une AMI provenant d'une source de confiance.

Les images publiques Amazon ont un propriétaire disposant d'un alias qui apparaît en tant que `amazon` dans le champ compte. Cela vous permet de trouver facilement des AMI provenant d'Amazon. Les autres utilisateurs ne peuvent attribuer un alias à leurs AMI.

Pour obtenir des informations sur la création d'une AMI, consultez Création d'une AMI Linux basée sur un stockage d'instance ou Création d'une AMI Linux basée sur Amazon EBS . Pour de plus amples informations sur la création, la livraison et la maintenance de vos applications sur AWS Marketplace, veuillez consulter la documentation AWS Marketplace.

Sommaire

- Rechercher des AMI partagées (p. 117)
- Rendre une AMI publique (p. 119)
- Partager une AMI avec des comptes AWS spécifiques (p. 121)
- Utiliser des signets (p. 123)
- Consignes pour les AMI Linux partagées (p. 123)

Rechercher des AMI partagées

Vous pouvez utiliser la console Amazon EC2 ou la ligne de commande pour trouver des AMI partagées.

Les AMI sont une ressource régionale. Par conséquent, lorsque vous recherchez une AMI partagée (publique ou privée), vous devez la rechercher dans la région à partir de laquelle elle est partagée. Pour rendre une AMI disponible dans une autre région, copiez-la dans la région souhaitée puis partagez-la. Pour plus d'informations, consultez Copie d'une AMI.

Rubriques

- Rechercher une AMI partagée (console) (p. 117)
- Rechercher une AMI partagée (AWS CLI) (p. 118)
- Utilisation des AMI partagées (p. 119)

Rechercher une AMI partagée (console)

Pour trouver une AMI privée partagée à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le volet de navigation, sélectionnez AMI.
3. Dans le premier filtre, choisissez Images privées. Toutes les AMI qui ont été partagées avec vous sont listées. Pour affiner votre recherche, choisissez la barre de recherche et utilisez les options de filtre offertes dans le menu.

Pour trouver une AMI publique partagée à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez AMI.
3. Dans le premier filtre, choisissez Images publiques. Pour affiner votre recherche, choisissez la barre de recherche et utilisez les options de filtre offertes dans le menu.
4. Utilisez les filtres pour lister uniquement les types d'AMI qui vous intéressent. Par exemple, choisissez Propriétaire, puis sélectionnez Images Amazon pour afficher uniquement les images publiques Amazon.

Rechercher une AMI partagée (AWS CLI)

Utilisez la commande describe-images (AWS CLI) pour répertorier les AMI. Vous pouvez parcourir la liste des types d'AMI qui vous intéressent, comme le montrent les exemples suivants.

Exemple : Affichage de toutes les AMI publiques

La commande suivante liste toutes les AMI publiques, y compris les AMI publiques que vous possédez.

```
aws ec2 describe-images --executable-users all
```

Exemple : Affichage des AMI avec des autorisations de lancement explicites

La commande suivante liste toutes les AMI pour lesquelles vous disposez d'autorisations de lancement explicites. Cette liste n'inclut pas les AMI publiques que vous possédez.

```
aws ec2 describe-images --executable-users self
```

Exemple : Affichage des AMI appartenant à Amazon

La commande suivante liste toutes les AMI possédées par Amazon. Les AMI publiques Amazon ont un propriétaire disposant d'un alias qui apparaît en tant que `amazon` dans le champ compte. Cela vous permet de trouver facilement des AMI provenant d'Amazon. Les autres utilisateurs ne peuvent attribuer un alias à leurs AMI.

```
aws ec2 describe-images --owners amazon
```

Exemple : Affichage des AMI appartenant à un compte

La commande suivante liste toutes les AMI possédées par le compte AWS spécifié.

```
aws ec2 describe-images --owners 123456789012
```

Exemple 1 : Limitation du nombre d'AMI affichées à l'aide d'un filtre

Pour réduire le nombre d'AMI affichées, utilisez un filtre pour lister uniquement les types d'AMI qui vous intéressent. Par exemple, utilisez le filtre suivant pour afficher uniquement les AMI basées sur EBS.

```
--filters "Name=root-device-type,Values=ebs"
```

Utilisation des AMI partagées

Avant d'utiliser une AMI partagée, suivez les étapes ci-après afin de vous assurer qu'il n'y a pas d'informations d'identification pré-installées qui permettraient un accès non désiré à votre instance par un tiers, ni de journalisation à distance préconfigurée susceptible de transmettre des données sensibles à un tiers. Consultez la documentation portant sur la distribution Linux utilisée par l'AMI pour en savoir plus sur la façon d'améliorer la sécurité du système.

Afin de vous assurer que vous ne perdiez pas accidentellement accès à votre instance, nous vous recommandons d'initier deux sessions SSH et de conserver la seconde session ouverte jusqu'à ce que vous retiriez les informations d'identification que vous ne reconnaissez pas, et que vous confirmiez que vous pouvez toujours vous connecter à votre instance à l'aide de SSH.

1. Identifiez et désactivez toute clé SSH publique non-autorisée. La seule clé dans le fichier devrait être la clé que vous avez utilisée pour lancer l'AMI. La commande suivante localise les fichiers `authorized_keys` :

   ```
   [ec2-user ~]$ sudo find / -name "authorized_keys" -print -exec cat {} \;
   ```

2. Désactivez l'authentification basée sur mot de passe pour l'utilisateur racine. Ouvrez le fichier `sshd_config` et éditez la ligne `PermitRootLogin` de la façon suivante :

   ```
   PermitRootLogin without-password
   ```

   L'alternative est de désactiver la possibilité de se connecter à l'instance en tant qu'utilisateur racine :

   ```
   PermitRootLogin No
   ```

   Redémarrez le service sshd.
3. Vérifiez si d'autres comptes utilisateur peuvent se connecter à votre instance. Les comptes possédant des privilèges super-utilisateur sont particulièrement dangereux. Supprimez ou verrouillez le mot de passe de tout compte inconnu.
4. Vérifiez s'il y a des ports ouverts que vous n'utilisez pas et des services de réseau en cours d'exécution en attente de connexions entrantes.
5. Pour éviter toute journalisation à distance préconfigurée, vous devriez supprimer le fichier de configuration existant et redémarrer le service rsyslog. Par exemple :

   ```
   [ec2-user ~]$ sudo rm /etc/rsyslog.conf
   [ec2-user ~]$ sudo service rsyslog restart
   ```

6. Vérifiez que toutes les tâches cron sont légitimes.

Si vous découvrez une AMI publique qui présente selon vous un risque de sécurité, contactez l'équipe de sécurité AWS. Pour plus d'informations, consultez le Centre de sécurité AWS.

Rendre une AMI publique

Amazon EC2 vous permet de partager vos AMI avec d'autres comptes AWS. Vous pouvez autoriser tous les comptes AWS à lancer l'AMI (rendre l'AMI publique) ou n'autoriser que quelques comptes spécifiques à le faire (consultez Partager une AMI avec des comptes AWS spécifiques (p. 121)). Vous n'êtes pas

facturé lorsque votre AMI est lancée par d'autres comptes AWS ; seuls les comptes lançant l'AMI sont facturés.

Les AMI avec des volumes chiffrés ne peuvent pas être rendues publiques.

Les AMI sont une ressource régionale. Par conséquent, lorsque vous partagez une AMI, celle-ci devient disponible dans la région concernée. Pour rendre une AMI disponible dans une autre région, copiez-la dans la région souhaitée puis partagez-la. Pour plus d'informations, consultez Copie une AMI (p. 173).

Pour éviter d'exposer des données sensibles lorsque vous partagez une AMI, consultez les normes de sécurités spécifiées ici Consignes pour les AMI Linux partagées (p. 123) et suivez les actions recommandées.

Si une AMI a un code produit ou contient un instantané ou un volume chiffré, vous pouvez la rendre publique. Vous pouvez partager l'AMI uniquement avec des comptes AWS spécifiques.

Rubriques

- Partager une AMI avec tous les comptes AWS (console) (p. 120)
- Partager une AMI avec tous les comptes AWS (AWS CLI) (p. 120)

Partager une AMI avec tous les comptes AWS (console)

Une fois que vous avez rendu une AMI publique, elle est disponible dans le champ AMI de la communauté lorsque vous lancez une instance dans la même région à l'aide de la console. Notez que cela peut prendre quelques instants pour qu'une AMI s'affiche dans le champ AMI de la communauté une fois que vous l'avez rendue publique. Cela peut également prendre quelques instants pour qu'une AMI soit supprimée du champ AMI de la communauté une fois que vous l'avez rendue privée à nouveau.

Pour partager une AMI publique à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez AMI.
3. Sélectionnez votre AMI dans la liste, puis Actions et Modifier des autorisations d'image.
4. Choisissez Public puis Enregistrer.

Partager une AMI avec tous les comptes AWS (AWS CLI)

Chaque AMI dispose d'une propriété `launchPermission` qui contrôle quels comptes AWS, hormis celui du propriétaire, sont autorisés à utiliser l'AMI pour lancer des instances. En modifiant la propriété `launchPermission` d'une AMI, vous pouvez la rendre publique (ce qui donne des autorisations de lancement à tous les comptes AWS), ou la partager uniquement avec les comptes AWS que vous spécifiez.

Vous pouvez ajouter ou supprimer des ID de compte de la liste des comptes disposant d'autorisations de lancement pour une AMI. Pour rendre l'AMI publique, spécifiez le groupe `all`. Vous pouvez spécifier à la fois des autorisations de lancement publiques et explicites.

Pour rendre une AMI publique

1. Utilisez la commande modify-image-attribute de la façon suivante pour ajouter le groupe `all` à la liste `launchPermission` pour l'AMI spécifiée.

```
aws ec2 modify-image-attribute \
```

```
    --image-id ami-0abcdef1234567890 \
    --launch-permission "Add=[{Group=all}]"
```

2. Pour vérifier les autorisations de lancement d'une AMI, utilisez la commande describe-image-attribute.

```
aws ec2 describe-image-attribute \
    --image-id ami-0abcdef1234567890 \
    --attribute launchPermission
```

3. (Facultatif) Pour rendre l'AMI de nouveau privée, supprimez le groupe `all` de ses autorisations de lancement. Veuillez noter que le propriétaire de l'AMI dispose toujours d'autorisations de lancement et n'est, par conséquent, pas affecté par cette commande.

```
aws ec2 modify-image-attribute \
    --image-id ami-0abcdef1234567890 \
    --launch-permission "Remove=[{Group=all}]"
```

Partager une AMI avec des comptes AWS spécifiques

Vous pouvez partager une AMI avec des comptes AWS spécifiques, sans rendre l'AMI publique. Vous avez uniquement besoin des ID des comptes AWS. Vous pouvez uniquement partager des AMI qui ont des volumes non chiffrés et des volumes chiffrés avec une CMK gérée par le client. Si vous partagez une AMI avec des volumes chiffrés, vous devez également partager les clés CMK utilisées pour les chiffrer. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143). Vous ne pouvez pas partager une AMI dont les volumes sont chiffrés avec une CMK gérée par AWS.

Les AMI sont une ressource régionale. Par conséquent, lorsque vous partagez une AMI, celle-ci devient disponible dans la région concernée. Pour rendre une AMI disponible dans une autre région, copiez-la dans la région souhaitée puis partagez-la. Pour plus d'informations, consultez Copie une AMI (p. 173).

Aucune limite n'est appliquée au nombre de comptes AWS avec lesquels une AMI peut être partagée. Les balises définies par l'utilisateur que vous attachez à une AMI partagée ne sont disponibles que pour votre compte AWS et non pour les autres comptes avec lesquels l'AMI est partagée.

une AMI (console)

Pour donner des autorisations de lancement explicites à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez AMI.
3. Sélectionnez votre AMI dans la liste, puis Actions et Modifier des autorisations d'image.
4. Spécifiez le numéro de compte AWS de l'utilisateur avec lequel vous souhaitez partager l'AMI dans le champ Numéro de compte AWS, puis choisissez Ajouter autorisation.

   Pour partager cette AMI avec plusieurs utilisateurs, répétez cette étape jusqu'à avoir ajouté tous les utilisateurs requis.
5. Pour donner des autorisations de création de volumes pour des instantanés, sélectionnez Ajouter des autorisations "Créer un volume" aux instantanés associés suivants lors de la création d'autorisations.

   Note

   Vous n'avez pas besoin de partager les instantanés (snapshots) Amazon EBS qu'une AMI référence afin de partager l'AMI. Seule l'AMI elle-même doit être partagée. Le système permet automatiquement à l'instance d'accéder aux instantanés (snapshots) Amazon EBS référencés pour le lancement. Toutefois, vous n’avez pas besoin de partager les clés CMK utilisées pour

chiffrer les instantanés référencés par l'AMI. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

6. Choisissez Enregistrer lorsque vous avez terminé.
7. (Facultatif) Pour afficher les ID de compte AWS avec lesquels vous avez partagé l'AMI, sélectionnez l'AMI dans la liste, puis choisissez l'onglet Autorisations. Pour rechercher les AMI partagées avec vous, consultez Rechercher des AMI partagées (p. 117).

une AMI (AWS CLI)

Utilisez la commande modify-image-attribute (AWS CLI) pour partager une AMI comme illustré dans les exemples suivants.

Pour donner des autorisations de lancement explicites

La commande suivante donne au compte AWS spécifié des autorisations de lancement pour l'AMI spécifiée.

```
aws ec2 modify-image-attribute \
    --image-id ami-0abcdef1234567890 \
    --launch-permission "Add=[{UserId=123456789012}]"
```

La commande suivante octroie une autorisation de création de volume pour un instantané.

```
aws ec2 modify-snapshot-attribute \
    --snapshot-id snap-1234567890abcdef0 \
    --attribute createVolumePermission \
    --operation-type add \
    --user-ids 123456789012
```

Note

Vous n'avez pas besoin de partager les instantanés (snapshots) Amazon EBS qu'une AMI référence afin de partager l'AMI. Seule l'AMI elle-même doit être partagée. Le système permet automatiquement à l'instance d'accéder aux instantanés (snapshots) Amazon EBS référencés pour le lancement. Toutefois, vous n'avez pas besoin de partager les clés CMK utilisées pour chiffrer les instantanés référencés par l'AMI. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

Pour supprimer des autorisations de lancement données à un compte

La commande suivante retire au compte AWS spécifié les autorisations de lancement pour l'AMI spécifiée :

```
aws ec2 modify-image-attribute \
    --image-id ami-0abcdef1234567890 \
    --launch-permission "Remove=[{UserId=123456789012}]"
```

La commande suivante supprime une autorisation de création de volume pour un instantané.

```
aws ec2 modify-snapshot-attribute \
    --snapshot-id snap-1234567890abcdef0 \
    --attribute createVolumePermission \
    --operation-type remove \
    --user-ids 123456789012
```

Pour supprimer toutes les autorisations de lancement

La commande suivante retire toutes les autorisations de lancement publiques et explicites pour l'AMI spécifiée. Veuillez noter que le propriétaire de l'AMI dispose toujours d'autorisations de lancement et n'est, par conséquent, pas affecté par cette commande.

```
aws ec2 reset-image-attribute \
    --image-id ami-0abcdef1234567890 \
    --attribute launchPermission
```

Utiliser des signets

Si vous avez créé une AMI publique ou si vous avez partagé une AMI avec un autre utilisateur AWS, vous pouvez créer un marque-page qui permet à un utilisateur d'accéder à votre AMI et de lancer immédiatement une instance dans leur propre compte. C'est une façon simple de partager des références d'AMI afin que les utilisateurs n'aient pas à passer du temps à trouver votre AMI en vue de l'utiliser.

Veuillez notez que votre AMI doit être publique, ou que vous devez l'avoir partagée avec l'utilisateur à qui vous souhaitez envoyer le marque-page.

Pour créer un marque-page pour votre AMI

1. Saisissez une URL avec les informations suivantes dans lesquelles Région correspond à la région dans laquelle votre AMI réside :

   ```
   https://console.aws.amazon.com/ec2/v2/home?
   region=region#LaunchInstanceWizard:ami=ami_id
   ```

   Par exemple, cette URL lance une instance depuis l'AMI ami-0abcdef1234567890 dans la région us-east-1 :

   ```
   https://console.aws.amazon.com/ec2/v2/home?region=us-
   east-1#LaunchInstanceWizard:ami=ami-0abcdef1234567890
   ```

2. Communiquez le lien aux utilisateurs qui souhaitent utiliser votre AMI.
3. Pour utiliser un marque-page, choisissez le lien ou effectuez un copier-coller dans votre navigateur. L'assistant de lancement s'ouvre avec l'AMI déjà sélectionnée.

Consignes pour les AMI Linux partagées

Utilisez les consignes suivantes pour réduire la surface d'attaque et améliorer la fiabilité des AMI que vous créez.

> **Important**
>
> Aucune liste de consignes de sécurité ne peut être exhaustive. Créez vos AMI partagées avec soin et prenez le temps d'étudier où vous exposez peut-être des données sensibles.

Sommaire

- Mise à jour des outils AMI avant leur utilisation (p. 124)
- Désactivation des connexions à distance basées sur mot de passe à la racine (p. 124)
- Désactivation de l'accès local à la racine (p. 125)
- Suppression des paires de clés de l'hôte SSH (p. 125)
- Installation d'informations d'identification publiques (p. 126)
- Désactivation des vérifications DNS sshd (facultatif) (p. 127)

- Vous identifier (p. 127)
- Vous protéger (p. 127)

Si vous créez des AMI pour AWS Marketplace, consultez Création d'AMIs pour AWS Marketplace pour obtenir des consignes, des stratégies et les bonnes pratiques.

Pour plus d'informations sur la façon de partager des AMI en toute sécurité, consultez les articles suivants :

- How To Share and Use Public AMIs in A Secure Manner
- Public AMI Publishing: Hardening and Clean-up Requirements

Mise à jour des outils AMI avant leur utilisation

Pour les AMI basées sur un stockage d'instance, nous recommandons que vos AMI téléchargent et mettent à jour les outils de création AMI Amazon EC2 avant de les utiliser. Cela garantit que les nouvelles AMI basées sur vos AMI partagées disposent des derniers outils AMI.

Pour Amazon Linux 2, installez le package `aws-amitools-ec2` et ajoutez les outils AMI à votre variable PATH avec la commande suivante. Pour Amazon Linux AMI, le package `aws-amitools-ec2` est déjà installé par défaut.

```
[ec2-user ~]$ sudo yum install -y aws-amitools-ec2 && export PATH=$PATH:/opt/aws/bin > /
etc/profile.d/aws-amitools-ec2.sh && . /etc/profile.d/aws-amitools-ec2.sh
```

Mettez à niveau les outils AMI avec la commande suivante :

```
[ec2-user ~]$ sudo yum upgrade -y aws-amitools-ec2
```

Pour les autres distributions, assurez-vous que vous disposez des derniers outils AMI.

Désactivation des connexions à distance basées sur mot de passe à la racine

En utilisant un mot de passe racine fixe pour une AMI publique, un risque de sécurité peut rapidement apparaître. Même le fait de compter sur les utilisateurs pour changer le mot de passe après leur première connexion laisse une petite place à une opportunité d'abus potentiel.

Pour résoudre ce problème, désactivez les connexions à distance basées sur mot de passe pour l'utilisateur racine.

Pour désactiver les connexions à distances basées sur mot de passe à la racine

1. Ouvrez le fichier `/etc/ssh/sshd_config` dans un éditeur de texte et localisez la ligne suivante :

   ```
   #PermitRootLogin yes
   ```

2. Changez la ligne en :

   ```
   PermitRootLogin without-password
   ```

   L'emplacement de ce fichier de configuration peut varier pour votre distribution, ou si vous n'exécutez pas OpenSSH. Si tel est le cas, consultez la documentation appropriée.

Désactivation de l'accès local à la racine

Lorsque vous travaillez avec des AMI partagées, une bonne pratique consiste à désactiver les connexions directes à la racine. Pour ce faire, connectez-vous à votre instance en cours d'exécution et entrez la commande suivante :

```
[ec2-user ~]$ sudo passwd -l root
```

> **Note**
>
> Cette commande n'a pas d'impact sur l'utilisation de `sudo`.

Suppression des paires de clés de l'hôte SSH

Si vous prévoyez de partager une AMI issue d'une AMI publique, supprimez les paires de clés de l'hôte SSH existantes situées dans `/etc/ssh`. Cela force SSH à générer de nouvelles paires de clés SSH uniques lorsque quelqu'un lance une instance utilisant votre AMI, ce qui améliore la sécurité et réduit la probabilité d'attaques MITM.

Supprimez tous les fichiers clés suivants présents dans votre système.

- ssh_host_dsa_key
- ssh_host_dsa_key.pub
- ssh_host_key
- ssh_host_key.pub
- ssh_host_rsa_key
- ssh_host_rsa_key.pub
- ssh_host_ecdsa_key
- ssh_host_ecdsa_key.pub
- ssh_host_ed25519_key
- ssh_host_ed25519_key.pub

Vous pouvez supprimer tous ces fichiers en toute sécurité avec la commande suivante.

```
[ec2-user ~]$ sudo shred -u /etc/ssh/*_key /etc/ssh/*_key.pub
```

> **Warning**
>
> Les utilitaires de suppression sécurisée tels que **`shred`** peuvent ne pas supprimer toutes les copies d'un fichier de vos supports de stockage. Des copies cachées de fichiers peuvent être crées par les systèmes de fichiers de journalisation (dont Amazon Linux default ext4), les instantanés (snapshots), les sauvegardes, RAID et la mise en cache temporaire. Pour plus d'informations, consultez la documentation **`shred`**.
>
> **Important**
>
> Si vous oubliez de supprimer les paires de clés de l'hôte SSH existantes de votre AMI publique, notre processus routinier d'audit vous informe ainsi que tous les clients exécutant des instances de votre AMI du risque de sécurité potentiel. Au terme d'une courte période de grâce, nous marquons l'AMI comme privée.

Installation d'informations d'identification publiques

Après avoir configuré l'AMI pour empêcher la connexion à l'aide d'un mot de passe, vous devez vous assurer que les utilisateurs peuvent se connecter à l'aide d'un autre mécanisme.

Amazon EC2 permet aux utilisateurs de spécifier un nom de paire de clés publique-privée au moment de lancer une instance. Lorsqu'un nom de paire de clés valide est fourni à l'appel de l'API `RunInstances` (ou par les outils API de ligne de commande), la clé publique (la portion de la paire de clés qu'Amazon EC2 conserve sur le serveur après un appel à `CreateKeyPair` ou `ImportKeyPair`) est rendue disponible pour l'instance via une requête HTTP sur les métadonnées d'instance.

Pour se connecter via SSH, votre AMI doit récupérer la valeur clé au moment du démarrage et la joindre à `/root/.ssh/authorized_keys` (ou l'équivalent pour tout autre compte utilisateur sur l'AMI). Les utilisateurs peuvent lancer des instances de votre AMI avec votre paire de clés et se connecter sans avoir besoin de mot de passe racine.

De nombreuses distributions, dont Amazon Linux et Ubuntu, utilisent le package `cloud-init` pour injecter des informations d'identification de clé publiques pour un utilisateur configuré. Si votre distribution ne prend pas en charge `cloud-init`, vous pouvez ajouter le code suivant à un script de démarrage système (tel que `/etc/rc.local`) pour extraire la clé publique que vous avez spécifiée au lancement pour l'utilisateur racine.

> **Note**
>
> Dans l'exemple suivant, l'adresse IP http://169.254.169.254/ est une adresse lien-local et elle n'est valide que depuis l'instance.

IMDSv2

```
if [ ! -d /root/.ssh ] ; then
        mkdir -p /root/.ssh
        chmod 700 /root/.ssh
fi
# Fetch public key using HTTP
TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-metadata-
token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/public-keys/0/openssh-key > /tmp/my-key
if [ $? -eq 0 ] ; then
        cat /tmp/my-key >> /root/.ssh/authorized_keys
        chmod 700 /root/.ssh/authorized_keys
        rm /tmp/my-key
fi
```

IMDSv1

```
if [ ! -d /root/.ssh ] ; then
        mkdir -p /root/.ssh
        chmod 700 /root/.ssh
fi
# Fetch public key using HTTP
curl http://169.254.169.254/latest/meta-data/public-keys/0/openssh-key > /tmp/my-key
if [ $? -eq 0 ] ; then
        cat /tmp/my-key >> /root/.ssh/authorized_keys
        chmod 700 /root/.ssh/authorized_keys
        rm /tmp/my-key
fi
```

Cela peut être appliqué à n'importe quel compte utilisateur, vous n'avez pas besoin de vous limiter à `root`.

Note

La création d'un nouveau bundle d'une instance basée sur cette AMI inclut la clé avec laquelle elle a été lancée. Pour éviter l'inclusion de la clé, vous devez vider (ou supprimer) le fichier `authorized_keys` ou exclure ce fichier du nouveau bundle.

Désactivation des vérifications DNS sshd (facultatif)

Désactiver les vérifications DNS sshd affaiblit quelque peu votre sécurité sshd. Toutefois, si la résolution DNS échoue, les connexions SSH continuent de fonctionner. Si vous ne désactivez pas les vérifications sshd, les échecs de résolution DNS empêchent toutes les connexions.

Pour désactiver les vérifications DNS sshd

1. Ouvrez le fichier `/etc/ssh/sshd_config` dans un éditeur de texte et localisez la ligne suivante :

```
#UseDNS yes
```

2. Changez la ligne en :

```
UseDNS no
```

Note

L'emplacement de ce fichier de configuration peut varier pour votre distribution, ou si vous n'exécutez pas OpenSSH. Si tel est le cas, consultez la documentation appropriée.

Vous identifier

A l'heure actuelle, il n'y a aucun moyen simple de savoir qui a fourni une AMI partagée, car chaque AMI est représentée par un ID de compte.

Nous vous recommandons de poster une description de votre AMI ainsi que l'ID de l'AMI dans le Amazon EC2 forum. Les utilisateurs qui souhaitent essayer de nouvelles AMI partagées disposent ainsi d'un emplacement central pratique.

Vous protéger

Nous déconseillons de stocker des données ou logiciels sensibles sur toute AMI que vous partagez. Les utilisateurs qui lancent une AMI partagée peuvent être en mesure de la regrouper et de l'enregistrer comme étant la leur. Suivez ces consignes pour vous permettre d'éviter de vous exposer à des risques de sécurité facilement négligés :

- Nous recommandons d'utiliser l'option `--exclude` *`directory`* sur `ec2-bundle-vol` pour éviter tout répertoire et sous-répertoire contenant des informations secrètes que vous ne souhaiteriez pas inclure dans votre regroupement. Excluez notamment toutes les paires de clés publiques/privées SSH appartenant à l'utilisateur, et les fichiers SSH `authorized_keys` lorsque vous créez un bundle de l'image. Les AMI publiques Amazon stockent ces éléments dans `/root/.ssh` pour le compte racine et dans `/home/`*`user_name`*`/.ssh/` pour les comptes utilisateur réguliers. Pour plus d'informations, consultez ec2-bundle-vol (p. 154).
- Supprimez toujours l'historique shell avant la création d'un bundle. Si vous essayez de réaliser plusieurs téléchargements de regroupement dans une même AMI, l'historique shell contient votre clé d'accès secrète. L'exemple ci-après devrait être la dernière commande exécutée avant de procéder à la création d'un bundle depuis l'instance.

```
[ec2-user ~]$ shred -u ~/.*history
```

> **Warning**
>
> Les limites de **shred** décrites dans l'avertissement ci-dessus s'appliquent également ici. Ayez à l'esprit que bash inscrit l'historique de la session en cours sur le disque au moment de quitter. Si vous vous déconnectez de votre instance après avoir supprimé `~/.bash_history` et si vous vous reconnectez ensuite, vous constaterez que `~/.bash_history` a été recréé et contient toutes les commandes exécutées durant votre session précédente.
> D'autres programmes en dehors de bash inscrivent les historiques sur le disque. Soyez prudent et retirez ou excluez tous les fichiers et répertoires dot superflus.

- La création d'un bundle pour une instance en cours d'exécution nécessite votre clé privée et votre certificat X.509. Mettez ces éléments et toutes les autres informations d'identification dans un endroit qui n'est pas regroupé (comme par exemple le stockage d'instance).

AMI payantes

Une AMI payante est une AMI que vous pouvez acheter auprès d'un développeur.

Amazon EC2 s'intègre à AWS Marketplace, ce qui permet aux développeurs de facturer l'utilisation de leurs AMI à d'autres utilisateurs Amazon EC2 ou d'assurer un support pour des instances.

AWS Marketplace est une boutique en ligne dans laquelle vous pouvez acheter des logiciels compatibles avec AWS, et notamment les AMI nécessaires pour lancer votre instance EC2. Les AMI AWS Marketplace sont organisées en catégories, par exemple Outils de développement, pour vous permettre de trouver des produits répondant à vos besoins. Pour plus d'informations sur AWS Marketplace, consultez le site AWS Marketplace.

Le lancement d'une instance à partir d'une AMI payante est identique au lancement d'une instance à partir de n'importe quelle AMI. Aucun paramètre supplémentaire n'est obligatoire. La facturation de l'instance correspond aux tarifs définis par le propriétaire de l'AMI, auxquels s'ajoutent les frais d'utilisation standard des sites web associés ; par exemple, le tarif horaire pour l'exécution d'un type d'instance m1.small dans Amazon EC2. Des taxes supplémentaires peuvent également être appliquées. Le propriétaire de l'AMI payante peut confirmer si une instance spécifique a été lancée à l'aide de cette AMI payante.

> **Important**
>
> Amazon DevPay n' accepte plus les nouveaux vendeurs ni les nouveaux produits. AWS Marketplace est désormais l'unique plateforme d'e-commerce en ligne unifiée pour la vente de logiciels et de services via AWS. Pour plus d'informations sur le déploiement et la vente de logiciels à partir d'AWS Marketplace, consultez la page relative à la vente sur AWS Marketplace. AWS Marketplace prend en charge les AMI basées sur des volumes Amazon EBS.

Sommaire

- Vendre votre AMI (p. 129)
- Rechercher une AMI payante (p. 129)
- Acheter une AMI payante (p. 130)
- Obtenir le code produit pour votre instance (p. 130)
- Utiliser le support payant (p. 131)
- Factures pour les AMI payantes et supportées (p. 131)
- Gérer vos abonnements AWS Marketplace (p. 131)

Vendre votre AMI

Vous pouvez vendre votre AMI à l'aide d'AWS Marketplace. AWS Marketplace offre une expérience d'achat organisée. AWS Marketplace prend également en charge les fonctions AWS telles que les AMI basées sur des volumes Amazon EBS, les Instances réservées et les Instances Spot.

Pour plus d'informations sur la vente de votre AMI sur AWS Marketplace, consultez la page relative à la vente sur AWS Marketplace.

Rechercher une AMI payante

Vous pouvez rechercher les AMI disponibles à l'achat de différentes façons. Par exemple, vous pouvez utiliser AWS Marketplace, la console Amazon EC2 ou la ligne de commande. Les développeurs peuvent également vous avertir eux-mêmes de la disponibilité d'AMI payantes.

Recherche d'une AMI payante à l'aide de la console

Pour rechercher une AMI payante à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez AMI.
3. Choisissez Images publiques comme premier filtre.
4. Dans la barre de recherche, choisissez Propriétaire, puis AWS Marketplace.
5. Si vous connaissez le code produit, choisissez Code Produit, puis tapez le code produit.

Rechercher une AMI payante avec AWS Marketplace

Pour rechercher une AMI payante à l'aide d'AWS Marketplace

1. Ouvrir AWS Marketplace.
2. Saisissez le nom du système d'exploitation dans la zone de recherche et cliquez sur Go (Accéder).
3. Pour affiner la recherche, utilisez l'une des catégories ou l'un des filtres.
4. Chaque produit est identifié par son type de produit : `AMI` ou `Software as a Service`.

Rechercher une AMI payante avec l'AWS CLI

Vous pouvez rechercher une AMI payante à l'aide de la commande describe-images suivante (AWS CLI).

```
aws ec2 describe-images
    --owners aws-marketplace
```

Cette commande renvoie un grand nombre d'informations qui décrivent chaque AMI, y compris le code produit d'une AMI payante. Le résultat de `describe-images` comprend une entrée pour le code produit, illustrée ici :

```
"ProductCodes": [
    {
        "ProductCodeId": "product_code",
        "ProductCodeType": "marketplace"
    }
],
```

Si vous connaissez le code produit, vous pouvez filtrer les résultats par code produit. Cet exemple renvoie l'AMI la plus récente ayant le code produit spécifié.

```
aws ec2 describe-images
    --owners aws-marketplace \
    --filters "Name=product-code,Values=product_code" \
    --query "sort_by(Images, &CreationDate)[-1].[ImageId]"
```

Acheter une AMI payante

Vous devez vous inscrire à (acheter) une AMI payante avant de pouvoir lancer une instance à l'aide de l'AMI.

Généralement, le vendeur d'une AMI payante vous présente les informations relatives à l'AMI, notamment le tarif et un lien auquel vous accédez pour l'acheter. Lorsque vous cliquez sur le lien, vous êtes invité à vous connecter à AWS, puis vous pouvez acheter l'AMI.

Acheter une AMI payante avec la console

Vous pouvez acheter une AMI payante à l'aide de l'assistant de lancement Amazon EC2. Pour plus d'informations, consultez Lancement d'une instance AWS Marketplace (p. 557).

S'abonner à un produit avec AWS Marketplace

Pour utiliser AWS Marketplace, vous devez posséder un compte AWS. Pour lancer des instances à partir de produits AWS Marketplace, vous devez être inscrit pour utiliser le service Amazon EC2 et être abonné au produit à partir duquel l'instance va être lancée. Vous pouvez vous abonner aux produits dans AWS Marketplace de deux façons :

- Site web AWS Marketplace : vous pouvez lancer rapidement le logiciel préconfiguré avec la fonctionnalité de déploiement 1-Click.
- Assistant de lancement Amazon EC2 : vous pouvez rechercher une AMI et lancer une instance directement à partir de l'assistant. Pour plus d'informations, consultez Lancement d'une instance AWS Marketplace (p. 557).

Obtenir le code produit pour votre instance

Vous pouvez récupérer le code produit AWS Marketplace pour votre instance à l'aide des métadonnées de l'instance. Pour obtenir plus d'informations sur la récupération des métadonnées, consultez Métadonnées d'instance et données utilisateur (p. 701).

Pour récupérer un code produit, utilisez la commande suivante :

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/product-codes
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/product-codes
```

Si l'instance comporte un code produit, Amazon EC2 le renvoie.

Utiliser le support payant

Amazon EC2 permet également aux développeurs d'offrir un support pour les logiciels (ou AMI dérivées). Les développeurs peuvent créer des produits de support que vous pouvez utiliser en vous y inscrivant. Pendant le processus d'inscription au produit de support, le développeur vous fournit un code produit, que vous devez ensuite associer à votre propre AMI. Le développeur est ainsi en mesure de confirmer que votre instance peut bénéficier du support. Cela garantit également que, lorsque vous exécutez des instances du produit, le tarif appliqué correspond aux conditions définies pour le produit par le développeur.

> Important
>
> Vous ne pouvez pas utiliser un produit de support avec les Instances réservées. Le tarif appliqué est toujours défini par le vendeur du produit de support.

Pour associer un code produit à votre AMI, utilisez l'une des commandes suivantes, dans lesquelles ami_id est l'ID de l'AMI et product_code est le code produit :

- modify-image-attribute (AWS CLI)

```
aws ec2 modify-image-attribute --image-id ami_id --product-codes "product_code"
```

- Edit-EC2ImageAttribute (Outils AWS pour Windows PowerShell)

```
PS C:\> Edit-EC2ImageAttribute -ImageId ami_id -ProductCode product_code
```

Une fois que vous avez défini l'attribut du code produit, il ne peut pas être modifié ni supprimé.

Factures pour les AMI payantes et supportées

A la fin de chaque mois, vous recevez un e-mail indiquant le montant prélevé sur votre carte de paiement pour l'utilisation des AMI payantes ou supportées au cours du mois. Cette facture est différente de votre facture Amazon EC2 habituelle. Pour plus d'informations, consultez Paiement des produits AWS Marketplace.

Gérer vos abonnements AWS Marketplace

Sur le site Web AWS Marketplace, vous pouvez vérifier les informations concernant votre abonnement, consulter les instructions d'utilisation du fournisseur, gérer vos abonnements et plus encore.

Pour vérifier les informations concernant votre abonnement

1. Connectez-vous à AWS Marketplace.
2. Choisissez Your Marketplace Account (Votre compte Marketplace).
3. Choisissez Manage your software subscriptions (Gérer vos abonnements logiciels).
4. Tous vos abonnements actuels sont répertoriés. Choisissez Usage Instructions (Instructions d'utilisation) pour consulter les instructions spécifiques à l'utilisation du produit, par exemple le nom d'utilisateur pour la connexion à votre instance en cours d'exécution.

Pour annuler un abonnement AWS Marketplace

1. Vérifiez que vous avez mis fin à toutes les instances en cours d'exécution à partir de l'abonnement.

   a. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
   b. Dans le volet de navigation, choisissez Instances.
   c. Sélectionnez l'instance et choisissez Actions, État de l'instance, Résilier.
   d. Sélectionnez Oui, résilier lorsque vous êtes invité à confirmer l'opération.
2. Connectez-vous à AWS Marketplace et choisissez Your Marketplace Account (Votre compte Marketplace), puis sur Manage your software subscriptions (Gérer vos abonnements logiciels).
3. Choisissez Cancel subscription (Annuler l'abonnement). Vous êtes invité à confirmer l'annulation.

   Note

   Après avoir annulé votre abonnement, vous ne pouvez plus lancer d'instances à partir de cette AMI. Pour réutiliser cette AMI, vous devez vous y abonner à nouveau, soit sur le site AWS Marketplace, soit via l'assistant de lancement de la console Amazon EC2.

Création d'une AMI Linux basée sur Amazon EBS

Pour créer une AMI basée sur Amazon EBS, démarrez à partir d'une instance que vous avez lancée depuis une AMI Linux existante basée sur Amazon EBS. Il peut s'agir d'une AMI que vous avez obtenue dans AWS Marketplace, que vous avez créée à l'aide d'AWS Server Migration Service ou de VM Import/Export ou bien de toute autre AMI à laquelle vous avez accès. Après avoir personnalisé l'instance pour répondre à vos besoins, créez et enregistrez une nouvelle AMI. Vous pouvez l'utiliser pour lancer de nouvelles instances avec ces personnalisations.

Les procédures décrites ci-dessous s'appliquent aux instances Amazon EC2 sauvegardées sur des volumes Amazon EBS chiffrés (notamment le volume racine) ainsi que pour les volumes non chiffrés.

Le processus de création d'une AMI est différent de celui des AMIs basées sur le stockage d'instance. Pour plus d'informations sur les différences entre les instances basées sur des volumes Amazon EBS et celles basées sur un stockage d'instance et les façons de déterminer le type de périphérique racine pour votre instance, consultez Stockage pour le périphérique racine (p. 106). Pour plus d'informations sur la création d'AMI Linux basées sur le stockage d'instance, consultez le didacticiel Création d'une AMI Linux basée sur le stockage d'instance (p. 136).

Pour plus d'informations sur la création d'une AMI Windows basée sur Amazon EBS, consultez la section Création d'une AMI Windows basée sur des volumes Amazon EBS du manuel Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Présentation de la création d'AMIs basées sur des volumes Amazon EBS

Tout d'abord, lancez une instance depuis une AMI qui est similaire à l'AMI que vous souhaiteriez créer. Vous pouvez vous connecter à votre instance et la personnaliser. Lorsque l'instance est configurée correctement, assurez l'intégrité des données en arrêtant l'instance avant de créer une AMI, puis créez l'image. Lorsque vous créez une AMI basée sur Amazon EBS, nous l'enregistrons automatiquement pour vous.

Amazon EC2 désactive l'instance avant de créer l'AMI pour s'assurer que tout le contenu de l'instance est arrêté et dans un état cohérent pendant le processus de création. Si vous êtes sûr que votre instance est dans un état cohérent approprié pour la création d'une AMI, vous pouvez indiquer à Amazon EC2 de ne pas procéder à la mise hors tension et redémarrer l'instance. Certains systèmes de fichiers, comme XFS, peuvent bloquer et débloquer l'activité ce qui sécurise la création de l'image sans redémarrer l'instance.

Pendant le processus de création d'AMI, Amazon EC2 crée des instantanés du volume racine de votre instance et de tout autre volume EBS attaché à cette dernière. Les instantanés vous sont facturés jusqu'à

ce que vous annuliez l'inscription de l'AMI et que vous les supprimiez. Pour plus d'informations, consultez Annulation de l'inscription de votre AMI Linux (p. 182). Si un volume attaché à l'instance est chiffré, la nouvelle AMI se lance uniquement avec succès sur les instances qui prennent en charge Chiffrement Amazon EBS. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

En fonction de la taille des volumes, le processus de création de l'AMI peut prendre quelques minutes pour se terminer (parfois jusqu'à 24 heures). Il se peut que la création d'instantanés de vos volumes avant de créer votre AMI vous paraisse plus efficace. De cette façon, seuls de petits instantanés incrémentiels doivent être formés lorsque l'AMI est créée, et le processus se termine plus rapidement (la durée totale de la création des instantanés reste la même). Pour plus d'informations, consultez Création d'instantanés Amazon EBS (p. 1132).

Une fois le processus terminé, vous disposez d'une nouvelle AMI et d'un instantané créés à partir du volume racine de l'instance. Lorsque vous lancez une instance à l'aide de la nouvelle AMI, nous créons un nouveau volume EBS pour son volume racine en utilisant l'instantané.

Si vous ajoutez des volumes EBS ou de stockage d'instance à votre instance en plus de du volume du périphérique racine, le mappage de périphérique de stockage en mode bloc pour la nouvelle AMI contient des informations pour ces volumes, et les mappages de périphérique de stockage en mode bloc pour les instances que vous lancez depuis la nouvelle AMI contient automatiquement des informations pour ces volumes. Les volumes de stockage d'instance spécifiés dans le mappage de périphérique de stockage en mode bloc pour la nouvelle instance sont nouveaux et ne contiennent aucune donnée des volumes de stockage d'instance de l'instance que vous avez utilisée pour créer l'AMI. Les données sur les volumes EBS persistent. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289).

> Note
>
> Lorsque vous créez une instance à partir d'une AMI basée sur EBS, vous devez initialiser son volume racine et tout stockage EBS supplémentaire avant de la mettre en production. Pour plus d'informations, consultez Initialisation des volumes Amazon EBS.

Création d'une AMI Linux à partir d'une instance

Vous pouvez créer une AMI à l'aide d'AWS Management Console ou de la ligne de commande. Le graphique suivant résume le processus de création d'une AMI basée sur Amazon EBS à partir d'une instance EC2 en cours d'exécution. Commencez par une AMI existante, lancez une instance, personnalisez-la, créez une autre AMI à partir de cela, puis lancez une instance de la nouvelle AMI. Les étapes du graphique suivant correspondent aux étapes de la procédure ci-dessous.

Pour créer une AMI à partir d'une instance en utilisant la console

1. Sélectionnez une AMI appropriée basée sur EBS à utiliser comme point de départ pour votre nouvelle AMI et configurez-la si nécessaire avant de la lancer. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).
2. Sélectionnez Lancement pour lancer une instance de l'AMI basée sur EBS que vous avez sélectionnée. Acceptez les valeurs par défaut alors que vous suivez les instructions de l'assistant.

Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

3. Pendant que l'instance est en cours d'exécution, connectez-vous à celle-ci. Vous pouvez effectuer toutes les actions suivantes sur votre instance pour la personnaliser afin de répondre à vos besoins :
   - Installer les logiciels et les applications
   - Copier les données
   - Réduire le délai de démarrage en effaçant les fichiers temporaires, en défragmentant le disque dur et en supprimant l'espace libre
   - Attacher des volumes Amazon EBS supplémentaires
4. (Facultatif) Créez des instantanés de l'ensemble des volumes attachés à votre instance. Pour plus d'informations sur la création des instantanés, consultez le didacticiel Création d'instantanés Amazon EBS (p. 1132).
5. Dans le panneau de navigation, choisissez Instances, sélectionnez votre instance, puis choisissez Actions, Image, Créer une image.

   Tip

   Si cette option est désactivée, votre instance n'est pas une instance basée sur Amazon EBS.
6. Dans la boîte de dialogue Créer une image, spécifiez les informations suivantes, puis choisissez Créer une image.
   - Nom de l'image – Nom unique de l'image.
   - Description de l'image – Description facultative de l'image de 255 caractères au maximum.
   - Pas de redémarrage – Cette option n'est pas sélectionnée par défaut. Amazon EC2 met hors tension l'instance, effectue des instantanés des volumes attachés, crée et enregistre l'AMI, puis redémarre l'instance. Sélectionnez Pas de redémarrage pour éviter l'arrêt de l'instance.

     Warning

     Si vous sélectionnez Pas de redémarrage, nous ne pouvons pas garantir l'intégrité du système de fichiers de l'image créée.
   - Volumes d'instance – Les champs de cette section vous permettent de modifier le volume racine et d'ajouter des volumes Amazon EBS et de stockage d'instance supplémentaires. Pour obtenir des informations sur les différents champs, positionnez le pointeur sur l'icône i en regard de chaque champ pour afficher l'info-bulle du champ. Certains points importants sont répertoriés ci-dessous.
     - Pour modifier la taille du volume racine, recherchez Racine dans la colonne Type de volume, puis saisissez la valeur requise pour Taille (Gio).
     - Si vous sélectionnez Supprimer à la résiliation, lorsque vous résiliez l'instance créée à partir de cette AMI, le volume EBS est supprimé. Si vous désélectionnez Supprimer à la résiliation, lorsque vous résiliez l'instance, le volume EBS n'est pas supprimé. Pour plus d'informations, consultez Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649).
     - Pour ajouter un volume Amazon EBS, choisissez Ajouter un volume (ce qui ajoute une ligne). Pour Type de volume, choisissez EBS et remplissez les champs de la ligne. Lorsque vous lancez une instance à partir de votre nouvelle AMI, des volumes supplémentaires sont automatiquement attachés à l'instance. Les volumes vides doivent être formatés et montés. Les volumes basés sur un instantané doivent être montés.
     - Pour ajouter un volume de stockage d'instance, consultez Ajout de volumes de stockage d'instance à une AMI (p. 1273). Lorsque vous lancez une instance à partir de votre nouvelle AMI, les volumes supplémentaires sont automatiquement initialisés et montés. Ces volumes ne contiennent pas les données des volumes de stockage d'instance de l'instance en cours d'exécution sur laquelle vous avez basé votre AMI.
7. Pour afficher le statut de votre AMI pendant sa création, dans le panneau de navigation, choisissez AMI. À l'origine, le statut est `pending` mais il doit être remplacé par `available` après quelques minutes.

(Facultatif) Pour afficher l'instantané qui a été créé pour la nouvelle AMI, choisissez Instantanés. Lorsque vous lancez une instance à partir de cette AMI, nous utilisons cet instantané pour créer son volume du périphérique racine.

8. Lancez une instance à partir de votre nouvelle AMI. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).
9. La nouvelle instance en cours d'exécution contient toutes les personnalisations que vous avez appliquées au cours des étapes précédentes.

Pour créer une AMI à partir d'une instance en utilisant la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-image (AWS CLI)
- New-EC2Image (Outils AWS pour Windows PowerShell)

Création d'une AMI Linux à partir d'un instantané

Si vous avez un instantané du volume du périphérique racine d'une instance, vous pouvez créer une AMI à partir de cet instantané en utilisant la AWS Management Console ou la ligne de commande.

Pour créer une AMI à partir d'un instantané en utilisant la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous Elastic Block Store, sélectionnez Snapshots.
3. Choisissez l'instantané et sélectionnez Actions, Create Image.
4. Dans la boîte de dialogue Create Image from EBS Snapshot, complétez les champs pour créer votre AMI, puis sélectionnez Create. Si vous recréez une instance parente, sélectionnez alors les mêmes options que l'instance parente.
   - Architecture : Sélectionnez i386 pour 32 bits ou x86_64 pour 64 bits.
   - Nom du périphérique racine : Saisissez le nom approprié du volume racine. Pour plus d'informations, consultez Noms d'appareil pour les instances Linux (p. 1287).
   - Virtualization type : Choisissez si les instances lancées à partir de cette AMI utilisent la virtualisation paravirtuelle ou la virtualisation HVM. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108).
   - (Type de virtualisation PV uniquement) ID du noyau et ID de disque RAM : choisissez l'AKI et l'ARI dans les listes. Si vous choisissez l'AKI par défaut ou si vous ne choisissez pas d'AKI, vous devez spécifier une AKI chaque fois que vous lancez une instance à l'aide de cette AMI. Par ailleurs, votre instance peut échouer aux vérifications de l'état si l'AKI par défaut est incompatible avec l'instance.
   - (Facultatif) Block Device Mappings : Ajoutez des volumes ou développez la taille par défaut du volume racine pour l'AMI. Pour plus d'informations sur le redimensionnement du système de fichiers sur votre instance pour un volume plus important, consultez Extension d'un système de fichiers Linux après redimensionnement d'un volume (p. 1190).

Pour créer une AMI à partir d'un instantané en utilisant la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- register-image (AWS CLI)
- Register-EC2Image (Outils AWS pour Windows PowerShell)

Lancement d'une instance à partir d'une AMI que vous avez créée

Vous pouvez lancer une instance à partir d'une AMI que vous avez créée à partir d'une instance ou d'un instantané.

Pour lancer une instance à partir de votre AMI

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous Images, choisissez AMIs (AMI).
3. Définissez le filtre sur Owned by me (M'appartenant) et sélectionnez votre AMI.
4. Choisissez Actions, Launch (Lancer).
5. Suivez l'assistant pour lancer l'instance. Pour de plus amples informations sur chaque étape de l'assistant, veuillez consulter Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Création d'une AMI Linux basée sur le stockage d'instance

L'AMI que vous spécifiez au lancement de votre instance détermine le type de volume du périphérique racine

Pour créer une AMI Linux basée sur le stockage d'instance, démarrez à partir d'une instance que vous avez lancée depuis une AMI Linux basée sur le stockage d'instance existante. Après avoir personnalisé l'instance pour répondre à vos besoins, créez un bundle du volume et inscrivez une nouvelle AMI que vous pouvez utiliser pour lancer de nouvelles instances avec ces personnalisations.

> Important
>
> Seuls les types d'instance suivants prennent en charge un volume de stockage d'instance en tant que périphérique racine : C3, D2, G2, I2, M3 et R3.

Le processus de création d'AMI est différent de celui des AMI basées sur les volumes Amazon EBS. Pour plus d'informations sur les différences entre les instances basées sur des volumes Amazon EBS et celles basées sur un stockage d'instance et les façons de déterminer le type de périphérique racine pour votre instance, consultez Stockage pour le périphérique racine (p. 106). Si vous devez créer une AMI Linux basée sur des volumes Amazon EBS, consultez le didacticiel Création d'une AMI Linux basée sur Amazon EBS (p. 132).

Présentation du processus de création des AMI basées sur le stockage d'instance

Le graphique suivant résume le processus de création d'une AMI à partir d'une instance basée sur le stockage d'instance.

Tout d'abord, lancez une instance depuis une AMI qui est similaire à l'AMI que vous souhaiteriez créer. Vous pouvez vous connecter à votre instance et la personnaliser. Lorsque l'instance est configurée comme vous le voulez, vous pouvezen créer un bundle. Le processus de création d'un bundle peut prendre plusieurs minutes. Après la fin du processus, vous avez un groupe qui se compose d'un manifeste d'image (`image.manifest.xml`) et de fichiers (`image.part.xx`) contenant un modèle pour le volume racine. Ensuite, vous chargez le bundle dans votre compartiment Amazon S3, puis vous inscrivez votre AMI.

Lorsque vous lancez une instance à l'aide de la nouvelle AMI, nous créons le volume racine pour l'instance avec le bundle que vous avez chargé sur Amazon S3. L'espace de stockage utilisé par le bundle dans Amazon S3 entraîne des frais sur votre compte jusqu'à ce que vous le supprimiez. Pour plus d'informations, consultez Annulation de l'inscription de votre AMI Linux (p. 182).

Si vous ajoutez des volumes de stockage d'instance à votre instance en plus de votre volume du périphérique racine, le mappage de périphérique de stockage en mode bloc pour la nouvelle AMI contient des informations pour ces volumes et les mappages de périphérique de stockage en mode bloc pour les instances que vous lancez depuis la nouvelle AMI contient automatiquement des informations pour ces volumes. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289).

Prérequis

Avant de créer un AMI, vous devez terminer les tâches suivantes :

- Installez les outils AMI. Pour plus d'informations, consultez Installation des outils AMI (p. 138).
- Installez l'AWS CLI. Pour plus d'informations, consultez la page Préparation de l'installation de l'AWS Command Line Interface.
- Assurez-vous d'avoir un compartiment Amazon S3 pour le bundle. Pour créer un compartiment Amazon S3, ouvrez la console Amazon S3 et cliquez sur Créer un compartiment. Vous pouvez également utiliser la commande mb de l'AWS CLI.
- Assurez-vous d'avoir un ID de compte AWS. Pour plus d'informations, consultez Identificateurs de compte AWS dans le document AWS General Reference.
- Assurez-vous d'avoir vos ID de clé d'accès et clé d'accès secrète. Pour plus d'informations, consultez Clés d'accès dans le document AWS General Reference.
- Assurez-vous d'avoir un certificat X.509 et la clé privée correspondante.
  - Si vous avez besoin créer un certificat X.509, consultez la section Gestion des certificats de signature (p. 140). Le certificat X.509 et la clé privée sont utilisés pour chiffrer et déchiffrer votre AMI.
  - [Chine (Pékin)] Utilisez le certificat `$EC2_AMITOOL_HOME/etc/ec2/amitools/cert-ec2-cn-north-1.pem`.

- [AWS GovCloud (US-West)] Utilisez le certificat `$EC2_AMITOOL_HOME/etc/ec2/amitools/cert-ec2-gov.pem`.
- Connectez-vous à votre instance et personnalisez-la. Par exemple, vous pouvez installer des logiciels et des applications, copier des données, supprimer des fichiers temporaires et modifier la configuration Linux.

Tâches

- Installation des outils AMI (p. 138)
- Création d'une AMI à partir d'une instance Amazon Linux basée sur le stockage d'instance (p. 141)
- Création d'une AMI à partir d'une instance Ubuntu basée sur le stockage d'instance (p. 144)
- Conversion d'une AMI basée sur un stockage d'instance en AMI basée sur des volumes Amazon EBS (p. 148)

Installation des outils AMI

Vous pouvez utiliser les outils AMI pour créer et gérer des AMIs Linux basées sur le stockage d'instance. Pour utiliser ces outils, vous devez les installer sur votre instance Linux. Les outils AMI sont disponibles sous forme de fichiers RPM et .zip pour les distributions Linux ne prenant pas en charge RPM.

Pour installer les outils AMI à l'aide d'un fichier RPM

1. Installez Ruby en utilisant le gestionnaire de package pour votre distribution de Linux, par exemple yum. Par exemple :

```
[ec2-user ~]$ sudo yum install -y ruby
```

2. Téléchargez le fichier RPM à l'aide d'un outil tel que wget ou curl. Par exemple :

```
[ec2-user ~]$ wget https://s3.amazonaws.com/ec2-downloads/ec2-ami-tools.noarch.rpm
```

3. Vérifiez la signature du fichier RPM en utilisant la commande suivante :

```
[ec2-user ~]$ rpm -K ec2-ami-tools.noarch.rpm
```

   La commande ci-dessus doit indiquer que les hachages SHA1 et MD5 du fichier sont `OK.` Si la commande indique que les hachages sont `NOT OK`, utilisez la commande suivante pour afficher les hachages SHA1 et MD5 d'en-tête du fichier :

```
[ec2-user ~]$ rpm -Kv ec2-ami-tools.noarch.rpm
```

   Ensuite, comparez les hachages SHA1 et MD5 d'en-tête du fichier avec les hachages des outils d'AMI vérifiés suivants pour confirmer l'authenticité du fichier :

   - SHA1 d'en-tête : a1f662d6f25f69871104e6a62187fa4df508f880
   - MD5 : 9faff05258064e2f7909b66142de6782

   Si les hachages SHA1 et MD5 d'en-tête du fichier correspondent aux hachages des outils d'AMI vérifiés, passez à l'étape suivante.

4. Installez le fichier RPM à l'aide de la commande suivante:

```
[ec2-user ~]$ sudo yum install ec2-ami-tools.noarch.rpm
```

5. Vérifiez l'installation de vos outils AMI avec la commande ec2-ami-tools-version (p. 152).

```
[ec2-user ~]$ ec2-ami-tools-version
```

Note

Si vous recevez une erreur de chargement comme « cannot load such file -- ec2/amitools/version (LoadError) », terminez l'étape suivante pour ajouter l'emplacement d'installation de vos outils AMI dans votre chemin d'accès `RUBYLIB`.

6. (Facultatif) Si vous avez reçu une erreur à l'étape précédente, ajoutez l'emplacement d'installation de vos outils AMI pour votre chemin d'accès `RUBYLIB`.

   a. Exécutez la commande suivante afin de déterminer les chemins à ajouter.

```
[ec2-user ~]$ rpm -qil ec2-ami-tools | grep ec2/amitools/version
/usr/lib/ruby/site_ruby/ec2/amitools/version.rb
/usr/lib64/ruby/site_ruby/ec2/amitools/version.rb
```

   Dans l'exemple ci-dessus, le fichier manquant à partir de l'erreur de chargement précédente est situé aux emplacements `/usr/lib/ruby/site_ruby` et `/usr/lib64/ruby/site_ruby`.

   b. Ajoutez les emplacements à partir de l'étape précédente pour votre chemin d'accès. `RUBYLIB`

```
[ec2-user ~]$ export RUBYLIB=$RUBYLIB:/usr/lib/ruby/site_ruby:/usr/lib64/ruby/
site_ruby
```

   c. Vérifiez l'installation de vos outils AMI avec la commande ec2-ami-tools-version (p. 152).

```
[ec2-user ~]$ ec2-ami-tools-version
```

Pour installer les outils AMI à l'aide du fichier .zip

1. Installez Ruby et décompressez en utilisant le gestionnaire de package pour votre distribution de Linux, comme apt-get. Par exemple :

```
[ec2-user ~]$ sudo apt-get update -y && sudo apt-get install -y ruby unzip
```

2. Téléchargez le fichier .zip à l'aide d'un outil tel que wget ou curl. Par exemple :

```
[ec2-user ~]$ wget https://s3.amazonaws.com/ec2-downloads/ec2-ami-tools.zip
```

3. Décompressez les fichiers dans un répertoire d'installation approprié, tel que `/usr/local/ec2`.

```
[ec2-user ~]$ sudo mkdir -p /usr/local/ec2
$ sudo unzip ec2-ami-tools.zip -d /usr/local/ec2
```

   Notez que le fichier .zip contient un dossier ec2-ami-tools-*x.x.x*, où *x.x.x* correspond au numéro de version des outils (par exemple, `ec2-ami-tools-1.5.7`).

4. Définissez la variable d'environnement `EC2_AMITOOL_HOME` sur le répertoire d'installation des outils. Par exemple :

```
[ec2-user ~]$ export EC2_AMITOOL_HOME=/usr/local/ec2/ec2-ami-tools-x.x.x
```

5. Ajoutez les outils à votre variable d'environnement `PATH`. Par exemple :

```
[ec2-user ~]$ export PATH=$EC2_AMITOOL_HOME/bin:$PATH
```

6. Vous pouvez vérifier l'installation de vos outils AMI avec la commande ec2-ami-tools-version (p. 152).

```
[ec2-user ~]$ ec2-ami-tools-version
```

Gestion des certificats de signature

Certaines commandes dans les outils AMI nécessitent un certificat de signature (également appelé certificat X.509). Vous devez créer et importer le certificat et le télécharger dans AWS. Par exemple, vous pouvez utiliser un outil tiers tel que OpenSSL pour créer le certificat.

Pour créer un certificat de signature

1. Installer et configurer OpenSSL.
2. Créez une clé privée à l'aide de la commande `openssl genrsa` et enregistrez la sortie dans un fichier `.pem`. Nous vous recommandons de créer une clé RSA 2048 bits ou 4096 bits.

```
openssl genrsa 2048 > private-key.pem
```

3. Générez un certificat à l'aide de la commande `openssl req`.

```
openssl req -new -x509 -nodes -sha256 -days 365 -key private-key.pem -outform PEM -
out certificate.pem
```

Pour charger le certificat dans AWS, utilisez la commande upload-signing-certificate.

```
aws iam upload-signing-certificate --user-name user-name --certificate-body file://path/to/
certificate.pem
```

Pour afficher les certificats d'un utilisateur, utilisez la commande list-signing-certificates :

```
aws iam list-signing-certificates --user-name user-name
```

Pour désactiver ou réactiver un certificat de signature pour un utilisateur, utilisez la commande update-signing-certificate. La commande suivante désactive le certificat :

```
aws iam update-signing-certificate --certificate-id OFHPLP4ZULTHYPMSYEX7O4BEXAMPLE --
status Inactive --user-name user-name
```

Pour supprimer un certificat, utilisez la commande delete-signing-certificate :

```
aws iam delete-signing-certificate --user-name user-name --certificate-
id OFHPLP4ZULTHYPMSYEX7O4BEXAMPLE
```

Création d'une AMI à partir d'une instance basée sur le stockage d'instance

Les procédures suivantes sont destinées à la création d'une AMI basée sur le stockage d'instance à partir d'une instance basée sur le stockage d'instance. Avant de commencer, prenez connaissance des prérequis (p. 137).

Rubriques

- Création d'une AMI à partir d'une instance Amazon Linux basée sur le stockage d'instance (p. 141)
- Création d'une AMI à partir d'une instance Ubuntu basée sur le stockage d'instance (p. 144)

Création d'une AMI à partir d'une instance Amazon Linux basée sur le stockage d'instance

Cette section décrit la création d'une AMI à partir d'une instance Amazon Linux. Il est possible que les procédures suivantes ne fonctionnent pas pour les instances exécutant d'autres distributions Linux. Pour les procédures spécifiques à Ubuntu, consultez Création d'une AMI à partir d'une instance Ubuntu basée sur le stockage d'instance (p. 144).

Pour se préparer à utiliser les outils AMI (instances HVM uniquement)

1. Les outils AMI ont besoin de GRUB Legacy pour démarrer correctement. Utilisez la commande suivante pour installer GRUB :

```
[ec2-user ~]$ sudo yum install -y grub
```

2. Installez les packages de gestion de partition à l'aide de la commande suivante :

```
[ec2-user ~]$ sudo yum install -y gdisk kpartx parted
```

Pour créer une AMI à partir d'une instance Amazon Linux basée sur un stockage d'instance

Cette procédure part du principe que vous avez respecté les prérequis dans Prérequis (p. 137).

1. Chargez vos informations d'identification sur votre instance. Nous utilisons ces informations d'identification pour garantir que seuls vous et Amazon EC2 peuvent accéder à votre AMI.

   a. Créez un répertoire temporaire sur votre instance pour vos informations d'identification en suivant ce qui suit :

   ```
   [ec2-user ~]$ mkdir /tmp/cert
   ```

   Ceci vous permet d'exclure vos informations d'identification de l'image créée.

   b. Copiez votre certificat X.509 et votre clé privée correspondante depuis votre ordinateur vers le répertoire `/tmp/cert` de votre instance en utilisant un outil de copie sécurisé tel que scp (p. 603). L'option `-i my-private-key.pem` de la commande scp suivante est la clé privée que vous utilisez pour vous connecter à votre instance avec SSH, et non la clé privée X.509. Par exemple :

   ```
   you@your_computer:~ $ scp -i my-private-key.pem /
   path/to/pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem /
   ```

```
path/to/cert-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem ec2-
user@ec2-203-0-113-25.compute-1.amazonaws.com:/tmp/cert/
pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem  100%  717     0.7KB/s   00:00
cert-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem  100%  685     0.7KB/s   00:00
```

Sinon, étant donné qu'il s'agit de fichiers de texte brut, vous pouvez ouvrir le certificat et la clé dans un éditeur de texte et copier leur contenu dans de nouveaux fichiers dans le répertoire `/tmp/cert`.

2. Préparez le bundle à charger sur Amazon S3 en exécutant la commande ec2-bundle-vol (p. 154) depuis votre instance. Assurez-vous de spécifier l'option `-e` pour exclure le répertoire où vos informations d'identification sont stockées. Par défaut, la création d'un bundle exclut les fichiers qui peuvent contenir des informations sensibles. Ces fichiers incluent `*.sw`, `*.swo`, `*.swp`, `*.pem`, `*.priv`, `*id_rsa*`, `*id_dsa*` `*.gpg`, `*.jks`, `*/.ssh/authorized_keys` et `*/.bash_history`. Pour inclure tous ces fichiers, utilisez l'option `--no-filter`. Pour inclure certains de ces fichiers, utilisez l'option `--include`.

   Important

   Par défaut, le processus de création d'un bundle d'AMI génère un ensemble de fichiers compressés et chiffrés dans le répertoire `/tmp` qui représente le volume racine. Si vous n'avez pas suffisamment d'espace disque libre dans `/tmp` pour stocker le groupe, vous devez spécifier un emplacement différent pour qu'il soit stocké avec l'option `-d` *`/path/to/bundle/storage`*. Certaines instances possèdent un stockage éphémère monté à l'emplacement `/mnt` ou `/media/ephemeral0` que vous pouvez utiliser, ou vous pouvez aussi créer (p. 1107), attacher (p. 1109) et monter (p. 1113) un nouveau volume Amazon EBS pour stocker le groupe.

   a. Vous devez exécuter la commande ec2-bundle-vol en tant que racine. Pour la plupart des commandes, vous pouvez utiliser sudo afin d'obtenir des autorisations d'un niveau élevé, mais dans ce cas, vous devriez exécuter sudo -E su pour conserver vos variables d'environnement.

   ```
   [ec2-user ~]$ sudo -E su
   ```

   Notez que l'invite de commande de Bash vous identifie maintenant en tant qu'utilisateur racine, et que le signe dollar a été remplacé par un hashtag, ce qui indique que vous êtes dans un shell racine :

   ```
   [root ec2-user]#
   ```

   b. Pour créer le bundle AMI, exécutez la commande ec2-bundle-vol (p. 154) comme suit :

   ```
   [root ec2-user]# ec2-bundle-vol -k /tmp/cert/pk-
   HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -c /tmp/cert/cert-
   HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -u 123456789012 -r x86_64 -e /tmp/cert --
   partition gpt
   ```

   Note

   Pour les régions Chine (Pékin) et AWS GovCloud (US-West), utilisez le paramètre `--ec2cert` et spécifiez les certificats en fonction des conditions préalables (p. 137).

   La création de l'image peut prendre quelques minutes. Lorsque cette commande se termine, le répertoire `/tmp` (ou votre répertoire personnalisé) contient le groupe (`image.manifest.xml`, ainsi que plusieurs fichiers `image.part.`*`xx`*).

   c. Quittez le shell racine.

```
[root ec2-user]# exit
```

3. (Facultatif) Pour ajouter davantage de volumes de stockage d'instance, modifiez les mappages de périphérique de stockage en mode bloc dans le fichier `image.manifest.xml` pour votre AMI. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289).

   a. Créez une sauvegarde de votre fichier `image.manifest.xml`.

   ```
   [ec2-user ~]$ sudo cp /tmp/image.manifest.xml /tmp/image.manifest.xml.bak
   ```

   b. Reformatez le fichier `image.manifest.xml` pour qu'il soit plus facile à lire et à modifier.

   ```
   [ec2-user ~]$ sudo xmllint --format /tmp/image.manifest.xml.bak > sudo /tmp/
   image.manifest.xml
   ```

   c. Modifiez les mappages de périphérique de stockage en mode bloc dans `image.manifest.xml` avec un éditeur de texte. L'exemple ci-dessous montre une nouvelle entrée pour le volume de stockage d'instance `ephemeral1`.

   > **Note**
   >
   > Pour obtenir la liste des fichiers exclus, consultez ec2-bundle-vol (p. 154).

   ```
   <block_device_mapping>
     <mapping>
       <virtual>ami</virtual>
       <device>sda</device>
     </mapping>
     <mapping>
       <virtual>ephemeral0</virtual>
       <device>sdb</device>
     </mapping>
     <mapping>
       <virtual>ephemeral1</virtual>
       <device>sdc</device>
     </mapping>
     <mapping>
       <virtual>root</virtual>
       <device>/dev/sda1</device>
     </mapping>
   </block_device_mapping>
   ```

   d. Enregistrez le fichier `image.manifest.xml` et quittez votre éditeur de texte.

4. Pour charger votre bundle sur Amazon S3, exécutez la commande ec2-upload-bundle (p. 165) comme suit.

   ```
   [ec2-user ~]$ ec2-upload-bundle -b my-s3-bucket/bundle_folder/bundle_name -m /tmp/
   image.manifest.xml -a your_access_key_id -s your_secret_access_key
   ```

   > **Important**
   >
   > Pour inscrire votre AMI dans une région autre que USA Est (Virginie du Nord), vous devez spécifier à la fois la région cible avec l'option `--region` et un chemin de compartiment qui existe déjà dans la région cible ou un chemin de compartiment unique qui peut être créé dans la région cible.

5. (Facultatif) Une fois que le groupe est chargé sur Amazon S3, vous pouvez le supprimer du répertoire `/tmp` sur l'instance en utilisant la commande rm suivante :

```
[ec2-user ~]$ sudo rm /tmp/image.manifest.xml /tmp/image.part.* /tmp/image
```

Important

Si vous avez spécifié un chemin avec l'option `-d /path/to/bundle/storage` dans Step 2 (p. 142), utilisez ce chemin à la place de `/tmp`.

6. Pour inscrire votre AMI, exécutez la commande register-image comme suit.

```
[ec2-user ~]$ aws ec2 register-image --image-location my-s3-
bucket/bundle_folder/bundle_name/image.manifest.xml --name AMI_name --virtualization-
type hvm
```

Important

Si vous avez précédemment spécifié une région pour la commande ec2-upload-bundle (p. 165), spécifiez de nouveau cette région pour cette commande.

Création d'une AMI à partir d'une instance Ubuntu basée sur le stockage d'instance

Cette section décrit la création d'une AMI à partir d'une instance Ubuntu Linux avec un volume de stockage d'instance comme volume racine. Il est possible que les procédures suivantes ne fonctionnent pas pour les instances exécutant d'autres distributions Linux. Pour les procédures spécifiques à Amazon Linux, consultez Création d'une AMI à partir d'une instance Amazon Linux basée sur le stockage d'instance (p. 141).

Pour se préparer à utiliser les outils AMI (instances HVM uniquement)

Les outils AMI ont besoin de GRUB Legacy pour démarrer correctement. Toutefois, Ubuntu est configuré pour utiliser GRUB 2. Vous devez vérifier si votre instance utilise GRUB Legacy. Si non, vous devez l'installer et le configurer.

Les instances HVM ont également besoin que des outils de partitionnement soient installés pour que les outils AMI fonctionnent bien.

1. GRUB Legacy (version 0.9*x* ou inférieure) doit être installé sur votre instance. Vérifiez si GRUB Legacy est présent et installez-le si nécessaire.

   a. Vérifiez la version de votre installation de GRUB.

   ```
   ubuntu:~$ grub-install --version
   grub-install (GRUB) 1.99-21ubuntu3.10
   ```

   Dans cet exemple, la version de GRUB est supérieure à 0.9*x*. Donc, GRUB Legacy doit être installée. Passez à l'Step 1.b (p. 144). Si GRUB Legacy est déjà présent, vous pouvez passer à l'Step 2 (p. 144).

   b. Installez le package `grub` à l'aide de la commande suivante.

   ```
   ubuntu:~$ sudo apt-get install -y grub
   ```

2. Installez les packages suivants de gestion de partition en utilisant le gestionnaire de package pour votre distribution.

   - `gdisk` (certaines distributions peuvent appeler ce package `gptfdisk` à la place)

- `kpartx`
- `parted`

Utilisez la commande suivante.

```
ubuntu:~$ sudo apt-get install -y gdisk kpartx parted
```

3. Vérifiez les paramètres du noyau pour votre instance.

```
ubuntu:~$ cat /proc/cmdline
BOOT_IMAGE=/boot/vmlinuz-3.2.0-54-virtual root=UUID=4f392932-ed93-4f8f-
aee7-72bc5bb6ca9d ro console=ttyS0 xen_emul_unplug=unnecessary
```

Notez les options suivant les paramètres du noyau et du périphérique racine : `ro`, `console=ttyS0` et `xen_emul_unplug=unnecessary`. Vos options peuvent différer.

4. Vérifiez les entrées du noyau dans `/boot/grub/menu.lst`.

```
ubuntu:~$ grep ^kernel /boot/grub/menu.lst
kernel  /boot/vmlinuz-3.2.0-54-virtual root=LABEL=cloudimg-rootfs ro console=hvc0
kernel  /boot/vmlinuz-3.2.0-54-virtual root=LABEL=cloudimg-rootfs ro single
kernel  /boot/memtest86+.bin
```

Notez que le paramètre `console` pointe vers `hvc0` au lieu de `ttyS0` et qu'il manque le paramètre `xen_emul_unplug=unnecessary`. Encore une fois, vos options peuvent différer.

5. Modifiez le fichier `/boot/grub/menu.lst` avec votre éditeur de texte préféré (comme vim ou nano) pour changer la console et ajoutez les paramètres que vous avez identifiés précédemment aux entrées de démarrage.

```
title           Ubuntu 12.04.3 LTS, kernel 3.2.0-54-virtual
root            (hd0)
kernel          /boot/vmlinuz-3.2.0-54-virtual root=LABEL=cloudimg-rootfs
 ro console=ttyS0 xen_emul_unplug=unnecessary
initrd          /boot/initrd.img-3.2.0-54-virtual

title           Ubuntu 12.04.3 LTS, kernel 3.2.0-54-virtual (recovery mode)
root            (hd0)
kernel          /boot/vmlinuz-3.2.0-54-virtual root=LABEL=cloudimg-rootfs ro
 single console=ttyS0 xen_emul_unplug=unnecessary
initrd          /boot/initrd.img-3.2.0-54-virtual

title           Ubuntu 12.04.3 LTS, memtest86+
root            (hd0)
kernel          /boot/memtest86+.bin
```

6. Vérifiez que les entrées de votre noyau contiennent maintenant les bons paramètres.

```
ubuntu:~$ grep ^kernel /boot/grub/menu.lst
kernel  /boot/vmlinuz-3.2.0-54-virtual root=LABEL=cloudimg-rootfs ro console=ttyS0
 xen_emul_unplug=unnecessary
kernel  /boot/vmlinuz-3.2.0-54-virtual root=LABEL=cloudimg-rootfs ro single
 console=ttyS0 xen_emul_unplug=unnecessary
kernel  /boot/memtest86+.bin
```

7. [Pour les versions Ubuntu 14.04 et ultérieures uniquement] Depuis la version Ubuntu 14.04, les AMI Ubuntu basées sur un stockage d'instance utilisent une table de partition GPT et une partition EFI séparée montée à l'emplacement `/boot/efi`. La commande ec2-bundle-vol ne regroupe pas

cette partition de démarrage. Vous devez donc mettre en commentaire l'entrée `/etc/fstab` pour la partition EFI comme indiqué dans l'exemple suivant.

```
LABEL=cloudimg-rootfs   /        ext4   defaults        0 0
#LABEL=UEFI     /boot/efi       vfat    defaults        0 0
/dev/xvdb       /mnt    auto    defaults,nobootwait,comment=cloudconfig 0       2
```

Pour créer une AMI à partir d'une instance Ubuntu basée sur un stockage d'instance

Cette procédure part du principe que vous avez respecté les prérequis dans Prérequis (p. 137).

1. Chargez vos informations d'identification sur votre instance. Nous utilisons ces informations d'identification pour garantir que seuls vous et Amazon EC2 peuvent accéder à votre AMI.

   a. Créez un répertoire temporaire sur votre instance pour vos informations d'identification en suivant ce qui suit :

   ```
   ubuntu:~$ mkdir /tmp/cert
   ```

   Ceci vous permet d'exclure vos informations d'identification de l'image créée.

   b. Copiez votre certificat X.509 et votre clé privée depuis votre ordinateur vers le répertoire `/tmp/cert` sur votre instance en utilisant un outil de copie sécurisé comme scp (p. 603). L'option `-i my-private-key.pem` de la commande scp suivante est la clé privée que vous utilisez pour vous connecter à votre instance avec SSH, et non la clé privée X.509. Par exemple :

   ```
   you@your_computer:~ $ scp -i my-private-key.pem /
   path/to/pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem /
   path/to/cert-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem ec2-
   user@ec2-203-0-113-25.compute-1.amazonaws.com:/tmp/cert/
   pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem  100%  717     0.7KB/s   00:00
   cert-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem  100%  685     0.7KB/s   00:00
   ```

   Sinon, étant donné qu'il s'agit de fichiers de texte brut, vous pouvez ouvrir le certificat et la clé dans un éditeur de texte et copier leur contenu dans de nouveaux fichiers dans le répertoire `/tmp/cert`.

2. Préparez le chargement du bundle sur Amazon S3 en exécutant la commande ec2-bundle-vol (p. 154) depuis votre instance. Assurez-vous de spécifier l'option `-e` pour exclure le répertoire où vos informations d'identification sont stockées. Par défaut, la création d'un bundle exclut les fichiers qui peuvent contenir des informations sensibles. Ces fichiers incluent `*.sw`, `*.swo`, `*.swp`, `*.pem`, `*.priv`, `*id_rsa*`, `*id_dsa* *.gpg`, `*.jks`, `*/.ssh/authorized_keys` et `*/.bash_history`. Pour inclure tous ces fichiers, utilisez l'option `--no-filter`. Pour inclure certains de ces fichiers, utilisez l'option `--include`.

   Important

   Par défaut, le processus de création d'un bundle d'AMI génère un ensemble de fichiers compressés et chiffrés dans le répertoire `/tmp` qui représente le volume racine. Si vous n'avez pas suffisamment d'espace disque libre dans `/tmp` pour stocker le groupe, vous devez spécifier un emplacement différent pour qu'il soit stocké avec l'option `-d /path/to/bundle/storage`. Certaines instances possèdent un stockage éphémère monté à l'emplacement `/mnt` ou `/media/ephemeral0` que vous pouvez utiliser, ou vous pouvez aussi créer (p. 1107), attacher (p. 1109) et monter (p. 1113) un nouveau volume Amazon EBS pour stocker le groupe.

a. Vous devez exécuter la commande ec2-bundle-vol en tant que racine. Pour la plupart des commandes, vous pouvez utiliser sudo afin d'obtenir des autorisations d'un niveau élevé, mais dans ce cas, vous devriez exécuter sudo -E su pour conserver vos variables d'environnement.

```
ubuntu:~$ sudo -E su
```

Notez que l'invite de commande de Bash vous identifie maintenant en tant qu'utilisateur racine, et que le signe dollar a été remplacé par un hashtag, ce qui indique que vous êtes dans un shell racine :

```
root@ubuntu:#
```

b. Pour créer le bundle AMI, exécutez la commande ec2-bundle-vol (p. 154) comme suit.

```
root@ubuntu:# ec2-bundle-vol -k /tmp/cert/pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem
 -c /tmp/cert/cert-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -u your_aws_account_id -r
 x86_64 -e /tmp/cert --partition gpt
```

> **Important**
>
> Pour les instances HVM de versions Ubuntu 14.04 et ultérieures, ajoutez l'indicateur `--partition mbr` pour regrouper correctement les instructions de démarrage. Sinon, votre AMI nouvellement créée ne démarrera pas.

La création de l'image peut prendre quelques minutes. Lorsque cette commande se termine, le répertoire `tmp` contient le groupe (`image.manifest.xml`, ainsi que plusieurs fichiers `image.part.xx`).

c. Quittez le shell racine.

```
root@ubuntu:# exit
```

3. (Facultatif) Pour ajouter davantage de volumes de stockage d'instance, modifiez les mappages de périphérique de stockage en mode bloc dans le fichier `image.manifest.xml` pour votre AMI. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289).

a. Créez une sauvegarde de votre fichier `image.manifest.xml`.

```
ubuntu:~$ sudo cp /tmp/image.manifest.xml /tmp/image.manifest.xml.bak
```

b. Reformatez le fichier `image.manifest.xml` pour qu'il soit plus facile à lire et à modifier.

```
ubuntu:~$ sudo xmllint --format /tmp/image.manifest.xml.bak > /tmp/
image.manifest.xml
```

c. Modifiez les mappages de périphérique de stockage en mode bloc dans `image.manifest.xml` avec un éditeur de texte. L'exemple ci-dessous montre une nouvelle entrée pour le volume de stockage d'instance `ephemeral1`.

```
    <block_device_mapping>
      <mapping>
        <virtual>ami</virtual>
        <device>sda</device>
      </mapping>
      <mapping>
        <virtual>ephemeral0</virtual>
        <device>sdb</device>
```

```
    </mapping>
    <mapping>
      <virtual>ephemeral1</virtual>
      <device>sdc</device>
    </mapping>
    <mapping>
      <virtual>root</virtual>
      <device>/dev/sda1</device>
    </mapping>
  </block_device_mapping>
```

d. Enregistrez le fichier `image.manifest.xml` et quittez votre éditeur de texte.

4. Pour charger votre bundle sur Amazon S3, exécutez la commande ec2-upload-bundle (p. 165) comme suit.

```
ubuntu:~$ ec2-upload-bundle -b my-s3-bucket/bundle_folder/bundle_name -m /tmp/
image.manifest.xml -a your_access_key_id -s your_secret_access_key
```

Important

Si vous tentez d'inscrire votre AMI dans une région autre que USA Est (Virginie du Nord), vous devez spécifier à la fois la région cible avec l'option `--region` et un chemin de compartiment qui existe déjà dans la région cible ou un chemin de compartiment unique qui peut être créé dans la région cible.

5. (Facultatif) Une fois que le groupe est chargé sur Amazon S3, vous pouvez le supprimer du répertoire `/tmp` sur l'instance en utilisant la commande rm suivante :

```
ubuntu:~$ sudo rm /tmp/image.manifest.xml /tmp/image.part.* /tmp/image
```

Important

Si vous avez spécifié un chemin avec l'option `-d /path/to/bundle/storage` dans Step 2 (p. 146), utilisez ce même chemin ci-dessous à la place de `/tmp`.

6. Pour inscrire votre AMI, exécutez la commande AWS CLI register-image comme suit.

```
ubuntu:~$ aws ec2 register-image --image-location my-s3-
bucket/bundle_folder/bundle_name/image.manifest.xml --name AMI_name --virtualization-
type hvm
```

Important

Si vous avez précédemment spécifié une région pour la commande ec2-upload-bundle (p. 165), spécifiez de nouveau cette région pour cette commande.

7. [Pour les versions Ubuntu 14.04 et supérieures] Supprimez la mise en commentaire de l'entrée EFI dans `/etc/fstab`. Sinon, votre instance en cours d'exécution ne pourra pas redémarrer.

Conversion d'une AMI basée sur un stockage d'instance en AMI basée sur des volumes Amazon EBS

Vous pouvez convertir une AMI Linux basée sur un stockage d'instance que vous possédez en AMI basée sur des volumes Amazon EBS.

> **Important**
>
> Vous ne pouvez pas convertir une AMI Windows basée sur un stockage d'instance en AMI Windows basée sur des volumes Amazon EBS, et vous ne pouvez pas convertir une AMI que vous ne possédez pas.

Pour convertir une AMI basée sur le stockage d'instance en une AMI basée sur des volumes Amazon EBS

1. Lancez une instance Amazon Linux à partir d'une AMI basée sur des volumes Amazon EBS. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530). L'AWS CLI et les outils AMI sont pré-installés sur les instances Amazon Linux.
2. Chargez la clé privée X.509 que vous avez utilisée pour grouper votre AMI basée sur le stockage d'instance vers votre instance. Nous utilisons cette clé pour garantir que seuls vous et Amazon EC2 peuvent accéder à votre AMI.

   a. Créez un répertoire temporaire sur votre instance pour votre clé privée X.509 en suivant ce qui suit :

   ```
   [ec2-user ~]$ mkdir /tmp/cert
   ```

   b. Copiez votre clé privée X.509 depuis votre ordinateur vers le répertoire `/tmp/cert` de votre instance en utilisant un outil de copie sécurisé comme scp (p. 603). Le paramètre *`my-private-key`* de la commande suivante est la clé privée que vous utilisez pour vous connecter à votre instance avec SSH. Par exemple:

   ```
   you@your_computer:~ $ scp -i my-private-key.pem /
   path/to/pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem ec2-
   user@ec2-203-0-113-25.compute-1.amazonaws.com:/tmp/cert/
   pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem  100%  717     0.7KB/s   00:00
   ```

3. Définissez des variables d'environnement pour votre clé d'accès et votre clé secrète AWS.

   ```
   [ec2-user ~]$ export AWS_ACCESS_KEY_ID=your_access_key_id
   [ec2-user ~]$ export AWS_SECRET_ACCESS_KEY=your_secret_access_key
   ```

4. Préparez un volume Amazon EBS pour votre nouvelle AMI.

   a. Créez un volume Amazon EBS vide dans la même zone de disponibilité que votre instance en utilisant la commande create-volume. Notez l'ID du volume dans la sortie de la commande.

   > **Important**
   >
   > Ce volume Amazon EBS doit avoir la même taille ou une taille plus importante que le volume racine de stockage d'instance original.

   ```
   [ec2-user ~]$ aws ec2 create-volume --size 10 --region us-west-2 --availability-
   zone us-west-2b
   ```

   b. Attachez le volume à votre instance basée sur Amazon EBS en utilisant la commande attach-volume.

   ```
   [ec2-user ~]$ aws ec2 attach-volume --volume-id volume_id --instance-id instance_id
    --device /dev/sdb --region us-west-2
   ```

5. Créez un dossier pour votre groupe.

   ```
   [ec2-user ~]$ mkdir /tmp/bundle
   ```

6. Téléchargez le groupe pour votre AMI basée sur le stockage d'instance sur `/tmp/bundle` en utilisant la commande ec2-download-bundle (p. 160).

```
[ec2-user ~]$ ec2-download-bundle -b my-s3-bucket/bundle_folder/bundle_name -m
 image.manifest.xml -a $AWS_ACCESS_KEY_ID -s $AWS_SECRET_ACCESS_KEY --privatekey /path/
to/pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -d /tmp/bundle
```

7. Reconstituez le fichier image à partir du groupe en utilisant la commande ec2-unbundle (p. 164).

   a. Déplacez les répertoires vers le dossier du groupe.

   ```
   [ec2-user ~]$ cd /tmp/bundle/
   ```

   b. Exécutez la commande ec2-unbundle (p. 164).

   ```
   [ec2-user bundle]$ ec2-unbundle -m image.manifest.xml --privatekey /path/to/pk-
   HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem
   ```

8. Copiez les fichiers depuis l'image dégroupée vers le nouveau volume Amazon EBS.

```
[ec2-user bundle]$ sudo dd if=/tmp/bundle/image of=/dev/sdb bs=1M
```

9. Examinez le volume pour voir si de nouvelles partitions ont été dégroupées.

```
[ec2-user bundle]$ sudo partprobe /dev/sdb1
```

10. Affichez les périphériques de stockage en mode bloc pour trouver le nom du périphérique à monter.

```
[ec2-user bundle]$ lsblk
NAME         MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
/dev/sda    202:0    0   8G  0 disk
##/dev/sda1 202:1    0   8G  0 part /
/dev/sdb    202:80   0  10G  0 disk
##/dev/sdb1 202:81   0  10G  0 part
```

    Dans cet exemple, la partition à monter est `/dev/sdb1`, mais le nom de votre périphérique sera probablement différent. Si votre volume n'est pas partitionné, l'appareil à monter sera similaire à `/dev/sdb` (sans chiffre de fin de partition de périphérique).

11. Créez un point de montage pour le nouveau volume Amazon EBS et montez le volume.

```
[ec2-user bundle]$ sudo mkdir /mnt/ebs
[ec2-user bundle]$ sudo mount /dev/sdb1 /mnt/ebs
```

12. Ouvrez le fichier `/etc/fstab` sur le volume EBS avec votre éditeur de texte préféré (comme vim ou nano) et supprimez toutes les entrées pour les volumes (éphémères) de stockage d'instance. Étant donné que le volume Amazon EBS est monté sur `/mnt/ebs`, le fichier `fstab` se situe à l'emplacement `/mnt/ebs/etc/fstab`.

```
[ec2-user bundle]$ sudo nano /mnt/ebs/etc/fstab
#
LABEL=/     /           ext4    defaults,noatime  1   1
tmpfs       /dev/shm    tmpfs   defaults        0   0
devpts      /dev/pts    devpts  gid=5,mode=620  0   0
sysfs       /sys        sysfs   defaults        0   0
proc        /proc       proc    defaults        0   0
/dev/sdb        /media/ephemeral0       auto    defaults,comment=cloudconfig    0
 2
```

Dans cet exemple, la dernière ligne devrait être supprimée.

13. Démontez le volume et détachez-le de l'instance.

```
[ec2-user bundle]$ sudo umount /mnt/ebs
[ec2-user bundle]$ aws ec2 detach-volume --volume-id volume_id --region us-west-2
```

14. Créez une AMI à partir du nouveau volume Amazon EBS de la façon suivante.

    a. Créez un instantané du nouveau volume Amazon EBS.

    ```
    [ec2-user bundle]$ aws ec2 create-snapshot --region us-west-2 --description
     "your_snapshot_description" --volume-id volume_id
    ```

    b. Vérifiez si votre instantané est terminé.

    ```
    [ec2-user bundle]$ aws ec2 describe-snapshots --region us-west-2 --snapshot-
    id snapshot_id
    ```

    c. Identifiez l'architecture de processeur, le type de virtualisation et l'image noyau (`aki`) utilisée sur l'AMI originale avec la commande describe-images. Pour cette étape, vous avez besoin de l'ID d'AMI de l'AMI d'origine basée sur un stockage d'instance.

    ```
    [ec2-user bundle]$ aws ec2 describe-images --region us-west-2 --image-id ami-id --
    output text
    IMAGES x86_64 amazon/amzn-ami-pv-2013.09.2.x86_64-s3 ami-8ef297be amazon available
     public machine aki-fc8f11cc instance-store paravirtual xen
    ```

    Dans cet exemple, l'architecture est `x86_64` et l'ID de l'image noyau est `aki-fc8f11cc`. Utilisez ces valeurs dans l'étape suivante. Si le résultat de la commande ci-dessus liste aussi un ID `ari`, prenez également note de cela.

    d. Enregistrez votre nouvelle AMI avec l'ID d'instantané de votre nouveau volume Amazon EBS et les valeurs de l'étape précédente. Si la sortie de la commande précédente a répertorié un ID `ari`, incluez-le dans la commande suivante avec `--ramdisk-id ari_id`.

    ```
    [ec2-user bundle]$ aws ec2 register-image --region us-west-2 --
    name your_new_ami_name --block-device-mappings DeviceName=device-
    name,Ebs={SnapshotId=snapshot_id} --virtualization-type paravirtual --architecture
     x86_64 --kernel-id aki-fc8f11cc --root-device-name device-name
    ```

15. (Facultatif) Après avoir vérifié que vous pouvez lancer une instance à partir de votre nouvelle AMI, vous pouvez supprimer le volume Amazon EBS que vous avez créé pour cette procédure.

```
aws ec2 delete-volume --volume-id volume_id
```

Référence des outils AMI

Vous pouvez utiliser les commandes des outils AMI pour créer et gérer des AMI Linux basées sur le stockage d'instance. Pour installer les outils, consultez Installation des outils AMI (p. 138).

Pour plus d'informations sur vos clés d'accès, consultez les bonnes pratiques en matière de gestion des clés d'accès AWS.

Commandes

- ec2-ami-tools-version (p. 152)

- ec2-bundle-image (p. 152)
- ec2-bundle-vol (p. 154)
- ec2-delete-bundle (p. 158)
- ec2-download-bundle (p. 160)
- ec2-migrate-manifest (p. 162)
- ec2-unbundle (p. 164)
- ec2-upload-bundle (p. 165)
- Options courantes pour les outils AMI (p. 168)

ec2-ami-tools-version

Description

Décrit la version des outils AMI.

Syntaxe

ec2-ami-tools-version

File d'attente

Informations de version.

Exemple

Cet exemple de commande affiche les informations de version des outils AMI que vous utilisez.

```
[ec2-user ~]$ ec2-ami-tools-version
1.5.2 20071010
```

ec2-bundle-image

Description

Crée une AMI Linux basée sur le stockage d'instance à partir d'une image du système d'exploitation créée dans un fichier de boucle.

Syntaxe

ec2-bundle-image -c *path* -k *path* -u *account* -i *path* [-d *path*] [--ec2cert *path*] [-r *architecture*] [--productcodes *code1*,*code2*,...] [-B *mapping*] [-p *prefix*]

Options

`-c, --cert` chemin

Fichier de certificat de clé publique RSA code PEM de l'utilisateur.

Obligatoire : oui

`-k, --privatekey` chemin

Chemin d'accès à un fichier de clé RSA codée PEM. Vous devrez également spécifier cette clé pour dissocier ce groupe, conservez-la dans un endroit sûr. Notez que la clé n'a pas besoin d'être inscrite dans votre compte AWS.

Obligatoire : oui

`-u, --user` compte

ID de compte AWS de l'utilisateur sans tirets.

Obligatoire : oui

`-i, --image` chemin

Chemin d'accès à l'image à grouper.

Obligatoire : oui

`-d, --destination` chemin

Répertoire dans lequel vous créez le groupe.

Par défaut: `/tmp`

Requis : non

`--ec2cert` chemin

Chemin d'accès au certificat de clé publique Amazon EC2 X.509 utilisé pour chiffrer le manifeste de l'image.

Les régions `us-gov-west-1`et `cn-north-1` utilisent un certificat de clé publique par défaut. Le chemin d'accès à ce certificat doit être spécifié avec cette option. Le chemin d'accès au certificat varie selon la méthode d'installation des outils AMI. Pour Amazon Linux, les certificats se trouvent à l'adresse `/opt/aws/amitools/ec2/etc/ec2/amitools/`. Si vous avez installé les outils AMI à partir du fichier RPM ou ZIP dans Installation des outils AMI (p. 138), les certificats se trouvent à l'adresse `$EC2_AMITOOL_HOME/etc/ec2/amitools/`.

Obligatoire : uniquement pour les régions `us-gov-west-1` et `cn-north-1`.

`-r, --arch` application

Architecture d'image. Si vous ne fournissez pas l'architecture dans la ligne de commande, vous serez invité à la saisir au début de la création du bundle.

Valeurs valides : `i386` | `x86_64`

Requis : non

`--productcodes` code1,code2,...

Codes de produit à attacher à l'image au moment de l'inscription, séparé par des virgules.

Requis : non

`-B, --block-device-mapping` mappage

Définit la façon dont les périphériques de stockage en mode bloc sont exposés à une instance de AMI si son type d'instance prend en charge le périphérique spécifié.

Spécifiez une liste séparée par des virgules de paires clé-valeur, où chaque clé est un nom virtuel et chaque valeur le nom de périphérique correspondant. Les noms virtuels incluent les éléments suivants :

- `ami`— Périphérique du système de fichiers racine, tel qu'il est vu par l'instance
- `root`— Périphérique du système de fichiers racine, tel qu'il est vu par le noyau
- `swap`— Périphérique d'échange, tel qu'il est vu par l'instance
- `ephemeralN`—Volume de stockage de la nième instance

Requis : non

`-p, --prefix` prefix

Préfixe du nom des fichiers AMI groupés.

Par défaut : nom du fichier image. Par exemple, si le chemin d'accès de l'image est `/var/spool/my-image/version-2/debian.img`, le préfixe par défaut est `debian.img`.

Requis : non

`--kernel` kernel_id

Obsolète. Utilisez register-image pour définir le noyau.

Requis : non

`--ramdisk` ramdisk_id

Obsolète. Utilisez register-image pour définir le disque RAM le cas échéant.

Requis : non

File d'attente

Messages d'état décrivant les étapes et le statut du processus de groupement.

Exemple

Cet exemple crée une AMI groupée à partir d'une image du système d'exploitation qui a été créée dans un fichier de boucle.

```
[ec2-user ~]$ ec2-bundle-image -k pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -c cert-
HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -u 111122223333 -i image.img -d bundled/ -r x86_64
Please specify a value for arch [i386]:
Bundling image file...
Splitting bundled/image.gz.crypt...
Created image.part.00
Created image.part.01
Created image.part.02
Created image.part.03
Created image.part.04
Created image.part.05
Created image.part.06
Created image.part.07
Created image.part.08
Created image.part.09
Created image.part.10
Created image.part.11
Created image.part.12
Created image.part.13
Created image.part.14
Generating digests for each part...
Digests generated.
Creating bundle manifest...
ec2-bundle-image complete.
```

ec2-bundle-vol

Description

Crée une AMI Linux basée sur le stockage d'instance par compression, chiffrement et signature d'une copie du volume du périphérique racine de l'instance.

Amazon EC2 tente d'hériter les codes de produit, les paramètres du noyau, les paramètres du disque RAM et les mappages du périphérique de stockage en mode bloc de l'instance.

Par défaut, la création d'un bundle exclut les fichiers qui peuvent contenir des informations sensibles. Ces fichiers incluent `*.sw`, `*.swo`, `*.swp`, `*.pem`, `*.priv`, `*id_rsa*`, `*id_dsa*` `*.gpg`, `*.jks`, `*/.ssh/`

`authorized_keys` et `*/.bash_history`. Pour inclure tous ces fichiers, utilisez l'option `--no-filter`. Pour inclure certains de ces fichiers, utilisez l'option `--include`.

Pour plus d'informations, consultez Création d'une AMI Linux basée sur le stockage d'instance (p. 136).

Syntaxe

```
ec2-bundle-vol -c path -k path -u account [-d path] [--ec2cert path] [-
r architecture] [--productcodes code1,code2,...] [-B mapping] [--all] [-e
directory1,directory2,...] [-i file1,file2,...] [--no-filter] [-p prefix] [-
s size] [--[no-]inherit] [-v volume] [-P type] [-S script] [--fstab path] [--
generate-fstab] [--grub-config path]
```

Options

`-c, --cert` chemin

Fichier de certificat de clé publique RSA code PEM de l'utilisateur.

Obligatoire : oui

`-k, --privatekey` chemin

Chemin d'accès au fichier de clé RSA codé PEM de l'utilisateur.

Obligatoire : oui

`-u, --user` compte

ID de compte AWS de l'utilisateur sans tirets.

Obligatoire : oui

`-d, --destination` destination

Répertoire dans lequel vous créez le groupe.

Par défaut: `/tmp`

Requis : non

`--ec2cert` chemin

Chemin d'accès au certificat de clé publique Amazon EC2 X.509 utilisé pour chiffrer le manifeste de l'image.

Les régions `us-gov-west-1`et `cn-north-1` utilisent un certificat de clé publique par défaut. Le chemin d'accès à ce certificat doit être spécifié avec cette option. Le chemin d'accès au certificat varie selon la méthode d'installation des outils AMI. Pour Amazon Linux, les certificats se trouvent à l'adresse `/opt/aws/amitools/ec2/etc/ec2/amitools/`. Si vous avez installé les outils AMI à partir du fichier RPM ou ZIP dans Installation des outils AMI (p. 138), les certificats se trouvent à l'adresse `$EC2_AMITOOL_HOME/etc/ec2/amitools/`.

Obligatoire : uniquement pour les régions `us-gov-west-1` et `cn-north-1`.

`-r, --arch` application

Architecture de l'image. Si vous ne fournissez pas cette ligne de commande, vous serez invité à la saisir au début de la création du bundle.

Valeurs valides : `i386` | `x86_64`

Requis : non

`--productcodes` code1,code2,...

Codes de produit à attacher à l'image au moment de l'inscription, séparé par des virgules.

Requis : non

`-B, --block-device-mapping` mappage

Définit la façon dont les périphériques de stockage en mode bloc sont exposés à une instance de AMI si son type d'instance prend en charge le périphérique spécifié.

Spécifiez une liste séparée par des virgules de paires clé-valeur, où chaque clé est un nom virtuel et chaque valeur le nom de périphérique correspondant. Les noms virtuels incluent les éléments suivants :

- `ami`— Périphérique du système de fichiers racine, tel qu'il est vu par l'instance
- `root`— Périphérique du système de fichiers racine, tel qu'il est vu par le noyau
- `swap`— Périphérique d'échange, tel qu'il est vu par l'instance
- `ephemeralN`—Volume de stockage de la nième instance

Requis : non

`-a, --all`

Groupez tous les répertoires, y compris ceux contenus dans les systèmes de fichiers montés à distance.

Requis : non

`-e, --exclude` directory1,directory2,...

Liste des chemins absolus de répertoires et fichiers à exclure de l'opération de groupement. Ce paramètre remplace l'option `--all`. Lorsque la commande exclude est spécifié, les répertoires et sous-répertoires répertoriés avec le paramètre ne sont pas groupés avec le volume.

Requis : non

`-i, --include` file1,file2,...

Liste des fichiers à inclure dans l'opération de groupement. Les fichiers spécifiés seraient autrement exclus de l' AMI car ils peuvent contenir des informations sensibles.

Requis : non

`--no-filter`

Si ce paramètre est spécifié, nous n'excluons pas les fichiers de l'AMI, car ils peuvent contenir des informations sensibles.

Requis : non

`-p, --prefix` prefix

Préfixe du nom des fichiers AMI groupés.

Par défaut: `image`

Requis : non

`-s, --size` taille

Taille, en Mo (1024 * 1024 octets), du fichier image à créer. La taille maximale est 10 240 Mo.

Par défaut: 10240

Requis : non

`--[no-]inherit`

Indique si l'image doit hériter des métadonnées de l'instance (la valeur par défaut consiste à hériter). Le groupement échoue si vous activez `--inherit`, mais les métadonnées d'instance ne sont pas accessibles.

Requis : non

`-v, --volume` volume

Chemin d'accès absolu au volume monté à partir duquel créer le groupe.

Par défaut : le répertoire racine (/)

Requis : non

`-P, --partition` type

Indique si l'image de disque doit utiliser une table de partition. Si vous ne spécifiez pas de type de table de partition, la valeur par défaut est le type utilisé sur le périphérique de stockage en mode bloc parent du volume, le cas échéant. Dans le cas contraire, la valeur par défaut est `gpt`.

Valeurs valides : `mbr` | `gpt` | `none`

Requis : non

`-S, --script` script

Script de personnalisation à exécuter juste avant de procéder à la création du bundle. Le script doit attendre un seul argument, le point de montage du volume.

Requis : non

`--fstab` chemin

Chemin d'accès au fichier fstab à grouper dans l'image. S'il n'est pas spécifié, Amazon EC2 groupe/etc/fstab.

Requis : non

`--generate-fstab`

Groupe le volume grâce au fichier fastab fourni par Amazon EC2.

Requis : non

`--grub-config`

Chemin d'accès à un autre fichier de configuration grub à grouper dans l'image. Par défaut, `ec2-bundle-vol` attend `/boot/grub/menu.lst` ou `/boot/grub/grub.conf` pour exister sur l'image clonée. Cette option vous permet de spécifier un chemin d'accès à un autre fichier de configuration grub, qui sera ensuite copié par-dessus les valeurs par défaut (le cas échéant).

Requis : non

`--kernel` kernel_id

Obsolète. Utilisez register-image pour définir le noyau.

Requis : non

`--ramdisk`ramdisk_id

Obsolète. Utilisez register-image pour définir le disque RAM le cas échéant.

Requis : non

File d'attente

Messages d'état décrivant les étapes et le statut de la création du bundle.

Exemple

Cet exemple crée un groupe AMI par compression, chiffrement et signature d'un instantané du système de fichiers racine de l'ordinateur local.

```
[ec2-user ~]$ ec2-bundle-vol -d /mnt -k pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -c cert-
HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -u 111122223333 -r x86_64
  Copying / into the image file /mnt/image...
  Excluding:
       sys
       dev/shm
       proc
       dev/pts
       proc/sys/fs/binfmt_misc
       dev
       media
       mnt
       proc
       sys
       tmp/image
       mnt/img-mnt
  1+0 records in
  1+0 records out
  mke2fs 1.38 (30-Jun-2005)
  warning: 256 blocks unused.

  Splitting /mnt/image.gz.crypt...
  Created image.part.00
  Created image.part.01
  Created image.part.02
  Created image.part.03
  ...
  Created image.part.22
  Created image.part.23
  Generating digests for each part...
  Digests generated.
  Creating bundle manifest...
  Bundle Volume complete.
```

ec2-delete-bundle

Description

Supprime le groupe spécifié du stockage Amazon S3. Une fois que vous supprimez un groupe, vous ne pouvez pas lancer d'instances à partir de l'AMI correspondante.

Syntaxe

ec2-delete-bundle -b *bucket* -a *access_key_id* -s *secret_access_key* [-t *token*] [--url *url*] [--region *region*] [--sigv *version*] [-m *path*] [-p *prefix*] [--clear] [--retry] [-y]

Options

`-b, --bucket` bucket

Le nom du compartiment Amazon S3 contenant l'AMI groupée, suivi d'un préfixe de chemin facultatif séparé par des « / »

Obligatoire : oui

`-a, --access-key` access_key_id

ID de clé d'accès rapide AWS.

Obligatoire : oui

`-s, --secret-key` secret_access_key

Clé d'accès secrète AWS.

Obligatoire : oui

`-t, --delegation-token` jeton

Jeton de délégation à transmettre à la demande AWS. Pour plus d'informations, consultez le Using Temporary Security Credentials.

Requis : uniquement lorsque vous utilisez des informations d'identification de sécurité temporaires.

Par défaut : valeur de la variable d'environnement `AWS_DELEGATION_TOKEN` (si elle est définie).

`--region`region

Région à utiliser dans la signature de la demande.

Par défaut: `us-east-1`

Requis : requis si vous utilisez Signature Version 4

`--sigv`version

Version de signature à utiliser lors de la signature de la demande.

Valeurs valides : `2` | `4`

Par défaut: `4`

Requis : non

`-m, --manifest`chemin

Chemin d'accès au fichier manifeste.

Requis : vous devez spécifier `--prefix` ou `--manifest`.

`-p, --prefix` prefix

Préfixe du nom de fichier AMI groupé. Fournissez le préfixe entier. Par exemple, si le préfixe est image.img, utilisez `-p image.img`, non `-p image`.

Requis : vous devez spécifier `--prefix` ou `--manifest`.

`--clear`

Supprime le compartiment Amazon S3 s'il est vide après la suppression du groupe spécifié.

Requis : non

`--retry`

Refait automatiquement des tentatives sur toutes les erreurs Amazon S3, jusqu'à cinq fois par opération.

Requis : non

`-y, --yes`

Suppose automatiquement que la réponse à toutes les invites est oui.

Requis : non

File d'attente

Amazon EC2 affiche les messages d'état indiquant les étapes et le statut du processus de suppression.

Exemple

Cet exemple supprime un groupe de Amazon S3.

```
[ec2-user ~]$ ec2-delete-bundle -b DOC-EXAMPLE-BUCKET1 -a your_access_key_id -
s your_secret_access_key
Deleting files:
DOC-EXAMPLE-BUCKET1/
image.manifest.xml
DOC-EXAMPLE-BUCKET1/
image.part.00
DOC-EXAMPLE-BUCKET1/
image.part.01
DOC-EXAMPLE-BUCKET1/
image.part.02
DOC-EXAMPLE-BUCKET1/
image.part.03
DOC-EXAMPLE-BUCKET1/
image.part.04
DOC-EXAMPLE-BUCKET1/
image.part.05
DOC-EXAMPLE-BUCKET1/image.part.06
Continue? [y/n]
y
Deleted DOC-EXAMPLE-BUCKET1/image.manifest.xml
Deleted DOC-EXAMPLE-BUCKET1/image.part.00
Deleted DOC-EXAMPLE-BUCKET1/image.part.01
Deleted DOC-EXAMPLE-BUCKET1/image.part.02
Deleted DOC-EXAMPLE-BUCKET1/image.part.03
Deleted DOC-EXAMPLE-BUCKET1/image.part.04
Deleted DOC-EXAMPLE-BUCKET1/image.part.05
Deleted DOC-EXAMPLE-BUCKET1/image.part.06
ec2-delete-bundle complete.
```

ec2-download-bundle

Description

Télécharge les AMIs Linux basées sur le stockage d'instance depuis le stockage Amazon S3.

Syntaxe

ec2-download-bundle -b *bucket* -a *access_key_id* -s *secret_access_key* -k *path* [--url *url*] [--region *region*] [--sigv *version*] [-m *file*] [-p *prefix*] [-d *directory*] [--retry]

Options

`-b, --bucket` bucket

Nom du compartiment Amazon S3 où se trouve le groupe, suivi d'un préfixe de chemin séparé par des « / »-facultatif.

Obligatoire : oui

`-a, --access-key` access_key_id

ID de clé d'accès rapide AWS.

Obligatoire : oui

`-s, --secret-key` secret_access_key

Clé d'accès secrète AWS.

Obligatoire : oui

`-k, --privatekey` chemin

Clé privée utilisée pour déchiffrer le manifeste.

Obligatoire : oui

`--url` url

URL du service Amazon S3.

Par défaut: `https://s3.amazonaws.com/`

Requis : non

`--region` région

Région à utiliser dans la signature de la demande.

Par défaut: `us-east-1`

Requis : requis si vous utilisez Signature Version 4

`--sigv` version

Version de signature à utiliser lors de la signature de la demande.

Valeurs valides : `2` | `4`

Par défaut: `4`

Requis : non

`-m, --manifest` file

Nom du fichier manifeste (sans le chemin d'accès). Nous vous recommandons de spécifier soit le manifeste (`-m`) soit un préfixe (`-p`).

Requis : non

`-p, --prefix` prefix

Préfixe du nom des fichiers AMI groupés.

Par défaut: `image`

Requis : non

`-d, --directory` directory

Répertoire dans lequel le groupe téléchargé est enregistré. Le répertoire doit exister.

Par défaut : le répertoire de travail actuel.

Requis : non

`--retry`

Refait automatiquement des tentatives sur toutes les erreurs Amazon S3, jusqu'à cinq fois par opération.

Requis : non

File d'attente

Les messages d'état indiquant les différentes étapes du processus de téléchargement s'affichent.

Exemple

Cet exemple crée le répertoire `bundled` (à l'aide de la commande Linux mkdir) et télécharge le groupe depuis le compartiment Amazon S3 *DOC-EXAMPLE-BUCKET1*.

```
[ec2-user ~]$ mkdir bundled
[ec2-user ~]$ ec2-download-bundle -b DOC-EXAMPLE-BUCKET1/bundles/bundle_name
 -m image.manifest.xml -a your_access_key_id -s your_secret_access_key -k pk-
HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -d mybundle
Downloading manifest image.manifest.xml from DOC-EXAMPLE-BUCKET1 to mybundle/
image.manifest.xml ...
Downloading part image.part.00 from DOC-EXAMPLE-BUCKET1/bundles/bundle_name to mybundle/
image.part.00 ...
Downloaded image.part.00 from DOC-EXAMPLE-BUCKET1
Downloading part image.part.01 from DOC-EXAMPLE-BUCKET1/bundles/bundle_name to mybundle/
image.part.01 ...
Downloaded image.part.01 from DOC-EXAMPLE-BUCKET1
Downloading part image.part.02 from DOC-EXAMPLE-BUCKET1/bundles/bundle_name to mybundle/
image.part.02 ...
Downloaded image.part.02 from DOC-EXAMPLE-BUCKET1
Downloading part image.part.03 from DOC-EXAMPLE-BUCKET1/bundles/bundle_name to mybundle/
image.part.03 ...
Downloaded image.part.03 from DOC-EXAMPLE-BUCKET1
Downloading part image.part.04 from DOC-EXAMPLE-BUCKET1/bundles/bundle_name to mybundle/
image.part.04 ...
Downloaded image.part.04 from DOC-EXAMPLE-BUCKET1
Downloading part image.part.05 from DOC-EXAMPLE-BUCKET1/bundles/bundle_name to mybundle/
image.part.05 ...
Downloaded image.part.05 from DOC-EXAMPLE-BUCKET1
Downloading part image.part.06 from DOC-EXAMPLE-BUCKET1/bundles/bundle_name to mybundle/
image.part.06 ...
Downloaded image.part.06 from DOC-EXAMPLE-BUCKET1
```

ec2-migrate-manifest

Description

Modifie une AMI Linux basée sur le stockage d'instance (par exemple, son certificat, son noyau et son disque RAM) de sorte qu'elle prenne en charge une autre région.

Syntaxe

ec2-migrate-manifest -c *path* -k *path* -m *path* {(-a *access_key_id* -s *secret_access_key* --region *region*) | (--no-mapping)} [--ec2cert *ec2_cert_path*] [--kernel *kernel-id*] [--ramdisk *ramdisk_id*]

Options

`-c, --cert` chemin

Fichier de certificat de clé publique RSA code PEM de l'utilisateur.

Obligatoire : oui

`-k, --privatekey` chemin

Chemin d'accès au fichier de clé RSA codé PEM de l'utilisateur.

Obligatoire : oui

`--manifest` chemin

Chemin d'accès au fichier manifeste.

Obligatoire : oui

`-a, --access-key` access_key_id

ID de clé d'accès rapide AWS.

Requis : requis si vous utilisez le mappage automatique.

`-s, --secret-key` secret_access_key

Clé d'accès secrète AWS.

Requis : requis si vous utilisez le mappage automatique.

`--region` région

Région à rechercher dans le fichier de mappage.

Requis : requis si vous utilisez le mappage automatique.

`--no-mapping`

Désactive le mappage automatique des noyaux et disques RAM.

Lors de la migration, Amazon EC2 remplace le noyau et le disque RAM dans le fichier manifeste par un noyau et un disque RAM conçus pour la région de destination. Si le paramètre `--no-mapping` n'est pas fourni, `ec2-migrate-bundle` peut utiliser les opérations `DescribeRegions` et `DescribeImages` pour effectuer les mappages automatiques.

Requis : requis si vous ne fournissez pas les options `-a`, `-s` et `--region` utilisées pour le mappage automatique.

`--ec2cert` chemin

Chemin d'accès au certificat de clé publique Amazon EC2 X.509 utilisé pour chiffrer le manifeste de l'image.

Les régions `us-gov-west-1`et `cn-north-1` utilisent un certificat de clé publique par défaut. Le chemin d'accès à ce certificat doit être spécifié avec cette option. Le chemin d'accès au certificat varie selon la méthode d'installation des outils AMI. Pour Amazon Linux, les certificats se trouvent à l'adresse `/opt/aws/amitools/ec2/etc/ec2/amitools/`. Si vous avez installé les outils AMI à partir du fichier ZIP dans Installation des outils AMI (p. 138), les certificats se trouvent à l'adresse `$EC2_AMITOOL_HOME/etc/ec2/amitools/`.

Obligatoire : uniquement pour les régions `us-gov-west-1` et `cn-north-1`.

`--kernel` kernel_id

ID du noyau à sélectionner.

> **Important**
>
> Nous vous recommandons d'utiliser PV-GRUB au lieu des noyaux et des disques RAM. Pour plus d'informations, consultez Activation de vos propres noyaux Linux (p. 205).

Requis : non

`--ramdisk` ramdisk_id

ID du disque RAM à sélectionner.

> **Important**
>
> Nous vous recommandons d'utiliser PV-GRUB au lieu des noyaux et des disques RAM. Pour plus d'informations, consultez Activation de vos propres noyaux Linux (p. 205).

Requis : non

File d'attente

Messages d'état décrivant les étapes et le statut du processus de groupement.

Exemple

Cet exemple copie l'AMI spécifiée dans le fichier manifeste `my-ami.manifest.xml` depuis les États-Unis vers l'Union européenne.

```
[ec2-user ~]$ ec2-migrate-manifest --manifest my-ami.manifest.xml --cert cert-
HKZYKTAIG2ECMXYIBH3HXV4ZBZQ55CLO.pem --privatekey pk-HKZYKTAIG2ECMXYIBH3HXV4ZBZQ55CLO.pem
 --region eu-west-1

Backing up manifest...
Successfully migrated my-ami.manifest.xml It is now suitable for use in eu-west-1.
```

ec2-unbundle

Description

Recrée le groupe à partir d'une AMI Linux basée sur le stockage d'instance.

Syntaxe

ec2-unbundle -k *path* -m *path* [-s *source_directory*] [-d *destination_directory*]

Options

`-k, --privatekey` chemin

Chemin d'accès à votre fichier de clé RSA codée PEM.

Obligatoire : oui

`-m, --manifest` chemin

Chemin d'accès au fichier manifeste.

Obligatoire : oui

`-s, --source` source_directory

Répertoire contenant le groupe.

Par défaut : le répertoire actuel.

Requis : non

`-d, --destination` destination_directory

Répertoire dans lequel dégrouper l'AMI. Le répertoire de destination doit exister.

Par défaut : le répertoire actuel.

Requis : non

Exemple

Cet exemple Linux et UNIX dégroupe l'AMI spécifiée dans le fichier `image.manifest.xml`.

```
[ec2-user ~]$ mkdir unbundled
$ ec2-unbundle -m mybundle/image.manifest.xml -k pk-HKZYKTAIG2ECMXYIBH3HXV4ZBEXAMPLE.pem -s
 mybundle -d unbundled
```

```
$ ls -l unbundled
total 1025008
-rw-r--r-- 1 root root 1048578048 Aug 25 23:46 image.img
```

File d'attente

Les messages d'état indiquant les différentes étapes du processus de dégroupement s'affichent.

ec2-upload-bundle

Description

Charge le bundle d'une AMI Linux basée sur le stockage d'instance vers Amazon S3 et définit les ACL appropriées au niveau des objets chargés. Pour plus d'informations, consultez Création d'une AMI Linux basée sur le stockage d'instance (p. 136).

Syntaxe

ec2-upload-bundle -b *bucket* -a *access_key_id* -s *secret_access_key* [-t *token*] -m *path* [--url *url*] [--region *region*] [--sigv *version*] [--acl *acl*] [-d *directory*] [--part *part*] [--retry] [--skipmanifest]

Options

`-b, --bucket` bucket

Nom du compartiment Amazon S3 dans lequel stocker le groupe, suivi d'un préfixe de chemin séparé par des « / » facultatif. Si le compartiment n'existe pas, il est créé si le nom de compartiment est disponible.

Obligatoire : oui

`-a, --access-key` access_key_id

Votre ID de clé d'accès rapide AWS.

Obligatoire : oui

`-s, --secret-key` secret_access_key

Clé d'accès secrète de votre compte AWS.

Obligatoire : oui

`-t, --delegation-token` jeton

Jeton de délégation à transmettre à la demande AWS. Pour plus d'informations, consultez le Using Temporary Security Credentials.

Requis : uniquement lorsque vous utilisez des informations d'identification de sécurité temporaires.

Par défaut : valeur de la variable d'environnement `AWS_DELEGATION_TOKEN` (si elle est définie).

`-m, --manifest` chemin

Chemin d'accès au fichier manifeste. Le fichier manifeste est créé pendant la création d'un bundle ; il est disponible dans le répertoire contenant le groupe.

Obligatoire : oui

`--url` url

Obsolète. Utilisez plutôt l'option `--region`, sauf si votre compartiment est limité à l'emplacement `EU` (et pas `eu-west-1`). L'indicateur `--location` est le seul moyen de cibler cette restriction d'emplacement spécifique.

URL du service de point de terminaison Amazon S3.

Par défaut: `https://s3.amazonaws.com/`

Requis : non

`--region` région

Région à utiliser dans la signature de la demande pour le compartiment de destination S3.

- Si le compartiment n'existe pas et que vous ne spécifiez pas une région, l'outil crée le compartiment sans contrainte d'emplacement (dans `us-east-1`).
- Si le compartiment n'existe pas et que vous spécifiez une région, l'outil crée le compartiment dans la région spécifiée.
- Si le compartiment existe et que vous ne spécifiez pas une région, l'outil utilise emplacement du compartiment.
- Si le compartiment existe et que vous spécifiez `us-east-1` comme région, l'outil utilise l'emplacement du compartiment sans aucun message d'erreur, tous les fichiers correspondants existants sont écrasés.
- Si le compartiment existe et que vous spécifiez une région (autre que `us-east-1`) qui ne correspond pas à l'emplacement du compartiment, l'outil se termine avec une erreur.

Si votre compartiment est limité à l'emplacement `EU` (et pas `eu-west-1`), utilisez plutôt l'indicateur `--location`. L'indicateur `--location` est le seul moyen de cibler cette restriction d'emplacement spécifique.

Par défaut: `us-east-1`

Requis : requis si vous utilisez Signature Version 4

`--sigv` version

Version de signature à utiliser lors de la signature de la demande.

Valeurs valides : `2` | `4`

Par défaut: `4`

Requis : non

`--acl` acl

Stratégie de liste de contrôle des accès de l'image groupée.

Valeurs valides : `public-read` | `aws-exec-read`

Par défaut: `aws-exec-read`

Requis : non

`-d, --directory` directory

Répertoire contenant les parties de l'AMI groupée.

Par défaut : le répertoire contenant le fichier manifeste (cf. l'option `-m`).

Requis : non

`--part` part

Commence le chargement de la partie spécifiée et de toutes les parties suivantes. Par exemple, `--part 04`.

Requis : non

`--retry`

Refait automatiquement des tentatives sur toutes les erreurs Amazon S3, jusqu'à cinq fois par opération.

Requis : non

`--skipmanifest`

Ne charge pas le fichier manifeste.

Requis : non

`--location` location

Obsolète. Utilisez plutôt l'option `--region`, sauf si votre compartiment est limité à l'emplacement `EU` (et pas `eu-west-1`). L'indicateur `--location` est le seul moyen de cibler cette restriction d'emplacement spécifique.

Contrainte d'emplacement du compartiment Amazon S3 de destination. Si le compartiment existe et que vous spécifiez un emplacement qui ne correspond pas à l'emplacement du compartiment, l'outil se termine avec une erreur. Si le compartiment existe et que vous ne spécifiez pas d'emplacement, l'outil utilise emplacement du compartiment. Si le compartiment n'existe pas et que vous spécifiez un emplacement, l'outil crée le compartiment dans l'emplacement spécifié. Si le compartiment n'existe pas et que vous ne spécifiez pas d'emplacement, l'outil crée le compartiment sans contrainte d'emplacement (dans `us-east-1`).

Par défaut : si `--region` est spécifié, l'emplacement est défini sur cette région spécifiée. Si `--regionn`'est pas spécifié, l'emplacement par défaut est `us-east-1`.

Requis : non

File d'attente

Amazon EC2 affiche les messages d'état qui indiquent les étapes et l'état du processus de chargement.

Exemple

Cet exemple télécharge le groupe spécifié par le fichier manifeste `image.manifest.xml`.

```
[ec2-user ~]$ ec2-upload-bundle -b DOC-EXAMPLE-BUCKET1/bundles/bundle_name -m
 image.manifest.xml -a your_access_key_id -s your_secret_access_key
Creating bucket...
Uploading bundled image parts to the S3 bucket DOC-EXAMPLE-BUCKET1 ...
Uploaded image.part.00
Uploaded image.part.01
Uploaded image.part.02
Uploaded image.part.03
Uploaded image.part.04
Uploaded image.part.05
Uploaded image.part.06
Uploaded image.part.07
Uploaded image.part.08
Uploaded image.part.09
Uploaded image.part.10
Uploaded image.part.11
Uploaded image.part.12
Uploaded image.part.13
Uploaded image.part.14
Uploading manifest ...
Uploaded manifest.
Bundle upload completed.
```

Options courantes pour les outils AMI

La plupart des outils AMI acceptent les paramètres facultatifs suivants.

`--help, -h`

Affiche le message d'aide.

`--version`

Affiche la version et l'avis de droit d'auteur.

`--manual`

Affiche l'entrée manuelle.

`--batch`

S'exécute en mode de traitement par lots et supprime les invites interactives.

`--debug`

Affiche les informations qui peuvent être utiles pour la résolution de problèmes.

Utiliser le chiffrement avec des AMI basées sur EBS

Les AMI basées sur des instantanés Amazon EBS peuvent tirer parti du chiffrement Amazon EBS. Les instantanés de volumes de données et racine peuvent être chiffrés et attachés à une AMI. Vous pouvez lancer des instances et copier des images avec une prise en charge complète du chiffrement EBS. Les paramètres de chiffrement de ces opérations sont pris en charge dans toutes les régions où AWS KMS est disponible.

Les instances EC2 avec des volumes EBS chiffrés sont lancées à partir des AMIs de la même manière que les autres instances. De plus, lorsque vous lancez une instance à partir d'une AMI basée sur des instantanés EBS non chiffrés, vous pouvez chiffrer une partie ou l'ensemble des volumes pendant le lancement.

Comme pour les volumes EBS, les instantanés des AMI peuvent être chiffrés avec votre clé principale client (CMK) AWS Key Management Service par défaut ou sous une clé gérée par le client que vous spécifiez. Dans tous les cas, vous devez être autorisé à utiliser la clé sélectionnée.

Les AMI avec des instantanés chiffrés peuvent être partagés entre plusieurs comptes AWS. Pour plus d'informations, consultez Shared AMIs (AMI partagées).

Rubriques relatives au chiffrement avec des AMI basées sur EBS

- Scénarios de lancement d'instances (p. 168)
- Scénarios de copie d'images (p. 171)

Scénarios de lancement d'instances

Les instances Amazon EC2 sont lancées à partir des AMI au moyen de l'action `RunInstances` avec les paramètres fournis via la propriété block device mapping, soit avec la AWS Management Console soit directement avec l'API Amazon EC2 ou l'interface de ligne de commande. Pour de plus amples informations sur la propriété block device mapping, veuillez consulter Block Device Mapping. Pour consultez des exemples de contrôle de la propriété block device mapping depuis l'AWS CLI, consultez Launch, List, and Terminate EC2 Instances (Lancer, répertorier et résilier des instances EC2).

Par défaut, sans paramètres de chiffrement explicites, une action `RunInstances` conserve l'état de chiffrement existant des instantanés source d'une AMI lors de la restauration des volumes EBS à partir

de ceux-ci. Si le Chiffrement par défaut (p. 1196) est activé, tous les volumes créés à partir de l'AMI (que ce soit à partir d'instantanés chiffrés ou non chiffrés) seront chiffrés. Si le chiffrement par défaut n'est pas activé, l'instance conserve l'état de chiffrement de l'AMI.

Vous pouvez également lancer une instance et, simultanément, appliquer un nouvel état de chiffrement aux volumes créés en spécifiant les paramètres de chiffrement. Dans un tel cas, les comportements suivants sont observés :

Lancement sans paramètres de chiffrement

- Un instantané non chiffré est restauré dans un volume non chiffré, sauf si le chiffrement par défaut est activé, auquel cas tous les volumes nouvellement créés seront chiffrés.
- Un instantané non chiffré que vous possédez est restauré dans un volume qui est chiffré avec la même clé CMK.
- Un instantané chiffré que vous ne possédez pas (par exemple, l'AMI est partagée avec vous) est restauré dans un volume qui est chiffré avec la clé CMK par défaut de votre compte AWS.

Les comportements par défaut peuvent être ignorés en spécifiant les paramètres de chiffrement. Les paramètres disponibles sont `Encrypted` et `KmsKeyId`. La définition du seul paramètre `Encrypted` produit les effets suivants :

Comportements en cas de lancement d'instance avec le paramètre **`Encrypted`** défini, mais sans spécifier le paramètre **`KmsKeyId`**

- Un instantané non chiffré est restauré dans un volume EBS qui est chiffré avec la clé CMK par défaut de votre compte AWS.
- Un instantané chiffré que vous possédez est restauré dans un volume EBS qui est chiffré avec la même clé CMK. (En d'autres mots, le paramètre `Encrypted` est sans effet.)
- Un instantané chiffré que vous ne possédez pas (autrement dit, l'AMI est partagée avec vous) est restauré dans un volume qui est chiffré avec la clé CMK par défaut de votre compte AWS. (En d'autres mots, le paramètre `Encrypted` est sans effet.)

La définition des paramètres `Encrypted` et `KmsKeyId` vous permet de spécifier une clé CMK autre que la clé par défaut pour une opération de chiffrement. Les comportements suivants sont observés :

Instance avec définition des paramètres **`Encrypted`** et **`KmsKeyId`**

- Un instantané non chiffré est restauré dans un volume EBS qui est chiffré avec la clé CMK spécifiée.
- Un instantané chiffré est restauré dans un volume EBS qui est chiffré non pas avec la clé CMK d'origine mais avec la clé CMK spécifiée.

L'envoi de `KmsKeyId` sans définir également le paramètre `Encrypted` génère une erreur.

Les sections suivantes fournissent des exemples de lancement d'instances à partir d'AMI avec des paramètres de chiffrement autres que les paramètres par défaut. Dans chacun de ces scénarios, les paramètres fournis à l'action `RunInstances` entraînent un changement de l'état de chiffrement pendant la restauration d'un volume à partir d'un instantané.

> **Note**
>
> Pour connaître les procédures détaillées de la console pour le lancement d'une instance à partir d'une AMI, veuillez consulter Lancer votre instance.
> Pour obtenir la documentation relative à l'API `RunInstances`, consultez RunInstances.
> Pour obtenir la documentation relative à la commande `run-instances` dans l'AWS Command Line Interface, veuillez consulter run-instances.

Chiffrement d'un volume pendant le lancement

Dans cet exemple, une AMI basée sur un instantané non chiffré est utilisée pour lancer une instance EC2 avec un volume EBS chiffré.

Le paramètre `Encrypted` seul entraîne le chiffrement du volume pour cette instance. Le paramètre `KmsKeyId` est facultatif. Si aucun ID de clé n'est spécifié, la clé CMK par défaut du compte AWS est utilisée pour chiffrer le volume. Pour chiffrer le volume avec une autre clé CMK que vous possédez, fournissez le paramètre `KmsKeyId`.

Rechiffrement d'un volume pendant le lancement

Dans cet exemple, une AMI basée sur un instantané chiffré est utilisée pour lancer une instance EC2 avec un volume EBS chiffré avec une nouvelle clé CMK.

Si vous possédez l'AMI et ne spécifiez pas de paramètres de chiffrement, l'instance obtenue a un volume chiffré avec la même clé que l'instantané. Si l'AMI est partagée avec vous mais que vous n'en êtes pas propriétaire, et si vous ne spécifiez pas de paramètres de chiffrement, le volume est chiffré avec votre clé CMK par défaut. Avec les paramètres de chiffrement fournis, comme illustré, le volume est chiffré avec la clé CMK spécifiée.

Modification de l'état de chiffrement de plusieurs volumes pendant le lancement

Dans cet exemple plus complexe, une AMI basée sur plusieurs instantanés (chacun avec son propre état de chiffrement) est utilisée pour lancer une instance EC2 avec un volume nouvellement chiffré et un volume rechiffré.

Dans ce scénario, l'action `RunInstances` reçoit des paramètres de chiffrement pour chacun des instantanés source. Lorsque tous les paramètres de chiffrement sont spécifiés, l'instance créée est la même, que vous possédiez ou non l'AMI.

Scénarios de copie d'images

Les AMIAmazon EC2 sont copiées au moyen de l'action `CopyImage`, soit via la AWS Management Console soit directement avec l'API Amazon EC2 ou l'interface de ligne de commande.

Par défaut, sans paramètres de chiffrement explicites, une action `CopyImage` conserve l'état de chiffrement existant des instantanés source d'une AMI lors de la copie. Vous pouvez également copier

une AMI et, simultanément, appliquer un nouvel état de chiffrement à ses instantanés EBS associés en spécifiant les paramètres de chiffrement. Dans un tel cas, les comportements suivants sont observés :

Copie sans paramètres de chiffrement

- Un instantané non chiffré est copié dans un autre instantané non chiffré, sauf si le chiffrement par défaut est activé, auquel cas tous les instantanés nouvellement créés seront chiffrés.
- Un instantané chiffré que vous possédez est copié dans un instantané chiffré avec la même clé.
- Un instantané chiffré que vous ne possédez pas (autrement dit, l'AMI est partagée avec vous) est copié dans un instantané qui est chiffré avec la clé CMK par défaut de votre compte AWS.

Tous ces comportements par défaut peuvent être ignorés en spécifiant les paramètres de chiffrement. Les paramètres disponibles sont `Encrypted` et `KmsKeyId`. La définition du seul paramètre `Encrypted` produit les effets suivants :

Comportements en cas de copie-image avec le paramètre **`Encrypted`** défini, mais pas le paramètre **`KmsKeyId`**

- Un instantané non chiffré est copié dans un instantané chiffré avec la clé CMK par défaut du compte AWS.
- Un instantané chiffré est copié dans un instantané chiffré avec la même clé CMK. (En d'autres mots, le paramètre `Encrypted` est sans effet.)
- Un instantané chiffré que vous ne possédez pas (autrement dit, l'AMI est partagée avec vous) est copié dans un volume qui est chiffré avec la clé CMK par défaut de votre compte AWS. (En d'autres mots, le paramètre `Encrypted` est sans effet.)

La définition des paramètres `Encrypted` et `KmsKeyId` vous permet de spécifier une clé CMK gérée par le client pour une opération de chiffrement. Les comportements suivants sont observés :

Comportements en cas de copie-image avec les paramètres **`Encrypted`** et **`KmsKeyId`** définis

- Un instantané non chiffré est copié dans un instantané chiffré avec la clé CMK spécifiée.
- Un instantané chiffré est copié dans un instantané qui est chiffré non pas avec la clé CMK d'origine mais avec la clé CMK spécifiée.

L'envoi de `KmsKeyId` sans définir également le paramètre `Encrypted` génère une erreur.

La section suivante fournit un exemple de copie d'une AMI avec des paramètres de chiffrement personnalisés, ce qui entraîne un changement de l'était de chiffrement.

> **Note**
>
> Pour connaître les procédures détaillées de la console afin de copier une AMI, consultez Copier une AMI.
> Pour obtenir la documentation relative à l'API `CopyImage`, consultez CopyImage.
> Pour obtenir la documentation relative à la commande `copy-image` dans l'AWS Command Line Interface, consultez copy-image.

Chiffrement d'une image non chiffrée pendant la copie

Dans ce scénario, une AMI basée sur un instantané racine non chiffré est copiée sur une AMI comportant un instantané racine chiffré. L'action `CopyImage` est appelée avec deux paramètres de chiffrement, notamment une clé CMK. Par conséquent, le statut de chiffrement de l'instantané racine est modifié. L'AMI cible est donc basée sur un instantané racine contenant les mêmes données que l'instantané source, mais chiffré à l'aide de la clé spécifiée. Vous payez des coûts de stockage pour les instantanés dans les deux AMIs, ainsi que des frais pour les instances que vous lancez à partir de l'une ou l'autre de ces AMI.

Note

L'activation du chiffrement par défaut (p. 1196) a le même effet que la définition du paramètre `Encrypted` sur `true` pour tous les instantanés de l'AMI.

La définition du paramètre `Encrypted` chiffre l'unique instantané pour cette instance. Si vous ne spécifiez pas le paramètre `KmsKeyId`, la clé CMK par défaut est utilisée pour chiffrer la copié d'instantané.

Note

Vous pouvez également copier une image avec plusieurs instantanés et configurer l'état de chiffrement de chacun individuellement.

Copie une AMI

Vous pouvez copier une Amazon Machine Image (AMI) au sein d'une région AWS ou de plusieurs régions à l'aide d'AWS Management Console, de l'AWS Command Line Interface ou des kits SDK, ou encore de l'API Amazon EC2, qui prennent tous en charge l'action `CopyImage`. Vous pouvez copier à la fois les AMIs basées sur Amazon EBS et les AMIs basées sur le stockage d'instance. Vous pouvez copier les AMI avec des instantanés chiffrés et également modifier le statut de chiffrement pendant le processus de copie.

La copie d'une AMI source crée une AMI cible identique mais distincte, avec son propre identificateur unique. Dans le cas d'une AMI basée sur des volumes Amazon EBS, chacun de ses instantanés est, par défaut, copié dans un instantané cible identique mais distinct. (Les seules exceptions correspondent aux cas où vous choisissez de chiffrer ou de rechiffrer l'instantané.) Vous pouvez modifier ou annuler l'enregistrement d'une AMI source sans que cela ait un impact sur l'AMI cible. L'inverse est également vrai.

Vous n'êtes pas facturé pour la copie d'une AMI. Toutefois, les taux standard de stockage et de transfert de données s'appliquent. Si vous copiez une AMI basée sur EBS, des frais seront facturés pour le stockage de tout instantané EBS supplémentaire.

AWS ne copie pas les autorisations de lancement, les balises définies par l'utilisateur ou les permissions de compartiment Amazon S3 de l'AMI source vers la nouvelle AMI. Une fois la copie terminée, vous

pouvez appliquer les autorisations de lancement, les balises définies par l'utilisateur et les permissions de compartiment Amazon S3 à la nouvelle AMI.

Si vous utilisez une AMI AWS Marketplace ou une AMI dérivée directement ou indirectement d'une AMI AWS Marketplace, vous ne pouvez pas la copier d'un compte à un autre. À la place, lancez une instance EC2 en utilisant l'AMI d'AWS Marketplace, puis créez une AMI à partir de cette instance. Pour plus d'informations, consultez Création d'une AMI Linux basée sur Amazon EBS (p. 132).

Rubriques

- Autorisations nécessaires pour copier une AMI basée sur le stockage d'instance (p. 174)
- Copie entre régions (p. 175)
- Copie entre comptes (p. 176)
- Chiffrement et copie (p. 176)
- Copie d'une AMI (p. 177)
- Arrêt d'une opération de copie d'une AMI en attente (p. 179)

Autorisations nécessaires pour copier une AMI basée sur le stockage d'instance

Si vous utilisez un utilisateur IAM pour copier une AMI basée sur le stockage d'instance, l'utilisateur doit disposer des autorisations Amazon S3 suivantes : `s3:CreateBucket`, `s3:GetBucketAcl`, `s3:ListAllMyBuckets`, `s3:GetObject`, `s3:PutObject` et `s3:PutObjectAcl`.

L'exemple de stratégie suivant permet à l'utilisateur de copier l'AMI source dans le compartiment spécifié de la région spécifiée.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "s3:ListAllMyBuckets",
            "Resource": [
                "arn:aws:s3:::*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": "s3:GetObject",
            "Resource": [
                "arn:aws:s3:::ami-source-bucket/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
              "s3:CreateBucket",
              "s3:GetBucketAcl",
              "s3:PutObjectAcl",
              "s3:PutObject"
            ],
            "Resource": [
                "arn:aws:s3:::amis-for-123456789012-in-us-east-1*"
            ]
        }
    ]
}
```

Pour trouver le nom de ressource Amazon (ARN) du compartiment source AMI, ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/, dans le volet de navigation, choisissez AMI, et recherchez le nom du compartiment dans la colonne Source .

> **Note**
>
> L'autorisation `s3:CreateBucket` n'est requise que la première fois que l'utilisateur IAM copie un magasin d'instance sauvegardé AMI dans une région donnée. Après cela, le compartiment Amazon S3 qui est déjà créé dans la région est utilisé pour stocker tous les futurs AMIs que vous copiez dans cette région.

Copie entre régions

La copie d'une AMI entre différentes régions géographiques offre les avantages suivants :

- Déploiement international uniforme : la copie d'une AMI d'une région à l'autre vous permet de lancer des instances uniformes reposant sur la même AMI dans différentes régions.
- Évolutivité : vous pouvez plus facilement concevoir et créer des applications d'envergure internationale répondant aux besoins de vos utilisateurs, quel que soit l'emplacement.
- Performances : vous pouvez accroître les performances en distribuant votre application, ainsi qu'en recherchant les composants critiques de votre application plus près de vos utilisateurs. Vous pouvez également bénéficier de fonctions propres aux régions, par exemple les types d'instance ou d'autres services AWS.
- Disponibilité élevée : vous pouvez concevoir et déployer des applications dans différentes régions AWS afin d'accroître leur disponibilité.

Le diagramme suivant représente les relations entre une AMI source et deux AMIs copiées dans différentes régions, ainsi que les instances EC2 lancées à partir de chacune d'entre elles. Lorsque vous lancez une instance à partir d'une AMI, elle se trouve dans la même région que l'AMI. Si vous modifiez l'AMI source et que vous souhaitez faire apparaître ces modifications dans les AMIs des régions cibles, vous devez recopier l'AMI source dans les régions cibles.

La première fois que vous copiez une AMI basée sur le stockage d'instance dans une région, nous créons un compartiment Amazon S3 pour les AMIs copiées dans cette région. Toutes les AMIs basées sur le stockage d'instance que vous copiez dans cette région sont stockées dans ce compartiment. Les noms des compartiments ont le format suivant : amis-for-*account*-in-*region*-*hash*. Par exemple : `amis-for-123456789012-in-us-east-2-yhjmxvp6`.

Prérequis

Avant de copier une AMI, vous devez veiller à ce que le contenu de l'AMI source ait été mis à jour afin de pouvoir être exécuté dans une région différente. Par exemple, vous devez mettre à jour toutes les chaînes de connexion à la base de données ou des données de configuration d'application similaires de façon à ce qu'elles pointent vers les ressources appropriées. Sinon, les instances lancées depuis la nouvelle AMI dans la région de destination seraient susceptibles d'utiliser encore les ressources de la région source, ce qui aurait des répercussions sur les performances et le coût.

Limites

- Les régions de destination sont limitées à 50 copies d'une AMI à la fois.
- Vous ne pouvez pas copier une AMI paravirtuelle (PV) vers une région qui ne prend pas en charge les AMI PV. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108).

Copie entre comptes

Vous pouvez partager une AMI avec un autre compte AWS. Le partage d'une AMI n'affecte pas la propriété de celle-ci. Le compte propriétaire est facturé pour le stockage dans la région. Pour plus d'informations, consultez Partager une AMI avec des comptes AWS spécifiques (p. 121).

Si vous copiez une AMI qui a été partagée avec votre compte, vous êtes le propriétaire de l'AMI cible de votre compte. Le propriétaire de l'AMI source se voir facturer des frais standard de transfert Amazon EBS ou Amazon S3, et vous devez régler le stockage de l'AMI cible dans la région de destination.

Autorisations d'accès aux ressources

Pour copier une AMI qui a été partagée avec vous à partir d'un autre compte, le propriétaire de l'AMI source doit vous accorder des autorisations de lecture pour le stockage sur lequel est basée l'AMI, soit l'instantané EBS associé (pour une AMI basée sur Amazon EBS) soit un compartiment S3 associé (pour une AMI basée sur le stockage d'instance). Si l'AMI partagée comporte des instantanés chiffrés, le propriétaire doit également partager la ou les clé(s) avec vous.

Chiffrement et copie

Le tableau suivant représente la prise en charge du chiffrement dans divers cas de figure de copie d'AMI. Bien qu'il soit possible de copier un instantané non chiffré pour créer un instantané chiffré, vous ne pouvez pas copier un instantané chiffré et en créer un qui ne soit pas chiffré.

Scénario	Description	Pris en charge
1	Non chiffré vers non chiffré	Oui
2	Chiffré vers chiffré	Oui
3	Non chiffré vers chiffré	Oui
4	Chiffré vers non chiffré	Non

Note

Le chiffrement pendant l'action `CopyImage` s'applique uniquement aux AMIs basées sur Amazon EBS. Dans la mesure où une AMI basée sur le stockage d'instance ne s'appuie pas sur les instantanés, vous ne pouvez pas utiliser la copie pour modifier son statut de chiffrement.

Par défaut (à savoir, sans spécifier les paramètres de chiffrement), l'instantané sur lequel repose une AMI est copié avec son statut de chiffrement initial. La copie d'une AMI reposant sur un instantané non chiffré crée un instantané cible identique qui n'est pas chiffré non plus. Si l'AMI source repose sur un instantané chiffré, sa copie crée un instantané cible identique qui est chiffré avec la même clé principale client (CMK). La copie d'une AMI basée sur plusieurs instantanés conserve, par défaut, le statut de chiffrement source dans chaque instantané cible.

Si vous spécifiez les paramètres de chiffrement lors de la copie d'une AMI, vous pouvez chiffrer ou rechiffrer ses instantanés. L'exemple suivant montre un cas (non par défaut) qui fournit les paramètres de chiffrement à l'action `CopyImage` dans le but de modifier l'état de chiffrement de l'AMI cible.

Copier une AMI source non chiffrée vers une AMI cible chiffrée

Dans ce scénario, une AMI basée sur un instantané racine non chiffré est copiée sur une AMI comportant un instantané racine chiffré. L'action `CopyImage` est appelée avec deux paramètres de chiffrement, notamment une clé CMK. Par conséquent, le statut de chiffrement de l'instantané racine est modifié. L'AMI cible est donc basée sur un instantané racine contenant les mêmes données que l'instantané source, mais chiffré à l'aide de la clé spécifiée. Vous payez des coûts de stockage pour les instantanés dans les deux AMIs, ainsi que des frais pour les instances que vous lancez à partir de l'une ou l'autre de ces AMI.

> **Note**
>
> L'activation du chiffrement par défaut (p. 1196) a le même effet que la définition du paramètre `Encrypted` sur `true` pour tous les instantanés de l'AMI.

La définition du paramètre `Encrypted` chiffre l'unique instantané pour cette instance. Si vous ne spécifiez pas le paramètre `KmsKeyId`, la clé CMK par défaut est utilisée pour chiffrer la copié d'instantané.

Pour plus d'informations sur la copie d'AMIs avec des instantanés chiffrés, consultez Utiliser le chiffrement avec des AMI basées sur EBS (p. 168).

Copie d'une AMI

Vous pouvez copier une AMI comme suit.

Prérequis

Créez ou obtenez une AMI basée sur un instantané Amazon EBS. Notez que vous pouvez utiliser la console Amazon EC2 pour effectuer une recherche parmi une grande variété d'AMI fournies par AWS. Pour de plus amples informations, veuillez consulter Création d'une AMI Linux basée sur Amazon EBS (p. 132) et Recherche d'une AMI.

Pour copier une AMI à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation de la console, sélectionnez la région comportant l'AMI. Dans le panneau de navigation, choisissez Images, puis AMI afin d'afficher la liste des AMIs disponibles pour vous dans la région.
3. Sélectionnez l'AMI à copier, puis choisissez Actions et Copier l'AMI.
4. Dans la boîte de dialogue Copier l'AMI, spécifiez les informations suivantes, puis choisissez Copier l'AMI :
   - Région de destination) : région dans laquelle vous souhaitez copier l'AMI.
   - Nom : nom de la nouvelle AMI. Vous pouvez inclure des informations sur le système d'exploitation dans le nom, car nous ne fournissons pas ces informations lors de l'affichage des détails relatifs à l'AMI.
   - Description : par défaut, la description inclut des informations sur l'AMI source afin que vous puissiez identifier une copie à partir de l'original. Vous pouvez modifier cette description si nécessaire.
   - Chiffrement : sélectionnez ce champ pour chiffrer les instantanés ou pour les rechiffrer à l'aide d'une clé différente. Si vous avez activé le chiffrement par défaut, l'option Chiffrement est activée et ne peut pas être désactivée depuis la console AMI.
   - Clé principale : clé KMS utilisée pour chiffrer les instantanés cibles.
5. Nous affichons une page de confirmation pour vous avertir que l'opération de copie a été lancée et pour vous communiquer l'ID de la nouvelle AMI.

   Pour vérifier la progression de l'opération de copie immédiatement, suivez le lien fourni. Pour vérifier la progression ultérieurement, choisissez Effectué puis, une fois que vous êtes prêt, utilisez la barre de navigation afin de passer à la région cible (le cas échéant) et de rechercher l'AMI dans la liste des AMI.

   Le statut initial de l'AMI cible est `pending` et l'opération est terminée lorsque le statut est `available`.

Pour copier une AMI à l'aide de l'AWS CLI

Vous pouvez copier une AMI à l'aide de la commande copy-image. Vous devez indiquer les régions source et de destination. Vous spécifiez la région source à l'aide du paramètre `--source-region`. Vous pouvez spécifier la région de destination à l'aide du paramètre `--region` ou d'une variable d'environnement. Pour plus d'informations, consultez Configuration de l'interface de ligne de commande AWS.

Lorsque vous chiffrez un instantané cible pendant la copie, vous devez spécifier ces paramètres supplémentaires : `--encrypted` et `--kms-key-id`.

Pour copier une AMI à l'aide des Outils pour Windows PowerShell

Vous pouvez copier une AMI à l'aide de la commande Copy-EC2Image. Vous devez indiquer les régions source et de destination. Vous spécifiez la région source à l'aide du paramètre `-SourceRegion`. Vous pouvez spécifier la région de destination à l'aide du paramètre `-Region` ou de la commande `Set-AWSDefaultRegion`. Pour plus d'informations, consultez Spécification des régions AWS.

Lorsque vous chiffrez un instantané cible pendant la copie, vous devez spécifier ces paramètres supplémentaires : `-Encrypted` et `-KmsKeyId`.

Arrêt d'une opération de copie d'une AMI en attente

Vous pouvez arrêter une copie d'AMI en attente comme suit.

Pour arrêter l'opération de copie d'une AMI à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région de destination dans le sélecteur de régions.
3. Dans le panneau de navigation, sélectionnez AMI.
4. Sélectionnez l'AMI à arrêter de copier et choisissez Actions et Annuler l'inscription.
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Continuer.

Pour arrêter une opération de copie d'une AMI à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- deregister-image (AWS CLI)
- Unregister-EC2Image (Outils AWS pour Windows PowerShell)

Obtenir des informations de facturation

Vous pouvez déterminer les détails de la plate-forme et les informations de facturation associées à une AMI avant de lancer une instance à la demande ou une instance Spot, ou d'acheter une Instance réservée. Pour les Instances Spot, vous pouvez utiliser les détails de la plate-forme pour confirmer que l'AMI est prise en charge pour les Instances Spot. Lors de l'achat d'une Instance réservée, vous pouvez vous assurer que, pour Plateforme, vous sélectionnez la valeur correcte qui correspond à Platform details (Détails de la plateforme) sur l'AMI. En connaissant les informations de facturation avant de lancer une instance ou d'acheter une Instance réservée, vous réduisez le risque de lancer des instances par erreur à partir d'AMI incorrectes et d'engager des coûts imprévus.

Pour plus d'informations sur la tarification des instances, consultez Tarification Amazon EC2.

Sommaire

- Champs d'informations de facturation d'AMI (p. 179)
- Détails de la plateforme et valeurs des opérations d'utilisation (p. 180)
- Affichage des détails de la plateforme et des valeurs des opérations d'utilisation (p. 180)
- Confirmation des informations de facturation sur votre facture (p. 182)

Champs d'informations de facturation d'AMI

Les champs suivants fournissent les informations de facturation associées à une AMI :

Platform details (Détails de la plateforme)

Détails de la plateforme associée au code de facturation de l'AMI. Par exemple, `Red Hat Enterprise Linux`.

Usage operation (Opération d'utilisation)

Opération de l'instance Amazon EC2 et code de facturation associé à l'AMI. Par exemple, `RunInstances:0010`. L'opération d'utilisation correspond à la colonne LineItem/Operation de votre

rapport de coût et d'utilisation AWS et de l'API AWS Price List. Pour obtenir la liste des codes Usage operation (Opération d'utilisation), veuillez consulter Détails de la plateforme et valeurs des opérations d'utilisation (p. 180) dans la section suivante.

Vous pouvez afficher ces champs sur la page Instances ou AMI de la console Amazon EC2, ou dans la réponse renvoyée par la commande describe-images.

Détails de la plateforme et valeurs des opérations d'utilisation

Le tableau suivant répertorie les détails de la plateforme et les valeurs des opérations d'utilisation qui peuvent être affichées sur la page Instances ou AMI de la console Amazon EC2, ou dans la réponse renvoyée par la commande describe-images .

Platform details (Détails de la plateforme)	Opération d'utilisation **
Linux/Unix	RunInstances
Red Hat BYOL Linux	RunInstances:00g0
Utilisation de Red Hat Enterprise Linux	RunInstances:0010
SQL Server Enterprise	RunInstances:0100
SQL Server Standard	RunInstances:0004
SQL Server Web	RunInstances:0200
SUSE Linux	RunInstances:000g
Windows	RunInstances:0002
Windows BYOL	RunInstances:0800
Windows avec SQL Server Enterprise *	RunInstances:0102
Windows avec SQL Server Standard *	RunInstances:0006
Windows avec SQL Server Web *	RunInstances:0202

* Si deux licences logicielles sont associées à une AMI, le champ Platform details (Détails de la plateforme) affiche les deux.

** Si vous exécutez des Instances Spot, la valeur `lineitem/Operation` de votre rapport de coût et d'utilisation AWS peut être différent de la valeur Usage operation (Opération d'utilisation) répertoriée ici. Par exemple, si `lineitem/Operation` affiche `RunInstances:0010:SV006`, cela signifie qu'Amazon EC2 exécute une heure d'instance Spot Red Hat Enterprise Linux dans la région US East (Virginie) dans le VPC Zone #6.

Affichage des détails de la plateforme et des valeurs des opérations d'utilisation

Vous pouvez afficher les détails de la plateforme et les valeurs des opérations d'utilisation associées à une AMI à partir de l'AMI ou de l'instance. Vous pouvez afficher ces valeurs dans la console Amazon EC2 ou à l'aide de la l'AWS CLI.

À partir de l'AMI

Pour afficher les détails de la plateforme et des opérations d'utilisation associées à une AMI (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez AMI, puis sélectionnez une AMI.
3. Sous l'onglet Details (Détails) vérifiez les valeurs de Platform details (Détails de la plateforme) et Usage operation (Opération d'utilisation).

Pour afficher les détails de la plateforme et des opérations d'utilisation associées à une AMI (AWS CLI)

Utilisez la commande describe-images.

```
$ aws ec2 describe-images --image-ids ami-0123456789EXAMPLE
```

L'exemple de sortie suivant montre les champs `PlatformDetails` et `UsageOperation`. Dans cet exemple, la plateforme ami-0123456789EXAMPLE est `Red Hat Enterprise Linux`, et la valeur de l'opération d'utilisation et du code de facturation est `RunInstances:0010`.

```
{
    "Images": [
        {
            "VirtualizationType": "hvm",
            "Description": "Provided by Red Hat, Inc.",
            "Hypervisor": "xen",
            "EnaSupport": true,
            "SriovNetSupport": "simple",
            "ImageId": "ami-0123456789EXAMPLE",
            "State": "available",
            "BlockDeviceMappings": [
                {
                    "DeviceName": "/dev/sda1",
                    "Ebs": {
                        "SnapshotId": "snap-111222333444aaabb",
                        "DeleteOnTermination": true,
                        "VolumeType": "gp2",
                        "VolumeSize": 10,
                        "Encrypted": false
                    }
                }
            ],
            "Architecture": "x86_64",
            "ImageLocation": "123456789012/RHEL-8.0.0_HVM-20190618-x86_64-1-Hourly2-GP2",
            "RootDeviceType": "ebs",
            "OwnerId": "123456789012",
            "PlatformDetails": "Red Hat Enterprise Linux",
            "UsageOperation": "RunInstances:0010",
            "RootDeviceName": "/dev/sda1",
            "CreationDate": "2019-05-10T13:17:12.000Z",
            "Public": true,
            "ImageType": "machine",
            "Name": "RHEL-8.0.0_HVM-20190618-x86_64-1-Hourly2-GP2"
        }
    ]
}
```

À partir de l'instance

Pour afficher les détails de la plateforme et des opérations d'utilisation associées à une AMI (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, puis sélectionnez l'instance.
3. Sous l'onglet Details (Détails) vérifiez les valeurs de Platform details (Détails de la plateforme) et Usage operation (Opération d'utilisation).

Pour afficher les détails de la plateforme et des opérations d'utilisation associées à une AMI (console)

Une fois que vous avez lancé une instance, vous pouvez rechercher les informations de facturation en inspectant le champ `billingProducts` dans les métadonnées de l'instance. Pour plus d'informations, consultez Documents d'identité d'instance (p. 728). Vous pouvez également utiliser la commande describe-instances pour obtenir l'ID d'AMI de l'instance, puis utiliser la commande describe-images comme décrit dans la procédure précédente, pour obtenir les informations de facturation à partir des champs `PlatformDetails` et `UsageOperation` dans la réponse.

Confirmation des informations de facturation sur votre facture

Pour vous assurer que vous n'encourez pas de coûts imprévus, vous pouvez confirmer que les informations de facturation concernant une instance et figurant dans votre rapport de coût et d'utilisation AWS correspondent aux informations de facturation associées à l'AMI que vous avez utilisée pour lancer l'instance. Pour confirmer les informations de facturation, recherchez l'ID d'instance dans votre rapport de coût et d'utilisation et vérifiez la valeur correspondante dans la colonne `lineitem/Operation`. La valeur doit correspondre à la valeur de Usage operation (Opération d'utilisation) associée à l'AMI.

Par exemple, l'AMI, `ami-0123456789EXAMPLE`, contient les informations de facturation suivantes : Platform details (Détails de la plateforme) = `Red Hat Enterprise Linux` et Usage operation (Opération d'utilisation) = `RunInstances:0010`. Si vous avez lancé une instance à l'aide de cette AMI, vous pouvez trouver l'ID d'instance dans votre rapport d'utilisation et de coût et vérifier la valeur correspondante dans la colonne `lineitem/Operation`. Dans cet exemple. la valeur devrait être `RunInstances:0010`.

Annulation de l'inscription de votre AMI Linux

Vous pouvez annuler l'inscription de votre AMI lorsque vous avez terminé de l'utiliser. Après cette opération, vous ne pouvez pas utiliser l'AMI pour lancer de nouvelles instances.

Lorsque vous annulez l'inscription d'une AMI, cela n'affecte pas les instances déjà lancées à partir de l'AMI. Les coûts d'utilisation seront toujours facturés pour ces instances. Ainsi, si vous n'avez plus besoin de ces instances, vous devez y mettre fin.

La procédure utilisée pour nettoyer votre AMI varie si elle est basée sur des volumes Amazon EBS ou sur un stockage d'instance. Pour plus d'informations, consultez Sélection du type de périphérique racine de votre AMI (p. 107).

> Note
>
> Une AMI doit appartenir à votre compte afin de pouvoir la désenregistrer.

Sommaire

- Nettoyage de vos AMI basées sur des volumes Amazon EBS (p. 183)
- Nettoyage de votre AMI basée sur le stockage d'instance (p. 184)

Nettoyage de vos AMI basées sur des volumes Amazon EBS

Lorsque vous annulez l'inscription d'une AMI basée sur des volumes Amazon EBS, cela n'affecte pas le ou les instantanés qui ont été créés pour le ou les volumes de l'instance pendant le processus de création de l'AMI. Les coûts de stockage continueront de vous être facturés pour cet instantané. Aussi, si vous n'avez plus besoin de ces instantanés, vous devez les supprimer.

Le schéma suivant illustre le processus de nettoyage de votre AMI basée sur des volumes Amazon EBS.

Pour nettoyer votre AMI basée sur des volumes Amazon EBS

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez AMI. Sélectionnez l'AMI et notez son ID, ce qui vous aidera à identifier l'instantané souhaité à la prochaine étape. Choisissez Actions, puis Annuler l'inscription. Lorsque vous êtes invité à confirmer l'opération, choisissez Continuer.

   Plusieurs minutes peuvent être nécessaires pour que la console supprime l'AMI de la liste. Choisissez Refresh pour actualiser le statut.
3. Dans le panneau de navigation, choisissez Instantanés, puis sélectionnez l'instantané (recherchez l'ID d'AMI dans la colonne Description). Choisissez Actions, puis Delete Snapshot. Lorsque vous êtes invité à confirmer l'opération, choisissez Yes, Delete.
4. (Facultatif) Lorsque vous n'avez plus besoin d'utiliser une instance lancée à partir de l'AMI, mettez-la hors service. Dans le panneau de navigation, choisissez Instances. Sélectionnez l'instance, choisissez État de l'instance, Résilier l'instance. Lorsque vous êtes invité à confirmer, choisissez Terminate (Mettre fin).

Nettoyage de votre AMI basée sur le stockage d'instance

Lorsque vous annulez l'inscription d'une AMI basée sur le stockage d'instance, cela n'affecte pas les fichiers que vous avez téléchargé sur Amazon S3 lorsque vous avez créé l'AMI. Les coûts d'utilisation seront toujours facturés pour ces fichiers dans Amazon S3. Ainsi, si vous n'avez plus besoin de ces fichiers, vous devez les supprimer.

Le schéma suivant illustre le processus de nettoyage de votre AMI basée sur le stockage d'instance.

Your AMI, its bundle in Amazon S3, and an instance launched from the AMI

Pour nettoyer votre AMI basée sur le stockage d'instance

1. Annulez l'inscription de l'AMI avec la commande deregister-image comme suit.

```
aws ec2 deregister-image --image-id ami_id
```

2. Supprimez le bundle dans Amazon S3 avec la commande ec2-delete-bundle (p. 158) (outils AMI) comme suit.

```
ec2-delete-bundle -b myawsbucket/myami -a your_access_key_id -s your_secret_access_key
 -p image
```

3. (Facultatif) Lorsque vous n'avez plus besoin d'utiliser une instance lancée à partir de l'AMI, vous pouvez la quitter avec la commande -terminate-instances comme suit.

```
aws ec2 terminate-instances --instance-ids instance_id
```

4. (Facultatif) Lorsque vous n'avez plus besoin d'utiliser le compartiment Amazon S3 dans lequel vous avez téléchargé le groupe, vous pouvez supprimer le compartiment. Pour supprimer un compartiment Amazon S3, ouvrez la console Amazon S3, sélectionnez le compartiment, choisissez Actions, puis Delete.

Amazon Linux

Amazon Linux est fourni par Amazon Web Services (AWS). Il est conçu pour offrir un environnement d'exécution stable, sécurisé et très performant pour des applications s'exécutant sur Amazon EC2. Il comporte aussi des packages qui permettent une intégration facile à AWS, notamment des outils de configuration du lancement et plusieurs bibliothèques et outils AWS populaires. AWS fournit des mises à jour continues de sécurité et de maintenance pour toutes les instances s'exécutant sur Amazon Linux. Beaucoup applications développées sur CentOS (et des distributions similaires) s'exécutent sur Amazon Linux.

Sommaire

- Disponibilité Amazon Linux (p. 186)
- Connexion à une instance Amazon Linux (p. 186)
- Identification des images Amazon Linux (p. 187)
- Outils de ligne de commande AWS (p. 188)
- Référentiel de package (p. 189)
- Bibliothèque Extras (Amazon Linux 2) (p. 191)
- Accès aux packages source à des fins de référence (p. 192)
- cloud-init (p. 192)
- Abonnement aux notifications Amazon Linux (p. 194)
- Exécution d'Amazon Linux 2 en tant que machine virtuelle sur site (p. 196)
- Kernel Live Patching sur Amazon Linux 2 (p. 200)

Disponibilité Amazon Linux

AWS fournit Amazon Linux 2 et l'Amazon Linux AMI. Si vous migrez à partir d'une autre distribution Linux vers Amazon Linux, nous vous recommandons de migrer vers Amazon Linux 2.

La dernière version de l'Amazon Linux AMI, 2018.03, atteindra la fin de la prise en charge standard le 31 décembre 2020. Pour plus d'informations, consultez le billet de blog sur la fin de vie de l'Amazon Linux AMI. Si vous utilisez actuellement l'Amazon Linux AMI, nous vous recommandons de migrer vers Amazon Linux 2. Pour migrer vers Amazon Linux 2, lancez une instance ou créez une machine virtuelle à l'aide de l'image Amazon Linux 2 actuelle. Installez vos applications, ainsi que tous les packages requis. Testez votre application et apportez les modifications requises pour que celle-ci s'exécute sur Amazon Linux 2.

Pour plus d'informations, consultez Amazon Linux 2 et Amazon Linux AMI. Pour les images de conteneur Docker Amazon Linux, consultez amazonlinux dans Docker Hub.

Connexion à une instance Amazon Linux

Amazon Linux n'autorise pas SSH racine à distance par défaut. De plus, l'authentification par mot de passe est désactivée pour empêcher les attaques de force sur les mots de passe. Pour activer les connexions

SSH à une instance Amazon Linux, vous devez fournir votre paire de clés à l'instance lors du lancement. Vous devez aussi définir le groupe de sécurité utilisé pour lancer votre instance afin d'autoriser l'accès SSH. Par défaut, le seul compte qui peut se connecter à distance en utilisant SSH est ec2-user. Ce compte possède également des privilèges sudo. Pour activer la connexion racine à distance, ayez à l'esprit qu'elle est moins sécurisée que l'utilisation de paires de clés et d'un utilisateur secondaire.

Identification des images Amazon Linux

Chaque image contient un fichier `/etc/image-id` unique qui l'identifie. Ce fichier contient les informations suivantes sur l'image :

- `image_name`, `image_version`, `image_arch` — Valeurs issues de la recette de création qu'Amazon a utilisée pour créer l'image.
- `image_stamp` — Valeur hexadécimale aléatoire unique qui a été générée pendant la création de l'image.
- `image_date` — Heure UTC de la création de l'image, au format AAAAMMJJhhmmss
- `recipe_name`, `recipe_id` — Nom et ID de la recette de création qu'Amazon a utilisée pour créer l'image.

Amazon Linux contient un fichier `/etc/system-release` qui spécifie la version actuelle qui est installée. Ce fichier est mis à jour à l'aide de la commande yum et fait partie du fichier RPM `system-release`.

Amazon Linux contient aussi une version lisible par la machine du fichier `/etc/system-release` qui suit la spécification CPE. Consultez `/etc/system-release-cpe`.

Amazon Linux 2

Voici un exemple de fichier `/etc/image-id` pour la version actuelle d'Amazon Linux 2 :

```
[ec2-user ~]$ cat /etc/image-id
image_name="amzn2-ami-hvm"
image_version="2"
image_arch="x86_64"
image_file="amzn2-ami-hvm-2.0.20180810-x86_64.xfs.gpt"
image_stamp="8008-2abd"
image_date="20180811020321"
recipe_name="amzn2 ami"
recipe_id="c652686a-2415-9819-65fb-4dee-9792-289d-1e2846bd"
```

Voici un exemple de fichier `/etc/system-release` pour la version actuelle d'Amazon Linux 2 :

```
[ec2-user ~]$ cat /etc/system-release
Amazon Linux 2
```

Voici un exemple de fichier `/etc/os-release` pour Amazon Linux 2 :

```
[ec2-user ~]$ cat /etc/os-release
NAME="Amazon Linux"
VERSION="2"
ID="amzn"
ID_LIKE="centos rhel fedora"
VERSION_ID="2"
PRETTY_NAME="Amazon Linux 2"
ANSI_COLOR="0;33"
CPE_NAME="cpe:2.3:o:amazon:amazon_linux:2"
```

```
HOME_URL="https://amazonlinux.com/"
```

Amazon Linux AMI

Voici un exemple de fichier `/etc/image-id` pour l'Amazon Linux AMI actuelle :

```
[ec2-user ~]$ cat /etc/image-id
image_name="amzn-ami-hvm"
image_version="2018.03"
image_arch="x86_64"
image_file="amzn-ami-hvm-2018.03.0.20180811-x86_64.ext4.gpt"
image_stamp="cc81-f2f3"
image_date="20180811012746"
recipe_name="amzn ami"
recipe_id="5b283820-dc60-a7ea-d436-39fa-439f-02ea-5c802dbd"
```

Voici un exemple de fichier `/etc/system-release` pour l'Amazon Linux AMI actuelle :

```
[ec2-user ~]$ cat /etc/system-release
Amazon Linux AMI release 2018.03
```

Outils de ligne de commande AWS

Les outils de ligne de commande suivants pour l'intégration et l'utilisation AWS sont inclus dans l'AMI Amazon Linux ou dans les référentiels par défaut pour Amazon Linux 2. Pour obtenir la liste de packages complète de l'Amazon Linux AMI, consultez Packages Amazon Linux AMI 2017.09.

- aws-amitools-ec2
- aws-apitools-as
- aws-apitools-cfn
- aws-apitools-elb
- aws-apitools-mon
- aws-cfn-bootstrap
- aws-cli

Amazon Linux 2 et les versions minimales d'Amazon Linux (`amzn-ami-minimal-*` et `amzn2-ami-minimal-*`) ne contiennent pas toujours tous ces packages. Par contre, vous pouvez les installer à partir des référentiels par défaut à l'aide de la commande suivante :

```
[ec2-user ~]$ sudo yum install -y package_name
```

Pour les instances lancées en utilisant les rôles IAM, un script simple a été inclus pour préparer `AWS_CREDENTIAL_FILE`, `JAVA_HOME`, `AWS_PATH`, `PATH` et les variables d'environnement spécifiques aux produits après l'installation d'un fichier d'identification pour simplifier la configuration de ces outils.

De plus, pour permettre l'installation de plusieurs version des outils API et AMI, nous avons placé des liens symboliques sur les versions souhaitées de ces outils dans `/opt/aws`, comme indiqué ici :

`/opt/aws/bin`

Liens symboliques vers les répertoires `/bin` dans chacun des répertoires des outils installés.

`/opt/aws/{apitools|amitools}`

Les produits sont installés dans les répertoires du type *nom-version* et un lien symbolique *nom* qui est attaché à la version la plus récemment installée.

`/opt/aws/{apitools|amitools}/name/environment.sh`

Utilisé par `/etc/profile.d/aws-apitools-common.sh` pour définir les variables d'environnement spécifiques aux produits, comme `EC2_HOME`.

Référentiel de package

Amazon Linux 2 et l'Amazon Linux AMI sont conçus pour être utilisés avec des référentiels de package en ligne hébergés dans chaque région AWS Amazon EC2. Ces référentiels fournissent les mises à jour continues aux packages dans Amazon Linux 2 et dans l'Amazon Linux AMI, ainsi que l'accès à des centaines d'applications serveur Open Source communes supplémentaires. Les référentiels sont disponibles dans toutes les régions et sont accessibles à l'aide des outils de mise à jour yum. L'hébergement de référentiels dans chaque région nous permet de déployer rapidement les mises à jour et sans aucuns frais de transfert de données.

Amazon Linux 2 et l'Amazon Linux AMI sont régulièrement mis à jour grâce avec des améliorations de la sécurité et des fonctions. Si vous n'avez pas besoin de conserver les données ou les personnalisations pour vos instances , il vous suffit de lancer de nouvelles instances à l'aide de l'AMI actuelle. Si vous devez conserver les données ou les personnalisations pour vos instances, vous pouvez maintenir ces instances via les référentiels de package d'Amazon Linux. Ces référentiels contiennent tous les packages mis à jour. Vous pouvez choisir d'appliquer ces mises à jour à vos instances en cours d'exécution. Les anciennes versions de l'AMI et des packages de mise à jour continuent d'être disponibles à l'utilisation, même lorsque de nouvelles versions sont proposées.

Important

Votre instance doit avoir un accès Internet sortant pour accéder au référentiel.

Pour installer des packages, utilisez la commande suivante :

```
[ec2-user ~]$ sudo yum install package
```

Pour l'Amazon Linux AMI, accédez au référentiel EPEL (Extra Packages for Enterprise Linux) qui est configuré, mais pas activé par défaut. Amazon Linux 2 n'est pas configuré pour utiliser le référentiel EPEL. Le référentiel EPEL fournit des packages tiers en plus de ceux qui sont dans les référentiels. Les packages tiers ne sont pas pris en charge par AWS. Vous pouvez activer le référentiel EPEL avec les commandes suivantes :

- Dans Amazon Linux 2:

```
[ec2-user ~]$ sudo yum install https://dl.fedoraproject.org/pub/epel/epel-release-
latest-7.noarch.rpm
```

- Pour l'Amazon Linux AMI :

```
[ec2-user ~]$ sudo yum-config-manager --enable epel
```

Si vous découvrez qu'Amazon Linux ne contient pas une application dont vous avez besoin, vous pouvez simplement installer l'application directement sur votre instance Amazon Linux. Amazon Linux utilise les fichiers RPM et la commande yum pour la gestion des packages. Il s'agit probablement de la façon la

plus simple d'installer de nouvelles applications. Vous devriez toujours vérifier si une application est déjà disponible dans notre référentiel central d'Amazon Linux, car beaucoup d'applications sont disponibles ici. Ces applications peuvent être facilement ajoutées à votre instance Amazon Linux.

Pour charger vos applications sur une instance Amazon Linux en cours d'exécution, utilisez scp ou sftp, puis configurez l'application en vous connectant à votre instance. Vos applications peuvent aussi être chargées pendant le lancement de l'instance en utilisant l'action PACKAGE_SETUP à partir du package cloud-init intégré. Pour plus d'informations, consultez cloud-init (p. 192).

Mises à jour de sécurité

Les mises à jour de sécurité sont fournies à l'aide des référentiels de package, ainsi que les AMI mises à jour. Les alertes de sécurité sont publiées dans le Centre de sécurité Amazon Linux. Pour plus d'informations sur les stratégies de sécurité AWS ou pour signaler un problème de sécurité, consultez le centre de sécurité AWS.

Amazon Linux est configuré pour télécharger et installer les mises à jour de sécurité critiques ou importantes au moment du lancement. Nous vous recommandons d'effectuer les mises à jour nécessaires pour votre cas d'utilisation après le lancement. Par exemple, vous pouvez souhaiter appliquer toutes les mises à jour (pas seulement les mises à jour de sécurité) au moment du lancement, ou évaluer chaque mise à jour et appliquer à votre système uniquement celles qui sont applicables. Ceci est contrôlé à l'aide du paramètre cloud-init suivant : `repo_upgrade`. L'extrait de configuration cloud-init suivant montre comment modifier les paramètres dans le texte de données utilisateur que vous transmettez à l'initialisation de votre instance :

```
#cloud-config
repo_upgrade: security
```

Les valeurs possibles pour `repo_upgrade` sont les suivantes :

`critical`

Appliquez les mises à jour de sécurité critiques en attente.

`important`

Appliquez les mises à jour de sécurité importantes et critiques.

`medium`

Appliquez les mises à jour de sécurité critiques, importantes et moyennes.

`low`

Appliquez toutes les mises à jour de sécurité en attente, y compris les mises à jour de sécurité de faible gravité.

`security`

Appliquez les mises à jour critiques ou importantes indiquées par Amazon comme étant des mises à jour de sécurité.

`bugfix`

Appliquez les mises à jour indiquées par Amazon comme étant des correctifs de bogues. Les correctifs de bogues constituent un ensemble plus important de mises à jour ce qui comprend des mises à jour de sécurité et des correctifs pour plusieurs autres bogues mineurs.

`all`

Appliquez toutes les mises à jour disponibles appropriées, peu importe leur classification.

`none`

N'appliquez aucune mise à jour à l'instance au démarrage.

Le paramètre par défaut pour `repo_upgrade` est la sécurité. En effet, si vous ne spécifiez pas une valeur différente dans vos données utilisateur, Amazon Linux effectue par défaut les mises à niveau de sécurité au lancement pour tous les packages installés à ce moment. Amazon Linux vous avertit aussi de toutes les mises à jour pour les packages installés en faisant la liste des mises à jour disponibles au moment de la connexion en utilisant le fichier `/etc/motd`. Pour installer ces mises à jour, vous devez exécuter sudo yum upgrade sur l'instance.

Configuration du référentiel

Avec Amazon Linux, les AMI sont traitées comme des instantanés ponctuels, avec une structure de référentiel et de mise à jour qui vous donne toujours les derniers packages lorsque vous exécutez yum update -y.

La structure du référentiel est configurée pour offrir un flux continu de mises à jour qui vous permettent de passer d'une version d'Amazon Linux à la suivante. Par exemple, si vous lancez une instance à partir d'une version plus ancienne de l'Amazon Linux AMI (comme 2017.09 ou antérieures) et exécutez yum update -y, vous obtenez les derniers packages.

Vous pouvez désactiver la propagation des mises à jour en activant la fonction de verrouillage au lancement. La fonction de verrouillage au lancement verrouille l'instance pour recevoir uniquement les mises à jour de la version spécifiée de l'AMI. Par exemple, vous pouvez lancer une AMI 2017.09 et faire en sorte qu'elle ne reçoive que les mises à jour qui ont été publiées avant l'AMI 2018.03 jusqu'à ce que vous soyez prêt à migrer vers l'AMI 2018.03.

> **Important**
>
> Si vous verrouillez une version de référentiels autre que la dernière version, vous ne recevez plus aucune mise à jour. Pour recevoir un flux continu de mises à jour, vous devez utiliser l'AMI la plus récente ou à mettre régulièrement à jour l'AMI avec les référentiels pointant vers la dernière version.

Pour activer la fonctionnalité de verrouillage au lancement dans les nouvelles instances, lancez-la avec les données utilisateur transmises à cloud-init :

```
#cloud-config
repo_releasever: 2017.09
```

Pour relier des instances existantes à leur version d'AMI actuelle

1. Modification `/etc/yum.conf`.
2. Mettez en commentaire `releasever=latest`.
3. Pour vider le cache, exécutez yum clean all.

Bibliothèque Extras (Amazon Linux 2)

Avec Amazon Linux 2, vous pouvez utiliser la bibliothèque Extras pour installer les mises à jour d'application et logicielles sur vos instances. Ces mises à jour logicielles sont appelées rubriques. Vous pouvez installer une version spécifique d'une rubrique ou omettre les informations de version pour utiliser la version la plus récente.

Pour répertorier les rubriques disponibles, utilisez la commande suivante :

```
[ec2-user ~]$ amazon-linux-extras list
```

Pour activer une rubrique et installer la dernière version de son package pour garantir son actualité, utilisez la commande suivante :

```
[ec2-user ~]$ sudo amazon-linux-extras install topic
```

Pour activer des rubriques et installer des versions spécifiques de leurs packages afin de garantir la stabilité, utilisez la commande suivante :

```
[ec2-user ~]$ sudo amazon-linux-extras install topic=version topic=version
```

Pour supprimer un package installé à partir d'une rubrique, utilisez la commande suivante :

```
[ec2-user ~]$ sudo yum remove $(yum list installed | grep amzn2extra-topic | awk '{ print
 $1 }')
```

> Note
>
> Cette commande ne supprime pas les paquets qui ont été installés en tant que dépendances de l'extra.

Pour désactiver une rubrique et rendre les packages inaccessibles au gestionnaire de paquets yum, utilisez la commande suivante :

```
[ec2-user ~]$ sudo amazon-linux-extras disable topic
```

> Important
>
> Cette commande est destinée aux utilisateurs avancés. Une utilisation incorrecte de cette commande peut entraîner des conflits de compatibilité de paquets.

Accès aux packages source à des fins de référence

Vous pouvez voir la source des packages que vous avez installés sur votre instance à des fins de référence en utilisant des outils fournis dans Amazon Linux. Les packages source sont disponibles pour tous les packages compris dans Amazon Linux et le référentiel de package en ligne. Il vous suffit de déterminer le nom du package pour le package source que vous voulez installer et d'utiliser la commande yumdownloader --source pour voir la source dans votre instance en cours d'exécution. Exemples :

```
[ec2-user ~]$ yumdownloader --source bash
```

Le fichier RPM source peut être décompressé et, à des fins de référence, vous pouvez voir l'arborescence source en utilisant les outils RPM standard. Après le débogage, le package peut être utilisé.

cloud-init

Le package cloud-init est une application open source réalisée par Canonical, qui est utilisée pour amorcer les images Linux dans un environnement de cloud computing, comme Amazon EC2. Amazon Linux contient une version personnalisée de cloud-init. Elle vous permet de spécifier des actions qui devraient arriver à votre instance au moment du démarrage. Vous pouvez transmettre les actions souhaitées à cloud-init via les champs de données utilisateur lors du lancement d'une instance. Cela signifie que vous pouvez

utiliser des AMI communes pour plusieurs cas d'utilisation et les configurer dynamiquement au démarrage. Amazon Linux utilise aussi cloud-init pour effectuer une configuration initiale du compte ec2-user.

Pour plus d'informations, consultez la documentation cloud-init.

Amazon Linux utilise les actions cloud-init trouvées dans `/etc/cloud/cloud.cfg.d` et `/etc/cloud/cloud.cfg`. Vous pouvez créer vos propres fichiers d'actions cloud-init dans `/etc/cloud/cloud.cfg.d`. Tous les fichiers dans ce répertoire sont lus par cloud-init. Ils sont lus en ordre lexical, et les fichiers plus récents remplacent les valeurs des fichiers plus anciens.

Le package cloud-init effectue des tâches de configuration communes (et d'autres tâches) pour les instances au démarrage :

- Définir les paramètres régionaux par défaut.
- Définir le nom d'hôte.
- Analyser et gérer les données utilisateur.
- Générer des clés SSH privées d'hôte.
- Ajouter des clés SSH publiques d'utilisateur à `.ssh/authorized_keys` pour une connexion et une administration faciles.
- Préparer les référentiels pour la gestion des packages.
- Gérer les actions de package définies dans les données utilisateur.
- Exécuter les scripts utilisateur trouvés dans les données utilisateur.
- Monter les volumes de stockage d'instance, le cas échéant.
  - Par défaut, le volume de stockage d'instance `ephemeral0` est monté sur `/media/ephemeral0` s'il est présent et contient un système de fichiers valide ; sinon, il n'est pas monté.
  - Par défaut, les volumes d'échange associés à l'instance sont montés (uniquement pour les types d'instance `m1.small` et `c1.medium`).
  - Vous pouvez remplacer le montage de volume de stockage d'instance par défaut avec la directive cloud-init suivante :

    ```
    #cloud-config
    mounts:
    - [ ephemeral0 ]
    ```

    Pour plus de contrôle sur les montages, consultez Mounts dans la documentation cloud-init.
  - Les volumes de stockage d'instance qui prennent en charge la commande TRIM ne sont pas formatés au lancement de l'instance. Vous devez donc les partitionner et les formater pour pouvoir les monter et les utiliser. Pour plus d'informations, consultez Prise en charge de TRIM sur les volumes de stockage d'instance (p. 1278). Vous pouvez utiliser le module `disk_setup` pour partitionner et formater vos volumes de stockage d'instance au démarrage. Pour plus d'informations, consultez Disk Setup dans la documentation cloud-init.

Formats de données utilisateur pris en charge

Le package cloud-init prend en charge la gestion des données utilisateur sous différents formats :

- Gzip
  - Si les données utilisateur sont compressés avec gzip, cloud-init décompresse les données et les gère de façon appropriée.
- Fichier MIME en plusieurs parties
  - En utilisant un fichier MIME en plusieurs parties, vous pouvez spécifier plus d'un type de données. Par exemple, vous pourriez spécifier à la fois un script de données utilisateur et un type de configuration

Cloud. Chaque partie de ce fichier peut être traitée par cloud-init s'il s'agit d'un des formats pris en charge.

- Décodage Base64
  - Si les données utilisateur sont sous la forme encodée base64, cloud-init détermine s'il peut comprendre les données décodées comme l'un des types pris en charge. S'il comprend les données décodées, il décode les données et les gère de façon appropriée. Si non, il renvoie les données base64 intactes.
- Script de données utilisateur
  - Commence par `#!` ou `Content-Type: text/x-shellscript`.
  - Le script est exécuté par `/etc/init.d/cloud-init-user-scripts` pendant le premier cycle de démarrage. Cela se produit tard dans le processus de démarrage (après l'exécution des actions de configuration initiales).
- Fichier d'inclusion
  - Commence par `#include` ou `Content-Type: text/x-include-url`.
  - Ce contenu est un fichier d'inclusion. Le fichier contient une liste d'URL, une par ligne. Chacune des URL est lue, et leur contenu est passé par le même ensemble de règles. Le contenu lu à partir de l'URL peut être compressé avec gzip, sous forme de fichier MIME en plusieurs parties ou de texte brut.
- Données de configuration Cloud
  - Commence par `#cloud-config` ou `Content-Type: text/cloud-config`.
  - Ce contenu correspond aux données de configuration Cloud. Pour obtenir un exemple commenté des formats de configuration pris en charge, consultez les exemples.
- Tâche Upstart
  - Commence par `#upstart-job` ou `Content-Type: text/upstart-job`.
  - Ce contenu est stocké dans un fichier dans `/etc/init`, et upstart consomme le contenu conformément aux autres tâches upstart.
- Cloud Boothook
  - Commence par `#cloud-boothook` ou `Content-Type: text/cloud-boothook`.
  - Ce contenu correspond aux données boothook. Il est stocké dans un fichier sous `/var/lib/cloud`, puis exécuté immédiatement.
  - Il s'agit du « hook » le plus récent disponible. Il n'existe aucun mécanisme proposé pour l'exécuter seulement une fois. Le boothook doit s'en occuper lui-même. Il est fourni avec l'ID d'instance dans la variable d'environnement `INSTANCE_ID`. Utilisez cette variable pour fournir un ensemble de données boothook à exécuter une fois par instance.

Abonnement aux notifications Amazon Linux

Pour être informé de la publication de nouvelles AMI, vous pouvez vous abonner à l'aide de Amazon SNS.

Pour s'abonner aux notifications Amazon Linux

1. Ouvrez la console Amazon SNS à l'adresse https://console.aws.amazon.com/sns/v3/home.
2. Dans la barre de navigation, remplacez la région par USA Est (Virginie du Nord) si besoin. Vous devez sélectionner la région dans laquelle la notification SNS à laquelle vous vous abonnez a été créée.
3. Dans le panneau de navigation, choisissez Abonnements, puis Créer un abonnement.
4. Dans la boîte de dialogue Créer un abonnement, procédez comme suit :
   a. [Amazon Linux 2] Pour ARN de la rubrique, copiez et collez l'Amazon Resource Name (ARN) suivant : **`arn:aws:sns:us-east-1:137112412989:amazon-linux-2-ami-updates`**.
   b. [Amazon Linux] Pour ARN de la rubrique, copiez et collez l'Amazon Resource Name (ARN) suivant : **`arn:aws:sns:us-east-1:137112412989:amazon-linux-ami-updates`**.

   c. Pour Protocole, choisissez E-mail.
   d. Pour Point de terminaison, entrez une adresse e-mail que vous pouvez utiliser pour recevoir les notifications.
   e. Choisissez Créer un abonnement.
5. Vous recevrez un e-mail de confirmation avec la ligne d'objet « AWS Notification - Subscription Confirmation » (Notification AWS - Confirmation de l'abonnement). Ouvrez l'e-mail et choisissez Confirm subscription (Confirmer l'abonnement) pour terminer votre abonnement.

Chaque fois que des AMI sont publiées, nous envoyons des notifications aux abonnés de la rubrique correspondante. Pour arrêter de recevoir ces notifications, utilisez la procédure suivante pour vous désabonner.

Pour annuler votre abonnement aux notifications Amazon Linux

1. Ouvrez la console Amazon SNS à l'adresse https://console.aws.amazon.com/sns/v3/home.
2. Dans la barre de navigation, remplacez la région par USA Est (Virginie du Nord) si besoin. Vous devez utiliser la région dans laquelle la notification SNS a été créée.
3. Dans le panneau de navigation, sélectionnez Abonnements, sélectionnez l'abonnement, puis Actions, Supprimer des abonnements.
4. Lorsque vous êtes invité à confirmer l'opération, choisissez Supprimer.

Format du message Amazon Linux AMI SNS

Le schéma du message SNS est le suivant.

```
{
    "description": "Validates output from AMI Release SNS message",
    "type": "object",
    "properties": {
        "v1": {
            "type": "object",
            "properties": {
                "ReleaseVersion": {
                    "description": "Major release (ex. 2018.03)",
                    "type": "string"
                },
                "ImageVersion": {
                    "description": "Full release (ex. 2018.03.0.20180412)",
                    "type": "string"
                },
                "ReleaseNotes": {
                    "description": "Human-readable string with extra information",
                    "type": "string"
                },
                "Regions": {
                    "type": "object",
                    "description": "Each key will be a region name (ex. us-east-1)",
                    "additionalProperties": {
                        "type": "array",
                        "items": {
                            "type": "object",
                            "properties": {
                                "Name": {
                                    "description": "AMI Name (ex. amzn-ami-
hvm-2018.03.0.20180412-x86_64-gp2)",
                                    "type": "string"
                                },
                                "ImageId": {
                                    "description": "AMI Name (ex.ami-467ca739)",
```

```
                                    "type": "string"
                                }
                            },
                            "required": [
                                "Name",
                                "ImageId"
                            ]
                        }
                    }
                }
            },
            "required": [
                "ReleaseVersion",
                "ImageVersion",
                "ReleaseNotes",
                "Regions"
            ]
        }
    },
    "required": [
        "v1"
    ]
}
```

Exécution d'Amazon Linux 2 en tant que machine virtuelle sur site

Utilisez les images de machine virtuelle (VM, virtual machine) Amazon Linux 2 pour le développement et les tests sur site. Ces images sont disponibles à l'utilisation sur les plateformes de virtualisation suivantes :

- VMWare
- KVM
- VirtualBox (Oracle VM)
- Microsoft Hyper-V

Pour utiliser les images de machine virtuelle Amazon Linux 2 avec l'une des plateformes de virtualisation prises en charge, procédez comme suit :

- Étape 1 : Préparer l'image de démarrage `seed.iso` (p. 196)
- Étape 2 : Télécharger l'image de la machine virtuelle Amazon Linux 2 (p. 198)
- Étape 3 : Démarrer et se connecter à votre nouvelle machine virtuelle (p. 198)

Étape 1 : Préparer l'image de démarrage `seed.iso`

L'image de démarrage `seed.iso` inclut les informations de configuration initiale requises pour démarrer votre nouvelle machine virtuelle, telles que la configuration réseau, le nom d'hôte et les données utilisateur.

> **Note**
>
> L'image de démarrage `seed.iso` inclut uniquement les informations de configuration requises pour démarrer la machine virtuelle. Elle n'inclut pas les fichiers de système d'exploitation Amazon Linux 2.

Pour générer l'image de démarrage `seed.iso`, vous avez besoin de deux fichiers de configuration :

- `meta-data` — Ce fichier inclut le nom d'hôte et les paramètres de réseau statique pour la machine virtuelle.

- `user-data` — Ce fichier configure les comptes utilisateur et spécifie leurs mots de passe, paires de clés et mécanismes d'accès. Par défaut, l'image de la machine virtuelle Amazon Linux 2 crée un compte utilisateur `ec2-user`. Vous utilisez le fichier de configuration `user-data` pour définir le mot de passe pour le compte utilisateur par défaut.

Pour créer le disque de démarrage **`seed.iso`**

1. Créez un dossier appelé `seedconfig` et accédez à celui-ci.
2. Créez le fichier de configuration `meta-data`.

   a. Créez un fichier nommé `meta-data`.

   b. Ouvrez le fichier `meta-data` à l'aide de l'éditeur de votre choix et ajoutez ce qui suit.

   ```
   local-hostname: vm_hostname
   # eth0 is the default network interface enabled in the image. You can configure
    static network settings with an entry like the following.
   network-interfaces: |
     auto eth0
     iface eth0 inet static
     address 192.168.1.10
     network 192.168.1.0
     netmask 255.255.255.0
     broadcast 192.168.1.255
     gateway 192.168.1.254
   ```

   Remplacez *`vm_hostname`* par le nom d'hôte d'une machine virtuelle de votre choix, et configurez les paramètres réseau comme requis.

   c. Enregistrez et fermez le fichier de configuration `meta-data`.

   Pour obtenir un exemple de fichier de configuration `meta-data` qui spécifie le nom d'hôte d'une machine virtuelle (`amazonlinux.onprem`), configure l'interface réseau par défaut (`eth0`) et spécifie les adresses IP statiques pour les périphériques réseau nécessaires, consultez l'exemple de fichier Seed.iso.

3. Créez le fichier de configuration `user-data`.

   a. Créez un fichier nommé `user-data`.

   b. Ouvrez le fichier `user-data` à l'aide de l'éditeur de votre choix et ajoutez ce qui suit.

   ```
   #cloud-config
   #vim:syntax=yaml
   users:
   # A user by the name `ec2-user` is created in the image by default.
     - default
   chpasswd:
     list: |
       ec2-user:plain_text_password
   # In the above line, do not add any spaces after 'ec2-user:'.
   ```

   Remplacez *`plain_text_password`* par un mot de passe de votre choix pour le compte d'utilisateur `ec2-user` par défaut.

   c. (Facultatif) Par défaut, cloud-init applique les paramètres réseau à chaque démarrage de la machine virtuelle. Ajoutez ce qui suit pour empêcher cloud-init d'appliquer les paramètres réseau à chaque démarrage et conserver les paramètres réseau appliqués lors du premier démarrage.

   ```
   # NOTE: Cloud-init applies network settings on every boot by default. To retain
    network settings from first
   ```

```
boot, add following 'write_files' section:
write_files:
  - path: /etc/cloud/cloud.cfg.d/80_disable_network_after_firstboot.cfg
    content: |
      # Disable network configuration after first boot
      network:
        config: disabled
```

d. Enregistrez et fermez le fichier de configuration `user-data`.

De même, vous pouvez créer des comptes d'utilisateur supplémentaires et spécifier leurs mécanismes d'accès, mots de passe et paires de clés. Pour plus d'informations sur les directives prises en charge, consultez Modules. Pour obtenir un exemple de fichier `user-data` permettant de créer trois utilisateurs supplémentaires et de spécifier un mot de passe personnalisé pour le compte d'utilisateur `ec2-user` par défaut, consultez le fichier d'exemple Seed.iso.

4. Créez l'image de démarrage `seed.iso` en utilisant les fichiers de configuration `meta-data` et `user-data`.

   Pour Linux, utilisez un outil tel que genisoimage. Naviguez dans le dossier `seedconfig` et exécutez la commande suivante.

```
$ genisoimage -output seed.iso -volid cidata -joliet -rock user-data meta-data
```

   Pour macOS, utilisez un outil tel que hdiutil. Remontez d'un niveau à partir du dossier `seedconfig` et exécutez la commande suivante.

```
$ hdiutil makehybrid -o seed.iso -hfs -joliet -iso -default-volume-name cidata
 seedconfig/
```

Étape 2 : Télécharger l'image de la machine virtuelle Amazon Linux 2

Nous proposons une image de la machine virtuelle Amazon Linux 2 différente pour chacune des plateformes de virtualisation prises en charges. Téléchargez l'image de la machine virtuelle adéquate pour la plateforme choisie :

- VMWare
- KVM
- Oracle VirtualBox
- Microsoft Hyper-V

Étape 3 : Démarrer et se connecter à votre nouvelle machine virtuelle

Pour démarrer et vous connecter à votre nouvelle machine virtuelle, vous devez avoir l'image de démarrage `seed.iso` (créée à l'étape 1) et une image de la machine virtuelle Amazon Linux 2 (téléchargée à l'étape 2). Cette procédure varie selon la plateforme de machine virtuelle que vous choisissez.

VMWare vSphere

L'image VM pour VMware est disponible au format OVF.

Pour démarrer la machine virtuelle à l'aide de VMware vSphere

1. Créez une banque de données pour le fichier `seed.iso` ou ajoutez-la à une banque de données existante.
2. Déployez le modèle OVF, mais ne démarrez pas encore la machine virtuelle.
3. Dans le panneau Navigateur, cliquez avec le bouton droit sur la nouvelle machine virtuelle et choisissez Modifier les paramètres.
4. Sous l'onglet Matériel virtuel pour Nouvel appareil, choisissez Lecteur de CD/DVD, puis Ajouter.
5. Pour Nouveau lecteur de CD/DVD, choisissez Fichier ISO de banque de données. Sélectionnez la banque de données à laquelle vous avez ajouté le fichier `seed.iso`, accédez au fichier `seed.iso` et sélectionnez-le, puis choisissez OK.
6. Pour Nouveau lecteur CD/DVD, sélectionnez Connexion, puis OK.

Après avoir associé la banque de données à la machine virtuelle, vous devriez pouvoir le démarrer.

KVM

Pour démarrer la machine virtuelle à l'aide de KVM

1. Ouvrez l'assistant Créer une machine virtuelle.
2. Pour l'étape 1, choisissez Importer une image disque existante.
3. Pour l'étape 2, accédez à l'image de la machine virtuelle et sélectionnez-la. Pour Type de système d'exploitation et Version, choisissez Linux et Red Hat Enterprise Linux 7.0, respectivement.
4. Pour l'étape 3, spécifiez la quantité de RAM et le nombre de processeurs à utiliser.
5. Pour l'étape 4, entrez un nom pour la nouvelle machine virtuelle et sélectionnez Personnaliser la configuration avant l'installation, puis choisissez Terminer.
6. Dans la fenêtre Configuration de la machine virtuelle, choisissez Ajouter du matériel.
7. Dans la fenêtre Ajouter un nouveau matériel virtuel, choisissez Stockage.
8. Dans la configuration de stockage, choisissez Sélectionner ou créer un stockage personnalisé. Pour Type d'appareil, choisissez Appareil CDROM. Choisissez Gérer, Parcourir en local, puis accédez au fichier `seed.iso` et sélectionnez-le. Choisissez Finish.
9. Choisissez Commencer l'installation.

Oracle VirtualBox

Pour démarrer la machine virtuelle à l'aide d'Oracle VirtualBox

1. Ouvrez Oracle VirtualBox et choisissez New (Nouveau).
2. Dans Name (Nom), entrez un nom descriptif pour la machine virtuelle et dans Type et Version, sélectionnez respectivement Linux et Red Hat (64 bit). Choisissez Continue.
3. Pour Memory size (Taille de la mémoire), spécifiez la quantité de mémoire à allouer à la machine virtuelle, puis choisissez Continue (Continuer).
4. Pour Hard disk (Disque dur), choisissez Use an existing virtual hard disk file (Utiliser un fichier de disque dur virtuel existant), recherchez et ouvrez l'image de machine virtuelle, puis choisissez Create (Créer).
5. Avant de démarrer la machine virtuelle, vous devez charger le fichier `seed.iso` dans le lecteur optique virtuel de la machine virtuelle :
   a. Sélectionnez la nouvelle machine virtuelle, choisissez Paramètres, puis Stockage.
   b. Dans la liste Storage Devices (Appareils de stockage), sous Controller: IDE (Contrôleur : IDE), choisissez le lecteur optique Empty (Vide).

c. Dans la section Attributs du lecteur optique, cliquez sur le bouton Parcourir, sélectionnez Choisir un fichier de disque optique virtuel, puis sélectionnez le fichier `seed.iso`. Cliquez sur OK pour appliquer les modifications et fermer les paramètres.

Après avoir ajouté le fichier `seed.iso` au lecteur optique virtuel, vous devriez pouvoir démarrer la machine virtuelle.

Microsoft Hyper-V

L'image VM pour Microsoft Hyper-V est compressée dans un fichier zip. Vous devez extraire le contenu du fichier .zip.

Pour démarrer la machine virtuelle à l'aide de Microsoft Hyper-V

1. Ouvrez New Virtual Machine Wizard (Nouvel assistant de machine virtuelle).
2. Lorsque vous êtes invité à sélectionner une génération, sélectionnez Génération 1.
3. Lorsque vous êtes invité à configurer la carte réseau, pour Connexion, choisissez Externe.
4. Lorsque vous êtes invité à connecter un disque dur virtuel, choisissez Utiliser un disque dur virtuel existant, choisissez Parcourir, puis accédez à et sélectionnez l'image de la machine virtuelle. Choisissez Terminer pour créer la machine virtuelle.
5. Cliquez avec le bouton droit sur la nouvelle machine virtuelle et choisissez Paramètres. Dans la fenêtre Paramètres, sous Contrôleur IDE 1, choisissez Lecteur de DVD.
6. Pour le lecteur de DVD, choisissez Fichier Image, puis recherchez et sélectionnez le fichier `seed.iso`.
7. Appliquez les modifications et démarrez la machine virtuelle.

Après le démarrage de la machine virtuelle, connectez-vous avec l'un des comptes d'utilisateur définis dans le fichier de configuration `user-data`. Pour les plateformes de virtualisation autres que VMWare, vous pouvez déconnecter l'image de démarrage `seed.iso` de la machine virtuelle après vous être connecté pour la première fois.

Kernel Live Patching sur Amazon Linux 2

Kernel Live Patching pour Amazon Linux 2 vous permet d'appliquer des correctifs de vulnérabilité de sécurité et de bogues critiques à un noyau Linux en cours d'exécution, sans redémarrer ni interrompre les applications en cours d'exécution. Cela vous permet de bénéficier d'une meilleure disponibilité des services et des applications, tout en gardant votre infrastructure sécurisée et à jour.

AWS propose deux types de correctifs à chaud du noyau pour Amazon Linux 2 :

- Mises à jour de sécurité — Inclut les mises à jour pour les vulnérabilités et expositions communes (CVE) Linux. Ces mises à jour sont généralement jugées importantes ou critiques à l'aide des évaluations Amazon Linux de sécurité. Elles correspondent généralement à un score CVSS (Common Vulnerability Scoring System) égal à 7 ou plus. Dans certains cas, AWS peut fournir des mises à jour avant qu'un CVE ne soit affecté. Dans ces cas, les correctifs peuvent apparaître comme des correctifs de bogues.
- Corrections de bugs — Inclut des correctifs pour les bogues critiques et les problèmes de stabilité qui ne sont pas associés aux CVE.

AWS fournit des correctifs Kernel Live Patching pour une version du noyau Amazon Linux 2 jusqu'à 3 mois après sa publication. Après la période de 3 mois, vous devez effectuer une mise à jour vers une version ultérieure du noyau pour continuer à recevoir les correctifs à chaud du noyau.

Les correctifs Kernel Live Patching pour Amazon Linux 2 sont disponibles sous forme de paquets RPM signés dans les référentiels Amazon Linux 2 existants. Les correctifs peuvent être installés sur des

instances individuelles à l'aide de workflows yum existants, ou sur un groupe d'instances gérées à l'aide d'AWS Systems Manager.

Kernel Live Patching sur Amazon Linux 2 est fourni sans frais supplémentaires.

Rubriques

- Configurations et conditions préalables prises en charge (p. 201)
- Utilisation de l'application Kernel Live Patching (p. 201)
- Limites (p. 205)
- Questions fréquentes (FAQ) (p. 205)

Configurations et conditions préalables prises en charge

Les correctifs Kernel Live Patching sont pris en charge sur les instances Amazon EC2 et les machines virtuelles locales (p. 196) exécutant Amazon Linux 2.

Pour utiliser Kernel Live Patching sur Amazon Linux 2, vous devez utiliser :

- Architecture 64 bits (x86_64) prise en charge par Amazon Linux 2
- Amazon Linux 2 avec la version du noyau `4.14.165-131.185` ou ultérieure

> **Note**
>
> L'architecture ARM 64 bits (arm64) n'est pas prise en charge.

Utilisation de l'application Kernel Live Patching

Vous pouvez activer et utiliser Kernel Live Patching sur des instances individuelles à l'aide de la ligne de commande de l'instance elle-même, ou activer et utiliser Kernel Live Patching sur un groupe d'instances gérées à l'aide d'AWS Systems Manager.

Les sections suivantes expliquent comment activer et utiliser Kernel Live Patching sur des instances individuelles à l'aide de la ligne de commande.

Pour plus d'informations sur l'activation et l'utilisation de Kernel Live Patching sur un groupe d'instances gérées, veuillez consulter Utilisation de Kernel Live Patching sur les instances Amazon Linux 2 dans le Guide de l'utilisateur AWS Systems Manager.

Rubriques

- Activation de l'application Kernel Live Patching (p. 201)
- Affichage des correctifs à chaud du noyau disponibles (p. 203)
- Application des correctifs à chaud du noyau (p. 203)
- Afficher les correctifs à chaud du noyau appliqués (p. 204)
- Désactivation de Kernel Live Patching (p. 205)

Activation de l'application Kernel Live Patching

Kernel Live Patching est désactivé par défaut sur Amazon Linux 2. Pour utiliser l'application Kernel Live Patching, vous devez installer le plug-in yum pour Kernel Live Patching et activer la fonctionnalité Kernel Live Patching.

Prérequis

Kernel Live Patching nécessite `binutils`. Si vous n'avez pas `binutils` installé, installez-le à l'aide de la commande suivante :

```
$ sudo yum install binutils
```

Pour activer Kernel Live Patching

1. Les correctifs Kernel Live Patching sont disponibles pour Amazon Linux 2 avec la version du noyau `4.14.165-131.185` ou ultérieure. Pour vérifier la version de votre noyau, exécutez la commande suivante.

   ```
   $ sudo yum list kernel
   ```

2. Si vous avez déjà une version du noyau prise en charge, ignorez cette étape. Si vous ne disposez pas d'une version du noyau prise en charge, exécutez les commandes suivantes pour mettre à jour le noyau vers la dernière version et pour redémarrer l'instance.

   ```
   $ sudo yum install -y kernel
   ```

   ```
   $ sudo reboot
   ```

3. Installez le plugin yum pour Kernel Live Patching.

   ```
   $ sudo yum install -y yum-plugin-kernel-livepatch
   ```

4. Activez le plugin yum pour Kernel Live Patching.

   ```
   $ sudo yum kernel-livepatch enable -y
   ```

   Cette commande installe également la dernière version du RPM du correctif à chaud du noyau à partir des référentiels configurés.

5. Pour confirmer que le plugin yum pour Kernel Live Patching a bien été installé, exécutez la commande suivante.

   ```
   $ rpm -qa | grep kernel-livepatch
   ```

   Lorsque vous activez Kernel Live Patching, un RPM vide du correctif à chaud du noyau est automatiquement appliqué. Si Kernel Live Patching a été activé avec succès, cette commande renvoie une liste qui inclut le RPM vide initial du correctif à chaud du noyau.

6. Mettez à jour et démarrez le service kpatch. Ce service charge tous les correctifs à chaud du noyau lors de l'initialisation ou au démarrage.

   ```
   $ sudo yum update kpatch-runtime
   ```

   ```
   $ sudo systemctl enable kpatch.service
   ```

7. Configurez le référentiel Amazon Linux 2 Kernel Live Patching, qui contient les correctifs à chaud du noyau.

   ```
   $ sudo amazon-linux-extras enable livepatch
   ```

Affichage des correctifs à chaud du noyau disponibles

Les alertes de sécurité Amazon Linux sont publiées dans le Centre de sécurité Amazon Linux. Pour de plus amples informations sur les alertes de sécurité Amazon Linux 2, qui incluent les alertes pour les correctifs à chaud du noyau, veuillez consulter le Centre de sécurité Amazon Linux. Les correctifs Kernel Live sont préfixés avec `ALASLIVEPATCH`. Le Centre de sécurité Amazon Linux peut ne pas répertorier les correctifs à chaud du noyau qui corrigent les bogues.

Vous pouvez également découvrir les correctifs à chaud du noyau disponibles pour les avis et les CVE à l'aide de la ligne de commande.

Pour répertorier tous les correctifs à chaud du noyau disponibles pour les avis

Utilisez la commande suivante.

```
$ yum updateinfo list
```

Voici un exemple de sortie.

```
Loaded plugins: extras_suggestions, kernel-livepatch, langpacks, priorities, update-motd
 ALAS2LIVEPATCH-2020-002 important/Sec. kernel-
livepatch-4.14.165-133.209-1.0-3.amzn2.x86_64
 ALAS2LIVEPATCH-2020-005 medium/Sec. kernel-livepatch-4.14.165-133.209-1.0-4.amzn2.x86_64
 updateinfo list done
```

Pour répertorier tous les correctifs à chaud du noyau disponibles pour les CVE

Utilisez la commande suivante.

```
$ yum updateinfo list cves
```

Voici un exemple de sortie.

```
Loaded plugins: extras_suggestions, kernel-livepatch, langpacks, priorities, update-
motdamzn2-core/2/x86_64 | 2.4 kB 00:00:00
 CVE-2019-15918 important/Sec. kernel-livepatch-4.14.165-133.209-1.0-3.amzn2.x86_64
 CVE-2019-20096 important/Sec. kernel-livepatch-4.14.165-133.209-1.0-3.amzn2.x86_64
 CVE-2020-8648 medium/Sec. kernel-livepatch-4.14.165-133.209-1.0-4.amzn2.x86_64
 updateinfo list done
```

Application des correctifs à chaud du noyau

Vous appliquez les correctifs à chaud du noyau en utilisant le gestionnaire de paquets yum de la même manière que vous appliquez les mises à jour régulières. Le plugin yum pour noyau Live Patching gère les correctifs à chaud du noyau qui doivent être appliqués et élimine le besoin de redémarrer.

> **Tip**
>
> Nous vous recommandons de mettre à jour votre noyau régulièrement à l'aide de Kernel Live Patching pour vous assurer qu'il reste sécurisé et à jour.

Vous pouvez choisir d'appliquer un correctif à chaud du noyau spécifique ou d'appliquer tous les correctifs à chaud du noyau disponibles avec vos mises à jour de sécurité régulières.

Pour appliquer un correctif à chaud du noyau spécifique

1. Obtenez la version du correctif à chaud du noyau à l'aide de l'une des commandes décrites à la section Affichage des correctifs à chaud du noyau disponibles (p. 203).

2. Appliquez le correctif à chaud du noyau pour votre noyau Amazon Linux 2.

```
$ sudo yum install kernel-livepatch-kernel_version.x86_64
```

Par exemple, la commande suivante applique un correctif à chaud du noyau pour la version du noyau Amazon Linux 2 `4.14.165-133.209`.

```
$ sudo yum install kernel-livepatch-4.14.165-133.209-1.0-4.amzn2.x86_64
```

Pour appliquer les correctifs à chaud du noyau disponibles avec vos mises à jour de sécurité régulières

Utilisez la commande suivante.

```
$ sudo yum update --security
```

Omettre l'option `--security`d'inclure les corrections de bogues.

Important

- La version du noyau n'est pas mise à jour après l'application des correctifs à chaud du noyau. La version est mise à jour vers la nouvelle version seulement après le redémarrage de l'instance.
- Un noyau Amazon Linux 2 reçoit des correctifs à chaud du noyau pendant une période de trois mois. Une fois la période de trois mois écoulée, aucun nouveau correctif à chaud du noyau n'est publié pour cette version du noyau. Pour continuer à recevoir les correctifs à chaud du noyau après la période de trois mois, vous devez redémarrer l'instance pour passer à la nouvelle version du noyau, qui continuera ensuite à recevoir les correctifs à chaud du noyau pendant les trois prochains mois. Pour vérifier la fenêtre de support de votre version du noyau, exécutez `yum kernel-livepatch supported`.

Afficher les correctifs à chaud du noyau appliqués

Pour afficher les correctifs à chaud du noyau appliqués

Utilisez la commande suivante.

```
$ kpatch list
```

La commande renvoie une liste des correctifs à chaud du noyau des mise à jour de sécurité chargés et installés. Voici un exemple de sortie.

```
Loaded patch modules:
 livepatch_cifs_lease_buffer_len [enabled]
 livepatch_CVE_2019_20096 [enabled]
 livepatch_CVE_2020_8648 [enabled]

 Installed patch modules:
 livepatch_cifs_lease_buffer_len (4.14.165-133.209.amzn2.x86_64)
 livepatch_CVE_2019_20096 (4.14.165-133.209.amzn2.x86_64)
 livepatch_CVE_2020_8648 (4.14.165-133.209.amzn2.x86_64)
```

Note

Un seul correctif à chaud du noyau peut inclure et installer plusieurs correctifs à chaud.

Désactivation de Kernel Live Patching

Si vous n'avez plus besoin d'utiliser Kernel Live Patching, vous pouvez le désactiver à tout moment.

Pour désactiver Kernel Live Patching

1. Supprimez les packages RPM pour les correctifs à chaud du noyau appliqués.

```
$ sudo yum kernel-livepatch disable
```

2. Désinstallez le plugin yum pour Kernel Live Patching.

```
$ sudo yum remove yum-plugin-kernel-livepatch
```

3. Redémarrez l'instance.

```
$ sudo reboot
```

Limites

Kernel Live Patching présente les limitations suivantes :

- Lors de l'application d'un correctif à chaud du noyau, vous ne pouvez pas effectuer de mise en veille prolongée, utiliser des outils de débogage avancés (tels que SystemTap, kprobes et EBPF) ou accéder aux fichiers de sortie ftrace utilisés par l'infrastructure noyau Live Patching.
- Les instancesAmazon Linux 2 avec architecture ARM (arm64) 64 bits ne sont pas prises en charge.

Questions fréquentes (FAQ)

Pour les questions fréquemment posées sur Kernel Live Patching pour Amazon Linux 2, veuillez consulter la FAQ sur Amazon Linux 2 Kernel Live Patching.

Noyaux fournis par l'utilisateur

Si vous avez besoin d'un noyau personnalisé sur vos instances Amazon EC2, vous pouvez commencer avec une AMI qui est proche de ce que vous voulez, compiler le noyau personnalisé sur votre instance et mettre à jour le chargeur de démarrage pour pointer vers le nouveau noyau. Ce processus varie en fonction du type de virtualisation qu'utilise votre AMI. Pour de plus amples informations, veuillez consulter Types de virtualisation AMI Linux (p. 108).

Sommaire

- AMIs HVM (GRUB) (p. 205)
- AMIs paravirtuelles (PV-GRUB) (p. 206)

AMIs HVM (GRUB)

Les volumes d'instance HVM sont traités comme des disques physiques réels. Le processus de démarrage est similaire à celui d'un système d'exploitation bare metal avec disque partitionné et programme d'amorçage, ce qui lui permet de travailler avec toutes les distributions Linux actuellement prises en charge. Le chargeur de démarrage le plus courant est GRUB ou GRUB2.

Par défaut, GRUB n'envoie pas ses données de sortie à la console de l'instance car il crée un délai de démarrage supplémentaire. Pour plus d'informations, consultez Sortie de la console de l'instance (p. 1361). Si vous installez un noyau personnalisé, vous devez envisager d'activer la sortie GRUB.

Vous n'avez pas besoin de spécifier un noyau de rechange, mais nous vous recommandons d'en avoir un lorsque vous testez un nouveau noyau. GRUB peut avoir recours à un autre noyau au cas où le nouveau noyau échoue. Le fait d'avoir un noyau de rechange permet à l'instance de démarrer même si le nouveau noyau n'est pas trouvé.

L'ancien GRUB pour Amazon Linux utilise `/boot/grub/menu.lst`. GRUB2 pour Amazon Linux 2 utilise `/etc/default/grub`. Pour de plus amples informations sur la mise à jour du noyau par défaut dans le chargeur d'amorçage, veuillez consulter la documentation de votre distribution Linux.

AMIs paravirtuelles (PV-GRUB)

Les Amazon Machine Images qui utilisent la virtualisation paravirtuelle ont recourt à un système appelé PV-GRUB pendant le processus de démarrage. PV-GRUB est un programme d'amorçage paravirtuel qui exécute une version corrigée de GNU GRUB 0.97. Lorsque vous lancez une instance, PV-GRUB commence le processus de démarrage, puis charge en chaîne le noyau spécifié par le fichier `menu.lst` de votre image.

PV-GRUB comprend les commandes `grub.conf` ou `menu.lst` standard qui lui permettent de fonctionner avec toutes les distributions Linux actuellement prises en charge. Les distributions plus anciennes comme Ubuntu 10.04 LTS, Oracle Enterprise Linux ou CentOS 5.x ont besoin d'un package noyau spécial « ec2 » ou « xen » alors les distributions les plus récentes comprennent les pilotes nécessaires dans le package noyau par défaut.

La plupart des AMI de virtualisation paravirtuelle utilisent une PV-GRUB AKI par défaut (notamment l'ensemble des AMI Linux de virtualisation paravirtuelle dans le menu de démarrage rapide de l'assistant de lancement Amazon EC2), donc il n'existe aucune autre étape supplémentaire que vous devez suivre pour utiliser un noyau différent sur votre instance, dans la mesure où le noyau que vous voulez utiliser est compatible avec votre distribution. La meilleure façon d'exécuter un noyau personnalisé sur une instance est de commencer avec une AMI qui est proche de ce que vous voulez, puis de compiler le noyau personnalisé sur votre instance et de modifier le fichier `menu.lst` pour démarrer avec ce noyau.

Vous pouvez vérifier que l'image du noyau d'une AMI est un AKI PV-GRUB. Exécutez la commande describe-images suivante (en indiquant votre ID d'image de noyau) vérifiez si le champ `Name` commence par `pv-grub`:

```
aws ec2 describe-images --filters Name=image-id,Values=aki-880531cd
```

Sommaire

- Restrictions de PV-GRUB (p. 206)
- Configuration de GRUB pour les AMIs paravirtuelles (p. 207)
- ID de l'image noyau PV-GRUB Amazon (p. 208)
- Mise à jour de PV-GRUB (p. 210)

Restrictions de PV-GRUB

PV-GRUB possède les restrictions suivantes :

- Vous ne pouvez pas utiliser la version 64 bit de PV-GRUB pour lancer un noyau 32 bits ou vice versa.
- Vous ne pouvez pas spécifier une image ramdisk Amazon (ARI) lorsque vous utilisez une PV-GRUB AKI.

- AWS a testé et vérifié que PV-GRUB fonctionne avec ces formats de système de fichiers : EXT2, EXT3, EXT4, JFS, XFS et ReiserFS. Il se peut que d'autres formats de système de fichiers ne fonctionnent pas.
- PV-GRUB peut démarrer les noyaux compressés à l'aide de formats de compression gzip, bzip2, lzo et xz.
- Les AMI de cluster ne prennent pas en charge ou n'ont pas besoin de PV-GRUB, car ils utilisent la virtualisation matérielle complète (HVM). Tandis que les instances paravirtuelles utilisent PV-GRUB pour le démarrage, les volumes d'instances HVM sont traités comme de véritables disques et le processus de démarrage est similaire au processus de démarrage d'un système d'exploitation bare metal avec un disque divisé et un chargeur de démarrage.
- Les versions PV-GRUB 1.03 et antérieures ne prennent pas en charge le partitionnement GPT. Elles prennent uniquement en charge le partitionnement MBR.
- Si vous comptez utiliser un gestionnaire par volumes logiques (LVM) avec des volumes Amazon EBS, vous avez besoin d'une partition de démarrage séparée en dehors de LVM. Puis, vous pouvez créer des volumes logiques avec le LVM.

Configuration de GRUB pour les AMIs paravirtuelles

Pour démarrer PV-GRUB, un fichier `menu.lst` GRUB doit exister dans l'image. L'emplacement le plus commun de ce fichier est `/boot/grub/menu.lst`.

Ce qui suit est un exemple d'un fichier de configuration `menu.lst` pour le démarrage d'une AMI avec une AKI PV-GRUB. Dans cet exemple, un choix de deux entées noyau est proposé : Amazon Linux 2018.03 (le noyau original pour cette AMI) et Vanilla Linux 4.16.4 (une version plus récente du noyau Vanilla Linux de https://www.kernel.org/). L'entrée Vanilla a été copiée de l'entrée originale pour cette AMI et les chemins `kernel` et `initrd` ont été mis à jour par rapport aux nouveaux emplacements. Le paramètre `default 0` pointe le programme d'amorçage vers la première entrée qu'il voit (dans ce cas, l'entrée Vanilla) et le paramètre `fallback 1` pointe le programme d'amorçage vers la prochaine entrée s'il existe un problème lors du démarrage du premier.

```
default 0
fallback 1
timeout 0
hiddenmenu

title Vanilla Linux 4.16.4
root (hd0)
kernel /boot/vmlinuz-4.16.4 root=LABEL=/ console=hvc0
initrd /boot/initrd.img-4.16.4

title Amazon Linux 2018.03 (4.14.26-46.32.amzn1.x86_64)
root (hd0)
kernel /boot/vmlinuz-4.14.26-46.32.amzn1.x86_64 root=LABEL=/ console=hvc0
initrd /boot/initramfs-4.14.26-46.32.amzn1.x86_64.img
```

Vous ne devez pas spécifier un noyau de rechange dans votre fichier `menu.lst`, mais nous vous recommandons d'en avoir un lorsque vous tester un nouveau noyau. PV-GRUB peut avoir recours à un autre noyau au cas où le nouveau noyau échoue. Le fait d'avoir un noyau de rechange permet à l'instance de démarrer même si le nouveau noyau n'est pas trouvé.

PV-GRUB vérifie les emplacements suivants pour `menu.lst` en utilisant le premier qu'il trouve :

- `(hd0)/boot/grub`
- `(hd0,0)/boot/grub`
- `(hd0,0)/grub`
- `(hd0,1)/boot/grub`
- `(hd0,1)/grub`

- `(hd0,2)/boot/grub`
- `(hd0,2)/grub`
- `(hd0,3)/boot/grub`
- `(hd0,3)/grub`

Notez que les versions PV-GRUB 1.03 et antérieures ne vérifient que l'un des deux premiers emplacements de cette liste.

ID de l'image noyau PV-GRUB Amazon

Les AKI PV-GRUB sont disponibles dans toutes les régions Amazon EC2. Il existe des AKI pour les types d'architecture 32 bits et 64 bits. La plupart des AMI modernes utilisent une AKI PV-GRUB par défaut.

Nous vous recommandons de toujours utiliser la dernière version de l'AKI PV-GRUB, car les versions de l'AKI PV-GRUB ne sont pas toutes compatibles avec les types d'instance. Utilisez la commande describe-images suivante pour obtenir une liste d'AKI PV-GRUB pour la région actuelle :

```
aws ec2 describe-images --owners amazon --filters Name=name,Values=pv-grub-*.gz
```

PV-GRUB est la seule AKI disponible dans la région `ap-southeast-2`. Vous devriez vérifier que toutes les AMI que vous voulez copier vers cette région utilisent une version de PV-GRUB qui est disponible dans cette région.

Ce qui suit correspond aux ID d'AKI actuels pour chaque région. Enregistrez les nouvelles AMI à l'aide d'une AKI hd0.

> **Note**
>
> Nous continuons de fournir des AKI hd00 pour la rétrocompatibilité dans les régions où elles étaient précédemment disponibles.

ap-northeast-1, Asie-Pacifique (Tokyo)

ID de l'image	Nom de l'image
`aki-f975a998`	`pv-grub-hd0_1.05-i386.gz`
`aki-7077ab11`	`pv-grub-hd0_1.05-x86_64.gz`

ap-southeast-1, Région Asie-Pacifique (Singapour)

ID de l'image	Nom de l'image
`aki-17a40074`	`pv-grub-hd0_1.05-i386.gz`
`aki-73a50110`	`pv-grub-hd0_1.05-x86_64.gz`

ap-southeast-2, Asie-Pacifique (Sydney)

ID de l'image	Nom de l'image
`aki-ba5665d9`	`pv-grub-hd0_1.05-i386.gz`
`aki-66506305`	`pv-grub-hd0_1.05-x86_64.gz`

eu-central-1, Europe (Francfort)

ID de l'image	Nom de l'image
aki-1419e57b	pv-grub-hd0_1.05-i386.gz
aki-931fe3fc	pv-grub-hd0_1.05-x86_64.gz

eu-west-1, Europe (Irlande)

ID de l'image	Nom de l'image
aki-1c9fd86f	pv-grub-hd0_1.05-i386.gz
aki-dc9ed9af	pv-grub-hd0_1.05-x86_64.gz

sa-east-1, Amérique du Sud (São Paulo)

ID de l'image	Nom de l'image
aki-7cd34110	pv-grub-hd0_1.05-i386.gz
aki-912fbcfd	pv-grub-hd0_1.05-x86_64.gz

us-east-1, USA Est (Virginie du Nord)

ID de l'image	Nom de l'image
aki-04206613	pv-grub-hd0_1.05-i386.gz
aki-5c21674b	pv-grub-hd0_1.05-x86_64.gz

us-gov-west-1, AWS GovCloud (US-West)

ID de l'image	Nom de l'image
aki-5ee9573f	pv-grub-hd0_1.05-i386.gz
aki-9ee55bff	pv-grub-hd0_1.05-x86_64.gz

us-west-1, USA Ouest (Californie du Nord)

ID de l'image	Nom de l'image
aki-43cf8123	pv-grub-hd0_1.05-i386.gz
aki-59cc8239	pv-grub-hd0_1.05-x86_64.gz

us-west-2, USA Ouest (Oregon)

ID de l'image	Nom de l'image
aki-7a69931a	pv-grub-hd0_1.05-i386.gz

ID de l'image	Nom de l'image
`aki-70cb0e10`	`pv-grub-hd0_1.05-x86_64.gz`

Mise à jour de PV-GRUB

Nous vous recommandons de toujours utiliser la dernière version de l'AKI PV-GRUB, car les versions de l'AKI PV-GRUB ne sont pas toutes compatibles avec les types d'instance. De plus, les versions les plus anciennes de PV-GRUB ne sont pas disponibles dans toutes les régions. Si vous copiez une AMI qui utilise une version plus ancienne pour une région que ne prend pas en charge cette version, vous ne pourrez donc pas démarrer des instances lancées à partir d'une AMI jusqu'à ce que vous mettiez à jour l'image noyau. Utilisez les procédures suivantes pour vérifier la version de PV-GRUB de votre instance et la mettre à jour si nécessaire.

Pour vérifier votre version de PV-GRUB

1. Trouvez l'ID noyau pour votre instance.

```
aws ec2 describe-instance-attribute --instance-id instance_id --attribute kernel --
region region

{
    "InstanceId": "instance_id",
    "KernelId": "aki-70cb0e10"
}
```

   L'ID noyau pour cette instance est `aki-70cb0e10`.

2. Consultez les informations sur la version de cet ID noyau.

```
aws ec2 describe-images --image-ids aki-70cb0e10 --region region

{
    "Images": [
        {
            "VirtualizationType": "paravirtual",
            "Name": "pv-grub-hd0_1.05-x86_64.gz",
            ...
            "Description": "PV-GRUB release 1.05, 64-bit"
        }
    ]
}
```

   Cette image noyau est PV-GRUB 1.05. Si votre version PV-GRUB n'est pas la plus récente (comme indiqué dans le didacticiel ID de l'image noyau PV-GRUB Amazon (p. 208)), vous devriez la mettre à jour en suivant la procédure ci-dessous.

Pour mettre à jour votre version de PV-GRUB

Si votre instance utilise une version de PV-GRUB plus ancienne, vous devriez la mettre à jour.

1. Identifiez le dernier PV-GRUB AKI pour votre région et l'architecture de processeur à partir de ID de l'image noyau PV-GRUB Amazon (p. 208).
2. Arrêtez votre instance. Votre instance doit être arrêtée pour modifier l'image noyau utilisée.

```
aws ec2 stop-instances --instance-ids instance_id --region region
```

3. Modifiez l'image noyau utilisée pour votre instance.

```
aws ec2 modify-instance-attribute --instance-id instance_id --kernel kernel_id --
region region
```

4. Redémarrez votre instance.

```
aws ec2 start-instances --instance-ids instance_id --region region
```

Utilisation de l'environnement de bureau MATE fourni avec Amazon Linux 2

L'environnement de bureau MATE est préinstallé et préconfiguré dans l'AMI avec la description suivante : Amazon Linux 2 avec .NET Core, PowerShell, Mono et l'environnement de bureau MATE. L'environnement fournit une interface utilisateur graphique intuitive pour administrer les instances Amazon Linux 2 sans utiliser la ligne de commande. L'interface utilise des représentations graphiques, telles que des icônes, des fenêtres, des barres d'outils, des dossiers, des fonds d'écran et des widgets de bureau. Des outils intégrés basés sur l'interface graphique sont disponibles pour effectuer des tâches courantes. Par exemple, il existe des outils pour ajouter et supprimer des logiciels, appliquer des mises à jour, organiser des fichiers, lancer des programmes et surveiller l'intégrité du système.

Suivez la procédure ci-dessous pour utiliser l'environnement de bureau MATE.

Configurer les connexions RDP (Remote Desktop Protocol) et définir un mot de passe

1. Utilisez la commande suivante describe-images pour obtenir l'ID de l'AMI pour Amazon Linux 2 qui inclut MATE dans le nom de l'AMI.

```
aws ec2 describe-images --filters "Name=name,Values=amzn2*MATE*" --query
 'Images[*].ImageId' --output text
```

   Voici un exemple de sortie :

```
ami-0abcdef1234567890
```

2. Lancez une instance EC2 avec l'AMI que vous avez localisée à l'étape précédente. Configurez le groupe de sécurité pour autoriser le trafic TCP entrant vers le port 3389. Pour de plus amples informations sur les groupes de sécurité, veuillez consulter Groupes de sécurité pour votre VPC. Cette configuration vous permet d'utiliser un client RDP pour vous connecter à l'instance.
3. Connectez-vous à l'instance à l'aide de SSH. Exécutez la commande suivante pour définir le mot de passe pour `ec2-user`.

```
[ec2-user ~]$ sudo passwd ec2-user
```

4. Ouvrez un client RDP sur l'ordinateur à partir duquel vous vous connecterez à l'instance (par exemple, Connexion Bureau à distance sur un ordinateur sous Microsoft Windows). Saisissez `ec2-user` comme nom d'utilisateur et entrez le mot de passe que vous avez défini à l'étape précédente.

Pour désactiver l'environnement de bureau MATE sur votre instance EC2

Vous pouvez désactiver l'environnement GUI à tout moment en exécutant l'une des commandes suivantes.

```
[ec2-user ~]$ sudo systemctl disable xrdp
```

```
[ec2-user ~]$ sudo systemctl stop xrdp
```

Pour activer l'environnement de bureau MATE sur votre instance EC2

Pour réactiver l'interface graphique, vous pouvez exécuter l'une des commandes suivantes.

```
[ec2-user ~]$ sudo systemctl enable xrdp
```

```
[ec2-user ~]$ sudo systemctl start xrdp
```

Instances Amazon EC2

Si vous découvrez Amazon EC2, consultez les rubriques suivantes pour démarrer :

- Présentation d'Amazon EC2 (p. 1)
- Configuration avec Amazon EC2 (p. 28)
- Didacticiel : Premiers pas avec les instances Amazon EC2 Linux (p. 32)
- Cycle de vie d'une instance (p. 524)

Avant que vous ne lanciez un environnement de production, vous devez répondre aux questions suivantes.

Q. Quel type d'instance répond le mieux à mes besoins ?

Amazon EC2 fournit différents types d'instance pour vous permettre de choisir les capacités d'UC, de mémoire, de stockage et de mise en réseau dont vous avez besoin pour exécuter vos applications. Pour plus d'informations, consultez Types d'instance (p. 213).

Q. Quelle option d'achat répond le mieux à mes besoins ?

Amazon EC2 prend en charge les Instances à la demande (instances par défaut), les Instances Spot et les Instances réservées. Pour plus d'informations, consultez Options d'achat d'instance (p. 323).

Q. Quel type de volume racine répond le mieux à mes besoins ?

Chaque instance est basée sur Amazon EBS ou sur le stockage d'instance. Sélectionnez une AMI en fonction du type de volume racine dont vous avez besoin. Pour plus d'informations, consultez Stockage pour le périphérique racine (p. 106).

Q. Est-ce que je peux gérer à distance un parc d'instances EC2 et des machines virtuelles dans mon environnement hybride ?

AWS Systems Manager vous permet de gérer, à distance et en toute sécurité la configuration de vos instances Amazon EC2, et de vos instances sur site et machines virtuelles (VM) dans les environnements hybrides, y compris des machines virtuelles d'autres fournisseurs de cloud. Pour plus d'informations, consultez le Guide de l'utilisateur AWS Systems Manager.

Types d'instance

Lorsque vous lancez une instance, le type d'instance que vous spécifiez détermine les capacités matérielles de l'ordinateur hôte utilisé pour votre instance. Chaque type d'instance propose différentes capacités de calcul, de mémoire et de stockage ; les types d'instance sont regroupés en familles d'instance en fonction de ces capacités. Sélectionnez un type d'instance en fonction des exigences de l'application ou du logiciel que vous prévoyez d'exécuter sur votre instance.

Amazon EC2 fournit à chaque instance une quantité cohérente et prévisible de capacité d'UC, indépendamment de son matériel sous-jacent.

Amazon EC2 dédie certaines ressources de l'ordinateur hôte (comme l'UC, la mémoire et le stockage d'instance) soient dédiées à une instance en particulier. Amazon EC2 partage d'autres ressources de l'ordinateur hôte (comme le réseau et le sous-système de disque) sont partagées entre les instances. Si chaque instance d'un ordinateur hôte essaie d'utiliser autant que possible de l'une de ces ressources partagées, chacun reçoit une part égale de cette ressource. Cependant, quand une ressource est sous-utilisée, une instance peut consommer une part plus important de cette ressource, tant qu'elle est disponible.

Chaque type d'instance offre des performances minimales plus ou moins élevées à partir d'une ressource partagée. Par exemple, les types d'instance avec des performances d'E/S élevées bénéficient d'une plus

grande allocation de ressources partagées. L'allocation d'une plus grande part de ressources partagées réduit aussi les écarts de performances d'E/S. Pour la plupart des applications, des performances d'E/S modérées sont plus que suffisantes. Cependant, pour les applications qui requièrent des performances d'E/S plus élevées ou plus régulières, envisagez un type d'instance avec des performances d'E/S supérieures.

Table des matières

- Types d'instance disponibles (p. 214)
- Spécifications matérielles (p. 217)
- Types de virtualisation AMI (p. 218)
- Instances reposant sur le système Nitro (p. 218)
- Fonctions de mise en réseau et de stockage (p. 219)
- Limites d'instance (p. 222)
- Instances à usage général (p. 222)
- Instances de calcul optimisé (p. 270)
- Instances de mémoire optimisée (p. 278)
- Instances de stockage optimisé (p. 289)
- Instances à calcul accéléré Linux (p. 296)
- Recherche d'un type d'instance Amazon EC2 (p. 312)
- Modification du type d'instance (p. 313)
- Obtention de recommandations pour un type d'instance (p. 319)

Types d'instance disponibles

Amazon EC2 fournit un large choix de types d'instance optimisés pour différents cas d'utilisation. Pour déterminer quels types d'instance répondent à vos besoins, tels que les régions prises en charge, les ressources de calcul ou les ressources de stockage, veuillez consulter Recherche d'un type d'instance Amazon EC2 (p. 312).

Instances de la génération actuelle

Pour obtenir les meilleures performances, nous vous recommandons d'utiliser les types d'instance suivants quand vous lancez de nouvelles instances. Pour de plus amples informations, veuillez consulter Types d'instance Amazon EC2.

Type	Tailles	Cas d'utilisation
A1	`a1.medium` | `a1.large` | `a1.xlarge` | `a1.2xlarge` | `a1.4xlarge` | `a1.metal`	Usage général (p. 222)
C4	`c4.large` | `c4.xlarge` | `c4.2xlarge` | `c4.4xlarge` | `c4.8xlarge`	Calcul optimisé (p. 270)
C5	`c5.large` | `c5.xlarge` | `c5.2xlarge` | `c5.4xlarge` | `c5.9xlarge` | `c5.12xlarge` | `c5.18xlarge` | `c5.24xlarge` | `c5.metal`	Calcul optimisé (p. 270)
C5a	`c5a.large` | `c5a.xlarge` | `c5a.2xlarge` | `c5a.4xlarge` | `c5a.8xlarge` | `c5a.12xlarge` | `c5a.16xlarge` | `c5a.24xlarge`	Calcul optimisé (p. 270)
C5ad	`c5ad.large` | `c5ad.xlarge` | `c5ad.2xlarge` | `c5ad.4xlarge` | `c5ad.8xlarge` | `c5ad.12xlarge` | `c5ad.16xlarge` | `c5ad.24xlarge`	Calcul optimisé (p. 270)

Type	Tailles	Cas d'utilisation
C5d	`c5d.large` | `c5d.xlarge` | `c5d.2xlarge` | `c5d.4xlarge` | `c5d.9xlarge` | `c5d.12xlarge` | `c5d.18xlarge` | `c5d.24xlarge` | `c5d.metal`	Calcul optimisé (p. 270)
C5n	`c5n.large` | `c5n.xlarge` | `c5n.2xlarge` | `c5n.4xlarge` | `c5n.9xlarge` | `c5n.18xlarge` | `c5n.metal`	Calcul optimisé (p. 270)
C6g	`c6g.medium` | `c6g.large` | `c6g.xlarge` | `c6g.2xlarge` | `c6g.4xlarge` | `c6g.8xlarge` | `c6g.12xlarge` | `c6g.16xlarge` | `c6g.metal`	Calcul optimisé (p. 270)
C6gd	`c6gd.medium` | `c6gd.large` | `c6gd.xlarge` | `c6gd.2xlarge` | `c6gd.4xlarge` | `c6gd.8xlarge` | `c6gd.12xlarge` | `c6gd.16xlarge` | `c6gd.metal`	Calcul optimisé (p. 270)
D2	`d2.xlarge` | `d2.2xlarge` | `d2.4xlarge` | `d2.8xlarge`	Stockage optimisé (p. 289)
F1	`f1.2xlarge` | `f1.4xlarge` | `f1.16xlarge`	Calcul accéléré (p. 296)
G3	`g3s.xlarge` | `g3.4xlarge` | `g3.8xlarge` | `g3.16xlarge`	Calcul accéléré (p. 296)
G4	`g4dn.xlarge` | `g4dn.2xlarge` | `g4dn.4xlarge` | `g4dn.8xlarge` | `g4dn.12xlarge` | `g4dn.16xlarge` | `g4dn.metal`	Calcul accéléré (p. 296)
H1	`h1.2xlarge` | `h1.4xlarge` | `h1.8xlarge` | `h1.16xlarge`	Stockage optimisé (p. 289)
I3	`i3.large` | `i3.xlarge` | `i3.2xlarge` | `i3.4xlarge` | `i3.8xlarge` | `i3.16xlarge` | `i3.metal`	Stockage optimisé (p. 289)
I3en	`i3en.large` | `i3en.xlarge` | `i3en.2xlarge` | `i3en.3xlarge` | `i3en.6xlarge` | `i3en.12xlarge` | `i3en.24xlarge` | `i3en.metal`	Stockage optimisé (p. 289)
Inf1	`inf1.xlarge` | `inf1.2xlarge` | `inf1.6xlarge` | `inf1.24xlarge`	Calcul accéléré (p. 296)
M4	`m4.large` | `m4.xlarge` | `m4.2xlarge` | `m4.4xlarge` | `m4.10xlarge` | `m4.16xlarge`	Usage général (p. 222)
M5	`m5.large` | `m5.xlarge` | `m5.2xlarge` | `m5.4xlarge` | `m5.8xlarge` | `m5.12xlarge` | `m5.16xlarge` | `m5.24xlarge` | `m5.metal`	Usage général (p. 222)
M5a	`m5a.large` | `m5a.xlarge` | `m5a.2xlarge` | `m5a.4xlarge` | `m5a.8xlarge` | `m5a.12xlarge` | `m5a.16xlarge` | `m5a.24xlarge`	Usage général (p. 222)
M5ad	`m5ad.large` | `m5ad.xlarge` | `m5ad.2xlarge` | `m5ad.4xlarge` | `m5ad.8xlarge` | `m5ad.12xlarge` | `m5ad.16xlarge` | `m5ad.24xlarge`	Usage général (p. 222)
M5d	`m5d.large` | `m5d.xlarge` | `m5d.2xlarge` | `m5d.4xlarge` | `m5d.8xlarge` | `m5d.12xlarge` | `m5d.16xlarge` | `m5d.24xlarge` | `m5d.metal`	Usage général (p. 222)
M5dn	`m5dn.large` | `m5dn.xlarge` | `m5dn.2xlarge` | `m5dn.4xlarge` | `m5dn.8xlarge` | `m5dn.12xlarge` | `m5dn.16xlarge` | `m5dn.24xlarge`	Usage général (p. 222)
M5n	`m5n.large` | `m5n.xlarge` | `m5n.2xlarge` | `m5n.4xlarge` | `m5n.8xlarge` | `m5n.12xlarge` | `m5n.16xlarge` | `m5n.24xlarge`	Usage général (p. 222)

Type	Tailles	Cas d'utilisation
M6g	`m6g.medium` | `m6g.large` | `m6g.xlarge` | `m6g.2xlarge` | `m6g.4xlarge` | `m6g.8xlarge` | `m6g.12xlarge` | `m6g.16xlarge` | `m6g.metal`	Usage général (p. 222)
M6gd	`m6gd.medium` | `m6gd.large` | `m6gd.xlarge` | `m6gd.2xlarge` | `m6gd.4xlarge` | `m6gd.8xlarge` | `m6gd.12xlarge` | `m6gd.16xlarge` | `m6gd.metal`	Usage général (p. 222)
P2	`p2.xlarge` | `p2.8xlarge` | `p2.16xlarge`	Calcul accéléré (p. 296)
P3	`p3.2xlarge` | `p3.8xlarge` | `p3.16xlarge`	Calcul accéléré (p. 296)
P3dn	`p3dn.24xlarge`	Calcul accéléré (p. 296)
R4	`r4.large` | `r4.xlarge` | `r4.2xlarge` | `r4.4xlarge` | `r4.8xlarge` | `r4.16xlarge`	Mémoire optimisée (p. 278)
R5	`r5.large` | `r5.xlarge` | `r5.2xlarge` | `r5.4xlarge` | `r5.8xlarge` | `r5.12xlarge` | `r5.16xlarge` | `r5.24xlarge` | `r5.metal`	Mémoire optimisée (p. 278)
R5a	`r5a.large` | `r5a.xlarge` | `r5a.2xlarge` | `r5a.4xlarge` | `r5a.8xlarge` | `r5a.12xlarge` | `r5a.16xlarge` | `r5a.24xlarge`	Mémoire optimisée (p. 278)
R5ad	`r5ad.large` | `r5ad.xlarge` | `r5ad.2xlarge` | `r5ad.4xlarge` | `r5ad.8xlarge` | `r5ad.12xlarge` | `r5ad.16xlarge` | `r5ad.24xlarge`	Mémoire optimisée (p. 278)
R5d	`r5d.large` | `r5d.xlarge` | `r5d.2xlarge` | `r5d.4xlarge` | `r5d.8xlarge` | `r5d.12xlarge` | `r5d.16xlarge` | `r5d.24xlarge` | `r5d.metal`	Mémoire optimisée (p. 278)
R5dn	`r5dn.large` | `r5dn.xlarge` | `r5dn.2xlarge` | `r5dn.4xlarge` | `r5dn.8xlarge` | `r5dn.12xlarge` | `r5dn.16xlarge` | `r5dn.24xlarge`	Mémoire optimisée (p. 278)
R5n	`r5n.large` | `r5n.xlarge` | `r5n.2xlarge` | `r5n.4xlarge` | `r5n.8xlarge` | `r5n.12xlarge` | `r5n.16xlarge` | `r5n.24xlarge`	Mémoire optimisée (p. 278)
R6g	`r6g.medium` | `r6g.large` | `r6g.xlarge` | `r6g.2xlarge` | `r6g.4xlarge` | `r6g.8xlarge` | `r6g.12xlarge` | `r6g.16xlarge` | `r6g.metal`	Mémoire optimisée (p. 278)
R6gd	`r6gd.medium` | `r6gd.large` | `r6gd.xlarge` | `r6gd.2xlarge` | `r6gd.4xlarge` | `r6gd.8xlarge` | `r6gd.12xlarge` | `r6gd.16xlarge` | `r6gd.metal`	Mémoire optimisée (p. 278)
T2	`t2.nano` | `t2.micro` | `t2.small` | `t2.medium` | `t2.large` | `t2.xlarge` | `t2.2xlarge`	Usage général (p. 222)
T3	`t3.nano` | `t3.micro` | `t3.small` | `t3.medium` | `t3.large` | `t3.xlarge` | `t3.2xlarge`	Usage général (p. 222)
T3a	`t3a.nano` | `t3a.micro` | `t3a.small` | `t3a.medium` | `t3a.large` | `t3a.xlarge` | `t3a.2xlarge`	Usage général (p. 222)
u-xtb1	`u-6tb1.metal` | `u-9tb1.metal` | `u-12tb1.metal` | `u-18tb1.metal` | `u-24tb1.metal`	Mémoire optimisée (p. 278)

Type	Tailles	Cas d'utilisation
X1	`x1.16xlarge` | `x1.32xlarge`	Mémoire optimisée (p. 278)
X1e	`x1e.xlarge` | `x1e.2xlarge` | `x1e.4xlarge` | `x1e.8xlarge` | `x1e.16xlarge` | `x1e.32xlarge`	Mémoire optimisée (p. 278)
z1d	`z1d.large` | `z1d.xlarge` | `z1d.2xlarge` | `z1d.3xlarge` | `z1d.6xlarge` | `z1d.12xlarge` | `z1d.metal`	Mémoire optimisée (p. 278)

Instances de la génération précédente

Amazon Web Services propose les types d'instance de la génération précédente aux utilisateurs qui ont optimisé leurs applications autour de ces instances, mais doivent encore les mettre à niveau. Nous vous encourageons à utiliser les types d'instance de la génération actuelle pour obtenir les meilleures performances, mais nous continuons à prendre en charge les types d'instance de la génération précédente suivants. Pour de plus amples informations sur le type d'instance de la génération actuelle qui constituerait une mise à niveau appropriée, veuillez consulter Instances de la génération précédente.

Type	Tailles
C1	`c1.medium` | `c1.xlarge`
C3	`c3.large` | `c3.xlarge` | `c3.2xlarge` | `c3.4xlarge` | `c3.8xlarge`
G2	`g2.2xlarge` | `g2.8xlarge`
I2	`i2.xlarge` | `i2.2xlarge` | `i2.4xlarge` | `i2.8xlarge`
M1	`m1.small` | `m1.medium` | `m1.large` | `m1.xlarge`
M2	`m2.xlarge` | `m2.2xlarge` | `m2.4xlarge`
M3	`m3.medium` | `m3.large` | `m3.xlarge` | `m3.2xlarge`
R3	`r3.large` | `r3.xlarge` | `r3.2xlarge` | `r3.4xlarge` | `r3.8xlarge`
T1	`t1.micro`

Spécifications matérielles

Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.

Pour que vous puissiez déterminer le type d'instance qui correspond le mieux à vos besoins, nous vous recommandons de lancer une instance et d'utiliser votre propre application de comparaison. Comme vous payez l'instance à la seconde, il est pratique et économique de tester plusieurs types d'instance avant de prendre une décision.

Si vos besoins évoluent, même après avoir pris une décision, vous pouvez par la suite redimensionner votre instance. Pour plus d'informations, consultez Modification du type d'instance (p. 313).

Note

Les instances Amazon EC2 s'exécutent généralement sur des processeurs Intel virtuels 64 bits, comme indiqué dans les pages produits de ces types d'instance. Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2. Toutefois, les conventions d'appellation du secteur concernant les processeurs 64 bits peuvent donner lieu à une certaine confusion. Le fabricant de puces Advanced Micro Devices (AMD) a présenté la première architecture 64 bits commercialement viable basée sur le jeu d'instructions Intel x86. Par conséquent, ce type d'architecture est souvent appelé AMD64, quel que soit le fabricant. C'est notamment le cas pour Windows et plusieurs distributions Linux. C'est ce qui explique que les informations système internes d'une instance EC2 Ubuntu ou Windows indiquent une architecture de processeur AMD64, même quand les instances s'exécutent sur du matériel Intel.

Types de virtualisation AMI

Le type de virtualisation de votre instance est déterminé par l'AMI que vous utilisez pour la lancer. Les types d'instance de la génération actuelle prennent uniquement en charge la virtualisation HVM. Certains types d'instance de la génération précédente prennent en charge la virtualisation paravirtuelle et certaines régions AWS prennent en charge les instances de virtualisation paravirtuelle. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108).

Pour de meilleures performances, nous vous recommandons d'utiliser une AMI HVM. De plus, les AMI HVM sont obligatoires pour tirer parti de la mise en réseau améliorée. La virtualisation HVM utilise une technologie à assistance matérielle fournie par la plateforme AWS. Avec la virtualisation HVM, la machine virtuelle invitée s'exécute comme si elle se trouvait sur une plateforme matérielle native, si ce n'est qu'elle continue d'utiliser les pilotes du stockage et le réseau de la virtualisation PV pour des performances améliorées.

Instances reposant sur le système Nitro

Le système Nitro est un ensemble de composants matériels et logiciels élaborés par AWS qui garantissent des performances élevées, une haute disponibilité et un niveau de sécurité élevé. En outre, le système Nitro propose des fonctionnalités de type matériel nu qui éliminent les frais associés à la virtualisation et prennent en charge les charges de travail qui nécessitent un accès complet au matériel hôte. Pour de plus amples informations, veuillez consulter Système Nitro AWS.

Composants Nitro

Les composants suivants font partie du système Nitro :

- Carte Nitro
  - Volumes de stockage NVMe locaux
  - Prise en charge du matériel de mise en réseau
  - Gestion
  - Surveillance
  - Sécurité
- Puce de sécurité Nitro, intégrée à la carte mère
- Hyperviseur Nitro : un hyperviseur léger qui gère l'allocation d'UC et de mémoire et offre des performances similaires au matériel nu pour la plupart des charges de travail.

Types d'instance

Les instances suivantes reposent sur le système Nitro :

- Virtualisé : A1, C5, C5a, C5ad, C5d, C5n, C6g, C6gd, G4, I3en, Inf1, M5, M5a, M5ad, M5d, M5dn, M5n, M6g, M6gd, `p3dn.24xlarge`, R5, R5a, R5ad, R5d, R5dn, R5n, R6g, R6gd, T3, T3a, et z1d
- Bare metal : `a1.metal`, `c5.metal`, `c5d.metal`, `c5n.metal`, `c6g.metal`, `c6gd.metal`, `i3.metal`, `i3en.metal`, `m5.metal`, `m5d.metal`, `m6g.metal`, `m6gd.metal`, `r5.metal`, `r5d.metal`, `r6g.metal`, `r6gd.metal`, `u-6tb1.metal`, `u-9tb1.metal`, `u-12tb1.metal`, `u-18tb1.metal`, `u-24tb1.metal`, and `z1d.metal`

En savoir plus

Pour plus d'informations, consultez les vidéos suivantes :

- AWS re:Invent 2017: The Amazon EC2 Nitro System Architecture
- AWS re:Invent 2017: Amazon EC2 Bare Metal Instances
- AWS re:Invent 2019: Powering next-gen Amazon EC2: Deep dive into the Nitro system
- AWS re:Inforce 2019: Security Benefits of the Nitro Architecture

Fonctions de mise en réseau et de stockage

Lorsque vous sélectionnez un type d'instance, celui-ci détermine les fonctionnalités de mise en réseau et de stockage disponibles. Pour décrire un type d'instance, utilisez la commande describe-instance-types.

Fonctionnalités de mise en réseau

- IPv6 est pris en charge sur tous les types d'instance de la génération actuelle, ainsi que sur les types d'instance de la génération précédente C3, R3 et I2.
- Afin d'optimiser la mise en réseau et les performances de bande passante de votre type d'instance, vous pouvez effectuer les opérations suivantes :
  - Lancez les types d'instance pris en charge dans un groupe de placement du cluster afin d'optimiser vos instances pour les applications de Calcul Haute Performance (HPC). Les instances contenues dans un groupe de placement de cluster commun peuvent profiter de la bande passante élevée et de la mise en réseau à faible latence. Pour plus d'informations, consultez Groupes de placement (p. 924).
  - Activez la mise en réseau améliorée pour les types d'instance de génération actuelle afin d'obtenir des performances de paquet par seconde (PPS) nettement plus élevées, une meilleure stabilité du réseau et une latence moindre. Pour plus d'informations, consultez Mise en réseau améliorée sur Linux (p. 864).
- Les types d'instance de la génération actuelle sur lesquels est activée la mise en réseau améliorée disposent des attributs de performance de mise en réseau suivants :
  - Le trafic au sein d'une même région sur une adresse IPv4 ou IPv6 privée peut prendre en charge 5 Gbit/s pour le trafic à flux unique et jusqu'à 25 Gbit/s pour le trafic à plusieurs flux (selon le type d'instance).
  - Le trafic vers et depuis des compartiments Amazon S3 de la même région via l'espace d'adressage IP public ou un point de terminaison d'un VPC peut utiliser la totalité de la bande passante cumulée disponible pour l'instance.
- L'unité de transmission maximale (MTU) prise en charge varie selon les types d'instance. Tous les types d'instance Amazon EC2 prennent en charge les délais MTU Ethernet V2 1500 standard. Toutes les instances de la génération actuelle prennent en charge 9001 MTU, ou jumbo frames, au même titre que certaines instances de la génération précédente. Pour plus d'informations, consultez Unité de transmission maximale (MTU) du réseau pour votre instance EC2 (p. 937).

Fonctionnalités de stockage

- Certains types d'instance prennent en charge les volumes EBS et les volumes de stockage d'instance, tandis que d'autres ne prennent en charge que les volumes EBS. Certains types d'instance qui prennent en charge les volumes de stockage d'instance utilisent les disques SSD (Solid State Drive) pour fournir des performances d'I/O aléatoires très élevées. Certains types d'instance prennent en charge les volumes de stockage d'instance NVMe. Certains types d'instance prennent en charge les volumes EBS NVMe. Pour plus d'informations, consultez Amazon EBS et NVMe sur des instances Linux (p. 1209) et Volumes SSD NVMe (p. 1276).
- Afin d'obtenir une capacité supplémentaire dédiée pour les I/O Amazon EBS, vous pouvez lancer certains types d'instance comme instances optimisées pour EBS. Certains types d'instance sont optimisés pour EBS par défaut. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Résumé des fonctions de réseautage et de stockage

Le tableau suivant récapitule les fonctions de mise en réseau et de stockage prises en charge par les types d'instance de la génération actuelle.

	EBS uniquement	EBS NVMe	Stockage d'instance	Groupe de placement	Mise en réseau améliorée
A1	Oui	Oui	Non	Oui	ENA
C4	Oui	Non	Non	Oui	Intel 82599 VF
C5	Oui	Oui	Non	Oui	ENA
C5a	Oui	Oui	Non	Oui	ENA
C5ad	Non	Oui	NVMe *	Oui	ENA
C5d	Non	Oui	NVMe *	Oui	ENA
C5n	Oui	Oui	Non	Oui	ENA
C6g	Oui	Oui	Non	Oui	ENA
C6gd	Non	Oui	NVME *	Oui	ENA
D2	Non	Non	HDD	Oui	Intel 82599 VF
F1	Non	Non	NVMe *	Oui	ENA
G3	Oui	Non	Non	Oui	ENA
G4	Non	Oui	NVMe *	Oui	ENA
HS1	Non	Non	HDD *	Oui	ENA
I3	Non	Non	NVMe *	Oui	ENA
I3en	Non	Oui	NVMe *	Oui	ENA
Inf1	Oui	Oui	Non	Oui	ENA
M4	Oui	Non	Non	Oui	`m4.16xlarge`: ENA

	EBS uniquement	EBS NVMe	Stockage d'instance	Groupe de placement	Mise en réseau améliorée
					Toutes les autres tailles : Intel 82599 VF
M5	Oui	Oui	Non	Oui	ENA
M5a	Oui	Oui	Non	Oui	ENA
M5ad	Non	Oui	NVMe *	Oui	ENA
M5d	Non	Oui	NVMe *	Oui	ENA
M5dn	Non	Oui	NVMe *	Oui	ENA
M5n	Oui	Oui	Non	Oui	ENA
M6g	Oui	Oui	Non	Oui	ENA
M6gd	Non	Oui	NVME *	Oui	ENA
P2	Oui	Non	Non	Oui	ENA
P3	Oui	Non	Non	Oui	ENA
P3dn	Non	Oui	NVMe *	Oui	ENA
R4	Oui	Non	Non	Oui	ENA
R5	Oui	Oui	Non	Oui	ENA
R5a	Oui	Oui	Non	Oui	ENA
R5ad	Non	Oui	NVMe *	Oui	ENA
R5d	Non	Oui	NVMe *	Oui	ENA
R5dn	Non	Oui	NVMe *	Oui	ENA
R5n	Oui	Oui	Non	Oui	ENA
R6g	Oui	Oui	Non	Oui	ENA
R6gd	Non	Oui	NVME *	Oui	ENA
T2	Oui	Non	Non	Non	Non
T3	Oui	Oui	Non	Non	ENA
T3a	Oui	Oui	Non	Non	ENA
u-xtb1.metal	Oui	Oui	Non	Non	ENA
X1	Non	Non	SSD *	Oui	ENA
X1e	Non	Non	SSD *	Oui	ENA
z1d	Non	Oui	NVMe *	Oui	ENA

* Le volume du périphérique racine doit être un volume Amazon EBS.

Le tableau suivant récapitule les fonctions de mise en réseau et de stockage prises en charge par les types d'instance de la génération précédente.

	Stockage d'instance	Groupe de placement	Mise en réseau améliorée
C3	SSD	Oui	Intel 82599 VF
G2	SSD	Oui	Non
I2	SSD	Oui	Intel 82599 VF
M3	SSD	Non	Non
R3	SSD	Oui	Intel 82599 VF

Limites d'instance

Le nombre total d'instances que vous pouvez lancer dans une région est soumis à une limite, de même qu'il existe des limites supplémentaires pour certains types d'instance.

Pour plus d'informations sur les limites par défaut, consultez Combien d'instances puis-je exécuter dans Amazon EC2 ?

Pour plus d'informations sur l'affichage de vos limites actuelles ou la demande d'augmentation de vos limites actuelles, consultez Quotas de service Amazon EC2 (p. 1322).

Instances à usage général

Les instances à usage général fournissent un équilibre entre les ressources de calcul, de mémoire et de mise réseau. Elles peuvent être utilisées pour un large éventail de charges de travail.

Instances A1

Ces instances sont parfaitement adaptées aux charges de travail adaptatives qui sont prises en charge par l'écosystème Arm. Ces instances conviennent à ce qui suit :

- Serveurs Web
- Microservices à conteneurs

Les instances de matériel nu, comme `a1.metal`, offrent à vos applications un accès direct aux ressources physiques du serveur hôte, comme les processeurs et la mémoire.

Pour de plus amples informations, veuillez consulter AWS Graviton Processor et Instances Amazon EC2 A1.

Instances M5 et M5a

Ces instances permettent de créer une infrastructure de cloud idéale en fournissant un équilibre entre ressources de calcul, de mémoire et de mise en réseau pour un large éventail d'applications déployées dans le cloud. Elles conviennent à ce qui suit :

- Bases de données de petite et moyenne taille
- Tâches de traitement des données nécessitant une mémoire supplémentaire
- Flottes de mise en cache
- Exécution de serveurs backend pour SAP, Microsoft SharePoint, le calcul en cluster et d'autres applications d'entreprise

Pour de plus amples informations, veuillez consulter InstancesAmazon EC2 M5 et M5a.

Les instances de matériel nu, comme `m5.metal`, offrent à vos applications un accès direct aux ressources physiques du serveur hôte, comme les processeurs et la mémoire. Ces instances conviennent à ce qui suit :

- Charges de travail nécessitant un accès à des fonctions matérielles de bas niveau (Intel VT, par exemple) qui ne sont pas disponibles ou entièrement prises en charge dans les environnements virtualisés
- Applications nécessitant un environnement non virtualisé pour des questions de licence ou d'assistance

Instances M6g et M6gD

Ces instances sont alimentées par des processeurs AWS Graviton2 et fournissent un calcul, une mémoire et une mise en réseau équilibrés pour une large gamme de charges de travail à usage général. Elles conviennent aux scénarios suivants :

- Serveurs d'applications
- Microservices
- Serveurs de jeu
- Stockages de données de taille moyenne
- Flottes de mise en cache

Les instances de matériel nu, comme `m6g.metal`, offrent à vos applications un accès direct aux ressources physiques du serveur hôte, comme les processeurs et la mémoire.

Pour de plus amples informations, veuillez consulter Instances M6g Amazon EC2.

Instances T2, T3

Ces instances offrent des performances d'UC de base et la possibilité d'atteindre un niveau supérieur lorsque votre charge de travail l'exige. Une instance en mode illimité peut maintenir des performances d'UC élevées pour toute période donnée en cas de nécessité. Pour plus d'informations, consultez Instances à capacité extensible (p. 233). Ces instances conviennent à ce qui suit :

- Sites et applications Web
- Référentiels de code
- Environnements de développement, de génération, de test ou de mise en place
- Microservices

Pour de plus amples informations, veuillez consulter Instances T2 Amazon EC2, Instances T3 Amazon EC2 et Instances T4g Amazon EC2.

Sommaire

- Spécifications matérielles (p. 224)
- Performances de l'instance (p. 227)
- Performances réseau (p. 227)
- Performances d'E/S sur SSD (p. 229)
- Fonctionnalités des instances (p. 231)
- Notes de mise à jour (p. 231)
- Instances à capacité extensible (p. 233)

Spécifications matérielles

Vous trouverez ci-dessous un résumé des spécifications matérielles relatives aux instances à usage général.

Type d'instance	vCPU par défaut	Mémoire (Gio)
`a1.medium`	1	2
`a1.large`	2	4
`a1.xlarge`	4	8
`a1.2xlarge`	8	16
`a1.4xlarge`	16	32
`a1.metal`	16	32
`m4.large`	2	8
`m4.xlarge`	4	16
`m4.2xlarge`	8	32
`m4.4xlarge`	16	64
`m4.10xlarge`	40	160
`m4.16xlarge`	64	256
`m5.large`	2	8
`m5.xlarge`	4	16
`m5.2xlarge`	8	32
`m5.4xlarge`	16	64
`m5.8xlarge`	32	128
`m5.12xlarge`	48	192
`m5.16xlarge`	64	256
`m5.24xlarge`	96	384
`m5.metal`	96	384
`m5a.large`	2	8
`m5a.xlarge`	4	16
`m5a.2xlarge`	8	32
`m5a.4xlarge`	16	64
`m5a.8xlarge`	32	128
`m5a.12xlarge`	48	192
`m5a.16xlarge`	64	256

Type d'instance	vCPU par défaut	Mémoire (Gio)
`m5a.24xlarge`	96	384
`m5ad.large`	2	8
`m5ad.xlarge`	4	16
`m5ad.2xlarge`	8	32
`m5ad.4xlarge`	16	64
`m5ad.8xlarge`	32	128
`m5ad.12xlarge`	48	192
`m5ad.16xlarge`	64	256
`m5ad.24xlarge`	96	384
`m5d.large`	2	8
`m5d.xlarge`	4	16
`m5d.2xlarge`	8	32
`m5d.4xlarge`	16	64
`m5d.8xlarge`	32	128
`m5d.12xlarge`	48	192
`m5d.16xlarge`	64	256
`m5d.24xlarge`	96	384
`m5d.metal`	96	384
`m5dn.large`	2	8
`m5dn.xlarge`	4	16
`m5dn.2xlarge`	8	32
`m5dn.4xlarge`	16	64
`m5dn.8xlarge`	32	128
`m5dn.12xlarge`	48	192
`m5dn.16xlarge`	64	256
`m5dn.24xlarge`	96	384
`m5n.large`	2	8
`m5n.xlarge`	4	16
`m5n.2xlarge`	8	32
`m5n.4xlarge`	16	64
`m5n.8xlarge`	32	128

Type d'instance	vCPU par défaut	Mémoire (Gio)
`m5n.12xlarge`	48	192
`m5n.16xlarge`	64	256
`m5n.24xlarge`	96	384
`m6g.medium`	1	4
`m6g.large`	2	8
`m6g.xlarge`	4	16
`m6g.2xlarge`	8	32
`m6g.4xlarge`	16	64
`m6g.8xlarge`	32	128
`m6g.12xlarge`	48	192
`m6g.16xlarge`	64	256
`m6g.metal`	64	256
`m6gd.medium`	1	4
`m6gd.large`	2	8
`m6gd.xlarge`	4	16
`m6gd.2xlarge`	8	32
`m6gd.4xlarge`	16	64
`m6gd.8xlarge`	32	128
`m6gd.12xlarge`	48	192
`m6gd.16xlarge`	64	256
`m6gd.metal`	64	256
`t2.nano`	1	0,5
`t2.micro`	1	1
`t2.small`	1	2
`t2.medium`	2	4
`t2.large`	2	8
`t2.xlarge`	4	16
`t2.2xlarge`	8	32
`t3.nano`	2	0,5
`t3.micro`	2	1
`t3.small`	2	2

Type d'instance	vCPU par défaut	Mémoire (Gio)
`t3.medium`	2	4
`t3.large`	2	8
`t3.xlarge`	4	16
`t3.2xlarge`	8	32
`t3a.nano`	2	0,5
`t3a.micro`	2	1
`t3a.small`	2	2
`t3a.medium`	2	4
`t3a.large`	2	8
`t3a.xlarge`	4	16
`t3a.2xlarge`	8	32

Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.

Pour plus d'informations sur la spécification des options d'UC, consultez Optimisations des options d'UC (p. 673).

Performances de l'instance

Les instances optimisées EBS vous permettent d'obtenir régulièrement des performances élevées pour vos volumes EBS en éliminant les conflits entre les E/S Amazon EBS et tout autre trafic réseau de votre instance Certaines instances à visée générale sont optimisées pour EBS par défaut sans frais supplémentaires. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Certains types d'instances à visée générale permettent de contrôler les états C et P du processeur sur Linux. Les états C contrôlent les niveaux de veille d'un noyau lorsqu'il est inactif, tandis que les états P contrôlent les performances attendues d'un noyau (en termes de fréquence d'UC). Pour plus d'informations, consultez Contrôle des états du processeur pour votre instance EC2 (p. 662).

Performances réseau

Vous pouvez activer la mise en réseau améliorée sur les types d'instance pris en charge pour fournir des latences plus faibles, une instabilité moindre sur le réseau et des performances de débit en paquets par seconde (PPS) plus élevées. La plupart des applications ne nécessitent pas en permanence un haut niveau de performances réseau, mais peuvent tirer profit d'un accès à une bande passante accrue lorsqu'elles envoient ou reçoivent des données. Pour plus d'informations, consultez Mise en réseau améliorée sur Linux (p. 864).

Vous trouverez ci-dessous un résumé des performances réseau relatives aux instances à usage général qui prennent en charge la mise en réseau améliorée.

Type d'instance	Performances réseau	Mise en réseau améliorée
`t2.nano` | `t2.micro` | `t2.small` | `t2.medium` |	Jusqu'à 1 Gbit/s	Non pris en charge

Type d'instance	Performances réseau	Mise en réseau améliorée
`t2.large` | `t2.xlarge` | `t2.2xlarge`		
`t3.nano` | `t3.micro` | `t3.small` | `t3.medium` | `t3.large` | `t3.xlarge` | `t3.2xlarge` | `t3a.nano` | `t3a.micro` | `t3a.small` | `t3a.medium` | `t3a.large` | `t3a.xlarge` | `t3a.2xlarge`	Jusqu'à 5 Gbit/s †	ENA (p. 865)
`m4.large`	Modérées	Intel 82599 VF (p. 878)
`m4.xlarge` | `m4.2xlarge` | `m4.4xlarge`	Elevé	Intel 82599 VF (p. 878)
`a1.4xlarge` et tailles inférieures | `a1.metal` | `m5.4xlarge` et tailles inférieures | `m5a.8xlarge` et tailles inférieures | `m5ad.8xlarge` et tailles inférieures | `m5d.4xlarge` et tailles inférieures | `m6g.4xlarge` et tailles inférieures | `m6gd.4xlarge` et tailles inférieures	Jusqu'à 10 Gbit/s †	ENA (p. 865)
`m4.10xlarge`	10 Gb/s	Intel 82599 VF (p. 878)
`m5.8xlarge` | `m5a.12xlarge` | `m5ad.12xlarge` | `m5d.8xlarge`	10 Gb/s	ENA (p. 865)
`m5.12xlarge` | `m5a.16xlarge` | `m5ad.16xlarge` | `m5d.12xlarge` | `m6g.8xlarge` | `m6gd.8xlarge`	12 Gb/s	ENA (p. 865)
`m5.16xlarge` | `m5a.24xlarge` | `m5ad.24xlarge` | `m5d.16xlarge` | `m6g.12xlarge` | `m6gd.12xlarge`	20 Gb/s	ENA (p. 865)
`m5dn.4xlarge` et plus petit | `m5n.4xlarge` et plus petit	Jusqu'à 25 Gbit/s †	ENA (p. 865)
`m4.16xlarge` | `m5.24xlarge` | `m5.metal` | `m5d.24xlarge` | `m5d.metal` | `m5dn.8xlarge` | `m5n.8xlarge` | `m6g.16xlarge` | `m6g.metal` | `m6gd.16xlarge` | `m6gd.metal`	25 Gb/s	ENA (p. 865)
`m5dn.12xlarge` | `m5n.12xlarge`	50 Gb/s	ENA (p. 865)

Type d'instance	Performances réseau	Mise en réseau améliorée
`m5dn.16xlarge` | `m5n.16xlarge`	75 Gb/s	ENA (p. 865)
`m5dn.24xlarge` | `m5n.24xlarge`	100 Gb/s	ENA (p. 865)

† Ces instances utilisent un mécanisme de crédit d'E/S réseau pour attribuer la bande passante réseau aux instances en fonction de l'utilisation moyenne de la bande passante. Elles accumulent des crédits lorsque la bande passante est inférieure à la bande passante de référence et peuvent les utiliser quand elles effectuent les transferts de données réseau. Pour de plus amples informations, ouvrez un dossier de support et renseignez-vous sur la bande passante de référence pour les types d'instance spécifiques qui vous intéressent.

Performances d'E/S sur SSD

Si vous utilisez une AMI Linux avec un noyau de version 4.4 ou ultérieure et que vous utilisez tous les volumes de stockage d'instance basés sur SSD disponibles pour votre instance, vous pouvez obtenir les performances d'IOPS (taille de bloc de 4 096 octets) répertoriées dans le tableau suivant (lorsque la profondeur de la file d'attente est saturée). Sinon, vous obtenez des performances d'I/O inférieures.

Taille d'instance	E/S par seconde en lecture aléatoires 100 %	E/S par seconde en écriture
`m5ad.large` *	30 000	15 000 USD
`m5ad.xlarge` *	59 000	29 000
`m5ad.2xlarge` *	117 000	57 000
`m5ad.4xlarge` *	234 000	114 000
`m5ad.8xlarge`	466 666	233 333
`m5ad.12xlarge`	700 000	340 000
`m5ad.16xlarge`	933 333	466 666
`m5ad.24xlarge`	1 400 000	680 000
`m5d.large` *	30 000	15 000 USD
`m5d.xlarge` *	59 000	29 000
`m5d.2xlarge` *	117 000	57 000
`m5d.4xlarge` *	234 000	114 000
`m5d.8xlarge`	466 666	233 333
`m5d.12xlarge`	700 000	340 000
`m5d.16xlarge`	933 333	466 666
`m5d.24xlarge`	1 400 000	680 000
`m5d.metal`	1 400 000	680 000

Taille d'instance	E/S par seconde en lecture aléatoires 100 %	E/S par seconde en écriture
`m5dn.large` *	30 000	15 000 USD
`m5dn.xlarge` *	59 000	29 000
`m5dn.2xlarge` *	117 000	57 000
`m5dn.4xlarge` *	234 000	114 000
`m5dn.8xlarge`	466 666	233 333
`m5dn.12xlarge`	700 000	340 000
`m5dn.16xlarge`	933 333	466 666
`m5dn.24xlarge`	1 400 000	680 000
`m6gd.medium`	13 438	5 625
`m6gd.large`	26 875	11 250
`m6gd.xlarge`	53 750	22 500
`m6gd.2xlarge`	107 500	45 000
`m6gd.4xlarge`	215 000	90 000
`m6gd.8xlarge`	430 000	180 000
`m6gd.12xlarge`	645 000	270 000
`m6gd.16xlarge`	860 000	360 000
`m6gd.metal`	860 000	360 000

* Pour ces instances, vous pouvez obtenir la performance spécifiée.

Au fur et à mesure que vous remplissez les volumes de stockage d'instance basés sur SSD pour votre instance, le nombre d'E/S en écriture que vous pouvez obtenir diminue. Ceci est dû au travail supplémentaire que le contrôleur SSD doit effectuer pour trouver de l'espace disponible, réécrire les données existantes et effacer l'espace non utilisé pour le rendre réinscriptible. Ce processus de nettoyage de la mémoire se traduit par une amplification d'écriture interne sur le disque SSD, exprimée sous la forme du rapport des opérations d'écriture SSD sur les opérations d'écriture utilisateur. Cette diminution des performances est encore plus importante si les opérations d'écriture ne sont pas exprimées en multiples de 4 096 octets ou ne sont pas alignées sur une limite de 4 096 octets. Si vous écrivez une quantité d'octets plus faible ou des octets qui ne sont pas alignés, le contrôleur SSD doit lire les données environnantes et stocker le résultat dans un nouvel emplacement. Ce modèle se traduit par une forte augmentation de l'amplification d'écriture, une latence accrue et une diminution considérable des performances d'E/S.

Les contrôleurs SSD peuvent utiliser plusieurs stratégies pour réduire l'impact de l'amplification d'écriture. Une telle stratégie consiste à réserver un espace dans le stockage d'instance SSD afin que le contrôleur puisse gérer efficacement l'espace disponible pour les opérations d'écriture. Cette solution est appelée sur-approvisionnement. Les volumes de stockage d'instance SSD fournis à une instance n'ont pas d'espace réservé pour le sur-approvisionnement. Pour réduire l'amplification d'écriture, nous vous conseillons de laisser 10 % du volume non partitionné que le contrôleur SSD pourra utiliser pour le sur-approvisionnement. Cela diminue le stockage que vous pouvez utiliser, mais augmente les performances même si le disque est proche de sa capacité maximale.

Pour les volumes de stockage d'instance qui prennent en charge TRIM, vous pouvez utiliser la commande TRIM pour informer le contrôleur SSD lorsque vous n'avez plus besoin des données que vous avez écrites. Cela fournit au contrôleur plus d'espace disponible, ce qui peut réduire l'amplification d'écriture et augmenter les performances. Pour plus d'informations, consultez Prise en charge de TRIM sur les volumes de stockage d'instance (p. 1278).

Fonctionnalités des instances

Voici un résumé des fonctions pour les instances à usage général :

	EBS uniquement	EBS NVMe	Stockage d'instance	Groupe de placement
A1	Oui	Oui	Non	Oui
M4	Oui	Non	Non	Oui
M5	Oui	Oui	Non	Oui
M5a	Oui	Oui	Non	Oui
M5ad	Non	Oui	NVMe *	Oui
M5d	Non	Oui	NVMe *	Oui
M5dn	Non	Oui	NVMe *	Oui
M5n	Oui	Oui	Non	Oui
M6g	Oui	Oui	Non	Oui
M6gd	Non	Oui	NVMe *	Oui
T2	Oui	Non	Non	Non
T3	Oui	Oui	Non	Non
T3a	Oui	Oui	Non	Non

* Le volume du périphérique racine doit être un volume Amazon EBS.

Pour plus d'informations, consultez les ressources suivantes :

- Amazon EBS et NVMe sur des instances Linux (p. 1209)
- Stockage d'instance Amazon EC2 (p. 1264)
- Groupes de placement (p. 924)

Notes de mise à jour

- Les instances M5, M5d et T3 sont équipées d'un processeur Intel Xeon Platinum 8000 de 3,1 GHz de la première génération (Skylake-SP) ou de la deuxième génération (Cascade Lake).
- Les instances M5a, M5ad et T3a utilisent un processeur AMD EPYC 7000 à 2,5 GHz.
- Les instances A1 utilisent un processeur AWS Graviton à 2,3 GHz basé sur une architecture Arm 64 bits.
- Les instances M6g et M6gd utilisent un processeur AWS Graviton2 basé sur l'architecture Arm 64 bits.
- Les instances de types M4, M5, M5a, M5ad, M5d, `t2.large` et de plus grande taille et `t3.large` et de plus grande taille, ainsi que `t3a.large` et de plus grande taille requièrent des AMIs HVM 64 bits. Elles

sont dotées d'une mémoire élevée et un système d'exploitation 64 bits est nécessaire pour tirer parti de cette capacité. Les AMI HVM offrent des performances supérieures par rapport aux AMI paravirtuelles (PV) sur les types d'instance à mémoire élevée. De plus, vous devez utiliser une AMI HVM pour tirer parti de la mise en réseau améliorée.

- Les instances reposant sur le système Nitro (p. 218) présentent les exigences suivantes :
  - Les pilotes NVMe (p. 1209) doivent être installés.
  - Les pilotes Elastic Network Adapter (ENA) (p. 865) doivent être installés.

  Les AMI Linux suivantes répondent aux critères suivants :
  - Amazon Linux 2
  - Amazon Linux AMI 2018.03
  - Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure
  - Red Hat Enterprise Linux 7.4 ou version ultérieure
  - SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
  - CentOS 7.4.1708 ou version ultérieure
  - FreeBSD 11.1 ou version ultérieure
  - Debian GNU/Linux 9 ou version ultérieure
- Les instances dotées de processeurs AWS Graviton présentent les exigences suivantes :
  - Utiliser une AMI pour l'architecture Arm 64 bits.
  - Prendre en charge le démarrage via UEFI avec les tables ACPI, ainsi que l'insertion à chaud ACPI des périphériques PCI.

  Les AMI Linux suivantes répondent à ces critères :
  - Amazon Linux 2 (Arm 64 bits)
  - Ubuntu 16.04 ou une version ultérieure (Arm 64 bits)
  - Red Hat Enterprise Linux 8.0 ou une version ultérieure (Arm 64 bits)
  - SUSE Linux Enterprise Server 15 ou version ultérieure (Arm 64 bits)
  - Debian 10 ou version ultérieure (Arm 64 bits)
- Les instances reposant sur le système Nitro prennent en charge un maximum de 28 attachements, y compris les interfaces réseau, les volumes EBS et les volumes de stockage d'instance NVMe. Pour plus d'informations, consultez Limites de volume du système Nitro (p. 1286).
- Le lancement d'une instance en matériel nu démarre le serveur sous-jacent, qui inclut la vérification de tous les composants du matériel et du microprogramme. Cela signifie que 20 minutes peuvent s'écouler entre le moment où l'instance passe à l'état d'exécution et le moment où elle devient disponible sur le réseau.
- Attacher ou détacher des volumes EBS ou des interfaces réseau secondaires à partir d'une instance en matériel nu requiert la prise en charge de l'enfichage à chaud natif de PCIe. Amazon Linux 2 et les dernières versions de l'Amazon Linux AMI prennent en charge l'enfichage à chaud natif de PCIe, ce qui n'est pas le cas des versions antérieures. Vous devez activer les options de configuration suivantes du noyau Linux :

  ```
  CONFIG_HOTPLUG_PCI_PCIE=y
  CONFIG_PCIEASPM=y
  ```

- Les instances en matériel nu utilisent un périphérique série basé sur PCI plutôt qu'un périphérique série basé sur le port d'E/S. Le noyau Linux kernel en amont et les dernières AMI Amazon Linux prennent en charge ce périphérique. Les instances en matériel nu fournissent également un tableau SPCR ACPI afin de permettre au système d'utiliser automatiquement le périphérique série basé sur PCI. Les dernières AMI Windows utilisent automatiquement le périphérique série basé sur PCI.
- system-logind ou acpid doit être installé sur les instances reposant sur le système Nitro afin que ces instances prennent en charge les fermetures propres via les demandes d'API.

- Le nombre total d'instances que vous pouvez lancer dans une région est soumis à une limite, et il existe des limites supplémentaires sur certains types d'instances. Pour de plus amples informations, veuillez consulter Combien d'instances est-il possible d'exécuter dans Amazon EC2 ? dans les questions fréquentes Amazon EC2.

Instances à capacité extensible

Les instances à capacité extensible sont conçues pour offrir des performances d'UC de base et la possibilité d'atteindre un niveau supérieur lorsque votre charge de travail l'exige. Les instances à capacité extensible conviennent bien à une large gamme d'applications à usage général. Des exemples incluent des microservices, des applications interactives à faible latence, de petites et moyennes bases de données, les postes de travail virtuels, les environnements de développement, de génération et de transit, les référentiels de code et les prototypes de produits.

Les instances à capacité extensible constituent les seuls types d'instances qui utilisent des crédits pour l'utilisation de l'UC. Pour plus d'informations sur la tarification des instances et sur le matériel, consultez Tarification Amazon EC2 et Types d'instance Amazon EC2.

Si votre compte a moins de 12 mois, vous pouvez utiliser une instance `t2.micro` gratuitement (ou une instance `t3.micro` dans les régions où `t2.micro` n'est pas disponible) dans certaines limites d'utilisation. Pour plus d'informations, consultez la page sur l'offre gratuite AWS.

Sommaire

- Exigences des instances à capacité extensible (p. 233)
- Bonnes pratiques (p. 234)
- Crédits UC et utilisation de référence pour les instances à capacité extensible (p. 234)
- Mode illimité pour les instances à capacité extensible (p. 237)
- Mode standard pour les instances à capacité extensible (p. 246)
- Utilisation des instances standard à capacité extensible (p. 261)
- Surveillance de vos crédits CPU (p. 266)

Exigences des instances à capacité extensible

Les exigences pour ces instances sont les suivantes :

- Les familles d'instances prises en charge sont : T2, T3.
- Les options d'achat prises en charge sont les suivantes : Instances à la demande, Instances réservées, Instances dédiées et Instances Spot. Ces instances ne sont pas prises en charge en tant qu'instances planifiées ou sur un Hôte dédié. Pour de plus amples informations, veuillez consulter Options d'achat d'instance (p. 323).
- Assurez-vous que la taille d'instance que vous choisissez correspond à la configuration minimum requise en matière de mémoire par votre système d'exploitation et vos applications. Les systèmes d'exploitation aux interfaces utilisateur graphiques qui consomment une quantité importante de mémoire et de ressources UC (par exemple Windows) peuvent nécessiter une taille d'instance `t2.micro` ou supérieure dans de nombreux cas d'utilisation. Si les exigences de votre charge de travail en matière de mémoire et d'UC augmentent au fil du temps, vous pouvez la mettre à l'échelle vers des instances de plus grande taille du même type d'instance ou d'un autre type d'instance.
- Pour des prérequis supplémentaires, consultez les notes de mise à jour relatives aux instances à usage général (p. 231).

Bonnes pratiques

Suivez ces bonnes pratiques pour tirer le meilleur profit et la plus grande satisfaction des instances à capacité extensible.

- Utiliser une AMI recommandée – Utilisez une AMI qui fournit les pilotes requis. Pour plus d'informations, consultez Notes de mise à jour (p. 231).
- Activer la récupération d'instance – Créez une alarme CloudWatch qui surveille une instance EC2 et la récupère automatiquement si elle est dégradée pour une raison quelconque. Pour plus d'informations, consultez Ajout d'actions de récupération aux alarmes Amazon CloudWatch (p. 789).

Crédits UC et utilisation de référence pour les instances à capacité extensible

Les types d'instances Amazon EC2 traditionnels fournissent une utilisation fixe de l'UC, tandis que les instances à capacité extensible fournissent un niveau d'utilisation de l'UC de référence avec la possibilité d'atteindre un niveau supérieur. L'utilisation de référence et la possibilité d'extension sont régies par les crédits UC.

Les crédits UC utilisés dépendent de l'utilisation de l'UC. Les scénarios suivants utilisent tous un crédit CPU :

- Un processeur virtuel fonctionnant à 100 % pendant une minute
- Un processeur virtuel fonctionnant à 50 % pendant deux minutes
- Deux processeurs virtuels fonctionnant à 25 % pendant deux minutes

Sommaire

- Obtention de crédits UC (p. 234)
- Taux d'obtention de crédits UC (p. 236)
- Limite d'accumulation de crédits UC (p. 236)
- Durée de vie des crédits UC accumulés (p. 236)
- Utilisation de référence (p. 237)

Obtention de crédits UC

Chaque instance à capacité extensible gagne continuellement (à une résolution de l'ordre de la milliseconde) un taux déterminé de crédits UC par heure, en fonction de sa taille. Le processus de comptabilité par lequel les crédits s'accumulent ou sont dépensés se fait également sur une résolution de l'ordre de la milliseconde. Vous n'avez donc pas à vous soucier de consommer trop de crédits UC ; une brève rafale dans l'utilisation de l'UC ne se sert que d'une petite quantité des crédits UC.

Si une instance à capacité extensible utilise moins de ressources d'UC que ne le requière son utilisation de référence (par exemple lorsqu'elle est inactive), les crédits UC inutilisés sont accumulés dans le solde de crédits UC. Si une instance à capacité extensible a besoin d'étendre l'utilisation au-dessus du niveau d'utilisation de référence, elle dépense les crédits accumulés. Plus une instance à capacité extensible accumule de crédits, plus elle peut dépasser son niveau d'utilisation de référence longtemps, quand l'UC le demande.

Le tableau ci-dessous répertorie les types d'instances à capacité extensible, le taux auquel les crédits UC sont gagnés par heure, le nombre maximal de crédits UC gagnés qu'une instance peut accumuler, le nombre de processeurs vCPU par instance et le niveau d'utilisation de référence sous la forme d'un pourcentage des performances d'un cœur complet (utilisant un seul processeur vCPU).

Type d'instance	Crédits UC gagnés par heure	Maximum de crédits gagnés pouvant être accumulés*	vCPU	Utilisation de référence par vCPU
T2				
`t2.nano`	3	72	1	5 %
`t2.micro`	6	144	1	10 %
`t2.small`	12	288	1	20 %
`t2.medium`	24	576	2	20%**
`t2.large`	36	864	2	30%**
`t2.xlarge`	54	1296	4	22,5%**
`t2.2xlarge`	81.6	1958.4	8	17%**
T3				
`t3.nano`	6	144	2	5%**
`t3.micro`	12	288	2	10%**
`t3.small`	24	576	2	20%**
`t3.medium`	24	576	2	20%**
`t3.large`	36	864	2	30%**
`t3.xlarge`	96	2304	4	40%**
`t3.2xlarge`	192	4608	8	40%**
T3a				
`t3a.nano`	6	144	2	5%**
`t3a.micro`	12	288	2	10%**
`t3a.small`	24	576	2	20%**
`t3a.medium`	24	576	2	20%**
`t3a.large`	36	864	2	30%**
`t3a.xlarge`	96	2304	4	40%**
`t3a.2xlarge`	192	4608	8	40%**

* Le nombre de crédits pouvant être accumulés est équivalent au nombre de crédits pouvant être gagnés en 24 heures.

** Le pourcentage d'utilisation de référence dans le tableau est par vCPU. Dans CloudWatch, l'utilisation de l'UC est indiquée par vCPU. Par exemple, l'utilisation de l'UC pour une instance `t3.large` fonctionnant au niveau de référence est de 30 % dans les métriques d'UC CloudWatch. Pour de plus amples informations sur le calcul de l'utilisation de référence, veuillez consulter Utilisation de référence (p. 237).

Taux d'obtention de crédits UC

Le nombre de crédits UC gagnés par heure est déterminé par la taille d'instance. Par exemple, une instance `t3.nano` gagne six crédits par heure, tandis qu'une instance `t3.small` en gagne 24 par heure. Le tableau précédent répertorie le taux d'obtention de crédits pour l'ensemble des instances.

Limite d'accumulation de crédits UC

Si les crédits gagnés n'expirent jamais sur une instance en cours d'exécution, il existe une limite pour le nombre de crédits gagnés pouvant être accumulés par une instance. Cette limite est déterminée par la limite du solde de crédits UC. Une fois la limite atteinte, les nouveaux crédits gagnés sont rejetés, comme l'indique l'image suivante. Le compartiment plein indique la limite du solde de crédits UC, tandis que le débordement signale les crédits excédant la limite qui viennent d'être gagnés.

La limite du solde de crédits UC diffère pour chaque taille d'instance Par exemple, une instance `t3.micro` peut accumuler un maximum de 288 crédits UC gagnés dans le solde de crédits UC. Le tableau précédent répertorie le nombre maximum de crédits gagnés pouvant être cumulés par instance

Les instances T2 standard gagnent également des crédits de lancement. Les crédits de lancement ne sont pas comptés dans la limite du solde de crédits UC. Si une instance T2 n'a pas dépensé ses crédits de lancement et reste inactive pendant 24 heures tout en accumulant des crédits gagnés, son solde de crédits d'UC est affiché comme dépassant la limite. Pour plus d'informations, consultez Crédits de lancement (p. 247).

Les instances T3 ne gagnent pas de crédits de lancement. Ces instances sont lancées en mode `unlimited` par défaut et peuvent par conséquent s'exécuter en mode rafale immédiatement après leur démarrage, sans avoir besoin de crédits de lancement.

Durée de vie des crédits UC accumulés

Les crédits UC sur une instance en cours d'exécution n'expirent pas.

Pour T2, le solde de crédits UC n'est pas conservé entre les arrêts et les démarrages des instances. Si vous arrêtez une instance T2, celle-ci perd tous ses crédits accumulés.

Pour les instances T3, le solde de crédits pour l'UC est conservé pendant sept jours après l'arrêt d'une instance ; les crédits sont perdus ensuite. Si vous démarrez l'instance dans les sept jours, aucun crédit n'est perdu.

Pour plus d'informations, consultez `CPUCreditBalance` dans le tableau des métriques CloudWatch (p. 266).

Utilisation de référence

L'utilisation de référence est le niveau auquel l'UC peut être utilisé pour un solde créditeur net de zéro, lorsque le nombre de crédits UC gagnés correspond au nombre de crédits UC utilisés. L'utilisation de référence est également appelée la référence.

L'utilisation de référence est exprimée en pourcentage de l'utilisation du vCPU, calculé comme suit :

```
(number of credits earned/number of vCPUs)/60 minutes = % baseline utilization
```

Par exemple, une instance `t3.nano`, avec 2 vCPU, gagne 6 crédits par heure, ce qui donne une utilisation de référence de 5 %, calculée comme suit :

```
(6 credits earned/2 vCPUs)/60 minutes = 5% baseline utilization
```

Une instance `t3.xlarge`, avec 4 vCPU, gagne 96 crédits par heure, ce qui donne une utilisation de référence de 40 % (`(96/4)/60`).

Mode illimité pour les instances à capacité extensible

Une instance à capacité extensible configurée en mode `unlimited` peut maintenir une utilisation d'UC élevée pour toute période donnée en cas de nécessité. Le prix horaire d'une instance couvre automatiquement tous les pics d'utilisation d'UC si l'utilisation moyenne de l'UC de l'instance est égale ou inférieure au niveau de base sur une période glissante de 24 heures ou pendant la durée de vie de l'instance si celle-ci est plus courte.

Pour la majorité des charges de travail à usage général, les instances configurées en mode `unlimited` fournissent d'excellentes performances sans frais supplémentaires. Si l'instance s'exécute avec une utilisation d'UC supérieure pendant une période prolongée, c'est possible moyennant des frais supplémentaires fixes par heure vCPU. Pour de plus amples informations sur la tarification des instances, veuillez consulter Tarification Amazon EC2 et la section relative à la tarification en mode illimité sur la page de tarification Amazon EC2 à la demande.

Si vous utilisez une instance `t2.micro` ou `t3.micro` relevant de l'offre gratuite AWS et si vous l'utilisez en mode `unlimited`, des frais peuvent s'appliquer si votre utilisation moyenne sur une période glissante de 24 heures excède l'utilisation de référence (p. 237) de l'instance.

Les instances T3 sont lancées en mode `unlimited` par défaut. Si l'utilisation moyenne de l'UC sur une période de 24 heures dépasse le niveau de référence, vous devrez payer des frais pour les crédits excédentaires. Si vous lancez des Instances Spot en mode `unlimited` et que vous prévoyez de les utiliser immédiatement et pour une courte durée, sans temps d'inactivité pour accumuler les crédits d'UC, vous devrez payer des frais pour les crédits excédentaires. Nous vous recommandons de lancer vos Instances Spot en mode standard (p. 246) pour éviter des coûts plus élevés. Pour de plus amples informations, veuillez consulter Les crédits excédentaires peuvent occasionner des frais (p. 241) et Instances à capacité extensible (p. 465).

Sommaire

- Concepts du mode illimité (Unlimited) (p. 238)
  - Fonctionnement des instances illimitées à capacité extensible (p. 238)
  - Quand utiliser le mode illimité/mode d'UC fixe ? (p. 238)
  - Les crédits excédentaires peuvent occasionner des frais (p. 241)
  - Pas de crédit de lancement pour les instances T2 illimitées (p. 241)
  - Activation du mode illimité (p. 241)
  - Comportement des crédits lors du basculement entre Illimité et Standard (p. 242)
  - Surveillance de l'utilisation du crédit (p. 242)
- Exemples de modes illimités (p. 242)
  - Exemple 1 : Explication de l'utilisation des crédits avec T3 illimité (p. 242)
  - Exemple 2 : Explication de l'utilisation des crédits avec T2 illimité (p. 244)

Concepts du mode illimité (Unlimited)

Le mode `unlimited` est une option de configuration de crédit pour les instances à capacité extensible. Il peut être activé ou désactivé à tout moment pour une instance en cours d'exécution ou arrêtée. Vous pouvez définir `unlimited` comme formule de crédit par défaut au niveau du compte par région AWS, par famille d'instances à capacité extensible, de sorte à lancer l'ensemble des nouvelles instances à capacité extensible du compte à l'aide de la formule de crédit par défaut.

Fonctionnement des instances illimitées à capacité extensible

Si une instance à capacité extensible configurée en mode `unlimited` épuise son solde de crédits UC, elle peut dépenser ses crédits excédentaires pour dépasser le niveau de référence (p. 237). Si son utilisation de l'UC chute au-dessous du niveau de référence, elle se sert des crédits UC gagnés pour rembourser progressivement les crédits excédentaires dépensés plus tôt. La possibilité de gagner des crédits UC pour rembourser progressivement des crédits excédentaires permet à Amazon EC2 d'obtenir l'utilisation moyenne de l'UC d'une instance sur une période de 24 heures. Si l'utilisation moyenne de l'UC dépasse le niveau de base pendant une période de 24 heures, l'utilisation supplémentaire est facturée pour l'instance selon un tarif supplémentaire fixe par heure de processeur virtuel.

Le graphique suivant montre l'utilisation de l'UC d'une instance `t3.large`. Le niveau de base de l'utilisation de l'UC pour une instance `t3.large` est de 30 %. Si l'instance s'exécute à un taux d'utilisation d'UC de 30 % ou moins en moyenne sur une période de 24 heures, aucun frais supplémentaire ne s'applique, car le coût est déjà couvert par le prix horaire de l'instance. Toutefois, si l'instance s'exécute à un taux d'utilisation d'UC de 40 % en moyenne sur une période de 24 heures, comme le montre le graphique, les 10 % supplémentaires d'utilisation de l'UC sont facturés pour l'instance selon un tarif supplémentaire fixe par heure de processeur virtuel.

Pour plus d'informations sur l'utilisation de référence par vCPU pour chaque type d'instance et sur le nombre de crédits gagnés par chaque type d'instance, consultez le tableau des crédits (p. 235).

Quand utiliser le mode illimité/mode d'UC fixe ?

Pour déterminer si vous devez utiliser une instance à capacité extensible en mode `unlimited`, telle qu'une instance T3, ou une instance à performance fixe, telle qu'une instance M5, vous devez déterminer l'utilisation d'UC de seuil de rentabilité. L'utilisation d'UC de seuil de rentabilité pour une instance à capacité extensible est le point où une instance à capacité extensible coûte autant qu'une instance à performance fixe. L'utilisation d'UC de seuil de rentabilité vous aide à déterminer les éléments suivants :

- Si l'utilisation moyenne de l'UC sur une période de 24 heures est égale ou inférieure à l'utilisation d'UC de seuil de rentabilité, utilisez une instance à capacité extensible en mode `unlimited` afin de pouvoir bénéficier du prix inférieur d'une instance à capacité extensible tout en profitant de la même performance que fournirait une instance à performance fixe.
- Si l'utilisation moyenne de l'UC sur une période de 24 heures est supérieure à l'utilisation d'UC de seuil de rentabilité, l'instance à capacité extensible coûtera plus qu'une instance à performance fixe de taille équivalente. Si une instance T3 fonctionne continuellement à un taux d'utilisation d'UC de 100 %, vous paierez en définitive environ 1,5 fois le prix d'une instance M5 de taille équivalente.

Le graphique suivant montre l'utilisation d'UC de seuil de rentabilité où une instance `t3.large` coûte autant qu'une instance `m5.large`. L'utilisation d'UC de seuil de rentabilité pour une instance `t3.large` est de 42,5 %. Si l'utilisation moyenne de l'UC est de 42,5 %, le coût de l'exécution de l'instance `t3.large` est identique à celui d'une instance `m5.large`, et il s'avère supérieur si l'utilisation moyenne de l'UC dépasse 42,5 %. Si la charge de travail nécessite une utilisation moyenne de l'UC inférieure à 42,5 %, vous pouvez tirer profit du prix inférieur de l'instance `t3.large` tout en obtenant la même performance que fournirait une instance `m5.large`.

Le tableau suivant indique comment calculer l'utilisation d'UC de seuil de rentabilité, qui vous permettra de déterminer quand il est moins onéreux d'utiliser une instance à capacité extensible en mode `unlimited` ou une instance à performance fixe. Les colonnes du tableau sont étiquetées de A à K.

Type d'instance	vCPU	T3 – Prix*/ heure	M5 – Prix*/ heure	Différence de prix	Utilisation de référence T3 par vCPU (%)	Frais par heure de processeur virtuel pour crédits excédentaire	Frais par minute de processeur virtuel	Minutes d'extension supplémentaires disponibles par processeur virtuel	% d'UC supplémentaire disponible	% d'UC de seuil de rentabilité
A	B	C	D	E = D - C	F	G	H = G / 60	I = E / H	J = (I / 60) / B	K = F + J
t3.large	2	0,0835 USD	0,096 USD	0,0125 USD	30 %	0,05 USD	0,000833 USD	15	12,5%	42,5 %

* Le prix se rapporte à la région us-east-1 et au système d'exploitation Linux.

Le tableau fournit les informations suivantes :

- La colonne A indique le type d'instance, `t3.large`.
- La colonne B indique le nombre de processeurs virtuels pour l'instance `t3.large`.
- La colonne C indique le prix d'une instance `t3.large` par heure.
- La colonne D indique le prix d'une instance `m5.large` par heure.
- La colonne E indique la différence de prix entre l'instance `t3.large` et l'instance `m5.large`.
- La colonne F indique l'utilisation de référence par vCPU de l'instance `t3.large`, qui est de 30 %. Au niveau de base, le coût horaire de l'instance couvre le coût de l'utilisation de l'UC.
- La colonne G indique les frais supplémentaires fixes par heure de processeur virtuel facturés pour une instance si elle passe à 100 % d'utilisation de l'UC après avoir épuisé ses crédits gagnés.
- La colonne H indique les frais supplémentaires fixes par minute de processeur virtuel facturés pour une instance si elle passe à 100 % d'utilisation de l'UC après avoir épuisé ses crédits gagnés.
- La colonne I indique le nombre de minutes supplémentaires pendant lesquelles l'instance `t3.large` peut fonctionner par heure à 100 % d'UC pour le même prix par heure qu'une instance `m5.large`.
- La colonne J indique l'utilisation d'UC supplémentaire (en %) par rapport à l'utilisation de base que l'instance peut assurer pour le même prix par heure qu'une instance `m5.large`.
- La colonne K indique l'utilisation d'UC de seuil de rentabilité (en %) que l'instance `t3.large` peut assurer sans générer plus de frais que l'instance `m5.large`. Au dessus de ce seuil, l'instance `t3.large` coûte plus que l'instance `m5.large`.

Le tableau ci-dessous indique l'utilisation d'UC de seuil de rentabilité (en %) des types d'instance T3 par rapport aux types d'instance M5 de taille équivalente.

Type d'instance T3	Utilisation d'UC de seuil de rentabilité (en %) de T3 par rapport à M5
`t3.large`	42,5 %
`t3.xlarge`	52,5 %
`t3.2xlarge`	52,5 %

Les crédits excédentaires peuvent occasionner des frais

Si l'utilisation moyenne de l'UC d'une instance est égale ou inférieure au niveau de base, aucuns frais supplémentaires ne sont appliqués à l'instance. Comme une instance gagne un nombre maximum de crédits (p. 235) sur une période de 24 heures (par exemple, une instance `t3.micro` peut acquérir un maximum de 288 crédits sur une période de 24 heures), elle peut dépenser des crédits excédentaires jusqu'à ce maximum sans être facturée immédiatement.

Cependant, si l'utilisation de l'UC reste supérieure au niveau de référence, l'instance ne peut pas gagner suffisamment de crédits pour rembourser progressivement les crédits excédentaires qu'elle a dépensés. Des frais supplémentaires fixes s'appliquent par heure vCPU aux crédits excédentaires qui ne sont pas remboursés progressivement.

Les crédits excédentaires qui ont été dépensés précédemment sont facturés lorsque l'une des situations suivantes se produit :

- Les crédits excédentaires dépensés dépassent le nombre maximum de crédits (p. 235) que l'instance peut gagner sur une période de 24 heures. Les crédits excédentaires dépensés au-dessus de ce maximum sont facturés à la fin de l'heure.
- L'instance est arrêtée ou résiliée.
- L'instance bascule du mode `unlimited` au mode `standard`.

La métrique `CPUSurplusCreditBalance` CloudWatch effectue le suivi des crédits excédentaires dépensés. Les crédits excédentaires facturés sont suivis par la métrique `CPUSurplusCreditsCharged` CloudWatch. Pour plus d'informations, consultez Métriques CloudWatch supplémentaires pour les instances à capacité extensible (p. 266).

Pas de crédit de lancement pour les instances T2 illimitées

Les instances T2 standard reçoivent des crédits de lancement (p. 247), mais les instances T2 illimité n'en reçoivent pas. Une instance T2 illimité peut dépasser le niveau de référence à tout moment sans frais supplémentaires tant que l'utilisation moyenne de son UC est égale ou inférieure au niveau de référence sur une période glissante de 24 heures ou pendant sa durée de vie (selon la période la plus courte). De ce fait, les instances T2 illimité ne nécessitent pas de crédits de lancement pour obtenir des performances élevées dès le lancement.

Si une instance T2 passe du mode `standard` au mode `unlimited`, tous les crédits de lancement accumulés sont supprimés de la métrique `CPUCreditBalance` avant que la métrique `CPUCreditBalance` restante soit reportée.

Les instances T3 ne reçoivent jamais de crédits de lancement.

Activation du mode illimité

Vous pouvez passer du mode `unlimited` au mode `standard` et du mode `standard` au mode `unlimited` à tout moment sur une instance en cours d'exécution ou arrêtée. Pour plus d'informations, consultez Lancement d'une instance à capacité extensible en mode Illimité ou Standard (p. 261) et Modification de la spécification de crédits d'une instance à capacité extensible (p. 264).

Vous pouvez définir `unlimited` comme formule de crédit par défaut au niveau du compte par région AWS, par famille d'instances à capacité extensible, de sorte à lancer l'ensemble des nouvelles instances à capacité extensible du compte à l'aide de la formule de crédit par défaut. Pour de plus amples informations, veuillez consulter Définition de la spécification de crédits par défaut pour le compte (p. 265).

Vous pouvez vérifier si une instance à capacité extensible est configurée en mode `unlimited` ou `standard` à l'aide de la console Amazon EC2 ou de l'AWS CLI. Pour plus d'informations, consultez Affichage de la spécification de crédits d'une instance à capacité extensible (p. 263) et Affichage de la spécification de crédits par défaut (p. 266).

Comportement des crédits lors du basculement entre Illimité et Standard

`CPUCreditBalance` est une métrique CloudWatch qui suit le nombre de crédits accumulés par une instance. `CPUSurplusCreditBalance` est une métrique CloudWatch qui suit le nombre de crédits excédentaires qu'une instance a dépensés.

Lorsque vous passez en mode `standard` une instance qui était configurée en mode `unlimited`, voici ce qui se produit :

- La valeur de `CPUCreditBalance` reste inchangée et est reportée.
- La valeur de `CPUSurplusCreditBalance` est immédiatement facturée.

Lorsqu'une instance `standard` passe à la configuration `unlimited`, la situation suivante se produit :

- La valeur de `CPUCreditBalance` contenant les crédits gagnés accumulés est reportée.
- Pour les instances T2 standard, tous les crédits de lancement sont supprimés de la valeur de `CPUCreditBalance` et la valeur de `CPUCreditBalance` restante contenant les crédits gagnés accumulés est reportée.

Surveillance de l'utilisation du crédit

Pour voir si votre instance dépense un nombre de crédits supérieur à celui correspondant au niveau de base, vous pouvez utiliser les métriques CloudWatch pour effectuer le suivi de l'utilisation et vous pouvez configurer des alarmes horaires pour être informé de l'utilisation des crédits. Pour plus d'informations, consultez Surveillance de vos crédits CPU (p. 266).

Exemples de modes illimités

Les exemples suivants expliquent l'utilisation des crédits lorsque des instances sont configurées en mode `unlimited`.

Exemples

- Exemple 1 : Explication de l'utilisation des crédits avec T3 illimité (p. 242)
- Exemple 2 : Explication de l'utilisation des crédits avec T2 illimité (p. 244)

Exemple 1 : Explication de l'utilisation des crédits avec T3 illimité

Cet exemple montre l'utilisation de l'UC d'une instance `t3.nano` lancée en mode `unlimited` et comment l'instance dépense les crédits gagnés et excédentaires pour maintenir l'utilisation de l'UC.

Une instance `t3.nano` gagne 144 crédits UC sur une période glissante de 24 heures, qu'elle peut rembourser pour 144 minutes d'utilisation de processeur vCPU. Lorsqu'elle épuise son solde de crédits UC (représenté par la métrique CloudWatch `CPUCreditBalance`), elle peut dépenser les crédits UC excédentaires – qu'elle —n'a pas encore gagnés – pour une— utilisation en mode rafale aussi longtemps que nécessaire. Comme une instance `t3.nano` gagne un maximum de 144 crédits sur une période de 24 heures, elle peut dépenser des crédits excédentaires jusqu'à ce maximum sans être facturée immédiatement. Si elle dépense plus de 144 crédits UC, la différence fait l'objet d'une facturation à la fin de l'heure.

L'exemple illustré par le graphique suivant a pour but de montrer comment une instance peut passer en mode rafale à l'aide des crédits excédentaires, même après avoir épuisé son `CPUCreditBalance`. Le flux de travail suivant référence les points numérotés sur le graphique :

P1 – À 0 heure sur le graphe, l'instance est lancée en mode `unlimited` et commence immédiatement à gagner des crédits. L'instance reste inactive après son lancement (l'utilisation de l'UC est de 0 %) et aucun crédit n'est dépensé. Tous les crédits non dépensés sont accumulés dans le solde de crédits. Pendant les

premières 24 heures, `CPUCreditUsage` est à 0 et la valeur de `CPUCreditBalance` atteint son maximum de 144.

P2 – Pendant les 12 heures suivantes, l'utilisation de l'UC est à 2,5 %, ce qui est inférieur au niveau de référence de 5 %. L'instance gagne plus de crédits qu'elle n'en dépense, mais la valeur de `CPUCreditBalance` ne peut pas dépasser son maximum de 144 crédits.

P3 – Pendant les 24 heures suivantes, l'utilisation de l'UC est de 7 % (au-dessus du niveau de référence), ce qui nécessite la dépense de 57,6 crédits. L'instance dépense plus de crédits qu'elle n'en gagne et la valeur de `CPUCreditBalance` baisse jusqu'à 86,4 crédits.

P4 – Pendant les 12 heures suivantes, l'utilisation de l'UC baisse jusqu'à 2,5 % (sous le niveau de référence), ce qui nécessite la dépense de 36 crédits. Au même moment, l'instance gagne 72 crédits. L'instance gagne plus de crédits qu'elle n'en dépense et la valeur `CPUCreditBalance` augmente jusqu'à 122 crédits.

P5 – Pendant les 5 heures suivantes, l'instance est à un pic de 100 % d'utilisation de l'UC et dépense un total de 570 crédits pour maintenir ce pic. Environ une heure après le début de cette période, l'instance épuise son solde `CPUCreditBalance` complet de 122 crédits et commence à dépenser les crédits excédentaires pour maintenir l'utilisation de l'UC élevée, totalisant 448 crédits excédentaires dans cette période (570-122=448). Lorsque la valeur de `CPUSurplusCreditBalance` atteint 144 crédits d'UC (maximum qu'une instance `t3.nano` peut gagner dans une période de 24 heures), les crédits excédentaires dépensés par la suite ne peuvent pas être compensés par les crédits gagnés. Les crédits excédentaires dépensés par la suite s'élèvent à 304 crédits (448-144=304), ce qui entraîne de faibles frais supplémentaires à la fin de l'heure pour 304 crédits.

P6 – Pendant les 13 heures suivantes, l'utilisation de l'UC est à 5 % (niveau de référence). L'instance gagne autant de crédits qu'elle en dépense, sans excès pour rembourser progressivement le solde `CPUSurplusCreditBalance`. La valeur de `CPUSurplusCreditBalance` reste à 144 crédits.

P7 – Pendant les dernières 24 heures de cet exemple, l'instance est en veille et l'utilisation de l'UC est de 0 %. Pendant ce temps, l'instance gagne 144 crédits, qu'elle utilise pour rembourser progressivement le solde `CPUSurplusCreditBalance`.

Exemple 2 : Explication de l'utilisation des crédits avec T2 illimité

Cet exemple montre l'utilisation de l'UC d'une instance `t2.nano` lancée en mode `unlimited` et comment l'instance dépense les crédits gagnés et excédentaires pour maintenir l'utilisation de l'UC.

Une instance `t2.nano` gagne 72 crédits UC sur une période glissante de 24 heures, qu'elle peut rembourser pour 72 minutes d'utilisation de processeur vCPU. Lorsqu'elle épuise son solde de crédits UC (représenté par la métrique CloudWatch `CPUCreditBalance`), elle peut dépenser les crédits UC excédentaires – qu'elle —n'a pas encore gagnés – pour une— utilisation en mode rafale aussi longtemps que nécessaire. Comme une instance `t2.nano` gagne un maximum de 72 crédits sur une période de 24 heures, elle peut dépenser des crédits excédentaires jusqu'à ce maximum sans être facturée immédiatement. Si elle dépense plus de 72 crédits UC, la différence fait l'objet d'une facturation à la fin de l'heure.

L'exemple illustré par le graphique suivant a pour but de montrer comment une instance peut passer en mode rafale à l'aide des crédits excédentaires, même après avoir épuisé son `CPUCreditBalance`. Vous pouvez supposer qu'au début de la ligne de temps du graphique, l'instance dispose d'un solde de crédits accumulés égal au nombre maximum de crédits qu'elle peut gagner en 24 heures. Le flux de travail suivant référence les points numérotés sur le graphique :

1 – Dans les 10 premières minutes, `CPUCreditUsage` est à 0, et la valeur de `CPUCreditBalance` reste à son maximum de 72.

2 – À 23h40, lorsque l'utilisation de l'UC augmente, l'instance dépense les crédits UC, et la valeur de `CPUCreditBalance` diminue.

3 – À 00 h 47, l'instance a épuisé l'intégralité de son `CPUCreditBalance` et commence à dépenser des crédits excédentaires pour maintenir l'utilisation élevée de l'UC.

4 – Les crédits excédentaires sont dépensés jusqu'à 01h55, lorsque la valeur de `CPUSurplusCreditBalance` atteint 72 crédits UC. Cela équivaut au nombre maximum de crédits qu'une instance `t2.nano` peut gagner sur une période de 24 heures. Les crédits excédentaires dépensés par la suite ne peuvent pas être compensés par les crédits gagnés au cours de la période de 24 heures, ce qui entraîne de faibles frais supplémentaires à la fin de l'heure.

5 – L'instance continue de dépenser les crédits excédentaires jusqu'à 02h20 environ. À ce moment-là, l'utilisation de l'UC chute au-dessous du niveau de base, et l'instance commence à gagner des crédits à raison de 3 crédits par heure (soit 0,25 crédit toutes les 5 minutes), qu'elle utilise pour rembourser progressivement le `CPUSurplusCreditBalance`. Une fois que la valeur de `CPUSurplusCreditBalance` est nulle, l'instance commence à accumuler les crédits gagnés dans son `CPUCreditBalance` à raison de 0,25 crédit toutes les 5 minutes.

Calcul de la facture

Les crédits excédentaires coûtent 0,05 USD par heure vCPU. L'instance a dépensé environ 25 crédits excédentaires entre 01h55 et 02h20, ce qui équivaut à 0,42 heure vCPU.

Pour cette instance, les frais supplémentaires se montent à 0,42 heure vCPU x 0,05 USD/heure vCPU = 0,021 USD, arrondi à 0,02 USD.

Facture de fin de mois correspondant à cette instance T2 illimité :

Amazon Elastic Compute Cloud running Linux/UNIX		
$0.0058 per On Demand Linux t2.nano Instance Hour	720.000 Hrs	$4.18

Amazon Elastic Compute Cloud T2 CPU Credits		
$0.05 per vCPU-Hour of T2 CPU credits	0.420 vCPU-Hours	$0.02

Vous pouvez définir des alertes de facturation pour être notifié toutes les heures des frais accumulés, puis prendre des mesures au besoin.

Mode standard pour les instances à capacité extensible

Une instance à capacité extensible configurée en mode `standard` est adaptée aux charges de travail avec une utilisation d'UC moyenne constamment inférieure à l'utilisation d'UC de référence de l'instance. Pour émettre en rafales au-dessus du niveau de base, l'instance dépense les crédits cumulés dans son solde de crédits UC. Si l'instance commence à manquer de crédits cumulés, son utilisation d'UC diminue progressivement pour atteindre le niveau d'utilisation de référence. Ainsi, l'instance ne subit pas une forte baisse des performances lorsque son solde de crédits UC est épuisé. Pour plus d'informations, consultez Crédits UC et utilisation de référence pour les instances à capacité extensible (p. 234).

Sommaire

- Concepts du mode standard (p. 246)
  - Fonctionnement des instances standard à capacité extensible (p. 246)
  - Crédits de lancement (p. 247)
  - Limites de crédits de lancement (p. 247)
  - Différences entre crédits de lancement et crédits gagnés (p. 248)
- Exemples de mode standard (p. 248)
  - Exemple 1 : Explication de l'utilisation des crédits avec une instance T3 Standard (p. 249)
  - Exemple 2 : Explication de l'utilisation des crédits avec une instance T2 standard (p. 250)
    - Période 1 : 1 – 24 heures (p. 251)
    - Période 2 : 25 – 36 heures (p. 252)
    - Période 3 : 37 – 61 heures (p. 253)
    - Période 4 : 62 – 72 heures (p. 255)
    - Période 5 : 73 – 75 heures (p. 256)
    - Période 6 : 76 – 90 heures (p. 258)
    - Période 7 : 91 – 96 heures (p. 259)

Concepts du mode standard

Le mode `standard` est une option de configuration pour les instances à capacité extensible. Il peut être activé ou désactivé à tout moment pour une instance en cours d'exécution ou arrêtée. Vous pouvez définir `standard` comme formule de crédit par défaut au niveau du compte par région AWS, par famille d'instances à capacité extensible, de sorte à lancer l'ensemble des nouvelles instances à capacité extensible du compte à l'aide de la formule de crédit par défaut.

Fonctionnement des instances standard à capacité extensible

Lorsqu'une instance à capacité extensible configurée en mode `standard` est en cours d'exécution, elle gagne continuellement (à une résolution de l'ordre de la milliseconde) un taux déterminé de crédits gagnés par heure. Lorsqu'une instance T2 Standard est arrêtée, elle perd tous ses crédits accumulés et le solde

de crédits est remis à zéro. Lorsqu'elle est redémarrée, elle reçoit un nouveau jeu de crédits de lancement, et commence à accumuler des crédits gagnés. Pour les instances T3 standard, le solde de crédits pour l'UC est conservé pendant sept jours après l'arrêt de l'instance et les crédits sont perdus ensuite. Si vous démarrez l'instance dans les sept jours, aucun crédit n'est perdu.

Les instances T2 standard reçoivent deux types de crédit d'UC : les crédits gagnés et les crédits de lancement. Lorsqu'une instance T2 Standard est en cours d'exécution, elle gagne continuellement (à une résolution de l'ordre de la milliseconde) un taux déterminé de crédits gagnés par heure. Au début, elle n'a pas de crédits gagnés pour une bonne expérience de démarrage ; elle reçoit donc, à cet effet, des crédits de lancement, qui sont dépensés pendant qu'elle accumule des crédits gagnés.

Les instances T3 standard ne reçoivent pas de crédits de lancement.

Crédits de lancement

Les instances T2 Standard obtiennent 30 crédits de lancement par processeur vCPU au lancement ou au démarrage. Par exemple, une instance `t2.micro` compte un processeur vCPU et obtient 30 crédits de lancement tandis qu'une instance `t2.xlarge` possède quatre processeurs vCPU et obtient 120 crédits de lancement. Les crédits de lancement sont conçus pour fournir une bonne expérience de démarrage et permettre aux instances de s'exécuter en mode rafale dès le lancement, avant qu'elles aient accumulé des crédits gagnés.

Les crédits de lancement sont dépensés en premier, avant les crédits gagnés. Les crédits de lancement non dépensés sont accumulés dans le solde de crédits UC, mais ne sont pas comptés dans la limite du solde de crédits UC. Par exemple, une instance `t2.micro` comporte une limite de solde de crédits UC de 144 crédits gagnés. Si elle est lancée et reste inactive pendant 24 heures, son solde de crédits UC atteint 174 (30 crédits de lancement + 144 crédits gagnés), ce qui se situe au-delà de la limite. Toutefois, une fois que l'instance a dépensé les 30 crédits de lancement, le solde de crédits ne peut pas excéder 144. Pour en savoir plus sur la limite du solde de crédits pour l'UC par rapport à chaque taille d'instance, consultez le tableau des crédits (p. 235).

Le tableau suivant répertorie l'allocation de crédits UC initiale reçus au lancement ou au démarrage, ainsi que le nombre de processeurs vCPU.

Type d'instance	Crédits de lancement	vCPU
`t1.micro`	15	1
`t2.nano`	30	1
`t2.micro`	30	1
`t2.small`	30	1
`t2.medium`	60	2
`t2.large`	60	2
`t2.xlarge`	120	4
`t2.2xlarge`	240	8

Limites de crédits de lancement

Le nombre de fois où les instances T2 Standard peuvent recevoir des crédits de lancement est limité. La limite par défaut est définie sur 100 lancements ou démarrages de toutes les instances T2 Standard combinées par compte, par région et par déploiement de 24 heures. Par exemple, la limite est atteinte lorsqu'une instance est arrêtée et démarrée 100 fois sur une période de 24 heures, ou lorsque

100 instances sont lancées sur une période de 24 heures, ou toute autre combinaison équivalente à 100 démarrages. Les nouveaux comptes peuvent présenter une limite inférieure qui augmentera au fil du temps en fonction de votre utilisation.

> Tip
>
> Pour vous assurer que vos charges de travail obtiennent toujours les performances nécessaires, passez à une instance Mode illimité pour les instances à capacité extensible (p. 237) ou utilisez une taille d'instance supérieure.

Différences entre crédits de lancement et crédits gagnés

Le tableau suivant répertorie les différences entre les crédits de lancement et les crédits gagnés.

	Crédits de lancement	Crédits gagnés
Taux d'obtention de crédits	Les instances T2 Standard obtiennent 30 crédits de lancement par processeur vCPU au lancement ou au démarrage. Si une instance T2 bascule du mode `unlimited` au mode `standard`, elle n'obtient pas de crédits de lancement au moment du basculement.	Chaque instance T2 gagne continuellement (à une résolution de l'ordre de la milliseconde) un taux déterminé de crédits UC par heure, en fonction de sa taille. Pour plus d'informations sur le nombre de crédits pour l'UC gagnés par taille d'instance, consultez le tableau des crédits (p. 235).
Limite d'obtention de crédits	La limite pour la réception de crédits de lancement est définie sur 100 lancements ou démarrages de toutes les instances T2 Standard combinées par compte, par région et par déploiement de 24 heures. Les nouveaux comptes peuvent présenter une limite inférieure qui augmentera au fil du temps en fonction de votre utilisation.	Une instance T2 ne peut pas accumuler davantage de crédits que la limite du solde de crédits UC. Si le solde de crédits UC a atteint sa limite, les crédits gagnés une fois que la limite a été atteinte sont détruits. Les crédits de lancement ne sont pas comptés dans la limite. Pour en savoir plus sur la limite du solde de crédits pour l'UC pour chaque taille d'instance T2, consultez le tableau des crédits (p. 235).
Utilisation des crédits	Les crédits de lancement sont dépensés en premier, avant les crédits gagnés.	Les crédits gagnés sont dépensés uniquement lorsque tous les crédits de lancement ont été dépensés.
Expiration des crédits	Les crédits de lancement d'une instance T2 Standard en cours d'exécution n'expirent pas. Lorsqu'une instance T2 Standard s'arrête ou passe à T2 illimité, tous les crédits de lancement sont perdus.	Lorsqu'une instance T2 est en cours d'exécution, les crédits gagnés qui ont été accumulés n'expirent pas. Lorsque l'instance T2 s'arrête, tous les crédits gagnés accumulés sont perdus.

Le suivi du nombre de crédits de lancement accumulés et de crédits gagnés accumulés est assuré par la métrique `CPUCreditBalance` CloudWatch. Pour plus d'informations, consultez `CPUCreditBalance` dans le tableau des métriques CloudWatch (p. 266).

Exemples de mode standard

Les exemples suivants expliquent l'utilisation des crédits lorsque des instances sont configurées en mode `standard`.

Exemples

- Exemple 1 : Explication de l'utilisation des crédits avec une instance T3 Standard (p. 249)
- Exemple 2 : Explication de l'utilisation des crédits avec une instance T2 standard (p. 250)

Exemple 1 : Explication de l'utilisation des crédits avec une instance T3 Standard

Cet exemple vous montre comment une instance `t3.nano` lancée en mode `standard` gagne, accumule et dépense des crédits gagnés. Vous pouvez voir que le solde de crédits reflète les crédits gagnés accumulés.

Une instance `t3.nano` en cours d'exécution gagne 144 crédits toutes les 24 heures. Sa limite de solde de crédits est de 144 crédits gagnés. Une fois que la limite est atteinte, les nouveaux crédits gagnés sont rejetés. Pour plus d'informations sur le nombre de crédits pour l'UC pouvant être gagnés et accumulés, consultez le tableau des crédits (p. 235).

Vous pouvez lancer une instance T3 Standard et l'utiliser immédiatement. Ou vous pouvez lancer une instance T3 Standard et la laisser inactive pendant quelques jours avant d'y exécuter des applications. L'utilisation ou l'inactivité d'une instance détermine si les crédits sont accumulés ou dépensés. Si une instance reste inactive pendant 24 heures après son lancement, le solde de crédits atteint sa limite, qui correspond au nombre maximal de crédits gagnés qui peuvent être accumulés.

Cet exemple décrit une instance qui reste inactive pendant 24 heures après son lancement, et explique sept périodes sur une plage de 96 heures. L'exemple illustre les taux d'obtention, d'accumulation, de dépense et de rejet de crédits, ainsi que la valeur du solde de crédits à la fin de chaque période.

Le flux de travail suivant référence les points numérotés sur le graphique :

P1 – À 0 heure sur le graphe, l'instance est lancée en mode `standard` et commence immédiatement à gagner des crédits. L'instance reste inactive après son lancement (l'utilisation de l'UC est de 0 %) et aucun crédit n'est dépensé. Tous les crédits non dépensés sont accumulés dans le solde de crédits. Pendant les premières 24 heures, `CPUCreditUsage` est à 0 et la valeur de `CPUCreditBalance` atteint son maximum de 144.

P2 – Pendant les 12 heures suivantes, l'utilisation de l'UC est à 2,5 %, ce qui est inférieur au niveau de référence de 5 %. L'instance gagne plus de crédits qu'elle n'en dépense, mais la valeur de `CPUCreditBalance` ne peut pas dépasser son maximum de 144 crédits. Tous les crédits gagnés au-delà de cette limite sont rejetés.

P3 – Pendant les 24 heures suivantes, l'utilisation de l'UC est de 7 % (au-dessus du niveau de référence), ce qui nécessite la dépense de 57,6 crédits. L'instance dépense plus de crédits qu'elle n'en gagne et la valeur de `CPUCreditBalance` baisse jusqu'à 86,4 crédits.

P4 – Pendant les 12 heures suivantes, l'utilisation de l'UC baisse jusqu'à 2,5 % (sous le niveau de référence), ce qui nécessite la dépense de 36 crédits. Au même moment, l'instance gagne 72 crédits. L'instance gagne plus de crédits qu'elle n'en dépense et la valeur `CPUCreditBalance` augmente jusqu'à 122 crédits.

P5 – Pendant les deux heures suivantes, l'instance est à un pic de 100 % d'utilisation de l'UC et épuise sa valeur de `CPUCreditBalance` complète de 122 crédits. À la fin de cette période, la valeur de `CPUCreditBalance` est nulle et l'utilisation de l'UC est obligée de baisser jusqu'au niveau d'utilisation de référence de 5 %. Au niveau de base, l'instance gagne autant de crédits qu'elle en dépense.

P6 – Pendant les 14 heures suivantes, l'utilisation de l'UC est à 5 % (niveau de référence). L'instance gagne autant de crédits qu'elle en dépense. La valeur de `CPUCreditBalance` reste à 0.

P7 – Pendant les dernières 24 heures de cet exemple, l'instance est en veille et l'utilisation de l'UC est de 0 %. Pendant ce temps, l'instance gagne 144 crédits, qu'elle accumule dans son solde `CPUCreditBalance`.

Exemple 2 : Explication de l'utilisation des crédits avec une instance T2 standard

Cet exemple vous montre comment une instance `t2.nano` lancée en tant que `standard` gagne, accumule et dépense des crédits de lancement et des crédits gagnés. Vous pouvez voir que le solde de crédits reflète non seulement les crédits gagnés accumulés, mais également les crédits de lancement accumulés.

Une instance `t2.nano` obtient 30 crédits de lancement lorsqu'elle est lancée, et gagne 72 crédits par 24 heures. Sa limite du solde de crédits est de 72 crédit gagnés ; les crédits de lancement ne sont pas comptés dans la limite. Une fois que la limite est atteinte, les nouveaux crédits gagnés sont rejetés. Pour plus d'informations sur le nombre de crédits pour l'UC pouvant être gagnés et accumulés, consultez le tableau des crédits (p. 235). Pour en savoir plus sur les limites, consultez Limites de crédits de lancement (p. 247).

Vous pouvez lancer une instance T2 Standard et l'utiliser immédiatement. Ou vous pouvez lancer une instance T2 Standard et la laisser inactive pendant quelques jours avant d'y exécuter des applications. L'utilisation ou l'inactivité d'une instance détermine si les crédits sont accumulés ou dépensés. Si une instance reste inactive pendant 24 heurs après son lancement, le solde de crédits est affiché comme dépassant sa limite, car le solde reflète à la fois les crédits gagnés accumulés et les crédits de lancement accumulés. Cependant, après l'utilisation de l'UC, les crédits de lancement sont dépensés en premier. Par la suite, la limite reflète toujours le nombre maximum de crédits gagnés pouvant être accumulés.

Cet exemple décrit une instance qui reste inactive pendant 24 heures après son lancement, et explique sept périodes sur une plage de 96 heures. L'exemple illustre les taux d'obtention, d'accumulation, de dépense et de rejet de crédits, ainsi que la valeur du solde de crédits à la fin de chaque période.

Période 1 : 1 – 24 heures

À 0 heure sur le graphe, l'instance T2 est lancée en tant que `standard` et obtient immédiatement 30 crédits de lancement. Elle gagne des crédits lorsqu'elle s'exécute. L'instance reste inactive après son — lancement – l'utilisation —de l'UC est de 0 % – et aucun crédit n'est dépensé. Tous les crédits non dépensés sont accumulés dans le solde de crédits. Environ 14 heures après le lancement, le solde de crédits est de72 (30 crédits de lancement + 42 crédits gagnés), ce qui équivaut à ce que l'instance peut gagner en 24 heures. 24 heures après le lancement, le solde de crédits dépasse 72 crédits, car les crédits de lancement non dépensés sont inclus dans le —solde de crédits. Le solde de crédits est de 102 crédits : 30 crédits de lancement + 72 crédits gagnés.

Taux de dépense de crédits	0 crédits par 24 heures (utilisation de l'UC 0 %)
Taux d'obtention de crédits	72 crédits par 24 heures
Taux de rejet de crédits	0 crédits par 24 heures

Solde de crédits	102 crédits (30 crédits de lancement + 72 crédits gagnés)

Conclusion

Si l'UC n'est pas utilisée après le lancement, l'instance accumule plus de crédits qu'elle ne peut en gagner en 24 heures (30 crédits de lancement + 72 crédits gagnés = 102 crédits).

Dans un scénario réel, une instance EC2 utilise quelques crédits pendant le lancement et l'exécution, ce qui évite que le solde atteigne la valeur théorique maximale dans cet exemple.

Période 2 : 25 – 36 heures

Pendant les 12 heures suivantes, l'instance reste encore inactive et gagne des crédits, mais le solde de crédits n'augmente pas. Il se stabilise à 102 crédits (30 crédits de lancement + 72 crédits gagnés). Le solde de crédits a atteint sa limite de 72 crédits gagnés accumulés. C'est pour cette raison que les nouveaux crédits gagnés sont rejetés.

Taux de dépense de crédits	0 crédits par 24 heures (utilisation de l'UC 0 %)
Taux d'obtention de crédits	72 crédits par 24 heures (3 crédits par heure)
Taux de rejet de crédits	72 crédits par 24 heures (100 % du taux d'obtention de crédits)
Solde de crédits	102 crédits (30 crédits de lancement + 72 crédits gagnés) — le solde est inchangé

Conclusion

Une instance gagne des crédits en permanence, mais elle ne peut pas accumuler des crédits gagnés au-delà de la limite du solde de crédits. Une fois que la limite est atteinte, les nouveaux crédits gagnés sont rejetés. Les crédits de lancement ne sont pas comptés dans la limite du solde de crédits. Si le solde comprend les crédits de lancement accumulés, il est affiché comme dépassant la limite.

Période 3 : 37 – 61 heures

Pendant les 25 heures suivantes, l'instance utilise 2 % d'UC, ce qui équivaut à 30 crédits. Pendant ce même laps de temps, elle gagne 75 crédits, mais le solde de crédits diminue. Le solde diminue car les crédits de lancement accumulés sont dépensés en premier, et les nouveaux crédits gagnés sont rejetés, car le solde de crédits a déjà atteint sa limite de 72 crédits gagnés.

Taux de dépense de crédits	28,8 crédits par 24 heures (1,2 crédits par heure, utilisation de l'UC de 2 %, 400 % du taux d'obtention de crédits) – 30— crédits sur 25 heures
Taux d'obtention de crédits	72 crédits par 24 heures
Taux de rejet de crédits	72 crédits par 24 heures (100 % du taux d'obtention de crédits)
Solde de crédits	72 crédits (30 crédits de lancement ont été dépensés ; 72 crédits gagnés n'ont pas été dépensé)

Conclusion

Une instance dépense les crédits de lancement en premier, avant les crédits gagnés. Les crédits de lancement ne sont pas comptés dans la limite de crédits. Lorsque les crédits de lancement sont dépensés,

le solde ne peut pas être plus élevé que ce qui peut être gagné en l'espace de 24 heures. De plus, lorsqu'une instance s'exécute, elle ne peut pas obtenir de nouveaux crédits de lancement.

Période 4 : 62 – 72 heures

Pendant les 11 heures suivantes, l'instance utilise 2 % d'UC, ce qui équivaut à 13.2 crédits. Cette utilisation de l'UC est identique à celle de la période précédente, mais le solde ne diminue pas. Il reste à 72 crédits.

Le solde ne diminue pas, car le taux d'obtention de crédits est supérieur à celui de dépense de crédits. Pendant que l'instance dépense 13.2 crédits, elle en gagne également 33. Cependant, la limite du solde étant de 72 crédits, les éventuels crédits gagnés au-delà de la limite sont rejetés. Le solde se stabilise à 72 crédits, et non à 102 crédits comme lors de la deuxième période, car il n'y a aucun crédit de lancement accumulé.

Taux de dépense de crédits	28,8 crédits par 24 heures (1,2 crédits par heure, utilisation de l'UC de 2 %, 400 % du taux d'obtention de crédits) — 13,2 crédits sur 11 heures)

Taux d'obtention de crédits	72 crédits par 24 heures
Taux de rejet de crédits	43.2 crédits par 24 heures (60 % du taux d'obtention de crédits)
Solde de crédits	72 crédits (0 crédit de lancement, 72 crédits gagnés) — le solde atteint sa limite

Conclusion

Une fois que les crédits de lancement sont dépensés, la limite du solde de crédits est déterminée par le nombre de crédits qu'une instance peut gagner en l'espace de 24 heures. Si l'instance gagne plus de crédits qu'elle n'en dépense, les nouveaux crédits gagnés au-delà de la limite sont rejetés.

Période 5 : 73 – 75 heures

Pendant les trois heures suivantes, l'utilisation de l'UC de l'instance passe à 20 %, ce qui équivaut à 36 crédits. L'instance gagne neuf crédits au cours de ces trois heures, ce qui entraîne une diminution du solde de 27 crédits. Au terme des trois heures, le solde de crédits est de 45 crédits gagnés accumulés.

Taux de dépense de crédits	288 crédits par 24 heures (12 crédits par heure, utilisation de l'UC de 20 %, 400 % du taux d'obtention de crédits) — 36 crédits sur 3 heures)
Taux d'obtention de crédits	72 crédits par 24 heures (9 crédits en 3 heures)
Taux de rejet de crédits	0 crédits par 24 heures
Solde de crédits	45 crédits (solde précédent (72) - crédits dépensés (36) + crédits gagnés (9)) — le solde diminue à 216 crédits par 24 heures (taux de dépense 288/24 + taux d'obtention 72/24 = taux de diminution du solde 216/24)

Conclusion

Si une instance dépense plus de crédits qu'elle n'en gagne, son solde de crédits diminue.

Période 6 : 76 – 90 heures

Pendant les 15 heures suivantes, l'instance utilise 2 % d'UC, ce qui équivaut à 18 crédits. L'utilisation est la même que celle des périodes 3 et 4. Cependant, le solde augmente au cours de cette période, alors qu'il avait diminué pendant la troisième période, et s'était stabilisé pendant la quatrième.

Pendant la troisième période, les crédits de lancement accumulés avaient été dépensés et les crédits gagnés au-delà de la limite de crédits avaient été rejetés, ce qui explique la diminution du solde de crédits. Pendant la quatrième période, l'instance avait dépensé moins de crédits qu'elle n'en avait gagné. Les crédits gagnés au-delà de la limite ont été rejetés, ce qui explique la stabilisation du solde à 72 crédits.

Au cours de cette nouvelle période, il n'y a aucun crédit de lancement accumulé, et le nombre de crédits gagnés accumulés du solde est inférieur à la limite. Aucun crédit gagné n'est rejeté. De plus, l'instance gagne plus de crédits qu'elle n'en dépense, ce qui entraîne une augmentation du solde de crédits.

Taux de dépense de crédits	28,8 crédits par 24 heures (1,2 crédits par heure, utilisation de l'UC de 2 %, 40 % du taux d'obtention de crédits) — 18 crédits sur 15 heures)

Taux d'obtention de crédits	72 crédits par 24 heures (45 crédits en 15 heures)
Taux de rejet de crédits	0 crédits par 24 heures
Solde de crédits	72 crédits (le solde augmente à un taux de 43,2 crédits par 24 heures — taux de variation = taux de dépense 28,8/24 + taux d'obtention 72/24)

Conclusion

Si une instance dépense moins de crédits qu'elle n'en gagne, son solde de crédits augmente.

Période 7 : 91 – 96 heures

Pendant les six heures suivantes, l'instance— reste inactive – l'utilisation— de l'UC est 0 % – et aucun crédit n'est dépensé. L'utilisation de l'UC est identique à celle de la deuxième période, mais le solde ne se stabilise pas à 102 crédits. Il se stabilise— à 72 crédits, soit la limite du solde de crédits de l'instance.

Au cours de la deuxième période, le solde de crédits comprenait 30 crédits de lancement accumulés. Les crédits de lancement ont été dépensés au cours de la troisième période. Une instance en cours d'exécution ne peut pas obtenir d'autres crédits de lancement. Lorsque la limite du solde de crédits est atteinte, les éventuels crédits gagnés au-delà de la limite sont rejetés.

Taux de dépense de crédits	0 crédits par 24 heures (utilisation de l'UC 0 %)
Taux d'obtention de crédits	72 crédits par 24 heures
Taux de rejet de crédits	72 crédits par 24 heures (100 % du taux d'obtention de crédits)
Solde de crédits	72 crédits (0 crédit de lancement + 72 crédits gagnés)

Conclusion

Une instance gagne des crédits en permanence, mais ne peut pas accumuler des crédits gagnés si la limite du solde de crédits est atteinte. Une fois que la limite est atteinte, les nouveaux crédits gagnés sont rejetés. La limite du solde de crédits est déterminée par le nombre de crédits qu'une instance peut gagner en l'espace de 24 heures. Pour plus d'informations sur les limites du solde de crédits, consultez le tableau des crédits (p. 235).

Utilisation des instances standard à capacité extensible

Les étapes de lancement, de surveillance et de modification de ces instances sont similaires. La différence clé est la spécification de crédits par défaut lors de leur lancement : Si vous ne modifiez pas la spécification de crédits par défaut, les valeurs par défaut sont les suivantes :

- Les instances T3 sont lancées en mode `unlimited`
- Les instances T2 dont lancées en mode `standard`

Sommaire

- Lancement d'une instance à capacité extensible en mode Illimité ou Standard (p. 261)
- Utilisation d'un groupe Auto Scaling pour lancer une instance à capacité extensible en mode Illimité (p. 262)
- Affichage de la spécification de crédits d'une instance à capacité extensible (p. 263)
- Modification de la spécification de crédits d'une instance à capacité extensible (p. 264)
- Définition de la spécification de crédits par défaut pour le compte (p. 265)
- Affichage de la spécification de crédits par défaut (p. 266)

Lancement d'une instance à capacité extensible en mode Illimité ou Standard

Vous pouvez lancer vos instances en mode `unlimited` ou `standard` à l'aide de la console Amazon EC2, d'un kit SDK AWS, d'un outil de ligne de commande ou d'un groupe Auto Scaling. Pour plus d'informations, consultez Utilisation d'un groupe Auto Scaling pour lancer une instance à capacité extensible en mode Illimité (p. 262).

Prérequis

- Vous devez lancer vos instances à l'aide d'un volume Amazon EBS comme périphérique racine. Pour plus d'informations, consultez Volume du périphérique racine Amazon EC2 (p. 21).
- Pour plus d'informations sur les exigences en matière d'AMI et de pilotes pour ces instances, consultez Notes de mise à jour (p. 231).

Pour lancer une instance à capacité extensible en mode Illimité ou Standard (console)

1. Suivez la procédure Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).
2. Sur la page Choisir un type d'instance, sélectionnez un type d'instance, puis Suivant : Configurer les détails de l'instance.
3. Choisissez une spécification de crédits.
   a. Pour lancer une instance T3 en mode `standard`, désactivez Unlimited (Illimité).
   b. Pour lancer une instance T2 en mode `unlimited`, sélectionnez Unlimited (Illimité).
4. Continuez comme indiqué par l'assistant. Lorsque vous avez terminé de vérifier vos options sur la page Examiner le lancement de l'instance, choisissez Lancer. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Pour lancer une instance à capacité extensible en mode Illimité ou Standard (AWS CLI)

Utilisez la commande run-instances pour lancer vos instances. Spécifiez la spécification de crédits à l'aide du paramètre `--credit-specification CpuCredits=`. Les spécifications de crédits valides sont `unlimited` et `standard`.

- Pour T3, si vous n'incluez pas le paramètre `--credit-specification`, l'instance est lancée en mode `unlimited` par défaut.
- Pour T2, si vous n'incluez pas le paramètre `--credit-specification`, l'instance est lancée en mode `standard` par défaut.

```
aws ec2 run-instances --image-id ami-abc12345 --count 1 --instance-type t3.micro --key-
name MyKeyPair --credit-specification "CpuCredits=unlimited"
```

Utilisation d'un groupe Auto Scaling pour lancer une instance à capacité extensible en mode Illimité

Lorsque les instances à capacité extensible sont lancées ou démarrées, elles requièrent des crédits UC pour assurer une bonne expérience d'action d'amorçage. Si vous utilisez un groupe Auto Scaling pour lancer vos instances, nous vous conseillons de configurer vos instances en mode `unlimited`. Dans ce cas, elles utilisent les crédits excédentaires en cas de lancement ou de redémarrage automatique par le groupe Auto Scaling. L'utilisation des crédits excédentaires empêche les restrictions de performances.

Création d'un modèle de lancement

Vous devez utiliser un modèle de lancement pour lancer les instances en mode `unlimited` dans un groupe Auto Scaling. Une configuration de lancement ne prend pas en charge le lancement des instances en mode `unlimited`.

Pour créer un modèle de lancement des instances en mode Illimité (console)

1. Suivez la procédure Création d'un modèle de lancement pour un groupe Auto Scaling.
2. Dans Launch template contents (Contenu du modèle de lancement), pour Instance type (Type d'instance), choisissez une taille d'instance.
3. Pour lancer des instances en mode `unlimited` dans un groupe Auto Scaling, sous Advanced details (Détails avancés), pour la Credit specification (Spécification de crédits), choisissez Unlimited (Illimité).
4. Lorsque vous avez fini de définir les paramètres de modèle de lancement, choisissez Créer un modèle de lancement. Pour plus d'informations, consultez Création d'un modèle de lancement pour un groupe Auto Scaling dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Pour créer un modèle de lancement des instances en mode Illimité (AWS CLI)

Utilisez la commande create-launch-template et spécifiez `unlimited` comme spécification de crédits.

- Pour T3, si vous n'incluez pas la valeur `CreditSpecification={CpuCredits=unlimited}`, l'instance est lancée en mode `unlimited` par défaut.
- Pour T2, si vous n'incluez pas la valeur `CreditSpecification={CpuCredits=unlimited}`, l'instance est lancée en mode `standard` par défaut.

```
aws ec2 create-launch-template --launch-template-name MyLaunchTemplate
 --version-description FirstVersion --launch-template-data
 ImageId=ami-8c1be5f6,InstanceType=t3.medium,CreditSpecification={CpuCredits=unlimited}
```

Association d'un groupe Auto Scaling avec un modèle de lancement

Pour associer le modèle de lancement à un groupe Auto Scaling, créez le groupe Auto Scaling à l'aide du modèle de lancement ou ajoutez le modèle de lancement à un groupe Auto Scaling existant.

Pour créer un groupe Auto Scaling avec un modèle de lancement (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans la barre de navigation située en haut de l'écran, sélectionnez la même région que celle utilisée lorsque vous avez créé le modèle de lancement.
3. Dans le panneau de navigation, choisissez Groupes Auto Scaling, puis Créer le groupe Auto Scaling.
4. Choisissez Modèle de lancement, sélectionnez votre modèle de lancement, puis choisissez Étape suivante.
5. Complétez les champs pour le groupe Auto Scaling. Lorsque vous avez fini de passer en revue vos paramètres de configuration sur la page Vérification, choisissez Créer le groupe Auto Scaling. Pour plus d'informations, consultez Création d'un groupe Auto Scaling à l'aide d'un modèle de lancement dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Pour créer un groupe Auto Scaling avec un modèle de lancement (AWS CLI)

Utilisez la commande create-auto-scaling-group de l'AWS CLI et spécifiez le paramètre `--launch-template`.

Pour ajouter un modèle de lancement à un groupe Auto Scaling existant (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation située en haut de l'écran, sélectionnez la même région que celle utilisée lorsque vous avez créé le modèle de lancement.
3. Dans le panneau de navigation, choisissez Groupes Auto Scaling.
4. Dans la liste des groupes Auto Scaling, sélectionnez un groupe Auto Scaling et choisissez Actions, Modifier.
5. Sous l'onglet Détails, pour Modèle de lancement, choisissez un modèle de lancement, puis choisissez Enregistrer.

Pour ajouter un modèle de lancement à un groupe Auto Scaling existant (AWS CLI)

Utilisez la commande update-auto-scaling-group de l'AWS CLI et spécifiez le paramètre `--launch-template`.

Affichage de la spécification de crédits d'une instance à capacité extensible

Vous pouvez afficher la spécification de crédits (`unlimited` ou `standard`) d'une instance en cours d'exécution ou arrêtée.

Nouvelle console

Pour afficher la spécification de crédits d'une instance à capacité extensible

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation de gauche, choisissez Instances.
3. Sélectionnez l'instance.
4. Choisissez Details (Détails) et affichez le champ Credit specification (Spécification de crédits). La valeur est `unlimited` ou `standard`.

Ancienne console

Pour afficher la spécification de crédits d'une instance à capacité extensible

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation de gauche, choisissez Instances.
3. Sélectionnez l'instance.

4. Choisissez Description et consultez le champ T2/T3 Unlimited (T2/T3 illimité).
   - Si la valeur est `Enabled`, votre instance est configurée en mode `unlimited`.
   - Si la valeur est `Disabled`, votre instance est configurée en mode `standard`.

Pour décrire la spécification de crédits d'une instance à capacité extensible (AWS CLI)

Utilisez la commande describe-instance-credit-specifications. Si vous ne spécifiez aucun ID d'instance, toutes les instances avec la spécification de crédits `unlimited` sont retournées, ainsi que les instances qui ont été précédemment configurées avec la spécification de crédits `unlimited`. Par exemple, si vous redimensionnez une instance T3 en instance M4 alors qu'elle est en mode `unlimited`, Amazon EC2 renvoie l'instance M4.

Example

```
aws ec2 describe-instance-credit-specifications --instance-id i-1234567890abcdef0
```

Voici un exemple de sortie :

```
{
  "InstanceCreditSpecifications": [
    {
        "InstanceId": "i-1234567890abcdef0",
        "CpuCredits": "unlimited"
    }
  ]
}
```

Modification de la spécification de crédits d'une instance à capacité extensible

À tout moment, vous pouvez permuter entre les spécifications de crédits `unlimited` et `standard`, pour une instance en cours d'exécution ou arrêtée.

Nouvelle console

Pour modifier la spécification de crédits d'une instance à capacité extensible

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation de gauche, choisissez Instances.
3. Sélectionnez l'instance. Pour modifier la spécification de crédits pour plusieurs instances à la fois, sélectionnez toutes les instances applicables.
4. Choisissez Actions, Instance settings (Paramètres de l'instance), Change credit specification (Modifier la spécification de crédits). Cette option n'est activée que si vous avez sélectionné une instance à capacité extensible.
5. Pour remplacer le mode de spécification de crédits par `unlimited`, activez la case à cocher en regard de l'ID de l'instance. Pour remplacer le mode de spécification de crédits par `standard`, désactivez la case à cocher en regard de l'ID de l'instance.

Ancienne console

Pour modifier la spécification de crédits d'une instance à capacité extensible

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation de gauche, choisissez Instances.

3. Sélectionnez l'instance. Pour modifier la spécification de crédits pour plusieurs instances à la fois, sélectionnez toutes les instances applicables.
4. Choisissez Actions, Paramètres de l'instance, Change T2/T3 Unlimited (Modifier T2/T3 illimité). Cette option n'est activée que si vous avez sélectionné une instance à capacité extensible.
5. La spécification de crédits actuelle apparaît entre parenthèses après l'ID de l'instance. Pour modifier la spécification de crédits en `unlimited`, choisissez Activer. Pour modifier la spécification de crédits en `standard`, choisissez Désactiver.

Pour modifier la spécification de crédits d'une instance à capacité extensible (AWS CLI)

Utilisez la commande modify-instance-credit-specification. Spécifiez l'instance et la spécification de crédits à l'aide du paramètre `--instance-credit-specification`. Les spécifications de crédits valides sont `unlimited` et `standard`.

Example

```
aws ec2 modify-instance-credit-specification --region us-east-1 --instance-credit-
specification "InstanceId=i-1234567890abcdef0,CpuCredits=unlimited"
```

Voici un exemple de sortie :

```
{
  "SuccessfulInstanceCreditSpecifications": [
    {
      "InstanceId": "i- 1234567890abcdef0"
    }
  ],
  "UnsuccessfulInstanceCreditSpecifications": []
}
```

Définition de la spécification de crédits par défaut pour le compte

Vous pouvez définir la spécification de crédits par défaut au niveau du compte pour chaque région AWS. Vous spécifiez la spécification de crédits par défaut pour chaque famille d'instances (par exemple, T2 ou T3).

Si vous utilisez l'assistant de lancement d'instance dans AWS Management Console pour lancer des instances, la valeur que vous sélectionnez pour la spécification de crédits remplace celle par défaut au niveau du compte. Si vous utilisez AWS CLI pour lancer des instances, toutes les nouvelles instances à capacité extensible du compte sont lancées à l'aide de la formule de crédit par défaut. La spécification de crédits pour les instances existantes en cours d'exécution ou arrêtées n'est pas affectée.

L'API `modify-default-credit-specification` est une opération asynchrone qui fonctionne au niveau d'une région AWS et qui modifie la formule de crédit pour chaque zone de disponibilité. Toutes les zones d'une région sont mises à jour en moins de cinq minutes. Toutefois, si des instances sont lancées au cours de cette opération, elles peuvent ne pas obtenir la nouvelle formule de crédit tant que la zone n'est pas mise à jour. Pour vérifier si la mise à jour s'est produite, vous pouvez appeler `get-default-credit-specification` et vérifier si des mises à jour sont disponibles pour la spécification de crédits par défaut. Pour de plus amples informations, veuillez consulter Affichage de la spécification de crédits par défaut (p. 266).

Considération

La spécification de crédits par défaut pour une famille d'instances ne peut être modifiée qu'une seule fois au cours d'une période continue de 5 minutes, et jusqu'à quatre fois au cours d'une période continue de 24 heures.

Pour définir la spécification de crédits par défaut au niveau du compte (AWS CLI)

Utilisez la commande modify-default-credit-specification. Spécifiez la région AWS, la famille d'instances et la spécification de crédits par défaut à l'aide du paramètre `--cpu-credits`. Les spécifications de crédits par défaut valides sont `unlimited` et `standard`.

```
aws ec2 modify-default-credit-specification --region us-east-1 --instance-family t2 --cpu-
credits unlimited
```

Affichage de la spécification de crédits par défaut

Vous pouvez afficher la spécification de crédits par défaut d'une famille d'instances à capacité extensible au niveau du compte pour chaque région AWS.

Pour afficher la spécification de crédits par défaut au niveau du compte (AWS CLI)

Utilisez la commande get-default-credit-specification. Spécifiez la région AWS et la famille d'instances.

```
aws ec2 get-default-credit-specification --region us-east-1 --instance-family t2
```

Surveillance de vos crédits CPU

Vous pouvez consulter le solde de crédits pour chaque instance dans les métriques par instance Amazon EC2 de la console CloudWatch.

Sommaire

- Métriques CloudWatch supplémentaires pour les instances à capacité extensible (p. 266)
- Calcul de l'utilisation des crédits UC (p. 268)

Métriques CloudWatch supplémentaires pour les instances à capacité extensible

Les instances à capacité extensible présentent ces métriques CloudWatch supplémentaires qui sont mises à jour toutes les cinq minutes :

- `CPUCreditUsage` – Nombre de crédits UC dépensés au cours de la période de mesure.
- `CPUCreditBalance` – Nombre de crédits UC qu'une instance a cumulés. Ce solde diminue lorsque les crédits UC sont dépensés plus rapidement qu'ils ne sont gagnés.
- `CPUSurplusCreditBalance` – Nombre de crédits UC excédentaires dépensés pour maintenir l'utilisation d'UC lorsque la métrique `CPUCreditBalance` est égale à zéro.
- `CPUSurplusCreditsCharged` – Nombre de crédits UC excédentaires qui dépassent le nombre maximal de crédits UC (p. 235) pouvant être gagnés en 24 heures, et qui génèrent donc des frais supplémentaires.

Les deux dernières métriques s'appliquent uniquement aux instances configurées en mode `unlimited`.

Le tableau suivant décrit les métriques CloudWatch correspondant aux instances à capacité extensible. Pour plus d'informations, consultez Répertorier les métriques CloudWatch disponibles pour vos instances (p. 764).

Métrique	Description
`CPUCreditUsage`	Nombre de crédits UC dépensés par l'instance pour l'utilisation de l'UC. Par exemple, un crédit UC est équivalant à un processeur virtuel fonctionnant à 100 % d'utilisation pendant une minute ou une combinaison équivalente de processeurs virtuels, d'utilisation et de temps (par exemple, un processeur virtuel fonctionnant à 50 %

Métrique	Description
	d'utilisation pendant deux minutes, ou deux processeurs virtuels fonctionnant à 25 % d'utilisation pendant deux minutes). Les métriques de crédits UC sont disponibles uniquement toutes les 5 minutes. Si vous spécifiez une période supérieure à cinq minutes, utilisez la statistique `Sum` au lieu de la statistique `Average`. Unités : crédits (minutes vCPU)
`CPUCreditBalance`	Nombre de crédits UC gagnés qu'une instance a accumulés depuis son lancement ou son démarrage. Pour les instances T2 Standard, le `CPUCreditBalance` inclut également le nombre de crédits de lancement qui ont été accumulés. Les crédits sont accumulés dans le solde de crédits quand ils sont gagnés et supprimés du solde de crédits lorsqu'ils sont dépensés. Le solde de crédits présente une limite maximum qui est déterminée par la taille de l'instance. Une fois que la limite est atteinte, tous les nouveaux crédits gagnés sont rejetés. Pour les instances T2 Standard, les crédits de lancement ne sont pas comptés dans la limite. L'instance peut dépenser les crédits figurant dans le `CPUCreditBalance` pour dépasser le niveau de base de l'utilisation de l'UC. Les crédits figurant dans le `CPUCreditBalance` d'une instance en cours d'exécution n'expirent pas. Lorsqu'une instance T3 s'arrête, la valeur `CPUCreditBalance` est conservée pendant sept jours. Au-delà, tous les crédits accumulés sont perdus. Lorsqu'une instance T2 s'arrête, la valeur de `CPUCreditBalance` n'est pas conservée, et tous les crédits accumulés sont perdus. Les métriques de crédits UC sont disponibles uniquement toutes les 5 minutes. Unités : crédits (minutes vCPU)
`CPUSurplusCreditBalance`	Nombre de crédits excédentaires ayant été dépensés par une instance `unlimited` lorsque la valeur `CPUCreditBalance` est nulle. La valeur de `CPUSurplusCreditBalance` est remboursée progressivement par les crédits UC gagnés. Si le nombre de crédits excédentaires dépasse le nombre maximum de crédits que l'instance peut gagner en 24 heures, les crédits excédentaires dépensés au-dessus du maximum génèrent des frais supplémentaires. Unités : crédits (minutes vCPU)

Métrique	Description
`CPUSurplusCreditsCharged`	Nombre de crédits excédentaires dépensés qui ne sont pas remboursés progressivement par les crédits UC gagnés et qui génèrent donc des frais supplémentaires. Les crédits excédentaires dépensés sont facturés lorsque l'une des situations suivantes se produit : • Les crédits excédentaires dépensés dépassent le nombre maximum de crédits que l'instance peut gagner sur une période de 24 heures. Les crédits excédentaires dépensés au-dessus de ce maximum sont facturés à la fin de l'heure. • L'instance est arrêtée ou résiliée. • L'instance bascule du mode `unlimited` au mode `standard`. Unités : crédits (minutes vCPU)

Calcul de l'utilisation des crédits UC

L'utilisation des crédits UC des instances est calculée à l'aide des métriques CloudWatch des instances décrites dans le tableau précédent.

Amazon EC2 envoie les métriques à CloudWatch toutes les cinq minutes. Une référence à la valeur antérieure d'une métrique à un moment donné désigne la valeur précédente de cette métrique, envoyée 5 minutes auparavant.

Calcul de l'utilisation des crédits UC pour les instances Standard

- Le solde de crédits UC augmente si l'utilisation de l'UC chute au-dessous du niveau de référence, lorsque les crédits dépensés sont inférieurs aux crédits gagnés au cours des cinq minutes précédentes.
- Le solde de crédits UC diminue si l'utilisation de l'UC est supérieure au niveau de référence, lorsque les crédits dépensés sont supérieurs aux crédits gagnés au cours des cinq minutes précédentes.

Cette description est illustrée d'un point de vue mathématique par l'équation suivante:

Example

```
CPUCreditBalance = prior CPUCreditBalance + [Credits earned per hour * (5/60) -
 CPUCreditUsage]
```

La taille de l'instance détermine le nombre de crédits que l'instance peut gagner par heure, ainsi que le nombre de crédits gagnés qu'elle peut accumuler dans le solde de crédits. Pour plus d'informations sur le nombre de crédits gagnés par heure et la limite du solde de crédits pour chaque taille d'instance, consultez le tableau des crédits (p. 235).

Exemple

Dans cet exemple, une instance `t3.nano` est utilisée. Pour calculer la valeur de `CPUCreditBalance` de l'instance, utilisez l'équation précédente comme suit :

- `CPUCreditBalance` – Solde de crédits actuel à calculer.
- `prior CPUCreditBalance` – Solde de crédits cinq minutes auparavant. Dans cet exemple, l'instance a accumulé deux crédits.
- `Credits earned per hour` – Une instance `t3.nano` gagne six crédits par heure.

- `5/60` – Représente l'intervalle de cinq minutes entre la publication des métriques CloudWatch. Multipliez les crédits gagnés par heure par 5/60 (cinq minutes) pour obtenir le nombre de crédits gagnés par l'instance au cours des cinq dernières minutes. Une instance `t3.nano` gagne 0,5 crédits toutes les cinq minutes.
- `CPUCreditUsage` – Nombre de crédits dépensés par l'instance au cours des cinq dernières minutes. Dans cet exemple, l'instance a dépensé un crédit au cours des cinq dernières minutes.

Vous pouvez calculer la valeur du `CPUCreditBalance` à l'aide de ces valeurs :

Example

```
CPUCreditBalance = 2 + [0.5 - 1] = 1.5
```

Calcul de l'utilisation des crédits UC pour les instances en mode Illimité

Lorsqu'une instance à capacité extensible doit dépasser le niveau de base, elle dépense toujours ses crédits accumulés avant de dépenser les crédits excédentaires. Si elle épuise le solde de ses crédits UC accumulés, elle peut dépenser les crédits excédentaires pour une utilisation en mode rafale de l'UC aussi longtemps que nécessaire. Si l'utilisation de l'UC chute au-dessous du niveau de base, les crédits excédentaires sont toujours remboursés avant que l'instance n'accumule des crédits gagnés.

Nous employons le terme `Adjusted balance` dans les équations suivantes pour refléter l'activité qui se produit dans cet intervalle de cinq minutes. Nous utilisons cette valeur pour obtenir les valeurs des métriques CloudWatch `CPUCreditBalance` et `CPUSurplusCreditBalance`.

Example

```
Adjusted balance = [prior CPUCreditBalance - prior CPUSurplusCreditBalance] + [Credits
 earned per hour * (5/60) - CPUCreditUsage]
```

La valeur `0` du `Adjusted balance` indique que l'instance a dépensé l'ensemble de ses crédits gagnés pour une utilisation en mode rafale et qu'aucun crédit excédentaire n'a été dépensé. Le `CPUCreditBalance` et le `CPUSurplusCreditBalance` sont donc tous deux définis sur `0`.

Une valeur positive pour le `Adjusted balance` indique que l'instance a accumulé des crédits gagnés, et que les crédits excédentaires précédents, le cas échéant, ont été remboursés. En conséquence, la valeur du `Adjusted balance` est attribuée au `CPUCreditBalance`, et le `CPUSurplusCreditBalance` est défini sur `0`. La taille de l'instance détermine le nombre maximal de crédits (p. 235) qu'elle peut accumuler.

Example

```
CPUCreditBalance = min [max earned credit balance, Adjusted balance]
CPUSurplusCreditBalance = 0
```

Une valeur négative pour le `Adjusted balance` indique que l'instance a dépensé tous les crédits gagnés qu'elle a accumulés, ainsi que des crédits excédentaires pour une utilisation en mode rafale. En conséquence, la valeur de `Adjusted balance` est attribuée à `CPUSurplusCreditBalance` et le `CPUCreditBalance` est défini sur `0`. Là encore, la taille de l'instance détermine le nombre maximal de crédits (p. 235) qu'elle peut accumuler.

Example

```
CPUSurplusCreditBalance = min [max earned credit balance, -Adjusted balance]
CPUCreditBalance = 0
```

Si les crédits excédentaires dépensés dépassent le nombre maximal de crédits que l'instance peut accumuler, le solde de crédits excédentaires est défini sur le maximum, comme le montre l'équation précédente. Les crédits excédentaires restants représentés par la métrique `CPUSurplusCreditsCharged` sont facturés.

Example

```
CPUSurplusCreditsCharged = max [-Adjusted balance - max earned credit balance, 0]
```

Pour finir, lorsque l'instance est résiliée, les crédits excédentaires suivis par le `CPUSurplusCreditBalance` sont facturés. Si l'instance bascule du mode `unlimited` au mode `standard`, tout solde `CPUSurplusCreditBalance` restant éventuel est également facturé.

Instances de calcul optimisé

Les instances optimisées pour le calcul sont particulièrement adaptées aux applications de calcul qui tirent parti de processeurs aux performances élevées.

Instances C5 et C5n

Ces instances conviennent à ce qui suit :

- Charges de travail de traitement par batch
- Transcodage multimédia
- Serveurs web hautes performances
- Calcul hautes performances (HPC)
- Modélisation scientifique
- Serveurs de jeux dédiés et moteurs de diffusion de publicités
- Inférence d'apprentissage machine et autre applications de calcul intensif

Les instances de matériel nu, comme `c5.metal`, offrent à vos applications un accès direct aux ressources physiques du serveur hôte, comme les processeurs et la mémoire.

Pour plus d'informations, consultez Instances C5 Amazon EC2.

Instances C6g et C6gd

Ces instances sont alimentées par des processeurs AWS Graviton2 et sont idéales pour exécuter des charges de travail avancées et exigeantes en calcul, telles que les suivantes :

- Calcul hautes performances (HPC)
- Traitement par lots
- Diffusion publicitaire
- Encodage vidéo
- Serveurs de jeu
- Modélisation scientifique
- Analyses distribuées
- Inférence de machine learning basée sur le processeur

Les instances de matériel nu, comme `c6g.metal`, offrent à vos applications un accès direct aux ressources physiques du serveur hôte, comme les processeurs et la mémoire.

Pour de plus amples informations, veuillez consulter Instances Amazon EC2 C6g.

Sommaire

- Spécifications matérielles (p. 271)
- Performances de l'instance (p. 273)
- Performances réseau (p. 273)
- Performances d'E/S sur SSD (p. 274)
- Fonctionnalités des instances (p. 276)
- Notes de mise à jour (p. 276)

Spécifications matérielles

Vous trouverez ci-dessous un résumé des spécifications matérielles relatives aux instances optimisées pour le calcul.

Type d'instance	vCPU par défaut	Mémoire (Gio)
`c4.large`	2	3,75
`c4.xlarge`	4	7,5
`c4.2xlarge`	8	15
`c4.4xlarge`	16	30
`c4.8xlarge`	36	60
`c5.large`	2	4
`c5.xlarge`	4	8
`c5.2xlarge`	8	16
`c5.4xlarge`	16	32
`c5.9xlarge`	36	72
`c5.12xlarge`	48	96
`c5.18xlarge`	72	144
`c5.24xlarge`	96	192
`c5.metal`	96	192
`c5a.large`	2	4
`c5a.xlarge`	4	8
`c5a.2xlarge`	8	16
`c5a.4xlarge`	16	32
`c5a.8xlarge`	32	64
`c5a.12xlarge`	48	96
`c5a.16xlarge`	64	128
`c5a.24xlarge`	96	192

Type d'instance	vCPU par défaut	Mémoire (Gio)
`c5ad.large`	2	4
`c5ad.xlarge`	4	8
`c5ad.2xlarge`	8	16
`c5ad.4xlarge`	16	32
`c5ad.8xlarge`	32	64
`c5ad.12xlarge`	48	96
`c5ad.16xlarge`	64	128
`c5ad.24xlarge`	96	192
`c5d.large`	2	4
`c5d.xlarge`	4	8
`c5d.2xlarge`	8	16
`c5d.4xlarge`	16	32
`c5d.9xlarge`	36	72
`c5d.12xlarge`	48	96
`c5d.18xlarge`	72	144
`c5d.24xlarge`	96	192
`c5d.metal`	96	192
`c5n.large`	2	5,25
`c5n.xlarge`	4	10,5
`c5n.2xlarge`	8	21
`c5n.4xlarge`	16	42
`c5n.9xlarge`	36	96
`c5n.18xlarge`	72	192
`c5n.metal`	72	192
`c6g.medium`	1	2
`c6g.large`	2	4
`c6g.xlarge`	4	8
`c6g.2xlarge`	8	16
`c6g.4xlarge`	16	32
`c6g.8xlarge`	32	64
`c6g.12xlarge`	48	96

Type d'instance	vCPU par défaut	Mémoire (Gio)
`c6g.16xlarge`	64	128
`c6g.metal`	64	128
`c6gd.medium`	1	2
`c6gd.large`	2	4
`c6gd.xlarge`	4	8
`c6gd.2xlarge`	8	16
`c6gd.4xlarge`	16	32
`c6gd.8xlarge`	32	64
`c6gd.12xlarge`	48	96
`c6gd.16xlarge`	64	128
`c6gd.metal`	64	128

Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.

Pour plus d'informations sur la spécification des options d'UC, consultez Optimisations des options d'UC (p. 673).

Performances de l'instance

Les instances optimisées EBS vous permettent d'obtenir régulièrement des performances élevées pour vos volumes EBS en éliminant les conflits entre les E/S Amazon EBS et tout autre trafic réseau de votre instance Certaines instances optimisées pour le calcul sont optimisées pour EBS par défaut sans frais supplémentaires. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Certains types d'instances optimisés pour le calcul offrent la possibilité de contrôler les états C et P du processeur sous Linux. Les états C contrôlent les niveaux de veille d'un noyau lorsqu'il est inactif, tandis que les états P contrôlent les performances attendues d'un noyau (en termes de fréquence d'UC). Pour plus d'informations, consultez Contrôle des états du processeur pour votre instance EC2 (p. 662).

Performances réseau

Vous pouvez activer la mise en réseau améliorée sur les types d'instance pris en charge pour fournir des latences plus faibles, une instabilité moindre sur le réseau et des performances de débit en paquets par seconde (PPS) plus élevées. La plupart des applications ne nécessitent pas en permanence un haut niveau de performances réseau, mais peuvent tirer profit d'un accès à une bande passante accrue lorsqu'elles envoient ou reçoivent des données. Pour plus d'informations, consultez Mise en réseau améliorée sur Linux (p. 864).

Vous trouverez ci-dessous un résumé des performances réseau relatives aux instances optimisées pour le calcul qui prennent en charge la mise en réseau améliorée.

Type d'instance	Performances réseau	Mise en réseau améliorée
`c5.4xlarge` et tailles inférieures | `c5d.4xlarge` et tailles	Jusqu'à 10 Gbit/s †	ENA (p. 865)

Type d'instance	Performances réseau	Mise en réseau améliorée
inférieures | `c6g.4xlarge` et tailles inférieures | `c6gd.4xlarge` et tailles inférieures		
`c5.9xlarge` | `c5d.9xlarge`	10 Gb/s	ENA (p. 865)
`c5.12xlarge` | `c5d.12xlarge` | `c6g.8xlarge` | `c6gd.8xlarge`	12 Gb/s	ENA (p. 865)
`c6g.12xlarge` | `c6gd.12xlarge`	20 Gb/s	ENA (p. 865)
`c5n.4xlarge` et tailles inférieures	Jusqu'à 25 Gbit/s †	ENA (p. 865)
`c5.18xlarge` | `c5.24xlarge` | `c5.metal` | `c5d.18xlarge` | `c5d.24xlarge` | `c5d.metal` | `c6g.16xlarge` | `c6g.metal` | `c6gd.16xlarge` | `c6gd.metal`	25 Gb/s	ENA (p. 865)
`c5n.9xlarge`	50 Gb/s	ENA (p. 865)
`c5n.18xlarge` | `c5n.metal`	100 Gb/s	ENA (p. 865)
`c4.large`	Modérées	Intel 82599 VF (p. 878)
`c4.xlarge` | `c4.2xlarge` | `c4.4xlarge`	Elevé	Intel 82599 VF (p. 878)
`c4.8xlarge`	10 Gb/s	Intel 82599 VF (p. 878)

† Ces instances utilisent un mécanisme de crédit d'E/S réseau pour attribuer la bande passante réseau aux instances en fonction de l'utilisation moyenne de la bande passante. Elles accumulent des crédits lorsque la bande passante est inférieure à la bande passante de référence et peuvent les utiliser quand elles effectuent les transferts de données réseau. Pour de plus amples informations, ouvrez un dossier de support et renseignez-vous sur la bande passante de référence pour les types d'instance spécifiques qui vous intéressent.

Performances d'E/S sur SSD

Si vous utilisez une AMI Linux avec un noyau de version 4.4 ou ultérieure et que vous utilisez tous les volumes de stockage d'instance basés sur SSD disponibles pour votre instance, vous pouvez obtenir les performances d'IOPS (taille de bloc de 4 096 octets) répertoriées dans le tableau suivant (lorsque la profondeur de la file d'attente est saturée). Sinon, vous obtenez des performances d'I/O inférieures.

Taille d'instance	E/S par seconde en lecture aléatoires 100 %	E/S par seconde en écriture
`c5ad.large`	16 283	7 105
`c5ad.xlarge`	32 566	14 211
`c5ad.2xlarge`	65 132	28 421

Taille d'instance	E/S par seconde en lecture aléatoires 100 %	E/S par seconde en écriture
`c5ad.4xlarge`	130 263	56 842
`c5ad.8xlarge`	260 526	113 684
`c5ad.12xlarge`	412 500	180 000
`c5ad.16xlarge`	521 053	227 368
`c5ad.24xlarge`	825 000	360 000
`c5d.large` *	20 000	9 000
`c5d.xlarge` *	40 000	18 000
`c5d.2xlarge` *	80 000	37 000
`c5d.4xlarge` *	175 000	75 000
`c5d.9xlarge`	350 000	170 000
`c5d.12xlarge`	700 000	340 000
`c5d.18xlarge`	700 000	340 000
`c5d.24xlarge`	1 400 000	680 000
`c5d.metal`	1 400 000	680 000
`c6gd.medium`	13 438	5 625
`c6gd.large`	26 875	11 250
`c6gd.xlarge`	53 750	22 500
`c6gd.2xlarge`	107 500	45 000
`c6gd.4xlarge`	215 000	90 000
`c6gd.8xlarge`	430 000	180 000
`c6gd.12xlarge`	645 000	270 000
`c6gd.16xlarge`	860 000	360 000
`c6gd.metal`	860 000	360 000

* Pour ces instances, vous pouvez obtenir la performance spécifiée.

Au fur et à mesure que vous remplissez les volumes de stockage d'instance basés sur SSD pour votre instance, le nombre d'E/S en écriture que vous pouvez obtenir diminue. Ceci est dû au travail supplémentaire que le contrôleur SSD doit effectuer pour trouver de l'espace disponible, réécrire les données existantes et effacer l'espace non utilisé pour le rendre réinscriptible. Ce processus de nettoyage de la mémoire se traduit par une amplification d'écriture interne sur le disque SSD, exprimée sous la forme du rapport des opérations d'écriture SSD sur les opérations d'écriture utilisateur. Cette diminution des performances est encore plus importante si les opérations d'écriture ne sont pas exprimées en multiples de 4 096 octets ou ne sont pas alignées sur une limite de 4 096 octets. Si vous écrivez une quantité d'octets plus faible ou des octets qui ne sont pas alignés, le contrôleur SSD doit lire les données environnantes

et stocker le résultat dans un nouvel emplacement. Ce modèle se traduit par une forte augmentation de l'amplification d'écriture, une latence accrue et une diminution considérable des performances d'E/S.

Les contrôleurs SSD peuvent utiliser plusieurs stratégies pour réduire l'impact de l'amplification d'écriture. Une telle stratégie consiste à réserver un espace dans le stockage d'instance SSD afin que le contrôleur puisse gérer efficacement l'espace disponible pour les opérations d'écriture. Cette solution est appelée sur-approvisionnement. Les volumes de stockage d'instance SSD fournis à une instance n'ont pas d'espace réservé pour le sur-approvisionnement. Pour réduire l'amplification d'écriture, nous vous conseillons de laisser 10 % du volume non partitionné que le contrôleur SSD pourra utiliser pour le sur-approvisionnement. Cela diminue le stockage que vous pouvez utiliser, mais augmente les performances même si le disque est proche de sa capacité maximale.

Pour les volumes de stockage d'instance qui prennent en charge TRIM, vous pouvez utiliser la commande TRIM pour informer le contrôleur SSD lorsque vous n'avez plus besoin des données que vous avez écrites. Cela fournit au contrôleur plus d'espace disponible, ce qui peut réduire l'amplification d'écriture et augmenter les performances. Pour plus d'informations, consultez Prise en charge de TRIM sur les volumes de stockage d'instance (p. 1278).

Fonctionnalités des instances

Voici un résumé des fonctionnalités destinées aux instances optimisées pour le calcul :

	EBS uniquement	EBS NVMe	Stockage d'instance	Groupe de placement
C4	Oui	Non	Non	Oui
C5	Oui	Oui	Non	Oui
C5a	Oui	Oui	Non	Oui
C5ad	Non	Oui	NVMe *	Oui
C5d	Non	Oui	NVMe *	Oui
C5n	Oui	Oui	Non	Oui
C6g	Oui	Oui	Non	Oui
C6gd	Non	Oui	NVMe *	Oui

* Le volume du périphérique racine doit être un volume Amazon EBS.

Pour plus d'informations, consultez les ressources suivantes :

- Amazon EBS et NVMe sur des instances Linux (p. 1209)
- Stockage d'instance Amazon EC2 (p. 1264)
- Groupes de placement (p. 924)

Notes de mise à jour

- Les instances C5 et C5d sont équipées d'un processeur Intel Xeon Platinum 8000 de 3,1 GHz de la première génération (Skylake-SP) ou de la deuxième génération (Cascade Lake).
- Les instances C5a et C5ad utilisent un processeur AMD EPYC de deuxième génération (Rome) fonctionnant à des fréquences allant jusqu'à 3,3. GHz.
- Les instances C6g et C6gd utilisent un processeur AWS Graviton2 basé sur l'architecture Arm 64 bits.

- Les instances C4 et les instances basées sur le système Nitro (p. 218) nécessitent des AMIs HVM 64 bits basées sur EBS. Elles sont dotées d'une mémoire élevée et un système d'exploitation 64 bits est nécessaire pour tirer parti de cette capacité. Les AMI HVM offrent des performances supérieures par rapport aux AMI paravirtuelles (PV) sur les types d'instance à mémoire élevée. De plus, vous devez utiliser une AMI HVM pour tirer parti de la mise en réseau améliorée.
- Les instances reposant sur le système Nitro présentent les exigences suivantes :
  - Les pilotes NVMe (p. 1209) doivent être installés.
  - Les pilotes Elastic Network Adapter (ENA) (p. 865) doivent être installés.

  Les AMI Linux suivantes répondent aux critères suivants :
  - Amazon Linux 2
  - Amazon Linux AMI 2018.03
  - Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure
  - Red Hat Enterprise Linux 7.4 ou version ultérieure
  - SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
  - CentOS 7.4.1708 ou version ultérieure
  - FreeBSD 11.1 ou version ultérieure
  - Debian GNU/Linux 9 ou version ultérieure
- Les instances dotées de processeurs AWS Graviton présentent les exigences suivantes :
  - Utiliser une AMI pour l'architecture Arm 64 bits.
  - Prendre en charge le démarrage via UEFI avec les tables ACPI, ainsi que l'insertion à chaud ACPI des périphériques PCI.

  Les AMI suivantes répondent aux critères suivants :
  - Amazon Linux 2 (Arm 64 bits)
  - Ubuntu 16.04 ou une version ultérieure (Arm 64 bits)
  - Red Hat Enterprise Linux 8.0 ou une version ultérieure (Arm 64 bits)
  - SUSE Linux Enterprise Server 15 ou version ultérieure (Arm 64 bits)
  - Debian 10 ou version ultérieure (Arm 64 bits)
- Les instances reposant sur le système Nitro prennent en charge un maximum de 28 attachements, y compris les interfaces réseau, les volumes EBS et les volumes de stockage d'instance NVMe. Pour plus d'informations, consultez Limites de volume du système Nitro (p. 1286).
- Le lancement d'une instance en matériel nu démarre le serveur sous-jacent, qui inclut la vérification de tous les composants du matériel et du microprogramme. Cela signifie que 20 minutes peuvent s'écouler entre le moment où l'instance passe à l'état d'exécution et le moment où elle devient disponible sur le réseau.
- Attacher ou détacher des volumes EBS ou des interfaces réseau secondaires à partir d'une instance en matériel nu requiert la prise en charge de l'enfichage à chaud natif de PCIe. Amazon Linux 2 et les dernières versions de l'Amazon Linux AMI prennent en charge l'enfichage à chaud natif de PCIe, ce qui n'est pas le cas des versions antérieures. Vous devez activer les options de configuration suivantes du noyau Linux :

  ```
  CONFIG_HOTPLUG_PCI_PCIE=y
  CONFIG_PCIEASPM=y
  ```
- Les instances en matériel nu utilisent un périphérique série basé sur PCI plutôt qu'un périphérique série basé sur le port d'E/S. Le noyau Linux kernel en amont et les dernières AMI Amazon Linux prennent en charge ce périphérique. Les instances en matériel nu fournissent également un tableau SPCR ACPI afin de permettre au système d'utiliser automatiquement le périphérique série basé sur PCI. Les dernières AMI Windows utilisent automatiquement le périphérique série basé sur PCI.
- acpid doit être installé sur les instances reposant sur le système Nitro afin que ces instances prennent en charge les fermetures nettes via les demandes d'API.

- Le nombre total d'instances que vous pouvez lancer dans une région est soumis à une limite, et il existe des limites supplémentaires sur certains types d'instances. Pour de plus amples informations, veuillez consulter Combien d'instances est-il possible d'exécuter dans Amazon EC2 ? dans les questions fréquentes Amazon EC2.

Instances de mémoire optimisée

Les instances à mémoire optimisée sont conçues pour garantir des performances rapides pour les charges de travail qui traitent de larges volumes de données en mémoire.

Instances R5, R5a, et R5n

Ces instances conviennent à ce qui suit :

- Bases de données relationnelles (MySQL) et NoSQL (MongoDB, Cassandra) hautes performances.
- Magasins de cache web distribués à l'échelle du web qui fournissent la mise en cache en mémoire des données de type clé-valeur (Memcached et Redis).
- Bases de données en mémoire utilisant les formats de stockage de données optimisés et l'analyse pour l'aide à la décision (par exemple, SAP HANA).
- Applications exécutant le traitement en temps réel des données non structurées volumineuses (services financiers, clusters Hadoop/Spark).
- Applications de calcul haute performance (HPC) et d'automatisation de la conception électronique (EDA).

Les instances de matériel nu, comme `r5.metal`, offrent à vos applications un accès direct aux ressources physiques du serveur hôte, comme les processeurs et la mémoire. Ces instances conviennent à ce qui suit :

- Charges de travail nécessitant un accès à des fonctions matérielles de bas niveau (Intel VT, par exemple) qui ne sont pas disponibles ou entièrement prises en charge dans les environnements virtualisés
- Applications nécessitant un environnement non virtualisé pour des questions de licence ou d'assistance

Pour plus d'informations, consultez Instances R5 Amazon EC2.

Instances R6g et R6gD

Ces instances sont alimentées par des processeurs AWS Graviton2 et sont idéales pour exécuter des charges de travail gourmandes en mémoire, telles que les suivantes :

- Bases de données open source (par exemple, MySQL, MariaDB et PostgreSQL)
- Caches en mémoire (par exemple, Memcached, Redis et KeyDB)

Les instances de matériel nu, comme `r6g.metal`, offrent à vos applications un accès direct aux ressources physiques du serveur hôte, comme les processeurs et la mémoire.

Pour de plus amples informations, veuillez consulter Instances Amazon EC2 R6g.

Instances à mémoire élevée

Les instances à mémoire élevée (`u-6tb1.metal`, `u-9tb1.metal`, `u-12tb1.metal`, `u-18tb1.metal` et `u-24tb1.metal`) offrent 6 Tio, 9 Tio, 12 Tio, 18 Tio et 24 Tio de mémoire par instance. Ces instances sont conçues pour exécuter d'importantes bases de données en mémoire, notamment des environnements de production de la base de données en mémoire SAP HANA, dans le Cloud. Elles offrent des performances bare metal avec accès direct au matériel hôte.

Pour de plus amples informations, veuillez consulter Instances à mémoire élevée Amazon EC2 et Configuration du stockage pour SAP HANA.

Instances X1

Ces instances conviennent à ce qui suit :

- Bases de données en mémoire telles que SAP HANA, y compris le support certifié SAP pour Business Suite S/4HANA, Business Suite on HANA (SoH), Business Warehouse on HANA (BW) et Data Mart Solutions on HANA. Pour plus d'informations, consultez SAP HANA sur le Cloud AWS.
- Moteurs de traitement Big Data, tels que Apache Spark ou Presto.
- Applications de calcul haute performance (HPC).

Pour plus d'informations, consultez Instances X1 Amazon EC2.

Instances X1e

Ces instances conviennent à ce qui suit :

- Bases de données hautes performances.
- Bases de données en mémoire telles que SAP HANA. Pour plus d'informations, consultez SAP HANA sur le Cloud AWS.
- Applications d'entreprise nécessitant beaucoup de mémoire.

Pour plus d'informations, consultez Instances X1e Amazon EC2.

Instances z1d

Ces instances offrent des performances de calcul et de mémoire élevées, et conviennent parfaitement à ce qui suit :

- CAO Eléctronique (EDA)
- Charges de travail de base de données relationnelle

Les instances `z1d.metal` offrent à vos applications un accès direct aux ressources physiques du serveur hôte, telles que les processeurs et la mémoire. Ces instances conviennent à ce qui suit :

- Charges de travail nécessitant un accès à des fonctions matérielles de bas niveau (Intel VT, par exemple) qui ne sont pas disponibles ou entièrement prises en charge dans les environnements virtualisés
- Applications nécessitant un environnement non virtualisé pour des questions de licence ou d'assistance

Pour plus d'informations, consultez Instances z1d Amazon EC2.

Sommaire

- Spécifications matérielles (p. 280)
- Performances de la mémoire (p. 283)
- Performances de l'instance (p. 283)
- Performances réseau (p. 283)
- Performances d'E/S sur SSD (p. 284)
- Fonctionnalités des instances (p. 286)
- Prise en charge de processeurs virtuels (p. 287)
- Notes de mise à jour (p. 288)

Spécifications matérielles

Vous trouverez ci-dessous un résumé des spécifications matérielles relatives aux instances optimisées pour la mémoire.

Type d'instance	vCPU par défaut	Mémoire (Gio)
`r4.large`	2	15,25
`r4.xlarge`	4	30,5
`r4.2xlarge`	8	61
`r4.4xlarge`	16	122
`r4.8xlarge`	32	244
`r4.16xlarge`	64	488
`r5.large`	2	16
`r5.xlarge`	4	32
`r5.2xlarge`	8	64
`r5.4xlarge`	16	128
`r5.8xlarge`	32	256
`r5.12xlarge`	48	384
`r5.16xlarge`	64	512
`r5.24xlarge`	96	768
`r5.metal`	96	768
`r5a.large`	2	16
`r5a.xlarge`	4	32
`r5a.2xlarge`	8	64
`r5a.4xlarge`	16	128
`r5a.8xlarge`	32	256
`r5a.12xlarge`	48	384
`r5a.16xlarge`	64	512
`r5a.24xlarge`	96	768
`r5ad.large`	2	16
`r5ad.xlarge`	4	32
`r5ad.2xlarge`	8	64
`r5ad.4xlarge`	16	128
`r5ad.8xlarge`	32	256

Type d'instance	vCPU par défaut	Mémoire (Gio)
`r5ad.12xlarge`	48	384
`r5ad.16xlarge`	64	512
`r5ad.24xlarge`	96	768
`r5d.large`	2	16
`r5d.xlarge`	4	32
`r5d.2xlarge`	8	64
`r5d.4xlarge`	16	128
`r5d.8xlarge`	32	256
`r5d.12xlarge`	48	384
`r5d.16xlarge`	64	512
`r5d.24xlarge`	96	768
`r5d.metal`	96	768
`r5dn.large`	2	16
`r5dn.xlarge`	4	32
`r5dn.2xlarge`	8	64
`r5dn.4xlarge`	16	128
`r5dn.8xlarge`	32	256
`r5dn.12xlarge`	48	384
`r5dn.16xlarge`	64	512
`r5dn.24xlarge`	96	768
`r5n.large`	2	16
`r5n.xlarge`	4	32
`r5n.2xlarge`	8	64
`r5n.4xlarge`	16	128
`r5n.8xlarge`	32	256
`r5n.12xlarge`	48	384
`r5n.16xlarge`	64	512
`r5n.24xlarge`	96	768
`r6g.medium`	1	8
`r6g.large`	2	16
`r6g.xlarge`	4	32

Type d'instance	vCPU par défaut	Mémoire (Gio)
`r6g.2xlarge`	8	64
`r6g.4xlarge`	16	128
`r6g.8xlarge`	32	256
`r6g.12xlarge`	48	384
`r6g.16xlarge`	64	512
`r6gd.medium`	1	8
`r6gd.large`	2	16
`r6gd.xlarge`	4	32
`r6gd.2xlarge`	8	64
`r6gd.4xlarge`	16	128
`r6gd.8xlarge`	32	256
`r6gd.12xlarge`	48	384
`r6gd.16xlarge`	64	512
`u-6tb1.metal`	448 *	6 144
`u-9tb1.metal`	448 *	9 216
`u-12tb1.metal`	448 *	12 288
`u-18tb1.metal`	448 *	18 432
`u-24tb1.metal`	448 *	24 576
`x1.16xlarge`	64	976
`x1.32xlarge`	128	1 952
`x1e.xlarge`	4	122
`x1e.2xlarge`	8	244
`x1e.4xlarge`	16	488
`x1e.8xlarge`	32	976
`x1e.16xlarge`	64	1 952
`x1e.32xlarge`	128	3 904
`z1d.large`	2	16
`z1d.xlarge`	4	32
`z1d.2xlarge`	8	64
`z1d.3xlarge`	12	96
`z1d.6xlarge`	24	192

Type d'instance	vCPU par défaut	Mémoire (Gio)
`z1d.12xlarge`	48	384
`z1d.metal`	48	384

* Chaque processeur logique est un hyperthread sur 224 cœurs.

Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.

Pour plus d'informations sur la spécification des options d'UC, consultez Optimisations des options d'UC (p. 673).

Performances de la mémoire

Les instances X1 comprennent des tampons de mémoire Intel scalablesqui fournissent 300 Gio/s de bande passante de mémoire durable en lecture et 140 Gio/s de bande passante de mémoire durable en écriture.

Pour plus d'informations sur le volume de mémoire RAM qu'il est possible d'activer pour les instances optimisées pour la mémoire, consultez Spécifications matérielles (p. 280).

Les instances à mémoire optimisée sont dotées d'une mémoire élevée et nécessitent des AMI HVM 64 bits pour tirer parti de cette capacité. Les AMI HVM offrent des performances supérieures par rapport aux AMI paravirtuelles (PV) sur les instances optimisées pour la mémoire. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108).

Performances de l'instance

Les instances à mémoire optimisée améliorent les performances de chiffrement grâce à la dernière fonction Intel AES-NI et prennent en charge les extensions de synchronisation transactionnelle (TSX) Intel pour améliorer les performances du traitement des données transactionnelles en mémoire, ainsi que les instructions de processeur Advanced Vector Extensions 2 (Intel AVX2) pour étendre la plupart des commandes entières à 256 bits.

Certaines instances à mémoire optimisée offrent la possibilité de contrôler les états C et les états P du processeur sur Linux. Les états C-states contrôlent les niveaux de veille d'un noyau lorsqu'il est inactif, tandis que les états P-states contrôlent les performances attendues d'un noyau (mesurées grâce à la fréquence d'UC). Pour plus d'informations, consultez Contrôle des états du processeur pour votre instance EC2 (p. 662).

Performances réseau

Vous pouvez activer la mise en réseau améliorée sur les types d'instance pris en charge pour fournir des latences plus faibles, une instabilité moindre sur le réseau et des performances de débit en paquets par seconde (PPS) plus élevées. La plupart des applications ne nécessitent pas en permanence un haut niveau de performances réseau, mais peuvent tirer profit d'un accès à une bande passante accrue lorsqu'elles envoient ou reçoivent des données. Pour plus d'informations, consultez Mise en réseau améliorée sur Linux (p. 864).

Vous trouverez ci-dessous un résumé des performances réseau relatives aux instances optimisées pour la mémoire qui prennent en charge la mise en réseau améliorée.

Type d'instance	Performances réseau	Mise en réseau améliorée
`r4.4xlarge` et tailles inférieures | `r5.4xlarge` et tailles inférieures | `r5a.8xlarge` et tailles	Jusqu'à 10 Gbit/s †	ENA (p. 865)

Type d'instance	Performances réseau	Mise en réseau améliorée
inférieures | `r5ad.8xlarge` et tailles inférieures | `r5d.4xlarge` et tailles inférieures | `r6g.4xlarge` et tailles inférieures | `r6gd.4xlarge` et tailles inférieures | `x1e.8large` et tailles inférieures | `z1d.3xlarge` et tailles inférieures		
`r4.8xlarge` | `r5.8xlarge` | `r5.12xlarge` | `r5a.12xlarge` | `r5ad.12xlarge` | `r5d.8xlarge` | `r5d.12xlarge` | `x1.16xlarge` | `x1e.16xlarge` | `z1d.6xlarge`	10 Gb/s	ENA (p. 865)
`r5a.16xlarge` | `r5ad.16xlarge` | `r6g.8xlarge` | `r6gd.8xlarge`	12 Gb/s	ENA (p. 865)
`r5.16xlarge` | `r5a.24xlarge` | `r5ad.24xlarge` | `r5d.16xlarge` | `r6g.12xlarge` | `r6gd.12xlarge`	20 Gb/s	ENA (p. 865)
`r5dn.4xlarge` et tailles inférieures | `r5n.4xlarge` et tailles inférieures	Jusqu'à 25 Gbit/s †	ENA (p. 865)
`r4.16xlarge` | `r5.24xlarge` | `r5.metal` | `r5d.24xlarge` | `r5d.metal` | `r5dn.8xlarge` | `r5n.8xlarge` | `r6g.16xlarge` | `r6g.metal` | `r6gd.16xlarge` | `r6gd.metal` | `x1.32xlarge` | `x1e.32xlarge` | `z1d.12xlarge` | `z1d.metal`	25 Gb/s	ENA (p. 865)
`r5dn.12xlarge` | `r5n.12xlarge`	50 Gb/s	ENA (p. 865)
`r5dn.16xlarge` | `r5n.16xlarge`	75 Gb/s	ENA (p. 865)
`r5dn.24xlarge` | `r5n.24xlarge` | `u-6tb1.metal` * | `u-9tb1.metal` * | `u-12tb1.metal` * | `u-18tb1.metal` | `u-24tb1.metal`	100 Gb/s	ENA (p. 865)

* Les instances de ce type lancées après le 12 mars 2020 fournissent des performances réseau de 100 Gbit/s. Les instances de ce type lancées avant le 12 mars 2020 ne peuvent fournir que des performances réseau de 25 Gbit/s. Pour vous assurer que les instances lancées avant le 12 mars 2020 ont une performance réseau de 100 Gbit/s, contactez votre équipe de compte pour mettre à niveau votre instance sans frais supplémentaires.

† Ces instances utilisent un mécanisme de crédit d'E/S réseau pour attribuer la bande passante réseau aux instances en fonction de l'utilisation moyenne de la bande passante. Elles accumulent des crédits lorsque la bande passante est inférieure à la bande passante de référence et peuvent les utiliser quand elles effectuent les transferts de données réseau. Pour de plus amples informations, ouvrez un dossier de support et renseignez-vous sur la bande passante de référence pour les types d'instance spécifiques qui vous intéressent.

Performances d'E/S sur SSD

Si vous utilisez une AMI Linux avec un noyau de version 4.4 ou ultérieure et que vous utilisez tous les volumes de stockage d'instance basés sur SSD disponibles pour votre instance, vous pouvez obtenir les performances d'IOPS (taille de bloc de 4 096 octets) répertoriées dans le tableau suivant (lorsque la profondeur de la file d'attente est saturée). Sinon, vous obtenez des performances d'I/O inférieures.

Taille d'instance	E/S par seconde en lecture aléatoires 100 %	E/S par seconde en écriture
`r5ad.large` *	30 000	15 000 USD
`r5ad.xlarge` *	59 000	29 000
`r5ad.2xlarge` *	117 000	57 000
`r5ad.4xlarge` *	234 000	114 000
`r5ad.8xlarge`	466 666	233 333
`r5ad.12xlarge`	700 000	340 000
`r5ad.16xlarge`	933 333	466 666
`r5ad.24xlarge`	1 400 000	680 000
`r5d.large` *	30 000	15 000 USD
`r5d.xlarge` *	59 000	29 000
`r5d.2xlarge` *	117 000	57 000
`r5d.4xlarge` *	234 000	114 000
`r5d.8xlarge`	466 666	233 333
`r5d.12xlarge`	700 000	340 000
`r5d.16xlarge`	933 333	466 666
`r5d.24xlarge`	1 400 000	680 000
`r5d.metal`	1 400 000	680 000
`r5dn.large` *	30 000	15 000 USD
`r5dn.xlarge` *	59 000	29 000
`r5dn.2xlarge` *	117 000	57 000
`r5dn.4xlarge` *	234 000	114 000
`r5dn.8xlarge`	466 666	233 333
`r5dn.12xlarge`	700 000	340 000
`r5dn.16xlarge`	933 333	466 666
`r5dn.24xlarge`	1 400 000	680 000
`r6gd.medium`	13 438	5 625
`r6gd.large`	26 875	11 250
`r6gd.xlarge`	53 750	22 500
`r6gd.2xlarge`	107 500	45 000
`r6gd.4xlarge`	215 000	90 000

Taille d'instance	E/S par seconde en lecture aléatoires 100 %	E/S par seconde en écriture
`r6gd.8xlarge`	430 000	180 000
`r6gd.12xlarge`	645 000	270 000
`r6gd.16xlarge`	860 000	360 000
`r6gd.metal`	860 000	360 000
`z1d.large` *	30 000	15 000 USD
`z1d.xlarge` *	59 000	29 000
`z1d.2xlarge` *	117 000	57 000
`z1d.3xlarge` *	175 000	75 000
`z1d.6xlarge`	350 000	170 000
`z1d.12xlarge`	700 000	340 000
`z1d.metal`	700 000	340 000

* Pour ces instances, vous pouvez obtenir la performance spécifiée.

Au fur et à mesure que vous remplissez les volumes de stockage d'instance basés sur SSD pour votre instance, le nombre d'E/S en écriture que vous pouvez obtenir diminue. Ceci est dû au travail supplémentaire que le contrôleur SSD doit effectuer pour trouver de l'espace disponible, réécrire les données existantes et effacer l'espace non utilisé pour le rendre réinscriptible. Ce processus de nettoyage de la mémoire se traduit par une amplification d'écriture interne sur le disque SSD, exprimée sous la forme du rapport des opérations d'écriture SSD sur les opérations d'écriture utilisateur. Cette diminution des performances est encore plus importante si les opérations d'écriture ne sont pas exprimées en multiples de 4 096 octets ou ne sont pas alignées sur une limite de 4 096 octets. Si vous écrivez une quantité d'octets plus faible ou des octets qui ne sont pas alignés, le contrôleur SSD doit lire les données environnantes et stocker le résultat dans un nouvel emplacement. Ce modèle se traduit par une forte augmentation de l'amplification d'écriture, une latence accrue et une diminution considérable des performances d'E/S.

Les contrôleurs SSD peuvent utiliser plusieurs stratégies pour réduire l'impact de l'amplification d'écriture. Une telle stratégie consiste à réserver un espace dans le stockage d'instance SSD afin que le contrôleur puisse gérer efficacement l'espace disponible pour les opérations d'écriture. Cette solution est appelée sur-approvisionnement. Les volumes de stockage d'instance SSD fournis à une instance n'ont pas d'espace réservé pour le sur-approvisionnement. Pour réduire l'amplification d'écriture, nous vous conseillons de laisser 10 % du volume non partitionné que le contrôleur SSD pourra utiliser pour le sur-approvisionnement. Cela diminue le stockage que vous pouvez utiliser, mais augmente les performances même si le disque est proche de sa capacité maximale.

Pour les volumes de stockage d'instance qui prennent en charge TRIM, vous pouvez utiliser la commande TRIM pour informer le contrôleur SSD lorsque vous n'avez plus besoin des données que vous avez écrites. Cela fournit au contrôleur plus d'espace disponible, ce qui peut réduire l'amplification d'écriture et augmenter les performances. Pour plus d'informations, consultez Prise en charge de TRIM sur les volumes de stockage d'instance (p. 1278).

Fonctionnalités des instances

Voici un résumé des fonctions pour les instances à mémoire optimisée :

	EBS uniquement	EBS NVMe	Stockage d'instance	Groupe de placement
R4	Oui	Non	Non	Oui
R5	Oui	Oui	Non	Oui
R5a	Oui	Oui	Non	Oui
R5ad	Non	Oui	NVME *	Oui
R5d	Non	Oui	NVME *	Oui
R5dn	Non	Oui	NVME *	Oui
R5n	Oui	Oui	Non	Oui
R6g	Oui	Oui	Non	Oui
R6gd	Non	Oui	NVMe *	Oui
`u-6tb1.metal`	Oui	Oui	Non	Non
`u-9tb1.metal`	Oui	Oui	Non	Non
`u-12tb1.metal`	Oui	Oui	Non	Non
`u-18tb1.metal`	Oui	Oui	Non	Non
`u-24tb1.metal`	Oui	Oui	Non	Non
X1	Non	Non	SSD	Oui
X1e	Non	Non	SSD *	Oui
z1d	Non	Oui	NVME *	Oui

* Le volume du périphérique racine doit être un volume Amazon EBS.

Pour plus d'informations, consultez les ressources suivantes :

- Amazon EBS et NVMe sur des instances Linux (p. 1209)
- Stockage d'instance Amazon EC2 (p. 1264)
- Groupes de placement (p. 924)

Prise en charge de processeurs virtuels

Les instances à mémoire optimisée offrent un grand nombre de processeurs virtuels, ce qui peut entraîner des problèmes de lancement avec les systèmes d'exploitation dont la limite de processeurs virtuels est inférieure. Il est fortement recommandé d'utiliser les dernières AMI lors du lancement d'instances à mémoire optimisée.

Les AMI suivantes prennent en charge le lancement des instances à mémoire optimisée :

- Amazon Linux 2 (HVM)
- Amazon Linux AMI 2016.03 (HVM) ou version ultérieure
- Ubuntu Server 14.04 LTS (HVM)
- Red Hat Enterprise Linux 7.1 (HVM)

- SUSE Linux Enterprise Server 12 SP1 (HVM)
- Windows Server 2019
- Windows Server 2016
- Windows Server 2012 R2
- Windows Server 2012
- Windows Server 2008 R2 64 bits
- Windows Server 2008 SP2 64 bits

Notes de mise à jour

- Les instances R4 comportent jusqu'à 64 vCPU et sont dotées de deux processeurs Intel Xeon AWS basés sur E5-2686v4 qui disposent d'une bande passante à mémoire élevée et de caches L3 plus étendus pour accroître les performances des applications en mémoire.
- Les instances R5, et R5d sont équipées d'un processeur Intel Xeon Platinum 8000 de 3,1 GHz de la première génération (Skylake-SP) ou de la deuxième génération (Cascade Lake).
- Les instances R5a et R5ad utilisent un processeur AMD EPYC 7000 à 2,5 GHz.
- Les instances R6g et R6gd disposent d'un processeur AWS Graviton2 basé sur l'architecture Arm 64 bits.
- Les instances à mémoire élevée (`u-6tb1.metal`, `u-9tb1.metal` et `u-12tb1.metal`) sont les premières instances alimentées par une plateforme à 8 sockets avec les processeurs Intel Xeon Platinum 8176M (Skylake) de dernière génération optimisés pour des charges de travail critiques de l'entreprise. Les instances à mémoire élevée avec 18 To et 24 To de mémoire (`u-18tb1.metal` et `u-24tb1.metal`) sont les premières instances alimentées par une plate-forme à 8 sockets avec des processeurs Intel Xeon 8280L (Cascade Lake) évolutifs de 2e génération.
- Les instances X1 et X1e comportent jusqu'à 128 vCPU et sont dotées de quatre processeurs Intel Xeon E7-8880 v3 qui disposent d'une bande passante à mémoire élevée et de caches L3 plus étendus pour accroître les performances des applications en mémoire.
- Les instances reposant sur le système Nitro présentent les exigences suivantes :
  - Les pilotes NVMe (p. 1209) doivent être installés.
  - Les pilotes Elastic Network Adapter (ENA) (p. 865) doivent être installés.

  Les AMI Linux suivantes répondent aux critères suivants :
  - Amazon Linux 2
  - Amazon Linux AMI 2018.03
  - Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure
  - Red Hat Enterprise Linux 7.4 ou version ultérieure
  - SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
  - CentOS 7.4.1708 ou version ultérieure
  - FreeBSD 11.1 ou version ultérieure
  - Debian GNU/Linux 9 ou version ultérieure
- Les instances dotées de processeurs AWS Graviton présentent les exigences suivantes :
  - Utiliser une AMI pour l'architecture Arm 64 bits.
  - Prendre en charge le démarrage via UEFI avec les tables ACPI, ainsi que l'insertion à chaud ACPI des périphériques PCI.

  Les AMI suivantes répondent aux critères suivants :
  - Amazon Linux 2 (Arm 64 bits)
  - Ubuntu 16.04 ou une version ultérieure (Arm 64 bits)
  - Red Hat Enterprise Linux 8.0 ou une version ultérieure (Arm 64 bits)

  - SUSE Linux Enterprise Server 15 ou version ultérieure (Arm 64 bits)
  - Debian 10 ou version ultérieure (Arm 64 bits)
- Les instances reposant sur le système Nitro prennent en charge un maximum de 28 attachements, y compris les interfaces réseau, les volumes EBS et les volumes de stockage d'instance NVMe. Pour plus d'informations, consultez Limites de volume du système Nitro (p. 1286).
- Le lancement d'une instance en matériel nu démarre le serveur sous-jacent, qui inclut la vérification de tous les composants du matériel et du microprogramme. Cela signifie que 20 minutes peuvent s'écouler entre le moment où l'instance passe à l'état d'exécution et le moment où elle devient disponible sur le réseau.
- Attacher ou détacher des volumes EBS ou des interfaces réseau secondaires à partir d'une instance en matériel nu requiert la prise en charge de l'enfichage à chaud natif de PCIe. Amazon Linux 2 et les dernières versions de l'Amazon Linux AMI prennent en charge l'enfichage à chaud natif de PCIe, ce qui n'est pas le cas des versions antérieures. Vous devez activer les options de configuration suivantes du noyau Linux :

```
CONFIG_HOTPLUG_PCI_PCIE=y
CONFIG_PCIEASPM=y
```

- Les instances en matériel nu utilisent un périphérique série basé sur PCI plutôt qu'un périphérique série basé sur le port d'E/S. Le noyau Linux kernel en amont et les dernières AMI Amazon Linux prennent en charge ce périphérique. Les instances en matériel nu fournissent également un tableau SPCR ACPI afin de permettre au système d'utiliser automatiquement le périphérique série basé sur PCI. Les dernières AMI Windows utilisent automatiquement le périphérique série basé sur PCI.
- Vous ne pouvez pas lancer d'instances X1 à l'aide d'une AMI Windows Server 2008 SP2 64 bits, sauf pour les instances `x1.16xlarge`.
- Vous ne pouvez pas lancer d'instances X1e à l'aide d'une AMI Windows Server 2008 SP2 64 bits.
- Vous ne pouvez pas lancer d'instances `r4.large` et `r4.4xlarge` à l'aide des versions précédentes de l'AMI Windows Server 2008 R2 64 bits. Si vous rencontrez ce problème, vous devez mettre à jour vers la dernière version de cette AMI.
- Le nombre total d'instances que vous pouvez lancer dans une région est soumis à une limite, et il existe des limites supplémentaires sur certains types d'instances. Pour de plus amples informations, veuillez consulter Combien d'instances est-il possible d'exécuter dans Amazon EC2 ? dans les questions fréquentes Amazon EC2.

Instances de stockage optimisé

Les instances optimisées pour le stockage sont destinées à des charges de travail qui nécessitent un accès séquentiel en lecture et écriture important, pour des ensembles de données très volumineux stockés localement. Elles sont optimisées de façon à fournir des dizaines de milliers d'opérations d'E/S aléatoires à faible latence par seconde (IOPS) aux applications.

Instances D2

Ces instances conviennent à ce qui suit :

- Entrepôt de données de traitement massivement parallèle (MPP)
- Informatique distribuée MapReduce et Hadoop
- Applications de consignation ou de traitement des données

Instances H1

Ces instances conviennent à ce qui suit :

- Charges de travail gourmandes en données comme MapReduce et les systèmes de fichiers distribués
- Applications nécessitant un accès séquentiel à d'importants volumes de données sur un stockage d'instance en attachement direct
- Applications nécessitant un accès à débit élevé aux grandes quantités de données

Instances I3 et I3en

Ces instances conviennent à ce qui suit :

- Systèmes de traitement de transaction en ligne (OLTP) à fréquence élevée
- Bases de données relationnelles
- Bases de données NoSQL
- Cache pour les bases de données en mémoire (Redis par exemple)
- Applications d'entreposage de données
- Systèmes de fichiers distribués

Les instances de matériel nu offrent à vos applications un accès direct aux ressources physiques du serveur hôte, telles que les processeurs et la mémoire. Ces instances conviennent à ce qui suit :

- Charges de travail nécessitant un accès à des fonctions matérielles de bas niveau (Intel VT, par exemple) qui ne sont pas disponibles ou entièrement prises en charge dans les environnements virtualisés
- Applications nécessitant un environnement non virtualisé pour des questions de licence ou d'assistance

Pour plus d'informations, consultez Instances I3 Amazon EC2.

Sommaire

- Spécifications matérielles (p. 290)
- Performances de l'instance (p. 291)
- Performances réseau (p. 292)
- Performances d'E/S sur SSD (p. 293)
- Fonctionnalités des instances (p. 294)
- Prise en charge de processeurs virtuels (p. 294)
- Notes de mise à jour (p. 295)

Spécifications matérielles

Le stockage de données principal des instances D2 est constitué de volumes de stockage d'instance HDD. Le stockage de données principal des instances I3 et I3en est constitué de volumes de stockage d'instance SSD NVMe (Non-Volatile Memory Express).

Les volumes de stockage d'instances ne persistent que pendant la durée de vie de l'instance. Quand vous arrêtez, mettez en veille prolongée ou résiliez une instance, les applications et les données figurant dans ses volumes de stockage d'instance sont effacées. Nous vous recommandons de sauvegarder ou de répliquer régulièrement les données importantes dans vos volumes de stockage d'instances. Pour plus d'informations, consultez Stockage d'instance Amazon EC2 (p. 1264) et Volumes de stockage d'instance SSD (p. 1276).

Vous trouverez ci-dessous un résumé des spécifications matérielles relatives aux instances optimisées pour le stockage.

Type d'instance	vCPU par défaut	Mémoire (Gio)
`d2.xlarge`	4	30,5
`d2.2xlarge`	8	61
`d2.4xlarge`	16	122
`d2.8xlarge`	36	244
`h1.2xlarge`	8	32
`h1.4xlarge`	16	64
`h1.8xlarge`	32	128
`h1.16xlarge`	64	256
`i3.large`	2	15,25
`i3.xlarge`	4	30,5
`i3.2xlarge`	8	61
`i3.4xlarge`	16	122
`i3.8xlarge`	32	244
`i3.16xlarge`	64	488
`i3.metal`	72	512
`i3en.large`	2	16
`i3en.xlarge`	4	32
`i3en.2xlarge`	8	64
`i3en.3xlarge`	12	96
`i3en.6xlarge`	24	192
`i3en.12xlarge`	48	384
`i3en.24xlarge`	96	768
`i3en.metal`	96	768

Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.

Pour plus d'informations sur la spécification des options d'UC, consultez Optimisations des options d'UC (p. 673).

Performances de l'instance

Pour garantir les meilleures performances de débit de disque à partir de votre instance sous Linux, nous vous recommandons d'utiliser la version la plus récente d'Amazon Linux 2 ou de l'Amazon Linux AMI.

Pour les instances avec des volumes de stockage d'instance NVMe, vous devez utiliser une AMI Linux avec la version du noyau 4.4 ou ultérieure. Sinon, votre instance n'obtiendra pas les performances IOPS maximales.

Les instances D2 offrent les meilleures performances de disque lorsque vous utilisez un noyau Linux qui prend en charge les accords persistants, une extension du protocole d'anneau par bloc Xen qui améliore considérablement le débit et l'évolutivité du disque. Pour plus d'informations sur les accords persistants, consultez cet article sur le blog du projet Xen.

Les instances optimisées EBS vous permettent d'obtenir régulièrement des performances élevées pour vos volumes EBS en éliminant les conflits entre les E/S Amazon EBS et tout autre trafic réseau de votre instance Certaines instances optimisées pour le stockage sont optimisées pour EBS par défaut sans frais supplémentaires. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Certains types d'instances optimisés pour le stockage offrent la possibilité de contrôler les états C et P du processeur sous Linux. Les états C contrôlent les niveaux de veille d'un noyau lorsqu'il est inactif, tandis que les états P contrôlent les performances attendues d'un noyau (en termes de fréquence d'UC). Pour plus d'informations, consultez Contrôle des états du processeur pour votre instance EC2 (p. 662).

Performances réseau

Vous pouvez activer la mise en réseau améliorée sur les types d'instance pris en charge pour fournir des latences plus faibles, une instabilité moindre sur le réseau et des performances de débit en paquets par seconde (PPS) plus élevées. La plupart des applications ne nécessitent pas en permanence un haut niveau de performances réseau, mais peuvent tirer profit d'un accès à une bande passante accrue lorsqu'elles envoient ou reçoivent des données. Pour plus d'informations, consultez Mise en réseau améliorée sur Linux (p. 864).

Vous trouverez ci-dessous un résumé des performances réseau relatives aux instances optimisées pour le stockage qui prennent en charge la mise en réseau améliorée.

Type d'instance	Performances réseau	Mise en réseau améliorée
`i3.4xlarge` et tailles inférieures	Jusqu'à 10 Gbit/s †	ENA (p. 865)
`i3.8xlarge` | `h1.8xlarge`	10 Gb/s	ENA (p. 865)
`i3en.3xlarge` et plus petits	Jusqu'à 25 Gbit/s †	ENA (p. 865)
`i3.16xlarge` | `i3.metal` | `i3en.6xlarge` | `h1.16xlarge`	25 Gb/s	ENA (p. 865)
`i3en.12xlarge`	50 Gb/s	ENA (p. 865)
`i3en.24xlarge` | `i3en.metal`	100 Gb/s	ENA (p. 865)
`d2.xlarge`	Modérées	Intel 82599 VF (p. 878)
`d2.2xlarge` | `d2.4xlarge`	Élevé	Intel 82599 VF (p. 878)
`d2.8xlarge`	10 Gb/s	Intel 82599 VF (p. 878)

† Ces instances utilisent un mécanisme de crédit d'E/S réseau pour attribuer la bande passante réseau aux instances en fonction de l'utilisation moyenne de la bande passante. Elles accumulent des crédits lorsque la bande passante est inférieure à la bande passante de référence et peuvent les utiliser quand elles effectuent les transferts de données réseau. Pour de plus amples informations, ouvrez un dossier de

support et renseignez-vous sur la bande passante de référence pour les types d'instance spécifiques qui vous intéressent.

Performances d'E/S sur SSD

Si vous utilisez une AMI Linux avec un noyau de version 4.4 ou ultérieure et que vous utilisez tous les volumes de stockage d'instance basés sur SSD disponibles pour votre instance, vous pouvez obtenir les performances d'IOPS (taille de bloc de 4 096 octets) répertoriées dans le tableau suivant (lorsque la profondeur de la file d'attente est saturée). Sinon, vous obtenez des performances d'I/O inférieures.

Taille d'instance	E/S par seconde en lecture aléatoires 100 %	E/S par seconde en écriture
`i3.large` *	100 125	35 000
`i3.xlarge` *	206 250	70 000
`i3.2xlarge`	412 500	180 000
`i3.4xlarge`	825 000	360 000
`i3.8xlarge`	1,65 million	720 000
`i3.16xlarge`	3.3 millions	1,4 million
`i3.metal`	3.3 millions	1,4 million
`i3en.large` *	42 500	32 500
`i3en.xlarge` *	85 000	65 000
`i3en.2xlarge` *	170 000	130 000
`i3en.3xlarge`	250 000	200 000
`i3en.6xlarge`	500 000	400 000
`i3en.12xlarge`	1 million	800 000
`i3en.24xlarge`	2 millions	1,6 million
`i3en.metal`	2 millions	1,6 million

* Pour ces instances, vous pouvez obtenir la performance spécifiée.

Au fur et à mesure que vous remplissez les volumes de stockage d'instance basés sur SSD, les performances d'I/O que vous obtenez diminuent. Ceci est dû au travail supplémentaire que le contrôleur SSD doit effectuer pour trouver de l'espace disponible, réécrire les données existantes et effacer l'espace non utilisé pour le rendre réinscriptible. Ce processus de nettoyage de la mémoire se traduit par une amplification d'écriture interne sur le disque SSD, exprimée sous la forme du rapport des opérations d'écriture SSD sur les opérations d'écriture utilisateur. Cette diminution des performances est encore plus importante si les opérations d'écriture ne sont pas exprimées en multiples de 4 096 octets ou ne sont pas alignées sur une limite de 4 096 octets. Si vous écrivez une quantité d'octets plus faible ou des octets qui ne sont pas alignés, le contrôleur SSD doit lire les données environnantes et stocker le résultat dans un nouvel emplacement. Ce modèle se traduit par une forte augmentation de l'amplification d'écriture, une latence accrue et une diminution considérable des performances d'E/S.

Les contrôleurs SSD peuvent utiliser plusieurs stratégies pour réduire l'impact de l'amplification d'écriture. Une telle stratégie consiste à réserver un espace dans le stockage d'instance SSD afin que le contrôleur

puisse gérer efficacement l'espace disponible pour les opérations d'écriture. Cette solution est appelée sur-approvisionnement. Les volumes de stockage d'instance SSD fournis à une instance n'ont pas d'espace réservé pour le sur-approvisionnement. Pour réduire l'amplification d'écriture, nous vous conseillons de laisser 10 % du volume non partitionné que le contrôleur SSD pourra utiliser pour le sur-approvisionnement. Cela diminue le stockage que vous pouvez utiliser, mais augmente les performances même si le disque est proche de sa capacité maximale.

Pour les volumes de stockage d'instance qui prennent en charge TRIM, vous pouvez utiliser la commande TRIM pour informer le contrôleur SSD lorsque vous n'avez plus besoin des données que vous avez écrites. Cela fournit au contrôleur plus d'espace disponible, ce qui peut réduire l'amplification d'écriture et augmenter les performances. Pour plus d'informations, consultez Prise en charge de TRIM sur les volumes de stockage d'instance (p. 1278).

Fonctionnalités des instances

Voici un résumé des fonctionnalités destinées aux instances optimisées pour le stockage :

	EBS uniquement	Stockage d'instance	Groupe de placement
D2	Non	HDD	Oui
H1	Non	HDD *	Oui
I3	Non	NVMe *	Oui
I3en	Non	NVMe *	Oui

* Le volume du périphérique racine doit être un volume Amazon EBS.

Pour plus d'informations, consultez les ressources suivantes :

- Amazon EBS et NVMe sur des instances Linux (p. 1209)
- Stockage d'instance Amazon EC2 (p. 1264)
- Groupes de placement (p. 924)

Prise en charge de processeurs virtuels

Le type d'instance `d2.8xlarge` fournit 36 processeurs virtuels, ce qui peut provoquer des problèmes de lancement sur certains systèmes d'exploitation Linux dont la limite de processeurs virtuels est fixée à 32. Il est fortement recommandé d'utiliser les dernières AMI lors du lancement d'instances `d2.8xlarge`.

Les AMI Linux suivantes prennent en charge le lancement d'instances `d2.8xlarge` avec 36 processeurs virtuels :

- Amazon Linux 2 (HVM)
- Amazon Linux AMI 2018.03 (HVM)
- Ubuntu Server 14.04 LTS (HVM) ou une version ultérieure
- Red Hat Enterprise Linux 7.1 (HVM)
- SUSE Linux Enterprise Server 12 (HVM)

Si vous devez utiliser une autre AMI pour votre application et que votre lancement d'instance `d2.8xlarge` ne se termine pas avec succès (par exemple, si le statut de votre instance se change en `stopped` pendant le lancement avec un motif de transition d'état `Client.InstanceInitiatedShutdown`), modifiez votre

instance comme décrit dans la procédure suivante pour prendre en charge plus de 32 processeurs virtuels de telle sorte que vous puissiez utiliser le type d'instance `d2.8xlarge`.

Pour mettre à jour une instance afin de prendre en charge plus de 32 vCPU

1. Lancez une instance D2 à l'aide de votre AMI, en choisissant un type d'instance D2 autre que `d2.8xlarge`.
2. Mettez à jour le noyau à l'aide de la dernière version en suivant les instructions propres à votre système d'exploitation. Par exemple, pour RHEL 6, utilisez la commande suivante :

   ```
   sudo yum update -y kernel
   ```

3. Arrêtez l'instance.
4. (Facultatif) Créez une AMI à partir de l'instance que vous utilisez pour lancer toute instance `d2.8xlarge` supplémentaire dont vous aurez besoin à l'avenir.
5. Remplacez le type d'instance de votre instance arrêtée par `d2.8xlarge` (choisissez successivement Actions, Paramètres de l'instance et Changer le type d'instance, puis suivez les instructions).
6. Démarrez l'instance. Si l'instance est lancée correctement, vous avez terminé. Si l'instance ne démarre toujours pas correctement, passez à l'étape suivante.
7. (Facultatif) Si l'instance ne démarre toujours pas correctement, le noyau de votre instance ne pourra peut-être pas prendre en charge plus de 32 vCPU. Toutefois, vous pourrez probablement redémarrer l'instance si vous limitez le nombre de vCPU.

   a. Remplacez le type d'instance de votre instance arrêtée par n'importe quel type d'instance D2 autre que `d2.8xlarge` (choisissez successivement Actions, Paramètres de l'instance et Changer le type d'instance, puis suivez les instructions).

   b. Ajoutez l'option `maxcpus=32` aux paramètres de votre noyau de démarrage en suivant les instructions propres à votre système d'exploitation. Par exemple, pour RHEL 6, modifiez le fichier `/boot/grub/menu.lst` et ajoutez l'option suivante à l'entrée `kernel` active la plus récente :

   ```
   default=0
   timeout=1
   splashimage=(hd0,0)/boot/grub/splash.xpm.gz
   hiddenmenu
   title Red Hat Enterprise Linux Server (2.6.32-504.3.3.el6.x86_64)
   root (hd0,0)
   kernel /boot/vmlinuz-2.6.32-504.3.3.el6.x86_64 maxcpus=32 console=ttyS0 ro
    root=UUID=9996863e-b964-47d3-a33b-3920974fdbd9 rd_NO_LUKS  KEYBOARDTYPE=pc
    KEYTABLE=us LANG=en_US.UTF-8 xen_blkfront.sda_is_xvda=1 console=ttyS0,115200n8
    console=tty0 rd_NO_MD SYSFONT=latarcyrheb-sun16 crashkernel=auto rd_NO_LVM
    rd_NO_DM
   initrd /boot/initramfs-2.6.32-504.3.3.el6.x86_64.img
   ```

   c. Arrêtez l'instance.

   d. (Facultatif) Créez une AMI à partir de l'instance que vous utilisez pour lancer toute instance `d2.8xlarge` supplémentaire dont vous aurez besoin à l'avenir.

   e. Remplacez le type d'instance de votre instance arrêtée par `d2.8xlarge` (choisissez successivement Actions, Paramètres de l'instance et Changer le type d'instance, puis suivez les instructions).

   f. Démarrez l'instance.

Notes de mise à jour

- Vous devez lancer les instances optimisées pour le stockage à l'aide d'une AMI HVM. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108).

- Les instances reposant sur le système Nitro (p. 218) présentent les exigences suivantes :
  - Les pilotes NVMe (p. 1209) doivent être installés.
  - Les pilotes Elastic Network Adapter (ENA) (p. 865) doivent être installés.

  Les AMI Linux suivantes répondent aux critères suivants :
  - Amazon Linux 2
  - Amazon Linux AMI 2018.03
  - Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure
  - Red Hat Enterprise Linux 7.4 ou version ultérieure
  - SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
  - CentOS 7.4.1708 ou version ultérieure
  - FreeBSD 11.1 ou version ultérieure
  - Debian GNU/Linux 9 ou version ultérieure
- Le lancement d'une instance en matériel nu démarre le serveur sous-jacent, qui inclut la vérification de tous les composants du matériel et du microprogramme. Cela signifie que 20 minutes peuvent s'écouler entre le moment où l'instance passe à l'état d'exécution et le moment où elle devient disponible sur le réseau.
- Attacher ou détacher des volumes EBS ou des interfaces réseau secondaires à partir d'une instance en matériel nu requiert la prise en charge de l'enfichage à chaud natif de PCIe. Amazon Linux 2 et les dernières versions de l'Amazon Linux AMI prennent en charge l'enfichage à chaud natif de PCIe, ce qui n'est pas le cas des versions antérieures. Vous devez activer les options de configuration suivantes du noyau Linux :

```
CONFIG_HOTPLUG_PCI_PCIE=y
CONFIG_PCIEASPM=y
```

- Les instances en matériel nu utilisent un périphérique série basé sur PCI plutôt qu'un périphérique série basé sur le port d'E/S. Le noyau Linux kernel en amont et les dernières AMI Amazon Linux prennent en charge ce périphérique. Les instances en matériel nu fournissent également un tableau SPCR ACPI afin de permettre au système d'utiliser automatiquement le périphérique série basé sur PCI. Les dernières AMI Windows utilisent automatiquement le périphérique série basé sur PCI.
- Avec les AMI FreeBSD, le démarrage des instances nues demande près d'une heure et les E/S vers le stockage NVMe local ne se terminent pas. Pour contourner ce problème, ajoutez la ligne suivante à `/boot/loader.conf` et redémarrez :

```
hw.nvme.per_cpu_io_queues="0"
```

- Le type d'instance `d2.8xlarge` fournit 36 processeurs virtuels, ce qui peut provoquer des problèmes de lancement sur certains systèmes d'exploitation Linux dont la limite de processeurs virtuels est fixée à 32. Pour plus d'informations, consultez Prise en charge de processeurs virtuels (p. 294).
- Le nombre total d'instances que vous pouvez lancer dans une région est soumis à une limite, et il existe des limites supplémentaires sur certains types d'instances. Pour de plus amples informations, veuillez consulter Combien d'instances est-il possible d'exécuter dans Amazon EC2 ? dans les questions fréquentes Amazon EC2.

Instances à calcul accéléré Linux

Les instances à calcul accéléré utilisent des accélérateurs matériels ou des coprocesseurs pour exécuter certaines fonctions, telles que le calcul en virgule flottante, le traitement graphique ou la correspondance de modèles de données, plus efficacement qu'il n'est possible avec les logiciels exécutés sur processeurs. Ces instances offrent de meilleures capacités de traitement en parallèle afin d'accélérer les charges de travail qui nécessitent beaucoup de ressources système.

Si vous avez besoin de capacités de traitement parallèle élevées, utilisez des instances à calcul accéléré qui vous donneront accès aux accélérateurs matériels tels que les unités GPU (Graphics Processing Units), FPGA (Field Programmable Gate Arrays) ou AWS Inferentia.

Sommaire

- Instances GPU (p. 297)
- Instances avec AWS Inferentia (p. 298)
- Instances FPGA (p. 299)
- Spécifications matérielles (p. 299)
- Performances de l'instance (p. 300)
- Performances réseau (p. 301)
- Fonctionnalités des instances (p. 301)
- Notes de mise à jour (p. 302)
- Installation des pilotes NVIDIA sur des instances Linux (p. 303)
- Activer les applications virtuelles NVIDIA GRID (p. 310)
- Optimisation des paramètres GPU (p. 311)

Instances GPU

Les instances GPU donnent accès aux unités GPU NVIDIA avec des milliers de cœurs de calcul. Vous pouvez utiliser ces instances pour accélérer de nombreuses applications scientifiques, d'ingénierie et de rendu en tirant parti de l'architecture CUDA ou d'infrastructures de calcul parallèle OpenCL (Open Computing Language). Vous pouvez également les utiliser pour des applications graphiques, notamment les jeux en streaming, les applications 3D en streaming, et d'autres charges de travail graphiques.

Instances G4

Les instances G4 utilisent les GPU NVIDIA Tesla et fournissent une plateforme hautes performances et économique pour les calculs génériques utilisant les infrastructures CUDA ou de machine learning, ainsi que des applications graphiques utilisant DirectX ou OpenGL. Les instances G4 offrent une mise en réseau avec bande passante élevée, des capacités en virgule flottante avec précision unique, ainsi que des précisions INT8 et INT4. Chaque GPU dispose de 16 Gio de mémoire GDDR6, ce qui permet aux instances G4 d'être adaptées aux inférences de machine learning, au transcodage vidéo et aux applications graphiques telles que les postes de travail graphiques à distance et le streaming de jeux dans le cloud.

Pour de plus amples informations, veuillez consulter Instances G4 Amazon EC2.

Les instances G4 prennent en charge le poste de travail virtuel NVIDIA GRID. Pour de plus amples informations, veuillez consulter les offres NVIDIA sur Marketplace.

Instances G3

Les instances G3 utilisent des GPU NVIDIA Tesla M60 et offrent une plateforme économique à hautes performances pour les applications graphiques qui utilisent DirectX ou OpenGL. Les instances G3 fournissent également des fonctions de station de travail virtuelle NVIDIA GRID, qui prennent en charge 4 écrans avec des résolutions pouvant atteindre 4096x2160 et des applications virtuelles NVIDIA GRID. Les instances G3 sont bien adaptées aux applications, telles que les visualisations 3D, les stations de travail distantes gourmandes en graphiques, le rendu 3D, l'encodage vidéo, la réalité virtuelle et autres charges de travail graphiques côté serveur nécessitant une puissance de traitement massivement parallèle.

Pour de plus amples informations, veuillez consulter Instances G3 Amazon EC2.

Les instances G3 prennent en charge les stations de travail virtuelles et les applications virtuelles NVIDIA GRID. Pour activer ces fonctionnalités, consultez Activer les applications virtuelles NVIDIA GRID (p. 310).

Instances G2

Les instances G2 utilisent des GPU NVIDIA GRID K520 et offrent une plate-forme économique à hautes performances pour les applications graphiques qui utilisent DirectX ou OpenGL. Les GPU NVIDIA GRID prennent également en charge la capture rapide de NVIDIA et encodent les opérations d'API. Les services de création vidéo, les visualisations 3D, la diffusion d'applications gourmandes en graphiques et autres charges de travail côté serveur sont des exemples d'applications.

Instances P3

Les instances P3 utilisent des GPU NVIDIA Tesla V100 et sont conçues pour le calcul GPU à usage général à l'aide des modèles de programmation CUDA ou OpenCL, ou via une infrastructure d'apprentissage automatique. Les instances P3 offrent une mise en réseau à bande passante élevée, de hautes capacités de calcul en virgule flottante de mi-précision, de simple précision et double précision, ainsi que jusqu'à 32 Gio de mémoire par GPU. Cela en fait un outil idéal dans les domaines du Deep Learning, de la mécanique des fluides numérique, des calculs financiers, de l'analyse sismique, de la modélisation moléculaire, de la génomique, du rendu et d'autres charges de travail de calcul GPU côté serveur. Les GPU Tesla V100 ne prennent pas en charge le mode graphique.

Pour plus d'informations, consultez Instances P3 Amazon EC2.

Les instances P3 prennent en charge les transferts entre homologues NVIDIA NVLink. Pour plus d'informations, consultez NVIDIA NVLink.

Instances P2

Les instances P2 utilisent des GPU NVIDIA Tesla K80 et sont conçues pour le calcul GPU à usage général à l'aide des modèles de programmation CUDA ou OpenCL. Les instances P2 offrent une mise en réseau à bande passante élevée, de hautes capacités de calcul en virgule flottante simple et double précision et 12 Gio de mémoire par GPU. Elles sont ainsi parfaitement adaptées au Deep Learning, aux bases de données graphiques, aux bases de données de haute performance, à la modélisation numérique en dynamique des fluides, aux calculs financiers, à l'analyse sismique, à la modélisation moléculaire, à la recherche génomique, aux tâches de rendu et aux autres charges de travail de calcul GPU côté serveur.

Les instances P2 prennent en charge les transferts pair à pair NVIDIA GPUDirect. Pour de plus amples informations, veuillez consulter NVIDIA GPUDirect.

Instances avec AWS Inferentia

Les instances Inf1 sont conçues pour accélérer le machine learning à l'aide d'AWS Inferentia, une puce IA/ML personnalisée d'Amazon qui fournit des performances élevées et une inférence de machine learning à faible latence. Ces instances sont optimisées pour déployer des modèles de Deep Learning (DL) pour des applications telles que le traitement du langage naturel, la détection et la classification des objets, la personnalisation et le filtrage du contenu et la reconnaissance vocale.

Il y a plusieurs façons de démarrer :

- Utilisez SageMaker, un service entièrement géré qui est le moyen le plus simple de démarrer avec les modèles de machine learning. Pour plus d'informations, consultez Compiler et déployer un modèle TensorFlow sur les instances Inf1 sur github.
- Lancez une instance Inf1 à l'aide de l'AMI Deep Learning. Pour plus d'informations, consultez AWS Inferentia avec DLAMI dans le Manuel du développeur AWS Deep Learning AMI.
- Lancez une instance Inf1 à l'aide de votre propre AMI et installez le kit SDK AWS Neuron, qui vous permet de compiler, d'exécuter et de profiler des modèles de deep learning pour AWS Inferentia.
- Lancez une instance de conteneur à l'aide d'une instance InF1 et d'une AMI optimisée par Amazon ECS. Pour plus d'informations, consultez (Inferentia) AMIAmazon Linux 2 dans le Amazon Elastic Container Service Developer Guide.

- Créez un cluster Amazon EKS avec des nœuds exécutant des instances Inf1. Pour plus d'informations, consultez Support Inferentia dans le Guide de l'utilisateur Amazon EKS.

Pour de plus amples informations, veuillez consulter Machine Learning sur AWS.

Instances Inf1

Les instances Inf1 utilisent des puces d'inférence de machine learning AWS Inferentia. Inferentia a été développé pour garantir des performances d'inférence à faible latence très rentables à n'importe quelle échelle.

Pour de plus amples informations, veuillez consulter Instances Inf1 Amazon EC2.

Instances FPGA

Les instances FPGA donnent accès à d'importants FPGA avec des millions de cellules logiques de système parallèle. Vous pouvez utiliser des instances de calcul accéléré FPGA pour accélérer des charges de travail comme l'analyse du génome, l'analyse financière, le traitement vidéo en temps réel, l'analyse du Big Data et les charges de travail de sécurité en tirant parti des accélérations matérielles personnalisées. Vous pouvez développer ces accélérations à l'aide des langages de description de matériel comme Verilog ou VHDL ou en utilisant des langages de niveau supérieur comme les infrastructures de calcul parallèle OpenCL (Open Computing Language). Vous pouvez également développer votre propre code d'accélération matérielle ou acheter des accélérations matérielles via AWS Marketplace.

L'AMI FPGA Developer AMI fournit les outils nécessaires pour développer, tester et créer des images AFI. Vous pouvez utiliser l'AMI des développeurs de FPGA sur n'importe quelle instance EC2 avec au moins 32 Go de mémoire système (instances C5, M4 et R4 par exemple).

Pour plus d'informations, consultez la documentation du kit de développement matériel FPGA AWS.

Instances F1

Les instances F1 utilisent les FPGA VU9P Xilinx UltraScale+ et sont conçues pour accélérer des algorithmes de calculs intensifs, comme les opérations de flux de données ou hautement parallèles non appropriées aux UC à usage général. Chaque FPGA dans une instance F1 contient environ 2,5 millions d'éléments logiques et approximativement 6 800 moteurs DSP (Digital Signal Processing) avec 64 Gio de mémoire protégée ECC DDR locale, connectés à l'instance par une connexion PCIe Gen3 x16 dédiée. Les instances F1 fournissent des volumes SSD NVMe locaux.

Les développeurs peuvent utiliser l'AMI des développeurs de FPGA et le kit de développement matériel AWS pour créer des accélérations matérielles personnalisées à utiliser sur des instances F1. L'AMI des développeurs de FPGA comprend des outils de développement pour le développement entier de FPGA dans le cloud. A l'aide de ces outils, les développeurs peuvent créer et partager des images AFI (Amazon FPGA Images) qui peuvent être chargées sur le FPGA d'une instance F1.

Pour plus d'informations, consultez Instances F1 Amazon EC2.

Spécifications matérielles

Vous trouverez ci-dessous un résumé des spécifications matérielles relatives aux instances à calcul accéléré.

Type d'instance	vCPU par défaut	Mémoire (Gio)	Accélérateurs
`p2.xlarge`	4	61	1
`p2.8xlarge`	32	488	8
`p2.16xlarge`	64	732	16

Type d'instance	vCPU par défaut	Mémoire (Gio)	Accélérateurs
`p3.2xlarge`	8	61	1
`p3.8xlarge`	32	244	4
`p3.16xlarge`	64	488	8
`p3dn.24xlarge`	96	768	8
`g2.2xlarge`	8	15	1
`g2.8xlarge`	32	60	4
`g3s.xlarge`	4	30,5	1
`g3.4xlarge`	16	122	1
`g3.8xlarge`	32	244	2
`g3.16xlarge`	64	488	4
`g4dn.xlarge`	4	16	1
`g4dn.2xlarge`	8	32	1
`g4dn.4xlarge`	16	64	1
`g4dn.8xlarge`	32	128	1
`g4dn.12xlarge`	48	192	4
`g4dn.16xlarge`	64	256	1
`g4dn.metal`	96	384	8
`f1.2xlarge`	8	122	1
`f1.4xlarge`	16	244	2
`f1.16xlarge`	64	976	8
`inf1.xlarge`	4	8	1
`inf1.2xlarge`	8	16	1
`inf1.6xlarge`	24	48	4
`inf1.24xlarge`	96	192	16

Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.

Pour plus d'informations sur la spécification des options d'UC, consultez Optimisations des options d'UC (p. 673).

Performances de l'instance

Il existe plusieurs optimisations de configuration GPU que vous pouvez effectuer pour obtenir les meilleures performances sur vos instances. Pour plus d'informations, consultez Optimisation des paramètres GPU (p. 311).

Les instances optimisées EBS vous permettent d'obtenir régulièrement des performances élevées pour vos volumes EBS en éliminant les conflits entre les E/S Amazon EBS et tout autre trafic réseau de votre instance Certaines instances optimisées à calcul accéléré sont optimisées pour EBS par défaut sans frais supplémentaires. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Certains types d'instances à calcul accéléré permettent de contrôler les états C et P du processeur sur Linux. Les états C contrôlent les niveaux de veille d'un noyau lorsqu'il est inactif, tandis que les états P contrôlent les performances attendues d'un noyau (en termes de fréquence d'UC). Pour plus d'informations, consultez Contrôle des états du processeur pour votre instance EC2 (p. 662).

Performances réseau

Vous pouvez activer la mise en réseau améliorée sur les types d'instance pris en charge pour fournir des latences plus faibles, une instabilité moindre sur le réseau et des performances de débit en paquets par seconde (PPS) plus élevées. La plupart des applications ne nécessitent pas en permanence un haut niveau de performances réseau, mais peuvent tirer profit d'un accès à une bande passante accrue lorsqu'elles envoient ou reçoivent des données. Pour plus d'informations, consultez Mise en réseau améliorée sur Linux (p. 864).

Vous trouverez ci-dessous un résumé des performances réseau relatives aux instances à calcul accéléré qui prennent en charge la mise en réseau améliorée.

Type d'instance	Performances réseau	Mise en réseau améliorée
`f1.2xlarge` | `f1.4xlarge` | `g3.4xlarge` | `p3.2xlarge`	Jusqu'à 10 Gbit/s †	ENA (p. 865)
`g3s.xlarge` | `g3.8xlarge` | `p2.8xlarge` | `p3.8xlarge`	10 Gb/s	ENA (p. 865)
`g4dn.xlarge` | `g4dn.2xlarge` | `g4dn.4xlarge` | `inf1.xlarge` | `inf1.2xlarge`	Jusqu'à 25 Gbit/s †	ENA (p. 865)
`f1.16xlarge` | `g3.16xlarge` | `inf1.6xlarge` | `p2.16xlarge` | `p3.16xlarge`	25 Gb/s	ENA (p. 865)
`g4dn.8xlarge` | `g4dn.12xlarge` | `g4dn.16xlarge`	50 Gb/s	ENA (p. 865)
`g4dn.metal` | `inf1.24xlarge` | `p3dn.24xlarge`	100 Gb/s	ENA (p. 865)

† Ces instances utilisent un mécanisme de crédit d'E/S réseau pour attribuer la bande passante réseau aux instances en fonction de l'utilisation moyenne de la bande passante. Elles accumulent des crédits lorsque la bande passante est inférieure à la bande passante de référence et peuvent les utiliser quand elles effectuent les transferts de données réseau. Pour de plus amples informations, ouvrez un dossier de support et renseignez-vous sur la bande passante de référence pour les types d'instance spécifiques qui vous intéressent.

Fonctionnalités des instances

Voici un résumé des fonctions pour les instances à calcul accéléré.

	EBS uniquement	EBS NVMe	Stockage d'instance	Groupe de placement
G2	Non	Non	SSD	Oui
G3	Oui	Non	Non	Oui
G4	Non	Oui	NVMe *	Oui
Inf1	Oui	Non	Non	Oui
P2	Oui	Non	Non	Oui
P3	`24xlarge` : non Toutes les autres tailles : oui	`24xlarge` : Oui Toutes les autres tailles : non	`24xlarge` : NVMe *	Oui
F1	Non	Non	NVMe *	Oui

* Le volume du périphérique racine doit être un volume Amazon EBS.

Pour plus d'informations, consultez les ressources suivantes :

- Amazon EBS et NVMe sur des instances Linux (p. 1209)
- Stockage d'instance Amazon EC2 (p. 1264)
- Groupes de placement (p. 924)

Notes de mise à jour

- Vous devez lancer l'instance à l'aide d'une AMI HVM.
- Les instances reposant sur le système Nitro (p. 218) présentent les exigences suivantes :
  - Les pilotes NVMe (p. 1209) doivent être installés.
  - Les pilotes Elastic Network Adapter (ENA) (p. 865) doivent être installés.

  Les AMI suivantes répondent à ces exigences :
  - Amazon Linux 2
  - Amazon Linux AMI 2018.03
  - Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure
  - Red Hat Enterprise Linux 7.4 ou version ultérieure
  - SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
  - CentOS 7.4.1708 ou version ultérieure
  - FreeBSD 11.1 ou version ultérieure
  - Debian GNU/Linux 9 ou version ultérieure
- Les instances GPU ne peuvent pas accéder à la GPU si les pilotes NVIDIA ne sont pas installés. Pour plus d'informations, consultez Installation des pilotes NVIDIA sur des instances Linux (p. 303).
- Le lancement d'une instance en matériel nu démarre le serveur sous-jacent, qui inclut la vérification de tous les composants du matériel et du microprogramme. Cela signifie que 20 minutes peuvent s'écouler entre le moment où l'instance passe à l'état d'exécution et le moment où elle devient disponible sur le réseau.
- Attacher ou détacher des volumes EBS ou des interfaces réseau secondaires à partir d'une instance en matériel nu requiert la prise en charge de l'enfichage à chaud natif de PCIe. Amazon Linux 2 et les dernières versions de l'Amazon Linux AMI prennent en charge l'enfichage à chaud natif de PCIe, ce qui

n'est pas le cas des versions antérieures. Vous devez activer les options de configuration suivantes du noyau Linux :

```
CONFIG_HOTPLUG_PCI_PCIE=y
CONFIG_PCIEASPM=y
```

- Les instances en matériel nu utilisent un périphérique série basé sur PCI plutôt qu'un périphérique série basé sur le port d'E/S. Le noyau Linux kernel en amont et les dernières AMI Amazon Linux prennent en charge ce périphérique. Les instances en matériel nu fournissent également un tableau SPCR ACPI afin de permettre au système d'utiliser automatiquement le périphérique série basé sur PCI. Les dernières AMI Windows utilisent automatiquement le périphérique série basé sur PCI.
- Il y a une limite de 100 images AFI par région.
- Le nombre total d'instances que vous pouvez lancer dans une région est soumis à une limite, et il existe des limites supplémentaires sur certains types d'instances. Pour plus d'informations, consultez Combien d'instances puis-je exécuter dans Amazon EC2 ? dans le Forum Aux Questions (FAQ) Amazon EC2.

Installation des pilotes NVIDIA sur des instances Linux

Une instance avec un GPU attaché, telle qu'une instance P3 ou G4, doit avoir le pilote NVIDIA approprié installé. En fonction du type d'instance, vous pouvez télécharger un pilote NVIDIA public, télécharger un pilote depuis Amazon S3 disponible uniquement pour les clients AWS ou utiliser une AMI avec le pilote préinstallé.

Table des matières

- Types de pilote NVIDIA (p. 303)
- Pilotes disponibles par type d'instance (p. 304)
- Options d'installation (p. 304)
  - Option 1 : AMI avec les pilotes NVIDIA installés (p. 304)
  - Option 2 : Pilotes NVIDIA publics (p. 305)
  - Option 3 : Pilotes GRID (instances G3 et G4) (p. 305)
  - Option 4 : Pilotes de jeu NVIDIA (instances G4) (p. 307)
- Installation d'une version supplémentaire de CUDA (p. 310)

Types de pilote NVIDIA

Voici les principaux types de pilote NVIDIA qui peuvent être utilisés avec des instances basées sur GPU.

Pilotes Tesla

Ces pilotes sont principalement destinés aux charges de travail de calcul, qui utilisent des GPU pour des tâches de calcul telles que les calculs parallélisés à virgule flottante pour le machine learning et les transformations de Fourier rapides pour les applications de calcul hautes performances.

Pilotes GRID

Ces pilotes sont certifiés pour fournir des performances optimales pour les applications de visualisation professionnelles qui traitent des contenus tels que des modèles 3D ou des vidéos haute résolution. Vous pouvez configurer les pilotes GRID pour prendre en charge deux modes. Les stations de travail virtuelles Quadro permettent d'accéder à quatre écrans 4K par GPU. Les vApps GRID fournissent des fonctionnalités d'hébergement RDSH App.

Pilotes de jeu

Ces pilotes contiennent des optimisations pour le jeu et sont fréquemment mis à jour pour améliorer les performances. Ils prennent en charge un seul écran 4K par GPU.

Panneau de configuration NVIDIA

Le panneau de commande NVIDIA est pris en charge avec les pilotes GRID et Gaming. Il n'est pas pris en charge avec les pilotes Tesla.

API prises en charge pour les pilotes Tesla, GRID et de jeu

- OpenCL, OpenGL et Vulkan
- NVIDIA CUDA et bibliothèques associées (par exemple, cuDNN, TensorRT, nvJPEG et cuBLAS)
- NVENC pour l'encodage vidéo et NVDEC pour le décodage vidéo

Pilotes disponibles par type d'instance

Le tableau suivant récapitule les pilotes NVIDIA pris en charge pour chaque type d'instance de GPU.

Type d'instance	Pilote Tesla	Pilote GRID	Pilote de jeu
G2	Non	Oui	Non
G3	Oui	Oui	Non
G4	Oui	Oui	Oui
P2	Oui	Non	Non
P3	Oui	Oui †	Non

† Utilisation d'AMI Marketplace uniquement

Options d'installation

Utilisez l'une des options suivantes pour obtenir les pilotes NVIDIA requis pour votre instance de GPU.

Options

- Option 1 : AMI avec les pilotes NVIDIA installés (p. 304)
- Option 2 : Pilotes NVIDIA publics (p. 305)
- Option 3 : Pilotes GRID (instances G3 et G4) (p. 305)
- Option 4 : Pilotes de jeu NVIDIA (instances G4) (p. 307)

Option 1 : AMI avec les pilotes NVIDIA installés

AWS et NVIDIA offrent différentes AMI 'Amazon Machine Images) fournies avec des pilotes NVIDIA installés.

- Offres Marketplace avec le pilote Tesla
- Offres Marketplace avec le pilote GRID
- Offres Marketplace avec le pilote de jeu

Pour mettre à jour la version du pilote installée à l'aide de l'une de ces AMI, vous devez désinstaller les packages NVIDIA de votre instance pour éviter les conflits de version. Utilisez cette commande pour désinstaller les packages NVIDIA :

```
[ec2-user ~]$ sudo yum erase nvidia cuda
```

Le package de boîte à outils CUDA fourni par Amazon comporte des dépendances sur les pilotes NVIDIA. La désinstallation des packages NVIDIA efface la boîte à outils CUDA. Vous devez réinstaller la boîte à outils CUDA après avoir installé le pilote NVIDIA.

Option 2 : Pilotes NVIDIA publics

Les options offertes par AWS sont fournies avec la licence nécessaire pour le pilote. Alternativement, vous pouvez installer les pilotes publics et apporter votre propre licence. Pour installer un pilote public, téléchargez-le à partir du site NVIDIA comme décrit ici.

Vous pouvez également utiliser les options offertes par AWS plutôt que les pilotes publics. Pour utiliser un pilote GRID sur une instance P3, utilisez les AMI AWS Marketplace comme décrit dans l'option 1 (p. 304). Pour utiliser un pilote GRID sur une instance G3 ou G4, utilisez les AMI AWS Marketplace, comme décrit dans l'option 1 ou installer les pilotes NVIDIA fournis par AWS comme décrit dans l'option 3 (p. 305).

Pour télécharger un pilote NVIDIA public

Connectez-vous à l'instance Linux et téléchargez le pilote NVIDIA 64 bits approprié à votre type d'instance à partir de http://www.nvidia.com/Download/Find.aspx. Pour Type de produit, Série de produits et Produit, utilisez les options du tableau suivant.

Instance	Type de produit	Série de produits	Produit
G2	GRID	Série GRID	GRID K520
G3	Tesla	M-Class	M60
G4 †	Tesla	T-Series	T4
P2	Tesla	Série K	K80
P3	Tesla	Série V	V100

Les instances † G4 ont besoin d'un pilote de version 418.87 ou ultérieure.

Pour installer le pilote NVIDIA sur Linux

Pour plus d'informations sur l'installation et la configuration du pilote, reportez-vous au Guide de démarrage rapide d'installation du pilote NVIDIA.

Option 3 : Pilotes GRID (instances G3 et G4)

Ces téléchargements sont disponibles uniquement pour les clients AWS. Si vous téléchargez le pilote, vous acceptez d'employer le logiciel téléchargé uniquement pour développer des AMIs à utiliser avec le matériel NVIDIA Tesla T4 ou NVIDIA Tesla M60. Dès l'installation du logiciel, vous êtes lié par les conditions du document Contrat de licence utilisateur final NVIDIA GRID Cloud.

Prérequis

- Installez l'AWS CLI sur votre instance Linux et configurez les informations d'identification par défaut. Pour de plus amples informations, veuillez consulter Installation de l'AWS CLI dans le AWS Command Line Interface Guide de l'utilisateur.
- Les utilisateurs IAM doivent disposer des autorisations accordées par la stratégie Amazons3ReadOnlyAccess.

Pour installer le pilote NVIDIA GRID sur votre instance Linux

1. Connectez-vous à votre instance Linux. Installez gcc et make, si ce n'est pas déjà fait.

2. Mettez à jour le cache de votre package et obtenez les mises à jour de packages nécessaires pour votre instance.

   - Pour Amazon Linux, CentOS, et Red Hat Enterprise Linux :

   ```
   [ec2-user ~]$ sudo yum update -y
   ```

   - Pour Ubuntu et Debian :

   ```
   $ sudo apt-get update -y
   ```

3. (Ubuntu 16.04 et versions ultérieures, avec le package `linux-aws`) Mettez à niveau le package `linux-aws` pour recevoir la dernière version.

   ```
   $ sudo apt-get upgrade -y linux-aws
   ```

4. Redémarrez votre instance pour charger la dernière version du noyau.

   ```
   [ec2-user ~]$ sudo reboot
   ```

5. Reconnectez-vous à votre instance après son redémarrage.
6. Installez le compilateur gcc et le package d'en-têtes de noyau correspondant à la version du noyau que vous utilisez actuellement.

   - Pour Amazon Linux, CentOS, et Red Hat Enterprise Linux :

   ```
   [ec2-user ~]$ sudo yum install -y gcc kernel-devel-$(uname -r)
   ```

   - Pour Ubuntu et Debian :

   ```
   $ sudo apt-get install -y gcc make linux-headers-$(uname -r)
   ```

7. [CentOS, Red Hat Enterprise Linux, Ubuntu, Debian] Désactivez le pilote open source `nouveau` pour les cartes graphiques NVIDIA.

   a. Ajoutez `nouveau` au fichier de liste noire `/etc/modprobe.d/blacklist.conf`. Copiez le bloc de code suivant et collez-le dans un terminal.

   ```
   [ec2-user ~]$ cat << EOF | sudo tee --append /etc/modprobe.d/blacklist.conf
   blacklist vga16fb
   blacklist nouveau
   blacklist rivafb
   blacklist nvidiafb
   blacklist rivatv
   EOF
   ```

   b. Modifiez le fichier `/etc/default/grub` et ajoutez la ligne suivante :

   ```
   GRUB_CMDLINE_LINUX="rdblacklist=nouveau"
   ```

   c. Générez à nouveau la configuration Grub.

   - Pour CentOS et Red Hat Enterprise Linux :

   ```
   [ec2-user ~]$ sudo grub2-mkconfig -o /boot/grub2/grub.cfg
   ```

   - Pour Ubuntu et Debian :

```
$ sudo update-grub
```

8. Téléchargez l'utilitaire d'installation du pilote GRID à l'aide de la commande suivante :

```
[ec2-user ~]$ aws s3 cp --recursive s3://ec2-linux-nvidia-drivers/latest/ .
```

   Plusieurs versions du pilote GRID sont stockées dans ce compartiment. Vous pouvez voir toutes les versions disponibles à l'aide de la commande suivante.

```
[ec2-user ~]$ aws s3 ls --recursive s3://ec2-linux-nvidia-drivers/
```

   À partir de GRID version 11.0, vous pouvez utiliser les packages de pilotes sous `latest` pour les instances G3 et G4. Nous n'ajouterons pas les versions postérieures à 11.0 à `g4/latest`, mais nous conserverons la version 11.0 et les versions antérieures spécifiques à G4 sous `g4/latest`.

9. Ajoutez les autorisations pour exécuter l'utilitaire d'installation du pilote à l'aide de la commande suivante.

```
[ec2-user ~]$ chmod +x NVIDIA-Linux-x86_64*.run
```

10. Exécutez le script d'auto-installation comme suit pour installer le pilote GRID que vous avez téléchargé. Exemples :

```
[ec2-user ~]$ sudo /bin/sh ./NVIDIA-Linux-x86_64*.run
```

   Lorsque vous y êtes invité, acceptez le contrat de licence et spécifiez les options d'installation comme requis (vous pouvez accepter les options par défaut).

11. Redémarrez l'instance.

```
[ec2-user ~]$ sudo reboot
```

12. Vérifiez que le pilote fonctionne. La sortie de la commande suivante affiche la version installée du pilote NVIDIA, ainsi que des détails sur les GPU.

```
[ec2-user ~]$ nvidia-smi -q | head
```

13. (Facultatif) Selon votre cas d'utilisation, vous pouvez effectuer les étapes facultatives suivantes. Si vous n'avez pas besoin de cette fonctionnalité, n'effectuez pas ces étapes.

    a. (Facultatif) Pour profiter des quatre écrans d'une résolution allant jusqu'à 4K, configurez le protocole d'affichage haute performance NICE DCV.

    b. Le mode de station de travail virtuelle NVIDIA Quadro est activé par défaut. Pour activer les fonctionnalités d'hébergement de GRID Virtual Applications for RDSH Application, suivez les étapes d'activation de GRID Virtual Applications dans Activer les applications virtuelles NVIDIA GRID (p. 310).

Option 4 : Pilotes de jeu NVIDIA (instances G4)

Ces pilotes sont disponibles uniquement pour les clients AWS. Si vous les téléchargez, vous acceptez d'employer le logiciel téléchargé uniquement pour développer des AMIs à utiliser avec le matériel NVIDIA Tesla T4. Dès l'installation du logiciel, vous êtes lié par les conditions du document Contrat de licence utilisateur final NVIDIA GRID Cloud.

Prérequis

- Installez l'AWS CLI sur votre instance Linux et configurez les informations d'identification par défaut. Pour de plus amples informations, veuillez consulter Installation de l'AWS CLI dans le AWS Command Line Interface Guide de l'utilisateur.
- Les utilisateurs IAM doivent disposer des autorisations accordées par la stratégie Amazons3ReadOnlyAccess.

Pour installer le pilote de jeu NVIDIA sur votre instance Linux

1. Connectez-vous à votre instance Linux. Installez gcc et make, si ce n'est pas déjà fait.
2. Mettez à jour le cache de votre package et obtenez les mises à jour de packages nécessaires pour votre instance.

   - Pour Amazon Linux, CentOS, et Red Hat Enterprise Linux :

   ```
   [ec2-user ~]$ sudo yum update -y
   ```

   - Pour Ubuntu et Debian :

   ```
   $ sudo apt-get update -y
   ```

3. (Ubuntu 16.04 et versions ultérieures, avec le package `linux-aws`) Mettez à niveau le package `linux-aws` pour recevoir la dernière version.

   ```
   $ sudo apt-get upgrade -y linux-aws
   ```

4. Redémarrez votre instance pour charger la dernière version du noyau.

   ```
   [ec2-user ~]$ sudo reboot
   ```

5. Reconnectez-vous à votre instance après son redémarrage.
6. Installez le compilateur gcc et le package d'en-têtes de noyau correspondant à la version du noyau que vous utilisez actuellement.

   - Pour Amazon Linux, CentOS, et Red Hat Enterprise Linux :

   ```
   [ec2-user ~]$ sudo yum install -y gcc kernel-devel-$(uname -r)
   ```

   - Pour Ubuntu et Debian :

   ```
   $ sudo apt-get install -y gcc make linux-headers-$(uname -r)
   ```

7. [CentOS, Red Hat Enterprise Linux, Ubuntu, Debian] Désactivez le pilote open source `nouveau` pour les cartes graphiques NVIDIA.

   a. Ajoutez `nouveau` au fichier de liste noire `/etc/modprobe.d/blacklist.conf`. Copiez le bloc de code suivant et collez-le dans un terminal.

   ```
   [ec2-user ~]$ cat << EOF | sudo tee --append /etc/modprobe.d/blacklist.conf
   blacklist vga16fb
   blacklist nouveau
   blacklist rivafb
   blacklist nvidiafb
   blacklist rivatv
   EOF
   ```

b. Modifiez le fichier `/etc/default/grub` et ajoutez la ligne suivante :

```
GRUB_CMDLINE_LINUX="rdblacklist=nouveau"
```

c. Générez à nouveau la configuration Grub.

- Pour CentOS et Red Hat Enterprise Linux :

```
[ec2-user ~]$ sudo grub2-mkconfig -o /boot/grub2/grub.cfg
```

- Pour Ubuntu et Debian :

```
$ sudo update-grub
```

8. Téléchargez l'utilitaire d'installation du pilote de jeu à l'aide de la commande suivante :

```
[ec2-user ~]$ aws s3 cp --recursive s3://nvidia-gaming/linux/latest/ .
```

Plusieurs versions du pilote de jeu sont stockées dans ce compartiment. Vous pouvez voir toutes les versions disponibles à l'aide de la commande suivante :

```
[ec2-user ~]$ aws s3 ls --recursive s3://nvidia-gaming/linux/
```

9. Ajoutez les autorisations pour exécuter l'utilitaire d'installation du pilote à l'aide de la commande suivante.

```
[ec2-user ~]$ chmod +x NVIDIA-Linux-x86_64*.run
```

10. Exécutez le programme d'installation à l'aide de la commande suivante :

```
[ec2-user ~]$ sudo ./NVIDIA-Linux-x86_64*.run
```

Lorsque vous y êtes invité, acceptez le contrat de licence et spécifiez les options d'installation comme requis (vous pouvez accepter les options par défaut).

11. Utilisez la commande suivante pour créer le fichier de configuration requis.

```
[ec2-user ~]$ cat << EOF | sudo tee -a /etc/nvidia/gridd.conf
vGamingMarketplace=2
EOF
```

12. Utilisez la commande suivante pour télécharger et renommer le fichier de certification.

- Pour une version 440.68 ou ultérieure :

```
[ec2-user ~]$ sudo curl -o /etc/nvidia/GridSwCert.txt "https://nvidia-
gaming.s3.amazonaws.com/GridSwCert-Archive/GridSwCert-Linux_2020_04.cert"
```

- Pour des versions antérieures :

```
[ec2-user ~]$ sudo curl -o /etc/nvidia/GridSwCert.txt "https://nvidia-
gaming.s3.amazonaws.com/GridSwCert-Archive/GridSwCert-Linux_2019_09.cert"
```

13. Redémarrez l'instance.

```
[ec2-user ~]$ sudo reboot
```

14. (Facultatif) Pour profiter d'un seul écran d'une résolution allant jusqu'à 4K, configurez le protocole d'affichage haute performance NICE DCV. Si vous n'avez pas besoin de cette fonctionnalité, n'effectuez pas cette étape.

Installation d'une version supplémentaire de CUDA

Après avoir installé un pilote graphique NVIDIA sur votre instance, vous pouvez installer une version de CUDA autre que celle fournie avec le pilote graphique. La procédure suivante montre comment configurer plusieurs versions de CUDA sur l'instance.

Pour installer la boîte à outils CUDA

1. Connectez-vous à votre instance Linux.
2. Ouvrez le site web NVIDIA et sélectionnez la version de CUDA dont vous avez besoin.
3. Sélectionnez l'architecture, la distribution et la version du système d'exploitation de votre instance. Pour Installer Type (Type de programme d'installation), sélectionnez runfile (local).
4. Suivez les instructions pour télécharger le script d'installation.
5. Ajoutez les autorisations d'exécution au script d'installation que vous avez téléchargé à l'aide de la commande suivante.

```
[ec2-user ~]$ chmod +x downloaded_installer_file
```

6. Exécutez le script d'installation comme suit pour installer la boîte à outils CUDA et ajouter le numéro de version CUDA au chemin d'accès de la boîte à outils.

```
[ec2-user ~]$ sudo downloaded_installer_file --silent --override --toolkit --samples --
toolkitpath=/usr/local/cuda-version --samplespath=/usr/local/cuda --no-opengl-lib
```

7. (Facultatif) Définissez la version CUDA par défaut comme suit.

```
[ec2-user ~]$ ln -s /usr/local/cuda-version /usr/local/cuda
```

Activer les applications virtuelles NVIDIA GRID

Pour activer les applications virtuelles GRID sur des instances G3 et G4 (la station de travail virtuelle NVIDIA GRID est activée par défaut), vous devez définir le type de produit pour le pilote dans le fichier `/etc/nvidia/gridd.conf` .

Pour activer les applications virtuelles GRID sur des instances Linux

1. Créez le fichier `/etc/nvidia/gridd.conf` à partir du modèle de fichier fourni.

```
[ec2-user ~]$ sudo cp /etc/nvidia/gridd.conf.template /etc/nvidia/gridd.conf
```

2. Ouvrez le fichier `/etc/nvidia/gridd.conf` dans votre éditeur de texte favori.
3. Accédez à la ligne `FeatureType` et affectez-lui la valeur 0. Puis ajoutez une ligne avec `IgnoreSP=TRUE`.

```
FeatureType=0
IgnoreSP=TRUE
```

4. Enregistrez le fichier et quittez l'éditeur.
5. Redémarrez l'instance pour récupérer la nouvelle configuration.

```
[ec2-user ~]$ sudo reboot
```

Optimisation des paramètres GPU

Il existe plusieurs optimisations de configuration GPU que vous pouvez effectuer pour obtenir des performances optimales sur les instances G3, G4, P2, P3 et P3dn. Par défaut, le pilote NVIDIA utilise une fonctionnalité autoboost, qui modifie les vitesses d'horloge GPU. En désactivant la fonctionnalité autoboost et en définissant les vitesses d'horloge GPU à leur fréquence maximale, vous pouvez obtenir les performances maximales de vos instances GPU. La procédure suivante vous permet de configurer la permanence des paramètres de GPU, de désactiver la fonctionnalité autoboost et de définir les vitesses d'horloge GPU à leur fréquence maximale.

Pour optimiser les paramètres GPU

1. Configurez les paramètres GPU de sorte qu'ils soient permanents. L'exécution de cette commande peut prendre plusieurs minutes.

   ```
   [ec2-user ~]$ sudo nvidia-persistenced
   ```

2. Désactivez la fonction autoboost pour toutes les GPU sur l'instance.

   ```
   [ec2-user ~]$ sudo nvidia-smi --auto-boost-default=0
   ```

   > **Note**
   >
   > Les GPU sur les instances P3, P3dn et G4 ne prennent pas en charge la fonction autoboost.

3. Définissez toutes les vitesses d'horloge GPU à leur fréquence maximale. Utilisez les vitesses d'horloge de mémoire et de graphiques spécifiées dans les commandes suivantes.

   > **Note**
   >
   > Certaines versions du pilote NVIDIA n'autorisent pas la définition de la vitesse d'horloge de l'application et déclenchent une erreur `"Setting applications clocks is not supported for GPU …"` que vous pouvez ignorer.

   - Instances G3 :

     ```
     [ec2-user ~]$ sudo nvidia-smi -ac 2505,1177
     ```

   - Instances G4 :

     ```
     [ec2-user ~]$ sudo nvidia-smi -ac 5001,1590
     ```

   - Instances P2 :

     ```
     [ec2-user ~]$ sudo nvidia-smi -ac 2505,875
     ```

   - Instances P3 et P3dn :

     ```
     [ec2-user ~]$ sudo nvidia-smi -ac 877,1530
     ```

Recherche d'un type d'instance Amazon EC2

Pour pouvoir lancer une instance, vous devez sélectionner un type d'instance à utiliser. Le type d'instance que vous choisissez peut différer selon vos exigences concernant les instances que vous allez lancer. Par exemple, vous pouvez choisir un type d'instance en fonction des exigences suivantes :

- Zone ou région de disponibilité
- Calcul
- Mémoire
- Mise en réseau
- Tarification
- Stockage

Recherche d'un type d'instance à l'aide de la console

Vous pouvez trouver un type d'instance qui répond à vos besoins à l'aide de la console Amazon EC2.

Recherche d'un type d'instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région dans laquelle lancer vos instances. Vous pouvez sélectionner n'importe quelle région disponible, quel que soit votre emplacement.
3. Dans le volet de navigation, choisissez Types d'instances.
4. (Facultatif) Choisissez l'icône de préférences pour sélectionner les attributs de type d'instance à afficher, tels que la tarification Linux à la demande, puis choisissez Valider. Vous pouvez également sélectionner un type d'instance et afficher tous les attributs à l'aide du volet Détails.
5. Utilisez les attributs de type d'instance pour filtrer la liste des types d'instance affichés uniquement aux types d'instance qui répondent à vos besoins. Par exemple, vous pouvez répertorier tous les types d'instances qui possèdent plus de huit UC virtuelles (vCPU) et qui prennent également en charge l'hibernation.
6. (Facultatif) Sélectionnez plusieurs types d'instances pour afficher une comparaison côte-à-côte de tous les attributs dans le volet Details (Détails).
7. (Facultatif) Pour enregistrer la liste des types d'instance dans un fichier de valeurs séparées par des virgules (.csv) pour un examen plus approfondi, choisissez Télécharger la liste CSV. Le fichier inclut tous les types d'instance qui correspondent aux filtres que vous avez définis.
8. Après avoir localisé les types d'instance qui répondent à vos besoins, vous pouvez les utiliser pour lancer des instances. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Recherche d'un type d'instance à l'aide de l'AWS CLI

Vous pouvez utiliser les commandes AWS CLI pour Amazon EC2 pour rechercher un type d'instance qui répond à vos besoins.

Pour rechercher un type d'instance à l'aide de l'AWS CLI

1. Si vous ne l'avez pas déjà fait, installez la section AWS CLI Pour de plus amples informations, veuillez consulter AWS Command Line Interface Guide de l'utilisateur.
2. Utilisez la commande describe-instance-types pour filtrer les types d'instance en fonction des attributs d'instance. Par exemple, vous pouvez utiliser la commande suivante pour afficher uniquement les types d'instances avec 48 UC virtuelles (vCPU).

```
aws ec2 describe-instance-types --filters "Name=vcpu-info.default-vcpus,Values=48"
```

3. Utilisez la commande describe-instance-type-offres pour filtrer les types d'instance proposés par emplacement (région ou zone de disponibilité). Par exemple, vous pouvez utiliser la commande suivante pour afficher les types d'instance proposés dans la zone de disponibilité spécifiée.

```
aws ec2 describe-instance-type-offerings --location-type "availability-zone" --filters
 Name=location,Values=us-east-2a --region us-east-2
```

4. Après avoir localisé les types d'instance qui répondent à vos besoins, notez-les afin que vous puissiez utiliser ces types d'instance lorsque vous lancez des instances. Pour de plus amples informations, veuillez consulter Lancement d'une instance à l'aide de l'interface de ligne de commande AWS dans le AWS Command Line Interface Guide de l'utilisateur.

Modification du type d'instance

Au fur et à mesure que vos besoins évoluent, il se peut que vous constatiez que votre instance est sur-utilisée (le type d'instance est trop petit) ou sous-utilisée (le type d'instance est trop grand). Si tel est le cas, vous pouvez modifier la taille de votre instance. Par exemple, si votre instance `t2.micro` est trop petite pour sa charge de travail, vous pouvez la remplacer par un autre type d'instance approprié pour la charge de travail.

Vous pouvez également migrer d'un type d'instance d'une génération précédente vers un type d'instance de génération actuelle pour tirer parti de certaines fonctionnalités, telles que la prise en charge d'IPv6.

Si le périphérique racine de votre instance est un volume EBS, vous pouvez modifier la taille de l'instance en modifiant simplement son type, opération appelée redimensionnement. Si le périphérique racine de votre instance est un volume de stockage d'instance, vous devez migrer votre application vers une nouvelle instance ayant le type d'instance nécessaire. Pour plus d'informations sur les volumes du périphérique racine, consultez Stockage pour le périphérique racine (p. 106).

Lorsque vous redimensionnez une instance, vous devez sélectionner un type d'instance compatible avec la configuration de l'instance. Si le type d'instance souhaité n'est pas compatible avec votre configuration d'instance, vous devez migrer votre application vers une nouvelle instance ayant le type nécessaire.

> **Important**
>
> Lors du redimensionnement d'une instance, l'instance redimensionnée a généralement le même nombre de volumes de stockage d'instance que celui spécifié lors du lancement de l'instance initiale. Avec les types d'instance qui prennent en charge les volumes de stockage d'instance NVMe (qui sont disponibles par défaut), l'instance redimensionnée peut avoir des volumes de stockage d'instance supplémentaires, selon l'IMA. Sinon, vous pouvez migrer manuellement votre application vers une instance avec un nouveau type d'instance, en spécifiant le nombre de volumes de stockage d'instance dont vous avez besoin lorsque vous lancez la nouvelle instance.

Sommaire

- Compatibilité pour le redimensionnement des instances (p. 313)
- Redimensionnement d'une instance basée sur Amazon EBS (p. 314)
- Migration d'une instance basée sur le stockage d'instance (p. 316)
- Migration vers une nouvelle configuration d'instance (p. 318)

Compatibilité pour le redimensionnement des instances

Vous ne pouvez redimensionner une instance que si son type d'instance en cours et le nouveau type souhaité sont compatibles comme suit :

- Type de virtualisation : les AMI Linux utilisent l'un des deux types de virtualisation : virtualisation paravirtuelle ou virtualisation HVM (Hardware Virtual Machine). Il n'est pas possible de redimensionner une instance qui a été lancée depuis une AMI de virtualisation paravirtuelle en un type d'instance qui n'utilise que la virtualisation HVM. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108). Pour vérifier le type de virtualisation de votre instance, consultez le champ Virtualisation dans le volet des détails de l'écran Instances dans la console Amazon EC2.
- Architecture : les AMI étant propres à l'architecture du processeur, vous devez sélectionner un type d'instance doté de la même architecture de processeur que le type d'instance en cours. Par exemple :
  - Si vous redimensionnez un type d'instance avec un processeur basé sur l'architecture Arm, vous êtes limité aux types d'instances qui prennent en charge un processeur basé sur l'architecture Arm, notamment A1 et M6g.
  - Les types d'instances suivants sont les seuls qui prennent en charge les AMIs 32 bits : `t2.nano`, `t2.micro`, `t2.small`, `t2.medium`, `c3.large`, `t1.micro`, `m1.small`, `m1.medium` et `c1.medium`. Si vous redimensionnez une instance 32 bits, vous êtes limité à ces types d'instance.
- Réseau : les types d'instance plus récents doivent être lancés dans un VPC. Par conséquent, vous ne pouvez pas redimensionner une instance dans la plateforme EC2-Classic en un type d'instance disponible uniquement dans un VPC à moins que vous n'ayez un VPC autre qu'un VPC par défaut. Pour vérifier si votre instance est dans un VPC, consultez la valeur de ID de VPC dans le volet des détails de l'écran Instances dans la console Amazon EC2. Pour plus d'informations, consultez Migration d'EC2-Classic vers un VPC (p. 961).
- Mise en réseau améliorée : les types d'instance prenant en charge la mise en réseau améliorée (p. 864) nécessitent l'installation des pilotes requis. Par exemple, les instances basées sur le système Nitro (p. 218) requièrent des AMI basées sur EBS avec installation des pilotes Elastic Network Adapter (ENA). Pour redimensionner une instance d'un type qui ne prend pas en charge la mise en réseau améliorée à un type qui prend en charge la mise en réseau améliorée, vous devez installer les pilotes ENA (p. 865) ou les pilotes ixgbevf (p. 878) sur l'instance, selon le cas.
- NVMe : les volumes EBS sont exposés sous forme de blocs NVMe sur des instances construites sur le Système Nitro (p. 218). Si vous redimensionnez une instance d'un type d'instance ne prenant pas en charge NVMe en un type d'instance prenant en charge NVMe, vous devez commencer par installer les pilotes NVMe (p. 1209) sur votre instance. En outre, les noms des périphériques que vous spécifiez dans le mappage de périphérique de stockage en mode bloc sont remplacés par les noms du périphérique NVMe (`/dev/nvme[0-26]n1`). Par conséquent, pour monter des systèmes de fichiers au moment du démarrage à l'aide de `/etc/fstab`, vous devez utiliser UUID/Label (UUID/Étiquette) au lieu des noms de périphériques.
- AMI : pour des informations au sujet des AMI requises par les types d'instance qui prennent en charge la mise en réseau améliorée et NVMe, consultez les notes de mise à jour dans la documentation suivante :
  - Instances à usage général (p. 222)
  - Instances de calcul optimisé (p. 270)
  - Instances de mémoire optimisée (p. 278)
  - Instances de stockage optimisé (p. 289)

Redimensionnement d'une instance basée sur Amazon EBS

Vous devez arrêter votre instance basée sur Amazon EBS avant de pouvoir modifier son type d'instance. Lorsque vous arrêtez et démarrez une instance, soyez conscient de ce qui suit :

- Nous déplaçons l'instance vers le nouveau matériel ; cependant, l'ID de l'instance ne change pas.
- Si votre instance possède une adresse IPv4, nous libérons l'adresse et lui attribuons une nouvelle adresse IPv4. L'instance conserve ses adresses IPv4 privées, les adresses IP Elastic et toutes les adresses IPv6.
- Si votre instance est dans un groupe Auto Scaling, le service Amazon EC2 Auto Scaling marque l'instance arrêtée comme étant défectueuse, et peut y mettre fin et lancer une instance de remplacement. Pour empêcher que cela ne se produise, vous pouvez suspendre les processus de dimensionnement

pour le groupe pendant que vous redimensionnez votre instance. Pour plus d'informations, consultez Suspension et reprise des processus de mise à l'échelle dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

- Si votre instance se trouve dans un groupe de placement de clusters (p. 925) et que le démarrage de l'instance échoue après une modification du type d'instance, procédez comme suit : arrêtez toutes les instances du groupe de placement de clusters, changez le type de l'instance affectée, puis redémarrez toutes les instances du groupe de placement de clusters.
- Veillez à prévoir un temps d'arrêt pendant que votre instance est arrêtée. L'arrêt et le redimensionnement d'une instance peuvent prendre quelques minutes, et la durée du redémarrage de votre instance peut varier en fonction des scripts de démarrage de votre application.

Pour plus d'informations, consultez Arrêt et démarrage de votre instance (p. 625).

Utilisez la procédure suivante pour redimensionner une instance basée sur Amazon EBS à l'aide de la AWS Management Console.

New console

Pour redimensionner une instance basée sur les Amazon EBS

1. (Facultatif) Si le nouveau type d'instance requiert des pilotes qui ne sont pas installés sur l'instance existante, vous devez vous connecter à votre instance et installer les pilotes. Pour plus d'informations, consultez Compatibilité pour le redimensionnement des instances (p. 313).
2. Ouvrez la console Amazon EC2.
3. Dans le panneau de navigation, choisissez Instances.
4. Sélectionnez l'instance et choisissez Actions, État de l'instance, Arrêter l'instance.
5. Dans la boîte de dialogue de confirmation, choisissez Arrêter. L'arrêt de l'instance peut prendre quelques minutes.
6. Tandis que l'instance est toujours sélectionnée, choisissez Actions, Paramètres de l'instance, puis Changer le type d'instance. Cette action est désactivée si l'état de l'instance n'est pas `stopped`.
7. Dans la boîte de dialogue Changer le type d'instance, exécutez l'une des actions suivantes :
   a. Dans Type d'instance, sélectionnez le type d'instance souhaité. Si le type d'instance que vous souhaitez n'apparaît pas dans la liste, il n'est pas compatible avec la configuration de votre instance (en raison du type de virtualisation, par exemple). Pour plus d'informations, consultez Compatibilité pour le redimensionnement des instances (p. 313).
   b. (Facultatif) Si le type d'instance que vous avez sélectionné prend en charge l'optimisation EBS, sélectionnez Optimisé pour EBS pour activer l'optimisation EBS ou annulez la sélection d'Optimisé pour EBS pour désactiver l'optimisation EBS. Si le type d'instance que vous avez sélectionné est optimisé pour EBS par défaut, Optimisé pour EBS est sélectionné et vous ne pouvez pas annuler la sélection.
   c. Choisissez Appliquer pour accepter les nouveaux paramètres.
8. Pour redémarrer l'instance arrêtée, sélectionnez l'instance et choisissez État de l'instance, Démarrer l'instance. Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`.
9. (Dépannage) Si votre instance ne démarre pas, il est possible qu'une des exigences pour le nouveau type d'instance n'ait pas été respectée. Pour plus d'informations, consultez la section relative à la raison pour laquelle mon instance Linux ne démarre pas après que j'ai modifié son type.

Old console

Pour redimensionner une instance basée sur les Amazon EBS

1. (Facultatif) Si le nouveau type d'instance requiert des pilotes qui ne sont pas installés sur l'instance existante, vous devez vous connecter à votre instance et installer les pilotes. Pour plus d'informations, consultez Compatibilité pour le redimensionnement des instances (p. 313).
2. Ouvrez la console Amazon EC2.
3. Dans le panneau de navigation, choisissez Instances.
4. Sélectionnez l'instance et choisissez Actions, État de l'instance, Arrêter.
5. Dans la boîte de dialogue de confirmation, sélectionnez Oui, arrêter. L'arrêt de l'instance peut prendre quelques minutes.
6. Tandis que l'instance est toujours sélectionnée, choisissez Actions, Paramètres de l'instance, puis Changer le type d'instance. Cette action est désactivée si l'état de l'instance n'est pas `stopped`.
7. Dans la boîte de dialogue Changer le type d'instance, exécutez l'une des actions suivantes :
   a. Dans Type d'instance, sélectionnez le type d'instance souhaité. Si le type d'instance que vous souhaitez n'apparaît pas dans la liste, il n'est pas compatible avec la configuration de votre instance (en raison du type de virtualisation, par exemple). Pour plus d'informations, consultez Compatibilité pour le redimensionnement des instances (p. 313).
   b. (Facultatif) Si le type d'instance que vous avez sélectionné prend en charge l'optimisation EBS, sélectionnez Optimisé pour EBS pour activer l'optimisation EBS ou annulez la sélection d'Optimisé pour EBS pour désactiver l'optimisation EBS. Si le type d'instance que vous avez sélectionné est optimisé pour EBS par défaut, Optimisé pour EBS est sélectionné et vous ne pouvez pas annuler la sélection.
   c. Choisissez Appliquer pour accepter les nouveaux paramètres.
8. Pour redémarrer l'instance arrêtée, sélectionnez l'instance et choisissez Actions, État de l'instance et Démarrer.
9. Dans la boîte de dialogue de confirmation, sélectionnez Oui, démarrer. Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`.
10. (Dépannage) Si votre instance ne démarre pas, il est possible qu'une des exigences pour le nouveau type d'instance n'ait pas été respectée. Pour plus d'informations, consultez la section relative à la raison pour laquelle mon instance Linux ne démarre pas après que j'ai modifié son type.

Migration d'une instance basée sur le stockage d'instance

Lorsque vous voulez déplacer votre instance d'une instance basée sur le stockage d'instance vers une instance basée sur le stockage d'instance mais d'un autre type, vous devez la migrer en créant une image à partir de votre instance, puis en lançant une nouvelle instance à partir de cette image avec le type d'instance dont vous avez besoin. Pour garantir que vos utilisateurs continuent à utiliser les applications que vous hébergez sur votre instance ininterrompue, vous devez choisir une adresse IP Elastic associée à votre instance originale et l'associer à la nouvelle instance. Vous pouvez alors mettre fin à l'instance d'origine.

New console

Pour migrer une instance basée sur le stockage d'instance

1. Sauvegardez les données qui se trouvent sur les volumes de stockage d'instance que vous devez conserver comme stockage permanent. Pour migrer les données sur les volumes EBS que vous devez conserver, prenez un instantané des volumes (consultez Création d'instantanés Amazon EBS (p. 1132)) ou détachez le volume de l'instance de façon à pouvoir l'attacher à la

nouvelle instance ultérieurement (consultez Détacher un volume Amazon EBS d'une instance Linux (p. 1127)).

2. Créez un AMI à partir de votre instance basée sur le stockage d'instance en remplissant les conditions requises et en suivant les procédures décrites dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136). Lorsque vous avez fini de créer un AMI à partir de votre instance, retournez à cette procédure.
3. Ouvrez la console Amazon EC2 et, dans le volet de navigation, sélectionnez AMI Dans les listes de filtres, choisissez M'appartenant, puis sélectionnez l'image que vous avez créée à l'étape précédente. Notez que Nom d'AMI est le nom que vous avez spécifié quand vous avez enregistré l'image et que Source est votre compartiment Amazon S3.

   Note

   Si l'AMI que vous avez créée dans l'étape précédente n'apparaît pas, assurez-vous d'avoir sélectionné la région dans laquelle vous avez créé votre AMI.
4. Choisissez Launch. Quand vous spécifiez les options de l'instance, veillez bien à sélectionner le nouveau type d'instance que vous voulez. Si le type d'instance que vous souhaitez ne peut pas être sélectionné, il n'est pas compatible avec la configuration de l'AMI que vous avez créée (en raison du type de virtualisation, par exemple). Vous pouvez aussi spécifier les volumes EBS que vous avez détachés de l'instance d'origine.

   Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`.
5. (Facultatif) Vous pouvez terminer l'instance avec laquelle vous avez démarré, si elle n'est plus nécessaire. Sélectionnez l'instance et vérifiez que vous êtes sur le point de terminer l'instance d'origine, et non la nouvelle instance (par exemple, vérifiez le nom ou l'heure du lancement). Choisissez État de l'instance, Résilier l'instance.

Old console

Pour migrer une instance basée sur le stockage d'instance

1. Sauvegardez les données qui se trouvent sur les volumes de stockage d'instance que vous devez conserver comme stockage permanent. Pour migrer les données sur les volumes EBS que vous devez conserver, prenez un instantané des volumes (consultez Création d'instantanés Amazon EBS (p. 1132)) ou détachez le volume de l'instance de façon à pouvoir l'attacher à la nouvelle instance ultérieurement (consultez Détacher un volume Amazon EBS d'une instance Linux (p. 1127)).
2. Créez un AMI à partir de votre instance basée sur le stockage d'instance en remplissant les conditions requises et en suivant les procédures décrites dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136). Lorsque vous avez fini de créer un AMI à partir de votre instance, retournez à cette procédure.
3. Ouvrez la console Amazon EC2 et, dans le volet de navigation, sélectionnez AMI Dans les listes de filtres, choisissez M'appartenant, puis choisissez l'image que vous avez créée à l'étape précédente. Notez que Nom d'AMI est le nom que vous avez spécifié quand vous avez enregistré l'image et que Source est votre compartiment Amazon S3.

   Note

   Si l'AMI que vous avez créée dans l'étape précédente n'apparaît pas, assurez-vous d'avoir sélectionné la région dans laquelle vous avez créé votre AMI.
4. Choisissez Launch. Quand vous spécifiez les options de l'instance, veillez bien à sélectionner le nouveau type d'instance que vous voulez. Si le type d'instance que vous souhaitez ne peut pas être sélectionné, il n'est pas compatible avec la configuration de l'AMI que vous avez créée (en raison du type de virtualisation, par exemple). Vous pouvez aussi spécifier les volumes EBS que vous avez détachés de l'instance d'origine.

Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`.

5. (Facultatif) Vous pouvez terminer l'instance avec laquelle vous avez démarré, si elle n'est plus nécessaire. Sélectionnez l'instance et vérifiez que vous êtes sur le point de terminer l'instance d'origine, et non la nouvelle instance (par exemple, vérifiez le nom ou l'heure du lancement). Choisissez Actions, État de l'instance, Résilier.

Migration vers une nouvelle configuration d'instance

Si la configuration active de votre instance est incompatible avec le nouveau type d'instance que vous voulez, vous ne pouvez pas redimensionner l'instance en ce type d'instance. A la place, vous pouvez migrer votre application vers une nouvelle instance ayant une configuration compatible avec le nouveau type d'instance que vous voulez.

Si vous voulez vous déplacer d'une instance lancée depuis une AMI de virtualisation paravirtuelle vers un type d'instance qui n'utilise que la virtualisation HVM, la procédure générale est la suivante :

Nouvelle console

Pour migrer votre application vers une instance compatible

1. Sauvegardez les données qui se trouvent sur les volumes de stockage d'instance que vous devez conserver comme stockage permanent. Pour migrer les données sur les volumes EBS que vous devez conserver, créez un instantané des volumes (consultez Création d'instantanés Amazon EBS (p. 1132)) ou détachez le volume de l'instance de façon à pouvoir l'attacher à la nouvelle instance ultérieurement (consultez Détacher un volume Amazon EBS d'une instance Linux (p. 1127)).
2. Lancez une nouvelle instance en sélectionnant les éléments suivants :
   - Un AMI HVM.
   - Le type d'instance HVM uniquement.
   - Si vous utilisez une adresse IP Elastic, sélectionnez le VPC dans lequel l'instance originale s'exécute.
   - Tous les volumes EBS que vous avez détachés de l'instance d'origine et que vous voulez attacher à la nouvelle instance, ou les nouveaux volumes EBS basés sur les instantanés que vous avez créés.
   - Si vous voulez autoriser le même trafic pour atteindre la nouvelle instance, sélectionnez le groupe de sécurité associé à l'instance d'origine.
3. Installez votre application et les logiciels requis sur l'instance.
4. Restaurez les données que vous avez sauvegardées depuis les volumes de stockage d'instance de l'instance d'origine.
5. Si vous utilisez une adresse IP Elastic, attribuez-la à la nouvelle instance lancée comme suit :
   a. Dans le panneau de navigation, sélectionnez Elastic IPs.
   b. Sélectionnez l'adresse IP Elastic associée à l'instance d'origine, choisissez Actions, puis Dissocier l'adresse IP Elastic. Sélectionnez Dissocier lorsque vous êtes invité à confirmer l'opération.
   c. L'adresse IP Elastic étant toujours sélectionnée, choisissez Actions, puis Associer l'adresse IP Elastic.
   d. Pour Resource type (Type de ressource), choisissez Instance.
   e. Pour Instance, choisissez l'instance à laquelle vous souhaitez associer l'adresse IP Elastic. Vous pouvez également entrer du texte pour rechercher une instance spécifique.
   f. (Facultatif) Pour Private IP address (Adresse IP privée), spécifiez une adresse IP privée à laquelle associer l'adresse IP Elastic.

   g. Choisissez Associate.
6. (Facultatif) Vous pouvez terminer l'instance d'origine si elle n'est plus nécessaire. Sélectionnez l'instance et vérifiez que vous êtes sur le point de terminer l'instance d'origine, et non la nouvelle instance (par exemple, vérifiez le nom ou l'heure du lancement). Choisissez État de l'instance, Résilier l'instance.

Ancienne console

Pour migrer votre application vers une instance compatible

1. Sauvegardez les données qui se trouvent sur les volumes de stockage d'instance que vous devez conserver comme stockage permanent. Pour migrer les données sur les volumes EBS que vous devez conserver, créez un instantané des volumes (consultez Création d'instantanés Amazon EBS (p. 1132)) ou détachez le volume de l'instance de façon à pouvoir l'attacher à la nouvelle instance ultérieurement (consultez Détacher un volume Amazon EBS d'une instance Linux (p. 1127)).
2. Lancez une nouvelle instance en sélectionnant les éléments suivants :
   - Un AMI HVM.
   - Le type d'instance HVM uniquement.
   - Si vous utilisez une adresse IP Elastic, sélectionnez le VPC dans lequel l'instance originale s'exécute.
   - Tous les volumes EBS que vous avez détachés de l'instance d'origine et que vous voulez attacher à la nouvelle instance, ou les nouveaux volumes EBS basés sur les instantanés que vous avez créés.
   - Si vous voulez autoriser le même trafic pour atteindre la nouvelle instance, sélectionnez le groupe de sécurité associé à l'instance d'origine.
3. Installez votre application et les logiciels requis sur l'instance.
4. Restaurez les données que vous avez sauvegardées depuis les volumes de stockage d'instance de l'instance d'origine.
5. Si vous utilisez une adresse IP Elastic, attribuez-la à la nouvelle instance lancée comme suit :
   a. Dans le panneau de navigation, sélectionnez Elastic IPs.
   b. Sélectionnez l'adresse IP Elastic associée à l'instance d'origine, choisissez Actions, puis Dissocier l'adresse. Sélectionnez Dissocier l'adresse lorsque vous êtes invité à confirmer l'opération.
   c. L'adresse IP Elastic étant toujours sélectionnée, choisissez Actions, puis Associer l'adresse.
   d. Dans Instance, sélectionnez la nouvelle instance, puis choisissez Associer.
6. (Facultatif) Vous pouvez terminer l'instance d'origine si elle n'est plus nécessaire. Sélectionnez l'instance et vérifiez que vous êtes sur le point de terminer l'instance d'origine, et non la nouvelle instance (par exemple, vérifiez le nom ou l'heure du lancement). Choisissez Actions, État de l'instance, Résilier.

Obtention de recommandations pour un type d'instance

AWS Compute Optimizer propose des recommandations d'instances Amazon EC2 pour vous aider à améliorer les performances, à économiser de l'argent, ou les deux. Vous pouvez utiliser ces recommandations pour décider de passer à un nouveau type d'instance.

Pour formuler des recommandations, Compute Optimizer analyse les spécifications et les métriques d'utilisation de vos instances existantes. Les données compilées sont ensuite utilisées pour recommander

les types d'instances Amazon EC2 qui sont le plus à même de gérer la charge de travail existante. Les recommandations sont renvoyées avec la tarification horaire des instances.

Cette rubrique explique comment afficher les recommandations via la console Amazon EC2. Pour de plus amples informations, consultez le Guide de l'utilisateur AWS Compute Optimizer.

> Note
>
> Pour obtenir des recommandations de Compute Optimizer, vous devez d'abord vous inscrire à Compute Optimizer. Pour plus d'informations, consultez Démarrez avec AWS Compute Optimizer dans le Guide de l'utilisateur AWS Compute Optimizer.

Table des matières

- Limites (p. 320)
- Conclusions (p. 320)
- Affichage des recommandations (p. 320)
- Considérations relatives à l'évaluation des recommandations (p. 322)

Limites

Compute Optimizer génère actuellement des recommandations pour les types d'instances M, C, R, T et X. Les autres types d'instance ne sont pas pris en compte par Compute Optimizer. Si vous utilisez d'autres types d'instances, ils ne seront pas répertoriés dans la vue des recommandations de Compute Optimizer. Pour de plus amples informations sur ces types d'instances et d'autres types, veuillez consulter Types d'instance (p. 213).

Conclusions

Compute Optimizer classe ses résultats pour les instances EC2 comme suit :

- Under-provisioned (Sous-allouée) – Une instance EC2 est considérée comme sous-allouée lorsqu'au moins une spécification de votre instance (l'UC, la mémoire ou le réseau, par exemple) ne répond pas aux exigences de performances de votre charge de travail. Les instances EC2 sous-allouées peuvent entraîner des performances d'application médiocres.
- Over-provisioned (Sur-allouée) – Une instance EC2 est considérée comme sur-allouée lorsque la taille d'au moins une spécification de votre instance (l'UC, la mémoire ou le réseau, par exemple) peut être réduite tout en répondant aux exigences de performances de votre charge de travail, et lorsqu'aucune spécification n'est sous-allouée. Les instances EC2 sur-allouées peuvent entraîner des coûts d'infrastructure inutiles.
- Optimized (Optimisée) – Une instance EC2 est considérée comme optimisée lorsque toutes les spécifications de votre instance (l'UC, la mémoire et le réseau, par exemple) répondent aux exigences de performances de votre charge de travail, et que l'instance n'est pas sur-allouée. Une instance EC2 optimisée exécute vos charges de travail avec des performances et des coûts d'infrastructure optimaux. Pour les instances optimisées, Compute Optimizer peut parfois recommander un type d'instance de nouvelle génération.
- None (Aucune) – Aucune recommandation n'est formulée pour cette instance. Cela peut se produire si vous êtes inscrit à Compute Optimizer depuis moins de 12 heures, ou lorsque l'instance s'exécute depuis moins de 30 heures, ou lorsque le type d'instance n'est pas pris en charge par Compute Optimizer. Pour de plus amples informations, consultez Limites (p. 320) dans la section précédente.

Affichage des recommandations

Une fois que vous avez choisi Compute Optimizer, vous pouvez afficher les résultats qu'il génère pour vos instances EC2 dans la console EC2. Vous pouvez ensuite accéder à la console Compute Optimizer pour

afficher les recommandations. Si vous vous êtes inscrit récemment, les résultats peuvent ne pas s'afficher dans la console EC2 avant 12 heures maximum.

Nouvelle console

Pour afficher une recommandation pour une instance EC2 via la console EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, puis choisissez l'instance EC2.
3. Sur la page récapitulative de l'instance, dans la bannière AWS Compute Optimizer située en bas de la page, choisissez Afficher les détails.

   L'instance s'ouvre dans Compute Optimizer, où elle est étiquetée comme instance Current (Actuelle) . Jusqu'à trois recommandations de type d'instance différentes, portant les noms Option 1, Option 2 et Option 3, sont fournies. La partie inférieure de la fenêtre affiche les données de métrique CloudWatch récentes pour l'instance actuelle : Utilisation de l'UC, Utilisation de la mémoire, Réseau entrant et Réseau sortant.
4. (Facultatif) Dans la console Compute Optimizer, choisissez l'icône des paramètres ( ), pour modifier les colonnes visibles du tableau ou pour afficher les informations de tarification publiques pour une option d'achat différente pour les types d'instances actuels et recommandés.

   > **Note**
   >
   > Si vous avez acheté une Instance réservée, votre instance à la demande peut être facturée au prix d'une Instance réservée. Avant de modifier votre type d'instance actuel, commencez par évaluer l'impact sur l'utilisation et la couverture de l'Instance réservée.

Ancienne console

Pour afficher une recommandation pour une instance EC2 via la console EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances.
3. Sélectionnez une instance et examinez le champ Résultat dans l'onglet Description. Choisissez Afficher les détails.

   L'instance s'ouvre dans Compute Optimizer, où elle est étiquetée comme instance Current (Actuelle) . Jusqu'à trois recommandations de type d'instance différentes, portant les noms Option 1, Option 2 et Option 3, sont fournies. La partie inférieure de la fenêtre affiche les données de métrique CloudWatch récentes pour l'instance actuelle : Utilisation de l'UC, Utilisation de la mémoire, Réseau entrant et Réseau sortant.
4. (Facultatif) Dans la console Compute Optimizer, choisissez l'icône Paramètres ( ), pour modifier les colonnes visibles du tableau ou pour afficher les informations de tarification publiques pour une option d'achat différente pour les types d'instances actuels et recommandés.

   > **Note**
   >
   > Si vous avez acheté une Instance réservée, votre instance à la demande peut être facturée au prix d'une Instance réservée. Avant de modifier votre type d'instance actuel, commencez par évaluer l'impact sur l'utilisation et la couverture de l'Instance réservée.

Déterminez si vous souhaitez utiliser l'une des recommandations. Décidez s'il convient d'améliorer les performances et/ou réduire les coûts. Pour plus d'informations, consultez Affichage des recommandations de ressources dans le Guide de l'utilisateur AWS Compute Optimizer.

Afficher les recommandations pour toutes les instances EC2 dans toutes les régions via la console Compute Optimizer

1. Ouvrez la console Compute Optimizer à l'adresse https://console.aws.amazon.com/compute-optimizer/.
2. Choisissez View recommendations for all EC2 instances (Afficher les recommandations pour toutes les instances EC2).
3. Vous pouvez effectuer les actions suivantes sur la page des recommandations :

   a. Pour filtrer les recommandations sur une ou plusieurs régions AWS, entrez le nom de la région dans la zone de texte Filter by one or more Regions (Filtrer par une ou plusieurs régions) ou choisissez une ou plusieurs régions dans la liste déroulante qui s'affiche.
   b. Pour afficher les recommandations relatives aux ressources d'un autre compte, choisissez Compte, puis sélectionnez un autre ID de compte.

      Cette option n'est disponible que si vous êtes connecté au compte principal d'une organisation et que vous vous êtes inscrit à tous les comptes membres de l'organisation.
   c. Pour effacer les filtres sélectionnés, choisissez Annuler les filtres.
   d. Pour modifier l'option d'achat affichée pour les types d'instances actuels et recommandés, choisissez l'icône Paramètres (), puis Instances à la demande, Reserved Instances, standard 1-year no upfront (Instances réservées, 1 an standard, sans frais initiaux) ou Reserved Instances, standard 3-year no upfront (Instances réservées, 3 ans standard, sans frais initiaux).
   e. Pour afficher des détails, tels que des recommandations supplémentaires et une comparaison des métriques d'utilisation, choisissez le résultat (Under-provisioned (Sous-allouée), Over-provisioned (Sur-allouée) ou Optimized (Optimisée)) en regard de l'instance souhaitée. Pour plus d'informations, consultez Affichage des détails de la ressource dans le Guide de l'utilisateur AWS Compute Optimizer.

Considérations relatives à l'évaluation des recommandations

Avant de modifier un type d'instance, tenez compte des éléments suivants :

- Les recommandations ne prévoient pas votre utilisation. Les recommandations sont basées sur votre historique d'utilisation au cours de la période de 14 jours la plus récente. Veillez à choisir un type d'instance censé répondre à vos futurs besoins en termes de ressources.
- Concentrez-vous sur le graphique des métriques pour déterminer si l'utilisation réelle est inférieure à la capacité d'instance. Vous pouvez également afficher les données de métriques (moyenne, pic, percentile) dans CloudWatch pour poursuivre l'évaluation de vos recommandations d'instances EC2. Par exemple, notez l'évolution des métriques de pourcentage d'UC pendant la journée et s'il y a des pics qui doivent être pris en compte. Pour de plus amples informations, veuillez consulter Affichage des métriques disponibles dans le Guide de l'utilisateur Amazon CloudWatch.
- Compute Optimizer peut fournir des recommandations pour les instances à capacité extensible, à savoir les instances T3, T3a et T2. Si vous dépassez régulièrement le niveau de base, assurez-vous que vous pouvez continuer à le faire en fonction des vCPU du nouveau type d'instance. Pour plus d'informations, consultez Crédits UC et utilisation de référence pour les instances à capacité extensible (p. 234).
- Si vous avez acheté une Instance réservée, votre instance à la demande peut être facturée au prix d'une Instance réservée. Avant de modifier votre type d'instance actuel, commencez par évaluer l'impact sur l'utilisation et la couverture de l'Instance réservée.
- Dans la mesure du possible, envisagez des conversions vers des instances de nouvelle génération.
- Lors de la migration vers une autre famille d'instances, assurez-vous que le type d'instance actuel et le nouveau type d'instance sont compatibles, en termes de virtualisation, d'architecture ou de type de réseau par exemple. Pour plus d'informations, consultez Compatibilité pour le redimensionnement des instances (p. 313).

- Enfin, tenez compte de la note de risque de performances fournie pour chaque recommandation. Le risque de performances correspond à l'effort que vous pourriez avoir à consacrer pour valider si le type d'instance recommandé répond aux exigences de performances de votre charge de travail. Nous recommandons également des tests rigoureux de charge et de performance avant et après toute modification.

Il existe d'autres considérations à prendre en compte lors du redimensionnement d'une instance EC2. Pour plus d'informations, consultez Modification du type d'instance (p. 313).

Ressources supplémentaires

- Types d'instance (p. 213)
- Guide de l'utilisateur AWS Compute Optimizer

Options d'achat d'instance

Amazon EC2 propose les options d'achat suivantes pour vous permettre d'optimiser vos coûts en fonction de vos besoins :

- Instances à la demande – Paiement à la seconde pour les instances que vous lancez.
- Plans d'économies – Réduisez vos coûts Amazon EC2 en vous engageant pour une utilisation continue, en USD par heure, pour une durée de 1 à 3 ans.
- Instances réservées – Réduisez vos coûts Amazon EC2 en vous engageant pour une configuration d'instance continue, incluant le type et la région, pour une durée de 1 à 3 ans.
- Instances planifiées – Achat d'instances qui sont toujours disponibles selon le planning récurent spécifié, pour une durée d'un an.
- Instances Spot – Elles vous permettent de demander des instances EC2 inutilisées, ce qui peut réduire considérablement vos coûts Amazon EC2.
- Hôtes dédiés – Paiement d'un hôte physique qui est entièrement dédié à l'exécution de vos instances et utilisation du modèle BYOL (Bring-Your-Own-License) pour vos licences logicielles par socket, par cœur ou par ordinateur virtuel afin de réduire les coûts.
- Instances dédiées – Paiement à l'heure pour les instances qui s'exécutent sur un matériel à client unique.
- Réservation de capacités – Réservez de la capacité pour vos instances EC2 dans une zone de disponibilité spécifique pendant la durée de votre choix.

Si vous avez besoin d'une réservation de capacité, achetez des Instances réservées ou des Réservations de capacité pour une zone de disponibilité spécifique, ou achetez des instances planifiées. Les Instances Spot constituent un choix économique si vous êtes flexible quant au moment où vos applications s'exécutent et à la possibilité de les interrompre. Les Hôtes dédiés ou les Instances dédiées peuvent vous aider à faire face aux exigences de conformité et à réduire les coûts en utilisant vos licences logicielles existantes liées au serveur. Pour en savoir plus, consultez la page Tarification Amazon EC2.

Pour de plus amples informations sur les plans d'économies, veuillez consulter le Guide de l'utilisateur des plans d'économies AWS.

Sommaire

- Détermination du cycle de vie de l'instance (p. 324)
- Instances à la demande (p. 325)
- Instances réservées (p. 329)
- Instances réservées planifiées (p. 369)
- Instances Spot (p. 373)

- Hôtes dédiés (p. 465)
- Instances dédiées (p. 498)
- Réservations de capacité à la demande (p. 504)

Détermination du cycle de vie de l'instance

Le cycle de vie d'une instance démarre au lancement de l'instance et prend fin à sa résiliation. L'option d'achat que vous choisissez affecte le cycle de vie de l'instance. Par exemple, une instance à la demande s'exécute lorsque vous la lancez et prend fin lorsque vous la résiliez. Une instance Spot s'exécute aussi longtemps que la capacité est disponible et que le prix maximum de votre offre est supérieur au prix Spot. Vous pouvez lancer une instance planifiée pendant sa durée planifiée ; Amazon EC2 lance les instances et les termine trois minutes avant que la période ne prenne fin.

Utilisez la procédure suivante pour déterminer le cycle de vie d'une instance.

Nouvelle console

Pour déterminer le cycle de vie d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Sous l'onglet Détails, sous Détails de l'instance, recherchez Cycle de vie. Si la valeur est `spot`, l'instance est une instance Spot. Si la valeur est `scheduled`, l'instance est une Instance réservée planifiée. Si la valeur est `normal`, l'instance est une instance à la demande ou une Instance réservée.
5. Sous l'onglet Détails, sous Hôte et groupe de placement, recherchez Locataire. Si la valeur est `host`, l'instance s'exécute sur un Hôte dédié. Si la valeur est `dedicated`, l'instance est une Instance dédiée.
6. (Facultatif) Si vous avez acheté une Instance réservée et que vous voulez vérifier qu'elle est appliquée, vous pouvez consulter les rapports d'utilisation pour Amazon EC2. Pour plus d'informations, consultez Rapports d'utilisation d'Amazon EC2 (p. 1324).

Ancienne console

Pour déterminer le cycle de vie d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Sous l'onglet Description, recherchez Tenancy. Si la valeur est `host`, l'instance s'exécute sur un Hôte dédié. Si la valeur est `dedicated`, l'instance est une Instance dédiée.
5. Sous l'ongletDescription, recherchez Lifecycle. Si la valeur est `spot`, l'instance est une instance Spot. Si la valeur est `scheduled`, l'instance est une instance planifiée. Si la valeur est `normal`, l'instance est une instance à la demande ou une Instance réservée.
6. (Facultatif) Si vous avez acheté une Instance réservée et que vous voulez vérifier qu'elle est appliquée, vous pouvez consulter les rapports d'utilisation pour Amazon EC2. Pour plus d'informations, consultez Rapports d'utilisation d'Amazon EC2 (p. 1324).

Pour déterminer le cycle de vie d'une instance à l'aide de l'AWS CLI

Utilisez la commande describe-instances suivante :

```
aws ec2 describe-instances --instance-ids i-1234567890abcdef0
```

Si l'instance s'exécute sur un Hôte dédié, la sortie contient les informations suivantes :

```
"Tenancy": "host"
```

Si l'instance est une Instance dédiée, la sortie contient les informations suivantes :

```
"Tenancy": "dedicated"
```

Si l'instance est une instance Spot, la sortie contient les informations suivantes :

```
"InstanceLifecycle": "spot"
```

Si l'instance est une instance planifiée, la sortie contient les informations suivantes :

```
"InstanceLifecycle": "scheduled"
```

Sinon, la sortie ne contient pas `InstanceLifecycle`.

Instances à la demande

Avec Instances à la demande, vous payez la capacité de calcul à la seconde à l' sans engagement à long terme. Vous bénéficiez d'un contrôle complet sur son cycle de vie : vous décidez quand la lancer, l'arrêter, la mettre en veille prolongée, la démarrer, la redémarrer ou la résilier.

Aucun engagement à long terme n'est requis lorsque vous achetez des Instances à la demande. Vous payez uniquement pour les secondes pendant lesquelles vos Instances à la demande sont à l'état `running`. Le prix par seconde pour une instance à la demande en cours d'exécution est fixe. Il figure sur la page Tarification Amazon EC2, Tarification à la demande.

Nous vous recommandons d'utiliser des Instances à la demande pour les applications avec des charges de travail irrégulières à court terme qui ne peuvent pas être interrompues.

Pour réaliser des économies importantes par rapport à Instances à la demande, utilisez Plans d'économies, Instances Spot (p. 373) ou Instances réservées (p. 329).

Table des matières

- Utilisation de Instances à la demande (p. 325)
- Limites instance à la demande (p. 326)
    - Calcul du nombre de vCPU dont vous avez besoin (p. 326)
    - Demande d'une augmentation de limite (p. 328)
    - Surveillance des limites et de l'utilisation d'instance à la demande (p. 328)
- Interrogation des prix des services AWS (p. 328)

Utilisation de Instances à la demande

Vous pouvez utiliser les Instances à la demande de l'une des façons suivantes :

- Lancer votre instance (p. 529)
- Connectez-vous à votre instance Linux (p. 598)
- Arrêt et démarrage de votre instance (p. 625)

- Mettre votre instance Linux en veille prolongée (p. 628)
- Redémarrer votre instance (p. 641)
- Mise hors service d'instance (p. 642)
- Résilier une instance (p. 645)
- Récupération de votre instance (p. 652)
- Personnalisation de votre instance Amazon Linux (p. 653)
- Identification des instances EC2 Linux (p. 736)

Si vous débutez avec Amazon EC2, consultez Comment démarrer avec Amazon EC2 (p. 1).

Limites instance à la demande

Il y a une limite au nombre d'Instances à la demande exécutés par compte AWS et par région. Les limites d'instance à la demande sont gérées en termes de nombre d'unités centrales virtuelles (vCPU) que les Instances à la demande en cours d'exécution utilisent quel que soit le type d'instance.

Il existe six limites d'instance à la demande, répertoriées dans le tableau suivant. Chaque limite spécifie la limite de vCPU pour une ou plusieurs familles d'instances. Pour plus d'informations sur les différentes familles, générations et tailles d'instances, consultez Types d'instance Amazon EC2.

Nom de la limite d'instance à la demande	Limite de vCPU par défaut
Toutes les instances standard à la demande (A, C, D, H, I, M, R, T, Z) en cours d'exécution	1152 vCPU
Toutes les instances F à la demande en cours d'exécution	128 vCPU
Toutes les instances G à la demande en cours d'exécution	128 vCPU
Toutes les instances Inf à la demande en cours d'exécution	128 vCPU
Toutes les instances P à la demande en cours d'exécution	128 vCPU
Toutes les instances X à la demande en cours d'exécution	128 vCPU

> **Note**
>
> Les nouveaux comptes AWS peuvent commencer avec des limites inférieures à celles présentées ici.

Avec les limites de vCPU, vous pouvez utiliser votre limite en fonction du nombre de vCPU nécessaires pour lancer toute combinaison de types d'instance qui répond à l'évolution de vos besoins en termes d'applications. Par exemple, avec une limite d'instances Standard de 256 vCPU, vous pouvez lancer 32 instances `m5.2xlarge` (32 x 8 vCPU) ou 16 instances `c5.4xlarge` (16 x 16 vCPU), ou une combinaison de n'importes quels types et tailles d'instances Standard pour un total de 256 vCPU. Pour plus d'informations, consultez Limites d'instance à la demande EC2.

Calcul du nombre de vCPU dont vous avez besoin

Vous pouvez utiliser le calculateur de limites de vCPU pour déterminer le nombre de vCPUs dont vous avez besoin pour vos applications.

Lorsque vous utilisez le calculateur, gardez ce qui suit à l'esprit : celui-ci considère que vous avez atteint votre limite actuelle. La valeur que vous entrez pour Instance Count (Nombre d'instances) correspond au nombre d'instances que vous devez lancer en plus de ce qui est autorisé par votre limite actuelle. Le calculateur ajoute votre limite actuelle à la valeur de Instance Count (Nombre d'instances) pour arriver à une nouvelle limite.

La capture d'écran suivante montre le calculateur de limites de vCPU.

Vous pouvez afficher et utiliser les informations et contrôles suivants :

- Instance type (Type d'instance) – Types d'instance que vous pouvez ajouter au calculateur de limites de vCPU.
- Instance count (Nombre d'instances) – Nombre d'instances dont vous avez besoin pour le type d'instance sélectionné.
- vCPU count (Nombre de vCPU) – Nombre de vCPU correspondant à Instance count (Nombre d'instances).
- Current limit (Limite actuelle) – Votre limite actuelle pour le type de limite auquel le type d'instance appartient. La limite s'applique à tous les types d'instance du même type de limite. Par exemple, dans la capture d'écran précédente, la limite actuelle pour `m5.2xlarge` et `c5.4xlarge` est de 1 920 vCPU, ce qui est la limite d'instances appartenant à la limite d'instances All Standard (Toutes Standard).
- New limit (Nouvelle limite) – Nouvelle limite, en nombre de vCPU, calculée en ajoutant la valeur de vCPU count (Nombre de vCPU) et celle de Current limit (Limite actuelle).
- X – Sélectionnez le X pour supprimer la ligne.
- Add instance type (Ajouter un type d'instance) – Choisissez Add instance type (Ajouter un type d'instance) pour ajouter un autre type d'instance au calculateur.
- Limits calculation (Calcul des limites) – Affiche la limite actuelle, les vCPU nécessaires et la nouvelle limite pour les types de limite.
  - Instance limit name (Nom de la limite d'instances) – Type de limite pour les types d'instance que vous avez sélectionnés.
  - Current limit (Limite actuelle) – Limite actuelle pour le type de limite.
  - vCPUs needed (vCPU nécessaires) – Nombre de vCPU correspondant au nombre d'instances que vous avez spécifié dans Instance count (Nombre d’instances). Pour le type de limite d'instances All Standard (Toutes Standard), le nombre de vCPU nécessaires est calculé en ajoutant les valeurs de vCPU count (Nombre de vCPU) pour tous les types d'instance de ce type de limite.

- New limit (Nouvelle limite) – La nouvelle limite est calculée en ajoutant les valeurs de Current limit (Limite actuelle) et vCPUs needed (vCPU nécessaires).
- Options – Choisissez Request limit increase afin de demander une augmentation de limite pour le type de limite correspondant.

Pour calculer le nombre de vCPU nécessaires

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez une région.
3. Dans le navigateur de gauche, choisissez Restrictions.
4. Choisissez Calculate vCPU limit (Calculer la limite de vCPU).
5. Choisissez Add instance type (Ajouter un type d'instance), sélectionnez le type d'instance requis, puis spécifiez le nombre d'instances nécessaires. Pour ajouter d'autres types d'instances, choisissez à nouveau Add instance type (Ajouter un type d'instance).
6. Affichez Limits calculation (Calcul des limites) pour la nouvelle limite requise.
7. Quand vous avez fini d'utiliser le calculateur, choisissez Close (Fermer).

Demande d'une augmentation de limite

Vous pouvez demander une augmentation de limite pour chaque type de limite d'instance à la demande sur la page des limites ou dans le calculateur de limites de vCPU dans la console Amazon EC2. Complétez les champs requis dans le formulaire d'augmentation de limite du Centre AWS Support avec votre cas d'utilisation. Pour Primary Instance Type (Type d'instance principal), sélectionnez le type de limite qui correspond au Instance limit name (Nom de la limite d'instances) dans le calculateur de limites de vCPU. Pour la valeur de nouvelle limite, utilisez la valeur qui apparaît dans la colonne New limit (Nouvelle limite) du calculateur de limites de vCPU. Pour plus d'informations sur comment demander une augmentation de limite, consultez Quotas de service Amazon EC2 (p. 1322).

Surveillance des limites et de l'utilisation d'instance à la demande

Vous pouvez afficher et gérer vos limites instance à la demande à l'aide des éléments suivants :

- La page Limites de la console Amazon EC2
- La page Quotas de service Amazon EC2 dans la console Quotas de service
- L'AWS CLI get-service-quota
- La page Limites de service dans la console AWS Trusted Advisor

Pour de plus amples informations, veuillez consulter Quotas de service Amazon EC2 (p. 1322) dans le Amazon EC2 Guide de l'utilisateur pour les instances Linux, Affichage d'un quota de service dans le guide de l'utilisateur Service Quotas et AWS Trusted Advisor.

Avec l'intégration des métriques Amazon CloudWatch, vous pouvez surveiller votre utilisation d'EC2 par rapport aux limites. Vous pouvez également configurer des alarmes pour vous avertir lorsque vous approchez des limites. Pour de plus amples informations, veuillez consulter Utilisation d'alarmes Amazon CloudWatch dans le guide de l'utilisateur Service Quotas.

Interrogation des prix des services AWS

Vous pouvez utiliser l'API du service de liste des prix ou l'API de liste de prix AWS pour interroger les prix d'Instances à la demande. Pour plus d'informations, consultez Utilisation de l'API de liste des prix AWS dans le AWS Billing and Cost Management Guide de l'utilisateur.

Instances réservées

Les Instances réservées vous permettent de réaliser des économies importantes sur vos coûts Amazon EC2 par rapport à la tarification instance à la demande. Les Instances réservées ne sont pas des instances physiques, mais plutôt une remise de facturation appliquée à l'utilisation de Instances à la demande dans votre compte. Ces Instances à la demande doivent correspondre à certains attributs, comme le type et la région de l'instance, afin d'entraîner une remise de facturation.

Les plans d'économies permettent également de réaliser des économies importantes sur vos coûts Amazon EC2 par rapport aux tarifs instance à la demande. Avec les plans d'économies, vous vous engagez pour une utilisation continue, mesurée en USD par heure. Cela vous donne la flexibilité d'utiliser les configurations d'instances répondant le mieux à vos besoins et de continuer à économiser de l'argent au lieu de vous engager pour une configuration d'instance spécifique. Pour de plus amples informations, veuillez consulter le Guide de l'utilisateur des plans d'économies AWS.

Rubriques Instances réservées

- Présentation de Instance réservée (p. 329)
- Variables clés déterminant la tarification d'une Instance réservée (p. 330)
- Limites d'une Instance réservée (p. 331)
- Instances réservées régionales et zonales (portée) (p. 332)
- Types d'Instances réservées (classes d'offres) (p. 333)
- Application des Instances réservées (p. 334)
- Principes de facturation (p. 339)
- Achat d'Instances réservées (p. 344)
- Place de marché des instances réservées (p. 350)
- Modification des Instances réservées (p. 357)
- Échange d'Instances réservées convertibles (p. 365)

Présentation de Instance réservée

Le schéma suivant montre une vue d'ensemble élémentaire de l'achat et de l'utilisation des Instances réservées.

Dans ce scénario, vous disposez dans votre compte d'une instance à la demande (T2) en cours d'exécution, qui vous est facturée au tarif à la demande. Vous achetez une Instance réservée qui correspond aux attributs de votre instance en cours d'exécution et l'avantage de facturation est immédiatement appliqué. Ensuite, vous achetez une Instance réservée pour une instance C4. Aucune instance en cours d'exécution dans votre compte ne correspond aux attributs de cette Instance réservée. Dans la dernière étape, vous lancez une instance qui correspond aux attributs de l'Instance réservée C4 et l'avantage de facturation est immédiatement appliqué.

Variables clés déterminant la tarification d'une Instance réservée

La tarification de Instance réservée est déterminée par les variables clés suivantes.

Attributs d'instance

Une Instance réservée dispose de quatre attributs d'instance qui déterminent son prix.

- Type d'instance : par exemple, `m4.large`. Il est composé de la famille d'instance (par exemple, `m4`) et de la taille de l'instance (par exemple, `large`).
- Région : Région dans laquelle l'Instance réservée a été achetée.
- Location : si votre instance est exécutée sur un matériel partagé (par défaut) ou à client unique (dédié). Pour plus d'informations, consultez Instances dédiées (p. 498).
- Plateforme : le système d'exploitation ; par exemple, Windows ou Linux/Unix. Pour plus d'informations, consultez Sélection d'une plateforme (p. 344).

Engagement de durée

Vous pouvez acheter une Instance réservée pour un engagement d'un ou de trois ans, avec une remise plus importante pour l'engagement de trois ans.

- Un an : un an correspond à 31536000 secondes (365 jours).

- Trois ans : trois ans correspondent à 94608000 secondes (1095 jours).

Les Instances réservées ne sont pas renouvelées automatiquement. Lorsqu'elles expirent, vous pouvez continuer à utiliser l'instance EC2 sans interruption, mais elle est facturée aux tarifs à la demande. Dans l'exemple ci-dessus, lorsque les Instances réservées qui couvrent les instances T2 et C4 expirent, les tarifs à la demande vous sont à nouveaux appliqués jusqu'à ce que vous mettiez les instances hors service ou que vous achetiez de nouvelles Instances réservées qui correspondent aux attributs de l'instance.

Options de paiement

Les options de paiement suivantes sont disponibles pour les Instances réservées :

- Tous les frais initiaux : le paiement est effectué en totalité au début de la période, sans aucun autre coût ou frais horaires supplémentaires pour le reste de la réservation, quel que soit le nombre d'heures utilisé.
- Frais initiaux partiels : une partie du coût doit être payée au départ et les heures restantes pendant la période sont facturées à un tarif horaire réduit, que la Instance réservée soit utilisée ou non.
- Sans frais initiaux : vous devez régler un taux horaire avec remise pour chaque heure entrant dans le cadre de l'abonnement, que la Instance réservée soit utilisée ou non. Aucun paiement initial n'est requis.

  Note

  Les Instances réservées sans frais initiaux sont basées sur une obligation contractuelle d'effectuer des paiements mensuels pendant toute la durée de la réservation. C'est la raison pour laquelle il est nécessaire de fournir un bon historique de facturation pour pouvoir acheter des Instances réservées sans frais initiaux.

En règle générale, l'option la plus économique consiste à acheter des Instances réservées en versant un paiement initial plus élevé. Vous pouvez aussi trouver des Instances réservées proposées par des vendeurs tiers à des prix inférieurs avec des durées de paiement plus courtes sur la Place de marché des instances réservées. Pour plus d'informations, consultez Place de marché des instances réservées (p. 350).

Classe d'offre

Si vos besoins informatiques évoluent, vous pourrez probablement modifier ou échanger votre Instance réservée, en fonction de la classe d'offre.

- Standard : proposent la réduction la plus importante, mais ne peut que se modifier.
- Convertible : proposent une réduction plus faible que les Instances réservées Standard, mais peut s'échanger contre une Instance réservée convertible avec différents attributs d'instance. Les Instances réservées convertibles peuvent également être modifiées.

Pour plus d'informations, consultez Types d'Instances réservées (classes d'offres) (p. 333).

Une fois que vous avez acheté une Instance réservée, vous ne pouvez pas annuler votre achat. Toutefois, vous pourrez probablement modifier (p. 357), échanger (p. 365) ou vendre (p. 350) votre Instance réservée si vos besoins évoluent.

Pour de plus informations, veuillez consulter la page relative à la tarification des instances réservées Amazon EC2.

Limites d'une Instance réservée

Il existe une limite au nombre d'Instances réservées que vous pouvez acheter par mois. Pour chaque région, vous pouvez acheter 20 Instances réservées régionales (p. 334) par mois, et 20 Instances réservées zonales (p. 334) supplémentaires par mois pour chaque zone de disponibilité.

Par exemple, dans une région comportant trois zones de disponibilité, la limite est de 80 Instances réservées par mois : 20 Instances réservées régionales pour la région et 20 Instances réservées zonales pour chacune des trois zones de disponibilité (20 x 3 = 60).

Une Instance réservée régionale applique une remise à une instance à la demande en cours d'exécution. La limite d'instance à la demande par défaut est de 20. Vous ne pouvez pas dépasser votre limite d'instance à la demande en cours d'exécution en achetant des Instances réservées régionales. Par exemple, si vous avez déjà 20 Instances à la demande en cours d'exécution, et que vous achetez 20 Instances réservées régionales, les 20 Instances réservées régionales sont utilisées pour appliquer une remise aux 20 Instances à la demande en cours d'exécution. Si vous achetez plus d'Instances réservées régionales, vous ne pourrez pas lancer plus d'instances parce que vous avez atteint votre limite d'instance à la demande.

Avant d'acheter des Instances réservées régionales, assurez-vous que votre limite d'instance à la demande atteint ou dépasse le nombre d'Instances réservées régionales que vous comptez posséder. Si nécessaire, assurez-vous de demander une augmentation de votre limite d'instance à la demande avant d'acheter des Instances réservées régionales supplémentaires.

Une Instance réservée zonale, c'est-à-dire une Instance réservée achetée pour une zone de disponibilité spécifique, offre une réservation de capacité ainsi qu'une remise. Vous pouvez dépasser votre limite d'instance à la demande en cours d'exécution en achetant des Instances réservées zonales. Par exemple, si vous avez déjà 20 Instances à la demande en cours d'exécution et que vous achetez 20 Instances réservées zonales, vous pouvez lancer 20 Instances à la demande supplémentaires qui correspondent aux spécifications de vos Instances réservées zonales, ce qui vous donne un total de 40 instances en cours d'exécution.

La console Amazon EC2 indique des informations relatives aux limites. Pour plus d'informations, consultez Affichages de vos limites actuelles (p. 1322).

Instances réservées régionales et zonales (portée)

Lorsque vous achetez une Instance réservée, vous déterminez la portée de la Instance réservée. La portée est régionale ou zonale.

- Régionale : lorsque vous achetez une Instance réservée pour une région, elle est appelée Instance réservée régionale.
- Zonale : lorsque vous achetez une Instance réservée pour une Zone de disponibilité spécifique, il s'agit d'une Instance réservée zonale.

Différences entre les Instances réservées régionales et zonales

Le tableau suivant souligne certaines différences essentielles entre les Instances réservées zonales et les Instances réservées régionales :

	Instances réservées régionales	Instances réservées zonales
Flexibilité des zones de disponibilité	La remise de Instance réservée s'applique à l'utilisation d'une instance dans n'importe quelle zone de disponibilité de la région spécifiée.	Aucune flexibilité de zone de disponibilité—la remise de Instance réservée s'applique à l'utilisation d'instance uniquement dans la zone de disponibilité spécifiée.
Possibilité de réserver de la capacité	Une Instance réservée de région ne réserve pas de capacité.	Une Instance réservée de zone réserve de la capacité dans la zone de disponibilité spécifiée.

	Instances réservées régionales	Instances réservées zonales
Flexibilité de la taille de l'instance	La remise Instance réservée s'applique à une utilisation d'instance, quelle que soit la taille, au sein de cette famille d'instances. Prise en charge uniquement sur les Instances réservées Amazon Linux/Unix avec location par défaut. Pour plus d'informations, consultez Flexibilité de taille d'instance déterminée par le facteur de normalisation (p. 334).	Aucune flexibilité de taille d'instance—la remise de Instance réservée s'applique pour l'utilisation d'instance uniquement pour la taille et le type d'instance spécifiés.

Pour plus d'informations et d'exemples, consultez Application des Instances réservées (p. 334).

Types d'Instances réservées (classes d'offres)

Lorsque vous achetez une Instance réservée, vous avez le choix entre les classes d'offres Standard ou Convertible. La Instance réservée s'applique à une famille d'instance d'un seul type, plateforme, portée et location pendant une période donnée. Si vos besoins informatiques évoluent, vous pourrez probablement modifier ou échanger votre Instance réservée, en fonction de la classe d'offre. Des restrictions ou restrictions supplémentaires peuvent aussi s'appliquer aux classes d'offres.

Les différences entre les classes d'offres Convertible et Standard sont les suivantes.

Instance réservée standard	Instance réservée convertible
Certains attributs, tels que la taille d'instance, peuvent être modifiés pendant cette période ; toutefois, la famille d'instance ne peut pas être modifiée. Vous ne pouvez pas échanger une Instance réservée standard, mais seulement la modifier. Pour plus d'informations, consultez Modification des Instances réservées (p. 357).	Peut être échangée, pendant la période de paiement, contre une autre Instance réservée convertible avec de nouveaux attributs tels que la famille d'instance, le type d'instance, la plateforme, la portée ou la location. Pour plus d'informations, consultez Échange d'Instances réservées convertibles (p. 365). Vous pouvez également modifier certains attributs d'une Instance réservée convertible. Pour plus d'informations, consultez Modification des Instances réservées (p. 357).
Peut être vendue sur le Place de marché des instances réservées.	Ne peut pas être vendue sur le Place de marché des instances réservées.

Les Instances réservées convertibles et standard peuvent être achetées pour s'appliquer aux instances d'une Zone de disponibilité (Instances réservées zonale) spécifique ou à des instances d'une région (Instances réservées régionale). Pour plus d'informations et d'exemples, consultez Application des Instances réservées (p. 334).

Si vous voulez acheter des réservations de capacités récurrentes sur une base quotidienne, hebdomadaire ou mensuelle, une Instance réservée planifiée peut répondre à vos besoins. Pour plus d'informations, consultez Instances réservées planifiées (p. 369).

Application des Instances réservées

Si vous achetez une Instance réservée et que vous avez déjà une instance en cours d'exécution qui correspond aux attributs de l'Instance réservée, l'avantage de facturation est immédiatement appliqué. Vous n'avez pas besoin de redémarrer vos instances. Si vous n'avez pas d'instance en cours d'exécution éligible, lancez une instance et veillez à respecter les mêmes critères que ceux spécifiés pour l'Instance réservée. Pour plus d'informations, consultez Utilisation de vos Instances réservées (p. 350).

Les Instances réservées s'appliquent à un usage identique, quel que soit le type d'offre (standard ou convertible) et sont automatiquement appliquées aux Instances à la demande en cours d'exécution avec des attributs correspondants.

Application des Instances réservées zonales

Les Instances réservées appliquées à une zone de disponibilité spécifique fournissent une remise d'Instance réservée pour l'utilisation d'instance correspondante dans cette zone de disponibilité. Par exemple, si vous achetez deux Instances réservées standard Linux/Unix à location par défaut `c4.xlarge` dans la zone de disponibilité us-east-1a, jusqu'à deux instances Linux/Unix à location par défaut `c4.xlarge` s'exécutant dans la zone de disponibilité us-east-1a peuvent bénéficier de la remise accordée aux Instance réservée. Les attributs (location, plateforme, zone de disponibilité, type d'instance et taille d'instance) des instances en cours d'exécution doivent correspondre à celles des Instances réservées.

Application des Instances réservées régionales

Les Instances réservées régionales sont achetées pour une région et assurent une flexibilité de Zone de disponibilité. La remise de Instance réservée s'applique à l'utilisation d'une instance dans n'importe quelle zone de disponibilité de la région spécifiée.

Les Instances réservées régionales assurent également une flexibilité d'instance où la remise Instance réservée s'applique à l'utilisation d'instance dans la famille d'instance, peu importe la taille.

Limites de la flexibilité de taille d'instance.

La flexibilité de taille d'instance ne s'applique pas aux Instances réservées suivantes :

- Les Instances réservées achetées pour une Zone de disponibilité spécifique (Instances réservées zonales)
- Instances réservées avec location dédiée
- Instances réservées pour Windows Server, Windows Server avec SQL Standard, Windows Server avec SQL Server Enterprise, Windows Server avec SQL Server Web, RHEL et SUSE Linux Enterprise Server
- Instances réservées pour les instances G4

Flexibilité de taille d'instance déterminée par le facteur de normalisation

La flexibilité de la taille d'instance est déterminée par le facteur de normalisation de la taille d'instance. La remise s'applique complètement ou partiellement aux instances en cours d'exécution d'une même famille d'instance, en fonction de la taille d'instance de la réservation, dans n'importe quelle zone de disponibilité de la région. Les seuls attributs qui doivent correspondre sont la famille d'instance, la location et la plate-forme.

La flexibilité de la taille d'instance est appliquée de la taille d'instance la plus petite à la taille d'instance la plus grande au sein de la famille d'instance, en fonction du facteur de normalisation.

Le tableau suivant décrit les différentes tailles au sein d'une famille d'instances et le facteur de normalisation correspondant par heure. Cette échelle est utilisée pour appliquer le taux avec remise des Instances réservées à l'utilisation normalisée de la famille d'instance.

Taille d'instance	Facteur de normalisation
nano	0.25
micro	0.5
small	1
medium	2
large	4
xlarge	8
2xlarge	16
3xlarge	24
4xlarge	32
6xlarge	48
8xlarge	64
9xlarge	72
10xlarge	80
12xlarge	96
16xlarge	128
18xlarge	144
24xlarge	192
32xlarge	256

Par exemple, le facteur de normalisation d'une instance `t2.medium` est 2. Si vous achetez une Instance réservée Amazon Linux/Unix `t2.medium` à location par défaut dans la région USA Est (Virginie du Nord) et que vous avez deux instances `t2.small` en cours d'exécution dans votre compte dans cette région, l'avantage de facturation est appliqué entièrement à ces deux instances.

Si vous avez une instance `t2.large` en cours d'exécution dans votre compte dans la région USA Est (Virginie du Nord), l'avantage de facturation est appliqué à 50 % de l'utilisation de l'instance.

Le facteur de normalisation est également appliqué lors de la modification d'Instances réservées standard. Pour plus d'informations, consultez Modification des Instances réservées (p. 357).

Facteur de normalisation pour les instances matériel nu

La flexibilité de taille d'instance s'applique également aux instances à matériel nu dans la famille d'instances. Si vous disposez de Instances réservées Amazon Linux/Unix régionales avec une location partagée sur des instances à matériel nu, vous pouvez profiter des économies Instance réservée avec la même famille d'instances. L'inverse est également vrai : si vous disposez de Instances réservées Amazon Linux/Unix régionales avec une location partagée sur des instances de la même famille que l'instance à matériel nu, vous pouvez profiter des économies Instance réservée sur l'instance à matériel nu.

Une instance à matériel nu fait la même taille que l'instance plus large dans la même famille d'instances. Par exemple, une `i3.metal` fait toujours la même taille qu'une `i3.16xlarge`. Elles disposent donc du même facteur de normalisation.

> **Note**
>
> Les tailles d'instances `.metal` ne disposent pas d'un seul et unique facteur de normalisation. Elles varient selon la famille d'instance spécifique.

Taille d'instance à matériel nu	Facteur de normalisation
`a1.metal`	32
`c5.metal`	192
`c5d.metal`	192
`c5n.metal`	144
`c6g.metal`	128
`c6gd.metal`	128
`g4dn.metal`	128
`i3.metal`	128
`i3en.metal`	192

Taille d'instance à matériel nu	Facteur de normalisation
`m5.metal`	192
`m5d.metal`	192
`m6g.metal`	128
`m6gd.metal`	128
`r5.metal`	192
`r5d.metal`	192
`r6g.metal`	128
`r6gd.metal`	128
`z1d.metal`	96

Par exemple, une instance `i3.metal` dispose d'un facteur de normalisation de 128. Si vous achetez un Instance réservée Amazon Linux/Unix à location par défaut `i3.metal` dans la USA Est (Virginie du Nord), l'avantage de facturation peut s'appliquer comme suit :

- Si vous disposez d'une `i3.16xlarge` en cours d'exécution dans votre compte pour cette région, l'avantage de facturation peut s'appliquer entièrement à l'instance `i3.16xlarge` (facteur de normalisation `i3.16xlarge` = 128).
- Sinon, si vous disposez de deux instances `i3.8xlarge` en cours d'exécution dans votre compte pour cette région, l'avantage de facturation peut s'appliquer entièrement aux deux instances `i3.8xlarge` (facteur de normalisation `i3.8xlarge` = 64).
- Sinon, si vous disposez de quatre instances `i3.4xlarge` en cours d'exécution dans votre compte pour cette région, l'avantage de facturation peut s'appliquer entièrement aux quatre instances `i3.4xlarge` (facteur de normalisation `i3.4xlarge` = 32).

L'inverse est également vrai. Par exemple, si vous achetez deux Instances réservées Amazon Linux/Unix à location par défaut `i3.8xlarge` dans la USA Est (Virginie du Nord) et que vous disposez d'une instance `i3.metal` dans cette région, l'avantage de facturation s'applique entièrement à l'instance `i3.metal`.

Exemples d'application des Instances réservées

Les scénarios suivants couvrent les façons dont les Instances réservées sont appliquées.

Example Scénario 1 : Instances réservées dans un compte unique

Vous exécutez les Instances à la demande suivantes dans le compte A :

- 4 instances `m3.large` Linux à location par défaut dans la zone de disponibilité us-east-1a
- 2 instances `m4.xlarge` Amazon Linux à location par défaut dans la zone de disponibilité us-east-1b
- 1 instance Amazon Linux `c4.xlarge` à location par défaut dans la zone de disponibilité us-east-1c

Vous achetez ensuite les Instances réservées suivantes dans le compte A :

- 4 Instances réservées `m3.large` Linux à location par défaut dans la zone de disponibilité us-east-1a (la capacité est réservée)
- 4 Instances réservées Amazon Linux `m4.large` à location par défaut dans la région us-east-1
- 1 Instances réservées Amazon Linux `c4.large` à location par défaut dans la région us-east-1

Les avantages de l'Instance réservée sont appliqués de la façon suivante :

- La remise et la réservation de capacité des quatre Instances réservées zonales `m3.large` sont utilisées par les quatre instances `m3.large`, car leurs attributs (taille de l'instance, région, plateforme, location) correspondent.
- Les Instances réservées régionales `m4.large` fournissent une flexibilité de zone de disponibilité et de taille d'instance, car il s'agit d'Instances réservées Amazon Linux régionales à location par défaut.

  Une instance `m4.large` est équivalente à 4 unités normalisées/heure.

  Vous avez acheté quatre Instances réservées régionales `m4.large` et, au total, celles-ci sont égales à 16 unités normalisées/heure (4x4). Le compte A comporte deux instances `m4.xlarge` en cours d'exécution, ce qui est équivalant à 16 unités normalisées/heure (2x8). Dans ce cas, les quatre Instances réservées régionales `m4.large` apportent l'avantage de facturation d'une heure d'utilisation complète de deux instances `m4.xlarge`.
- L'Instance réservée régionale `c4.large` dans la région us-east-1 fournit une flexibilité de zone de disponibilité et de taille d'instance, car il s'agit d'une Instance réservée régionale Amazon Linux à location par défaut et elle s'applique à l'instance `c4.xlarge`. Une instance `c4.large` est équivalente à 4 unités normalisées/heure et une instance `c4.xlarge` est équivalente à 8 unités normalisées/heure.

  Dans ce cas, l'Instance réservée régionale `c4.large` apporte un avantage partiel à l'utilisation de `c4.xlarge`. Cela est dû au fait qu'une Instance réservée `c4.large` est équivalente à 4 unités normalisées/heure d'utilisation, mais qu'une instance `c4.xlarge` requiert 8 unités normalisées/heure. Par conséquent, la remise de facturation de l'Instance réservée `c4.large` s'applique à 50 % de l'utilisation de `c4.xlarge`. L'utilisation `c4.xlarge` restante est facturée au taux à la demande.

Example Scénario 2 : Instances réservées régionales dans des comptes liés

Les Instances réservées sont d'abord appliquées à une utilisation au sein du compte d'achat, puis à l'utilisation éligible dans tout autre compte au sein de l'organisation. Pour plus d'informations, consultez Instances réservées et la facturation consolidée (p. 341). Pour les Instances réservées régionales qui offrent la flexibilité de la taille d'instance, l'avantage est appliqué de la taille d'instance la plus petite à la taille d'instance la plus grande au sein de la famille d'instance.

Vous exécutez les Instances à la demande suivantes dans le compte A (le compte d'achat) :

- 2 instances `m4.xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1a
- 1 instances `m4.2xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1b
- 2 instances `c4.xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1a
- 1 instances `c4.2xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1b

Un autre client exécute les Instances à la demande suivantes dans le compte B (un compte lié) :

- 2 instances `m4.xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1a

Vous achetez ensuite les Instances réservées régionales suivantes dans le compte A :

- 4 Instances réservées Linux `m4.xlarge` à location par défaut dans la région us-east-1
- 2 Instances réservées Linux `c4.xlarge` à location par défaut dans la région us-east-1

Les avantages de l'Instance réservée régionale sont appliqués de la façon suivante :

- La remise des quatre Instances réservées `m4.xlarge` est utilisée par les deux instances `m4.xlarge` et par l'instance `m4.2xlarge` unique dans le compte A (compte d'achat). Les trois instances correspondent

toutes aux attributs (famille d'instance, région, plate-forme, location). La remise s'applique d'abord aux instances dans le compte d'achat (compte A), même si le compte B (compte lié) dispose de deux `m4.xlarge` qui correspondent également aux Instances réservées. Il n'y a pas de réservation de capacité, car les Instances réservées sont des Instances réservées régionales.

- La remise des deux Instances réservées `c4.xlarge` s'applique aux deux instances `c4.xlarge`, car elles ont une taille d'instance plus petite que l'instance `c4.2xlarge`. Il n'y a pas de réservation de capacité, car les Instances réservées sont des Instances réservées régionales.

Example Scénario 3 : Instances réservées zonales dans un compte lié

En général, les Instances réservées appartenant à un compte sont appliquées en premier à l'utilisation dans ce compte. Cependant, s'il existe des Instances réservées éligibles non utilisées pour une zone de disponibilité spécifique (Instances réservées zonales) dans d'autres comptes de l'organisation, elles sont appliquées au compte avant les Instances réservées régionales appartenant au compte. Ceci vise à garantir une utilisation maximale des Instance réservée et une facture moins élevée. A des fins de facturation, tous les comptes de l'organisation sont traités comme s'il s'agissait d'un seul compte. L'exemple suivant peut contribuer à en apporter l'explication.

Vous exécutez l'instance à la demande suivante dans le compte A (le compte d'achat) :

- 1 instance `m4.xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1a

Un client exécute l'instance à la demande suivante dans le compte B lié :

- 1 instance `m4.xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1b

Vous achetez ensuite les Instances réservées régionales suivantes dans le compte A :

- 1 Instance réservée Linux `m4.xlarge` à location par défaut dans la région us-east-1

Un client achète également les Instances réservées zonales suivantes dans le compte C lié :

- 1 Instances réservées `m4.xlarge` Linux à location par défaut dans la zone de disponibilité us-east-1a

Les avantages de l'Instance réservée sont appliqués de la façon suivante :

- La remise de l'Instance réservée zonale `m4.xlarge` appartenant au compte C est appliquée à l'utilisation de `m4.xlarge` dans le compte A.
- La remise de l'Instance réservée régionale `m4.xlarge` appartenant au compte A est appliquée à l'utilisation de `m4.xlarge` dans le compte B.
- Si l'Instance réservée régionale appartenant au compte A est appliquée d'abord à l'utilisation dans le compte A, l'Instance réservée zonale appartenant au compte C reste inutilisée et l'utilisation dans le compte B est facturée aux tarifs à la demande.

Pour plus d'informations, consultez Instances réservées dans le rapport Billing and Cost Management.

Principes de facturation

Toutes les Instances réservées vous permettent de bénéficier d'une remise par rapport à la tarification à la demande. Avec les Instances réservées, vous payez pour toute la durée de l'abonnement et non en fonction de l'utilisation réelle. Vous pouvez choisir d'effectuer un paiement initial ou un paiement initial partiel, ou mensuel pour votre Instance réservée, en fonction de l'option de paiement (p. 331) spécifiée pour l'Instance réservée.

Lorsque des Instances réservées expirent, le tarif à la demande est facturé pour l'utilisation d'instance EC2. Vous pouvez mettre en file d'attente l'achat d'une Instance réservée jusqu'à trois ans en avance. Cela peut vous aider à garantir une couverture ininterrompue. Pour de plus amples informations, veuillez consulter Mise en file d'attente de votre achat (p. 345).

L'offre gratuite AWS est disponible pour les nouveaux comptes AWS. Si vous avez recours à l'offre gratuite AWS pour exécuter les instances Amazon EC2 et que vous achetez ensuite une Instance réservée, vous serez facturé conformément aux consignes de tarification standard. Pour plus d'informations, consultez Offre gratuite AWS.

Sommaire

- Facturation de l'utilisation (p. 340)
- Affichage d'une facture (p. 341)
- Instances réservées et la facturation consolidée (p. 341)
- Niveaux de tarification avec remise d'Instance réservée (p. 342)

Facturation de l'utilisation

Les Instances réservées sont facturées toutes les heures d'horloge au cours de la réservation sélectionnée, que l'instance soit exécutée. Chaque heure d'horloge commence à l'heure (zéro minute et zéro seconde après l'heure) d'une horloge standard de 24 heures. Par exemple, 1:00:00 à 1:59:59 est une heure horloge. Pour plus d'informations sur les états de l'instance, consultez Cycle de vie d'une instance (p. 524).

L'avantage de facturation d'une Instance réservée peut être appliqué à une instance en cours d'exécution sur une base par seconde. La facturation par seconde est disponible pour les instances qui utilisent une distribution Linux en open source, telle que Amazon Linux et Ubuntu. La facturation par heure est utilisée pour les distributions Linux commerciales, telles que Red Hat Enterprise Linux et SUSE Linux Enterprise Server.

L'avantage de facturation d'une Instance réservée peut s'appliquer à un maximum de 3 600 secondes (une heure) d'utilisation d'instance par heure d'horloge. Vous pouvez exécuter plusieurs instances simultanément, mais vous ne pouvez bénéficier de l'avantage de la remise d'Instance réservée que pour un total de 3600 secondes par heure d'horloge ; l'utilisation d'instance qui dépasse 3600 secondes dans une heure d'horloge est facturée au tarif à la demande.

Par exemple, si vous achetez une Instance réservée `m4.xlarge` et que vous exécutez simultanément quatre instances `m4.xlarge` pendant une heure, une instance est facturée au tarif d'une heure d'utilisation d'Instance réservée et les trois autres instances sont facturées au tarif de trois heures d'utilisation à la demande.

Par contre, si vous achetez une Instance réservée `m4.xlarge` et que vous exécutez simultanément quatre instances `m4.xlarge` pendant 15 minutes (900 secondes), chacune au cours de la même heure, la durée d'exécution totale pour les instances est d'une heure, ce qui se traduit par une heure d'utilisation d'Instance réservée et 0 heure d'utilisation à la demande.

Si plusieurs instances éligibles s'exécutent simultanément, l'avantage de facturation d'Instance réservée est appliqué à toutes les instances en même temps pour un maximum de 3600 secondes dans une heure d'horloge ; ensuite ce sont les tarifs à la demande qui s'appliquent.

Cost Explorer dans la console Billing and Cost Management vous permet d'analyser les économies réalisées par rapport à l'exécution d'Instances à la demande. Le Forum Aux Questions (FAQ) Instances réservées inclut un exemple de calcul de valeur de liste.

Si vous fermez votre compte AWS, la facturation à la demande de vos ressources est interrompue. Toutefois, si vous avez des Instances réservées dans votre compte, vous continuez à recevoir une facture pour ces instances jusqu'à ce qu'elles expirent.

Affichage d'une facture

Vous pouvez consulter les frais et tarifs appliqués à votre compte sur la page de la console AWS Billing and Cost Management.

- Le Tableau de bord affiche un récapitulatif des dépenses de votre compte.
- Sur la page Factures, sous Détails, développez la section Elastic Compute Cloud et la région pour obtenir des informations de facturation sur vos Instances réservées.

Vous pouvez consulter les frais en ligne ou télécharger un fichier CSV.

Vous pouvez également assurer le suivi de l'utilisation de vos Instance réservées à l'aide du rapport d'utilisation et de coût AWS. Pour plus d'informations, consultez Instances réservées dans le rapport d'utilisation et de coût du AWS Billing and Cost Management Guide de l'utilisateur.

Instances réservées et la facturation consolidée

Les avantages de tarification des Instances réservées sont partagés lorsque le compte d'achat fait partie d'un ensemble de comptes facturés réunis sous un même compte payeur de facturation consolidée. L'utilisation d'instance pour tous les comptes membres est regroupé dans le compte souscripteur tous les mois. Cette fonctionnalité est généralement utile dans le cadre des sociétés disposant de plusieurs équipes ou groupes fonctionnels. Ensuite, la logique standard des Instance réservées est appliquée pour calculer le montant de la facture. Pour plus d'informations, consultez Facturation consolidée et AWS Organizations dans le Manuel de l'utilisateur AWS Organizations.

Si vous fermez le compte qui a acheté l'Instance réservée, le compte payeur continuera d'être débité pour cette Instance réservée jusqu'à ce que l'Instance réservée expire ou que le compte clôturé soit

définitivement supprimé. Le compte fermé est définitivement supprimé après 90 jours. Une fois supprimé, les comptes membres cessent de bénéficier de la réduction de facturation pour l'Instance réservée. Pour plus d'informations sur la fermeture d'un compte, consultez Clôture d'un compte AWS dans le Manuel de l'utilisateur AWS Organizations.

Niveaux de tarification avec remise d'Instance réservée

Si votre compte est éligible pour bénéficier d'un niveau de tarification avec remise, il bénéficie automatiquement des remises dès le départ et le tarif d'utilisation des instances pour tous les achats d'Instance réservée effectuées dans le cadre de ce niveau à partir de ce moment-là. Pour que votre compte soit éligible, la valeur de la liste répertoriant vos Instances réservées dans la région doit s'élever à 500 000 USD au minimum.

Les règles suivantes s'appliquent :

- Les niveaux de tarification et les remises associées s'appliquent uniquement aux achats d'Instances réservées standard Amazon EC2.
- Les niveaux de tarification ne s'appliquent pas aux Instances réservées pour Windows avec SQL Server Standard, SQL Server Web et SQL Server Enterprise.
- Les niveaux de tarification ne s'appliquent pas aux Instances réservées pour Linux avec SQL Server Standard, SQL Server Web et SQL Server Enterprise.
- Les remises en fonction des niveaux de tarification s'appliquent uniquement aux achats effectués auprès d'AWS. Elles ne s'appliquent pas aux achats d'Instances réservées tierces.
- Les achats d'Instance réservée convertible ne bénéficient pas actuellement de niveaux de tarification avec remise.

Rubriques

- Calcul des remises de tarification d'une Instance réservée (p. 342)
- Achat avec un niveau de remise (p. 343)
- Changement de niveau de tarification (p. 343)
- Facturation consolidée pour les niveaux de tarification (p. 344)

Calcul des remises de tarification d'une Instance réservée

Vous pouvez déterminer le niveau de tarification de votre compte en calculant la valeur de la liste répertoriant toutes vos Instances réservées dans une région. Multipliez le taux horaire récurrent de chaque réservation par le nombre total d'heures pour l'abonnement et ajoutez le tarif initial avant remise (également connu sous le nom de tarif fixe) au moment de l'achat. Dans la mesure où la valeur de la liste repose sur le tarif avant remise (public), elle ne change pas si vous êtes éligible pour une remise sur le volume ou si le tarif chute une fois que vous avez acheté vos Instances réservées.

```
List value = fixed price + (undiscounted recurring hourly price * hours in term)
```

Par exemple, pour une Instance réservée `t2.small` avec frais initiaux partiels d'une année, supposons que le prix initial est de 60,00 USD et que le tarif horaire est de 0,007 USD. Cela donne une valeur de liste de 121,32 USD.

```
121.32 = 60.00 + (0.007 * 8760)
```

Pour afficher les valeurs du tarif fixe des Instances réservées à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, choisissez Instances réservées.
3. Affichez la colonne Prix initial en choisissant Afficher / Masquer les colonnes (icône en forme d'engrenage) dans le coin supérieur droit.

Pour afficher les valeurs du tarif fixe des Instances réservées à l'aide de la ligne de commande

- describe-reserved-instances (AWS CLI)
- Get-EC2ReservedInstance (Outils AWS pour Windows PowerShell)
- DescribeReservedInstances (API Amazon EC2)

Achat avec un niveau de remise

Lorsque vous achetez des Instances réservées, Amazon EC2 applique automatiquement les remises à la partie de votre achat se trouvant au niveau de la tarification avec remise. Vous ne devez rien faire de particulier et vous pouvez acheter des Instances réservées à l'aide de n'importe quel outil Amazon EC2. Pour plus d'informations, consultez Achat d'Instances réservées (p. 344).

Une fois que la valeur de la liste répertoriant vos Instances réservées actives dans une région a atteint le niveau de tarification avec remise, tous les achats suivants d'Instances réservées dans cette région sont facturés au tarif réduit. Si un seul achat d'Instances réservées dans une région vous permet de dépasser le seuil d'un niveau de remise, la partie de l'achat qui dépasse ce seuil est facturée au tarif réduit. Pour plus d'informations sur les ID temporaires d'Instance réservée qui sont créés au cours du processus d'achat, consultez Changement de niveau de tarification (p. 343).

Si votre valeur de liste tombe en dessous du seuil minimum pour ce niveau de tarification avec remise (par exemple, lorsque certaines Instances réservées arrivent à expiration), les achats suivants d'Instances réservées dans la région ne sont pas facturés au tarif réduit. Toutefois, vous continuez à bénéficier de la remise appliquée aux Instances réservées initialement achetées dans le cadre du niveau de tarification avec remise.

Lorsque vous achetez des Instances réservées, quatre scénarios peuvent se produire :

- Aucune remise : votre achat dans une région se trouve toujours en dessous du seuil de remise.
- Remise partielle : votre achat dans une région dépasse le seuil du premier niveau de tarification avec remise. Aucune remise n'est appliquée à une ou plusieurs réservations et le taux avec remise est appliqué aux réservations restantes.
- Remise complète : tous vos achats au sein d'une région relèvent d'un niveau de tarification avec remise et sont en conséquence facturés au tarif réduit.
- Deux taux avec remise : votre achat dans une région vous permet de passer d'un niveau de tarification inférieur avec remise à un niveau de tarification supérieur avec remise. Deux taux différents sont facturés : une ou plusieurs réservations au taux avec remise inférieur et les réservations restantes au taux avec remise supérieur.

Changement de niveau de tarification

Si votre achat vous fait passer à un niveau de tarification avec remise, vous voyez plusieurs entrées pour cet achat : une première correspondant à la partie de l'achat facturée au prix standard et une deuxième correspondant à la partie de l'achat facturée au taux avec remise applicable.

Le service des Instance réservée génère plusieurs ID d'Instance réservée dans la mesure où votre achat vous permet de passer à un niveau avec remise ou d'un niveau avec remise inférieur à un niveau avec remise supérieur. Un ID est attribué à chaque ensemble de réservations d'un niveau. C'est pourquoi l'ID retourné par la commande CLI ou l'action d'API correspondant à votre achat est différent du véritable ID des nouvelles Instances réservées.

Facturation consolidée pour les niveaux de tarification

Un compte de facturation consolidée regroupe la valeur de liste des comptes membres au sein d'une région. Lorsque la valeur de la liste de toutes les Instances réservées actives du compte de facturation consolidée atteint un niveau de tarification avec remise, toute Instances réservées achetée après ce stade par un membre du compte de facturation consolidée est facturée au tarif avec remise (tant que la valeur de la liste associée à ce compte consolidé reste au-dessus du seuil du niveau de tarification avec remise). Pour plus d'informations, consultez Instances réservées et la facturation consolidée (p. 341).

Achat d'Instances réservées

Pour acheter une Instance réservée, recherchez des offres d'Instance réservée d'AWS et de vendeurs tiers, en modifiant les paramètres jusqu'à ce que vous trouviez la correspondance qui vous intéresse.

Lorsque vous recherchez des Instances réservées à acheter, vous recevez un devis avec le coût des offres renvoyées. Lorsque vous validez l'achat, AWS associe automatiquement un tarif limite au prix d'achat. Le coût total de vos Instances réservées ne dépasse pas le montant du devis.

Si le tarif augmente ou change pour quelque raison que ce soit, l'achat n'est pas validé. Au moment de l'achat, si des offres similaires à votre choix sont disponibles pour un prix inférieur, AWS vous vend les offres à ce prix inférieur.

Avant de valider votre achat, vérifiez les détails des Instance réservées que vous avez l'intention d'acheter et veillez à ce que tous les paramètres soient exacts. Après avoir acheté une Instance réservée (auprès d'un vendeur tiers sur la Place de marché des instances réservées ou sur AWS), vous ne pouvez plus annuler votre achat.

> **Note**
>
> Pour acheter et modifier les Instances réservées, vérifiez que votre compte d'utilisateur IAM dispose des autorisations appropriées, telles que la possibilité de décrire des Zones de disponibilité. Pour de plus amples informations, veuillez consulter Exemples de stratégies à utiliser avec l'interface de ligne de commande AWS ou un kit AWS SDK et Exemples de stratégies à utiliser dans la console Amazon EC2.

Tâches

- Sélection d'une plateforme (p. 344)
- Mise en file d'attente de votre achat (p. 345)
- Achat d'Instances réservées standard (p. 345)
- Achat d'Instances réservées convertibles (p. 347)
- Affichage de vos Instances réservées (p. 349)
- Annulation d'un achat mis en file d'attente (p. 349)
- Renouvellement d'une Instance réservée (p. 350)
- Utilisation de vos Instances réservées (p. 350)

Sélection d'une plateforme

Amazon EC2 prend en charge les plateformes Linux suivantes pour les Instances réservées :

- Linux/Unix
- Linux avec SQL Server Standard
- Linux avec SQL Server Web
- Linux avec SQL Server Enterprise

- SUSE Linux
- Utilisation de Red Hat Enterprise Linux

Lorsque vous achetez une Instance réservée, vous devez choisir une offre pour une plateforme qui correspond au système d'exploitation de votre instance.

- Pour les distributions SUSE Linux et RHEL, vous devez choisir les offres correspondant à ces plateformes spécifiques, c'est-à-dire pour les plateformes SUSE Linux ou Red Hat Enterprise Linux.
- Pour toutes les autres distributions Linux (y compris Ubuntu), choisissez une offre pour la plateforme Linux/UNIX.
- Si vous apportez votre abonnement RHEL existant, vous devez choisir une offre pour la plateforme Linux/UNIX et non une offre pour la plateforme Red Hat Enterprise Linux.

> **Important**
>
> Si vous prévoyez d'acheter une Instance réservée à appliquer à une instance à la demande qui a été lancée à partir d'un AWS Marketplace AMI, vérifiez d'abord le champ `PlatformDetails` de l'AMI. Le champ `PlatformDetails` indique quelle Instance réservée acheter. Les détails de la plate-forme de l'AMI doivent correspondre à la plate-forme de l'Instance réservée, sinon l'Instance réservée ne sera pas appliquée à l'instance à la demande. Pour de plus amples informations sur la façon d'afficher les détails de la plate-forme de l'AMI, reportez-vous à la section Obtenir des informations de facturation (p. 179).

Pour de plus amples informations sur les plateformes prises en charge pour Windows, veuillez consulter Sélection d'une plateforme dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Mise en file d'attente de votre achat

Par défaut, lorsque vous achetez une Instance réservée, l'achat s'exécute immédiatement. Vous pouvez également mettre vos achats en file d'attente pour une date et une heure futures. Par exemple, vous pouvez mettre un achat en file d'attente jusqu'à ce qu'une Instance réservée existante expire. Cela peut vous aider à garantir une couverture ininterrompue.

Vous pouvez mettre en file d'attente des achats pour une Instances réservées régionale, mais pas pour une Instances réservées de zone ou une Instances réservées d'autres vendeurs. Vous pouvez mettre un achat en file d'attente jusqu'à trois ans en avance. À l'heure et la date prévues, l'achat est exécuté à l'aide du mode de paiement par défaut. Une fois le paiement réussi, l'avantage de facturation est appliqué.

Vous pouvez afficher vos achats mis en file d'attente dans la console Amazon EC2. Le statut d'un achat mis en file d'attente est queued. Vous pouvez annuler un achat mis en file d'attente à tout moment avant son heure planifiée. Pour plus d'informations, consultez Annulation d'un achat mis en file d'attente (p. 349).

Achat d'Instances réservées standard

Vous pouvez acheter des Instances réservées standard dans une zone de disponibilité spécifique et obtenir une réservation de capacité. Vous avez également la possibilité de renoncer à la réservation de capacité et d'acheter une Instance réservée standard régionale.

Pour acheter des Instances réservées standard à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées, puis Acheter des Instances réservées.
3. Pour Classe d'offre, choisissez Standard pour afficher les Instances réservées standard.

4. Pour acheter une réservation de capacité, choisissez Ne montre que les offres réservant une capacité dans le coin supérieur droit de l'écran d'achat. Pour acheter une Instance réservée régionale, laissez la case à cocher désactivée.
5. Sélectionnez d'autres configurations en fonction de vos besoins, puis choisissez Recherche.

   Pour acheter une Instance réservée standard sur la Place de marché des instances réservées, recherchez Tiers dans la colonne Vendeur des résultats de recherche. La colonne Durée affiche des durées non standard.
6. Sélectionnez les Instances réservées à acheter, entrez la quantité, puis choisissez Ajouter au panier.
7. Pour afficher un récapitulatif des Instances réservées sélectionnées, choisissez Afficher le panier.
8. Si Order On (Commander le) est Maintenant, l'achat est terminé immédiatement. Pour mettre un achat en file d'attente, choisissez Maintenant et sélectionnez une date. Vous pouvez sélectionner une date différente pour chaque offre éligible dans le panier. L'achat est mis en file d'attente jusqu'à 00 h 00, dans le fuseau horaire de votre navigateur, à la date choisie.
9. Pour valider la commande, choisissez Order (Commander).

   Au moment de la commande, si des offres similaires à votre choix sont disponibles pour un prix inférieur, AWS vous vend les offres à ce prix inférieur.
10. L'état de votre commande figure dans la colonne État. Une fois votre commande terminée, la valeur État passe de `payment-pending` à `active`. Lorsque l'Instance réservée est `active`, elle est prête à être utilisée.

    Note

    Si l'état passe à `retired`, AWS n'a peut-être pas reçu votre paiement.

Pour acheter une Instance réservée standard à l'aide de l'AWS CLI

1. Pour rechercher les Instances réservées disponibles, utilisez la commande describe-reserved-instances-offerings. Spécifiez `standard` pour le paramètre `--offering-class` afin de renvoyer uniquement des Instances réservées standard. Vous pouvez appliquer des paramètres supplémentaires pour affiner vos résultats. Par exemple, si vous souhaitez acheter une Instance réservée `t2.large` régionale avec une location par défaut pour `Linux/UNIX` pour une durée d'un an seulement :

```
aws ec2 describe-reserved-instances-offerings \
    --instance-type t2.large \
    --offering-class standard \
    --product-description "Linux/UNIX" \
    --instance-tenancy default \
    --filters Name=duration,Values=31536000 Name=scope,Values=Region
```

   Pour rechercher des Instances réservées sur la Place de marché des instances réservées uniquement, utilisez le filtre `marketplace` et ne spécifiez pas de durée dans la demande, puisque la durée peut être inférieure à 1 ou 3 ans.

```
aws ec2 describe-reserved-instances-offerings \
    --instance-type t2.large \
    --offering-class standard \
    --product-description "Linux/UNIX" \
    --instance-tenancy default \
    --filters Name=marketplace,Values=true
```

   Lorsque vous trouvez une Instance réservée qui correspond à vos besoins, notez l'ID de l'offre. Exemples :

```
"ReservedInstancesOfferingId": "bec624df-a8cc-4aad-a72f-4f8abc34caf2"
```

2. Utilisez la commande purchase-reserved-instances-offering pour acheter une Instance réservée. Vous devez spécifier l'ID d'offre d'Instance réservée que vous avez obtenu à l'étape précédente et indiquer le nombre d'instances pour la réservation.

```
aws ec2 purchase-reserved-instances-offering \
    --reserved-instances-offering-id bec624df-a8cc-4aad-a72f-4f8abc34caf2 \
    --instance-count 1
```

   Par défaut, l'achat est terminé immédiatement. Pour mettre l'achat en file d'attente, vous pouvez également ajouter le paramètre suivant à l'appel précédent.

```
--purchase-time "2020-12-01T00:00:00Z"
```

3. Utilisez la commande describe-reserved-instances pour obtenir le statut de votre Instance réservée.

```
aws ec2 describe-reserved-instances
```

Vous pouvez également utiliser les commandes Outils AWS pour Windows PowerShell suivantes :

- Get-EC2ReservedInstancesOffering
- New-EC2ReservedInstance
- Get-EC2ReservedInstance

Une fois l'achat terminé, si vous avez déjà une instance en cours d'exécution qui correspond aux attributs de l'Instance réservée, l'avantage de facturation est immédiatement appliqué. Vous n'avez pas besoin de redémarrer vos instances. Si vous n'avez pas d'instance en cours d'exécution adéquate, lancez une instance et veillez à respecter les mêmes critères que ceux spécifiés pour l'Instance réservée. Pour plus d'informations, consultez Utilisation de vos Instances réservées (p. 350).

Pour des exemples de la façon dont les Instances réservées sont appliquées à vos instances en cours d'exécution, consultez Application des Instances réservées (p. 334).

Achat d'Instances réservées convertibles

Vous pouvez acheter des Instances réservées convertibles dans une zone de disponibilité spécifique et obtenir une réservation de capacité. Vous avez également la possibilité de renoncer à la réservation de capacité et d'acheter une Instance réservée convertible régionale.

Pour acheter des Instances réservées convertibles à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées, puis Acheter des Instances réservées.
3. Pour Classe d'offre, choisissez Convertible pour afficher des Instances réservées convertibles.
4. Pour acheter une réservation de capacité, choisissez Ne montre que les offres réservant une capacité dans le coin supérieur droit de l'écran d'achat. Pour acheter une Instance réservée régionale, laissez la case à cocher désactivée.
5. Sélectionnez d'autres configurations en fonction de vos besoins, puis choisissez Recherche.
6. Sélectionnez les Instances réservées convertibles à acheter, entrez la quantité, puis choisissez Ajouter au panier.
7. Pour afficher un résumé de votre sélection, choisissez Afficher le panier.

8. Si Order On (Commander le) est Maintenant, l'achat est terminé immédiatement. Pour mettre un achat en file d'attente, choisissez Maintenant et sélectionnez une date. Vous pouvez sélectionner une date différente pour chaque offre éligible dans le panier. L'achat est mis en file d'attente jusqu'à 00 h 00, dans le fuseau horaire de votre navigateur, à la date choisie.
9. Pour valider la commande, choisissez Order (Commander).

   Au moment de la commande, si des offres similaires à votre choix sont disponibles pour un prix inférieur, AWS vous vend les offres à ce prix inférieur.
10. L'état de votre commande figure dans la colonne État. Une fois votre commande terminée, la valeur État passe de `payment-pending` à `active`. Lorsque l'Instance réservée est `active`, elle est prête à être utilisée.

    Note

    Si l'état passe à `retired`, AWS n'a peut-être pas reçu votre paiement.

Pour acheter une Instance réservée convertible à l'aide de l'AWS CLI

1. Pour rechercher les Instances réservées disponibles, utilisez la commande describe-reserved-instances-offerings. Spécifiez `convertible` pour le paramètre `--offering-class` afin de renvoyer uniquement des Instances réservées convertibles. Vous pouvez appliquer des paramètres supplémentaires pour affiner vos résultats. Par exemple, si vous voulez acheter une Instance réservée `t2.large` régionale à location par défaut pour `Linux/UNIX` :

   ```
   aws ec2 describe-reserved-instances-offerings \
       --instance-type t2.large \
       --offering-class convertible \
       --product-description "Linux/UNIX" \
       --instance-tenancy default \
       --filters Name=scope,Values=Region
   ```

   Lorsque vous trouvez une Instance réservée qui correspond à vos besoins, notez l'ID de l'offre. Exemples :

   ```
   "ReservedInstancesOfferingId": "bec624df-a8cc-4aad-a72f-4f8abc34caf2"
   ```

2. Utilisez la commande purchase-reserved-instances-offering pour acheter une Instance réservée. Vous devez spécifier l'ID d'offre d'Instance réservée que vous avez obtenu à l'étape précédente et indiquer le nombre d'instances pour la réservation.

   ```
   aws ec2 purchase-reserved-instances-offering \
       --reserved-instances-offering-id bec624df-a8cc-4aad-a72f-4f8abc34caf2 \
       --instance-count 1
   ```

   Par défaut, l'achat est terminé immédiatement. Pour mettre l'achat en file d'attente, vous pouvez également ajouter le paramètre suivant à l'appel précédent.

   ```
   --purchase-time "2020-12-01T00:00:00Z"
   ```

3. Utilisez la commande describe-reserved-instances pour obtenir le statut de votre Instance réservée.

   ```
   aws ec2 describe-reserved-instances
   ```

Vous pouvez également utiliser les commandes Outils AWS pour Windows PowerShell suivantes :

- Get-EC2ReservedInstancesOffering
- New-EC2ReservedInstance
- Get-EC2ReservedInstance

Si vous avez déjà une instance en cours d'exécution qui correspond aux attributs de l'Instance réservée, l'avantage de facturation est immédiatement appliqué. Vous n'avez pas besoin de redémarrer vos instances. Si vous n'avez pas d'instance en cours d'exécution adéquate, lancez une instance et veillez à respecter les mêmes critères que ceux spécifiés pour l'Instance réservée. Pour plus d'informations, consultez Utilisation de vos Instances réservées (p. 350).

Pour des exemples de la façon dont les Instances réservées sont appliquées à vos instances en cours d'exécution, consultez Application des Instances réservées (p. 334).

Affichage de vos Instances réservées

Vous pouvez afficher les Instances réservées que vous avez achetées à l'aide de la console Amazon EC2 ou d'un outil de ligne de commande.

Pour afficher vos Instances réservées sur la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées.
3. Vos Instances réservées actives et mises hors service sont répertoriées. La colonne État indique l'état.
4. Si vous êtes vendeur sur la Place de marché des instances réservées, l'onglet Mes listes indique le statut d'une réservation répertoriée sur le Place de marché des instances réservées (p. 350). Pour plus d'informations, consultez États de la liste des éléments Instance réservée (p. 355).

Pour afficher vos Instances réservées à l'aide de la ligne de commande

- describe-reserved-instances (AWS CLI)
- Get-EC2ReservedInstance (Outils pour Windows PowerShell)

Annulation d'un achat mis en file d'attente

Vous pouvez mettre un achat en file d'attente jusqu'à trois ans en avance. Vous pouvez annuler un achat mis en file d'attente à tout moment avant son heure planifiée.

Pour annuler un achat mis en file d'attente

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées.
3. Sélectionnez une ou plusieurs Instances réservées.
4. Sélectionnez Actions, Delete Queued Reserved Instance (Supprimer les instances réservées mises en file d'attente).
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Yes, Delete.

Pour annuler un achat en file d'attente à l'aide de la ligne de commande

- delete-queued-reserved-instances (AWS CLI)
- Remove-EC2QueuedReservedInstance (Outils pour Windows PowerShell)

Renouvellement d'une Instance réservée

Vous pouvez renouveler une Instance réservée avant qu'elle n'entre en phase d'expiration. Le renouvellement d'une Instance réservée met en file d'attente l'achat d'une Instance réservée possédant la même configuration jusqu'à ce que l'Instance réservée actuelle expire.

Pour renouveler une Instance réservée à l'aide d'un achat en file d'attente

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées.
3. Sélectionnez une ou plusieurs Instances réservées.
4. Choisissez Actions, Renew Reserved Instances (Renouveler les instances réservées).
5. Pour valider la commande, choisissez Order (Commander).

Utilisation de vos Instances réservées

Les Instances réservées sont appliquées automatiquement aux Instances à la demande en cours d'exécution correspondant aux spécifications. Si vous n'avez pas d'Instances à la demande en cours d'exécution qui correspond aux spécifications de votre Instance réservée, l'Instance réservée est inutilisée jusqu'à ce que vous lanciez une instance avec les spécifications requises.

Si vous lancez une instance pour bénéficier de l'avantage de facturation d'une Instance réservée, veillez à spécifier les informations suivantes lors du lancement :

- Plateforme : vous devez choisir une Amazon Machine Image (AMI) qui correspond à la plateforme (description du produit) de votre Instance réservée. Par exemple, si vous avez spécifié `Linux/UNIX`, vous pouvez lancer une instance à partir d'une AMI Amazon Linux ou d'une AMI Ubuntu.
- Type d'instance : spécifiez le même type d'instance que pour votre Instance réservée ; par exemple, `t2.large`.
- Zone de disponibilité : si vous avez acheté une Instance réservée pour une zone de disponibilité spécifique, vous devez lancer l'instance dans la même zone de disponibilité. Si vous avez acheté une Instance réservée régionale, vous pouvez lancer votre instance dans n'importe quelle zone de disponibilité.
- Location : la location de votre instance doit correspondre à celle de l'Instance réservée ; par exemple, `dedicated` ou `shared`. Pour plus d'informations, consultez Instances dédiées (p. 498).

Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530). Pour des exemples de la façon dont les Instances réservées sont appliquées à vos instances en cours d'exécution, consultez Application des Instances réservées (p. 334).

Vous pouvez utiliser Amazon EC2 Auto Scaling ou d'autres services AWS pour lancer les Instances à la demande qui utilisent les avantages de vos Instance réservée. Pour plus d'informations, consultez le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Place de marché des instances réservées

La Place de marché des instances réservées est une plateforme qui prend en charge la vente d'Instances réservées standard inutilisées appartenant à des tiers et à des clients AWS. La durée et les options de tarification de ces instances sont donc susceptibles de varier. Par exemple, vous pouvez décider de vendre des Instances réservées après avoir transféré des instances vers une nouvelle région AWS, changé de type d'instance ou terminé un projet avant la fin de l'abonnement, lorsque vos besoins professionnels évoluent ou si vous avez une capacité superflue.

Si vous voulez vendre vos Instances réservées inutilisées sur la Place de marché des instances réservées, vous devez respecter certains critères d'éligibilité.

Table des matières

- Vente sur le Place de marché des instances réservées. (p. 351)
- Achat sur le Marketplace Instance réservée (p. 357)

Vente sur le Place de marché des instances réservées.

Dès que vous faites apparaître vos Instances réservées sur la Place de marché des instances réservées, elles deviennent disponibles et des acheteurs potentiels peuvent se les procurer. Toutes les Instances réservées sont regroupées selon la durée de réservation restante et le taux horaire.

Pour répondre à la demande d'un acheteur, AWS commence par vendre l'Instance réservée associée au tarif initial le plus faible dans le groupe spécifié, puis vend l'Instance réservée avec le tarif suivant le plus faible, et ainsi de suite jusqu'à ce que toute la commande de l'acheteur ait été exécutée. AWS traite ensuite les transactions et transfère la propriété des Instances réservées à l'acheteur.

Vous êtes le propriétaire de l'Instance réservée jusqu'à ce qu'elle soit vendue. Une fois la vente conclue, vous ne disposez plus de la réservation de capacité et vous n'êtes plus soumis aux frais récurrents avec remise. Si vous continuez à utiliser votre instance, AWS vous facturera le tarif à la demande à partir du moment où l'Instance réservée aura été vendue.

Sommaire

- Limites et restrictions (p. 351)
- Enregistrement en tant que vendeur (p. 352)
- Compte bancaire pour les remboursements (p. 352)
- Informations fiscales (p. 353)
- Évaluation de la tarification de vos Instances réservées (p. 353)
- Liste de vos Instances réservées (p. 354)
- États de la liste des éléments Instance réservée (p. 355)
- Cycle de vie d'une liste (p. 355)
- Après la vente de votre Instance réservée (p. 356)
- Obtention du paiement (p. 356)
- Communication des informations à l'acheteur (p. 357)

Limites et restrictions

Avant de pouvoir vendre vos réservations inutilisées, vous devez vous inscrire en tant que vendeur sur le Place de marché des instances réservées. Pour plus d'informations, consultez Enregistrement en tant que vendeur (p. 352).

Les restrictions et restrictions suivantes s'appliquent à la vente d'Instances réservées :

- Seules les Instances réservées Amazon EC2 standard peuvent être vendues sur le Place de marché des instances réservées. Les Instances réservées convertibles Amazon EC2 ne peuvent pas être vendues. Les Instances réservées pour les autres services AWS, tels que Amazon RDS et Amazon ElastiCache, ne peuvent pas être vendues.
- L'Instance réservée standard doit être valable pendant encore au moins un mois.
- Vous ne pouvez pas vendre une Instance réservée standard dans une région désactivée par défaut (p. 8).
- Le tarif minimum autorisé sur le Place de marché des instances réservées est de 0 USD.
- Vous pouvez vendre des Instances réservées avec aucun paiement initial, un paiement initial ou un paiement initial complet sur la Place de marché des instances réservées. S'il existe un paiement initial sur une Instance réservée, elle ne peut être vendue que quand AWS a reçu le paiement initial et que la réservation a été active (vous en avez été le propriétaire) pendant au moins 30 jours.

- Vous ne pouvez pas modifier directement votre liste sur le Place de marché des instances réservées. Toutefois, vous pouvez la changer en commençant par l'annuler, puis en créant une autre liste avec de nouveaux paramètres. Pour plus d'informations, consultez Évaluation de la tarification de vos Instances réservées (p. 353). Vous pouvez également modifier vos Instances réservées avant de les inclure dans votre liste. Pour plus d'informations, consultez Modification des Instances réservées (p. 357).
- AWS applique des frais de service s'élevant à 12 % du prix initial total de chaque Instance réservée standard vendue sur la Place de marché des instances réservées. Le prix initial correspond au prix demandé par le vendeur pour l'Instance réservée standard.
- Lorsque vous vous inscrivez en tant que vendeur, la banque que vous spécifiez doit avoir une adresse aux États-Unis. Pour de plus amples informations, veuillez consulter Exigences supplémentaires du vendeur pour les produits payés dans le Guide du vendeur AWS Marketplace.
- Les clients Amazon Internet Services Private Limited (AISPL) ne peuvent pas vendre Instances réservées sur le Place de marché des instances réservées, même s'ils ont un compte bancaire aux États-Unis. Pour de plus amples informations, veuillez consulter Quelles sont les différences entre les comptes AWS et AISPL ?

Enregistrement en tant que vendeur

> **Note**
>
> Seul l'utilisateur racine d'un compte AWS peut enregistrer un compte en tant que vendeur.

Pour vendre sur le Place de marché des instances réservées, vous devez d'abord vous enregistrer en tant que vendeur. Lors de l'enregistrement, vous devez fournir les informations suivantes lors de l'enregistrement :

- Informations bancaires : AWS doit disposer de vos informations bancaires afin de pouvoir rembourser les montants collectés lorsque vous vendez vos réservations. La banque que vous spécifiez doit avoir une adresse aux États-Unis. Pour plus d'informations, consultez Compte bancaire pour les remboursements (p. 352).
- Questionnaire fiscal : tous les vendeurs doivent répondre à un questionnaire fiscal afin de déterminer les obligations de déclaration fiscale éventuelles. Pour plus d'informations, consultez Informations fiscales (p. 353).

Une fois qu'AWS aura reçu votre enregistrement de vendeur, vous recevrez un e-mail confirmant votre enregistrement et vous informant que vous pouvez commencer à vendre sur la Place de marché des instances réservées.

Compte bancaire pour les remboursements

AWS doit disposer de vos informations bancaires afin de pouvoir rembourser les montants collectés lorsque vous vendez votre Instance réservée. La banque que vous spécifiez doit avoir une adresse aux États-Unis. Pour de plus amples informations, veuillez consulter Exigences supplémentaires du vendeur pour les produits payés dans le Guide du vendeur AWS Marketplace.

Pour enregistrer un compte par défaut destiné aux décaissements

1. Ouvrez la page Place de marché des instances réservées Seller Registration et connectez-vous à l'aide de vos informations d'identification AWS.
2. Sur la page Manage Bank Account (Gérer le compte bancaire), entrez les informations suivantes concernant la banque qui recevra vos paiements :
   - Nom du titulaire du compte bancaire
   - Code d'acheminement
   - Numéro de compte

- Type de compte bancaire

  Note

  Si vous utilisez le compte bancaire de votre société, vous êtes invité à envoyer les informations relatives au compte bancaire par télécopie au 1-206-765-3424.

Une fois l'enregistrement terminé, le compte bancaire spécifié est utilisé par défaut, dans l'attente d'une vérification auprès de la banque. Cette opération peut prendre jusqu'à deux semaines, une période au cours de laquelle vous ne pouvez pas recevoir de décaissements. Pour un compte établi, deux jours sont généralement nécessaires à l'exécution d'un décaissement.

Pour modifier le compte bancaire par défaut utilisé pour les décaissements

1. Sur la page Place de marché des instances réservées Seller Registration, connectez-vous avec le compte utilisé lors de votre inscription.
2. Sur la page Manage Bank Account (Gérer le compte bancaire), ajoutez un nouveau compte bancaire ou modifiez le compte défini par défaut.

Informations fiscales

Votre vente d'Instances réservées peut être soumise à une taxe appliquée aux transactions, telle qu'une taxe de vente ou une TVA. Vérifiez auprès du service fiscal, juridique, financier ou comptable de votre entreprise afin de déterminer si des taxes sont applicables aux transactions concernées. Il vous incombe de collecter et d'envoyer les taxes applicables aux transactions à l'administration fiscale appropriée.

Dans le cadre du processus d'enregistrement du vendeur, vous devez remplir un questionnaire d'ordre fiscal dans le Seller Registration Portal. Le questionnaire collecte vos informations fiscales et remplit un formulaire IRS W-9, W-8BEN ou W-8BEN-E, utilisé pour déterminer les éventuelles obligations de déclaration fiscale.

Les informations fiscales que vous renseignez dans le questionnaire peuvent différer selon que vous œuvrez comme personne morale ou physique, et que votre entreprise est une entité ou personne américaine ou non. En remplissant ce questionnaire, gardez les points suivants à l'esprit :

- Les informations fournies par AWS, notamment celles présentes dans cette rubrique, ne constituent pas des conseils d'ordre fiscal, juridique ou professionnel. Pour découvrir en quoi les obligations de déclaration imposées par l'IRS affectent votre entreprise, ou pour toute autre question, veuillez contacter votre conseiller fiscal, juridique ou autre.
- Pour vous conformer aux exigences de l'IRS en matière de déclarations aussi efficacement que possible, répondez à toutes les questions et entrez toutes les informations demandées au cours du questionnaire.
- Vérifiez vos réponses. Évitez les fautes de frappe ou la saisie de numéros d'identification fiscale inexacts. Ces erreurs risqueraient d'entraîner le refus de votre formulaire fiscal.

Selon vos réponses au questionnaire et les seuils de déclaration de l'IRS, Amazon peut soumettre le formulaire 1099-K. Vous en recevrez une copie par voie postale au plus tard le 31 janvier de l'année suivant celle où votre compte fiscal a atteint les niveaux de seuil. Par exemple, si votre compte fiscal atteint le seuil en 2018, vous recevrez le formulaire 1099-K le 31 janvier 2019 au plus tard.

Pour en savoir plus sur les exigences de l'IRS et sur le formulaire 1099-K, consultez le site web de l' IRS.

Évaluation de la tarification de vos Instances réservées

Les frais initiaux sont les seuls que vous puissiez spécifier pour l'Instance réservée que vous vendez. Les frais initiaux correspondent aux frais ponctuels que l'acheteur règle lorsqu'il achète une Instance réservée.

Vous devez garder à l'esprit les limites suivantes :

- Vous pouvez vendre jusqu'à 50 000 USD d'Instances réservées. Pour augmenter cette limite, remplissez le formulaire Ventes d'Instance réservée EC2.
- Vous pouvez vendre jusqu'à 5 000 Instances réservées. Pour augmenter cette limite, remplissez le formulaire Ventes d'Instance réservée EC2.
- Le tarif minimum est de 0 USD. Le tarif minimum autorisé sur le Place de marché des instances réservées est de 0 USD.

Vous ne pouvez pas modifier votre liste directement. Toutefois, vous pouvez la changer en commençant par l'annuler, puis en créant une autre liste avec de nouveaux paramètres.

Vous pouvez annuler votre liste à tout moment tant qu'elle est `active`. Vous ne pouvez pas annuler une liste si elle fait déjà l'objet d'une correspondance ou si sa vente est en cours de traitement. Si certaines instances de votre liste font l'objet d'une correspondance et que vous annulez la liste, seules les instances restantes qui ne font pas l'objet d'une correspondance sont supprimées de la liste.

Dans la mesure où la valeur des Instances réservées baisse régulièrement, AWS peut, par défaut, définir les prix de façon à ce qu'ils baissent d'un montant uniforme mois après mois. Toutefois, vous pouvez définir des tarifs initiaux différents en fonction du moment de vente de votre réservation.

Par exemple, si votre Instance réservée est encore valide pendant neuf mois, vous pouvez indiquer le montant que vous accepteriez si un client achetait cette Instance réservée au cours des neuf mois restants. Vous pouvez définir un autre prix avec cinq mois restants, et encore un autre avec un mois restant.

Liste de vos Instances réservées

En tant que vendeur enregistré, vous pouvez choisir de vendre une ou plusieurs de vos Instances réservées. Vous pouvez choisir de les vendre toutes sur une même liste ou par sections. En outre, vous pouvez ajouter à la liste les Instances réservées avec n'importe quelle configuration de type d'instance, plateforme et portée.

La console détermine une suggestion de prix. Elle vérifie les offres qui correspondent à votre Instance réservée et sélectionne celle dont le prix est le plus bas. Sinon, elle calcule un prix suggéré basé sur le coût de l’Instance réservée pour le temps restant. Si la valeur calculée est inférieure à 1,01 USD, le prix suggéré est de 1,01 USD.

Si vous annulez votre liste et qu'une partie de celle-ci a déjà été vendue, l'annulation ne s'applique pas à la partie déjà vendue. Seule la partie de la liste non encore vendue n'est plus disponible sur le Place de marché des instances réservées.

Pour ajouter une Instance réservée à la liste de la Place de marché des instances réservées à l'aide de AWS Management Console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées.
3. Sélectionnez les Instances réservées à répertorier, puis choisissez Actions, Vendre des Instances réservées.
4. Sur la page Configuration de votre liste d'Instance réservée définissez le nombre d'instances à vendre et le prix initial pour la durée restante dans les colonnes appropriées. Pour afficher l'évolution de la valeur de votre réservation au cours de la durée restante, sélectionnez la flèche en regard de la colonne Mois restant.
5. Si vous êtes un utilisateur avancé et que vous souhaitez personnaliser la tarification, vous pouvez entrer différentes valeurs pour les mois suivants. Pour revenir à la baisse de prix linéaire par défaut, choisissez Réinitialiser.

6. Choisissez Continuer une fois la configuration de la liste terminée.
7. Vérifiez les détails de votre liste sur la page Configuration de votre liste d'Instance réservée. Si vous n'avez rien à modifier, choisissez Répertorier l'instance réservée.

Pour afficher vos listes sur la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées.
3. Sélectionnez l'Instance réservée que vous avez répertoriée et choisissez l'onglet Mes listes en bas de la page.

Pour gérer Instances réservées dans Place de marché des instances réservées avec AWS CLI

1. Obtenez la liste de vos Instances réservées en utilisant la commande describe-reserved-instances.
2. Notez l'ID de l'Instance réservée que vous souhaitez répertorier et appelez create-reserved-instances-listing. Vous devez spécifier l'ID de l'Instance réservée, le nombre d'instances et le barème de tarification.
3. Pour afficher votre liste, utilisez la commande describe-reserved-instances-listings.
4. Pour annuler votre liste, utilisez la commande cancel-reserved-instances-listings.

États de la liste des éléments Instance réservée

État de la liste de l'onglet Mes listes de la page des Instances réservées affiche le statut actuel de vos listes :

Les informations figurant dans État de la liste concernent l'état de votre liste sur la Place de marché des instances réservées. Elles diffèrent des informations d'état affichées par la colonne État de la page Instances réservées. Ces informations d'État concernent votre réservation.

- active : la liste peut être achetée.
- canceled (annulée) : la liste a été annulée et ne peut plus être achetée sur la Place de marché des instances réservées.
- closed (fermée) : l'Instance réservée figure pas sur la liste. Une Instance réservée peut être `closed` parce que la vente de la liste est terminée.

Cycle de vie d'une liste

Lorsque toutes les instances d'une liste correspondent aux besoins d'un acheteur et sont vendues, l'onglet Mes listes indique que votre Total instance count (Nombre total d'instances) correspond au nombre indiqué sous Vendue. Il n'y a plus aucune instance avec le statut Disponible pour votre liste dont le Statut est désormais `closed`.

Lorsqu'une seule partie de votre liste est vendue, AWS supprime les Instances réservées concernées de la liste et crée le nombre d'Instances réservées égal aux Instances réservées restantes. Par conséquent, l'ID de liste et la liste qu'il représente, et qui a désormais moins de réservations en vente, restent actifs.

Toute vente ultérieure d'Instances réservées figurant sur la liste est traitée de cette façon. Une fois que toutes les Instances réservées de la liste ont été vendues, AWS indique que la liste est `closed`.

Par exemple, vous créez une liste ID de liste d'Instances réservées 5ec28771-05ff-4b9b-aa31-9e57dexample. Cette liste comporte 5 instances.

L'onglet Mes listes de la page de console Instance réservée affiche la liste de cette façon :

ID de liste d'Instance réservée 5ec28771-05ff-4b9b-aa31-9e57dexample

- Total reservation count = 5
- Sold = 0
- Available = 5
- Status = active

Un acheteur achète deux de ces réservations, ce qui laisse trois réservations encore disponibles à la vente. En raison de cette vente partielle, AWS crée une réservation avec trois instances correspondant à celles qui peuvent encore être achetées.

Voici comment votre liste apparaît sous l'onglet Mes listes :

ID de liste d'Instance réservée 5ec28771-05ff-4b9b-aa31-9e57dexample

- Total reservation count = 5
- Sold = 2
- Available = 3
- Status = active

Si vous annulez votre liste et qu'une partie de celle-ci a déjà été vendue, l'annulation ne s'applique pas à la partie déjà vendue. Seule la partie de la liste non encore vendue n'est plus disponible sur le Place de marché des instances réservées.

Après la vente de votre Instance réservée

Une fois que votre Instance réservée a été vendue, AWS vous envoie une notification par e-mail. Vous êtes averti par e-mail de toutes les activités quotidiennes vous concernant. Parmi les activités peuvent figurer le moment où vous créez ou vendez un liste, ou celui où AWS envoie des fonds à votre compte.

Pour suivre le statut d'une liste d'Instance réservée dans la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances réservées.
3. Cliquez sur l'onglet Mes listes.

   L'onglet Mes listes contient la valeur État de la liste. Il contient aussi des informations sur la durée, le tarif et le nombre d'instances disponibles, en attente, vendues ou annulées.

Vous pouvez également utiliser la commande describe-reserved-instances-listings avec le filtre approprié pour obtenir des informations sur vos listes.

Obtention du paiement

Dès qu'AWS reçoit le paiement de l'acheteur, un message est envoyé à l'e-mail de compte associé au compte enregistré en tant que propriétaire de l'Instance réservée vendue.

AWS envoie un virement ACH au compte bancaire spécifié. En règle générale, ce virement est effectué entre 1 à 3 jours après la vente de l'Instance réservée. Les décaissements se déroulent une fois par jour. Vous recevrez un e-mail avec un rapport de remboursement une fois que les fonds auront été débloqués. N'oubliez pas que vous pourrez recevoir les décaissements uniquement une fois qu'AWS aura reçu la validation de la part de votre banque. Cela peut prendre jusqu'à deux semaines.

L'Instance réservée que vous avez vendue continue à apparaître lorsque vous décrivez vos Instances réservées.

Vous recevez un décaissement pour vos Instances réservées via un virement qui arrive directement sur votre compte bancaire. AWS applique des frais de service s'élevant à 12 % du prix initial total de chaque Instance réservée vendue sur la Place de marché des instances réservées.

Communication des informations à l'acheteur

Lorsque vous vendez des instances sur la Place de marché des instances réservées, AWS communique le nom légal de votre société sur le relevé de l'acheteur conformément aux réglementations américaines. En outre, si l'acheteur appelle AWS Support parce qu'il a besoin de vous contacter au sujet d'une facture ou pour tout autre motif fiscal, AWS peut être amené à lui communiquer votre adresse e-mail afin qu'il puisse vous contacter directement.

De la même manière, le code postal et le pays de résidence de l'acheteur sont communiqués au vendeur dans le rapport de décaissement. En tant que vendeur, vous aurez parfois besoin de joindre ces informations aux taxes que vous remettez au gouvernement (par exemple, les taxes de vente ou la TVA) pour ces transactions.

AWS ne propose aucun conseil fiscal, mais si votre expert fiscal pense que vous avez besoin d'informations supplémentaires spécifiques, veuillez contacter AWS Support.

Achat sur le Marketplace Instance réservée

Vous pouvez acheter des Instances réservées auprès de vendeurs tiers qui possèdent des Instances réservées dont ils n'ont plus besoin sur la Place de marché des instances réservées. Vous pouvez effectuer cette opération à l'aide de la console Amazon EC2 ou d'un outil de ligne de commande. Le processus est semblable à l'achat d'Instances réservées auprès d'AWS. Pour plus d'informations, consultez Achat d'Instances réservées (p. 344).

Il existe peu de différences entre les Instances réservées achetées sur la Place de marché des instances réservées et les Instances réservées achetées directement auprès d'AWS :

- Durée : les Instances réservées que vous achetez auprès de tiers ont une durée inférieure à la durée standard complète. La durée standard complète proposée par AWS va de un à trois ans.
- Prix initial : les Instances réservées tierces peuvent être vendues à différents prix initiaux. Les frais d'utilisation ou récurrents restent les mêmes que ceux déterminés lorsque les Instances réservées ont été achetées initialement auprès d'AWS.
- Types d'Instances réservées : seules les Instances réservées standard Amazon EC2 peuvent être achetées sur la Place de marché des instances réservées. Les Instances réservées Instances réservées convertibles, Amazon RDS et Amazon ElastiCache ne sont pas disponibles à l'achat sur la Place de marché des instances réservées.

Les informations principales vous concernant sont communiquées au vendeur, par exemple votre code postal et votre pays de résidence.

Ces informations permettent au vendeur de calculer toutes les taxes destinées au gouvernement qui sont susceptibles d'être appliquées aux transactions (par exemple, les taxes de vente ou la TVA). Elles sont communiquées sous la forme d'un rapport de décaissement. Dans de rares cas de figure, il peut être demandé à AWS de fournir votre adresse de messagerie au vendeur afin que celui-ci puisse vous poser des questions sur la vente (par exemple, des questions relatives aux taxes).

Pour les mêmes raisons, AWS peut être amené à communiquer le nom de l'entité juridique du vendeur sur la facture d'achat de l'acheteur. Si vous avez besoin d'informations supplémentaires sur le vendeur pour des raisons fiscales ou autres, contactez AWS Support.

Modification des Instances réservées

Lorsque vos besoins évoluent, vous pouvez modifier vos Instances réservées convertibles et continuer à bénéficier de votre avantage de facturation. Vous pouvez modifier des attributs tels que la zone de

disponibilité , la taille d'instance (au sein de la même famille d'instances) et la portée de votre Instance réservée.

> **Note**
>
> Vous pouvez également échanger une Instance réservée convertible contre une autre Instance réservée convertible avec une configuration différente. Pour plus d'informations, consultez Échange d'Instances réservées convertibles (p. 365).

Vous pouvez modifier toutes vos Instances réservées ou un sous-ensemble. Vous pouvez séparer les Instances réservées initiales en deux nouvelles Instances réservées ou plus. Par exemple, si vous avez une réservation pour 10 instances dans `us-east-1a` et que vous décidez de déplacer 5 instances vers `us-east-1b`, la demande de modification entraîne la création de deux réservations : une pour 5 instances dans `us-east-1a` et l'autre pour 5 instances dans `us-east-1b`.

Vous pouvez aussi fusionner deux Instances réservées ou plus dans une Instance réservée unique. Par exemple, si vous avez quatre Instances réservées `t2.small` d'une instance chacune, vous pouvez les fusionner pour créer une Instance réservée `t2.large`. Pour plus d'informations, consultez Prise en charge de la modification de tailles d'instances (p. 360).

Après une modification, la tarification des Instances réservées est appliquée uniquement aux instances qui correspondent aux nouveaux paramètres. Par exemple, si vous modifiez la zone de disponibilité d'une réservation, les avantages de réservation de capacité et de tarification sont appliqués automatiquement à l'utilisation d'instance dans la nouvelle zone de disponibilité. Les instances qui ne correspondent plus aux nouveaux paramètres sont facturées au taux à la demande à moins que votre compte n'ait d'autres réservations applicables.

Si votre demande de modification a été appliquée :

- La réservation modifiée devient effective immédiatement et l'avantage de tarification est appliqué aux nouvelles instances à partir de l'heure de la demande de modification. Par exemple, si vous avez modifié vos réservations à 21 h 15, l'avantage de tarification est appliqué à votre nouvelle instance à partir de 21 h 00. Vous pouvez obtenir la date d'effet des Instances réservées modifiées en utilisant la commande describe-reserved-instances.
- La réservation initiale est mise hors service. Sa date de fin est la date de début de la nouvelle réservation et la date de fin de la nouvelle réservation est identique à la date de fin de l'Instance réservée initiale. Si vous modifiez une réservation d'une durée de trois ans avec 16 mois restants, la réservation modifiée a une durée de 16 mois, avec la même date de fin que la réservation initiale.
- La réservation modifiée indique un tarif fixe s'élevant à 0 USD et non le tarif fixe de la réservation initiale.
- Le tarif fixe de la réservation modifiée n'a aucune répercussion sur les calculs du niveau tarifaire avec remise appliqué à votre compte. Ces calculs reposent en effet sur le tarif fixe de la réservation initiale.

Si votre demande de modification échoue, vos Instances réservées conservent leur configuration d'origine et sont immédiatement disponibles pour une autre demande de modification.

Il n'y a aucun frais pour les modifications et vous ne recevez pas de nouvelles factures.

Vous pouvez modifier vos réservations aussi souvent que vous le souhaitez. Toutefois, vous ne pouvez pas modifier ou annuler une demande de modification en attente une fois que vous l'avez envoyée. Une fois la modification appliquée, vous pouvez envoyer une autre demande de modification afin d'annuler des modifications précédentes, si nécessaire.

Sommaire

- Conditions obligatoires et restrictions pour toute modification (p. 359)
- Prise en charge de la modification de tailles d'instances (p. 360)
- Envoi de demandes de modification (p. 363)
- Résolution des problèmes liés aux demandes de modification (p. 364)

Conditions obligatoires et restrictions pour toute modification

Vous pouvez modifier ces attributs comme suit.

Attribut modifiable	Plateformes prises en charge	Limites
Changer de zones de disponibilité au sein de la même région	Linux et Windows	-
Modifier la portée pour passer de Zone de disponibilité à Région et inversement	Linux et Windows	Si vous remplacez la portée Zone de disponibilité par Région, vous ne bénéficiez plus de l'avantage de réservation de capacité. Si vous remplacez la portée Région par Zone de disponibilité, vous perdez la flexibilité de zone de disponibilité et la flexibilité de taille d'instance (le cas échéant). Pour plus d'informations, consultez Application des Instances réservées (p. 334).
Modifiez la taille d'instance dans la même famille d'instances.	Linux/UNIX uniquement La flexibilité de taille d'instance n'est pas disponible pour les Instances réservées sur les autres plateformes, notamment Linux avec SQL Server Standard, Linux avec SQL Server Web, Linux avec SQL Server Enterprise, Red Hat Enterprise Linux, SUSE Linux, Windows, Windows avec SQL Standard, Windows avec SQL Server Enterprise et Windows avec SQL Server Web.	La réservation doit utiliser la location par défaut. Certaines familles d'instances ne sont pas prises en charge dans la mesure où aucune autre taille n'est disponible. Pour de plus amples informations, veuillez consulter Prise en charge de la modification de tailles d'instances (p. 360).
Changer de réseau pour passer de EC2-Classic à Amazon VPC et inversement	Linux et Windows	La plateforme réseau doit être disponible dans votre compte AWS. Si vous avez créé votre compte AWS après le 04/12/2013, il ne prend pas en charge EC2-Classic.

Prérequis

Amazon EC2 traite votre demande de modification si votre capacité est suffisante pour la configuration cible (le cas échéant) et si les conditions suivantes sont respectées :

- Vous ne pouvez pas modifier la Instance réservée avant ou au moment même de son achat.
- La Instance réservée doit être active.
- Il ne peut pas y avoir de demande de modification en attente
- La Instance réservée n'est pas répertoriée dans le Place de marché des instances réservées

- Il doit exister une correspondance entre la couverture de la taille d'instance associée à la réservation active et la configuration cible. Pour plus d'informations, consultez Prise en charge de la modification de tailles d'instances (p. 360).
- Les Instances réservées d'entrée sont toutes des Instances réservées standard ou des Instances réservées convertibles, non quelques-unes de chaque sorte
- Les Instances réservées d'entrée doivent expirer dans la même heure si ce sont des Instances réservées standard
- Le Instance réservée n'est pas une instance G4.

Prise en charge de la modification de tailles d'instances

Vous pouvez modifier la taille d'instance d'une Instance réservée si les conditions suivantes sont remplies.

Prérequis

- La plateforme est Linux/UNIX.
- Vous devez sélectionner une autre taille d'instance dans la même famille d'instances. Par exemple, vous ne pouvez pas modifier une Instance réservée `t2` en `t3`, que vous utilisiez la même taille ou une taille différente.

  Vous ne pouvez pas modifier la taille d'instance des Instances réservées pour les instances suivantes, car chacune de ces familles d'instances n'a qu'une taille :
  - `cc2.8xlarge`
  - `cr1.8xlarge`
  - `hs1.8xlarge`
  - `t1.micro`
- La Instance réservée originale et modifiée doit avoir la même couverture de taille d'instance.

Sommaire

- Couverture de taille d'instance (p. 360)
- Facteur de normalisation pour les instances à matériel nu (p. 362)

Couverture de taille d'instance

Chaque Instance réservée a une couverture de taille d'instance qui est déterminée par le facteur de normalisation de taille d'instance et par le nombre d'instances dans la réservation. Lorsque vous modifiez les tailles des instances dans une Instance réservée, la couverture de la configuration cible doit correspondre à celle de la configuration d'origine, sinon la demande de modification n'est pas traitée.

Pour calculer la couverture de la taille d'une Instance réservée, multipliez le nombre d'instances par le facteur de normalisation. Dans la console Amazon EC2, le facteur de normalisation est mesuré en unités. Le tableau suivant décrit le facteur de normalisation pour les tailles d'instance dans une famille d'instances. Par exemple, une instance `t2.medium` dispose d'un facteur de normalisation de 2, ce qui implique qu'une réservation de 4 instances `t2.medium` dispose d'une couverture de 8 unités.

Taille d'instance	Facteur de normalisation
nano	0.25
micro	0.5
small	1

Taille d'instance	Facteur de normalisation
medium	2
large	4
xlarge	8
2xlarge	16
4xlarge	32
8xlarge	64
9xlarge	72
10xlarge	80
12xlarge	96
16xlarge	128
18xlarge	144
24xlarge	192
32xlarge	256

Vous pouvez allouer vos réservations en utilisant différentes tailles d'instance sur la même famille d'instance tant que la couverture de taille d'instance de votre réservation reste la même. Par exemple, vous pouvez diviser une réservation pour une instance `t2.large` (1 @ 4 unités) en quatre instances `t2.small` (4 @ 1 unité). De même, vous pouvez combiner une réservation pour quatre instances `t2.small` en une seule instance `t2.large`. Toutefois, vous ne pouvez pas remplacer votre réservation de deux instances `t2.small` par une seule instance `t2.large`, car la couverture de la réservation modifiée (4 unités) est plus grande que celle de la réservation existante (2 unités).

Dans l'exemple suivant, vous avez une réservation avec deux instances `t2.micro` (1 unité) et une réservation avec une instance `t2.small` (1 unité). Si vous fusionnez ces deux réservations en une seule avec une instance `t2.medium` (2 unités), la couverture de la réservation modifiée est égale à la couverture des réservations combinées.

Vous pouvez aussi modifier une réservation pour la diviser en deux réservations ou plus. Dans l'exemple suivant, vous disposez d'une réservation avec une instance `t2.medium` (2 unités). Vous pouvez diviser la réservation en deux, l'une avec deux instances `t2.nano` (0,5 unités) et l'autre avec trois instances `t2.micro` (1,5 unité).

Facteur de normalisation pour les instances à matériel nu

Vous pouvez modifier une réservation avec des instances `metal` en utilisant d'autres tailles au sein de la même famille d'instances. De même, vous pouvez modifier une réservation avec des instances autres que des instances à matériel nu en utilisant la taille `metal` de la même famille d'instances. Généralement, une instance à matériel nu a la même taille que la plus grande taille d'instance disponible au sein de la même famille d'instances. Par exemple, une instance `i3.metal` a la même taille qu'une instance `i3.16xlarge`, de sorte qu'elles ont le même facteur de normalisation.

Le tableau suivant décrit le facteur de normalisation pour les tailles d'instance à matériel nu dans les familles d'instances qui ont des instances à matériel nu. Le facteur de normalisation des instances `metal` dépend de la famille d'instances, contrairement aux autres tailles d'instance.

Taille d'instance à matériel nu	Facteur de normalisation
`c5.metal`	192
`i3.metal`	128
`r5.metal`	192
`r5d.metal`	192
`z1d.metal`	96
`m5.metal`	192
`m5d.metal`	192

Par exemple, une instance `i3.metal` dispose d'un facteur de normalisation de 128. Si vous achetez une Instance réservée Amazon Linux/Unix à location par défaut `i3.metal`, vous pouvez diviser la réservation comme suit :

- Une `i3.16xlarge` fait toujours la même taille qu'une instance `i3.metal`. Il dispose donc d'un facteur de normalisation de 128 (128/1). La réservation pour une instance `i3.metal` peut être modifiée en une instance `i3.16xlarge`.
- Une `i3.8xlarge` fait toujours la moitié de la taille d'une instance `i3.metal`. Il dispose donc d'un facteur de normalisation de 64 (128/2). La réservation pour une instance `i3.metal` peut être divisée en deux instances `i3.8xlarge`.
- Une `i3.4xlarge` fait toujours le quart de la taille d'une instance `i3.metal`. Il dispose donc d'un facteur de normalisation de 32 (128/4). La réservation pour une instance `i3.metal` peut être divisée en quatre instances `i3.4xlarge`.

Envoi de demandes de modification

Avant de modifier vos Instances réservées, veillez à lire les restrictions (p. 359) applicables. Avant de modifier la taille d'instance, calculez la couverture de la taille d'instance (p. 360) totale des réservations que vous voulez modifier et vérifiez qu'elle correspond à la couverture de taille d'instance totale de vos configurations cibles.

Pour modifier vos Instances réservées à l'aide d'AWS Management Console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sur la page Instances réservées, sélectionnez une ou plusieurs Instances réservées à modifier, puis choisissez Actions, Modifier des instances réservées.

   Note

   Si vos Instances réservées ne sont pas actives ou si elles ne peuvent pas être modifiées, Modifier des Instances réservées est désactivé.
3. La première entrée du tableau de modification indique les attributs des Instances réservées sélectionnées, et au moins une configuration cible en dessous. La colonne Unités indique la couverture de taille d'instance totale. Choisissez Ajouter pour chaque nouvelle configuration à ajouter. Modifiez les attributs de chaque configuration selon vos besoins, puis choisissez Continuer :

- Portée : indiquez si la configuration s'applique à une zone de disponibilité ou à l'ensemble de la région.
- Zone de disponibilité : choisissez la zone de disponibilité requise. Ne s'applique pas aux Instances réservées régionales.
- Type d'instance : sélectionnez le type d'instance requis. Les configurations combinées doivent être égales à la couverture de taille d'instance de vos configurations d'origine.
- Nombre : spécifiez le nombre d'instances. Pour fractionner les Instances réservées en plusieurs configurations, réduisez leur nombre, choisissez Ajouter et spécifiez un nombre pour la configuration supplémentaire. Par exemple, si vous disposez d'une configuration unique comportant 10 instances réservées, vous pouvez redéfinir ce nombre sur 6 et ajouter une configuration avec un nombre de 4. Ce processus supprime l'Instance réservée d'origine une fois les nouvelles Instances réservées activées.

4. Pour valider vos choix de modification une fois que vous avez terminé la définition des configurations cibles, choisissez Soumettre des modifications.
5. Vous pouvez consulter l'état de votre demande de modification en observant la colonne État de l'écran des Instances réservées. Les états possibles sont les suivants :
   - active (en attente de modification) : État de transition pour les Instances réservées initiales
   - hors service (en attente de modification) : État de transition pour les Instances réservées initiales pendant que les nouvelles Instances réservées sont créées
   - hors service : Instances réservées modifiées et remplacées avec succès.
   - active : L'un des statuts suivants :
     - Nouvelles Instances réservées créées à la suite d'une demande de modification
     - Instances réservées initiales après l'échec d'une demande de modification

Pour modifier vos Instances réservées à l'aide de la ligne de commande

1. Pour modifier vos Instances réservées, vous pouvez utiliser l'une des commandes suivantes :
   - modify-reserved-instances (AWS CLI)
   - Edit-EC2ReservedInstance (Outils AWS pour Windows PowerShell)
2. Pour obtenir le statut de votre demande modification (`processing`, `fulfilled` ou `failed`), utilisez une des commandes suivantes :
   - describe-reserved-instances-modifications (AWS CLI)
   - Get-EC2ReservedInstancesModification (Outils AWS pour Windows PowerShell)

Résolution des problèmes liés aux demandes de modification

Si les paramètres que vous avez demandés pour la configuration cible sont uniques, vous recevez un message indiquant que votre demande est en cours de traitement. A ce stade, Amazon EC2 a uniquement déterminé que les paramètres de votre demande de modification étaient valides. Votre demande de modification peut encore échouer au cours du traitement si la capacité nécessaire n'est pas disponible.

Dans certains cas, vous ne recevrez pas de confirmation, mais un message indiquant que la demande de modification a échoué ou est incomplète. Utilisez les informations de ces messages comme point de départ pour soumettre une nouvelle demande de modification. Veillez à lire les restrictions (p. 359) applicables avant d'envoyer la demande.

Certaines Instances réservées sélectionnées ne peuvent pas faire l'objet d'une modification

Amazon EC2 identifie et répertorie les Instances réservées qui ne peuvent pas être modifiées. Si vous recevez un message de ce type, accédez à la page Instances réservées de la console Amazon EC2 et consultez les informations pour les Instances réservées.

Erreur lors du traitement de votre demande de modification

Vous avez demandé la modification d'une ou de plusieurs Instances réservées, mais aucune de ces demandes ne peut être traitée. Selon le nombre de réservations que vous modifiez, vous pouvez obtenir différentes versions de ce message.

Amazon EC2 affiche les raisons pour lesquelles votre demande ne peut pas être traitée. Par exemple, vous pouvez avoir spécifié la même configuration cible (une combinaison de zone de disponibilité et de plateforme) pour une ou plusieurs parties des Instances réservées que vous modifiez. Essayez de soumettre à nouveau les demandes de modification, mais veillez à ce que les détails d'instance des réservations soient corrects et à ce que les configurations cibles pour toutes les parties modifiées soient uniques.

Échange d'Instances réservées convertibles

Vous pouvez échanger une ou plusieurs Instances réservées convertibles contre une autre Instance réservée convertible avec une configuration différente, y compris la famille d'instance, le système d'exploitation et la location. Il n'y a pas de limite au nombre d'échanges que vous pouvez effectuer, tant que l'Instance réservée convertible cible est de valeur égale ou plus élevée que les Instances réservées convertibles que vous échangez.

Lorsque vous échangez votre Instance réservée convertible, le nombre d'instances de votre réservation actuelle est remplacé par un nombre d'instances qui couvre une valeur égale ou supérieure à celle de la configuration de l'Instance réservée convertible cible. Amazon EC2 calcule le nombre d'Instances réservées que vous pouvez recevoir comme résultat de l'échange.

Sommaire

- Exigences pour l'échange d'Instances réservées convertibles (p. 365)
- Calcul des échanges de Instances réservées convertibles (p. 366)
- Fusion de Instances réservées convertibles (p. 367)
- Échanger une partie d'une Instance réservée convertible (p. 367)
- Envoi de demandes d'échange (p. 368)

Exigences pour l'échange d'Instances réservées convertibles

Si les conditions suivantes sont remplies, Amazon EC2 traite votre demande d'échange. Votre Instance réservée convertible doit être :

- Actif
- Libre de toute demande d'échange précédente

Les règles suivantes s'appliquent :

- Les Instances réservées convertibles ne peuvent être échangées que contre d'autres Instances réservées convertibles actuellement proposées par AWS.
- Les Instances réservées convertibles sont associées à une région spécifique, qui reste la même pendant la durée de la période de réservation. Vous ne pouvez pas échanger une Instance réservée convertible par une Instance réservée convertible d'une autre région.
- Vous pouvez échanger une ou plusieurs Instances réservées convertibles à la fois contre une seule Instance réservée convertible.
- Pour échanger une partie d'une Instance réservée convertible, vous pouvez la modifier en deux réservations ou plus, avant d'en échanger une ou plusieurs contre une nouvelle Instance réservée convertible. Pour plus d'informations, consultez Échanger une partie d'une Instance réservée convertible (p. 367). Pour plus d'informations sur la modification de vos Instances réservées, consultez Modification des Instances réservées (p. 357).

- Les Instances réservées convertibles avec tous les frais initiaux peuvent être échangées contre des Instances réservées convertibles avec frais initiaux partiels, et inversement.

  Note

  Si le paiement total des frais initiaux requis pour l'échange (coût de régularisation) est inférieur à $ 0,00, AWS vous donne automatiquement une quantité d'instances dans la Instance réservée convertible qui garantit que le coût d'alignement est de $ 0,00 ou plus.

  Note

  Si la valeur totale (prix initial + prix horaire * nombre d'heures restantes) de la nouvelle Instance réservée convertible est inférieure à la valeur totale de la Instance réservée convertible échangée, AWS vous attribue automatiquement une quantité d'instances dans la Instance réservée convertible qui garantit que la valeur totale est égale ou supérieure à celle de la Instance réservée convertible échangée.
- Pour bénéficier d'un meilleur tarif, vous pouvez échanger une Instance réservée convertible sans paiement initial pour une Instance réservée convertible avec tous les frais totaux ou avec frais initiaux partiels.
- Vous ne pouvez pas échanger de Instances réservées convertibles avec tous les frais initiaux ou avec frais initiaux partiels contre des Instances réservées convertibles. sans frais initiaux.
- Vous pouvez échanger une Instance réservée convertible sans frais initiaux pour une autre Instance réservée convertible sans frais initiaux uniquement si le tarif horaire de la nouvelle Instance réservée convertible est égal ou supérieur au prix horaire de la Instance réservée convertible échangée.

  Note

  Si la valeur totale (prix horaire * nombre d'heures restantes) de la nouvelle Instance réservée convertible est inférieure à la valeur totale de la Instance réservée convertible échangée, AWS vous attribue automatiquement une quantité d'instances dans la Instance réservée convertible qui garantit que la valeur totale est égale ou supérieure à celle de la Instance réservée convertible échangée.
- Si vous échangez plusieurs Instances réservées convertibles avec différentes dates d'expiration, la date d'expiration de la nouvelle Instance réservée convertible est la plus lointaine dans le futur.
- Si vous échangez une Instance réservée convertible unique, elle doit avoir la même durée que la nouvelle Instance réservée convertible (1 an ou 3 ans). Si vous fusionnez plusieurs Instances réservées convertibles avec différentes durées, la nouvelle Instance réservée convertible aura une durée de 3 ans. Pour plus d'informations, consultez Fusion de Instances réservées convertibles (p. 367).
- Une fois que vous avez échangé un Instance réservée convertible, la réservation originale est retirée. Sa date de fin est la date de début de la nouvelle réservation et la date de fin de la nouvelle réservation est identique à la date de fin de la Instance réservée convertible initiale. Par exemple, si vous modifiez une réservation d'une durée de trois ans avec 16 mois restants, la réservation modifiée a une durée de 16 mois, avec la même date de fin que la réservation initiale.

Calcul des échanges de Instances réservées convertibles

L'échange de Instances réservées convertibles est gratuit. Toutefois, vous pouvez être tenu de payer des frais de régularisation calculés au prorata du paiement comptant de la différence entre les Instances réservées convertibles que vous aviez et les Instances réservées convertibles que vous recevez de l'échange.

Chaque Instance réservée convertible dispose d'une liste de valeurs. Cette valeur de liste est comparée à la valeur de liste des Instances réservées convertibles que vous voulez pour déterminer combien de réservations d'instances vous pouvez recevoir de l'échange.

Par exemple : vous avez une Instance réservée convertible avec une valeur de liste de 35 $ que vous voulez échanger contre un nouveau type d'instance avec une valeur de liste de 10 $.

```
$35/$10 = 3.5
```

Vous pouvez échanger votre Instance réservée convertible contre trois Instances réservées convertibles de 10 USD. Étant donné qu'il n'est pas possible d'acheter des moitiés de réservation, vous devez acheter une Instance réservée convertible supplémentaire pour couvrir le reste :

```
3.5 = 3 whole Instances réservées convertibles + 1 additional Instance réservée
 convertible.
```

La quatrième Instance réservée convertible a la même date de fin que les trois autres. Vous payez la régularisation correspondant à la quatrième réservation si vous échangez des Instances réservées convertibles à paiement initial partiel ou comptant. Si le reste du paiement en amont de vos Instances réservées convertibles est de 500 USD et que la réservation cible coûterait normalement 600 USD au prorata, vous êtes facturé 100 USD.

```
$600 prorated upfront cost of new reservations - $500 remaining upfront cost of original
 reservations = $100 difference.
```

Fusion de Instances réservées convertibles

Si vous fusionnez deux Instances réservées convertibles ou plus, le terme de la Instance réservée convertible obtenue doit être le même que celui des Instances réservées convertibles originales ou celui de la plus grande des Instances réservées convertibles originales. La date d'expiration de la nouvelle Instance réservée convertible est la plus lointaine dans le futur.

Par exemple, si vous possédez les Instances réservées convertibles suivantes sur votre compte :

ID Instance réservée	Durée	Date d'expiration
aaaa1111	1 an	31-12-2018
bbbb2222	1 an	31-07-2018
cccc3333	3 ans	30-06-2018
dddd4444	3 ans	31-12-2019

- Vous pouvez fusionner `aaaa1111` et `bbbb2222` et les échanger contre une Instance réservée convertible valable 1 an. Vous ne pouvez pas les échanger contre une Instance réservée convertible valable trois ans. La date d'expiration de la nouvelle Instance réservée convertible est 2018-12-31.
- Vous pouvez fusionner `bbbb2222` et `cccc3333` et les échanger contre une Instance réservée convertible valable 3 ans. Vous ne pouvez pas les échanger contre une Instance réservée convertible valable un an. La date d'expiration de la nouvelle Instance réservée convertible est 2018-07-31.
- Vous pouvez fusionner `cccc3333` et `dddd4444` et les échanger contre une Instance réservée convertible valable 3 ans. Vous ne pouvez pas les échanger contre une Instance réservée convertible valable un an. La date d'expiration de la nouvelle Instance réservée convertible est 2019-12-31.

Échanger une partie d'une Instance réservée convertible

Vous pouvez utiliser le processus de modification pour diviser votre Instance réservée convertible en plus petites réservations, avant d'en échanger une ou plusieurs contre une nouvelle Instance réservée convertible. Les exemples suivant montrent comment procéder.

Example Exemple : Instance réservée convertible avec plusieurs instances

Dans cet exemple, vous disposez d'une Instance réservée convertible `t2.micro` avec quatre instances dans la réservation. Pour échanger deux instances `t2.micro` contre une instance `m4.xlarge` :

1. Modifiez la Instance réservée convertible `t2.micro` en la divisant en deux Instances réservées convertibles `t2.micro` avec deux instances chacune.
2. Échangez l'une des nouvelles Instances réservées convertibles `t2.micro` obtenues contre une Instance réservée convertible `m4.xlarge`.

Example Exemple : Instance réservée convertible avec une seule instance

Dans cet exemple, vous disposez d'une `t2.large` Instance réservée convertible. Pour la changer en une instance `t2.medium` plus petite et une instance `m3.medium` :

1. Modifiez l'Instance réservée convertible `t2.large` en la divisant en deux Instances réservées convertibles `t2.medium`. Une seule instance `t2.large` a la même couverture de taille d'instance que les deux instances `t2.medium`.
2. Échangez l'une des nouvelles Instances réservées convertibles `t2.medium` obtenues contre une Instance réservée convertible `m3.medium`.

Pour plus d'informations, consultez Prise en charge de la modification de tailles d'instances (p. 360) et Envoi de demandes d'échange (p. 368).

Envoi de demandes d'échange

Vous pouvez échanger vos Instances réservées convertibles à l'aide de la console Amazon EC2 ou d'un outil de ligne de commande.

Recherche d'une Instance réservée convertible à l'aide de la console

Vous pouvez rechercher des offres de Instances réservées convertibles et sélectionner votre nouvelle configuration parmi les choix fournis.

Pour échanger des Instances réservées convertibles à l'aide de la console Amazon EC2 :

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instances réservées, sélectionnez les Instances réservées convertibles à échanger, puis choisissez Actions, Échange de l'Instance réservée.
3. Sélectionnez les attributs de la configuration souhaitée à l'aide des menus déroulants et choisissez Trouver une offre.
4. Sélectionnez une nouvelle Instance réservée convertible. La colonne Nombre d'instances indique le nombre d'Instances réservées que vous recevez pour l'échange. Lorsque vous avez sélectionné une Instance réservée convertible qui répond à vos besoins, choisissez Échange.

Les Instances réservées qui ont été échangées sont mises hors service et les nouvelles Instances réservées s'affichent dans la console Amazon EC2. Ce processus peut prendre quelques minutes pour se propager.

Échange d'uneInstance réservée convertible à l'aide de l'interface de ligne de commande

Pour échanger une Instance réservée convertible, commencez par rechercher une Instance réservée convertible cible qui répond à vos besoins :

- describe-reserved-instances-offerings (AWS CLI)
- Get-EC2ReservedInstancesOffering (Outils pour Windows PowerShell)

Obtenez un devis pour l'échange, qui inclut le nombre de Instances réservées que vous obtenez lors de l'échange et les frais de régularisation pour l'échange :

- get-reserved-instances-exchange-quote (AWS CLI)
- GetEC2-ReservedInstancesExchangeQuote (Outils pour Windows PowerShell)

Enfin, effectuez l'échange :

- accept-reserved-instances-exchange-quote (AWS CLI)
- Confirm-EC2ReservedInstancesExchangeQuote (Outils pour Windows PowerShell)

Instances réservées planifiées

Les Instances réservées planifiées (instances planifiées) vous permettent d'acheter des réservations de capacités récurrentes sur une base quotidienne, hebdomadaire ou mensuelle, avec une date de début et une durée pour une période d'un an. Comme vous réservez la capacité à l'avance, vous savez qu'elle est disponible quand vous en avez besoin. Vous payez pour la durée pour laquelle les instances sont planifiées, même si vous ne les utilisez pas.

Les instances planifiées conviennent plus particulièrement aux charges de travail que vous n'exécutez pas en continu, mais selon une planification régulière. Par exemple, vous pouvez utiliser des instances planifiées pour une application qui s'exécute pendant des heures d'ouverture ou pour un traitement par lots qui s'exécute en fin de semaine.

Si vous avez besoin d'une réservation de capacité de manière récurrente, les Instances réservées peuvent répondre à vos besoins et réduire les coûts. Pour plus d'informations, consultez Instances

réservées (p. 329). Si vous êtes flexible quant au moment où vos instances s'exécutent, les Instances Spot peuvent répondre à vos besoins et réduire les coûts. Pour plus d'informations, consultez Instances Spot (p. 373).

Sommaire

- Fonctionnement des instances planifiées (p. 370)
- Rôles liés à un service pour les Instances planifiées (p. 371)
- Achat d'une instance planifiée (p. 371)
- Lancement d'une instance planifiée (p. 372)
- Limites des instances planifiées (p. 373)

Fonctionnement des instances planifiées

Amazon EC2 met de côté des pools d'instances EC2 dans chaque zone de disponibilité pour une utilisation en tant qu'instances planifiées. Chaque groupe prend en charge une combinaison spécifique de type d'instance, de système d'exploitation et de réseau.

Pour commencer, vous devez rechercher un planning disponible. Vous pouvez effectuer une recherche dans un ou plusieurs pools. Une fois que vous avez trouvé un planning qui vous convient, achetez-le.

Vous devez lancer vos instances planifiées pendant les périodes prévues, à l'aide d'une configuration de lancement correspondant aux attributs suivants du programme que vous avez acheté : type d'instance, zone de disponibilité, réseau et plateforme. Quand vous procédez ainsi, Amazon EC2 lance des instances EC2 en votre nom, en fonction de la spécification de lancement indiquée. Amazon EC2 doit s'assurer que les instances EC2 ont pris fin d'ici la fin de la période planifiée en cours de sorte que la capacité soit disponible pour les autres instances planifiées pour lesquelles elle est réservée. Par conséquent, Amazon EC2 met fin aux instances EC2 trois minutes avant la fin de la période planifiée en cours.

Vous ne pouvez pas arrêter ou redémarrer des instances planifiées, mais vous pouvez y mettre fin manuellement si nécessaire. Si vous mettez fin à une instance planifiée avant la fin de sa période planifiée en cours, vous pouvez la relancer après quelques minutes. Sinon, vous devez attendre jusqu'à la période planifiée suivante.

Le graphique suivant illustre le cycle de vie d'une instance planifiée.

Rôles liés à un service pour les Instances planifiées

Amazon EC2 crée un rôle lié à un service lorsque vous achetez une Instance planifiée. Un rôle lié à un service comprend toutes les autorisations requises par Amazon EC2 pour appeler d'autres services AWS en votre nom. Pour plus d'informations, consultez Utilisation des rôles liés à un service dans le IAM Guide de l'utilisateur.

Amazon EC2 se sert du rôle lié à un service nommé AWSServiceRoleForEC2ScheduledInstances pour accomplir les actions suivantes :

- `ec2:TerminateInstances` : Résilier des Instances planifiées après leur planification
- `ec2:CreateTags` : Ajouter des balises système aux Instances planifiées

Si vous avez acheté des Instances planifiées avant octobre 2017 au moment où Amazon EC2 a commencé à prendre en charge ce rôle lié aux services, Amazon EC2 a créé le rôle AWSServiceRoleForEC2ScheduledInstances dans votre compte AWS. Pour plus d'informations, consultez Un nouveau rôle est apparu dans mon compte dans le IAM Guide de l'utilisateur.

Si vous n'avez plus besoin d'utiliser d'instances planifiées, nous vous recommandons de supprimer le rôle AWSServiceRoleForEC2ScheduledInstances. Après sa suppression de votre compte, Amazon EC2 créera de nouveau le rôle si vous achetez des Instances planifiées.

Achat d'une instance planifiée

Pour acheter une instance planifiée, utilisez l'assistant de réservation des instances réservées planifiées.

> Warning
>
> Une fois que vous avez acheté une instance planifiée, vous ne pouvez pas annuler, modifier ou revendre votre achat.

Pour acheter une instance planifiée (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous INSTANCES, sélectionnez Instances planifiées. Si la région actuellement sélectionnée ne prend pas en charge les instances planifiées, la page n'est pas disponible. En savoir plus
3. Sélectionnez Purchase Scheduled Instances (Acheter des instances planifiées).
4. Sur la page Find available schedules (Rechercher les programmes disponibles), procédez de la façon suivante :

   a. Sous Create a schedule (Créer un planning), sélectionnez la date de début dans Starting on (Début le), la récurrence de planification (quotidienne, hebdomadaire ou mensuelle) dans Recurring (Récurrent) et la durée minimale dans for duration (pour une durée). Notez que la console s'assure que vous spécifiez une valeur de durée minimale répondant à l'utilisation requise minimale pour votre instance planifiée (1 200 heures par an).

b. Sous Détails de l'instance, sélectionnez le système d'exploitation et le réseau dans Plateforme. Pour limiter les résultats, sélectionnez un ou plusieurs types d'instance dans Type d'instance ou une ou plusieurs zones de disponibilité dans Zone de disponibilité.

c. Sélectionnez Find schedules (Rechercher les plannings).

d. Sous Available schedules (Plannings disponibles), sélectionnez un ou plusieurs plannings. Pour chaque planning que vous sélectionnez, définissez la quantité d'instances et sélectionnez Ajouter au panier.

e. Votre panier est affiché en bas de la page. Lorsque vous avez fini d'ajouter ou de supprimer des plannings dans votre panier, sélectionnez Review and purchase (Vérifier et acheter).

5. Sur la page Review and purchase (Vérifier et acheter), vérifiez vos sélections et modifiez-les si nécessaire. Lorsque vous avez terminé, sélectionnez Acheter.

Pour acheter une instance planifiée (AWS CLI)

Utilisez la commande describe-scheduled-instance-availability pour afficher les horaires disponibles qui répondent à vos besoins, puis la commande purchase-scheduled-instances pour valider l'achat.

Lancement d'une instance planifiée

Une fois que vous achetez une instance planifiée, vous pouvez la lancer lors des périodes planifiées correspondantes.

Pour lancer une instance planifiée (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous INSTANCES, sélectionnez Instances planifiées. Si la région actuellement sélectionnée ne prend pas en charge les instances planifiées, la page n'est pas disponible. En savoir plus (p. 373)
3. Sélectionnez l'instance planifiée, puis Launch Scheduled Instances (Lancer les instance planifiées).
4. Sur la page Configurer, complétez la spécification de lancement pour vos instances planifiées et sélectionnez Vérification.

   Important

   La spécification de lancement doit correspondre aux attributs de type d'instance, de zone de disponibilité, de réseau et de plateforme du programme que vous avez acheté.

5. Sur la page Vérification, vérifiez la configuration du lancement et modifiez-la si nécessaire. Lorsque vous avez terminé, sélectionnez Lancer.

Pour lancer une instance planifiée (AWS CLI)

Utilisez la commande describe-scheduled-instances pour afficher vos instances planifiées, puis la commande run-scheduled-instances pour lancer chaque instance planifiée au cours de ses périodes planifiées.

Limites des instances planifiées

Les instances planifiées sont soumises aux limites suivantes :

- Les seuls types d'instance pris en charge sont les suivants : C3, C4, M4 et R3.
- La durée requise est 365 jours (un an).
- L'utilisation requise minimale est 1 200 heures par an.
- Vous pouvez acheter une instance planifiée jusqu'à 3 mois à l'avance.
- Elles sont disponibles dans les régions suivantes : USA Est (Virginie du Nord), USA Ouest (Oregon) et UE (Irlande).

Instances Spot

Une instance Spot est une instance EC2 non utilisée qui est disponible pour un prix inférieur à celui du prix à la demande. Comme une Instances Spot vous permet de demander des instances EC2 inutilisées avec de fortes remises, vous pouvez réduire considérablement vos coûts Amazon EC2. Le prix horaire d'une instance Spot est appelé prix spot. Le prix Spot de chaque type d'instance dans chaque zone de disponibilité est défini par Amazon EC2 et varie en fonction de l'offre et de la demande à long terme pour les Instances Spot. Votre instance Spot s'exécute chaque fois que la capacité est disponible et que le taux horaire maximal de votre demande dépasse le prix spot.

Les Instances Spot constituent un choix économique si vous êtes flexible quant au moment où vos applications s'exécutent et à la possibilité de les interrompre. Par exemple, les Instances Spot sont particulièrement adaptées à l'analyse de données, aux travaux par lots, au traitement en arrière-plan et aux tâches facultatives. Pour de plus amples informations, veuillez consulter Amazon EC2 Instances Spot.

Rubriques

- Concepts (p. 373)
- Comment démarrer (p. 375)
- Services connexes (p. 375)
- Tarification et économies (p. 376)

Concepts

Avant de commencer à utiliser les Instances Spot, vous devez connaître les concepts suivants :

- Pool d'instance Spot – Ensemble d'instances EC2 inutilisées ayant les mêmes type d'instance (`m5.large` par exemple), système d'exploitation, zone de disponibilité et plateforme réseau.
- Prix spot – Prix horaire actuel d'une instance Spot.
- Demande instance Spot – demande un instance Spot. La demande fournit le prix maximum par heure que vous êtes prêt à payer pour un instance Spot. Si vous ne spécifiez pas de prix maximum, la valeur par défaut est le prix À la demande. Lorsque le prix maximum par heure de votre demande dépasse le prix Spot, Amazon EC2 satisfait cette dernière si la capacité est disponible. Une demande instance Spot est unique ou persistante. Amazon EC2 soumet automatiquement à nouveau une demande instance Spot persistante après la fin de l'instance Spot associée à la demande résiliée. Votre demande d'instance Spot peut éventuellement indiquer une durée pour les Instances Spot.
- Parc d'instances Spot – Ensemble d'Instances Spot lancé en fonction des critères que vous spécifiez. Le Parc d'instances Spot sélectionne les groupes d'instance Spot correspondant à vos besoins et lance les Instances Spot de façon à répondre à la capacité cible du parc. Par défaut, un Parcs d'instances Spot est configuré pour maintenir la capacité cible en lançant des instances de remplacement après la résiliation

d'Instances Spot du parc. Vous pouvez soumettre un Parc d'instances Spot comme demande unique qui ne persiste pas après que les instances ont été résiliées. Vous pouvez inclure les demandes d'instance à la demande dans une demande de Parc d'instances Spot.

- Interruption d'instance Spot – Amazon EC2 supprime, arrête ou met en veille prolongée votre instance Spot lorsque le prix spot dépasse le prix maximum pour votre demande ou votre capacité n'est plus disponible. Amazon EC2 fournit un avis d'interruption instance Spot, qui prévient l'instance deux minutes avant qu'elle soit interrompue.

Principales différences entre les Instances Spot et les Instances à la demande

Le tableau suivant répertorie les principales différences entre les Instances Spot et les Instances à la demande.

	Instances Spot	Instances à la demande
Heure de lancement	Peut uniquement être lancé immédiatement si la demande Spot est active et que la capacité est disponible.	Peut uniquement être lancé immédiatement si vous émettez une demande de lancement manuel et que la capacité est disponible.
Capacité disponible	Si la capacité n'est pas disponible, la demande Spot continue à effectuer automatiquement la demande de lancement jusqu'à ce que la capacité devienne disponible.	Si la capacité n'est pas disponible lorsque vous effectuez une demande de lancement, vous obtenez une erreur de capacité insuffisante (ICE).
Tarif horaire	Le prix horaire pour les Instances Spot varie en fonction de la demande.	Le prix horaire pour les Instances à la demande est statique.
Interruption d'instance	Vous pouvez arrêter et démarrer une instance Spot basée sur Amazon EBS. De plus, le service Spot Amazon EC2 peut interrompre (p. 452) une instance Spot individuelle si la capacité n'est plus disponible, si le prix Spot dépasse votre prix maximum ou si la demande pour les Instances Spot augmente.	Vous déterminez le moment où une instance à la demande est interrompue (arrêtée, mise en veille prolongée ou résiliée).

Stratégies d'utilisation des Instances Spot

Pour maintenir un niveau minimal de ressources de calcul garanties pour vos applications, vous pouvez utiliser la stratégie qui consiste à lancer un groupe principal d'Instances à la demande et les compléter par des Instances Spot lorsque l'occasion se présente.

Une autre stratégie consiste à lancer Instances Spot avec une durée spécifiée (également connus sous le nom de blocs Spot), qui sont conçus pour ne pas être interrompus et fonctionneront en continu pendant la durée que vous choisissez. Dans de rares cas, les blocs d'instances Spot peuvent être interrompus en raison des besoins en capacités Amazon EC2. Dans ces cas-là, nous fournirons un avertissement de deux minutes avant de mettre fin à une instance, et vous ne serez pas facturé pour les instances résiliées, même si vous les avez utilisées. Pour plus d'informations, consultez Définition d'une durée pour vos Instances Spot (p. 392).

Comment démarrer

Vous devez commencer par préparer l'utilisation d'Amazon EC2. Il peut être utile d'avoir déjà eu l'occasion de lancer des Instances à la demande avant de lancer des Instances Spot.

Etre opérationnel

- Configuration avec Amazon EC2 (p. 28)
- Didacticiel : Premiers pas avec les instances Amazon EC2 Linux (p. 32)

Bases des instances Spot

- Fonctionnement des Instances Spot (p. 379)
- Fonctionnement de Parc d'instances Spot (p. 380)

Utilisation de Instances Spot

- Préparation aux interruptions (p. 456)
- Création d'une demande d'instance Spot (p. 395)
- Obtention d'informations sur le statut de la demande (p. 450)

Utilisation de Parcs d'instances Spot

- Autorisations Parc d'instances Spot (p. 411)
- Création d'une demande de Parc d'instances Spot (p. 416)

Services connexes

Vous pouvez allouer des Instances Spot directement à partir d'Amazon EC2. Vous pouvez également mettre en service des Instances Spot à partir d'autres services AWS. Pour plus d'informations, consultez la documentation suivante.

Amazon EC2 Auto Scaling et Instances Spot

Vous pouvez créer des modèles ou des configurations de lancement avec le prix maximum que vous êtes disposé à payer, de façon à ce qu'Amazon EC2 Auto Scaling puisse lancer des Instances Spot. Pour plus d'informations, voir Lancement Instances Spot dans votre Auto Scaling groupe et Utilisation de plusieurs types d'instance et options d'achat dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Amazon EMR et Instances Spot

Dans certains cas de figure, il peut être utile d'exécuter des Instances Spot dans un cluster Amazon EMR. Pour plus d'informations, consultez Instances Spot et Quand faut-il utiliser des Instances Spot ? dans le Amazon EMR Guide de gestion.

Modèles AWS CloudFormation

AWS CloudFormation vous permet de créer et de gérer un ensemble de ressources AWS à l'aide d'un modèle au format JSON. Les modèles AWS CloudFormation peuvent inclure le prix maximum que vous êtes disposé à payer. Pour plus d'informations, consultez EC2 instance Spot Updates - Auto Scaling and CloudFormation Integration.

AWS SDK for Java

Vous pouvez utiliser le langage de programmation Java afin de gérer vos Instances Spot. Pour plus d'informations, consultez Didacticiel : Instances Spot Amazon EC2 et Didacticiel : Gestion avancée des demandes Spot Amazon EC2.

Kit AWS SDK pour .NET

Vous pouvez utiliser l'environnement de programmation .NET pour gérer vos Instances Spot. Pour plus d'informations, consultez Didacticiel : Instances Spot Amazon EC2.

Tarification et économies

Vous payez le prix spot des Instances Spot qui est défini par Amazon EC2 et varie régulièrement en fonction de l'offre et de la demande à long terme pour les Instances Spot. Si le prix maximum par heure de votre demande dépasse le prix Spot actuel, Amazon EC2 satisfait cette dernière si la capacité est disponible. Vos Instances Spot sont exécutées jusqu'à ce que vous les arrêtiez, dès que la capacité n'est plus disponible, que le prix Spot dépasse votre prix maximum ou que votre groupe Amazon EC2 Auto Scaling les arrête durant le dimensionnement.

Les Instances Spot à durée prédéfinie utilisent un tarif horaire fixe qui reste en vigueur pour l'instance Spot pendant la durée de son exécution.

Si vous-même ou Amazon EC2 interrompez une instance Spot en cours d'exécution, vous serez facturé pour les secondes utilisées ou l'heure complète, ou vous ne serez pas facturé, selon le système d'exploitation utilisé et qui a interrompu l'instance Spot. Pour plus d'informations, consultez Facturation des Instances Spot interrompues (p. 460).

Consulter les tarifs

Pour consulter le prix spot actuellement le moins élevé par région AWS et type d'instance (information mise à jour toutes les cinq minutes), lisez la page Tarification des Instances Spot.

Pour consulter l'historique du prix spot au cours des trois derniers mois, utilisez la console Amazon EC2 ou la commande describe-Spot-price-history (AWS CLI). Pour plus d'informations, consultez Historique de tarification d'instance Spot (p. 387).

Nous mappons indépendamment les zones de disponibilité avec les codes pour chaque compte AWS. Par conséquent, vous pouvez obtenir des résultats variables pour un même code de zone de disponibilité (par exemple, `us-west-2a`) entre différents comptes.

Consulter les économies

Vous pouvez afficher les économies réalisées à partir de l'utilisation de Instances Spot pour un Parc d'instances Spot unique ou pour toutes les Instances Spot. Vous pouvez consulter les économies réalisées au cours de la dernière heure ou des trois derniers jours, ainsi que le coût moyen par heure vCPU et par heure de mémoire (Gio). Les économies sont des estimations et peuvent différer de vos économies réelles, car elles ne tiennent pas compte des ajustements de facturation en fonction de votre utilisation. Pour plus d'informations sur la consultation des informations sur les économies, consultez Économies réalisées grâce à l'achat d'Instances Spot (p. 388).

Consulter les factures

Pour consulter votre facture, accédez à la page Activité du compte AWS. Votre facture contient des liens vers les rapports d'utilisation qui fournissent des détails sur votre facture. Pour plus d'informations, consultez Facturation du compte AWS.

Pour toute question relative à la facturation, aux comptes et aux événements AWS, contactez AWS Support.

Bonnes pratiques pour EC2 Spot

Amazon EC2 Instances Spot sont des capacités de calcul EC2 de rechange dans le Cloud AWS qui vous sont offertes avec jusqu'à 90 % de réduction par rapport aux prix à la demande. La seule différence entre Instances à la demande et Instances Spot est que Instances Spot peut être interrompu par Amazon EC2, avec deux minutes de notification, quand Amazon EC2 a à nouveau besoin de la capacité.

Les Instances Spot sont recommandés pour les applications flexibles sans état, tolérantes aux pannes. Par exemple, Instances Spot fonctionne bien pour le Big Data, les charges de travail conteneurisées, les CI/CD, les serveurs Web sans état, le calcul haute performance (HPC) et les charges de travail de rendu.

En cours d'exécution, les Instances Spot sont exactement les mêmes que Instances à la demande. Toutefois, Spot ne garantit pas que vous pouvez conserver vos instances en cours d'exécution suffisamment longtemps pour terminer vos charges de travail. Spot ne garantit pas non plus que vous pouvez obtenir la disponibilité immédiate des instances que vous recherchez, ou que vous pouvez toujours obtenir la capacité globale que vous avez demandée. De plus, les interruptions et la capacité instance Spot peuvent changer au fil du temps, car la disponibilité instance Spot varie en fonction de l'offre et de la demande, et les performances passées ne sont pas une garantie de résultats futurs.

Les Instances Spot ne conviennent pas aux charges de travail inflexibles, dynamiques, intolérantes aux pannes ou étroitement couplées entre des nœuds d'instance. Il n'est pas non plus recommandé pour les charges de travail qui ne tolèrent pas les périodes occasionnelles où la capacité cible n'est pas complètement disponible. Nous mettons fortement en garde contre l'utilisation de Instances Spot pour ces charges de travail ou la tentative de basculement vers Instances à la demande pour gérer les interruptions.

Que vous soyez un utilisateur Spot expérimenté ou un nouvel utilisateur Instances Spot, si vous rencontrez actuellement des problèmes avec les interruptions ou la disponibilité de instance Spot, nous vous recommandons de suivre ces meilleures pratiques pour bénéficier de la meilleure expérience d'utilisation du service Spot.

Bonnes pratiques en matière d'instances Spot

- Préparer des instances individuelles pour les interruptions (p. 377)
- Soyez flexible en ce qui concerne les types d'instance et les zones de disponibilité (p. 378)
- Utilisez les groupes Auto Scaling EC2 ou Parc d'instances Spot pour gérer votre capacité agrégée (p. 378)
- Utiliser la stratégie d'allocation optimisée pour la capacité (p. 378)
- Utilisez des services AWS intégrés pour gérer votre Instances Spot (p. 378)

Préparer des instances individuelles pour les interruptions

La meilleure façon pour vous de gérer gracieusement les interruptions instance Spot consiste à concevoir votre application pour qu'elle soit tolérante aux pannes. Pour ce faire, vous pouvez profiter des avis d'interruption instance Spot. Un avis d'interruption instance Spot est un avertissement émis deux minutes avant que Amazon EC2 interrompe instance Spot. Nous vous recommandons de créer une règle Amazon EventBridge qui capture la notification d'interruption, puis déclenche un point de contrôle pour la progression de votre charge de travail ou gère l'interruption de manière gracieuse. Pour obtenir un exemple détaillé qui vous explique comment créer et utiliser des règles d'événement, veuillez consulter Tirer parti des avis d'interruption d'instance Spot Amazon EC2.

Si votre charge de travail est « flexible dans le temps », vous pouvez également configurer votre Instances Spot pour qu'elle soit arrêtée ou mise en veille prolongée lorsqu'elle est interrompue. Amazon EC2 arrête ou met en hibernation automatiquement votre Instances Spot lors de l'interruption et reprend automatiquement les instances lorsque nous disposons de la capacité disponible.

Pour plus d'informations, consultez Interruptions d'une instance Spot (p. 452).

Soyez flexible en ce qui concerne les types d'instance et les zones de disponibilité

Un pool instance Spot est un ensemble d'instances EC2 inutilisées avec le même type d'instance (par exemple, m5.large) et la même zone de disponibilité (par exemple, us-east-1a). Vous devez être flexible quant aux types d'instance que vous demandez et aux zones de disponibilité dans lesquelles vous pouvez déployer votre charge de travail. Cela donne à Spot une meilleure chance de trouver et d'allouer la quantité requise de capacité de calcul. Par exemple, ne demandez pas simplement `c5.large` si vous seriez prêt à utiliser des larges des familles c4, m5 et m4.

En fonction de vos besoins spécifiques, vous pouvez évaluer les types d'instance que vous pouvez utiliser pour répondre à vos besoins de calcul. Si une charge de travail peut être mise à l'échelle verticale, vous devez inclure des types d'instance plus importants (plus de vCPU et de mémoire) dans vos requêtes. Si vous ne pouvez évoluer qu'horizontalement, vous devez inclure des types d'instance de génération plus ancienne car ils sont moins demandés par les clients à la demande.

Une bonne règle générale est d'être flexible sur au moins 10 types d'instance pour chaque charge de travail. En outre, assurez-vous que toutes les zones de disponibilité sont configurées pour être utilisées dans votre VPC et sélectionnées pour votre charge de travail.

Utilisez les groupes Auto Scaling EC2 ou Parc d'instances Spot pour gérer votre capacité agrégée

Spot vous permet de penser en termes de capacité agrégée, dans des unités comprenant des processeurs virtuels, de la mémoire, du stockage ou du débit réseau, plutôt que de penser en termes d'instances individuelles. Les groupes Auto Scaling et Parc d'instances Spot vous permettent de lancer et de gérer une capacité cible, et de demander automatiquement des ressources pour remplacer celles qui sont interrompues ou terminées manuellement. Lorsque vous configurez un groupe Auto Scaling ou un Parc d'instances Spot, il vous suffit de spécifier les types d'instance et la capacité cible en fonction des besoins de votre application. Pour de plus amples informations, veuillez consulter Groupes de mise à l'échelle automatique dans le Amazon EC2 Auto Scaling Guide de l'utilisateur et Création d'une demande de Parc d'instances Spot (p. 416) dans ce guide de l'utilisateur.

Utiliser la stratégie d'allocation optimisée pour la capacité

Les stratégies d'allocation dans les groupes Auto Scaling vous aident à provisionner votre capacité cible sans avoir à rechercher manuellement les pools instance Spot avec une capacité de réserve. Nous vous recommandons d'utiliser la stratégie `capacity optimized` car elle provisionne automatiquement les instances des pools instance Spot les plus disponibles. Vous pouvez également profiter de la stratégie d'allocation `capacity optimized` dans Parc d'instances Spot. Étant donné que votre capacité instance Spot provient de pools avec une capacité optimale, cela réduit la possibilité que votre capacité Instances Spot soit récupérée. Pour de plus amples informations sur les stratégies d'allocation, veuillez consulter Instances Spot dans le Amazon EC2 Auto Scaling Guide de l'utilisateur et Configuration du Parc d'instances Spot pour l'optimisation de la capacité (p. 382) dans ce guide de l'utilisateur.

Utilisez des services AWS intégrés pour gérer votre Instances Spot

D'autres services AWS s'intègrent à Spot pour réduire les coûts de calcul globaux sans avoir à gérer les instances ou les parcs individuels. Nous vous recommandons d'envisager les solutions suivantes pour vos charges de travail applicables : Amazon EMR, Amazon ECS, AWS Batch, Amazon EKS, SageMaker, AWS Elastic Beanstalk et Amazon GameLift. Pour de plus amples informations sur les meilleures pratiques Spot avec ces services, veuillez consulter le site Web sur les ateliers Amazon EC2Instances Spot.

Fonctionnement des Instances Spot

Pour lancer une instance Spot, vous créez une demande d'instance Spot ou Amazon EC2 crée une demande d'instance Spot en votre nom. L'instance Spot est lancée lorsque la demande d'instance Spot est satisfaite.

Vous pouvez lancer une instance Spot en utilisant plusieurs services différents. Pour de plus amples informations, veuillez consulter Mise en route avec les instances ponctuelles Amazon EC2. Dans ce guide de l'utilisateur, nous décrivons les façons suivantes de lancer une instance Spot à l'aide d'EC2 :

- Vous pouvez créer une demande d'instance Spot. Pour de plus amples informations, veuillez consulter Création d'une demande d'instance Spot (p. 395).
- Vous pouvez créer une demande de Parc d'instances Spot, dans laquelle vous spécifiez le nombre souhaité d'Instances Spot. Amazon EC2 crée une demande d'instance Spot en votre nom pour chaque instance Spot spécifiée dans la demande de Parc d'instances Spot. Pour de plus amples informations, veuillez consulter Création d'une demande de Parc d'instances Spot (p. 416).
- Vous pouvez créer une Flotte EC2, dans laquelle vous spécifiez le nombre souhaité d'Instances Spot. Amazon EC2 crée une demande d'instance Spot en votre nom pour chaque instance Spot spécifiée dans la Flotte EC2. Pour de plus amples informations, veuillez consulter Création d'un Flotte EC2 (p. 579).

La demande d'instance Spot doit inclure le prix maximum que vous êtes prêt à payer par heure et par instance. Si vous ne spécifiez pas de prix, le prix par défaut est le prix À la demande. La demande peut inclure d'autres contraintes telles que le type d'instance et la zone de disponibilité.

Votre instance Spot est lancée si le prix maximum que vous êtes prêt à payer dépasse le prix Spot et si des capacités sont disponibles. Si le prix maximum que vous êtes prêt à payer est inférieur au prix Spot, votre instance n'est pas lancée. Toutefois, étant donné que le service Amazon EC2 ajuste progressivement le prix Spot en fonction de l'offre et de la demande d'Instances Spot à long terme, le prix maximum que vous êtes prêt à payer peut finalement dépasser le prix Spot. Dans ce cas, votre instance sera lancée.

Votre instance Spot s'exécute jusqu'à ce que vous l'arrêtiez ou la résiliiez, ou jusqu'à ce que le service Amazon EC2 l'interrompe (il s'agit alors d'une interruption d'instance Spot).

Lorsque vous utilisez Instances Spot, vous devez être préparé aux interruptions. Amazon EC2 peut interrompre votre instance Spot dès que le prix spot de cette dernière dépasse le prix maximum, lorsque la demande pour des Instances Spot augmente ou lorsque l'offre d'Instances Spot baisse. Lorsqu'Amazon EC2 interrompt une instance Spot, il communique un avis d'interruption d'instance Spot, ce qui avertit l'instance qu'Amazon EC2 va l'interrompe dans deux minutes. Vous ne pouvez pas activer la protection de la résiliation pour les Instances Spot. Pour plus d'informations, consultez Interruptions d'une instance Spot (p. 452).

Vous pouvez arrêter, démarrer, redémarrer ou résilier une instance Spot basée sur Amazon EBS. Le service Spot peut arrêter, résilier ou mettre en veille prolongée une instance Spot lorsqu'il l'interrompt.

Sommaire

- Lancement d'Instances Spot dans un groupe de lancement (p. 379)
- Lancement d'Instances Spot dans un groupe de zone de disponibilité (p. 380)
- Lancement d'Instances Spot sur un VPC (p. 380)

Lancement d'Instances Spot dans un groupe de lancement

Spécifiez un groupe de lancement dans votre demande d'instance Spot pour indiquer à Amazon EC2 de lancer un ensemble d'Instances Spot uniquement s'il peut toutes les lancer. De plus, si le service Spot doit mettre hors service l'une des instances du groupe de lancement (par exemple, si le prix Spot est supérieur au prix maximum), il doit toutes les mettre hors service. Toutefois, si vous mettez hors service

une ou plusieurs instances d'un groupe de lancement, Amazon EC2 ne met pas hors service les instances restantes du groupe de lancement.

Même si cette option peut être utile, l'ajout d'une contrainte de ce type peut réduire les chances de voir votre demande d'instance Spot satisfaite et accroître les risques de suppression de vos Instances Spot. Par exemple, votre groupe de lancement inclut des instances figurant dans plusieurs zones de disponibilité. Si la capacité dans l'une de ces zones de disponibilité diminue et n'est plus disponible, Amazon EC2 résilie alors toutes les instances du groupe de lancement.

Si vous créez une autre demande d'instance Spot réussie qui spécifie le même groupe de lancement (existant) qu'une demande précédente réussie, les nouvelles instances sont ajoutées au groupe de lancement. Par conséquent, si une instance de ce groupe de lancement est mise hors service, toutes les instances du groupe de lancement sont également mises hors service, ce qui inclut les instances lancées par les première et deuxième demandes.

Lancement d'Instances Spot dans un groupe de zone de disponibilité

Spécifiez un groupe de zone de disponibilité dans votre demande d'instance Spot afin de demander au service d'Instances Spot de lancer un ensemble d'instances Spot dans la même zone de disponibilité. Amazon EC2 n'a pas besoin de mettre hors service toutes les instances d'un groupe de zone de disponibilité en même temps. Si Amazon EC2 doit interrompre l'une des instances d'un groupe de zone de disponibilité, les autres continuent à être exécutées.

Même si cette option peut s'avérer utile, l'ajout d'une contrainte de ce type peut réduire les chances de voir votre demande d'instance Spot satisfaite.

Si vous spécifiez un groupe de zone de disponibilité, mais que vous n'indiquez aucune zone de disponibilité dans la demande d'instance Spot, le résultat dépend du réseau que vous avez spécifié.

VPC par défaut

Amazon EC2 utilise la zone de disponibilité pour le sous-réseau spécifié. Si vous ne spécifiez pas de sous-réseau, le service sélectionne une zone de disponibilité et son sous-réseau par défaut, mais pas nécessairement la zone ayant le prix le plus bas. Si vous avez supprimé le sous-réseau par défaut pour une zone de disponibilité, vous devez spécifier un autre sous-réseau.

VPC personnalisé

Amazon EC2 utilise la zone de disponibilité pour le sous-réseau spécifié.

Lancement d'Instances Spot sur un VPC

Vous spécifiez un sous-réseau pour vos Instances Spot de la même façon que vous spécifiez un sous-réseau pour vos Instances à la demande.

- Vous devez utiliser le prix maximum par défaut (le prix à la demande) ou indiquer le prix maximum de l'historique des prix spot des Instances Spot d'un VPC.
- [VPC par défaut] Si vous souhaitez que votre instance Spot soit lancée dans une zone de disponibilité à faible prix, vous devez spécifier le sous-réseau correspondant dans votre demande d'instance Spot. Si vous ne spécifiez pas de sous-réseau, Amazon EC2 en sélectionne un pour vous et la zone de disponibilité de ce sous-réseau ne correspondra peut-être pas au prix Spot le plus faible.
- [VPC personnalisé] Vous devez spécifier le sous-réseau de votre instance Spot.

Fonctionnement de Parc d'instances Spot

Un Parc d'instances Spot est un ensemble, ou parc, d'Instances Spot et, le cas échéant, d'Instances à la demande.

Le Parc d'instances Spot tente de lancer le nombre d'Instances Spot et d'Instances à la demande pour satisfaire la capacité cible que vous avez spécifiée dans la demande du Parc d'instances Spot. La demande des Instances Spot est satisfaite si le prix maximal que vous avez spécifié dans la demande dépasse le prix spot actuel et si la capacité est disponible. Parc d'instances Spot tente également de préserver le parc de capacité cible si Instances Spot est interrompu.

Vous pouvez également définir le montant maximum que vous être prêt à payer par heure pour votre parc et Parc d'instances Spot lance les instances jusqu'à ce que le montant maximum soit atteint. Une fois le montant maximum que vous être prêt à payer atteint, le parc arrête de lancer des instances même si la capacité cible n'a pas été atteinte.

Un pool d'instance Spot est un ensemble d'instances EC2 inutilisées ayant les mêmes type d'instance (`m5.large` par exemple), système d'exploitation, zone de disponibilité et plateforme réseau. Lorsque vous effectuez une demande de Parc d'instances Spot, vous pouvez inclure plusieurs spécifications de lancement qui varient selon le type d'instance, l'AMI, la zone de disponibilité ou le sous-réseau. Le Parc d'instances Spot sélectionne les groupes d'instance Spot servant à satisfaire la demande, selon les spécifications de lancement incluses dans votre demande Parc d'instances Spot et la configuration de cette demande Parc d'instances Spot. Les Instances Spot proviennent des groupes sélectionnés.

Sommaire

- À la demande dans Parc d'instances Spot (p. 381)
- Stratégie d'allocation pour les Instances Spot (p. 381)
- Remplacements du prix Spot (p. 383)
- Contrôle des dépenses (p. 383)
- Pondération d'instance de Parc d'instances Spot (p. 384)
- Démonstration : Utilisation d'un Parc d'instances Spot avec pondération des instances (p. 385)

À la demande dans Parc d'instances Spot

Pour garantir que vous avez toujours la capacité d'instance, vous pouvez inclure une capacité à la demande dans votre demande de Parc d'instances Spot. Dans votre demande de Parc d'instances Spot, vous spécifiez votre capacité cible souhaitée et quelle quantité de la capacité doit être à la demande. Le solde inclut la capacité Spot qui est lancée si une capacité et une disponibilité Amazon EC2 sont disponibles. Par exemple, si dans votre demande de Parc d'instances Spot, vous spécifiez une capacité de cible de 10 et une capacité à la demande de 8, Amazon EC2 lance 8 unités de capacité à la demande et 2 unités de capacité (10 - 8 = 2) comme unités Spot.

Hiérarchisation des types d'instance pour la capacité à la demande

Lorsque le Parc d'instances Spot essaie de traiter l'affectation de capacité à la demande, il lance par défaut le type d'instance dont le prix est le plus bas en premier. Si `OnDemandAllocationStrategy` a pour valeur `prioritized`, le Parc d'instances Spot utilise la priorité pour déterminer quel type d'instance utiliser en premier afin de traiter l'affectation de capacité à la demande. La priorité est affectée au remplacement du modèle de lancement, et la priorité la plus élevée est lancée en premier.

Par exemple, vous avez configuré trois remplacements du modèle de lancement, ayant chacun un type d'instance différent : `c3.large`, `c4.large` et `c5.large`. Le prix à la demande d'une instance `c5.large` est inférieur à celui d'une instance `c4.large`. L'instance `c3.large` est la moins chère. Si vous n'utilisez pas la priorité pour déterminer l'ordre, le parc traite l'affectation de capacité à la demande en commençant par `c3.large`, puis `c5.large`. Étant donné que vous avez souvent des Instances réservées non utilisées pour `c4.large`, vous pouvez définir la priorité du remplacement du modèle de lancement afin que l'ordre soit `c4.large`, `c3.large`, puis `c5.large`.

Stratégie d'allocation pour les Instances Spot

La stratégie d'allocation des Instances Spot de votre Parc d'instances Spot détermine la façon dont la demande de Parc d'instances Spot est satisfaite à partir des groupes d'instance Spot indiqués dans les

spécifications de lancement. Voici les stratégies d'allocation que vous pouvez spécifier dans une demande de Parc d'instances Spot :

`lowestPrice`

Les Instances Spot proviennent du groupe offrant le prix le plus bas. Il s'agit de la stratégie par défaut.

`diversified`

Les Instances Spot sont réparties entre tous les groupes.

`capacityOptimized`

Les Instances Spot proviennent du pool d'instances Spot avec une capacité optimale pour le nombre d'instances qui sont lancées.

`InstancePoolsToUseCount`

Les Instances Spot sont réparties entre le nombre de groupes d'instances Spot que vous spécifiez. Ce paramètre n'est valide que s'il est utilisé conjointement avec `lowestPrice`.

Maintien de la capacité cible

Une fois les Instances Spot résiliées en raison d'un changement de prix spot ou de la modification de la capacité disponible d'un groupe d'instance Spot, un Parc d'instances Spot de type `maintain` lance des Instances Spot de remplacement. Si la stratégie d'allocation `lowestPrice` est sélectionnée, le parc d'instances lance les instances de remplacement dans le pool où le prix Spot est actuellement le plus faible. Si la stratégie d'allocation est `diversified`, le parc d'instances répartit les Instances Spot de remplacement entre les groupes restants. Si la stratégie d'allocation est `lowestPrice` conjointement avec `InstancePoolsToUseCount`, le parc sélectionne les groupes d'instances Spot ayant le prix le plus bas et lance les Instances Spot sur le nombre de groupes d'instances Spot que vous spécifiez.

Configuration du Parc d'instances Spot pour l'optimisation des coûts

Pour optimiser les coûts relatifs à votre utilisation des Instances Spot, spécifiez la stratégie d'allocation `lowestPrice` pour que le Parc d'instances Spot déploie automatiquement la combinaison la plus économique de types d'instance et de zones de disponibilité en fonction du prix Spot actuel.

Pour la capacité cible des instance à la demande, le Parc d'instances Spot sélectionne toujours le type d'instance le moins cher en fonction du prix à la demande public, tout en continuant à suivre la stratégie d'allocation (`lowestPrice`, `capacityOptimized` ou `diversified`) pour les Instances Spot.

Configuration du Parc d'instances Spot pour l'optimisation des coûts et la diversification

Pour créer un parc d'Instances Spot à la fois peu onéreux et diversifié, utilisez la stratégie d'allocation `lowestPrice` conjointement avec `InstancePoolsToUseCount`. Le Parc d'instances Spot déploie automatiquement la combinaison la plus économique de types d'instance et de zones de disponibilité en fonction du prix spot actuel sur le nombre de groupes d'instances Spot que vous spécifiez. Il est possible d'utiliser cette combinaison pour éviter les Instances Spot les plus onéreuses.

Configuration du Parc d'instances Spot pour l'optimisation de la capacité

Avec les Instances Spot, la tarification change lentement au fil du temps en fonction des tendances à long terme en matière d'offre et de demande, mais la capacité fluctue en temps réel. La stratégie `capacityOptimized` lance automatiquement des Instances Spot dans les pools les plus disponibles en examinant les données de capacité en temps réel et en prédisant les instances les plus disponibles. Cela convient parfaitement aux charges de travail telles que le Big Data et l'analyse, le rendu d'images et de médias, l'apprentissage automatique et le calcul haute performance qui peuvent avoir un coût d'interruption plus élevé, associé au redémarrage des tâches et aux points de contrôle. En offrant la possibilité de moins d'interruptions, la stratégie `capacityOptimized` peut réduire le coût global de votre charge de travail.

Choix d'une stratégie d'allocation appropriée

Vous pouvez optimiser vos Parcs d'instances Spot en fonction de votre cas d'utilisation.

Si votre parc est de petite taille ou qu'il s'exécute pendant un temps limité, il y a peu de risques que vos Instances Spot soient interrompues, même si toutes les instances se trouvent dans un même groupe d'instance Spot. C'est pourquoi la stratégie `lowestPrice` répondra certainement à vos besoins pour le prix le plus faible.

Si votre parc est important ou qu'il s'exécute pendant une longue durée, vous pouvez améliorer sa disponibilité en répartissant les Instances Spot entre plusieurs groupes. Par exemple, si votre demande de Parc d'instances Spot spécifie 10 groupes et une capacité cible de 100 instances, le parc lance 10 Instances Spot dans chaque groupe. Si le prix Spot d'un pool dépasse le prix maximum de ce pool, seul 10 % de votre parc est touché. Avec cette stratégie, votre parc est également moins affecté par les augmentations du prix Spot dans un pool au fil du temps.

Avec la stratégie `diversified`, le Parc d'instances Spot ne lance pas d'Instances Spot dans des groupes dont le prix spot est supérieur ou égal au prix à la demande.

Pour créer un parc peu onéreux et diversifié, utilisez la stratégie `lowestPrice` conjointement avec `InstancePoolsToUseCount`. Vous pouvez utiliser un nombre plus ou moins élevé de groupes d'instances Spot auxquels allouer vos Instances Spot. Par exemple, si vous exécutez un traitement par lots, nous vous recommandons de spécifier un petit nombre de groupes d'instances Spot (par exemple, `InstancePoolsToUseCount=2`) pour garantir que votre file d'attente aura toujours une capacité de calcul suffisante tout en optimisant les économies. Si vous exécutez un service web, nous vous recommandons de spécifier un nombre élevé de groupes d'instances Spot (par exemple, `InstancePoolsToUseCount=10`) pour minimiser l'impact si un groupe d'instance Spot devient temporairement indisponible.

Si votre parc exécute des charges de travail dont l'interruption entraîne des coûts plus élevés associés au redémarrage du travail et aux points de contrôle, utilisez la stratégie `capacityOptimized`. Cette stratégie offre la possibilité de moins d'interruptions, ce qui peut réduire le coût global de votre charge de travail.

Remplacements du prix Spot

Chaque demande de Parc d'instances Spot peut inclure un prix maximum global ou utiliser la valeur par défaut (prix à la demande). Le Parc d'instances Spot utilise ce prix comme prix maximum par défaut pour chacun de ses spécifications de lancement.

Si vous le souhaitez, vous pouvez également spécifier un prix maximum dans une ou plusieurs spécifications de lancement. Ce prix est propre à la spécification de lancement. Si une spécification de lancement comprend un prix spécifique, le Parc d'instances Spot utilise ce prix maximum à la place du prix maximum global. Toute autre spécification de lancement qui ne comprend pas de prix maximum spécifique continue à utiliser le prix maximum global.

Contrôle des dépenses

Le Parc d'instances Spot arrête le lancement des instances une fois la capacité cible ou le montant maximum que vous êtes prêt à payer atteint(e). Pour contrôler le montant payé par heure pour votre parc, vous pouvez spécifier `SpotMaxTotalPrice` pour Instances Spot et `OnDemandMaxTotalPrice` pour Instances à la demande. Une fois le prix total atteint, le Parc d'instances Spot arrête de lancer des instances même si la capacité cible n'a pas été atteinte.

Les exemples suivants montrent deux manières de le faire : Dans le premier, le Parc d'instances Spot arrête de lancer des instances une fois la capacité cible atteinte. Dans le deuxième, le Parc d'instances Spot arrête le lancement des instances une fois le montant maximum que vous êtes prêt à payer atteint.

Exemple : Arrêt du lancement des instances lorsque la capacité cible est atteinte

Prenons l'exemple d'une demande pour `m4.large` Instances à la demande, avec :

- Prix à la demande : 0,10 USD par heure
- `OnDemandTargetCapacity` : 10
- `OnDemandMaxTotalPrice` : 1,50 USD

Parc d'instances Spot lance 10 Instances à la demande car le total de 1 USD (10 instances x 0,10 USD) ne dépasse pas le `OnDemandMaxTotalPrice` de 1,50 USD.

Exemple : Arrêt du lancement des instances lorsque le prix total maximum est atteint

Prenons l'exemple d'une demande pour `m4.large` Instances à la demande, avec :

- Prix à la demande : 0,10 USD par heure
- `OnDemandTargetCapacity` : 10
- `OnDemandMaxTotalPrice` : 0,80 USD

Si Parc d'instances Spot lance la capacité cible à la demande (10 Instances à la demande), le coût total par heure est de 1 USD. Ce montant est supérieur à celui (0,80 USD) spécifié pour `OnDemandMaxTotalPrice`. Afin d'éviter de dépenser plus que vous le souhaitez, Parc d'instances Spot lance uniquement 8 Instances à la demande (ce qui est inférieur à la capacité cible à la demande) car le lancement d'instances supplémentaires dépasserait `OnDemandMaxTotalPrice`.

Pondération d'instance de Parc d'instances Spot

Lorsque vous demandez un parc d'Instances Spot, vous pouvez définir les unités de capacité que chaque type d'instance apporterait aux performances de votre application et ajuster votre prix maximum pour chaque groupe d'instance Spot en conséquence à l'aide de la pondération d'instance.

Par défaut, le prix que vous spécifiez représente le prix par heure d'instance. Lorsque vous utilisez la fonction de pondération d'instance, le prix que vous spécifiez correspond au prix par heure d'unité. Vous pouvez calculer le prix de l'offre par heure d'unité en divisant le prix de l'offre pour un type d'instance par le nombre d'unités qu'il représente. Le Parc d'instances Spot calcule le nombre d'Instances Spot à lancer en divisant la capacité cible par la pondération de l'instance. Si le résultat n'est pas un entier, le Parc d'instances Spot l'arrondit à l'entier suivant afin que la taille de votre parc ne soit pas inférieure à sa capacité cible. Le Parc d'instances Spot peut sélectionner n'importe quel groupe indiqué dans votre spécification de lancement, même si la capacité des instances lancées dépasse la capacité cible demandée.

Les tableaux suivants présentent des exemples de calcul afin de déterminer le prix par unité pour une demande de Parc d'instances Spot ayant une capacité cible de 10.

Type d'instance	Pondération de l'instance	Prix par heure d'instance	Prix par heure d'unité	Nombre d'instances lancées
`r3.xlarge`	2	0,05 USD	0,025 (0,05 divisé par 2)	5 (10 divisé par 2)

Type d'instance	Pondération de l'instance	Prix par heure d'instance	Prix par heure d'unité	Nombre d'instances lancées
`r3.8xlarge`	8	0,10 USD	0,0125 (0,10 divisé par 8)	2 (10 divisé par 8, résultat arrondi)

Utilisez la pondération d'instance de Parc d'instances Spot comme suit, afin de mettre en service la capacité cible que vous voulez dans les groupes selon le prix par unité le plus bas au moment de l'exécution :

1. Définissez la capacité cible de votre Parc d'instances Spot en instances (valeur par défaut) ou dans les unités de votre choix, par exemple les UC, la mémoire, le stockage ou le débit virtuels.
2. Définissez le prix par unité.
3. Pour chaque configuration de lancement, spécifiez la pondération, c'est-à-dire le nombre d'unités du type d'instance par rapport à la capacité cible.

Exemple de pondération d'instance

Prenons l'exemple d'une demande de Parc d'instances Spot avec la configuration suivante :

- Capacité cible de 24
- Spécification de lancement avec le type d'instance `r3.2xlarge` et une pondération de 6
- Spécification de lancement avec le type d'instance `c3.xlarge` et une pondération de 5

La pondération correspond au nombre d'unités du type d'instance par rapport à la capacité cible. Si la première spécification de lancement fournit le prix par unité le plus faible (prix pour `r3.2xlarge` par heure d'instance divisé par 6), le parc d'instances Spot lance quatre de ces instances (24 divisé par 6).

Si la deuxième spécification de lancement fournit le prix par unité le plus bas (prix pour `c3.xlarge` par heure d'instance divisé par 5), le Parc d'instances Spot lance cinq de ces instances (24 divisé par 5, résultat arrondi).

Pondération d'instance et stratégie d'attribution

Prenons l'exemple d'une demande de Parc d'instances Spot avec la configuration suivante :

- Capacité cible de 30
- Spécification de lancement avec le type d'instance `c3.2xlarge` et une pondération de 8
- Spécification de lancement avec le type d'instance `m3.xlarge` et une pondération de 8
- Spécification de lancement avec le type d'instance `r3.xlarge` et une pondération de 8

Le Parc d'instances Spot lancerait quatre instances (30 divisé par 8, résultat arrondi). Avec la stratégie `lowestPrice`, les quatre instances sont issues du pool d'instances Spot qui fournit le prix par unité le plus bas. Avec la stratégie `diversified`, le Parc d'instances Spot lance une instance dans chacun des trois groupes, et lance la quatrième instance dans l'un des groupes ayant le prix par unité le plus bas.

Démonstration : Utilisation d'un Parc d'instances Spot avec pondération des instances

Cet exemple utilise une société fictive nommée Example Corp afin d'illustrer le processus de demande d'un Parc d'instances Spot utilisant la pondération d'instance.

Objectif

Example Corp est une entreprise pharmaceutique qui souhaite utiliser la puissance de calcul d'Amazon EC2 pour contrôler les composants chimiques susceptibles d'être utilisés afin de lutter contre le cancer.

Planification

Example Corp commence par examiner les bonnes pratiques en matière d'instances Spot. Ensuite, Example Corp détermine les exigences suivantes pour son Parc d'instances Spot.

Types d'instance

Example Corp a une application qui exige beaucoup de calculs et de mémoire. Pour un fonctionnement optimal, cette application a besoin d'au moins 60 Go de mémoire et de huit UC virtuelles (vCPU). L'entreprise souhaite optimiser ces ressources pour l'application au prix le plus bas possible. Example Corp décide que l'un des types d'instance EC2 suivants est capable de répondre à ses besoins :

Type d'instance	Mémoire (Gio)	vCPU
r3.2xlarge	61	8
r3.4xlarge	122	16
r3.8xlarge	244	32

Capacité cible en unités

Avec la pondération d'instance, la capacité cible peut être égale à un certain nombre d'instances (valeur par défaut) ou à une combinaison de critères, par exemple les noyaux (vCPU), la mémoire (GiO) et le stockage (Go). En considérant que la base de son application (60 Go de RAM et huit vCPU) compte pour 1 unité, Example Corp décide que 20 fois cette quantité suffirait pour répondre à ses besoins. L'entreprise définit donc la capacité cible de sa demande de Parc d'instances Spot sur 20.

Pondérations d'instance

Après avoir déterminé sa capacité cible, Example Corp calcule ses pondérations d'instance. Pour calculer la pondération de chaque type d'instance, l'entreprise détermine les unités de chaque type d'instance nécessaires pour atteindre la capacité cible de la façon suivante :

- r3.2xlarge (61 Go, 8 vCPU) = 1 unité de 20
- r3.4xlarge (122 Go, 16 vCPU) = 2 unités de 20
- r3.8xlarge (244 Go, 32 vCPU) = 4 unités de 20

Par conséquent, Example Corp assigne des pondérations d'instance de 1, 2 et 4 aux configurations de lancement respectives dans sa demande de Parc d'instances Spot.

Prix par heure d'unité

Example Corp utilise le prix à la Demande par heure d'instance comme point de départ de son prix. Elle peut également utiliser les prix Spot récents ou une combinaison des deux. Pour calculer le prix par heure d'unité, elle divise le prix de départ basé sur l'heure d'instance par la pondération. Par exemple :

Type d'instance	Prix à la Demande	Pondération de l'instance	Prix par heure d'unité
r3.2xLarge	0,7 USD	1	0,7 USD
r3.4xLarge	1,4 USD	2	0,7 USD
r3.8xLarge	$2.8	4	0,7 USD

Example Corp peut utiliser un prix global par heure d'unité s'élevant à 0,7 USD et rester concurrentielle pour les trois types d'instance. Elle peut également utiliser un prix global par heure d'unité s'élevant à 0,7 USD et un prix spécifique par heure d'unité de 0,9 USD dans la spécification de lancement du type d'instance `r3.8xlarge`.

Vérification des autorisations

Avant de créer une demande de Parc d'instances Spot, Example Corp vérifie qu'elle dispose d'un rôle IAM avec les autorisations requises. Pour plus d'informations, consultez Autorisations Parc d'instances Spot (p. 411).

Création de la demande

Example Corp crée un fichier, `config.json`, avec la configuration suivante pour sa demande de Parc d'instances Spot :

```
{
  "SpotPrice": "0.70",
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
    {
      "ImageId": "ami-1a2b3c4d",
      "InstanceType": "r3.2xlarge",
      "SubnetId": "subnet-482e4972",
      "WeightedCapacity": 1
    },
    {
      "ImageId": "ami-1a2b3c4d",
      "InstanceType": "r3.4xlarge",
      "SubnetId": "subnet-482e4972",
      "WeightedCapacity": 2
    },
    {
      "ImageId": "ami-1a2b3c4d",
      "InstanceType": "r3.8xlarge",
      "SubnetId": "subnet-482e4972",
      "SpotPrice": "0.90",
      "WeightedCapacity": 4
    }
  ]
}
```

Example Corp crée la demande de Parc d'instances Spot à l'aide de la commande request-spot-fleet.

```
aws ec2 request-spot-fleet --spot-fleet-request-config file://config.json
```

Pour plus d'informations, consultez Parc d'instances Spot requêtes (p. 409).

Exécution

La stratégie d'allocation détermine de quels groupes d'instance Spot sont issues vos Instances Spot.

Avec la stratégie `lowestPrice` (qui est la stratégie par défaut), les instances Spot sont issues du pool d'instances Spot ayant le prix par unité le plus bas au moment de l'exécution. Pour fournir 20 unités de capacité, le Parc d'instances Spot lance 20 instances `r3.2xlarge` (20 divisé par 1), 10 instances `r3.4xlarge` (20 divisé par 2) ou 5 instances `r3.8xlarge` (20 divisé par 4).

Si Example Corp avait utilisé la stratégie `diversified`, les Instances Spot auraient été issues des trois groupes. Le Parc d'instances Spot aurait lancé 6 instances `r3.2xlarge` (soit 6 unités), 3 instances `r3.4xlarge` (soit 6 unités) et 2 instances `r3.8xlarge` (soit 8 unités), pour un total de 20 unités.

Historique de tarification d'instance Spot

Au moment d'effectuer des demandes d'Instances Spot, nous vous recommandons d'utiliser le prix maximum par défaut, qui correspond au prix à la demande. Avant de spécifier un prix maximum, nous vous

recommandons de consulter l'historique des prix Spot. Vous pouvez consulter l'historique des prix Spot pour les 90 derniers jours en filtrant par type d'instance, système d'exploitation et zone de disponibilité.

Les prix instance Spot sont fixés Amazon EC2 et ajustés progressivement en fonction des tendances à long terme de l'offre et de la demande de capacité instance Spot. Pour connaître les prix actuels de instance Spot, veuillez consulter Tarification des instances Spot Amazon EC2.

Pour afficher l'historique des prix Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Si vous utilisez les Instances Spot pour la première fois, une page de bienvenue s'affiche. Sélectionnez Mise en route, faites défiler l'écran vers le bas, puis choisissez Annuler.
4. Sélectionnez Historique de tarification.
5. Choisissez le système d'exploitation (Product (Produit)), le Instance type (Type d'instance), et la Date range (Plage de dates) pour lesquels afficher l'historique des prix. Survolez le graphique avec le pointeur de la souris pour afficher les prix à des moments donnés dans plage de dates sélectionnée.

6. (Facultatif) Pour consulter l'historique des prix ponctuels d'une zone de disponibilité spécifique, vous pouvez filtrer les zones de disponibilité en supprimant les zones de disponibilité du graphique. Pour supprimer une zone de disponibilité du graphique, sélectionnez la zone à supprimer. Vous pouvez également sélectionner un produit, un type d'instance ou une plage de dates différent(e).

Pour afficher l'historique des prix Spot à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour plus d'informations, consultez Accès à Amazon EC2 (p. 3).

- describe-Spot-price-history (AWS CLI)
- Get-EC2SpotPriceHistory (Outils AWS pour Windows PowerShell)

Économies réalisées grâce à l'achat d'Instances Spot

Vous pouvez visualiser les informations relatives à l'utilisation et aux économies réalisées grâce aux Instances Spot pour chaque parc ou pour l'ensemble des Instances Spot en cours d'exécution. Les informations relatives à l'utilisation et aux économies pour chaque parc incluent l'ensemble des instances

lancées et résiliées par le parc. Ces informations peuvent être consultées pour la dernière heure ou pour les trois derniers jours.

La capture d'écran suivante de la page Demandes Spot illustre les informations relatives à l'utilisation d'instances Spot et aux économies associées pour un Parc d'instances Spot.

Les informations suivantes relatives à l'utilisation et aux économies sont disponibles :

- Instances Spot – Nombre d'Instances Spot lancées et résiliées par le Parc d'instances Spot. Le nombre qui apparaît dans le récapitulatif des économies représente l'ensemble de vos Instances Spot en cours d'exécution.
- vCPU-hours (Heures vCPU) : nombre d'heures vCPU utilisées pour l'ensemble des Instances Spot sur la période sélectionnée.
- Mem(GiB)-hours (Heures de mémoire (Gio)) : nombre d'heures Gio utilisées pour l'ensemble des Instances Spot sur la période sélectionnée.
- On-Demand total (Total à la demande) : montant total que vous auriez payé pour la période sélectionnée si vous aviez lancé ces instances en tant qu'Instances à la demande.
- Spot total (Total Spot) : montant total à payer pour la période sélectionnée.
- Économie : pourcentage que vous économisez en ne payant pas le prix à la demande.
- Average cost per vCPU-hour (Coût moyen par heure vCPU) : coût horaire moyen de l'utilisation des vCPU pour l'ensemble des Instances Spot sur la période sélectionnée, calculé comme suit : Coût moyen par heure vCPU = Total Spot / Heures vCPU.
- Average cost per mem(GiB)-hour (Coût moyen par heure de mémoire (Gio)) : coût horaire moyen de l'utilisation des Gio pour l'ensemble des Instances Spot sur la période sélectionnée, calculé comme suit : Coût moyen par heure de mémoire (Gio) = Total Spot / Heures de mémoire (Gio).
- Tableau Détails : différents types d'instances (le nombre d'instances par type d'instance est placé entre parenthèses) compris dans le Parc d'instances Spot. Le récapitulatif des économies comprend l'ensemble de vos Instances Spot en cours d'exécution.

Vous ne pouvez consulter les informations relatives aux économies qu'à l'aide de la console Amazon EC2.

Pour afficher les informations relatives aux économies pour un Parc d'instances Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez une demande de Parc d'instances Spot et choisissez Économie.
4. Par défaut, la page affiche les informations relatives à l'utilisation et aux économies de ces trois derniers jours. Vous pouvez choisir last hour (dernière heure) ou last three days (trois derniers jours). Pour les Parcs d'instances Spot qui ont été lancés il y a moins d'une heure, la page affiche une estimation des économies réalisées sur cette heure.

Pour afficher les informations relatives aux économies pour l'ensemble des Instances Spot en cours d'exécution (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Choisissez Savings Summary (Récapitulatif des économies).

instance Spot requêtes

Pour utiliser des Instances Spot, vous créez une demande d'instance Spot qui inclut le nombre d'instances souhaité, le type d'instance, la zone de disponibilité et le prix maximum que vous êtes prêt à payer par heure d'instance. Si votre prix maximum dépasse le prix Spot actuel, Amazon EC2 satisfait votre demande immédiatement si la capacité est disponible. Dans le cas contraire, Amazon EC2 attend jusqu'à ce que votre demande soit exécutée ou jusqu'à l'annulation de celle-ci.

L'illustration suivante représente le fonctionnement des instances Spot. Notez que le type de demande (unique ou persistant) détermine si la demande est ouverte à nouveau lorsqu'Amazon EC2 interrompt une instance Spot ou que vous arrêtez une instance Spot. Si la demande est persistante, elle est rouverte après que votre instance Spot est interrompue. Si la demande est persistante et que vous arrêtez votre instance Spot, la demande s'ouvre seulement après que votre instance Spot a démarré.

Sommaire

- États des demandes d'instance Spot (p. 391)
- Définition d'une durée pour vos Instances Spot (p. 392)
- Définition d'une location pour vos Instances Spot (p. 393)
- Rôle lié à un service pour les demandes d'instance Spot (p. 393)
- Création d'une demande d'instance Spot (p. 395)
- Recherche d'Instances Spot en cours d'exécution (p. 398)
- Balisage des demandes d'instance Spot (p. 399)

- Annulation d'une demande d'instance Spot (p. 404)
- Arrêt d'une instance Spot (p. 404)
- Démarrage d'une instance Spot (p. 405)
- Résiliation d'une instance Spot (p. 406)
- Exemple de spécifications de lancement d'une demande d'instance Spot (p. 407)

États des demandes d'instance Spot

Une demande d'instance Spot peut avoir l'un des états suivants :

- `open` : la demande est en attente d'exécution.
- `active` – La demande a été satisfaite et une instance Spot lui est associée.
- `failed` : la demande a un ou plusieurs paramètres erronés.
- `closed` : l'instance Spot a été interrompue ou résiliée.
- `disabled` – Vous avez arrêté l'instance Spot.
- `cancelled` : vous avez annulé la demande ou elle est arrivée à expiration.

L'illustration suivante représente les transitions entre les états de la demande. Remarquez que les transitions dépendent du type de demande (exceptionnelle ou persistante).

Une demande d'instance Spot unique reste active jusqu'à ce qu'Amazon EC2 lance l'instance Spot, que la demande arrive à expiration ou que vous annuliez la demande. Si le prix spot dépasse votre prix maximum ou que la capacité n'est pas disponible, votre instance Spot est supprimée et la demande d'instance Spot est close.

Une demande d'instance Spot persistante reste active jusqu'à ce qu'elle arrive à expiration ou que vous l'annuliez, même si la demande est satisfaite. Si le prix Spot dépasse votre prix maximum ou que la capacité n'est pas disponible, votre instance Spot est interrompue. Une fois que votre instance a été interrompue, lorsque votre prix maximum dépasse le prix Spot ou que la capacité redevient disponible, l'instance Spot est démarrée si elle a été arrêtée, ou reprise si elle a été mise en veille prolongée. Vous pouvez arrêter une instance Spot et la redémarrer si la capacité est disponible et que votre prix maximum

dépasse le prix Spot actuel. Si l'instance Spot est résiliée (que l'instance Spot soit à l'état arrêté ou en cours d'exécution), la demande d'instance Spot est rouverte et Amazon EC2 lance une nouvelle instance Spot. Pour plus d'informations, consultez Arrêt d'une instance Spot (p. 404), Démarrage d'une instance Spot (p. 405) et Résiliation d'une instance Spot (p. 406).

Vous pouvez effectuer le suivi du statut de vos demandes d'instance Spot, ainsi que celui des Instances Spot lancées, via le statut. Pour de plus amples informations, veuillez consulter Statut des demandes Spot (p. 446).

Définition d'une durée pour vos Instances Spot

Les Instances Spot avec une durée spécifiée (également connues sous le nom de blocs Spot) sont conçues pour ne pas être interrompues et fonctionneront en continu pendant la durée que vous choisissez. C'est pourquoi elles conviennent parfaitement aux tâches associées à une durée définie, par exemple les traitements par lots, le codage et le rendu, la modélisation et l'analyse, ainsi que l'intégration continue.

Vous pouvez utiliser une durée de 1, 2, 3, 4, 5 ou 6 heures. Le prix que vous payez dépend de la durée spécifiée. Pour prendre connaissance des prix actuels pour une durée d'une heure ou de 6 heures, veuillez consulter Prix des instance Spot. Vous pouvez utiliser ces prix pour estimer le coût sur des durées de 2, 3, 4 et 5 heures. Lorsqu'une demande associée à une durée est satisfaite, le prix de l'instance Spot est fixe et reste en vigueur jusqu'à la résiliation de l'instance. Vous êtes facturé à ce prix pour chaque heure complète ou partielle pendant laquelle l'instance s'exécute. Une heure d'instance partielle est facturée à la seconde près.

Lorsque vous définissez une durée dans votre demande d'instance Spot, la période de chaque instance Spot commence dès que l'instance reçoit son ID. L'instance Spot s'exécute jusqu'à ce que vous la résiliiez ou que la période se termine. À la fin de la durée spécifiée, Amazon EC2 marque l'instance Spot pour résiliation et communique un avis de résiliation d'instance Spot , qui avertit l'instance qu'elle va être résiliée dans un délai de deux minutes. Dans de rares cas, les blocs d'instances Spot peuvent être interrompus en raison des besoins en capacités Amazon EC2. Dans ces cas-là, nous fournirons un avertissement de deux minutes avant de mettre fin à une instance, et vous ne serez pas facturé pour les instances résiliées, même si vous les avez utilisées.

Les nouveaux comptes ou les comptes sans historique de facturation avec AWS ne sont pas éligibles pour Instances Spot avec une durée définie (également appelés blocs ponctuels).

Pour lancer des Instances Spot avec une durée définie (console)

Suivez la procédure Création d'une demande de Parc d'instances Spot (p. 416). Pour lancer Instances Spot avec une durée définie, pour Nous indiquer votre besoin en application en tâche, choisissez Charges de travail avec une durée définie.

Pour lancer des Instances Spot avec une durée définie (AWS CLI)

Pour associer une durée à vos Instances Spot, intégrez l'option `--block-duration-minutes` à la commande request-spot-instances. Par exemple, la commande suivante crée une demande Spot qui lance des Instances Spot qui s'exécutent pendant deux heures.

```
aws ec2 request-spot-instances \
    --instance-count 5 \
    --block-duration-minutes 120 \
    --type "one-time" \
    --launch-specification file://specification.json
```

Pour extraire le coût des Instances Spot avec une durée définie (AWS CLI)

Utilisez la commande describe-spot-instance-requests afin d'extraire le coût fixe de vos Instances Spot associées à une durée spécifique. Ces informations se trouvent dans le champ `actualBlockHourlyPrice`.

Définition d'une location pour vos Instances Spot

Vous pouvez exécuter une instance Spot sur du matériel à un seul locataire. Les Instances Spot dédiées sont physiquement isolées des instances qui appartiennent à d'autres comptes AWS. Pour plus d'informations, consultez Instances dédiées (p. 498) et la page produit Instances dédiées Amazon EC2.

Pour exécuter une instance Spot dédiée, effectuez l'une des actions suivantes :

- Spécifiez une location `dedicated` au moment de créer la demande d'instance Spot. Pour plus d'informations, consultez Création d'une demande d'instance Spot (p. 395).
- Demandez une instance Spot sur un VPC avec une location d'instance `dedicated`. Pour plus d'informations, consultez Création d'un VPC avec une location d'instance dédiée (p. 501). Vous ne pouvez pas demander une instance Spot avec une location `default` si vous la demandez sur un VPC avec une location d'instance `dedicated`.

Les types d'instance suivants prennent en charge les Instances Spot dédiées.

Génération actuelle

- `c4.8xlarge`
- `d2.8xlarge`
- `i3.16xlarge`
- `m4.10xlarge`
- `m4.16xlarge`
- `p2.16xlarge`
- `r4.16xlarge`
- `x1.32xlarge`

Génération précédente

- `c3.8xlarge`
- `cc2.8xlarge`
- `cr1.8xlarge`
- `g2.8xlarge`
- `i2.8xlarge`
- `r3.8xlarge`

Rôle lié à un service pour les demandes d'instance Spot

Amazon EC2 utilise des rôles liés à un service pour les autorisations requises pour appeler d'autres services AWS en votre nom. Un rôle lié à un service est un type unique de rôle IAM directement lié à un service AWS. Les rôles liés à un service offrent une manière sécurisée d'accorder des autorisations aux services AWS, car seul le service lié peut assumer un rôle lié à un service. Pour plus d'informations, consultez Utilisation des rôles liés à un service dans le IAM Guide de l'utilisateur.

Amazon EC2 se sert du rôle lié à un service nommé AWSServiceRoleForEC2Spot pour lancer et gérer Instances Spot en votre nom.

Autorisations octroyées par AWSServiceRoleForEC2Spot

Amazon EC2 utilise AWSServiceRoleForEC2Spot pour réaliser les actions suivantes :

- `ec2:DescribeInstances` – Décrire Instances Spot

- `ec2:StopInstances` – Arrêter Instances Spot
- `ec2:StartInstances` – Démarrer Instances Spot

Création du rôle lié à un service

Dans la plupart des cas, vous n'avez pas besoin de créer manuellement un rôle lié à un service. Amazon EC2 crée le rôle lié à un service AWSServiceRoleForEC2Spot la première fois que vous demandez un instance Spot à l'aide de la console.

Si vous aviez une demande d'instance Spot active avant octobre 2017, moment à partir duquel Amazon EC2 a commencé à prendre en charge ce rôle lié à un service, Amazon EC2 a créé le rôle AWSServiceRoleForEC2Spot dans votre compte AWS. Pour de plus amples informations, veuillez consulter Un nouveau rôle est apparu dans mon compte dans le IAM Guide de l'utilisateur.

Assurez-vous que ce rôle existe avant d'utiliser l'AWS CLI ou une API pour créer un instance Spot. Pour créer le rôle, utilisez la console IAM comme suit.

Pour créer manuellement le rôle lié à un service AWSServiceRoleForEC2Spot

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le volet de navigation, choisissez Rôles.
3. Sélectionnez Créer un rôle.
4. Sur la page Select type of trusted entity (Sélectionner le type d'entité de confiance), choisissez EC2, EC2 - Spot Instances (EC2 - Instances Spot), Suivant : Autorisations.
5. Sur la page suivante, choisissez Suivant : Vérification.
6. Sur la page Vérification, choisissez Create Role (Créer un rôle).

Si vous n'avez plus besoin d'utiliser des Instances Spot, nous vous recommandons de supprimer le rôle AWSServiceRoleForEC2Spot. Une fois ce rôle supprimé de votre compte, Amazon EC2 crée de nouveau le rôle si vous effectuez une demande d'Instances Spot.

Octroi d'un accès aux CMK en vue de leur utilisation avec les AMI chiffrées et les instantanés EBS

Si vous spécifiez une AMI chiffrée (p. 168) ou un instantané Amazon EBS chiffré (p. 1194) pour Instances Spot et que vous utilisez la clé principale client gérée (CMK) pour le chiffrement, vous devez autoriser le rôle AWSServiceRoleForEC2Spot à utiliser la CMK afin que Amazon EC2 puisse lancer Instances Spot en votre nom. Pour cela, vous devez ajouter une autorisation à la CMK, comme indiqué dans la procédure suivante.

Lorsque vous définissez les autorisations, les octrois constituent une alternative aux stratégies de clé. Pour de plus amples informations, veuillez consulter Utilisation des autorisations et Utilisation des stratégies de clé dans AWS KMS dans le AWS Key Management Service Developer Guide.

Pour autoriser le rôle AWSServiceRoleForEC2Spot à utiliser la CMK

- Utilisez la commande create-grant pour ajouter un octroi à la CMK et spécifier le mandataire (le rôle lié à un service AWSServiceRoleForEC2Spot) qui reçoit l'autorisation d'effectuer les opérations autorisées par l'octroi. La CMK est spécifiée par le paramètre `key-id` et l'ARN de la CMK. Le mandataire est spécifié par le paramètre `grantee-principal` et l'ARN du rôle lié à un service AWSServiceRoleForEC2Spot.

```
aws kms create-grant \
    --region us-east-1 \
    --key-id arn:aws:kms:us-
east-1:444455556666:key/1234abcd-12ab-34cd-56ef-1234567890ab \
    --grantee-principal arn:aws:iam::111122223333:role/AWSServiceRoleForEC2Spot \
```

```
   --operations "Decrypt" "Encrypt" "GenerateDataKey"
"GenerateDataKeyWithoutPlaintext" "CreateGrant" "DescribeKey" "ReEncryptFrom"
"ReEncryptTo"
```

Création d'une demande d'instance Spot

La procédure de demande d'un instance Spot est similaire à la procédure de lancement d'un instance à la demande. Vous pouvez demander un instance Spot de l'une des façons suivantes :

- Pour demander un instance Spot à l'aide de la console, utilisez l'assistant de lancement de l'instance. Pour de plus amples informations, veuillez consulter Pour créer une demande d'instance Spot (console) (p. 395).
- Pour demander un instance Spot à l'aide de l'interface de ligne de commande, utilisez la commande request-spot-instances ou la commande run-instances. Pour plus d'informations, consultez Pour créer une demande instance Spot à l'aide de request-spot-instances (AWS CLI) et Pour créer une demande instance Spot à l'aide de run-instances (AWS CLI).
- Pour demander un instance Spot avec une durée définie à l'aide de la console, suivez la procédure Création d'une demande de Parc d'instances Spot (p. 416). Pour Nous indiquer votre besoin en application ou tâche, choisissez Charges de travail avec une durée définie. Pour de plus amples informations, veuillez consulter Définition d'une durée pour vos Instances Spot (p. 392).
- Pour demander une instance Spot avec une durée définie à l'aide de l'interface de ligne de commande, utilisez la commande request-spot-instances et spécifiez le paramètre `--block-duration-minutes`. Pour de plus amples informations, veuillez consulter Définition d'une durée pour vos Instances Spot (p. 392).

Après avoir soumis votre demande instance Spot, vous ne pouvez pas modifier les paramètres de la demande. Cela signifie que vous ne pouvez pas modifier le prix maximum que vous êtes prêt à payer.

Si vous demandez plusieurs Instances Spot à la fois, Amazon EC2 crée des demandes d'instance Spot distinctes pour vous permettre de suivre le statut de chaque demande séparément. Pour de plus amples informations sur le suivi des demandes d'instance Spot, veuillez consulter Statut des demandes Spot (p. 446).

Pour lancer un parc comprenant Instances Spot et Instances à la demande, veuillez consulter Création d'une demande de Parc d'instances Spot (p. 416).

> **Note**
>
> Vous ne pouvez pas lancer un instance Spot et un instance à la demande dans le même appel à l'aide de l'assistant de lancement d'instance ou de la commande run-instances.

Prérequis

Avant de commencer, déterminez votre prix maximum, le nombre d'Instances Spot souhaité et le type d'instance à utiliser. Pour passer en revue les tendances de prix Spot, consultez Historique de tarification d'instance Spot (p. 387).

Pour créer une demande d'instance Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, en haut de l'écran, sélectionnez une région.
3. Sur le tableau de bord de la console Amazon EC2, sélectionnez Lancer une instance.
4. Sur la page Sélection d'une Amazon Machine Image (AMI), choisissez une AMI de la façon suivante : Pour de plus amples informations, veuillez consulter Étape 1 : Sélection d'une Amazon Machine Image (AMI) (p. 531).

5. Sur la page Choisir un type d'instance, sélectionnez la configuration matérielle et la taille de l'instance à lancer, puis choisissez Suivant : Configurer les détails de l'instance. Pour de plus amples informations, veuillez consulter Étape 2 : Choisir un type d'instance (p. 532).
6. Sur la page Configurer les détails de l'instance, configurez la demande instance Spot comme suit :
   - Nombre d'instances : entrez le nombre d'instances à lancer.

     Note

     Amazon EC2 crée une demande distincte pour chaque instance Spot.
   - (Facultatif) Pour vous assurer de conserver le nombre correct d'instances pour gérer la demande de votre application, vous pouvez choisir Lancer dans le groupe Auto Scaling pour créer une configuration de lancement et un groupe Auto Scaling. Auto Scaling met à l'échelle le nombre d'instances du groupe en fonction de vos spécifications. Pour plus d'informations, consultez l'Amazon EC2 Auto Scaling Guide de l'utilisateur.
   - Option d'achat : sélectionnez Demander des instances Spot pour lancer une instance Spot. Lorsque vous choisissez cette option, les champs suivants s'affichent.
   - Prix actuel : le prix Spot actuel dans chaque zone de disponibilité s'affiche pour le type d'instance sélectionné.
   - (Facultatif) Prix maximum : vous pouvez laisser le champ vide ou spécifier le montant maximum que vous êtes prêt à payer.
     - Si vous laissez le champ vide, le prix maximal est par défaut le prix à la demande actuel. Vos lancements instance Spot au prix Spot actuel, ne dépassant pas le prix à la demande.
     - Si vous spécifiez un prix maximum supérieur au prix Spot actuel, vos lancements instance Spot sont facturés au prix Spot actuel.
     - Si vous spécifiez un prix maximum inférieur au prix Sot, votre instance Spot n'est pas lancé.
   - Demande persistante : choisissez Demande persistante pour soumettre à nouveau la demande instance Spot si votre instance Spot est interrompu.
   - Comportement d'interruption : par défaut, le service Spot arrête un instance Spot lorsqu'il est interrompu. Si vous choisissez Demande persistante, vous pouvez spécifier que le service Spot arrête ou hiberne votre instance Spot lorsque vous l'interrompez. Pour de plus amples informations, veuillez consulter Comportements d'interruption (p. 453).
   - (Facultatif) Demande valide pour : choisissez Modifier pour spécifier l'expiration de la demande instance Spot.

   Pour de plus amples informations sur la configuration de instance Spot, veuillez consulter Étape 3 : Configurer les détails de l'instance (p. 532).
7. L'AMI sélectionnée inclut un ou plusieurs volumes de stockage, notamment le volume du périphérique racine. Sur la page Ajouter le stockage, vous pouvez spécifier des volumes supplémentaires à attacher à l'instance en choisissant Ajouter un nouveau volume. Pour de plus amples informations, veuillez consulter Étape 4 : Ajouter du stockage (p. 535).
8. Sur la page Ajouter des balises, spécifiez les balises (p. 1309) en fournissant les combinaisons clé et valeur. Pour de plus amples informations, veuillez consulter Étape 5 : Ajouter des balises (p. 536).
9. Sur la page Configurer le groupe de sécurité, utilisez un groupe de sécurité afin de définir les règles de pare-feu de votre instance. Ces règles déterminent le trafic réseau entrant acheminé vers votre instance. Le reste du trafic est ignoré. Pour plus d'informations sur les groupes de sécurité, consultez Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062). Sélectionnez ou créez un groupe de sécurité de la façon suivante, puis choisissez Vérifier et lancer. Pour de plus amples informations, veuillez consulter Étape 6 : Configurer un groupe de sécurité (p. 536).
10. Sur la page Examiner le lancement de l'instance, vérifiez les détails de votre instance, puis effectuez les modifications nécessaires en sélectionnant le lien Modifier approprié. Une fois que vous êtes prêt, choisissez Lancer. Pour de plus amples informations, veuillez consulter Étape 7 : Vérifier le lancement de l'instance et sélectionner une paire de clés (p. 536).

11. Dans la boîte de dialogue Select an existing key pair or create a new key pair (Sélectionner une paire de clés existante ou créer une nouvelle paire de clés), vous pouvez choisir une paire de clés existante ou en créer une nouvelle. Par exemple, sélectionnez Choisir une paire de clés existante, puis choisissez la paire de clés que vous avez créée lors de la configuration. Pour de plus amples informations, veuillez consulter Paires de clés Amazon EC2 et instances Linux (p. 1047).

    Important

    Si vous sélectionnez l'option Proceed without key pair (Continuer sans paire de clé), vous ne pourrez pas vous connecter à l'instance à moins de choisir une AMI configurée de façon à autoriser les utilisateurs à se connecter d'une autre façon.

12. Pour lancer votre instance, activez la case à cocher de confirmation, puis choisissez Lancer des instances.

    Si l'instance ne peut pas être lancée ou que l'état passe immédiatement à `terminated` au lieu de `running`, consultez Résolution des problèmes de lancement d'instance (p. 1325).

Pour créer une demande instance Spot à l'aide de request-spot-instances (AWS CLI)

Utilisez la commande request-Spot-instances pour créer une demande unique.

```
aws ec2 request-spot-instances \
    --instance-count 5 \
    --type "one-time" \
    --launch-specification file://specification.json
```

Utilisez la commande request-Spot-instances pour créer une demande persistante.

```
aws ec2 request-spot-instances \
    --instance-count 5 \
    --type "persistent" \
    --launch-specification file://specification.json
```

Pour accéder à des exemples de fichiers de spécification à utiliser avec ces commandes, consultez Exemple de spécifications de lancement d'une demande d'instance Spot (p. 407). Si vous téléchargez un fichier de spécification de lancement à partir de la console, vous devez plutôt utiliser la commande request-spot-fleet (la console spécifie une demande Spot utilisant un Parc d'instances Spot).

Pour créer une demande instance Spot à l'aide de run-instances (AWS CLI)

Utilisez la commande run-instances et spécifiez les options instance Spot dans le paramètre `--instance-market-options`.

```
aws ec2 run-instances \
    --image-id ami-0abcdef1234567890 \
    --instance-type t2.micro \
    --count 5 \
    --subnet-id subnet-08fc749671b2d077c \
    --key-name MyKeyPair \
    --security-group-ids sg-0b0384b66d7d692f9 \
    --instance-market-options file://spot-options.json
```

Voici la structure de données à spécifier dans le fichier JSON pour `--instance-market-options`. Vous pouvez également spécifier `BlockDurationMinutes`, `ValidUntil` et `InstanceInterruptionBehavior`. Si vous ne spécifiez pas de champ dans la structure de données, la valeur par défaut est utilisée. Cet exemple crée une demande `one-time` et spécifie `0.02` comme prix maximum que vous êtes prêt à payer pour le instance Spot.

```
{
```

```
  "MarketType": "spot",
  "SpotOptions": {
    "MaxPrice": "0.02",
    "SpotInstanceType": "one-time"
  }
}
```

Recherche d'Instances Spot en cours d'exécution

Amazon EC2 lance une instance Spot lorsque le prix maximum dépasse le prix spot et que la capacité est disponible. Une instance Spot s'exécute jusqu'à ce qu'elle soit interrompue ou que vous la résiliiez. Si votre prix maximum est parfaitement identique au prix Spot, il y a des chances pour que votre instance Spot continue de s'exécuter, selon la demande.

Pour rechercher des Instances Spot en cours d'exécution (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot. Vous pouvez voir à la fois les demandes d'instance Spot et les demandes de Parc d'instances Spot. Si une demande d'instance Spot a été satisfaite, Capacité a pour valeur l'ID de l'instance Spot. Pour un Parc d'instances Spot, le champ Capacité indique la part de la capacité demandée qui a été satisfaite. Pour afficher les ID des instances d'un Parc d'instances Spot, choisissez la flèche de développement, ou sélectionnez le parc et choisissez Instances.

   Note

   Pour les demandes instance Spot créées par un Parc d'instances Spot, les demandes ne sont pas étiquetées instantanément avec la balise système qui indique la Parc d'instances Spot à laquelle elles appartiennent, et pendant un certain temps peuvent apparaître séparément de la demande Parc d'instances Spot.

   Sinon, dans le panneau de navigation, sélectionnez Instances. Dans le coin supérieur droit, choisissez l'icône des paramètres (⚙), puis sous Colonnes attributaires, sélectionnez Cycle de vie de l'instance. Pour chaque instance, la valeur de Cycle de vie est `normal`, `spot` ou `scheduled`.

Pour rechercher des Instances Spot en cours d'exécution (AWS CLI)

Pour énumérer les Instances Spot, utilisez la commande describe-spot-instance-requests avec l'option `--query`.

```
aws ec2 describe-spot-instance-requests \
    --query "SpotInstanceRequests[*].{ID:InstanceId}"
```

Voici un exemple de sortie :

```
[
    {
        "ID": "i-1234567890abcdef0"
    },
    {
        "ID": "i-0598c7d356eba48d7"
    }
]
```

Vous pouvez aussi énumérer vos Instances Spot en utilisant la commande describe-instances avec l'option `--filters`.

```
aws ec2 describe-instances \
```

```
    --filters "Name=instance-lifecycle,Values=spot"
```

Pour décrire une seule instance instance Spot, utilisez la commande describe-spot-instance-requests avec l'option `--spot-instance-request-ids`.

```
aws ec2 describe-spot-instance-requests \
    --spot-instance-request-ids sir-08b93456
```

Balisage des demandes d'instance Spot

Pour vous aider à classer et à gérer vos demandes de instance Spot, vous pouvez les baliser avec des métadonnées personnalisées. Vous pouvez affecter une balise à une demande de instance Spot lorsque vous la créez, ou après. Vous pouvez attribuer des balises à l'aide de la console Amazon EC2 ou d'un outil de ligne de commande.

Lorsque vous balisez une demande instance Spot, les instances et les volumes lancés par le instance Spot ne sont pas automatiquement balisés. Vous devez baliser explicitement les instances et les volumes lancés par la demande instance Spot. Vous pouvez affecter une balise à instance Spot et à des volumes pendant ou après le lancement.

Pour plus d'informations sur le fonctionnement des balises, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Table des matières

- Prérequis (p. 399)
- Balisage d'une nouvelle demande instance Spot (p. 401)
- Balisage d'une demande instance Spot existante (p. 402)
- Affichage des balises de demande instance Spot (p. 402)

Prérequis

Octroyez à l'utilisateur IAM l'autorisation de baliser les ressources. Pour de plus amples informations sur les stratégies IAM et les exemples de stratégies, veuillez consulter Exemple : Balisage de vos ressources (p. 1018).

La stratégie IAM que vous créez est déterminée par la méthode que vous utilisez pour créer une demande instance Spot.

- Si vous utilisez l'assistant de lancement d'instance ou `run-instances` pour demander Instances Spot, veuillez consulter Pour accorder à un utilisateur IAM l'autorisation de baliser des ressources lors de l'utilisation de l'assistant d'instance de lancement ou des instances d'exécution.
- Si vous utilisez la console Ponctuelle pour demander Instances Spot avec une durée définie ou si vous utilisez la commande `request-spot-instances` pour demander Instances Spot, veuillez consulter Pour accorder à un utilisateur IAM l'autorisation de baliser des ressources lors de l'utilisation de request-spot-instances.

Pour accorder à un utilisateur IAM l'autorisation de baliser des ressources lors de l'utilisation de l'assistant d'instance de lancement ou des instances d'exécution

Créez une stratégie IAM qui inclut les éléments suivants :

- L'action `ec2:RunInstances`. Cela accorde à l'utilisateur IAM l'autorisation de lancer une instance.
- Pour `Resource`, spécifiez `spot-instances-request`. Cela permet aux utilisateurs de créer des demandes instance Spot, qui demandent Instances Spot.
- L'action `ec2:CreateTags`. Celle-ci accorde à l'utilisateur IAM l'autorisation de créer des balises.

- Pour `Resource`, spécifiez `*`. Cela permet aux utilisateurs de baliser toutes les ressources créées lors du lancement de l'instance.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowLaunchInstances",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*",
                "arn:aws:ec2:us-east-1:*:spot-instances-request/*"
            ]
        },
        {
            "Sid": "TagSpotInstanceRequests",
            "Effect": "Allow",
            "Action": "ec2:CreateTags",
            "Resource": "*"
        }
    ]
}
```

Note

Lorsque vous utilisez l'action RunInstances pour créer des demandes instance Spot et que vous balisez les demandes instance Spot lors de la création, vous devez savoir comment Amazon EC2 évalue la ressource `spot-instances-request` dans l'instruction RunInstances.
La ressource `spot-instances-request` est évaluée dans la stratégie IAM comme suit :

- Si vous ne balisez pas une demande instance Spot lors de la création, Amazon EC2 n'évalue pas la ressource `spot-instances-request` dans l'instruction RunInstances.
- Si vous balisez une demande instance Spot lors de la création, Amazon EC2 évalue la ressource `spot-instances-request` dans l'instruction RunInstances.

Par conséquent, pour la ressource `spot-instances-request`, les règles suivantes s'appliquent à la stratégie IAM :

- Si vous utilisez RunInstances pour créer une demande instance Spot et que vous n'avez pas l'intention de baliser la demande instance Spot lors de la création, vous n'avez pas besoin d'autoriser explicitement la ressource `spot-instances-request` ; l'appel réussira.
- Si vous utilisez RunInstances pour créer une demande instance Spot et que vous avez l'intention de baliser la demande instance Spot lors de la création, vous devez inclure la ressource `spot-instances-request` dans l'instruction RunInstances allow, sinon l'appel échouera.
- Si vous utilisez RunInstances pour créer une demande instance Spot et que vous avez l'intention de baliser la demande instance Spot lors de la création, vous devez spécifier la ressource `spot-instances-request` ou inclure un caractère générique `*` dans l'instruction CreateTags allow, sinon l'appel échouera.

Par exemple, pour les stratégies IAM, y compris les stratégies qui ne sont pas prises en charge pour les demandes instance Spot, veuillez consulter Utilisation de Instances Spot (p. 1012).

Pour accorder à un utilisateur IAM l'autorisation de baliser des ressources lors de l'utilisation d'instances request-spot-instances

Créez une stratégie IAM qui inclut les éléments suivants :

- L'action `ec2:RequestSpotInstances`. Celle-ci accorde à l'utilisateur IAM l'autorisation de créer une demande de instance Spot.
- L'action `ec2:CreateTags`. Celle-ci accorde à l'utilisateur IAM l'autorisation de créer des balises.
- Pour `Resource`, spécifiez `spot-instances-request`. Cela permet aux utilisateurs de baliser uniquement la demande instance Spot.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "TagSpotInstanceRequest",
            "Effect": "Allow",
            "Action": [
                "ec2:RequestSpotInstances",
                "ec2:CreateTags"
            ],
            "Resource": "arn:aws:ec2:us-east-1:111122223333:spot-instances-request/*"
}
```

Balisage d'une nouvelle demande instance Spot

Pour baliser une nouvelle demande instance Spot à l'aide de la console

1. Suivez la procédure Création d'une demande d'instance Spot (p. 395).
2. Pour ajouter une balise, sur la page Ajouter des balises , choisissez Ajouter une balise, puis entrez la clé et la valeur de la balise. Choisissez Ajouter une autre balise pour chaque balise supplémentaire.

   Pour chaque balise, vous pouvez baliser la demande instance Spot, l'Instances Spot et les volumes avec la même balise. Pour baliser les trois, assurez-vous que les instances, les volumes, et les demandes instance Spot sont sélectionnés. Pour n'en baliser qu'une ou deux, assurez-vous que les ressources que vous souhaitez baliser sont sélectionnées et que les autres ressources sont effacées.
3. Remplissez les champs obligatoires pour créer une demande de instance Spot, puis choisissez Lancer. Pour plus d'informations, consultez Création d'une demande d'instance Spot (p. 395).

Pour baliser une nouvelle demande instance Spot à l'aide de l'AWS CLI

Pour baliser une demande instance Spot lors de sa création, configurez la demande instance Spot comme suit :

- Spécifiez les balises de la demande instance Spot à l'aide du paramètre `--tag-specification`.
- Pour `ResourceType`, spécifiez `spot-instances-request`. Si vous indiquez une autre valeur, la demande instance Spot échouera.
- Pour `Tags`, spécifiez la paire clé-valeur. Vous pouvez définir plusieurs paires clé-valeur.

Dans l'exemple suivant, la demande instance Spot est marquée par deux balises : Key=Environment et Value=Production, ainsi que Key=Cost-Center et Value=123.

```
aws ec2 request-spot-instances \
```

```
    --instance-count 5 \
    --type "one-time" \
    --launch-specification file://specification.json \
    --tag-specification 'ResourceType=spot-instances-
request,Tags=[{Key=Environment,Value=Production},{Key=Cost-Center,Value=123}]'
```

Balisage d'une demande instance Spot existante

Pour baliser une demande instance Spot existante à l'aide de la console

Après avoir créé une demande instance Spot, vous pouvez ajouter des balises à la demande instance Spot à l'aide de la console.

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot.
2. Sélectionnez votre demande instance Spot.
3. Choisissez l'onglet Tags (Balises), puis Create Tag (Créer une balise).

Pour baliser une instance Spot existante à l'aide de la console

Une fois que votre demande instance Spot a lancé votre instance Spot, vous pouvez ajouter des balises à l'instance à l'aide de la console. Pour de plus amples informations, veuillez consulter Ajout et suppression de balises sur une ressource individuelle (p. 1316).

Pour baliser une demande instance Spot existante ou instance Spot à l'aide de AWS CLI

Utilisez la commande create-tags pour baliser les ressources existantes. Dans l'exemple suivant, la demande instance Spot existante et le instance Spot sont balisés avec Key=purpose et Value=test.

```
aws ec2 create-tags \
    --resources sir-08b93456 i-1234567890abcdef0 \
    --tags Key=purpose,Value=test
```

Affichage des balises de demande instance Spot

Pour afficher les balises d'une demande instance Spot à l'aide de la console

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot.
2. Sélectionnez votre demande instance Spot et choisissez l'onglet Balises.

Pour décrire les balises de la demande instance Spot

Utilisez la commande describe-tags pour afficher les balises de la ressource spécifiée. Dans l'exemple suivant, vous décrivez les balises de la demande spécifiée.

```
aws ec2 describe-tags \
    --filters "Name=resource-id,Values=sir-11112222-3333-4444-5555-66666EXAMPLE"
```

```
{
    "Tags": [
        {
            "Key": "Environment",
            "ResourceId": "sir-11112222-3333-4444-5555-66666EXAMPLE",
            "ResourceType": "spot-instances-request",
            "Value": "Production"
        },
        {
            "Key": "Another key",
```

```
            "ResourceId": "sir-11112222-3333-4444-5555-66666EXAMPLE",
            "ResourceType": "spot-instances-request",
            "Value": "Another value"
        }
    ]
}
```

Vous pouvez également afficher les balises d'une demande instance Spot en décrivant la demande instance Spot.

Utilisez la commande describe-spot-instance-requests pour afficher la configuration de la demande instance Spot spécifiée, qui inclut toutes les balises définies pour la demande.

```
aws ec2 describe-spot-instance-requests \
    --spot-instance-request-ids sir-11112222-3333-4444-5555-66666EXAMPLE
```

```
{
    "SpotInstanceRequests": [
        {
            "CreateTime": "2020-06-24T14:22:11+00:00",
            "InstanceId": "i-1234567890EXAMPLE",
            "LaunchSpecification": {
                "SecurityGroups": [
                    {
                        "GroupName": "launch-wizard-6",
                        "GroupId": "sg-1234567890EXAMPLE"
                    }
                ],
                "BlockDeviceMappings": [
                    {
                        "DeviceName": "/dev/xvda",
                        "Ebs": {
                            "DeleteOnTermination": true,
                            "VolumeSize": 8,
                            "VolumeType": "gp2"
                        }
                    }
                ],
                "ImageId": "ami-1234567890EXAMPLE",
                "InstanceType": "t2.micro",
                "KeyName": "my-key-pair",
                "NetworkInterfaces": [
                    {
                        "DeleteOnTermination": true,
                        "DeviceIndex": 0,
                        "SubnetId": "subnet-11122233"
                    }
                ],
                "Placement": {
                    "AvailabilityZone": "eu-west-1c",
                    "Tenancy": "default"
                },
                "Monitoring": {
                    "Enabled": false
                }
            },
            "LaunchedAvailabilityZone": "eu-west-1c",
            "ProductDescription": "Linux/UNIX",
            "SpotInstanceRequestId": "sir-1234567890EXAMPLE",
            "SpotPrice": "0.012600",
            "State": "active",
            "Status": {
                "Code": "fulfilled",
```

```
                "Message": "Your spot request is fulfilled.",
                "UpdateTime": "2020-06-25T18:30:21+00:00"
            },
            "Tags": [
                {
                    "Key": "Environment",
                    "Value": "Production"
                },
                {
                    "Key": "Another key",
                    "Value": "Another value"
                }
            ],
            "Type": "one-time",
            "InstanceInterruptionBehavior": "terminate"
        }
    ]
}
```

Annulation d'une demande d'instance Spot

Si vous n'avez plus besoin de votre demande d'instance Spot, vous pouvez la supprimer. Vous pouvez seulement annuler des demandes d'instance Spot à l'état `open`, `active` ou `disabled`.

- Votre demande d'instance Spot est `open` lorsqu'elle n'a pas encore été satisfaite et qu'aucune instance n'a été lancée.
- Votre demande d'instance Spot est `active` lorsqu'elle a été satisfaite et que des Instances Spot ont été lancées en conséquence.
- Votre demande d'instance Spot est `disabled` lorsque vous arrêtez votre instance Spot.

Si votre demande d'instance Spot est à l'état `active` et qu'elle est associée à une instance Spot en cours d'exécution, l'annulation de la demande ne résilie pas l'instance. Pour plus d'informations sur l'arrêt d'une instance Spot, consultez Résiliation d'une instance Spot (p. 406).

Pour annuler une demande d'instance Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Demandes Spot et sélectionnez la demande d'instance Spot.
3. Choisissez Actions, Annuler la demande.
4. (Facultatif) Si vous n'avez plus besoin d'utiliser les Instances Spot associées, vous pouvez les résilier. Dans la boîte de dialogue Annuler la demande Spot sélectionnez Terminer les instances, puis choisissez Confirmer.

Pour annuler une demande d'instance Spot (AWS CLI)

- Utilisez la commande cancel-spot-instance-requests pour annuler la demande d'instance Spot spécifiée.

```
aws ec2 cancel-spot-instance-requests --spot-instance-request-ids sir-08b93456
```

Arrêt d'une instance Spot

Si vous n'avez pas besoin de vos Instances Spot actuellement, mais que vous souhaitez les redémarrer plus tard sans perdre les données persistantes du volume Amazon EBS, vous pouvez les arrêter. Les

étapes d'arrêt d'une instance Spot sont similaires aux étapes d'arrêt d'une instance à la demande. Vous ne pouvez arrêter une instance Spot si l'instance Spot a été lancé à partir d'une demande d'instance Spot `persistent`.

> **Note**
>
> Pendant qu'une instance Spot est arrêtée, vous pouvez modifier certains de ses attributs, mais pas le type d'instance.
> Nous ne vous facturons ni l'utilisation d'une instance Spot arrêtée, ni les frais de transfert de données. Par contre, nous facturons le stockage des volumes Amazon EBS.

Limites

- Vous ne pouvez pas arrêter une instance Spot si elle fait partie d'un parc ou d'un groupe de lancement, d'un groupe de zone de disponibilité ou d'un bloc Spot.

Nouvelle console

Pour arrêter une instance Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis sélectionnez l'instance Spot.
3. Choisissez État de l'instance, Arrêter l'instance.
4. Lorsque vous êtes invité à confirmer l'opération, choisissez Arrêter.

Ancienne console

Pour arrêter une instance Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis sélectionnez l'instance Spot.
3. Choisissez Actions, Instance State, Stop.

AWS CLI

Pour arrêter une instance Spot (AWS CLI)

- Utilisez la commande stop-instances pour arrêter manuellement une ou plusieurs Instances Spot.

```
aws ec2 stop-instances --instance-ids i-1234567890abcdef0
```

Démarrage d'une instance Spot

Vous pouvez démarrer une instance Spot que vous avez précédemment arrêtée. Les étapes de démarrage d'une instance Spot sont similaires à celles du démarrage d'une instance à la demande.

Conditions préalables requises

Vous pouvez démarrer une instance Spot uniquement si :

- Vous avez arrêté manuellement l'instance Spot.
- L'instance Spot est une instance basée sur EBS.
- La capacité de l'instance Spot est disponible.
- Le prix Spot est inférieur à votre prix maximum.

Limites

- Vous ne pouvez pas démarrer une instance Spot qui fait partie d'un parc ou d'un groupe de lancement, d'un groupe de zone de disponibilité ou d'un bloc Spot.

Nouvelle console

Pour démarrer une instance Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis sélectionnez l'instance Spot.
3. Choisissez État de l'instance, Démarrer l'instance.

Ancienne console

Pour démarrer une instance Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis sélectionnez l'instance Spot.
3. Choisissez Actions, État de l'instance, Début.

AWS CLI

Pour démarrer une instance Spot (AWS CLI)

- Utilisez la commande start-instances pour démarrer manuellement une ou plusieurs Instances Spot.

```
aws ec2 start-instances --instance-ids i-1234567890abcdef0
```

Résiliation d'une instance Spot

Si votre demande d'instance Spot est `active` et qu'elle est associée à une instance Spot en cours d'exécution, ou si votre demande d'instance Spot est `disabled` et qu'elle est associée à une instance Spot arrêtée, l'annulation de la demande ne résilie pas l'instance ; vous devez résilier manuellement l'exécution de l'instance Spot.

Si vous résiliez une instance Spot en cours d'exécution ou arrêtée qui a été lancée par une demande d'instance Spot persistante, la demande d'instance Spot revient à l'état `open` pour qu'une nouvelle instance Spot puisse être lancée. Pour annuler une demande Spot persistante et mettre hors service ses Instances Spot, vous devez commencer par annuler la demande, puis résilier les Instances Spot qui lui sont associées. Si vous ne procédez pas de cette façon, la demande d'instance Spot persistante peut lancer une nouvelle instance. Pour plus d'informations sur l'annulation d'une demande d'instance Spot, veuillez consulter Annulation d'une demande d'instance Spot (p. 404).

Nouvelle console

Pour résilier manuellement une instance Spot à l'aide de la console

1. Avant de résilier une instance, vérifiez que vous ne perdrez aucune donnée en vous assurant que vos volumes Amazon EBS ne seront pas supprimés lors de la résiliation et que vous avez copié les données dont vous avez besoin des volumes du stockage d'instance vers un stockage persistant, par exemple Amazon EBS ou Amazon S3.

2. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
3. Dans le panneau de navigation, choisissez Instances.
4. Pour confirmer que l'instance est une instance Spot, vérifiez que spot s'affiche dans la colonne Instance lifecycle (Cycle de vie de l'instance).
5. Sélectionnez l'instance, puis choisissez Actions, État de l'instance, Résilier l'instance.
6. Choisissez Résilier lorsque vous êtes invité à confirmer.

Ancienne console

Pour résilier manuellement une instance Spot à l'aide de la console

1. Avant de résilier une instance, vérifiez que vous ne perdrez aucune donnée en vous assurant que vos volumes Amazon EBS ne seront pas supprimés lors de la résiliation et que vous avez copié les données dont vous avez besoin des volumes du stockage d'instance vers un stockage persistant, par exemple Amazon EBS ou Amazon S3.
2. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
3. Dans le panneau de navigation, choisissez Instances.
4. Pour confirmer que l'instance est une instance Spot, vérifiez que spot s'affiche dans la colonne Lifecycle (Cycle de vie).
5. Sélectionnez l'instance et choisissez Actions, État de l'instance, Résilier.
6. Sélectionnez Oui, résilier lorsque vous êtes invité à confirmer l'opération.

AWS CLI

Pour résilier manuellement une instance Spot à l'aide de l'AWS CLI

- Utilisez la commande terminate-instances pour résilier manuellement des Instances Spot.

```
aws ec2 terminate-instances --instance-ids i-1234567890abcdef0 i-0598c7d356eba48d7
```

Exemple de spécifications de lancement d'une demande d'instance Spot

Les exemples suivants montrent les configurations de lancement que vous pouvez utiliser avec la commande request-spot-instances afin de créer une demande d'instance Spot. Pour plus d'informations, consultez Création d'une demande d'instance Spot (p. 395).

1. Lancement d'Instances Spot (p. 407)
2. Lancement d'Instances Spot dans la zone de disponibilité spécifiée (p. 408)
3. Lancement d'Instances Spot dans le sous-réseau spécifié (p. 408)
4. Lancement d'une instance Spot dédiée (p. 409)

Exemple 1 : Lancement d'Instances Spot

L'exemple suivant ne comprend aucune zone de disponibilité ni aucun sous-réseau. Amazon EC2 sélectionne une zone de disponibilité pour vous. Amazon EC2 lance les instances sur le sous-réseau par défaut de la zone de disponibilité sélectionnée.

```
{
  "ImageId": "ami-1a2b3c4d",
  "KeyName": "my-key-pair",
  "SecurityGroupIds": [ "sg-1a2b3c4d" ],
```

```
  "InstanceType": "m3.medium",
  "IamInstanceProfile": {
      "Arn": "arn:aws:iam::123456789012:instance-profile/my-iam-role"
  }
}
```

Exemple 2 : Lancement d'Instances Spot dans la zone de disponibilité spécifiée

L'exemple suivant comprend une zone de disponibilité. Amazon EC2 lance les instances sur le sous-réseau par défaut de la zone de disponibilité spécifiée.

```
{
  "ImageId": "ami-1a2b3c4d",
  "KeyName": "my-key-pair",
  "SecurityGroupIds": [ "sg-1a2b3c4d" ],
  "InstanceType": "m3.medium",
  "Placement": {
    "AvailabilityZone": "us-west-2a"
  },
  "IamInstanceProfile": {
      "Arn": "arn:aws:iam::123456789012:instance-profile/my-iam-role"
  }
}
```

Exemple 3 : Lancement d'Instances Spot dans le sous-réseau spécifié

L'exemple suivant comprend un sous-réseau. Amazon EC2 lance les instances sur le sous-réseau spécifié. S'il s'agit d'un VPC personnalisé, l'instance ne reçoit pas d'adresse IPv4 publique par défaut.

```
{
  "ImageId": "ami-1a2b3c4d",
  "SecurityGroupIds": [ "sg-1a2b3c4d" ],
  "InstanceType": "m3.medium",
  "SubnetId": "subnet-1a2b3c4d",
  "IamInstanceProfile": {
      "Arn": "arn:aws:iam::123456789012:instance-profile/my-iam-role"
  }
}
```

Pour attribuer une adresse IPv4 publique à une instance sur un VPC personnalisé, spécifiez le champ `AssociatePublicIpAddress` tel qu'illustré à l'exemple suivant. Lorsque vous spécifiez une interface réseau, vous devez inclure l'ID de sous-réseau et l'ID du groupe de sécurité à l'aide de l'interface réseau au lieu d'utiliser les champs `SubnetId` et `SecurityGroupIds` illustrés à l'exemple 3.

```
{
  "ImageId": "ami-1a2b3c4d",
  "KeyName": "my-key-pair",
  "InstanceType": "m3.medium",
  "NetworkInterfaces": [
    {
      "DeviceIndex": 0,
      "SubnetId": "subnet-1a2b3c4d",
      "Groups": [ "sg-1a2b3c4d" ],
      "AssociatePublicIpAddress": true
    }
  ],
  "IamInstanceProfile": {
      "Arn": "arn:aws:iam::123456789012:instance-profile/my-iam-role"
  }
}
```

Exemple 4 : Lancement d'une instance Spot dédiée

L'exemple suivant demande une instance Spot avec une location `dedicated`. Une instance Spot dédiée doit être lancée sur un VPC.

```
{
  "ImageId": "ami-1a2b3c4d",
  "KeyName": "my-key-pair",
  "SecurityGroupIds": [ "sg-1a2b3c4d" ],
  "InstanceType": "c3.8xlarge",
  "SubnetId": "subnet-1a2b3c4d",
  "Placement": {
    "Tenancy": "dedicated"
  }
}
```

Parc d'instances Spot requêtes

Pour utiliser un Parc d'instances Spot, vous créez une demande de Parc d'instances Spot comprenant la capacité cible, une portion à la demande facultative, une ou plusieurs spécifications de lancement pour les instances et le prix maximum que vous êtes prêt à payer. Amazon EC2 tente de maintenir la capacité cible de votre Parc d'instances Spot au fur et à mesure de l'évolution des prix spot. Pour plus d'informations, consultez Fonctionnement de Parc d'instances Spot (p. 380).

Il existe deux types de demande de Parc d'instances Spot : `request` et `maintain`. Vous pouvez créer un Parc d'instances Spot pour soumettre une demande unique pour la capacité souhaitée ou pour lui demander de maintenir une capacité cible dans le temps. Les deux types de demande bénéficient de la stratégie d'allocation du Parc d'instances Spot.

Lorsque vous effectuez une demande unique, le Parc d'instances Spot passe les demandes requises, mais n'essaie pas de réapprovisionner les Instances Spot si la capacité est réduite. Si la capacité n'est pas disponible, le Parc d'instances Spot n'envoie pas de demandes à d'autres groupes d'instances Spot.

Pour maintenir une capacité cible, le Parc d'instances Spot passe les demandes pour satisfaire cette capacité cible et réapprovisionne automatiquement les instances interrompues.

Il n'est pas possible de modifier la capacité cible d'une demande unique une fois qu'elle a été soumise. Pour modifier la capacité cible, annulez la demande et soumettez-en une nouvelle.

Une demande de Parc d'instances Spot reste active jusqu'à ce qu'elle arrive à expiration ou que vous l'annuliez. Lorsque vous annulez une demande de Parc d'instances Spot, vous pouvez préciser si l'annulation de votre demande de Parc d'instances Spot résilie les Instances Spot de votre Parc d'instances Spot.

Chaque spécification de lancement comprend les informations dont Amazon EC2 a besoin pour lancer une instance : par exemple, une AMI, un type d'instance, un sous-réseau ou une zone de disponibilité, et un ou plusieurs groupes de sécurité.

Sommaire

- États des demandes Parc d'instances Spot (p. 410)
- Vérifications de l'état du Parc d'instances Spot (p. 410)
- Planification d'une demande de Parc d'instances Spot (p. 411)
- Autorisations Parc d'instances Spot (p. 411)
- Création d'une demande de Parc d'instances Spot (p. 416)
- Balisage d'un Parc d'instances Spot (p. 420)
- Surveillance de votre Parc d'instances Spot (p. 426)
- Modification d'une demande de Parc d'instances Spot (p. 427)

- Annulation d'une demande de Parc d'instances Spot (p. 428)
- Exemples de configuration d'un Parc d'instances Spot (p. 429)

États des demandes Parc d'instances Spot

Une demande d'Parc d'instances Spot peut avoir l'un des états suivants :

- `submitted` – La demande de Parc d'instances Spot est en cours d'évaluation et Amazon EC2 se prépare à lancer le nombre cible d'instances.
- `active` – La demande de Parc d'instances Spot a été validée et Amazon EC2 tente de conserver le nombre cible d'Instances Spot en cours d'exécution. La demande conserve cet état jusqu'à ce qu'elle soit modifiée ou annulée.
- `modifying` – La demande de Parc d'instances Spot est en cours de modification. La demande conserve cet état jusqu'à ce que la modification soit totalement traitée ou que le Parc d'instances Spot soit annulé. Il n'est pas possible de modifier une demande (`request`) unique, et cet état ne s'applique pas à ce type de demande d'instance Spot.
- `cancelled_running` – Le Parc d'instances Spot est annulé et ne lance pas d'Instances Spot supplémentaires. Ses Instances Spot existantes continuent de s'exécuter jusqu'à ce qu'elles soient interrompues ou mises hors service. La demande conserve cet état jusqu'à ce que toutes les instances soient interrompues ou mises hors service.
- `cancelled_terminating` – Le Parc d'instances Spot est annulé et ses Instances Spot sont résiliées. La demande conserve cet état jusqu'à ce que toutes les instances soient mises hors service.
- `cancelled` – Le Parc d'instances Spot est annulé et n'a aucune Instances Spot en cours d'exécution. La demande de Parc d'instances Spot est supprimée deux jours après la résiliation de ses instances.

L'illustration suivante représente les transitions entre les états de la demande. Si vous dépassez les limites du Parc d'instances Spot, la demande est annulée immédiatement.

Vérifications de l'état du Parc d'instances Spot

Le Parc d'instances Spot vérifie le statut d'état des Instances Spot dans le parc toutes les deux minutes. Le statut d'état d'une instance est `healthy` ou `unhealthy`. Le Parc d'instances Spot détermine le statut

d'état d'une instance à l'aide des contrôles de statut fournis par Amazon EC2. Si le contrôle du statut de l'instance ou de celui du système est `impaired` pendant trois vérifications de l'état consécutives, le statut de l'état de l'instance est `unhealthy`. Sinon, le statut d'état est `healthy`. Pour plus d'informations, consultez Contrôles de statut pour vos instances (p. 742).

Vous pouvez configurer votre Parc d'instances Spot pour qu'il remplace les instances non saines. Après activation du remplacement lié à la vérification de l'état, une instance est remplacée lorsque son statut d'état est `unhealthy`. La taille du Parc d'instances Spot peut être inférieure à sa capacité cible pendant quelques minutes, le temps que l'instance non saine soit remplacée.

Prérequis

- Le remplacement lié à la vérification de l'état est pris en charge uniquement avec les Parcs d'instances Spot qui maintiennent une capacité cible, pas avec les Parcs d'instances Spot uniques.
- Vous pouvez configurer votre Parc d'instances Spot pour qu'il remplace les instances non saines au moment de sa création uniquement.
- Les utilisateurs IAM peuvent utiliser le remplacement lié à la vérification de l'état seulement s'ils sont autorisés à appeler l'action `ec2:DescribeInstanceStatus`.

Planification d'une demande de Parc d'instances Spot

Avant de créer une demande de Parc d'instances Spot, passez en revue les bonnes pratiques en matière d'instances Spot. Appuyez-vous sur ces bonnes pratiques lorsque vous planifiez votre demande de Parc d'instances Spot de façon à provisionner le type d'instance souhaité au prix le plus bas possible. Nous vous recommandons également d'effectuer les opérations suivantes :

- Déterminez si vous souhaitez créer un Parc d'instances Spot qui envoie une demande unique pour la capacité cible souhaitée ou qui doit maintenir une capacité cible dans le temps.
- Déterminez les types d'instance qui correspondent aux exigences de votre application.
- Déterminez la capacité cible de votre demande de Parc d'instances Spot. Vous pouvez définir la capacité cible en instances ou en unités personnalisées. Pour plus d'informations, consultez Pondération d'instance de Parc d'instances Spot (p. 384).
- Déterminez quelle portion de la capacité cible du Parc d'instances Spot doit correspondre à la capacité à la demande. Vous pouvez spécifier une capacité à la demande égale à 0.
- Déterminez le prix par unité si vous avez recours à la pondération d'instance. Pour calculer le prix par unité, divisez le prix pour une heure d'instance par le nombre d'unités (ou pondération) que cette instance représente. Si vous n'utilisez pas la pondération d'instance, le prix par unité défini par défaut est le prix par heure d'instance.
- Passez en revue les options possibles pour votre demande de Parc d'instances Spot. Pour plus d'informations, consultez la section sur la commande request-Spot-fleet dans le AWS CLI Command Reference. Pour accéder à des exemples supplémentaires, consultez Exemples de configuration d'un Parc d'instances Spot (p. 429).

Autorisations Parc d'instances Spot

Si vos utilisateurs IAM sont appelés à créer ou à gérer un Parc d'instances Spot, veillez à leur accorder les autorisations nécessaires comme suit.

Si vous utilisez la console Amazon EC2 pour créer un Parc d'instances Spot, cela crée un rôle lié au service nommé `AWSServiceRoleForEC2SpotFleet` et un rôle nommé `aws-ec2-spot-fleet-tagging-role` qui octroie à Parc d'instances Spot les autorisations de demander, de lancer, de terminer et de baliser des ressources en votre nom. Si vous utilisez AWS CLI ou une API, vous devez vous assurer que ces rôles existent.

Suivez les instructions ci-dessous pour accorder les autorisations requises et créer les rôles.

Autorisations et rôles

- Octroi de l'autorisation aux utilisateurs IAM pour Parc d'instances Spot (p. 412)
- Rôle lié à un service pour Parc d'instances Spot (p. 413)
- Rôle IAM pour Parc d'instances Spot (p. 415)

Octroi de l'autorisation aux utilisateurs IAM pour Parc d'instances Spot

Si vos utilisateurs IAM sont appelés à créer ou à gérer un Parc d'instances Spot, veillez à leur accorder les autorisations nécessaires comme suit.

Pour accorder à un utilisateur IAM des autorisations pour un Parc d'instances Spot

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le panneau de navigation, choisissez Policies (Stratégies), puis Create policy (Créer une stratégie).
3. Sur la page Créer une stratégie, choisissez JSON, puis remplacez le texte par ce qui suit.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances",
                "ec2:CreateTags",
                "ec2:RequestSpotFleet",
                "ec2:ModifySpotFleetRequest",
                "ec2:CancelSpotFleetRequests",
                "ec2:DescribeSpotFleetRequests",
                "ec2:DescribeSpotFleetInstances",
                "ec2:DescribeSpotFleetRequestHistory"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "arn:aws:iam::*:role/aws-ec2-spot-fleet-tagging-role"
        },
        {
            "Effect": "Allow",
            "Action": [
                "iam:CreateServiceLinkedRole",
                "iam:ListRoles",
                "iam:ListInstanceProfiles"
            ],
            "Resource": "*"
        }
    ]
}
```

L'exemple de stratégie précédent accorde à un utilisateur IAM les autorisations requises pour la plupart des cas d'utilisation Parc d'instances Spot. Pour limiter l'utilisateur à des actions d'API données, spécifiez celles qui sont autorisées.

EC2 et API IAM requises

Les API suivantes doivent être incluses dans la stratégie :

- `ec2:RunInstances` – Obligatoire pour lancer des instances dans un Parc d'instances Spot

- `ec2:CreateTags` – Obligatoire pour baliser la demande, les instances ou les volumes Parc d'instances Spot
- `iam:PassRole` – Obligatoire pour spécifier le rôle Parc d'instances Spot
- `iam:CreateServiceLinkedRole` – Obligatoire pour créer le rôle lié au service
- `iam:ListRoles` – Obligatoire pour énumérer les rôles IAM existants
- `iam:ListInstanceProfiles` – Obligatoire pour énumérer les profils d'instance existants

> **Important**
>
> Si vous spécifiez un rôle pour le profil d'instance IAM dans la spécification ou le modèle de lancement, vous devez accorder à l'utilisateur IAM l'autorisation de transmettre le rôle au service. Pour ce faire, dans la stratégie IAM, incluez `"arn:aws:iam::*:role/IamInstanceProfile-role"` comme ressource pour l'action `iam:PassRole`. Pour plus d'informations, consultez Octroi d'autorisations à un utilisateur pour transférer un rôle à un service AWS dans le IAM Guide de l'utilisateur.

API Parc d'instances Spot

Ajoutez les actions d'API Parc d'instances Spot suivantes à votre stratégie, selon vos besoins :

- `ec2:RequestSpotFleet`
- `ec2:ModifySpotFleetRequest`
- `ec2:CancelSpotFleetRequests`
- `ec2:DescribeSpotFleetRequests`
- `ec2:DescribeSpotFleetInstances`
- `ec2:DescribeSpotFleetRequestHistory`

API IAM facultatives

(Facultatif) Pour autoriser un utilisateur IAM à créer des rôles ou des profils d'instances à l'aide de la console IAM, vous devez aussi ajouter les actions suivantes à la stratégie :

- `iam:AddRoleToInstanceProfile`
- `iam:AttachRolePolicy`
- `iam:CreateInstanceProfile`
- `iam:CreateRole`
- `iam:GetRole`
- `iam:ListPolicies`

4. Choisissez Examiner une stratégie.
5. Sur la page Review Policy (Vérifier la stratégie), saisissez un nom et une description pour la stratégie, puis choisissez Create policy (Créer une stratégie).
6. Dans le panneau de navigation, choisissez Utilisateurs et sélectionnez l'utilisateur.
7. Dans Autorisations, choisissez Ajouter des autorisations.
8. Choisissez Attacher directement les stratégies existantes. Sélectionnez la stratégie que vous avez créée précédemment, puis choisissez Suivant : Vérification.
9. Choisissez Add permissions.

Rôle lié à un service pour Parc d'instances Spot

Amazon EC2 utilise des rôles liés à un service pour les autorisations requises pour appeler d'autres services AWS en votre nom. Un rôle lié à un service est un type unique de rôle IAM directement lié à

un service AWS. Les rôles liés à un service offrent une manière sécurisée d'accorder des autorisations aux services AWS, car seul le service lié peut assumer un rôle lié à un service. Pour plus d'informations, consultez Utilisation des rôles liés à un service dans le IAM Guide de l'utilisateur.

Amazon EC2 utilise le rôle lié à un service nommé AWSServiceRoleForEC2SpotFleet pour lancer et gérer des instances en votre nom.

> **Important**
>
> Si vous spécifiez une AMI chiffrée (p. 168) ou un instantané Amazon EBS chiffré (p. 1194) dans votre Parc d'instances Spot, vous devez accorder au rôle AWSServiceRoleForec2SpotFleet l'autorisation d'utiliser la CMK afin que Amazon EC2 puisse lancer des instances en votre nom. Pour de plus amples informations, veuillez consulter Octroi d'un accès aux CMK en vue de leur utilisation avec les AMI chiffrées et les instantanés EBS (p. 415).

Autorisations octroyées par AWSServiceRoleForEC2SpotFleet

Amazon EC2 utilise AWSServiceRoleForEC2SpotFleet pour réaliser les actions suivantes :

- `ec2:RequestSpotInstances` - Demander des Instances Spot
- `ec2:RunInstances` - Lancer des instances
- `ec2:TerminateInstances` - Résilier des instances
- `ec2:DescribeImages` - Décrire des images Amazon Machine Images (AMI) pour les instances
- `ec2:DescribeInstanceStatus` - Décrire le statut des instances.
- `ec2:DescribeSubnets` - Décrire les sous-réseaux des instances
- `ec2:CreateTags` - Ajouter des balises à la demande, aux instances et aux volumes Parc d'instances Spot
- `elasticloadbalancing:RegisterInstancesWithLoadBalancer` - Ajouter les instances spécifiées à l'équilibreur de charge indiqué.
- `elasticloadbalancing:RegisterTargets` - Enregistrer les cibles spécifiées auprès du groupe cible indiqué.

Création du rôle lié à un service

Dans la plupart des cas, vous n'avez pas besoin de créer manuellement un rôle lié à un service. Amazon EC2 crée le rôle lié à un service AWSServiceRoleForEC2SpotFleet la première fois que vous créez un Parc d'instances Spot à l'aide de la console.

Si vous utilisez AWS CLI ou une API, vous devez vous assurer que ce rôle existe.

Si vous aviez une demande d'Parc d'instances Spot active avant novembre 2017, moment à partir duquel Amazon EC2 a commencé à prendre en charge ce rôle lié à un service, Amazon EC2 a créé le rôle AWSServiceRoleForEC2SpotFleet dans votre compte AWS. Pour plus d'informations, consultez Un nouveau rôle est apparu dans Mon Compte AWS dans le IAM Guide de l'utilisateur.

Assurez-vous que ce rôle existe avant d'utiliser l'AWS CLI ou une API pour créer un Parc d'instances Spot. Pour créer le rôle, utilisez la console IAM comme suit.

Pour créer manuellement le rôle lié à un service AWSServiceRoleForEC2SpotFleet

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le volet de navigation, choisissez Rôles.
3. Choisissez Créer un rôle.
4. Pour Select type of trusted entity (Sélectionner le type d'entité de confiance), choisissez Service AWS.

5. Dans la liste des services, choisissez EC2.
6. Dans la section Select your use case (Sélectionner votre cas d'utilisation), choisissez EC2 - Spot Fleet (EC2 - Parc d'instances Spot).
7. Choisissez Étape suivante : autorisations.
8. Sur la page suivante, choisissez Suivant : Vérification.
9. Sur la page Vérification, choisissez Create Role (Créer un rôle).

Si vous n'avez plus besoin d'utiliser le Parc d'instances Spot, nous vous recommandons de supprimer le rôle AWSServiceRoleForEC2SpotFleet. Après sa suppression de votre compte, Amazon EC2 créera de nouveau le rôle si vous effectuez une demande Parc d'instances Spot à l'aide de la console. Pour de plus amples informations, veuillez consulter Suppression d'un rôle lié à un service dans le IAM Guide de l'utilisateur.

Octroi d'un accès aux CMK en vue de leur utilisation avec les AMI chiffrées et les instantanés EBS

Si vous spécifiez une AMI chiffrée (p. 168) ou un instantané Amazon EBS chiffré (p. 1194) pour votre demande Parc d'instances Spot et que vous utilisez la clé principale client gérée (CMK) pour le chiffrement, vous devez autoriser le rôle AWSServiceRoleForEC2SpotFleet à utiliser la CMK afin que Amazon EC2 puisse lancer des instances en votre nom. Pour cela, vous devez ajouter une autorisation à la CMK, comme indiqué dans la procédure suivante.

Lorsque vous définissez les autorisations, les octrois constituent une alternative aux stratégies de clé. Pour de plus amples informations, veuillez consulter Utilisation des autorisations et Utilisation des stratégies de clé dans AWS KMS dans le AWS Key Management Service Developer Guide.

Pour autoriser le rôle AWSServiceRoleForEC2SpotFleet à utiliser la CMK

- Utilisez la commande create-grant pour ajouter un octroi à la CMK et spécifier le mandataire (le rôle lié à un service AWSServiceRoleForEC2SpotFleet) qui reçoit l'autorisation d'effectuer les opérations autorisées par l'octroi. La CMK est spécifiée par le paramètre `key-id` et l'ARN de la CMK. Le mandataire est spécifié par le paramètre `grantee-principal` et l'ARN du rôle lié à un service AWSServiceRoleForEC2SpotFleet.

```
aws kms create-grant \
    --region us-east-1 \
    --key-id arn:aws:kms:us-
east-1:444455556666:key/1234abcd-12ab-34cd-56ef-1234567890ab \
    --grantee-principal arn:aws:iam::111122223333:role/AWSServiceRoleForEC2SpotFleet \
    --operations "Decrypt" "Encrypt" "GenerateDataKey"
 "GenerateDataKeyWithoutPlaintext" "CreateGrant" "DescribeKey" "ReEncryptFrom"
 "ReEncryptTo"
```

Rôle IAM pour Parc d'instances Spot

Le rôle `aws-ec2-spot-fleet-tagging-role` IAM accorde à Parc d'instances Spot l'autorisation de baliser la demande, les instances et les volumes Parc d'instances Spot. Pour de plus amples informations, veuillez consulter Balisage d'un Parc d'instances Spot (p. 420).

> **Important**
>
> Si vous choisissez de baliser des instances dans le parc et que vous choisissez de maintenir la capacité cible (la demande de Parc d'instances Spot est de type `maintain`), les différences entre les autorisations de l'utilisateur IAM et de `IamFleetRole` peuvent entraîner un comportement de balisage incohérent pour les instances du parc. Si le rôle `IamFleetRole` n'inclut pas l'autorisation `CreateTags`, il se peut que certaines instances lancées par le parc ne soient pas balisées. En attendant que cette incohérence soit corrigée, pour vous assurer que toutes les instances lancées

par le parc sont marquées, nous vous recommandons d'utiliser le rôle `aws-ec2-spot-fleet-tagging-role` pour `IamFleetRole`. Autre option : pour utiliser un rôle existant, attachez la stratégie gérée AWS `AmazonEC2SpotFleetTaggingRole` au rôle existant. Sinon, vous devrez ajouter manuellement l'autorisation `CreateTags` à votre stratégie.

Pour créer le rôle IAM de balisage de Parc d'instances Spot

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le volet de navigation, choisissez Rôles.
3. Sur la page Select type of trusted entity (Sélectionner le type d'entité de confiance), choisissez AWS service (Service AWS), EC2, EC2 - Spot Fleet Tagging (EC2 - Balisage du parc d'instances Spot), Suivant : Autorisations.
4. Dans la page Attached permissions policy (Stratégie d'autorisations attachée), choisissez Next:Review (Suivant : Vérification).
5. Sur la page Vérification, attribuez un nom au rôle (**`aws-ec2-spot-fleet-tagging-role`** par exemple), puis choisissez Créer un rôle.

Création d'une demande de Parc d'instances Spot

À l'aide d'AWS Management Console, créez rapidement une demande de Parc d'instances Spot en choisissant uniquement vos besoins pour votre application ou votre tâche et les spécifications minimales de calcul. Amazon EC2 configure un parc qui répond le mieux à vos besoins et qui est conforme aux bonnes pratiques en matière d'instances Spot. Pour plus d'informations, consultez Création rapide d'une demande de Parc d'instances Spot (console) (p. 416). Sinon, vous pouvez modifier l'un des paramètres par défaut. Pour plus d'informations, consultez Création d'une demande de Parc d'instances Spot à l'aide des paramètres définis (console) (p. 417).

Création rapide d'une demande de Parc d'instances Spot (console)

Pour créer rapidement une demande de Parc d'instances Spot, procédez comme suit.

Pour créer une demande de Parc d'instances Spot à l'aide des paramètres recommandés (console)

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot.
2. Si vous utilisez les instances Spot pour la première fois, sélectionnez Mise en route. Sinon, sélectionnez Demander des Instances Spot.
3. Pour Indiquer vos besoins en application ou en tâche, choisissez Charges de travail d'équilibrage de charge, Charges de travail flexibles, Charges de travail Big Data ou Charges de travail avec une durée définie.
4. Sous Configure your instances (Configurer vos instances), pour Minimum compute unit (Unité de calcul minimale), choisissez les spécifications matérielles minimales (vCPU, mémoire et stockage) dont vous avez besoin pour votre application ou votre tâche, as specs (en tant que spécifications) ou as an instance type (en tant que type d'instance).
   - Pour as specs (en tant que spécifications), indiquez le nombre de vCPU requis et la quantité de mémoire requise.
   - Pour as an instance type (en tant que type d'instance), acceptez le type d'instance par défaut ou choisissez Changer le type d'instance pour choisir un autre type d'instance.
5. Sous Tell us how much capacity you need (Indiquez la capacité dont vous avez besoin), pour Total target capacity (Capacité cible totale), indiquez le nombre d'unités à demander pour la capacité cible. Vous pouvez choisir des instances ou des vCPU.
6. Passez en revue la section Fleet request settings (Paramètres de demande de parc) selon l'application ou la tâche sélectionnée, puis choisissez Lancer.

Création d'une demande de Parc d'instances Spot à l'aide des paramètres définis (console)

Vous pouvez créer un Parc d'instances Spot à l'aide des paramètres que vous définissez.

Pour créer une demande de Parc d'instances Spot à l'aide des paramètres définis (console)

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot.
2. Si vous utilisez les instances Spot pour la première fois, sélectionnez Mise en route. Sinon, sélectionnez Demander des Instances Spot.
3. Pour Indiquer vos besoins en application ou en tâche, choisissez Charges de travail d'équilibrage de charge, Charges de travail flexibles, Charges de travail Big Data ou Charges de travail avec une durée définie.
4. Pour Configure your instances (Configurer vos instances), procédez comme suit :

   a. (Facultatif) Pour Launch template (Modèle de lancement), choisissez un modèle de lancement. Le modèle de lancement doit spécifier une Amazon Machine Image (AMI), car vous ne pouvez pas remplacer l'AMI par un Parc d'instances Spot si vous spécifiez un modèle de lancement.

      > **Important**
      >
      > Si vous prévoyez de spécifier Optional On-Demand portion (Partie à la demande facultative), vous devez choisir un modèle de lancement.

   b. Pour AMI, choisissez l'une des AMI basiques fournies par AWS, ou choisissez Search for AMI (Rechercher une AMI) pour utiliser une AMI issue de notre communauté d'utilisateurs ou d'AWS Marketplace, ou une AMI vous appartenant.

   c. Pour Minimum compute unit (Unité de calcul minimale), choisissez les spécifications matérielles minimales (vCPU, mémoire et stockage) dont vous avez besoin pour votre application ou votre tâche, as specs (en tant que spécifications) ou as an instance type (en tant que type d'instance).

      - Pour as specs (en tant que spécifications), indiquez le nombre de vCPU requis et la quantité de mémoire requise.
      - Pour as an instance type (en tant que type d'instance), acceptez le type d'instance par défaut ou choisissez Changer le type d'instance pour choisir un autre type d'instance.

   d. Pour Réseau, choisissez un VPC existant ou créez-en un.

      [VPC existant] Choisissez le VPC.

      [Nouveau VPC] Choisissez Créer un nouveau VPC pour accéder à la console Amazon VPC. Lorsque vous avez terminé, revenez dans l'assistant et actualisez la liste.

   e. (Facultatif) Pour Zone de disponibilité, laissez AWS choisir les zones de disponibilité de vos Instances Spot ou indiquez une ou plusieurs zones de disponibilité.

      Si vous avez plusieurs sous-réseaux dans une zone de disponibilité, choisissez le sous-réseau approprié dans Sous-réseau. Pour ajouter des sous-réseaux, choisissez Créer un nouveau sous-réseau pour accéder à la console Amazon VPC. Lorsque vous avez terminé, revenez dans l'assistant et actualisez la liste.

   f. (Facultatif) Pour Nom de la paire de clés, choisissez une paire de clés existante ou créez-en une.

      [Paire de clés existante] Choisissez la paire de clés.

      [Nouvelle paire de clés] Choisissez Create new key pair (Créer une nouvelle paire de clés) pour accéder à la console Amazon VPC. Lorsque vous avez terminé, revenez dans l'assistant et actualisez la liste.

5. (Facultatif) Pour Additional configurations (Configurations supplémentaires), procédez comme suit :

   a. (Facultatif) Pour ajouter de l'espace de stockage supplémentaire, spécifiez des volumes de stockage d'instance ou des volumes Amazon EBS supplémentaires, selon le type d'instance.

b. (Facultatif) Pour activer l'optimisation Amazon EBS, choisissez Launch EBS-optimized instances (Lancer les instances optimisées pour EBS) pour Optimisé pour EBS.

c. (Facultatif) Pour ajouter de l'espace de stockage temporaire de niveau bloc pour vos instances, choisissez Attach at launch (Attacher au lancement) pour Stockage d'instance.

d. (Facultatif) Par défaut, la surveillance basique est activée pour vos instances. Pour activer la surveillance détaillée, pour Monitoring (Surveillance), choisissez Enable CloudWatch detailed monitoring (Activer la surveillance détaillée CW).

e. (Facultatif) Pour remplacer des instances non saines, dans Health check (Contrôle d'intégrité), choisissez Replace unhealthy instances (Remplacer les instances non saines). Pour activer cette option, vous devez d'abord choisir Maintain target capacity (Maintenir la capacité cible).

f. (Facultatif) Pour exécuter une instance Spot dédiée, pour Location, choisissez Dédié - Exécute une instance dédiée.

g. (Facultatif) Pour Groupes de sécurité, choisissez un ou plusieurs groupes de sécurité ou créez-en un.

   [Groupe de sécurité existant] Choisissez un ou plusieurs groupes de sécurité.

   [Nouveau groupe de sécurité] Choisissez Create new security group (Créer un nouveau groupe de sécurité) pour accéder à la console Amazon VPC. Lorsque vous avez terminé, revenez dans l'assistant et actualisez la liste.

h. (Facultatif) Pour rendre vos instances accessibles depuis Internet, choisissez Activer pour Auto-assign IPv4 Public IP (Attribuer automatiquement une adresse IP publique IPv4).

i. (Facultatif) Pour lancer vos Instances Spot avec un rôle IAM, pour IAM instance profile (Profil d'instance IAM), choisissez le rôle.

j. (Facultatif) Pour exécuter un script de démarrage, copiez-le dans Données utilisateur.

k. (Facultatif) Pour ajouter une balise, choisissez Ajouter une balise et entrez la clé et la valeur de la balise. Répétez l'opération pour chaque balise.

   Pour baliser les instances et la demande de Parc d'instances Spot avec la même balise, assurez-vous que les balises d'instance et les balises de parc sont sélectionnées. Pour balisées uniquement les instances lancées par le parc, désactivez les balises de parc. Pour baliser uniquement la demande de Parc d'instances Spot, désactivez les balises d'instance.

6. Pour Tell us how much capacity you need (Indiquez la capacité dont vous avez besoin), procédez comme suit :

   a. Pour Total target capacity (Capacité cible totale), indiquez le nombre d'unités à demander pour la capacité cible. Vous pouvez choisir des instances ou des vCPU. Pour spécifier une capacité cible de 0 afin d'ajouter une capacité ultérieurement, choisissez Maintain target capacity (Maintenir la capacité cible).

   b. (Facultatif) Pour Optional On-Demand portion (Partie à la demande facultative), indiquez le nombre d'unités à la demande à demander. Ce nombre doit être inférieur à la valeur du champ Total target capacity (Capacité cible totale). Amazon EC2 calcule la différence et l'alloue aux unités Spot à demander.

      > **Important**
      >
      > Pour spécifier une partie à la demande facultative, vous devez commencer par choisir un modèle de lancement.

   c. (Facultatif) Par défaut, le service Spot résilie les Instances Spot lorsqu'elles sont interrompues. Pour maintenir la capacité cible, sélectionnez Maintain target capacity (Maintenir la capacité cible). Vous pouvez ensuite spécifier que le service Spot résilie, arrête ou met en veille prolongée les Instances Spot lorsqu'elles sont interrompues. Pour ce faire, choisissez l'option correspondante à partir de Interruption behavior (Comportement d'interruption).

   d. (Facultatif) Pour contrôler le montant que vous payez par heure pour l'ensemble des Instances Spot de votre parc, sélectionnez Conserver le coût cible pour Spot (avancé - facultatif), puis

entrez le montant total maximal que vous êtes prêt à payer par heure. Une fois le prix total atteint, Parc d'instances Spot arrête de lancer des Instances Spot même si la capacité cible n'a pas été atteinte. Pour de plus amples informations, veuillez consulter Contrôle des dépenses (p. 383).

7. Pour Fleet request settings (Paramètres de demande de parc), procédez comme suit :
   a. Passez en revue la demande de parc et la stratégie d'allocation de parc selon l'application ou la tâche sélectionnée. Pour modifier les types d'instances ou la stratégie d'allocation, décochez la case Apply recommendations (Appliquer les recommandations).
   b. (Facultatif) Pour supprimer des types d'instances, choisissez Supprimer pour Fleet request (Demande de parc). Pour ajouter des types d'instances, choisissez Select instance types (Sélectionner des types d'instances).
   c. (Facultatif) Pour Fleet allocation strategy (Stratégie d'allocation de parc), choisissez la stratégie qui répond à vos besoins. Pour plus d'informations, consultez Stratégie d'allocation pour les Instances Spot (p. 381).
8. Pour Additional request details (Détails de la demande supplémentaire), procédez comme suit :
   a. Vérifiez les détails de la demande supplémentaire. Pour effectuer des modifications, décochez la case Apply defaults (Appliquer les valeurs par défaut).
   b. (Facultatif) Pour IAM fleet role (Rôle de parc IAM), vous pouvez utiliser le rôle par défaut ou choisir un autre rôle. Choisissez Use default role (Utiliser le rôle par défaut) pour utiliser le rôle par défaut après avoir changé de rôle.
   c. (Facultatif) Pour Prix maximum, vous pouvez utiliser le prix maximum par défaut (prix à la demande) ou indiquer le prix maximum que vous êtes prêt à payer. Vos Instances Spot ne sont pas lancées si votre prix maximal est inférieur au prix spot pour les types d'instance que vous avez sélectionnés.
   d. (Facultatif) Pour créer une demande valide uniquement pendant une période spécifique, modifiez les valeurs des champs Demande valide du et Demande valide jusqu'au.
   e. (Facultatif) Par défaut, nous résilions vos Instances Spot à l'expiration de la demande. Si vous souhaitez qu'elles continuent de s'exécuter après l'expiration de votre demande, décochez la case Terminate the instances when the request expires (Résilier les instances lorsque la demande expire).
   f. (Facultatif) Pour enregistrer vos Instances Spot auprès d'un équilibreur de charge, choisissez Receive traffic from one or more load balancers (Recevoir le trafic d'un ou plusieurs équilibreurs de charge) et choisissez un ou plusieurs Equilibreurs de charge classiques ou groupes cibles.
9. (Facultatif) Pour télécharger une copie de la configuration de lancement à utiliser avec l'AWS CLI, sélectionnez JSON Config (Configuration JSON).
10. Choisissez Launch.

    Le type de demande de Parc d'instances Spot est `fleet`. Une fois la demande exécutée, les demandes de type `instance` sont ajoutées, avec l'état `active` et le statut `fulfilled`.

Pour créer une demande de Parc d'instances Spot à partir de l'AWS CLI

- Utilisez la commande request-spot-fleet pour créer une demande de Parc d'instances Spot.

```
aws ec2 request-spot-fleet --spot-fleet-request-config file://config.json
```

Pour accéder à des exemples de fichiers de configuration, consultez Exemples de configuration d'un Parc d'instances Spot (p. 429).

Voici un exemple de sortie :

```
{
```

```
    "SpotFleetRequestId": "sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE"
}
```

Balisage d'un Parc d'instances Spot

Pour vous aider à classer et à gérer vos demandes de Parc d'instances Spot, vous pouvez les baliser avec des métadonnées personnalisées. Vous pouvez affecter une balise à une demande de Parc d'instances Spot lorsque vous la créez, ou après. Vous pouvez attribuer des balises à l'aide de la console Amazon EC2 ou d'un outil de ligne de commande.

Lorsque vous balisez une demande Parc d'instances Spot, les instances et les volumes lancés par le Parc d'instances Spot ne sont pas automatiquement balisés. Vous devez baliser explicitement les instances et les volumes lancés par le Parc d'instances Spot. Vous pouvez choisir d'affecter des balises uniquement à la demande Parc d'instances Spot, ou uniquement aux instances lancées par le parc, ou uniquement aux volumes attachés aux instances lancées par le parc, ou aux trois.

> **Note**
>
> Les balises de volume ne sont prises en charge que pour les volumes attachés à Instances à la demande. Vous ne pouvez pas baliser les volumes attachés à Instances Spot.

Pour plus d'informations sur le fonctionnement des balises, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Sommaire

- Prérequis (p. 420)
- Balisage d'un nouveau Parc d'instances Spot (p. 421)
- Balisage d'un nouveau Parc d'instances Spot et des instances et volumes qu'il lance (p. 422)
- Balisage d'un Parc d'instances Spot existant (p. 424)
- Affichage des balises de demande Parc d'instances Spot (p. 425)

Prérequis

Octroyez à l'utilisateur IAM l'autorisation de baliser les ressources. Pour plus d'informations, consultez Exemple : Balisage de vos ressources (p. 1018).

Pour accorder à un utilisateur IAM l'autorisation de baliser les ressources

Créez une stratégie IAM qui inclut les éléments suivants :

- L'action `ec2:CreateTags`. Celle-ci accorde à l'utilisateur IAM l'autorisation de créer des balises.
- L'action `ec2:RequestSpotFleet`. Celle-ci accorde à l'utilisateur IAM l'autorisation de créer une demande de Parc d'instances Spot.
- Pour `Resource`, vous devez spécifier `"*"`. Cela permet aux utilisateurs de baliser tous les types de ressources.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "TagSpotFleetRequest",
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags",
                "ec2:RequestSpotFleet"
            ],
            "Resource": "*"
```

```
}
```

> **Important**
>
> Actuellement, nous ne prenons pas en charge les autorisations de niveau ressource pour la ressource `spot-fleet-request`. Si vous spécifiez `spot-fleet-request` en tant que ressource, vous recevrez une exception de non-autorisation lorsque vous tenterez de baliser le parc. L'exemple suivant illustre comment ne pas définir la stratégie.

```
{
    "Effect": "Allow",
    "Action": [
        "ec2:CreateTags",
        "ec2:RequestSpotFleet"
    ],
    "Resource": "arn:aws:ec2:us-east-1:111122223333:spot-fleet-request/*"
}
```

Balisage d'un nouveau Parc d'instances Spot

Pour baliser une nouvelle demande de Parc d'instances Spot à l'aide de la console

1. Suivez la procédure Création d'une demande de Parc d'instances Spot à l'aide des paramètres définis (console) (p. 417).
2. Pour ajouter une balise, développez Additional configurations (Configurations supplémentaires), choisissez Add new tag (Ajouter une nouvelle balise), puis entrez la clé et la valeur de la balise. Répétez l'opération pour chaque balise.

   Pour chaque balise, vous pouvez marquer la demande de Parc d'instances Spot et les instances avec la même balise. Pour baliser les deux, assurez-vous que les balises d'instance et les balises de parc sont sélectionnées. Pour baliser uniquement la demande de Parc d'instances Spot, désactivez les balises d'instance. Pour balisées uniquement les instances lancées par le parc, désactivez les balises de parc.
3. Remplissez les champs obligatoires pour créer une demande de Parc d'instances Spot, puis choisissez Launch (Lancer). Pour plus d'informations, consultez Création d'une demande de Parc d'instances Spot à l'aide des paramètres définis (console) (p. 417).

Pour baliser une nouvelle demande de Parc d'instances Spot à l'aide de l'AWS CLI

Pour baliser une demande de Parc d'instances Spot lors de sa création, configurez la demande de Parc d'instances Spot comme suit :

- Spécifiez les balises de la demande de Parc d'instances Spot dans `SpotFleetRequestConfig`.
- Pour `ResourceType`, spécifiez `spot-fleet-request`. Si vous indiquez une autre valeur, la demande de parc échouera.
- Pour `Tags`, spécifiez la paire clé-valeur. Vous pouvez définir plusieurs paires clé-valeur.

Dans l'exemple suivant, la demande de Parc d'instances Spot est marquée par deux balises : Key=Environment et Value=Production, ainsi que Key=Cost-Center et Value=123.

```
{
    "SpotFleetRequestConfig": {
        "AllocationStrategy": "lowestPrice",
        "ExcessCapacityTerminationPolicy": "default",
        "IamFleetRole": "arn:aws:iam::111122223333:role/aws-ec2-spot-fleet-tagging-role",
        "LaunchSpecifications": [
```

```
            {
                "ImageId": "ami-0123456789EXAMPLE",
                "InstanceType": "c4.large"
            }
        ],
        "SpotPrice": "5",
        "TargetCapacity": 2,
        "TerminateInstancesWithExpiration": true,
        "Type": "maintain",
        "ReplaceUnhealthyInstances": true,
        "InstanceInterruptionBehavior": "terminate",
        "InstancePoolsToUseCount": 1,
        "TagSpecifications": [
            {
                "ResourceType": "spot-fleet-request",
                "Tags": [
                    {
                        "Key": "Environment",
                        "Value":"Production"
                    },
                    {
                        "Key": "Cost-Center",
                        "Value":"123"
                    }
                ]
            }
        ]
    }
}
```

Balisage d'un nouveau Parc d'instances Spot et des instances et volumes qu'il lance

Pour baliser une nouvelle demande Parc d'instances Spot et les instances qu'elle lance à l'aide de l'AWS CLI

Pour baliser une demande Parc d'instances Spot lors de sa création et pour baliser les instances lorsqu'elles sont lancées par le parc, configurez la demande Parc d'instances Spot comme suit :

Balises de la demande deParc d'instances Spot :

- Spécifiez les balises de la demande de Parc d'instances Spot dans `SpotFleetRequestConfig`.
- Pour `ResourceType`, spécifiez `spot-fleet-request`. Si vous indiquez une autre valeur, la demande de parc échouera.
- Pour `Tags`, spécifiez la paire clé-valeur. Vous pouvez définir plusieurs paires clé-valeur.

Balises d'instance :

- Spécifiez les balises des instances dans `LaunchSpecifications`.
- Pour `ResourceType`, spécifiez `instance`. Si vous indiquez une autre valeur, la demande de parc échouera.
- Pour `Tags`, spécifiez la paire clé-valeur. Vous pouvez définir plusieurs paires clé-valeur.

  Vous pouvez également spécifier les balises de l'instance dans le modèle de lancement (p. 539) référencé dans la demande de Parc d'instances Spot.

Balises de volume :

- Spécifiez les balises des volumes dans le modèle de lancement (p. 539) référencé dans la demande Parc d'instances Spot. Le balisage de volume dans `LaunchSpecifications` n'est pas pris en charge.

Dans l'exemple suivant, la demande de Parc d'instances Spot est marquée par deux balises : Key=Environment et Value=Production, ainsi que Key=Cost-Center et Value=123. Les instances qui sont lancées par le parc sont marquées avec une balise (qui est la même que l'une des balises de la demande de Parc d'instances Spot) : Key=Cost-Center et Value=123.

```
{
    "SpotFleetRequestConfig": {
        "AllocationStrategy": "lowestPrice",
        "ExcessCapacityTerminationPolicy": "default",
        "IamFleetRole": "arn:aws:iam::111122223333:role/aws-ec2-spot-fleet-tagging-role",
        "LaunchSpecifications": [
            {
                "ImageId": "ami-0123456789EXAMPLE",
                "InstanceType": "c4.large",
                "TagSpecifications": [
                    {
                        "ResourceType": "instance",
                        "Tags": [
                            {
                                "Key": "Cost-Center",
                                "Value": "123"
                            }
                        ]
                    }
                ]
            }
        ],
        "SpotPrice": "5",
        "TargetCapacity": 2,
        "TerminateInstancesWithExpiration": true,
        "Type": "maintain",
        "ReplaceUnhealthyInstances": true,
        "InstanceInterruptionBehavior": "terminate",
        "InstancePoolsToUseCount": 1,
        "TagSpecifications": [
            {
                "ResourceType": "spot-fleet-request",
                "Tags": [
                    {
                        "Key": "Environment",
                        "Value":"Production"
                    },
                    {
                        "Key": "Cost-Center",
                        "Value":"123"
                    }
                ]
            }
        ]
    }
}
```

Pour baliser les instances lancées par un Parc d'instances Spot à l'aide de l'AWS CLI

Pour balisées les instances lorsqu'elles sont lancées par le parc, vous pouvez spécifier les balises dans le modèle de lancement (p. 539) référencé dans la demande de Parc d'instances Spot ou dans la configuration de la demande de Parc d'instances Spot comme suit :

- Spécifiez les balises des instances dans `LaunchSpecifications`.
- Pour `ResourceType`, spécifiez `instance`. Si vous indiquez une autre valeur, la demande de parc échouera.
- Pour `Tags`, spécifiez la paire clé-valeur. Vous pouvez définir plusieurs paires clé-valeur.

Dans l'exemple suivant, les instances lancées par le parc sont marquées avec une balise : Key=Cost-Center et Value=123.

```
{
    "SpotFleetRequestConfig": {
        "AllocationStrategy": "lowestPrice",
        "ExcessCapacityTerminationPolicy": "default",
        "IamFleetRole": "arn:aws:iam::111122223333:role/aws-ec2-spot-fleet-tagging-role",
        "LaunchSpecifications": [
            {
                "ImageId": "ami-0123456789EXAMPLE",
                "InstanceType": "c4.large",
                "TagSpecifications": [
                    {
                        "ResourceType": "instance",
                        "Tags": [
                            {
                                "Key": "Cost-Center",
                                "Value": "123"
                            }
                        ]
                    }
                ]
            }
        ],
        "SpotPrice": "5",
        "TargetCapacity": 2,
        "TerminateInstancesWithExpiration": true,
        "Type": "maintain",
        "ReplaceUnhealthyInstances": true,
        "InstanceInterruptionBehavior": "terminate",
        "InstancePoolsToUseCount": 1
    }
}
```

Pour baliser les volumes attachés à Instances à la demande lancés par un Parc d'instances Spot l'aide de AWS CLI

Pour baliser des instances lorsqu'elles sont lancées par le parc d'instances, spécifiez les balises dans le modèle de lancement (p. 539) référencé dans la demande Parc d'instances Spot.

> **Note**
>
> Les balises de volume ne sont prises en charge que pour les volumes attachés à Instances à la demande. Vous ne pouvez pas baliser les volumes attachés à Instances Spot.
> Le balisage de volume dans `LaunchSpecifications` n'est pas pris en charge.

Balisage d'un Parc d'instances Spot existant

Pour baliser une demande de Parc d'instances Spot existante à l'aide de la console

Après avoir créé une demande de Parc d'instances Spot, vous pouvez ajouter des balises à la demande de parc à l'aide de la console.

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot.
2. Sélectionnez votre demande de Parc d'instances Spot.
3. Choisissez l'onglet Tags (Balises), puis Create Tag (Créer une balise).

Pour baliser une demande de Parc d'instances Spot existante à l'aide de l'AWS CLI

Utilisez la commande create-tags pour baliser les ressources existantes. Dans l'exemple suivant, la demande de Parc d'instances Spot existante est balisée avec Key=purpose et Value=test.

```
aws ec2 create-tags \
    --resources sfr-11112222-3333-4444-5555-66666EXAMPLE \
    --tags Key=purpose,Value=test
```

Affichage des balises de demande Parc d'instances Spot

Pour afficher les balises d'une demande de Parc d'instances Spot à l'aide de la console

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot.
2. Sélectionnez votre demande de Parc d'instances Spot et choisissez l'onglet Tags (Balises).

Pour décrire les balises de la demande de Parc d'instances Spot

Utilisez la commande describe-tags pour afficher les balises de la ressource spécifiée. Dans l'exemple suivant, vous décrivez les balises de la demande de Parc d'instances Spot spécifiée.

```
aws ec2 describe-tags \
    --filters "Name=resource-id,Values=sfr-11112222-3333-4444-5555-66666EXAMPLE"
```

```
{
    "Tags": [
        {
            "Key": "Environment",
            "ResourceId": "sfr-11112222-3333-4444-5555-66666EXAMPLE",
            "ResourceType": "spot-fleet-request",
            "Value": "Production"
        },
        {
            "Key": "Another key",
            "ResourceId": "sfr-11112222-3333-4444-5555-66666EXAMPLE",
            "ResourceType": "spot-fleet-request",
            "Value": "Another value"
        }
    ]
}
```

Vous pouvez également afficher les balises d'une demande de Parc d'instances Spot en décrivant la demande Parc d'instances Spot.

Utilisez la commande describe-spot-fleet-requests pour afficher la configuration de la demande Parc d'instances Spot spécifiée, qui inclut toutes les balises définies pour la demande de parc.

```
aws ec2 describe-spot-fleet-requests \
    --spot-fleet-request-ids sfr-11112222-3333-4444-5555-66666EXAMPLE
```

```
{
    "SpotFleetRequestConfigs": [
        {
            "ActivityStatus": "fulfilled",
            "CreateTime": "2020-02-13T02:49:19.709Z",
            "SpotFleetRequestConfig": {
                "AllocationStrategy": "capacityOptimized",
                "OnDemandAllocationStrategy": "lowestPrice",
                "ExcessCapacityTerminationPolicy": "Default",
                "FulfilledCapacity": 2.0,
                "OnDemandFulfilledCapacity": 0.0,
                "IamFleetRole": "arn:aws:iam::111122223333:role/aws-ec2-spot-fleet-tagging-
role",
                "LaunchSpecifications": [
```

```
                    {
                        "ImageId": "ami-0123456789EXAMPLE",
                        "InstanceType": "c4.large"
                    }
                ],
                "TargetCapacity": 2,
                "OnDemandTargetCapacity": 0,
                "Type": "maintain",
                "ReplaceUnhealthyInstances": false,
                "InstanceInterruptionBehavior": "terminate"
            },
            "SpotFleetRequestId": "sfr-11112222-3333-4444-5555-66666EXAMPLE",
            "SpotFleetRequestState": "active",
            "Tags": [
                {
                    "Key": "Environment",
                    "Value": "Production"
                },
                {
                    "Key": "Another key",
                    "Value": "Another value"
                }
            ]
        }
    ]
}
```

Surveillance de votre Parc d'instances Spot

Le Parc d'instances Spot lance des Instances Spot lorsque votre prix maximum dépasse le prix spot et que la capacité est disponible. Les Instances Spot s'exécutent jusqu'à ce qu'elles soient interrompues ou que vous les résiliiez.

Pour surveiller votre Parc d'instances Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez votre demande de Parc d'instances Spot. Pour afficher les détails de la configuration, choisissez Description.
4. Pour répertorier les Instances Spot du Parc d'instances Spot, choisissez Instances.
5. Pour afficher l'historique du Parc d'instances Spot, choisissez Historique.

Pour surveiller votre Parc d'instances Spot (AWS CLI)

Utilisez la commande describe-spot-fleet-requests pour décrire vos demandes de Parc d'instances Spot.

```
aws ec2 describe-spot-fleet-requests
```

Utilisez la commande describe-spot-fleet-instances pour décrire les Instances Spot du Parc d'instances Spot spécifié.

```
aws ec2 describe-spot-fleet-instances \
    --spot-fleet-request-id sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE
```

Utilisez la commande describe-spot-fleet-request-history pour décrire l'historique de la demande de Parc d'instances Spot spécifiée.

```
aws ec2 describe-spot-fleet-request-history \
    --spot-fleet-request-id sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
```

```
    --start-time 2015-05-18T00:00:00Z
```

Modification d'une demande de Parc d'instances Spot

Vous pouvez modifier une demande de Parc d'instances Spot active pour effectuer les tâches suivantes :

- Augmenter la capacité cible et la partie à la demande
- Diminuer la capacité cible et la partie à la demande

> **Note**
>
> Vous ne pouvez pas modifier une demande unique de Parc d'instances Spot. Vous pouvez uniquement modifier une demande Parc d'instances Spot si vous avez sélectionné Maintain target capacity (Maintenir la capacité cible) au moment de la création de la demande Parc d'instances Spot.

Lorsque vous augmentez la capacité cible, le Parc d'instances Spot lance des Instances Spot supplémentaires. Lorsque vous augmentez la partie à la demande, le Parc d'instances Spot lance des Instances à la demande supplémentaires.

Lorsque vous augmentez la capacité cible, le Parc d'instances Spot lance les Instances Spot supplémentaires en fonction de la stratégie d'allocation de sa demande de Parc d'instances Spot. Si la stratégie d'allocation est `lowestPrice`, le Parc d'instances Spot lance les instances du groupe d'instance Spot offrant le tarif le moins élevé de la demande de Parc d'instances Spot. En revanche, si la stratégie d'allocation est `diversified`, le Parc d'instances Spot répartit les instances entre les groupes de la demande de Parc d'instances Spot.

Lorsque vous diminuez la capacité cible, le Parc d'instances Spot annule les demandes ouvertes qui dépassent la nouvelle capacité cible. Vous pouvez demander à ce que le Parc d'instances Spot résilie les Instances Spot tant que la taille du parc n'atteint pas la nouvelle capacité cible. Si la stratégie d'allocation est `lowestPrice`, le Parc d'instances Spot résilie les instances ayant le prix par unité le plus élevé. Si la stratégie d'allocation est `diversified`, le Parc d'instances Spot résilie les instances de tous les groupes. Vous pouvez aussi demander à ce que le Parc d'instances Spot conserve sa taille actuelle, mais sans remplacer les Instances Spot interrompues ou que vous résiliez manuellement.

Lorsqu'un Parc d'instances Spot résilie une instance du fait de la diminution de la capacité cible, l'instance reçoit un avis d'interruption d'instance Spot.

Pour modifier une demande de Parc d'instances Spot (console)

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot/home/fleet.
2. Sélectionnez votre demande de Parc d'instances Spot.
3. Choisissez Actions, Modify target capacity (Modifier la capacité cible).
4. Dans Modify target capacity (Modifier la capacité cible), effectuez les opérations suivantes :
   a. Entrez la nouvelle capacité cible et la partie à la demande.
   b. (Facultatif) Si vous diminuez la capacité cible, mais que vous souhaitez conserver la taille actuelle du parc, décochez la case Terminate instances (Résilier les instances).
   c. Choisissez Submit.

Pour modifier une demande de Parc d'instances Spot à partir de l'AWS CLI

Utilisez la commande modify-spot-fleet-request pour mettre à jour la capacité cible de la demande de Parc d'instances Spot spécifiée.

```
aws ec2 modify-spot-fleet-request \
    --spot-fleet-request-id sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
```

```
    --target-capacity 20
```

Vous pouvez modifier la commande précédente comme suit de façon à diminuer la capacité cible du Parc d'instances Spot spécifié sans que cela n'ait pour effet de résilier les Instances Spot.

```
aws ec2 modify-spot-fleet-request \
    --spot-fleet-request-id sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
    --target-capacity 10 \
    --excess-capacity-termination-policy NoTermination
```

Annulation d'une demande de Parc d'instances Spot

Lorsque vous n'avez plus besoin d'utiliser votre Parc d'instances Spot, vous pouvez annuler la demande de Parc d'instances Spot. Cette action annule toutes les demandes Spot associées au Parc d'instances Spot, si bien qu'aucune nouvelle instance Spot n'est lancée pour votre Parc d'instances Spot. Vous devez indiquer si le Parc d'instances Spot doit résilier ses Instances Spot. Si vous résiliez les instances, la demande de Parc d'instances Spot a l'état `cancelled_terminating`. Sinon, la demande de parc d'instances Spot passe à l'état `cancelled_running` et les instances continuent à être exécutées jusqu'à ce qu'elles soient interrompues ou jusqu'à ce que vous les mettiez hors service manuellement.

Pour annuler une demande de Parc d'instances Spot (console)

1. Ouvrez la console des instances Spot à l'adresse https://console.aws.amazon.com/ec2spot/home/fleet.
2. Sélectionnez votre demande de Parc d'instances Spot.
3. Choisissez Actions, Cancel spot request (Annuler la demande d'instance Spot).
4. Dans Cancel spot request (Annuler la demande Spot), confirmez que vous souhaitez annuler le Parc d'instances Spot. Pour conserver la taille actuelle du parc, décochez la case Terminate instances (Résilier les instances). Lorsque vous êtes prêt, sélectionnez Confirm (Confirmer).

Pour annuler une demande de Parc d'instances Spot à partir de l'AWS CLI

Utilisez la commande cancel-spot-fleet-requests pour annuler la demande de Parc d'instances Spot spécifiée et résilier les instances.

```
aws ec2 cancel-spot-fleet-requests \
    --spot-fleet-request-ids sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
    --terminate-instances
```

Voici un exemple de sortie :

```
{
    "SuccessfulFleetRequests": [
        {
            "SpotFleetRequestId": "sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE",
            "CurrentSpotFleetRequestState": "cancelled_terminating",
            "PreviousSpotFleetRequestState": "active"
        }
    ],
    "UnsuccessfulFleetRequests": []
}
```

Vous pouvez modifier la commande précédente comme suit afin d'annuler la demande de Parc d'instances Spot spécifiée sans résilier les instances.

```
aws ec2 cancel-spot-fleet-requests \
    --spot-fleet-request-ids sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
    --no-terminate-instances
```

Voici un exemple de sortie :

```
{
    "SuccessfulFleetRequests": [
        {
            "SpotFleetRequestId": "sfr-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE",
            "CurrentSpotFleetRequestState": "cancelled_running",
            "PreviousSpotFleetRequestState": "active"
        }
    ],
    "UnsuccessfulFleetRequests": []
}
```

Exemples de configuration d'un Parc d'instances Spot

Les exemples suivants montrent les configurations de lancement que vous pouvez utiliser avec la commande request-spot-fleet afin de créer une demande de Parc d'instances Spot. Pour plus d'informations, consultez Création d'une demande de Parc d'instances Spot (p. 416).

> **Note**
>
> Pour Parc d'instances Spot, vous ne pouvez pas spécifier d'ID d'interface réseau dans une spécification de lancement. Veillez à omettre le paramètre `NetworkInterfaceID` dans votre spécification de lancement.

1. Lancement d'Instances Spot en utilisant la zone de disponibilité ou le sous-réseau offrant le prix le moins élevé de la région (p. 429)
2. Lancement d'Instances Spot en utilisant la zone de disponibilité ou le sous-réseau offrant le prix le moins élevé dans une liste spécifiée (p. 430)
3. Lancement d'Instances Spot en utilisant le type d'instance offrant le prix le moins élevé dans une liste spécifiée (p. 431)
4. Remplacement du prix pour la demande (p. 432)
5. Lancement d'un Parc d'instances Spot en utilisant la stratégie d'allocation diversifiée (p. 433)
6. Lancement d'un Parc d'instances Spot en utilisant la pondération d'instance (p. 435)
7. Lancement d'un Parc d'instances Spot avec une capacité à la demande (p. 436)

Exemple 1 : Lancement d'Instances Spot en utilisant la zone de disponibilité ou le sous-réseau offrant le prix le moins élevé de la région

L'exemple suivant spécifie une seule spécification de lancement sans Zone de disponibilité ou sous-réseau. Le Parc d'instances Spot lance les instances dans la zone de disponibilité ayant le prix le moins élevé qui a un sous-réseau par défaut. Le prix que vous payez ne dépasse pas le prix à la demande.

```
{
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "KeyName": "my-key-pair",
          "SecurityGroups": [
              {
                  "GroupId": "sg-1a2b3c4d"
              }
          ],
          "InstanceType": "m3.medium",
          "IamInstanceProfile": {
              "Arn": "arn:aws:iam::123456789012:instance-profile/my-iam-role"
          }
```

```
        }
    ]
}
```

Exemple 2 : Lancement d'Instances Spot en utilisant la zone de disponibilité ou le sous-réseau offrant le prix le moins élevé dans une liste spécifiée

Les exemples suivants spécifient deux spécifications de lancement avec différents sous-réseaux ou zones de disponibilité, mais avec les mêmes types d'instance et AMI.

Zones de disponibilité

Le Parc d'instances Spot lance les instances dans le sous-réseau par défaut de la zone de disponibilité ayant le prix le moins élevé que vous avez spécifié.

```
{
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "KeyName": "my-key-pair",
          "SecurityGroups": [
              {
                  "GroupId": "sg-1a2b3c4d"
              }
          ],
          "InstanceType": "m3.medium",
          "Placement": {
              "AvailabilityZone": "us-west-2a, us-west-2b"
          },
          "IamInstanceProfile": {
              "Arn": "arn:aws:iam::123456789012:instance-profile/my-iam-role"
          }
      }
  ]
}
```

Sous-réseaux

Vous pouvez spécifier des sous-réseaux par défaut ou personnalisés, les derniers pouvant être issus d'un VPC par défaut ou personnalisé. Le service d'instances Spot lance les instances sur n'importe quel réseau se trouvant dans la zone de disponibilité ayant le prix le moins élevé.

Vous ne pouvez pas spécifier plusieurs sous-réseaux d'une même zone de disponibilité dans une demande de Parc d'instances Spot.

```
{
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "KeyName": "my-key-pair",
          "SecurityGroups": [
              {
                  "GroupId": "sg-1a2b3c4d"
              }
          ],
          "InstanceType": "m3.medium",
          "SubnetId": "subnet-a61dafcf, subnet-65ea5f08",
          "IamInstanceProfile": {
              "Arn": "arn:aws:iam::123456789012:instance-profile/my-iam-role"
```

```
            }
        }
    ]
}
```

Si les instances sont lancées sur un VPC par défaut, elles reçoivent une adresse IPv4 publique par défaut. Si les instances sont lancées sur un VPC personnalisé, elles ne reçoivent pas d'adresse IPv4 publique par défaut. Utilisez une interface réseau dans la spécification de lancement afin d'attribuer une adresse IPv4 publique aux instances lancées dans un VPC personnalisé. Lorsque vous spécifiez une interface réseau, vous devez inclure l'ID de sous-réseau et l'ID du groupe de sécurité à l'aide de l'interface réseau.

```
  ...
      {
          "ImageId": "ami-1a2b3c4d",
          "KeyName": "my-key-pair",
          "InstanceType": "m3.medium",
          "NetworkInterfaces": [
              {
                  "DeviceIndex": 0,
                  "SubnetId": "subnet-1a2b3c4d",
                  "Groups": [ "sg-1a2b3c4d" ],
                  "AssociatePublicIpAddress": true
              }
          ],
          "IamInstanceProfile": {
              "Arn": "arn:aws:iam::880185128111:instance-profile/my-iam-role"
          }
      }
  ...
```

Exemple 3 : Lancement d'Instances Spot en utilisant le type d'instance offrant le prix le plus bas dans une liste spécifiée

Les exemples suivants spécifient deux configurations de lancement avec différents types d'instance, mais la même AMI et la même zone de disponibilité ou le même sous-réseau. Le Parc d'instances Spot lance les instances en utilisant le type d'instance spécifié offrant le prix le plus bas.

Zone de disponibilité

```
{
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "SecurityGroups": [
              {
                  "GroupId": "sg-1a2b3c4d"
              }
          ],
          "InstanceType": "cc2.8xlarge",
          "Placement": {
            "AvailabilityZone": "us-west-2b"
          }
      },
      {
          "ImageId": "ami-1a2b3c4d",
          "SecurityGroups": [
              {
                  "GroupId": "sg-1a2b3c4d"
              }
          ],
```

```
            "InstanceType": "r3.8xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2b"
            }
        }
    ]
}
```

Sous-réseau

```
{
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "SecurityGroups": [
              {
                  "GroupId": "sg-1a2b3c4d"
              }
          ],
          "InstanceType": "cc2.8xlarge",
          "SubnetId": "subnet-1a2b3c4d"
      },
      {
          "ImageId": "ami-1a2b3c4d",
          "SecurityGroups": [
              {
                  "GroupId": "sg-1a2b3c4d"
              }
          ],
          "InstanceType": "r3.8xlarge",
          "SubnetId": "subnet-1a2b3c4d"
      }
  ]
}
```

Exemple 4 : Remplacement du prix pour la demande

Nous vous avons recommandé d'utiliser le prix maximum par défaut, qui correspond au prix à la demande. Si vous préférez, vous pouvez indiquer un prix maximum pour la demande du parc, et les prix maximum des spécifications de lancement individuelles.

Les exemples suivants indiquent le prix maximum pour la demande du parc, et les prix maximum pour deux des trois spécifications de lancement. Le prix maximum de la demande de parc est utilisé pour toutes les spécifications de lancement qui ne spécifient aucun prix maximum. Le Parc d'instances Spot lance les instances en utilisant le type d'instance offrant le prix le plus bas.

Zone de disponibilité

```
{
  "SpotPrice": "1.00",
  "TargetCapacity": 30,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "c3.2xlarge",
          "Placement": {
              "AvailabilityZone": "us-west-2b"
          },
          "SpotPrice": "0.10"
      },
```

```
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "c3.4xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2b"
            },
            "SpotPrice": "0.20"
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "c3.8xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2b"
            }
        }
    ]
}
```

Sous-réseau

```
{
  "SpotPrice": "1.00",
  "TargetCapacity": 30,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "c3.2xlarge",
          "SubnetId": "subnet-1a2b3c4d",
          "SpotPrice": "0.10"
      },
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "c3.4xlarge",
          "SubnetId": "subnet-1a2b3c4d",
          "SpotPrice": "0.20"
      },
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "c3.8xlarge",
          "SubnetId": "subnet-1a2b3c4d"
      }
  ]
}
```

Exemple 5 : Lancement d'un Parc d'instances Spot en utilisant la stratégie d'allocation diversifiée

L'exemple suivant utilise la stratégie d'attribution `diversified`. Les spécifications de lancement ont différents types d'instance, mais la même AMI et la même zone de disponibilité ou le même sous-réseau. Le Parc d'instances Spot répartit les 30 instances entre les trois spécifications de lancement de sorte qu'il existe 10 instances de chaque type. Pour plus d'informations, consultez Stratégie d'allocation pour les Instances Spot (p. 381).

Zone de disponibilité

```
{
  "SpotPrice": "0.70",
  "TargetCapacity": 30,
  "AllocationStrategy": "diversified",
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
```

```
            "InstanceType": "c4.2xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2b"
            }
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "m3.2xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2b"
            }
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "r3.2xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2b"
            }
        }
    ]
}
```

Sous-réseau

```
{
    "SpotPrice": "0.70",
    "TargetCapacity": 30,
    "AllocationStrategy": "diversified",
    "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
    "LaunchSpecifications": [
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "c4.2xlarge",
            "SubnetId": "subnet-1a2b3c4d"
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "m3.2xlarge",
            "SubnetId": "subnet-1a2b3c4d"
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "r3.2xlarge",
            "SubnetId": "subnet-1a2b3c4d"
        }
    ]
}
```

Une bonne pratique pour augmenter les chances qu'une demande au comptant puisse être satisfaite par une capacité EC2 en cas de panne dans l'une des zones de disponibilité est de diversifier ces dernières. Pour ce scénario, incluez chaque zone disponibilité à votre disposition dans les spécifications de lancement. Et, au lieu d'utiliser le même sous-réseau à chaque fois, utilisez trois sous-réseaux uniques (chacun correspondant à une zone différente).

Zone de disponibilité

```
{
  "SpotPrice": "0.70",
  "TargetCapacity": 30,
  "AllocationStrategy": "diversified",
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
```

```
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "c4.2xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2a"
            }
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "m3.2xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2b"
            }
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "r3.2xlarge",
            "Placement": {
                "AvailabilityZone": "us-west-2c"
            }
        }
    ]
}
```

Sous-réseau

```
{
    "SpotPrice": "0.70",
    "TargetCapacity": 30,
    "AllocationStrategy": "diversified",
    "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
    "LaunchSpecifications": [
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "c4.2xlarge",
            "SubnetId": "subnet-1a2b3c4d"
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "m3.2xlarge",
            "SubnetId": "subnet-2a2b3c4d"
        },
        {
            "ImageId": "ami-1a2b3c4d",
            "InstanceType": "r3.2xlarge",
            "SubnetId": "subnet-3a2b3c4d"
        }
    ]
}
```

Exemple 6 : Lancement d'un Parc d'instances Spot en utilisant la pondération d'instance

Les exemples suivants utilisent la pondération d'instance, ce qui signifie que le prix est déterminé par heure d'unité, et non par heure d'instance. Chaque configuration de lancement répertorie un type d'instance différent et une pondération différente. Le Parc d'instances Spot sélectionne le type d'instance ayant le prix par heure d'unité le plus bas. Le Parc d'instances Spot calcule le nombre d'Instances Spot à lancer en divisant la capacité cible par la pondération d'instance. Si le résultat n'est pas un entier, le Parc d'instances Spot l'arrondit à l'entier suivant afin que la taille de votre parc ne soit pas inférieure à sa capacité cible.

Si la demande `r3.2xlarge` est satisfaite, le parc d'instances Spot met en service 4 de ces instances. Divisez 20 par 6 pour un total de 3,33 instances, puis arrondissez à 4 instances.

Si la demande `c3.xlarge` est satisfaite, le parc d'instances Spot met en service 7 de ces instances. Divisez 20 par 3 pour un total de 6,66 instances, puis arrondissez à 7 instances.

Pour plus d'informations, consultez Pondération d'instance de Parc d'instances Spot (p. 384).

Zone de disponibilité

```
{
  "SpotPrice": "0.70",
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "r3.2xlarge",
          "Placement": {
              "AvailabilityZone": "us-west-2b"
          },
          "WeightedCapacity": 6
      },
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "c3.xlarge",
          "Placement": {
              "AvailabilityZone": "us-west-2b"
          },
          "WeightedCapacity": 3
      }
    ]
}
```

Sous-réseau

```
{
  "SpotPrice": "0.70",
  "TargetCapacity": 20,
  "IamFleetRole": "arn:aws:iam::123456789012:role/aws-ec2-spot-fleet-tagging-role",
  "LaunchSpecifications": [
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "r3.2xlarge",
          "SubnetId": "subnet-1a2b3c4d",
          "WeightedCapacity": 6
      },
      {
          "ImageId": "ami-1a2b3c4d",
          "InstanceType": "c3.xlarge",
          "SubnetId": "subnet-1a2b3c4d",
          "WeightedCapacity": 3
      }
  ]
}
```

Exemple 7 : Lancement d'un Parc d'instances Spot avec une capacité à la demande

Pour garantir que vous avez toujours la capacité d'instance, vous pouvez inclure une capacité à la demande dans votre demande de Parc d'instances Spot. S'il y a la capacité nécessaire, la demande à la demande est toujours satisfaite. Le solde de la capacité cible est assuré en tant que Spot s'il existe une capacité et une disponibilité.

L'exemple suivant spécifie la capacité cible souhaitée de 10 instances, dont 5 correspondent à une capacité à la demande. La capacité Spot n'est pas spécifiée : elle est impliquée dans le solde de la capacité cible moins la capacité à la demande. Amazon EC2 lance 5 unités de capacité à la demande et 5 unités de capacité (10 - 5 = 5) Spot s'il existe une capacité et une disponibilité Amazon EC2.

Pour plus d'informations, consultez À la demande dans Parc d'instances Spot (p. 381).

```
{
  "IamFleetRole": "arn:aws:iam::781603563322:role/aws-ec2-spot-fleet-tagging-role",
  "AllocationStrategy": "lowestPrice",
  "TargetCapacity": 10,
  "SpotPrice": null,
  "ValidFrom": "2018-04-04T15:58:13Z",
  "ValidUntil": "2019-04-04T15:58:13Z",
  "TerminateInstancesWithExpiration": true,
  "LaunchSpecifications": [],
  "Type": "maintain",
  "OnDemandTargetCapacity": 5,
  "LaunchTemplateConfigs": [
    {
      "LaunchTemplateSpecification": {
        "LaunchTemplateId": "lt-0dbb04d4a6cca5ad1",
        "Version": "2"
      },
      "Overrides": [
        {
          "InstanceType": "t2.medium",
          "WeightedCapacity": 1,
          "SubnetId": "subnet-d0dc51fb"
        }
      ]
    }
  ]
}
```

Métriques CloudWatch pour un Parc d'instances Spot

Amazon EC2 fournit des métriques Amazon CloudWatch qui permettent de surveiller un Parc d'instances Spot.

> **Important**
>
> Pour garantir la précision des informations, nous vous recommandons d'activer la surveillance détaillée lorsque vous utilisez ces métriques. Pour plus d'informations, consultez Activer ou désactiver la surveillance détaillée pour vos instances (p. 762).

Pour plus d'informations sur les métriques CloudWatch fournies par Amazon EC2, consultez Surveillance de vos instances avec CloudWatch (p. 761).

Métriques pour un Parc d'instances Spot

L'espace de noms `AWS/EC2Spot` inclut les métriques suivantes, ainsi que les métriques CloudWatch pour les Instances Spot de votre parc. Pour plus d'informations, consultez Métriques des instances (p. 764).

Métrique	Description
`AvailableInstancePoolsCount`	Groupes d'instance Spot spécifiés dans la demande de Parc d'instances Spot. Unités : nombre
`BidsSubmittedForCapacity`	Capacité pour laquelle Amazon EC2 a soumis des demandes de Parc d'instances Spot. Unités : nombre
`EligibleInstancePoolCount`	Groupes d'instance Spot spécifiés dans la demande de Parc d'instances Spot où Amazon EC2 peut répondre aux demandes.

Métrique	Description
	Amazon EC2 ne répond pas aux demandes dans les groupes où le prix maximum que vous acceptez de payer pour des Instances Spot est inférieur au prix Spot ou lorsque le prix Spot est supérieur au prix des Instances à la demande. Unités : nombre
`FulfilledCapacity`	Capacité exécutée par Amazon EC2. Unités : nombre
`MaxPercentCapacityAllocation`	Valeur maximale de `PercentCapacityAllocation` pour tous les groupes de Parc d'instances Spot spécifiés dans la demande de Parc d'instances Spot. Unités : pourcentage
`PendingCapacity`	Différence entre `TargetCapacity` et `FulfilledCapacity`. Unités : nombre
`PercentCapacityAllocation`	Capacité allouée pour le groupe d'instance Spot pour les dimensions spécifiées. Pour obtenir la valeur maximale enregistrée pour tous les groupes d'instance Spot, utilisez `MaxPercentCapacityAllocation`. Unités : pourcentage
`TargetCapacity`	Capacité cible de la demande de Parc d'instances Spot. Unités : nombre
`TerminatingCapacity`	Capacité résiliée car la capacité allouée est supérieure à la capacité cible. Unités : nombre

Si l'unité de mesure d'une métrique est `Count`, la statistique la plus utile est`Average`.

Dimensions d'un Parc d'instances Spot

Pour filtrer les données de votre Parc d'instances Spot, utilisez les dimensions suivantes.

Dimensions	Description
`AvailabilityZone`	Filtrer les données par Zone de disponibilité.
`FleetRequestId`	Filtrer les données demande de parc d'instances Spot.
`InstanceType`	Filtrer les données par type d'instance.

Affichage des métriques CloudWatch pour votre Parc d'instances Spot

Vous pouvez afficher les métriques CloudWatch pour votre Parc d'instances Spot à partir de la console Amazon CloudWatch. Ces métriques s'affichent sous forme de graphiques de surveillance. Ces graphiques affichent des points de données si le Parc d'instances Spot est actif.

Les métriques sont d'abord regroupées par espace de noms, puis par les différentes combinaisons de dimensions au sein de chaque espace de noms. Par exemple, vous pouvez afficher toutes les métriques de Parc d'instances Spot ou des groupes de métriques de Parc d'instances Spot par ID de demande de Parc d'instances Spot, type d'instance ou zone de disponibilité.

Pour afficher les métriques de Parc d'instances Spot

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, choisissez Métriques.
3. Choisissez l'espace de noms EC2 Spot.

   Note

   Si l'espace de noms EC2 Spot n'est pas affiché, c'est pour deux raisons. Soit vous n'avez pas encore utilisé Parc d'instances Spot — seuls les services AWS que vous utilisez envoient des métriques à Amazon CloudWatch. Soit, si vous n'avez pas utilisé Parc d'instances Spot au cours des deux dernières semaines, l'espace de noms n'apparaît pas.
4. (Facultatif) Pour filtrer les métriques par dimension, sélectionnez l'une des options suivantes :
   - Métriques de demande de parc : regroupement par demande de Parc d'instances Spot
   - Par zone de disponibilité : regroupement par demande de Parc d'instances Spot et zone de disponibilité
   - Par type d'instance : regroupement par demande de Parc d'instances Spot et type d'instance
   - Par zone de disponibilité/Type d'instance : regroupement par demande de Parc d'instances Spot, zone de disponibilité et type d'instance
5. Pour afficher les données d'une métrique, cochez la case en regard de la métrique.

Dimensionnement automatique pour un Parc d'instances Spot

La mise à l'échelle automatique est l'aptitude à augmenter ou à diminuer automatiquement la capacité cible de votre Parc d'instances Spot en fonction de la demande. Un Parc d'instances Spot peut lancer des instances (montée en charge) ou résilier des instances (diminution en charge), dans la plage que vous choisissez, en réponse à une ou plusieurs stratégies de dimensionnement.

Le Parc d'instances Spot prend en charge les types de mise à l'échelle automatique suivants :

- Target tracking scaling (p. 441) – Augmente ou réduit la capacité actuelle du parc en fonction d'une valeur cible pour une métrique spécifique. Cette option est similaire à la façon dont votre thermostat maintient la température de votre domicile : vous sélectionnez une température et le thermostat se charge du reste.

- Step scaling (p. 442) – Augmente ou réduit la capacité actuelle du parc en fonction d'un ensemble d'ajustements du dimensionnement, appelés ajustements d'étape, qui varient en fonction de la valeur du seuil de l'alarme.
- Scheduled scaling (p. 445) – Augmente ou réduit la capacité actuelle du parc en fonction de la date et de l'heure.

Si vous utilisez une pondération d'instance (p. 384), gardez à l'esprit que le Parc d'instances Spot peut dépasser la capacité cible si nécessaire. La capacité fournie peut correspondre à un nombre à virgule flottante, mais la capacité cible doit être un nombre entier pour que le Parc d'instances Spot puisse l'arrondir au nombre entier suivant. Vous devez prendre ces comportements en compte lorsque vous examinez les résultats d'une stratégie de dimensionnement lorsqu'une alarme se déclenche. Par exemple, supposons que la capacité cible est 30, que la capacité fournie est 30,1 et que la stratégie de dimensionnement soustrait 1. Lorsque l'alarme se déclenche, le processus de scalabilité automatique soustrait 1 de 30,1 pour obtenir 29,1, puis arrondit la valeur à 30. Aucune action de mise à l'échelle n'est alors effectuée. Pour prendre un autre exemple, supposons que vous avez sélectionné des pondérations d'instance de 2, 4 et 8, et une capacité cible de 10, mais qu'aucune instance de pondération 2 n'était disponible, si bien que le Parc d'instances Spot a provisionné des instances de pondération 4 et 8 pour une capacité fournie de 12. Si la stratégie de mise à l'échelle réduit la capacité cible de 20 % et qu'une alarme se déclenche, le processus de scalabilité automatique soustrait 12 x 0,2 de 12 pour obtenir 9,6, puis arrondit la valeur à 10. Aucune action de mise à l'échelle n'est alors effectuée.

Les stratégies de mise à l'échelle que vous créez pour Parc d'instances Spot prennent en charge un temps de stabilisation. C'est le nombre de secondes après la fin d'une activité de dimensionnement au cours desquelles les activités de dimensionnement précédentes liées à un déclencheur peuvent influer sur les événements de dimensionnement futurs. Pour les stratégies de montée en charge (scale-out), pendant la durée du temps de stabilisation, la capacité qui a été ajoutée par l'événement de montée en charge précédent qui a lancé la stabilisation est calculée dans le cadre de la capacité souhaitée pour la montée en charge suivante. L'objectif est d'effectuer une montée en charge continue (mais pas excessive). Pour les stratégies de diminution de charge, la période de récupération est utilisée pour bloquer les demandes de montée en charge suivantes jusqu'à leur expiration. L'objectif est de diminuer la charge avec prudence afin de protéger la disponibilité de votre application. Toutefois, si une autre alarme déclenche une stratégie de montée en charge pendant le temps de stabilisation après une diminution en charge (scale-in), la scalabilité automatique monte immédiatement en charge votre cible scalable.

Nous vous recommandons de dimensionner sur des métriques d’instance à une fréquence de 1 minute, car cela permet de réagir plus rapidement aux modifications d'utilisation. Un dimensionnement sur des métriques à une fréquence de 5 minutes peut entraîner des temps de réponse plus lents et un dimensionnement sur des données de métrique obsolètes. Pour envoyer les données des métriques de votre instance à CloudWatch toutes les minutes, vous pouvez activer la surveillance détaillée sur l'instance. Pour plus d'informations, consultez Activer ou désactiver la surveillance détaillée pour vos instances (p. 762) et Création d'une demande de Parc d'instances Spot à l'aide des paramètres définis (console) (p. 417).

Pour plus d'informations sur la configuration du dimensionnement pour Parc d'instances Spot, consultez les ressources suivantes :

- Sectionapplication-autoscaling du AWS CLI Command Reference
- Référence d'API Application Auto Scaling
- Guide de l'utilisateur Application Auto Scaling

Autorisations IAM requises pour le dimensionnement automatique d'un Parc d'instances Spot

Le dimensionnement automatique pour Parc d'instances Spot est rendu possible en combinant Amazon EC2, Amazon CloudWatch et les API Application Auto Scaling. Les demandes de Parc d'instances Spot

sont créées avec Amazon EC2, les alarmes sont créées avec CloudWatch et les stratégies de mise à l'échelle sont créées avec Application Auto Scaling.

En plus des autorisations IAM pour le Parc d'instances Spot (p. 412) et Amazon EC2, l'utilisateur IAM qui accède aux paramètres de mise à l'échelle du parc doit avoir les autorisations appropriées pour les services qui prennent en charge la mise à l'échelle dynamique. Les utilisateurs IAM doivent avoir les autorisations d'utiliser les actions dans l'exemple de stratégie suivant.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "application-autoscaling:*",
                "ec2:DescribeSpotFleetRequests",
                "ec2:ModifySpotFleetRequest",
                "cloudwatch:DeleteAlarms",
                "cloudwatch:DescribeAlarmHistory",
                "cloudwatch:DescribeAlarms",
                "cloudwatch:DescribeAlarmsForMetric",
                "cloudwatch:GetMetricStatistics",
                "cloudwatch:ListMetrics",
                "cloudwatch:PutMetricAlarm",
                "cloudwatch:DisableAlarmActions",
                "cloudwatch:EnableAlarmActions",
                "iam:CreateServiceLinkedRole",
                "sns:CreateTopic",
                "sns:Subscribe",
                "sns:Get*",
                "sns:List*"
            ],
            "Resource": "*"
        }
    ]
}
```

Vous pouvez également créer vos propres stratégies IAM qui permettent de créer des autorisations plus détaillées pour les appels vers l'API Application Auto Scaling. Pour de plus amples informations, veuillez consulter Authentification et contrôle d'accès dans le Guide de l'utilisateur Application Auto Scaling.

Le service Application Auto Scaling nécessite également l'autorisation de décrire votre Parc d'instances Spot et vos alarmes CloudWatch, ainsi que des autorisations lui permettant de modifier votre capacité Parc d'instances Spot cible en votre nom. Si vous activez le dimensionnement automatique pour votre Parc d'instances Spot, il crée un rôle lié à un service nommé `AWSServiceRoleForApplicationAutoScaling_EC2SpotFleetRequest`. Le rôle lié à un service donne à Application Auto Scaling l'autorisation de décrire les alarmes de vos stratégies, de surveiller la capacité actuelle du parc et éventuellement de la modifier. Le rôle de parc d'instances Spot géré original pour Application Auto Scaling était `aws-ec2-spot-fleet-autoscale-role`, mais il n'est plus nécessaire. Le rôle lié à un service est le rôle par défaut pour Application Auto Scaling. Pour de plus amples informations, veuillez consulter Rôles liés à un service dans le Guide de l'utilisateur Application Auto Scaling.

Mise à l'échelle d'un Parc d'instances Spot en utilisant une stratégie Suivi de la cible

Grâce aux stratégies de mise à l'échelle Suivi de la cible, vous sélectionnez une métrique et définissez une valeur cible. Le Parc d'instances Spot crée et gère les alarmes CloudWatch qui déclenchent la stratégie de mise à l'échelle, et calcule l'ajustement de la mise à l'échelle en fonction de la métrique et de la valeur cible. La stratégie de dimensionnement ajoute ou supprime de la capacité si nécessaire pour maintenir la métrique à la valeur cible spécifiée ou proche de celle-ci. En plus de maintenir la métrique proche de la

valeur cible, une stratégie de dimensionnement Suivi de la cible s'ajuste également aux fluctuations de la métrique dues à un modèle de charge fluctuant, et minimise les fluctuations rapides dans la capacité du parc.

Vous pouvez créer plusieurs stratégies de dimensionnement Suivi de la cible pour un Parc d'instances Spot dans la mesure où chacune d'elles utilise une métrique différente. Le parc est dimensionné selon la stratégie qui fournit la plus grande capacité de parc. Cela vous permet de couvrir plusieurs scénarios et de toujours disposer d'une capacité suffisante pour traiter vos charges de travail d'application.

Pour garantir la disponibilité de l'application, le parc augmente proportionnellement aux métriques aussi rapidement que possible, mais diminue plus progressivement.

Lorsqu'un Parc d'instances Spot résilie une instance du fait de la diminution de la capacité cible, l'instance reçoit un avis d'interruption d'instance Spot.

Évitez de modifier ou de supprimer les alarmes CloudWatch que le Parc d'instances Spot gère pour une stratégie de mise à l'échelle Suivi de la cible. Le Parc d'instances Spot supprime les alarmes automatiquement lorsque vous supprimez la stratégie de mise à l'échelle Suivi de la cible.

Limitation

- La demande de Parc d'instances Spot doit être de type `maintain`. Le dimensionnement automatique n'est pas pris en charge pour les demandes uniques et les blocs d'instances Spot.

Pour configurer une stratégie de suivi de cible (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez votre demande de Parc d'instances Spot et choisissez Auto Scaling.
4. Si la mise à l'échelle automatique n'est pas configurée, sélectionnez Configurer.
5. Utilisez le champ Scale capacity between (Mettre à l'échelle la capacité entre) pour définir les capacités minimale et maximale de votre parc. Avec le dimensionnement automatique, votre parc n'aura jamais une capacité inférieure ou supérieure aux limites fixées.
6. Dans le champ Policy Name (Nom de stratégie), attribuez un nom à la stratégie.
7. Choisissez une valeur Target Metric (Métrique cible).
8. Spécifiez une valeur Target Value (Valeur cible) pour la métrique.
9. (Facultatif) Pour modifier le temps de stabilisation par défaut, définissez la valeur Cooldown Period (Temps de stabilisation).
10. (Facultatif) Sélectionnez Disable Scale-in (Désactiver la diminution en charge) pour ignorer la création d'une stratégie de diminution en charge sur la base de la configuration en cours. Vous pouvez créer une stratégie d'ajustement à la baisse à l'aide d'une autre configuration.
11. Choisissez Save.

Pour configurer une stratégie de dimensionnement Suivi de la cible à l'aide de l'AWS CLI

1. Inscrivez la demande de Parc d'instances Spot en tant que cible scalable à l'aide de la commande register-scalable-target.
2. Créez une stratégie de mise à l'échelle à l'aide de la commande put-scaling-policy.

Mise à l'échelle d'un Parc d'instances Spot en utilisant des stratégies de dimensionnement d'étape

Ces stratégies permettent de dimensionner les alarmes CloudWatch pour déclencher le processus de dimensionnement. Par exemple, si vous souhaitez augmenter la capacité du parc quand l'utilisation de l'UC

atteint un niveau donné, créez une alarme en utilisant la métrique `CPUUtilization` fournie par Amazon EC2.

Lorsque vous créez une stratégie de dimensionnement d'étape, vous devez indiquer l'un des types d'ajustement suivants :

- Ajouter – Augmentez la capacité cible du parc selon un nombre donné d'unités de capacité ou un pourcentage de la capacité actuelle.
- Supprimer – Réduisez la capacité cible du parc selon un nombre donné d'unités de capacité ou un pourcentage de la capacité actuelle.
- Définir sur – Définissez la capacité cible du parc selon un nombre précis d'unités de capacité.

Lorsqu'une alarme se déclenche, le processus de scalabilité automatique calcule la nouvelle capacité cible d'après la capacité fournie et la stratégie de mise à l'échelle, puis met à jour la capacité cible en conséquence. Par exemple, supposons que la capacité cible et la capacité fournie sont égales à 10 et que la stratégie de dimensionnement ajoute 1. Lorsque l'alarme se déclenche, le processus de scalabilité automatique ajoute 1 à 10 pour obtenir 11, pour que le Parc d'instances Spot lance 1 instance.

Lorsqu'un Parc d'instances Spot résilie une instance du fait de la diminution de la capacité cible, l'instance reçoit un avis d'interruption d'instance Spot.

Limitation

- La demande de Parc d'instances Spot doit être de type `maintain`. Le dimensionnement automatique n'est pas pris en charge pour les demandes uniques et les blocs d'instances Spot.

Prérequis

- Identifiez les métriques CloudWatch importantes pour votre application. Vous pouvez créer des alarmes CloudWatch à partir des métriques fournies par AWS ou de vos propres métriques personnalisées.
- Pour les métriques AWS utilisées dans vos stratégies de dimensionnement, activez la collecte de métriques CloudWatch si le service qui les fournit ne le fait pas par défaut.

Pour créer une alarme CloudWatch

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, cliquez sur Alarms.
3. Sélectionnez Créer une alarme.
4. Sur la page Specify metric and conditions (Spécifier une métrique et des conditions), sélectionnez Sélectionner une métrique.
5. Sélectionnez EC2 Spot (Parc d'instances EC2 Spot), Métriques de demande de parc, sélectionnez une métrique (par exemple, CPUUtilization), puis Sélectionner une métrique.

   La page Specify metric and conditions (Spécifier les métriques et les conditions) apparaît, présentant un graphique et d'autres informations sur la métrique sélectionnée.
6. Sous Période, choisissez la période d'évaluation de l'alarme, par exemple, 1 minute. Lors de l'évaluation de l'alarme, chaque période est regroupée en un point de données.

   > **Note**
   >
   > Une période plus courte crée une alarme plus sensible.
7. Sous Conditions, définissez l'alarme en définissant la condition de seuil. Par exemple, vous pouvez définir un seuil pour déclencher l'alarme lorsque la valeur de la métrique est supérieure ou égale à 80 %.

8. Sous Additional configuration (Configuration supplémentaire), pour Datapoints to alarm (Points de données pour l'alarme), spécifiez le nombre de points de données (périodes d'évaluation) qui doivent être dans l'état ALARME pour déclencher l'alarme, par exemple, 1 sur 2. Cela crée une alarme qui passe à l'état ALARME si le seuil est dépassé par ce nombre de périodes consécutives. Pour plus d'informations, consultez Évaluation d'une alarme dans le Guide de l'utilisateur Amazon CloudWatch.
9. Pour Missing data treatment (Traitement des données manquantes), choisissez l'une des options (ou conservez la valeur par défaut Treat missing data as missing (Traiter les données manquantes comme manquantes). Pour plus d'informations, consultez Configuration de la manière dont les alarmes CloudWatch traitent les données manquantes dans le Guide de l'utilisateur Amazon CloudWatch.
10. Choisissez Suivant.
11. (Facultatif) Pour recevoir une notification d'un événement de mise à l'échelle, pour Notification, vous pouvez sélectionner ou créer la rubrique Amazon SNS que vous voulez utiliser pour recevoir des notifications. Sinon, vous pouvez supprimer la notification maintenant et en ajouter une plus tard si nécessaire.
12. Choisissez Suivant.
13. Sous Add a description (Ajouter une description), entrez un nom et une description pour l'alarme et choisissez Suivant.
14. Sélectionnez Créer une alarme.

Pour configurer une stratégie de mise à l'échelle par étapes pour votre Parc d'instances Spot (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez votre demande de Parc d'instances Spot et choisissez Auto Scaling.
4. Si la mise à l'échelle automatique n'est pas configurée, sélectionnez Configurer.
5. Utilisez le champ Scale capacity between (Mettre à l'échelle la capacité entre) pour définir les capacités minimale et maximale de votre parc. Avec le dimensionnement automatique, votre parc n'aura jamais une capacité inférieure ou supérieure aux limites fixées.
6. À l'origine, la section Scaling policies (Stratégies de mise à l'échelle) contient des stratégies nommées ScaleUp et ScaleDown. Vous pouvez compléter ces stratégies ou cliquer sur Remove policy (Supprimer la stratégie) pour les supprimer. Vous pouvez également choisir Add policy (Ajouter une stratégie).
7. Pour définir une stratégie, procédez comme suit :
   a. Dans le champ Policy Name (Nom de stratégie), attribuez un nom à la stratégie.
   b. Dans le champ Policy trigger (Déclencheur de stratégie), sélectionnez une alarme existante ou choisissez Create new alarm (Créer une alarme) pour ouvrir la console Amazon CloudWatch et créer une alarme.
   c. Pour Modify capacity (Modifier la capacité), sélectionnez le type d'ajustement de la mise à l'échelle, un nombre, puis une unité.
   d. (Facultatif) Pour effectuer une mise à l'échelle à étapes, cliquez sur Define steps (Définir des étapes). Par défaut, une stratégie d'ajout a une limite inférieure correspondant à l'infini négatif et une limite supérieure correspondant au seuil d'alarme. Par défaut, une stratégie de suppression a une limite inférieure correspondant au seuil d'alarme et une limite supérieure correspondant à l'infini positif. Pour ajouter une autre étape, cliquez sur Add step (Ajouter une étape).
   e. (Facultatif) Pour modifier la valeur par défaut concernant le temps de stabilisation, sélectionnez un nombre dans le champ Cooldown period (Temps de stabilisation).
8. Choisissez Save.

Pour configurer des stratégies de mise à l'échelle d'étape pour votre Parc d'instances Spot à partir de l'AWS CLI

1. Inscrivez la demande de Parc d'instances Spot en tant que cible scalable à l'aide de la commande register-scalable-target.
2. Créez une stratégie de mise à l'échelle à l'aide de la commande put-scaling-policy.
3. Créez une alarme qui déclenche la stratégie de mise à l'échelle à l'aide de la commande put-metric-alarm.

Mise à l'échelle d'un Parc d'instances Spot à l'aide de la mise à l'échelle planifiée

La mise à l'échelle en fonction d'une planification vous permet de mettre à l'échelle l'application en réponse aux changements de demande. Pour utiliser la mise à l'échelle planifiée, vous créez des actions planifiées, qui indiquent au Parc d'instances Spot d'effectuer des activités de mise à l'échelle à des heures spécifiques. Lorsque vous créez une action planifiée, vous spécifiez le Parc d'instances Spot, quand l'activité de mise à l'échelle doit avoir lieu, la capacité minimale et la capacité maximale. Vous pouvez créer des actions planifiées pour une mise à l'échelle unique ou selon une planification récurrente.

Limites liés aux s

- La demande de Parc d'instances Spot doit être de type `maintain`. Le dimensionnement automatique n'est pas pris en charge pour les demandes uniques et les blocs d'instances Spot.

Pour créer une action planifiée unique

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez votre demande de Parc d'instances Spot et choisissez Scheduled Scaling (Mise à l'échelle planifiée).
4. Choisissez Create Scheduled Action (Créer une action planifiée).
5. Pour Nom, spécifiez un nouveau nom pour l'action planifiée.
6. Saisissez une valeur pour Minimum capacity (Capacité minimum), Maximum capacity (Capacité maximum), ou les deux.
7. Pour Recurrence (Récurrence), choisissez Once (Une fois).
8. (Facultatif) Choisissez la date et l'heure pour Heure de début, Heure de fin, ou les deux.
9. Choisissez Submit.

Pour mettre à l'échelle selon un calendrier récurrent

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez votre demande de Parc d'instances Spot et choisissez Scheduled Scaling (Mise à l'échelle planifiée).
4. Pour Recurrence (Récurrence), choisissez un des calendriers prédéfinis (par exemple, Every day (Chaque jour)), ou choisissez Personnalisé et saisissez une expression CRON. Pour plus d'informations sur les expressions CRON prises en charge par la mise à l'échelle planifiée, consultez Expressions CRON dans le Guide de l'utilisateur Amazon CloudWatch Events.
5. (Facultatif) Choisissez la date et l'heure pour Heure de début, Heure de fin, ou les deux.
6. Choisissez Submit.

Pour modifier une action planifiée

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez votre demande de Parc d'instances Spot et choisissez Scheduled Scaling (Mise à l'échelle planifiée).
4. Sélectionnez l'action planifiée et choisissez Actions, Modifier.
5. Apportez les modifications nécessaires et choisissez Soumettre.

Pour supprimer une action planifiée

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Demandes Spot.
3. Sélectionnez votre demande de Parc d'instances Spot et choisissez Scheduled Scaling (Mise à l'échelle planifiée).
4. Sélectionnez l'action planifiée et choisissez Actions, Supprimer.
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Supprimer.

Pour gérer la mise à l'échelle planifiée à l'aide de l'AWS CLI

Utilisez les commandes suivantes :

- put-scheduled-action
- describe-scheduled-actions
- delete-scheduled-action

Statut des demandes Spot

Pour vous aider à assurer le suivi de vos demandes d'instance Spot et à planifier l'utilisation de vos Instances Spot, utilisez le statut de demande fourni par Amazon EC2. Par exemple, le statut de la demande peut indiquer pourquoi votre demande d'instance Spot n'a pas encore été satisfaite, ou répertorier les contraintes qui empêchent l'exécution de votre demande d'instance Spot.

À chaque étape du processus, c'est-à-dire au cours du cycle de vie d'une demande Spot, des événements spécifiques déterminent les états successifs de la demande.

Sommaire

- Cycle de vie d'une demande Spot (p. 446)
- Obtention d'informations sur le statut de la demande (p. 450)
- Codes de statut des demandes Spot (p. 450)

Cycle de vie d'une demande Spot

Le diagramme suivant illustre les étapes suivies par votre demande d'instance Spot au cours de son cycle de vie, de la soumission à la mise hors service. Chaque étape est représentée sous forme d'un nœud et le code de statut de chaque nœud décrit le statut de la demande Spot et de l'instance Spot.

Évaluation en attente

Dès que vous créez une demande d'instance Spot, celle-ci passe à l'état `pending-evaluation` à moins qu'un ou plusieurs paramètres de demande ne soient pas valides (`bad-parameters`).

Code d'état	État de la demande	État de l'instance
`pending-evaluation`	`open`	Non applicable
`bad-parameters`	`closed`	Non applicable

En attente

Si une ou plusieurs contraintes de demande sont valides mais ne peuvent pas encore être respectées ou s'il n'y a pas suffisamment de capacité, la demande se voit attribuer l'état En attente jusqu'à ce que les contraintes soient respectées. Les options de la demande ont un impact sur les possibilités d'exécution de la demande. Par exemple, si vous spécifiez un prix maximum inférieur au prix Spot actuel, votre demande conserve l'état En attente jusqu'à ce que le prix Spot passe en dessous du prix maximum. Si vous spécifiez un groupe de zone de disponibilité, la demande conserve l'état En attente jusqu'à ce que la contrainte de zone de disponibilité soit respectée.

En cas de panne de l'une des zones de disponibilité, il est possible que la capacité EC2 disponible pour les demandes d'instance Spot dans d'autres zones de disponibilité puisse être affectée.

Code d'état	État de la demande	État de l'instance
`capacity-not-available`	`open`	Non applicable
`price-too-low`	`open`	Non applicable
`not-scheduled-yet`	`open`	Non applicable
`launch-group-constraint`	`open`	Non applicable
`az-group-constraint`	`open`	Non applicable
`placement-group-constraint`	`open`	Non applicable
`constraint-not-fulfillable`	`open`	Non applicable

Fin de l'évaluation/exécution en attente

Votre demande d'instance Spot peut passer à l'état `terminal` si vous créez une demande valide uniquement pendant une durée spécifique et que cette durée arrive à expiration avant que votre demande atteigne la phase d'exécution en attente. Cela peut également se produire si vous annulez la demande ou si une erreur système se produit.

Code d'état	État de la demande	État de l'instance
`schedule-expired`	`cancelled`	Non applicable
`canceled-before-fulfillment*`	`cancelled`	Non applicable
`bad-parameters`	`failed`	Non applicable
`system-error`	`closed`	Non applicable

* Si vous annulez la demande.

Exécution en attente

Lorsque les contraintes que vous avez spécifiées (le cas échéant) sont respectées et si le prix maximum est égal ou supérieur au prix Spot actuel, votre demande Spot se voit attribuer l'état `pending-fulfillment`.

A ce stade, Amazon EC2 est prêt à mettre en service les instances que vous avez demandées. Si le processus s'arrête à ce stade, il a probablement été annulé par l'utilisateur avant le lancement d'une instance Spot. Cela peut aussi être dû à une erreur système inattendue.

Code d'état	État de la demande	État de l'instance
`pending-fulfillment`	`open`	Non applicable

Exécutée

Lorsque toutes les spécifications de vos Instances Spot sont respectées, votre demande d'instance Spot est exécutée. Amazon EC2 lance les Instances Spot, ce qui peut prendre quelques minutes. Si une

instance Spot est mise en veille prolongée ou arrêtée lorsqu'elle est interrompue, elle reste dans cet état jusqu'à ce que la demande puisse être de nouveau satisfaite ou qu'elle soit annulée.

Code d'état	État de la demande	État de l'instance
`fulfilled`	`active`	`pending` → `running`
`fulfilled`	`active`	`stopped` → `running`

Si vous arrêtez une instance Spot, votre demande Spot passe à l'état `instance-stopped-by-user` ou `marked-for-stop` jusqu'à ce que l'instance Spot puisse être redémarrée ou que la demande soit annulée.

Code d'état	État de la demande	État de l'instance
`marked-for-stop`	`active`	`stopping`
`instance-stopped-by-user`*	`disabled` ou `cancelled`**	`stopped`

* Une instance Spot passe à l'état `instance-stopped-by-user` si vous arrêtez l'instance ou si vous exécutez la commande shutdown à partir de l'instance. Une fois l'instance arrêtée, vous pouvez la redémarrer. Au redémarrage, la demande d'instance Spot revient à l'état `pending-evaluation`, puis Amazon EC2 lance une nouvelle instance Spot lorsque les contraintes sont remplies.

** L'état de la demande Spot est `disabled` si vous arrêtez l'instance Spot sans annuler la demande. L'état de la demande est `cancelled` si votre instance Sport est arrêtée et que la demande expire.

Fin d'exécution

Vos Instances Spot continuent de s'exécuter tant que le prix maximum est égal ou supérieur au prix spot, qu'il existe de la capacité pour votre type d'instance et que vous ne résiliez pas l'instance. Si, en raison d'une évolution du prix spot ou de la capacité disponible, Amazon EC2 doit résilier vos Instances Spot, la demande Spot se voit attribuer l'état terminal. Une demande se voit attribuer l'état terminal si vous annulez la demande Spot ou si vous résiliez les Instances Spot.

Code d'état	État de la demande	État de l'instance
`request-canceled-and-instance-running`	`cancelled`	`running`
`marked-for-stop`	`active`	`running`
`marked-for-termination`	`active`	`running`
`instance-stopped-by-price`	`disabled`	`stopped`
`instance-stopped-by-user`	`disabled`	`stopped`
`instance-stopped-no-capacity`	`disabled`	`stopped`
`instance-terminated-by-price`	`closed` (exceptionnelle), `open` (persistante)	`terminated`

Code d'état	État de la demande	État de l'instance
`instance-terminated-by-schedule`	`closed`	`terminated`
`instance-terminated-by-service`	`cancelled`	`terminated`
`instance-terminated-by-user`	`closed` ou `cancelled` *	`terminated`
`instance-terminated-no-capacity`	`closed` (exceptionnelle), `open` (persistante)	`terminated`
`instance-terminated-launch-group-constraint`	`closed` (exceptionnelle), `open` (persistante)	`terminated`

* L'état de la demande est `closed` si vous mettez hors service l'instance, mais que vous n'annulez pas la demande. L'état de la demande est `cancelled` si vous mettez l'instance hors service et que vous annulez la demande. Même si vous résiliez une instance Spot avant d'annuler sa demande, un certain laps de temps peut s'écouler avant qu'Amazon EC2 ne détecte la résiliation de votre instance Spot. Le cas échéant, l'état `closed` ou `cancelled` est attribué à la demande.

Demandes persistantes

Lorsque vos Instances Spot sont résiliées (soit par vous, soit par Amazon EC2), si la demande Spot est persistante, l'état `pending-evaluation` lui est attribué à nouveau et Amazon EC2 peut lancer une nouvelle instance Spot lorsque les contraintes sont respectées.

Obtention d'informations sur le statut de la demande

Vous pouvez obtenir des informations sur le statut de la demande à l'aide d'AWS Management Console ou d'un outil de ligne de commande.

Pour obtenir des informations sur le statut d'une demande (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Demandes Spot et sélectionnez la demande d'instance Spot.
3. Pour vérifier l'état, sous l'onglet Description, cochez le champ Statut.

Pour obtenir des informations sur le statut de la demande à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-Spot-instance-requests (AWS CLI)
- Get-EC2SpotInstanceRequest (Outils AWS pour Windows PowerShell)

Codes de statut des demandes Spot

Les informations sur le statut des demandes Spot sont composées d'un code de statut, de l'heure de mise à jour et d'un message de statut. Toutes ces informations vous permettent de savoir où en est votre demande d'instance Spot.

Voici les codes de statut des demandes Spot :

`az-group-constraint`

Amazon EC2 ne peut pas lancer toutes les instances que vous avez demandées dans la même zone de disponibilité.

`bad-parameters`

Un ou plusieurs paramètres de votre demande d'instance Spot ne sont pas valides (par exemple, l'AMI que vous avez spécifiée n'existe pas). Le message de statut indique quel paramètre n'est pas valide.

`canceled-before-fulfillment`

L'utilisateur a annulé la demande d'instance Spot avant son exécution.

`capacity-not-available`

Il n'y a pas suffisamment de capacité disponible pour les instances que vous avez demandées.

`constraint-not-fulfillable`

La demande d'instance Spot ne peut pas être satisfaite dans la mesure où une ou plusieurs contraintes ne sont pas valides (par exemple, la zone de disponibilité n'existe pas). Le message de statut indique quelle contrainte n'est pas valide.

`fulfilled`

La demande d'instance Spot est `active` et Amazon EC2 lance vos instances Spot.

`instance-stopped-by-price`

Votre instance a été arrêtée car le prix Spot a dépassé votre prix maximum.

`instance-stopped-by-user`

Votre instance a été arrêtée car un utilisateur l'a arrêtée ou a exécuté la commande shutdown à partir de l'instance.

`instance-stopped-no-capacity`

Votre instance a été arrêtée car il n'y avait plus de capacité Spot disponible pour l'instance.

`instance-terminated-by-price`

Votre instance a été supprimée car le prix Spot a dépassé votre prix maximum. Si votre demande est une offre persistante, le processus redémarre et votre demande se retrouve en attente d'évaluation.

`instance-terminated-by-schedule`

Votre instance Spot a été résiliée à la fin de sa durée planifiée.

`instance-terminated-by-service`

Votre instance a été mise hors service à partir d'un état d'arrêt.

`instance-terminated-by-user` ou `spot-instance-terminated-by-user`

Étant donné que vous avez résilié une instance Spot qui a été exécutée, l'état de la demande est `closed` (sauf s'il s'agit d'une demande persistante) et l'état de l'instance est `terminated`.

`instance-terminated-launch-group-constraint`

Une ou plusieurs instances de votre groupe de lancement ont été mises hors service, c'est pourquoi la contrainte du groupe de lancement n'est plus respectée.

`instance-terminated-no-capacity`

Votre instance a été supprimée car il n'y a plus de capacité Spot disponible pour l'instance.

`launch-group-constraint`

Amazon EC2 ne peut pas lancer toutes les instances que vous avez demandées en même temps. Toutes les instances d'un groupe de lancement sont démarrées et mises hors service ensemble.

`limit-exceeded`

La limite du nombre de volumes EBS ou du stockage de volume total a été dépassée. Pour plus d'informations sur ces limites et pour savoir comment demander une augmentation, consultez Limites Amazon EBS dans Référence générale d'Amazon Web Services.

`marked-for-stop`

L'instance Spot est marquée pour être arrêtée.

`marked-for-termination`

L'instance Spot est marquée pour être résiliée.

`not-scheduled-yet`

La demande d'instance Spot n'est pas évaluée avant la date prévue.

`pending-evaluation`

Une fois que vous avez effectué une demande d'instance Spot, elle passe à l'état `pending-evaluation` le temps que le système évalue les paramètres de votre demande.

`pending-fulfillment`

Amazon EC2 tente d'allouer vos Instances Spot.

`placement-group-constraint`

La demande Spot ne peut pas encore être satisfaite, car une instance Spot ne peut pas être ajoutée au groupe de placement à ce stade.

`price-too-low`

La demande ne peut pas encore être exécutée, car le prix maximum est inférieur au prix Spot. Dans le cas présent, aucune instance n'est lancée et votre demande reste à l'état `open`.

`request-canceled-and-instance-running`

Vous avez annulé la demande Spot alors que les Instances Spot sont toujours en cours d'exécution. La demande est `cancelled`, tandis que les instances conservent l'état `running`.

`schedule-expired`

La demande d'instance Spot est arrivée à expiration car elle n'a pas été exécutée avant la date spécifiée.

`system-error`

Il y a eu une erreur système inattendue. S'il s'agit d'un problème récurrent, veuillez contacter AWS Support pour obtenir de l'aide.

Interruptions d'une instance Spot

Vous pouvez lancer des Instances Spot sur la capacité EC2 inutilisée et bénéficier de remises importantes si vous les renvoyez lorsque le service Amazon EC2 a à nouveau besoin de la capacité. Lorsque le service Amazon EC2 réclame un Instances Spot, nous appelons cet événement une interruption d'instance Spot.

La demande d'Instances Spot peut varier considérablement d'un instant à l'autre et la disponibilité des Instances Spot peut également varier sensiblement selon le nombre d'instances EC2 disponibles qui ne sont pas utilisées. Par ailleurs, votre instance Spot est susceptible d'être interrompue à tout moment. Par conséquent, vous devez veiller à ce que votre application soit préparée à une interruption d'instance Spot.

Une instance à la demande spécifiée dans un Flotte EC2 ou Parc d'instances Spot ne peut pas être interrompue.

Sommaire

- Raisons d'interruption (p. 453)
- Comportements d'interruption (p. 453)
- Spécification du comportement d'interruption (p. 456)
- Préparation aux interruptions (p. 456)
- Préparation de la mise en veille d'une instance (p. 456)
- Avis d'interruption d'une instance Spot (p. 457)
- Recherche interrompue Instances Spot (p. 459)
- Déterminer si Amazon EC2 a interrompu une instance Spot (p. 459)
- Facturation des Instances Spot interrompues (p. 460)

Raisons d'interruption

Voici les raisons pour lesquelles Amazon EC2 est susceptible d'interrompre vos Instances Spot :

- Prix – le prix spot est supérieur à votre prix maximum.
- Capacité – S'il n'y a pas suffisamment d'instances EC2 inutilisées pour répondre à la demande d'Instances à la demande, Amazon EC2 interrompt les Instances Spot. L'ordre dans lequel les instances sont interrompues est déterminé par Amazon EC2.
- Contraintes – si votre demande comprend une contrainte telle qu'un groupe de lancement ou un groupe de zone de disponibilité, ces Instances Spot sont résiliées en tant que groupe lorsque la contrainte n'est plus respectée.

Comportements d'interruption

Vous pouvez spécifier que Amazon EC2 doit effectuer l'une des opérations suivantes lorsqu'il interrompt un instance Spot :

- Arrêter le instance Spot
- Mettre en veille prolongée la instance Spot
- Résilier la instance Spot

Par défaut, les Instances Spot interrompues sont résiliées. Pour modifier le comportement d'interruption, veuillez consulter Spécification du comportement d'interruption (p. 456).

Arrêt des Instances Spot interrompues

Vous pouvez spécifier le comportement d’interruption de sorte qu'Amazon EC2 arrête les Instances Spot lorsqu'elles sont interrompues si les exigences suivantes sont satisfaites.

Exigences

- Pour une demande d'instance Spot, le type doit être `persistent`. Vous ne pouvez pas spécifier un groupe de lancement dans la demande d'instance Spot.
- Pour une demande Flotte EC2 ou Parc d'instances Spot, le type doit être `maintain`.
- Le volume racine doit être un volume EBS, et non un volume de stockage d'instance.

Une fois qu'une instance Spot a été arrêtée par le service Spot, seul ce service peut redémarrer l'instance Spot et la même configuration de lancement doit être utilisée.

Pour une instance Spot lancée par une demande d'instance Spot `persistent`, le service Spot redémarre l'instance arrêtée quand la capacité est disponible dans la même zone de disponibilité et pour le même type d'instance que l'instance arrêtée.

Si des instances figurant dans un Flotte EC2 ou Parc d'instances Spot sont arrêtées et que le parc est de type `maintain`, le service Spot lance des instances de remplacement pour maintenir la capacité cible. Le service Spot trouve le ou les meilleurs pool(s) en fonction de la stratégie d'allocation spécifiée (`lowestPrice`, `diversified` ou `InstancePoolsToUseCount`). Il ne donne pas la priorité au pool contenant les instances arrêtées le plus tôt. Ultérieurement, si la stratégie d'allocation conduit à un pool contenant les instances arrêtées le plus tôt, le service Spot redémarre les instances arrêtées pour assurer la capacité cible.

Par exemple, considérons un Parc d'instances Spot avec la stratégie d'allocation `lowestPrice`. Lors du lancement initial, un pool `c3.large` répond aux critères `lowestPrice` pour la spécification de lancement. Ultérieurement, quand les instances `c3.large` sont interrompues, le service Spot arrête les instances et réapprovisionne la capacité à partir d'un autre pool adapté à la stratégie `lowestPrice`. Cette fois, le pool est un pool `c4.large` et le service Spot lance des instances `c4.large` pour assurer la capacité cible. De même, le Parc d'instances Spot peut devenir un pool `c5.large` la prochaine fois. Dans chacune de ces transitions, le service Spot ne donne pas la priorité aux pools contenant les instances arrêtées le plus tôt, mais définit les priorités uniquement en fonction de la stratégie d'allocation spécifiée. La stratégie `lowestPrice` peut ramener aux pools contenant les instances arrêtées le plus tôt. Par exemple, si des instances sont interrompues dans le pool `c5.large` et que la stratégie `lowestPrice` ramène aux pools `c3.large` ou `c4.large`, les instances arrêtées le plus tôt sont redémarrées pour assurer la capacité cible.

Pendant qu'une instance Spot est arrêtée, vous pouvez modifier certains de ses attributs, mais pas le type d'instance. Si vous détachez ou supprimez un volume EBS, celui-ci n'est pas attaché lorsque l'instance Spot est démarrée. Si vous détachez le volume racine et que le service Spot tente de démarrer l'instance Spot, l'instance ne peut pas démarrer et le service Spot résilie l'instance arrêtée.

Vous pouvez résilier une instance Spot pendant qu'elle est arrêtée. Si vous annulez une demande Spot, un Flotte EC2 ou un Parc d'instances Spot, le service Spot résilie les Instances Spot associées qui sont arrêtées.

Pendant qu'une instance Spot est arrêtée, seuls les volumes EBS préservés vous sont facturés. Avec Flotte EC2 et Parc d'instances Spot, si vous avez de nombreuses instances arrêtées, vous pouvez dépasser la limite du nombre de volumes EBS pour votre compte.

Mise en veille prolongée des Instances Spot interrompues

Vous pouvez spécifier le comportement d’interruption de sorte qu'Amazon EC2 mette les Instances Spot en veille prolongée lorsqu'elles sont interrompues si les exigences suivantes sont satisfaites.

Exigences

- Pour une demande d'instance Spot, le type doit être `persistent`. Vous ne pouvez pas spécifier un groupe de lancement dans la demande d'instance Spot.
- Pour une demande Flotte EC2 ou Parc d'instances Spot, le type doit être `maintain`.
- Le volume racine doit correspondre à un volume EBS, et non à un volume de stockage d'instance. Il doit être suffisamment volumineux pour stocker la mémoire de l'instance (RAM) lors de la mise en veille prolongée.
- Les instances suivantes sont prises en charge : C3, C4, C5, M4, M5, R3 et R4, avec moins de 100 Go de mémoire.
- Les systèmes d'exploitation pris en charge sont les suivants : Amazon Linux 2, Amazon Linux AMI, Ubuntu avec un noyau Ubuntu adapté à AWS (linux-aws) d'une version supérieure à 4.4.0-1041, ainsi que Windows Server 2008 R2 et versions ultérieures.
- Installez l'agent de mise en veille prolongée sur un système d'exploitation pris en charge, ou utilisez l'un des AMI suivants, qui incluent déjà l'agent :
  - Amazon Linux 2
  - Amazon Linux AMI 2017.09.1 ou version ultérieure

  - Ubuntu Xenial 16.04 20171121 ou version ultérieure
  - AMI Windows Server 2008 R2 2017.11.19 ou version ultérieure
  - AMI Windows Server 2012 ou Windows Server 2012 R2 2017.11.19 ou version supérieure
  - AMI Windows Server 2016 2017.11.19 ou version supérieure
  - Windows Server 2019
- Démarrez l'agent. Nous vous conseillons d'utiliser les données utilisateur pour démarrer l'agent lors du lancement de l'instance. Vous pouvez également démarrer l'agent manuellement.

Recommandation

- Nous vous recommandons vivement d'utiliser un volume Amazon EBS chiffré en tant que volume racine, car la mémoire de l'instance est stockée sur le volume racine lors de la mise en veille. Cela permet de garantir que le contenu de la mémoire (RAM) est chiffré lorsque les données sont au repos sur le volume, ou lorsqu'elles sont déplacées entre l'instance et le volume. L'une des trois options suivantes permet de s'assurer que le volume racine est une volume Amazon EBS chiffré :
  - Chiffrement EBS en une étape : dans un seul appel d'API d'exécution d'instances, vous pouvez lancer des instances EC2 soutenues par EBS depuis une AMI chiffrée . Pour de plus amples informations, veuillez consulter Utiliser le chiffrement avec des AMI basées sur EBS (p. 168).
  - Chiffrement EBS par défaut : vous pouvez activer le chiffrement EBS par défaut afin de vous assurer que tous les nouveaux volumes EBS de votre compte AWS sont chiffrés. Pour de plus amples informations, veuillez consulter Chiffrement par défaut (p. 1196).
  - AMI chiffrée : vous pouvez activer le chiffrement EBS en utilisant une AMI chiffrée pour lancer votre instance. Si votre AMI ne dispose d'aucun volume racine chiffré, vous pouvez le copier sur le nouvel AMI et demander son chiffrement. Pour de plus amples informations, veuillez consulter Chiffrement d'une image non chiffrée pendant la copie (p. 172) et Copie d'une AMI (p. 177).

Lorsqu'une instance Spot est mise en veille prolongée par le service Spot, les volumes EBS sont conservés et la mémoire d'instance (RAM) est préservée sur le volume racine. Les adresses IP privées de l'instance sont également conservées. Par contre, les volumes de stockage d'instance et les adresses IP publiques, autres que les adresses IP Elastic, ne sont pas conservés. Lorsque l'instance est mise en veille, seuls les volumes EBS qui sont préservés vous sont facturés. Avec Flotte EC2 et Parc d'instances Spot, si vous avez de nombreuses instances en veille prolongée, vous pouvez dépasser la limite du nombre de volumes EBS pour votre compte.

L'agent invite le système d'exploitation à se mettre en veille lorsque l'instance reçoit un signal du service d'instances Spot. Si l'agent n'est pas installé, si le système d'exploitation sous-jacent ne prend pas en charge la mise en veille, ou si l'espace sur le volume est insuffisant pour l'enregistrement de la mémoire de l'instance, la mise en veille échoue et le service d'instances Spot arrête l'instance à la place.

Lorsque le service Spot met une instance Spot en veille prolongée, vous recevez un avis d'interruption, mais l'instance Spot est interrompue en moins de deux minutes. La mise en veille commence immédiatement. Lorsque l'instance est en cours de mise en veille, la vérification de son état peut échouer. Lorsque le processus de mise en veille se termine, l'état de l'instance est défini sur `stopped`.

Reprise d'une mise en veille prolongée instance Spot

Une fois qu'une instance Spot est mise en veille prolongée par le service Spot, ce dernier est le seul à pouvoir la reprendre. Le service d'instances Spot relance l'instance lorsque la capacité est disponible, avec un prix Spot inférieur au prix maximum spécifié.

Pour de plus amples informations, veuillez consulter Préparation de la mise en veille d'une instance (p. 456).

Pour de plus amples informations sur l'hibernation de Instances à la demande, veuillez consulter Mettre votre instance Linux en veille prolongée (p. 628).

Spécification du comportement d'interruption

Si vous ne spécifiez pas de comportement d'interruption, par défaut les Instances Spot interrompues sont résiliées. Vous pouvez spécifier le comportement d'interruption lorsque vous créez une demande Spot. La façon dont vous spécifiez le comportement d'interruption est différente selon la façon dont vous demandez Instances Spot.

Si vous faites une demande Instances Spot à l'aide de l'assistant de lancement d'instance (p. 530), vous pouvez spécifier le comportement d'interruption comme suit : activez la case à cocher Demande persistante, puis, dans Comportement d'interruption, choisissez un comportement d'interruption.

Si vous faites une demande Instances Spot à l'aide de la console Spot (p. 416), vous pouvez spécifier le comportement d'interruption comme suit : activez la case à cocher Maintenir la capacité cible puis, dans Comportement d'interruption, choisissez un comportement d'interruption.

Si vous configurez Instances Spot dans un modèle de lancement (p. 539), vous pouvez spécifier le comportement d'interruption comme suit : dans le modèle de lancement, développez Détails avancés et cochez la case Demande Instances Spot. Choisissez Personnaliser, puis, dans Comportement d'interruption, choisissez un comportement d'interruption.

Si vous configurez Instances Spot dans une configuration de lancement lors de l'utilisation de l'interface de ligne de commande request-spot-fleet, vous pouvez spécifier le comportement d'interruption comme suit : pour `InstanceInterruptionBehavior`, spécifiez un comportement d'interruption.

Si vous configurez Instances Spot à l'aide de l'interface de ligne de commande request-spot-instances vous pouvez spécifier le comportement d'interruption comme suit : pour `--instance-interruption-behavior`, spécifiez un comportement d'interruption.

Préparation aux interruptions

Voici quelques bonnes pratiques à suivre lorsque vous utilisez des Instances Spot :

- Utilisez le prix maximum par défaut, qui correspond au prix à la demande.
- Veillez à ce que votre instance soit prête pour le lancement dès que la demande est exécutée en utilisant une Amazon Machine Image (AMI) comportant la configuration logicielle requise. Vous pouvez également utiliser les données utilisateur afin d'exécuter les commandes lors du démarrage.
- Stockez les données importantes régulièrement à un emplacement qui n'est pas touché par la résiliation d'une instance Spot. Par exemple, vous pouvez utiliser Amazon S3, Amazon EBS ou DynamoDB.
- Divisez le travail en petites tâches (à l'aide d'une architecture Grid, Hadoop ou reposant sur les files d'attente) ou utilisez des points de contrôle afin de pouvoir enregistrer votre travail fréquemment.
- Utilisez les avis d'interruption d'instance Spots pour surveiller le statut de vos instances Spot.
- Même si nous nous efforçons de vous communiquer cet avertissement dès que possible, il se peut que votre instance Spot soit résiliée avant que l'avertissement puisse être mis à disposition. Testez votre application afin de vous assurer qu'elle peut gérer correctement une mise hors service inattendue d'une instance, même si vous testez les avis d'interruption. Pour cela, exécutez l'application en utilisant une instance à la demande, puis résiliez vous-même cette instance à la demande.

Préparation de la mise en veille d'une instance

Vous devez installer un agent de mise en veille sur votre instance, sauf si vous avez utilisé un AMI qui inclut déjà cet agent. Exécutez l'agent lors du démarrage de l'instance, que l'agent soit inclus dans votre AMI ou que vous l'ayez installé vous-même.

Les procédures suivantes vous aident à préparer une instance Linux. Pour savoir comment préparer une instance Windows, consultez Préparation de la mise en veille prolongée d'une instance dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Pour préparer une instance Amazon Linux

1. Vérifiez que le Kernel prend en charge la mise en veille. Mettez-le à jour, le cas échéant.
2. Si votre AMI n'inclut aucun agent, installez-le à l'aide de la commande suivante.

```
sudo yum update; sudo yum install hibagent
```

3. Ajoutez la commande suivante aux données utilisateur :

```
#!/bin/bash
/usr/bin/enable-ec2-spot-hibernation
```

Pour préparer une instance Ubuntu

1. Si votre AMI n'inclut aucun agent, installez-le à l'aide de la commande suivante. L'agent de mise en veille prolongée est disponible uniquement sur Ubuntu 16.04 ou version ultérieure.

```
sudo apt-get install hibagent
```

2. Ajoutez la commande suivante aux données utilisateur.

```
#!/bin/bash
/usr/bin/enable-ec2-spot-hibernation
```

Avis d'interruption d'une instance Spot

La meilleure façon pour vous de gérer gracieusement les interruptions instance Spot consiste à concevoir votre application pour qu'elle soit tolérante aux pannes. Pour ce faire, vous pouvez profiter des avis d'interruption instance Spot. Un avis d'interruption instance Spot est un avertissement qui est émis deux minutes avant l'arrêt de Amazon EC2 ou la fin de votre instance Spot. Lorsque vous spécifiez la mise en veille comme comportement d'interruption, vous recevez un avis d'interruption, mais vous ne recevez pas d'avertissement de deux minutes car le processus de mise en veille commence immédiatement.

Nous vous recommandons de vérifier ces avis d'interruption toutes les 5 secondes.

Cet avertissement est rendu disponible comme événement CloudWatch et comme élément des métadonnées de l'instance (p. 701) sur l'instance Spot.

EC2 instance Spot interruption notice

Quand Amazon EC2 va interrompre votre instance Spot, il émet un événement deux minutes avant l'interruption réelle (sauf pour l'hibernation, qui reçoit l'avis d'interruption, mais pas deux minutes à l'avance, car l'hibernation commence immédiatement). Cet événement peut être détecté par Amazon CloudWatch Events. Pour de plus amples informations sur les événements CloudWatch, veuillez consulter Guide de l'utilisateur Amazon CloudWatch Events. Pour obtenir un exemple détaillé qui vous explique comment créer et utiliser des règles d'événement, veuillez consulter Tirer parti des avis d'interruption d'instance Spot Amazon EC2.

Vous trouverez ci-dessous un exemple d'événement pour une interruption d'instance Spot. Les valeurs possibles pour `instance-action` sont `hibernate`, `stop` et `terminate`.

```
{
    "version": "0",
    "id": "12345678-1234-1234-1234-123456789012",
    "detail-type": "EC2 Spot Instance Interruption Warning",
    "source": "aws.ec2",
```

```
    "account": "123456789012",
    "time": "yyyy-mm-ddThh:mm:ssZ",
    "region": "us-east-2",
    "resources": ["arn:aws:ec2:us-east-2:123456789012:instance/i-1234567890abcdef0"],
    "detail": {
        "instance-id": "i-1234567890abcdef0",
        "instance-action": "action"
    }
}
```

instance-action

Si votre instance Spot est marquée comme devant être arrêtée ou résiliée par le service Spot, l'élément `instance-action` est présent dans les métadonnées de l'instance (p. 701). Sinon, il n'est pas présent. Vous pouvez extraire la valeur `instance-action` comme suit.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/spot/instance-action
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/spot/instance-action
```

L'élément `instance-action` spécifie l'action et l'heure approximative (UTC) à laquelle l'action aura lieu.

L'exemple suivant indique la date et l'heure auxquelles cette instance sera arrêtée.

```
{"action": "stop", "time": "2017-09-18T08:22:00Z"}
```

L'exemple suivant indique la date et l'heure auxquelles cette instance sera résiliée.

```
{"action": "terminate", "time": "2017-09-18T08:22:00Z"}
```

Si Amazon EC2 ne s'apprête pas à supprimer ou arrêter l'instance, ou si vous avez suspendu vous-même l'instance, `instance-action` n'est pas présent et vous recevez une erreur HTTP 404 lorsque vous tentez de la récupérer.

termination-time

Cet élément est conservé à des fins de compatibilité descendante ; nous vous invitons à utiliser `instance-action` à la place.

Si votre instance Spot est marquée pour être résiliée par le service Spot, l'élément `termination-time` est présent dans les métadonnées de l'instance. Sinon, il n'est pas présent. Vous pouvez extraire la valeur `termination-time` comme suit.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"`
[ec2-user ~]$ if curl -H "X-aws-ec2-metadata-token: $TOKEN" -s http://169.254.169.254/
latest/meta-data/spot/termination-time | grep -q .*T.*Z; then echo terminated; fi
```

IMDSv1

```
[ec2-user ~]$ if curl -s http://169.254.169.254/latest/meta-data/spot/termination-time
 | grep -q .*T.*Z; then echo terminated; fi
```

L'élément `termination-time` spécifie l'heure approximative (au format UTC) à laquelle l'instance reçoit le signal d'arrêt. Par exemple :

```
2015-01-05T18:02:00Z
```

Si Amazon EC2 ne s'apprête pas à mettre hors service l'instance ou si vous avez mis hors service vous-même l'instance Spot, l'élément `termination-time` est absent (et vous recevez une erreur HTTP 404) ou il contient une valeur qui n'est pas temporelle.

Si Amazon EC2 ne parvient pas à mettre hors service l'instance, le statut de la demande est défini sur `fulfilled`. La valeur `termination-time` reste dans les métadonnées de l'instance avec l'heure approximative initiale, qui se trouve maintenant dans le passé.

Recherche interrompue Instances Spot

Dans la console, le volet Instances affiche toutes les instances, y compris Instances Spot. Vous pouvez identifier une instance Spot à partir de la valeur `spot` de la colonne Cycle de vie de l'instance. La colonne État de l'instance indique si l'instance est `pending`, `running`, `stopping`, `stopped`, `shutting-down` ou `terminated`. Pour une instance Spot mise en veille de manière prolongée, l'état de l'instance indique `stopped`.

Pour identifier une instance Spot interrompue (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances. Dans le coin supérieur droit, choisissez l'icône Paramètres (⚙), puis sous Colonnes Attribut, sélectionnez Cycle de vie de l'instance. Pour Instances Spot, le cycle de vie de l'instance est `spot`.

   Sinon, dans le panneau de navigation, sélectionnez Demandes Spot. Vous pouvez voir à la fois les demandes d'instance Spot et les demandes de Parc d'instances Spot. Pour afficher les ID des instances, sélectionnez une demande instance Spot ou Parc d'instances Spot, puis choisissez l'onglet Instances. Choisissez un ID pour afficher l'instance correspondante dans le volet Instances.
3. Pour chaque instance Spot, vous pouvez afficher son état dans la colonne État de l'instance.

Pour identifier des Instances Spot interrompues (AWS CLI)

Vous pouvez répertorier les Instances Spot interrompues à l'aide de la commande describe-instances avec le paramètre `--filters`. Pour répertorier uniquement les ID d'instance dans la sortie, ajoutez le paramètre `--query`.

```
aws ec2 describe-instances \
    --filters Name=instance-lifecycle,Values=spot Name=instance-state-
name,Values=terminated,stopped \
    --query Reservations[*].Instances[*].InstanceId
```

Déterminer si Amazon EC2 a interrompu une instance Spot

Si une instance Spot est arrêtée, mise en veille prolongée ou résiliée, vous pouvez utiliser CloudTrail pour voir si Amazon EC2 a interrompu la instance Spot. Dans CloudTrail, le nom de l'événement

`BidEvictedEvent` indique que Amazon EC2 a interrompu la instance Spot. Pour plus d'informations sur l'utilisation de CloudTrail, consultez Journalisation des appels d'API Amazon EC2 et Amazon EBS avec AWS CloudTrail (p. 804).

Facturation des Instances Spot interrompues

Lorsqu'une instance Spot (non intégrée à un bloc Spot) est interrompue, vous êtes facturé comme suit.

Qui interrompt l'instance Spot	Système d'exploitation	Interrompue au cours de la première heure	Interrompue après la première heure
Si vous arrêtez ou résiliez l'instance Spot	Linux (sauf RHEL et SUSE)	Les secondes utilisées sont facturées	Les secondes utilisées sont facturées
	Windows, RHEL, SUSE	L'heure complète est facturée même si vous n'en avez utilisé qu'une partie	Les heures complètes utilisées sont facturées, et une heure complète est facturée pour l'heure partielle interrompue
Si Amazon EC2 interrompt l'instance Spot	Linux (sauf RHEL et SUSE)	Aucuns frais.	Les secondes utilisées sont facturées
	Windows, RHEL, SUSE	Aucuns frais.	Les heures complètes utilisées sont facturées, mais l'heure partielle interrompue n'est pas facturée

Lorsqu'une instance Spot intégrée à un bloc Spot est interrompue, vous êtes facturé comme suit.

Qui interrompt l'instance Spot	Système d'exploitation	Interrompue au cours de la première heure	Interrompue après la première heure
Si vous arrêtez ou résiliez l'instance Spot	Linux (sauf RHEL et SUSE)	Les secondes utilisées sont facturées	Les secondes utilisées sont facturées
	Windows, RHEL, SUSE	L'heure complète est facturée même si vous n'en avez utilisé qu'une partie	Les heures complètes utilisées sont facturées, et une heure complète est facturée pour l'heure partielle interrompue
Si Amazon EC2 interrompt l'instance Spot	Linux (sauf RHEL et SUSE)	Aucuns frais.	Aucuns frais.
	Windows, RHEL, SUSE	Aucuns frais.	Aucuns frais.

Flux de données d'une instance Spot

Pour vous aider à comprendre les frais associés à vos Instances Spot, Amazon EC2 fournit un flux de données qui décrit votre utilisation d'instance Spot et la tarification. Ce flux de données est envoyé vers un compartiment Amazon S3 que vous spécifiez lorsque vous vous abonnez au flux de données.

Les fichiers de flux de données arrivent généralement dans votre compartiment toutes les heures et chaque heure d'utilisation est généralement couverte dans un seul fichier de données. Ces fichiers sont

compressés (gzip) avant d'être transmis à votre compartiment. Amazon EC2 peut inscrire les données dans plusieurs fichiers pour une heure spécifique d'utilisation lorsque les fichiers sont volumineux (par exemple, si le contenu du fichier pour cette heure dépasse les 50 Mo avant compression).

> **Note**
>
> Si vous n'avez aucune instance Spot en cours d'exécution à une certaine heure, vous ne recevez pas de fichier de flux de données pour cette heure.

Sommaire

- Nom et format du fichier de flux de données (p. 461)
- Conditions requises pour le compartiment Amazon S3 (p. 462)
- Abonnement à votre flux de données d'une instance Spot (p. 462)
- Suppression de votre flux de données d'une instance Spot (p. 462)

Nom et format du fichier de flux de données

Le nom du fichier de flux de données d'instance Spots utilise le format suivant (avec la date et l'heure UTC) :

```
bucket-name.s3.amazonaws.com/optional-prefix/aws-account-id.YYYY-MM-DD-HH.n.unique-id.gz
```

Par exemple, si le nom de votre compartiment est **`my-bucket-name`** et que votre préfixe est **`my-prefix`**, vos noms de fichier ont le format suivant :

```
my-bucket-name.s3.amazonaws.com/my-prefix/111122223333.2019-03-17-20.001.pwBdGTJG.gz
```

Pour de plus amples informations sur les noms de compartiment, veuillez consulter Règles relatives à l'attribution des noms de compartiments dans le Amazon Simple Storage Service Manuel du développeur.

Les fichiers de flux de données d'instance Spots sont délimités par des tabulations. Chaque ligne du fichier de données correspond à une heure d'instance et contient les champs répertoriés dans le tableau suivant.

Champ	Description
`Timestamp`	Horodatage utilisé pour déterminer le prix facturé pour cette utilisation d'instance.
`UsageType`	Type d'utilisation et type d'instance associés à la facturation. Pour les instances Spot `m1.small`, ce champ est défini sur `SpotUsage`. Pour tous les autres types d'instance, ce champ est défini sur `SpotUsage:`{instance-type. Par exemple, `SpotUsage:c1.medium`.
`Operation`	Le produit faisant l'objet d'une facturation. Pour les Instances Spot Linux, ce champ est défini sur `RunInstances`. Pour les Instances Spot Windows, ce champ est défini sur `RunInstances:0002`. L'utilisation des instances Spot est regroupée par zone de disponibilité.
`InstanceID`	ID de l'instance Spot qui a généré cette utilisation d'instance.
`MyBidID`	ID de la demande d'instance Spot qui a généré cette utilisation d'instance.
`MyMaxPrice`	Prix maximum spécifié pour cette demande d'instance Spot.
`MarketPrice`	Prix Spot au moment spécifié dans le champ `Timestamp`.
`Charge`	Prix facturé pour cette utilisation d'instance.

Champ	Description
`Version`	Version incluse dans le nom du fichier de flux de données pour cet enregistrement.

Conditions requises pour le compartiment Amazon S3

Lorsque vous vous abonnez au flux de données, vous devez spécifier un compartiment Amazon S3 afin de stocker les fichiers de flux de données. Avant de choisir un compartiment Amazon S3 pour le flux de données, tenez compte des points suivants :

- Vous devez disposer de l'autorisation `FULL_CONTROL` sur le compartiment, qui englobe l'autorisation pour les actions `s3:GetBucketAcl` et `s3:PutBucketAcl`.

  Si vous êtes le propriétaire du compartiment, vous disposez de cette autorisation par défaut. Sinon, le propriétaire du compartiment doit accorder cette autorisation à votre compte AWS.
- Lorsque vous vous abonnez à un flux de données, ces autorisations servent à mettre à jour la liste ACL du compartiment pour donner l'autorisation `FULL_CONTROL` au compte de flux de données AWS. Le compte de flux de données AWS écrit des fichiers de flux de données dans le compartiment. Si votre compte ne dispose pas des autorisations nécessaires, les fichiers de flux de données ne peuvent pas être écrits dans le compartiment.

  > **Note**
  >
  > Si vous mettez à jour la liste ACL et que vous supprimez les autorisations pour le compte de flux de données AWS, les fichiers de flux de données ne peuvent pas être écrits dans le compartiment. Vous devez vous réabonner au flux de données pour recevoir les fichiers de flux de données.
- Chaque fichier de flux de données a son propre ACL (distinct de celui du compartiment). Le propriétaire du compartiment bénéficie de l'autorisation `FULL_CONTROL` pour les fichiers de données. Le compte de flux de données AWS a des autorisations de lecture et d'écriture.
- Si vous supprimez votre abonnement au flux de données, Amazon EC2 ne supprime pas les autorisations de lecture et d'écriture pour le compte de flux de données AWS sur le compartiment ou les fichiers de données. Vous devez supprimer ces autorisations vous-même.

Abonnement à votre flux de données d'une instance Spot

Pour vous abonner à votre flux de données, utilisez la commande create-Spot-datafeed-subscription.

```
aws ec2 create-spot-datafeed-subscription \
    --bucket my-bucket-name \
    [--prefix my-prefix]
```

Voici un exemple de sortie :

```
{
    "SpotDatafeedSubscription": {
        "OwnerId": "111122223333",
        "Bucket": "my-bucket-name",
        "Prefix": "my-prefix",
        "State": "Active"
    }
}
```

Suppression de votre flux de données d'une instance Spot

Pour supprimer votre flux de données, utilisez la commande delete-Spot-datafeed-subscription.

```
aws ec2 delete-spot-datafeed-subscription
```

Limites d'une instance Spot

Il y a une limite sur le nombre d'Instances Spot demandées et en cours d'exécution par compte AWS et par région. Les limites d'instance Spot sont gérées en fonction du nombre d'unités de traitement centralisées virtuelles (vCPU) que votre Instances Spot en cours d'exécution utilise ou utilisera en attendant le traitement des demandes d'instance Spot ouvertes. Si vous résiliez votre Instances Spot, mais que vous n'annulez pas les demandes d'instance Spot, ces dernières sont comptabilisées par rapport à votre limite de demandes d'instance Spot dynamiques jusqu'à ce que Amazon EC2 détecte les résiliations instance Spot et clôture les demandes.

Rubriques

- Limites d'une instance Spot (p. 463)
- Demande d'une augmentation de limite d'instance Spot (p. 463)
- Surveillance des limites et de l'utilisation d'instance Spot (p. 464)
- Limites d'un Parc d'instances Spot (p. 464)

Limites d'une instance Spot

Il existe six limites d'instance Spot, répertoriées dans le tableau suivant. Chaque limite spécifie la limite de vCPU pour une ou plusieurs familles d'instances. Pour plus d'informations sur les différentes familles, générations et tailles d'instances, consultez Types d'instance Amazon EC2.

Nom de la limite d'instance Spot	Limite de vCPU par défaut
Toutes les demandes d'instance Spot standard (A, C, D, H, I, M, R, T, Z)	1 440 vCPU
Toutes les demandes d'instance Spot F	11 vCPU
Toutes les demandes d'instance Spot G	11 vCPU
Toutes les demandes d'instance Spot Inf	64 vCPU
Toutes les demandes instance Spot P	16 vCPU
Toutes les demandes instance Spot X	21 vCPU

> **Note**
>
> Les nouveaux comptes AWS peuvent commencer avec des limites inférieures à celles présentées ici. Ces limites peuvent augmenter au fil du temps.

Avec les limites de vCPU, vous pouvez utiliser votre limite en fonction du nombre de vCPU nécessaires pour lancer toute combinaison de types d'instance qui répond à l'évolution de vos besoins en termes d'applications. Par exemple, avec une limite Toutes les demandes d'instance Spot Standard de 256 vCPU, vous pouvez demander 32 Instances Spot `m5.2xlarge` (32 x 8 vCPU) ou 16 Instances Spot `c5.4xlarge` (16 x 16 vCPU), ou une combinaison de tous les types et tailles d'instance Spot standard totalisant 256 vCPU.

Demande d'une augmentation de limite d'instance Spot

Même si Amazon EC2 augmente automatiquement vos limites d'instance Spot en fonction de votre utilisation, vous pouvez demander une augmentation de limite si nécessaire. Par exemple, si vous avez

l'intention de lancer plus d'Instances Spot que celles autorisées par votre limite actuelle, vous pouvez demander une augmentation de limite. Si vous soumettez une demande d'instance Spot et que vous recevez l'erreur `Max spot instance count exceeded`, vous pouvez demander une augmentation de limite.

Pour demander une augmentation de limite d'instance Spot

1. Ouvrez Créer un cas, Augmentation de limite de service dans la console du Centre de support à l'adresse https://console.aws.amazon.com/support/home#/case/create.
2. Pour Limit type (Type de limite), choisissez EC2 Spot Instances (Instances Spot EC2).
3. Pour Région, sélectionnez la région requise.
4. Pour Type d'instance principale, sélectionnez la limite d'instance Spot pour laquelle vous souhaitez demander une augmentation de limite.
5. Pour Nouvelle valeur limite, saisissez le nombre total de vCPU que vous souhaitez exécuter simultanément. Pour déterminer le nombre total de vCPU dont vous avez besoin, consultez Types d'instances Amazon EC2 pour trouver le nombre de vCPU de chaque type d'instance.
6. (Conditionnel) Vous devez créer une demande de limite distincte pour chaque limite d'instance Spot. Pour demander une augmentation pour une autre limite d'instance Spot, choisissez Ajouter une autre demande et répétez les étapes 4 et 5 de cette procédure.
7. Pour Description du cas d'utilisation, entrez votre cas d'utilisation, puis choisissez Soumettre.

Pour plus d'informations sur l'affichage des limites et la demande d'augmentation des limites, veuillez consulter Quotas de service Amazon EC2 (p. 1322).

Surveillance des limites et de l'utilisation d'instance Spot

Vous pouvez afficher et gérer les limites de votre instance Spot à l'aide des éléments suivants :

- La page Limites de la console Amazon EC2
- La page Quotas de service Amazon EC2 dans la console Quotas de service
- L'AWS CLI get-service-quota

Pour plus d'informations, consultez Quotas de service Amazon EC2 (p. 1322) dans le Amazon EC2 Guide de l'utilisateur pour les instances Linux et Afficher un quota de service dans le Guide de l'utilisateur Service Quotas.

Avec l'intégration des métriques Amazon CloudWatch, vous pouvez surveiller votre utilisation d'EC2 par rapport aux limites. Vous pouvez également configurer des alarmes pour vous avertir lorsque vous approchez des limites. Pour de plus amples informations, veuillez consulter Utilisation d'alarmes Amazon CloudWatch dans le guide de l'utilisateur Service Quotas.

Limites d'un Parc d'instances Spot

Les limites d'Amazon EC2 habituelles s'appliquent aux instances lancées par un Parc d'instances Spot ou un Flotte EC2, telles que les limites d'instance Spot et les limites de volume. En outre, les limites suivantes s'appliquent :

- Nombre de Parcs d'instances Spot et de Flottes EC2 actifs par région : 1 000*
- Nombre de groupes instance Spot (combinaison unique de type d'instance et de sous-réseau) : 300*
- Taille des données utilisateur dans une spécification de lancement : 16 Ko*
- Capacité cible par Parc d'instances Spot ou Flotte EC2 : 10 000
- Capacité cible pour tous les Parcs d'instances Spot et Flottes EC2 d'une région : 100 000

- Une demande de Parc d'instances Spot ou de Flotte EC2 ne peut pas couvrir plusieurs régions.
- Une demande de Parc d'instances Spot ou de Flotte EC2 ne peut pas couvrir différents sous-réseaux de la même zone de disponibilité.

S'il vous faut une capacité cible supérieure aux limites par défaut, remplissez le formulaire Créer une demande du centre AWS Support pour demander l'augmentation de la limite. Pour Limit type (Type de limite), choisissez EC2 Fleet (Flotte EC2), choisissez une région, puis choisissez Target Fleet Capacity per Fleet (in units) (Capacité de flotte cible par flotte (en unités)), Target Fleet Capacity per Region (in units) (Capacité de flotte cible par région (en unités)) ou les deux.

* Ces limites sont finales. Vous ne pouvez pas demander une augmentation de ces limites.

Instances à capacité extensible

Si vous lancez vos Instances Spot à l'aide d'un type d'instance à capacité extensible (p. 233), et si vous prévoyez d'utiliser vos Instances Spot à capacité extensible immédiatement et pour une courte durée, sans temps d'inactivité pour accumuler des crédits UC, nous vous recommandons de les lancer en mode standard (p. 246) pour éviter de payer des coûts plus élevés. Si vous lancez vos Instances Spot à capacité extensible en mode Illimité (p. 237) et que vous étendez immédiatement l'utilisation de l'UC, vous dépensez des crédits excédentaires pour cette extension d'utilisation. Si vous utilisez l'instance pour une courte durée, elle n'a pas le temps d'accumuler des crédits UC pour rembourser les crédits excédentaires, et ces derniers vous sont facturés lorsque vous résiliez l'instance.

Le mode Illimité convient aux Instances Spot à capacité extensible uniquement si l'instance s'exécute suffisamment longtemps pour accumuler des crédits UC pour l'extension d'utilisation. Sinon, payer des crédits excédentaires rend les Instances Spot à capacité extensible plus coûteuses que les autres instances. Pour de plus amples informations, veuillez consulter Quand utiliser le mode illimité/mode d'UC fixe ? (p. 238).

Les crédits de lancement visent à optimiser la productivité du lancement initial des instances T2 en leur fournissant suffisamment de ressources de calcul pour pouvoir configurer l'instance. Il est interdit de procéder à des lancements répétés d'instances T2 pour bénéficier de nouveaux crédits de lancement. Si vous avez besoin de performances soutenues de l'UC, vous pouvez obtenir des crédits (en restant inactif pendant un certain temps) : utilisez le mode Illimité (p. 237) pour les Instances Spot T2 ou un type d'instance avec UC dédiée.

Hôtes dédiés

Un Hôte dédié Amazon EC2 est un serveur physique avec une capacité d'instance EC2 entièrement dédiée à votre utilisation. Les Hôtes dédiés vous permettent d'utiliser vos licences logicielles existantes par socket, par cœur ou par machine virtuelle, parmi lesquelles figurent Microsoft Windows Server, Microsoft SQL Server, SUSE et Linux Enterprise Server.

Pour de plus amples informations sur les configurations prises en charge sur Hôtes dédiés, veuillez consulter le tableau de configuration des hôtes dédiés.

Sommaire

- Différences entre les Hôtes dédiés et les Instances dédiées (p. 466)
- Bring Your Own License (BYOL, licence à fournir) (p. 466)
- Capacité d'instance d'un Hôte dédié (p. 467)
- Restrictions Hôtes dédiés (p. 467)
- Tarification et facturation (p. 468)
- Utilisation d'Hôtes dédiés (p. 469)
- Utilisation d'Hôtes dédiés partagés (p. 487)

- Récupération de l'hôte (p. 492)
- Suivi des changements de configuration (p. 497)

Différences entre les Hôtes dédiés et les Instances dédiées

Les Hôtes dédiés et les Instances dédiées peuvent être utilisés pour lancer des instances Amazon EC2 sur des serveurs physiques qui sont dédiés à votre utilisation.

Il n'existe pas de différence physique, de sécurité ou de performance entre les Instances dédiées et les instances des Hôtes dédiés. Cependant, il y a quelques différences entre les deux. Le tableau suivant souligne certaines différences essentielles entre les Hôtes dédiés et les Instances dédiées :

	Hôte dédié	Instance dédiée
Facturation	Facturation par hôte	Facturation par instance
Visibilité des sockets, cœurs et ID d'hôte	Offre une visibilité sur le nombre de sockets et de cœurs physiques	Aucune visibilité
Affinité de l'hôte et de l'instance	Permet de déployer vos instances de façon cohérente sur le même serveur physique au fil du temps	Non pris en charge
Placement ciblé d'instances	Offre une visibilité supplémentaire et un contrôle sur la façon dont les instances sont placées sur un serveur physique	Non pris en charge
Récupération automatique des instances	Pris en charge. Pour plus d'informations, consultez Récupération de l'hôte (p. 492).	Pris en charge
Bring Your Own License (Licence à fournir)	Pris en charge	Non pris en charge

Bring Your Own License (BYOL, licence à fournir)

Les Hôtes dédiés vous permettent d'utiliser vos licences logicielles existantes par socket, par cœur ou par machine virtuelle. Lorsque vous utilisez vos propres licences, vous êtes responsable de leur gestion. Toutefois, Amazon EC2 comporte des fonctionnalités qui vous aident à assurer la conformité de vos licences, telles que l'affinité d'instance et le placement ciblé.

Voici les grandes étapes que vous devez suivre afin d'utiliser votre propre image de machine virtuelle sous licence en volume dans Amazon EC2.

1. Assurez-vous que les termes du contrat de licence régissant l'utilisation de vos images de machine permettent l'utilisation dans un environnement cloud virtualisé.
2. Après avoir vérifié que votre image de machine peut être utilisée dans Amazon EC2, importez-la avec VM Import/Export. Pour plus d'informations sur la procédure à suivre pour importer votre image de machine, consultez le manuel VM Import/Export Guide de l'utilisateur.
3. Une fois votre image de machine importée, vous pouvez lancer des instances depuis cette image sur des Hôtes dédiés actifs de votre compte.

4. Lorsque vous exécutez ces instances, en fonction du système d'exploitation, vous pouvez être contraint d'activer ces instances sur votre propre serveur KMS.

   Note

   Pour suivre la façon dont vos images sont utilisées dans AWS, activez l'enregistrement de l'hôte dans AWS Config. Vous pouvez utiliser AWS Config pour enregistrer les changements de configuration d'un Hôte dédié et utiliser le résultat comme source de données pour les rapports d'utilisation des licences. Pour plus d'informations, consultez Suivi des changements de configuration (p. 497).

Capacité d'instance d'un Hôte dédié

La prise en charge de plusieurs types d'instance sur le même hôte dédié est disponible pour les familles d'instances suivantes : `c5`, `m5`, `r5`, `c5n`, `r5n` et `m5n`. Par exemple, lorsque vous allouez un Hôte dédié `r5`, vous pouvez utiliser un hôte avec 2 sockets et 48 cœurs physiques sur lesquels vous pouvez exécuter différents types d'instances, tels que `r5.2xlarge` et `r5.4xlarge`. Vous pouvez exécuter n'importe quel nombre d'instances, dans la limite de la capacité de cœur associée à l'hôte. Par exemple, le tableau ci-dessous montre les différentes combinaisons de types d'instances que vous pouvez exécuter sur un Hôte dédié.

Famille d'instances	Exemples de combinaisons de types d'instances	
R5	• Exemple 1 : 4 x `r5.4xlarge` + 4 x `r5.2xlarge` • Exemple 2 : 1 x `r5.12xlarge` + 1 x `r5.4xlarge` + 1 x `r5.2xlarge` + 5 x `r5.xlarge` + 2 x `r5.large`	
C5	• Exemple 1 : 1 x `c5.9xlarge` + 2 x `c5.4xlarge` + 1 x `c5.xlarge` • Exemple 2 : 4 x `c5.4xlarge` + 1 x `c5.xlarge` + 2 x `c5.large`	
M5	• Exemple 1 : 4 x `m5.4xlarge` + 4 x `m5.2xlarge` • Exemple 2 : 1 x `m5.12xlarge` + 1 x `m5.4xlarge` + 1 x `m5.2xlarge` + 5 x `m5.xlarge` + 2 x `m5.large`	

Les autres familles d'instances ne prennent en charge qu'un seul type d'instance sur le même Hôte dédié. Pour plus d'informations sur les familles d'instances et les configurations de type d'instance prises en charge sur Hôtes dédiés, veuillez consulter Tarification Amazon EC2 Hôte dédié.

Restrictions Hôtes dédiés

Avant d'allouer des Hôtes dédiés, prenez note des restrictions suivantes :

- Pour exécuter RHEL, SUSE Linux et SQL Server sur Hôtes dédiés, vous devez apporter vos propres AMI. Les AMI RHEL, SUSE Linux et SQL Server proposées par AWS ou disponibles sur AWS Marketplace ne peuvent pas être utilisées avec Hôtes dédiés. Pour plus d'informations sur la création de votre propre AMI, veuillez consulter Bring Your Own License (BYOL, licence à fournir) (p. 466).
- Il est possible d'allouer jusqu'à deux Hôtes dédiés à la demande par famille d'instance et par région. Il est possible de faire une demande d'augmentation de limite : Demande d'augmentation de la limite d'allocation sur des Hôtes dédiés Amazon EC2.
- Les instances qui fonctionnent sur un Hôte dédié ne peuvent être lancées que dans un VPC.
- Les groupes Auto Scaling sont pris en charge lors de l'utilisation d'un modèle de lancement qui spécifie un groupe de ressources hôte. Pour de plus amples informations, veuillez consulter Création d'un

modèle de lancement pour un groupe Auto Scaling dans le Guide de l'utilisateur Amazon EC2 Auto Scaling.

- Les instances Amazon RDS ne sont pas prises en charge.
- Le niveau d'offre gratuite AWS n'est pas disponible pour un Hôtes dédiés.
- Le contrôle de placement d'instance fait référence à la gestion du lancement d'instances sur les Hôtes dédiés. Vous ne pouvez pas lancer Hôtes dédiés dans des groupes de placement.

Tarification et facturation

Le prix d'un Hôte dédié varie selon l'option de paiement.

Options de paiement

- Hôtes dédiés à la demande (p. 468)
- Réservations d'hôtes dédiés (p. 468)
- Plans d'économies (p. 469)
- Tarification pour Windows Server sur les Hôtes dédiés (p. 469)

Hôtes dédiés à la demande

La facturation à la demande est automatiquement activée lorsque vous allouez un Hôte dédié à votre compte.

Le prix à la demande pour un Hôte dédié varie par famille d'instance et par région. Vous payez par seconde (avec un minimum de 60 secondes) pour l'Hôte dédié actif, quelle que soit la quantité ou la taille des instances que vous choisissez de lancer dessus. Pour de plus amples informations sur la tarification à la demande, veuillez consulter Tarification à la demande des Hôtes dédiés Amazon EC2.

Vous pouvez libérer un Hôte dédié à la demande à tout moment pour arrêter d'accumuler des frais dessus. Pour plus d'informations sur la libération d'un Hôte dédié, consultez Libération d'Hôtes dédiés (p. 483).

Réservations d'hôtes dédiés

Les Réservations d'hôtes dédiés permettent de bénéficier d'une remise sur la facturation par rapport à l'exécution d'Hôtes dédiés à la demande. Trois options de paiement sont disponibles pour les réservations :

- Aucun paiement initial — Les réservations sans aucun paiement initial vous offrent une remise sur votre utilisation d'un Hôte dédié pendant une période donnée et ne nécessitent aucun paiement initial. Disponible pour une période d'un an uniquement.
- Paiement initial partiel — Une partie de la réservation doit être payée au départ et les heures restantes pendant la période sont facturées à un tarif réduit. Disponible pour une période d'un an ou de trois ans.
- Paiement initial complet — Offre le coût effectif le plus bas. Disponible pour une période d'un an et de trois ans et couvre le coût intégral de la période à l'avance, sans plus aucuns frais supplémentaire futurs.

Vous devez disposer d'un Hôtes dédiés actif sur votre compte pour pouvoir acheter des réservations. Chaque réservation concerne un Hôte dédié spécifique et unique sur votre compte. Les réservations sont appliquées à la famille d'instance sur l'hôte, et non à la taille de l'instance. Si vous avez trois Hôtes dédiés avec des tailles d'instance différentes (`m4.xlarge`, `m4.medium` et `m4.large`), vous pouvez associer une même réservation `m4` à tous ces Hôtes dédiés. La famille d'instance et la région de la réservation doivent correspondre à celles des hôtes dédiés que vous souhaitez leur associer.

Lorsqu'une réservation est associée à un Hôte dédié, cet Hôte dédié ne peut pas être libéré avant la fin de la période de la réservation.

Pour plus d'informations sur la tarification de réservation, consultez Tarification d'un Hôtes dédiés Amazon EC2.

Plans d'économies

Les plans d'économies sont un modèle de tarification flexible qui offre des économies importantes par rapport aux Instances à la demande. Avec les plans d'économies, vous pouvez vous engager pour une utilisation continue, en USD par heure, pour une durée de un à trois ans. Cela vous donne la flexibilité d'utiliser le Hôtes dédiés répondant le mieux à vos besoins et de continuer à économiser de l'argent au lieu de vous engager pour un Hôte dédié spécifique. Pour de plus amples informations, veuillez consulter le Guide de l'utilisateur des plans d'économies AWS.

Tarification pour Windows Server sur les Hôtes dédiés

Sous réserve des conditions de licence de Microsoft, vous pouvez importer vos licences Windows Server et SQL Server existantes vers les Hôtes dédiés. Aucuns frais supplémentaires ne s'applique à l'utilisation du logiciel si vous choisissez de réutiliser vos licences.

Vous pouvez également utiliser les AMI Windows Server fournies par Amazon pour exécuter les dernières versions de Windows Server sur des Hôtes dédiés. Il s'agit d'une pratique courante pour les scénarios où vous possédez des licences SQL Server existantes pouvant être exécutées sur des Hôtes dédiés, mais qui ont besoin de Windows Server pour exécuter la charge de travail SQL Server. Les AMI Windows Server fournies par Amazon sont prises en charge uniquement sur les types d'instances de génération actuelle (p. 214). Pour de plus amples informations, veuillez consulter Tarification des hôtes dédiés Amazon EC2.

Utilisation d'Hôtes dédiés

Pour utiliser un Hôte dédié, vous devez d'abord allouer des hôtes à utiliser sur votre compte. Vous pouvez ensuite lancer des instances sur les hôtes en spécifiant une location d'hôte pour l'instance. Vous devez sélectionner un hôte spécifique sur lequel lancer l'instance, ou vous pouvez autoriser son lancement sur n'importe quel hôte sur lequel le placement automatique est activé et qui correspond à son type d'instance. Lorsqu'une instance est arrêtée puis redémarrée, le paramètre d'affinité de l'hôte détermine si elle est redémarrée sur le même hôte ou sur un autre.

Si vous n'avez plus besoin d'un hôte à la demande, vous pouvez arrêter les instances en cours d'exécution sur celui-ci, configurer leur lancement sur un autre hôte, puis libérer l'hôte.

Les Hôtes dédiés sont également intégrés à AWS License Manager. Grâce à License Manager, vous pouvez créer un groupe de ressources hôte, qui est un ensemble d'Hôtes dédiés gérés en tant qu'entité unique. Lors de la création d'un groupe de ressources hôte, vous spécifiez les préférences de gestion de l'hôte, telles que l'allocation automatique et la libération automatique, pour les Hôtes dédiés. Vous pouvez ainsi lancer des instances sur les Hôtes dédiés sans allouer ni gérer manuellement ces hôtes. Pour de plus amples informations, veuillez consulter Groupes de ressources hôte dans le Guide de l'utilisateur AWS License Manager.

Table des matières

- Allocation d'Hôtes dédiés (p. 470)
- Lancement d'instances sur un Hôte dédié (p. 472)
- Lancement d'instances dans un groupe de ressources hôte (p. 474)
- Présentation du placement automatique et de l'affinité (p. 475)
- Modification du placement automatique d'un Hôte dédié (p. 475)
- Modification des types d'instances pris en charge. (p. 476)
- Modification de l'affinité et de la location d'une instance (p. 478)
- Affichage des Hôtes dédiés (p. 480)

- Balisage des entités Hôte dédié (p. 481)
- Surveillance des Hôtes dédiés (p. 482)
- Libération d'Hôtes dédiés (p. 483)
- Achat d'Réservations d'hôtes dédiés (p. 484)
- Affichage de réservations d'Hôte dédié. (p. 486)
- Balisage d'Réservations d'hôtes dédiés (p. 487)

Allocation d'Hôtes dédiés

Pour commencer à utiliser des Hôtes dédiés, vous devez allouer des Hôtes dédiés à votre compte à l'aide de la console Amazon EC2 ou des outils de ligne de commande. Lorsque vous allouez l'Hôte dédié, la capacité de l'Hôte dédié est immédiatement mise à disposition dans votre compte et vous pouvez commencer à lancer des instances sur l'Hôte dédié.

La prise en charge de plusieurs types d'instance sur le même hôte dédié est disponible pour les familles d'instances suivantes : `c5`, `m5`, `r5`, `c5n`, `r5n` et `m5n`. Les autres familles d'instances ne prennent en charge qu'un seul type d'instance sur le même Hôte dédié.

Vous pouvez allouer un Hôte dédié à l'aide des méthodes suivantes.

Nouvelle console

Pour allouer un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Hôtes dédiés, puis Allouer un Hôte dédié.
3. Pour Famille d'instances, choisissez la famille d'instance de l'Hôte dédié.
4. Indiquez si l'Hôte dédié prend en charge plusieurs types d'instances dans la famille d'instances sélectionnée ou uniquement un type d'instance spécifique. Effectuez l'une des actions suivantes :
   - Pour configurer l'Hôte dédié afin qu'il prenne en charge plusieurs types d'instances dans la famille d'instances sélectionnée, pour Support multiple instance types (Prendre en charge plusieurs types d'instances), choisissez Activer. La sélection de cette option vous permet de lancer différents types d'instances d'une même famille d'instances sur l'Hôte dédié. Par exemple, si vous choisissez la famille d'instances `m5` et que vous choisissez cette option, vous pouvez lancer les instances `m5.xlarge` et `m5.4xlarge` sur l'Hôte dédié.
   - Pour configurer lHôte dédié afin qu'il prenne en charge un type d'instance spécifique dans la famille d'instances sélectionnée, désélectionnez Support multiple instance types (Prendre en charge plusieurs types d'instances), puis, dans Instance type (Type d'instance), choisissez le type d'instance à prendre en charge. Cette action vous permet de lancer un seul type d'instance sur l'Hôte dédié. Par exemple, si vous choisissez cette option et spécifiez `m5.4xlarge` comme type d'instance pris en charge, vous pouvez uniquement lancer des instances `m5.4xlarge` sur l'Hôte dédié.
5. Pour Zone de disponibilité, choisissez la zone de disponibilité dans laquelle allouer l'Hôte dédié.
6. Pour autoriser l'Hôte dédié à accepter les lancements d'instance non ciblés correspondant à son type d'instance, pour Instance auto-placement (Placement automatique d'instance), choisissez Enable (Autoriser). Pour en savoir plus sur le placement automatique, consultez Présentation du placement automatique et de l'affinité (p. 475).
7. Pour autoriser la récupération d'hôte pour l'Hôte dédié, pour Host recovery (Récupération de l'hôte), choisissez Activer. Pour de plus amples informations, veuillez consulter Récupération de l'hôte (p. 492).
8. Pour Quantité, entrez le nombre d'Hôtes dédiés à allouer.
9. (Facultatif) Choisissez Ajouter une balise et entrez une clé et une valeur de balise.

10. Choisissez Allocate.

Ancienne console

Pour allouer un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Hôtes dédiés, Allouer Hôte dédié.
3. Pour Famille d'instances, choisissez la famille d'instance de l'Hôte dédié.
4. Indiquez si l'Hôte dédié prend en charge plusieurs types d'instances dans la famille d'instances sélectionnée ou uniquement un type d'instance spécifique. Effectuez l'une des actions suivantes :
   - Pour configurer l'Hôte dédié afin qu'il prenne en charge plusieurs types d'instances dans la famille d'instances sélectionnée, sélectionnez Support multiple instance types (Prendre en charge plusieurs types d'instances). La sélection de cette option vous permet de lancer différents types d'instances d'une même famille d'instances sur l'Hôte dédié. Par exemple, si vous choisissez la famille d'instances `m5` et que vous choisissez cette option, vous pouvez lancer les instances `m5.xlarge` et `m5.4xlarge` sur l'Hôte dédié. La famille d'instances doit être alimentée par le système Nitro.
   - Pour configurer lHôte dédié afin qu'il prenne en charge un type d'instance spécifique dans la famille d'instances sélectionnée, désélectionnez Support multiple instance types (Prendre en charge plusieurs types d'instances), puis, dans Instance type (Type d'instance), choisissez le type d'instance à prendre en charge. Cette action vous permet de lancer un seul type d'instance sur l'Hôte dédié. Par exemple, si vous choisissez cette option et spécifiez `m5.4xlarge` comme type d'instance pris en charge, vous pouvez uniquement lancer des instances `m5.4xlarge` sur l'Hôte dédié.
5. Pour Zone de disponibilité, choisissez la zone de disponibilité dans laquelle allouer l'Hôte dédié.
6. Pour autoriser l'Hôte dédié à accepter les lancements d'instance non ciblés correspondant à son type d'instance, pour Instance auto-placement (Placement automatique d'instance), choisissez Enable (Autoriser). Pour en savoir plus sur le placement automatique, consultez Présentation du placement automatique et de l'affinité (p. 475).
7. Pour autoriser la récupération d'hôte pour l'Hôte dédié, pour Host recovery (Récupération de l'hôte), choisissez Enable (Autoriser). Pour de plus amples informations, veuillez consulter Récupération de l'hôte (p. 492).
8. Pour Quantité, entrez le nombre d'Hôtes dédiés à allouer.
9. (Facultatif) Choisissez Ajouter une balise et entrez une clé et une valeur de balise.
10. Choisissez Allouer un hôte.

AWS CLI

Pour allouer un Hôte dédié

Utilisez la commande allocate-hosts de l'AWS CLI. La commande suivante alloue un Hôte dédié qui prend en charge plusieurs types d'instances de la famille d'instances `m5` dans la zone de disponibilité `us-east-1a`. La fonction de récupération de l'hôte est activée et la fonction de placement automatique est désactivée sur l'hôte.

```
aws ec2 allocate-hosts --instance-family "m5" --availability-zone "us-east-1a" --auto-
placement "off" --host-recovery "on" --quantity 1
```

La commande suivante alloue un Hôte dédié qui prend en charge des lancements d'instance `m4.large` non ciblés dans la zone de disponibilité `eu-west-1a`, autorise la récupération de l'hôte et applique une balise avec une clé `purpose` et une valeur `production`.

```
aws ec2 allocate-hosts --instance-type "m4.large" --availability-zone "eu-west-1a"
 --auto-placement "on" --host-recovery "on" --quantity 1 --tag-specifications
 'ResourceType=dedicated-host,Tags=[{Key=purpose,Value=production}]'
```

PowerShell

Pour allouer un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Host. La commande suivante alloue un Hôte dédié qui prend en charge plusieurs types d'instances de la famille d'instances `m5` dans la zone de disponibilité `us-east-1a`. La fonction de récupération de l'hôte est activée et la fonction de placement automatique est désactivée sur l'hôte.

```
PS C:\> New-EC2Host -InstanceFamily m5 -AvailabilityZone us-east-1a -AutoPlacement Off
 -HostRecovery On -Quantity 1
```

Les commandes suivantes allouent un Hôte dédié qui prend en charge des lancements d'instance non ciblés `m4.large` dans la zone de disponibilité `eu-west-1a` et appliquent une balise avec une clé `purpose` et une valeur `production`.

Le paramètre `TagSpecification` utilisé pour baliser un Hôte dédié à la création requiert un objet qui spécifie le type de ressource à baliser, ainsi que la clé et la valeur de balise. Les commandes suivantes permettent de créer l'objet requis.

```
PS C:\> $tag = @{ Key="purpose"; Value="production" }
PS C:\> $tagspec = new-object Amazon.EC2.Model.TagSpecification
PS C:\> $tagspec.ResourceType = "dedicated-host"
PS C:\> $tagspec.Tags.Add($tag)
```

La commande suivante alloue le Hôte dédié et applique la balise spécifiée dans l'objet `$tagspec`.

```
PS C:\> New-EC2Host -InstanceType m4.large -AvailabilityZone eu-west-1a -
AutoPlacement On -HostRecovery On -Quantity 1 -TagSpecification $tagspec
```

Lancement d'instances sur un Hôte dédié

Une fois que vous avez alloué un Hôte dédié, vous pouvez lancer des instances sur cet hôte. Vous ne pouvez pas lancer des instances avec la location `host` si vous n'avez pas d'Hôtes dédiés actifs avec suffisamment de capacité disponible pour le type d'instance que vous lancez.

Note

Les instances lancées sur un Hôtes dédiés ne peuvent être lancées que dans un VPC. Pour plus d'informations, consultez Présentation de VPC.

Avant de lancer vos instances, prenez note des restrictions. Pour plus d'informations, consultez Restrictions Hôtes dédiés (p. 467).

Vous pouvez lancer une instance dans un Hôte dédié à l'aide des méthodes suivantes.

Console

Pour lancer une instance sur un Hôte dédié spécifique depuis la page Hôtes dédiés

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Hôtes dédiés dans le volet de navigation.
3. Dans la page Hôtes dédiés, sélectionnez un hôte et choisissez Actions, Lancement d'une ou de plusieurs instances sur l'hôte.

4. Sélectionnez une AMI dans la liste. Les AMI SQL Server, SUSE et RHEL fournies par Amazon EC2 ne peuvent pas être utilisées avec des Hôtes dédiés.
5. Sur la page Choisir un type d'instance, sélectionnez le type d'instance à lancer, puis choisissez Suivant : Configurer les détails de l'instance.

   Si l'Hôte dédié ne prend en charge qu'un seul type d'instance, ce type est sélectionné par défaut et ne peut pas être modifié.

   Si l'Hôte dédié prend en charge plusieurs types d'instances, vous devez sélectionner un type d'instance dans la famille d'instances prise en charge en fonction de la capacité d'instance disponible de l'Hôte dédié. Nous vous recommandons de lancer d'abord les tailles d'instance plus grandes, puis de remplir la capacité d'instance restante avec les tailles d'instance plus petites, si nécessaire.
6. Sur la page Configurer les détails de l'instance, configurez les paramètres d'instance en fonction de vos besoins, puis, pour Affinité, choisissez l'une des options suivantes :

   - Désactivé — L'instance est lancée sur l'hôte spécifié, mais il n'est pas garanti qu'elle redémarre sur le même Hôte dédié si elle est arrêtée.
   - Hôte — Si l'instance est arrêtée, elle redémarre toujours sur cet hôte spécifique.

   Pour en savoir plus sur l'affinité, consultez Présentation du placement automatique et de l'affinité (p. 475).

   Les options Location et Hôte sont préconfigurées en fonction de l'hôte que vous avez sélectionné.
7. Choisissez Vérifier et lancer.
8. Sur la page Review Instance Launch, sélectionnez Launch.
9. Lorsque vous y êtes invité, sélectionnez une paire de clés existante ou créez-en une autre, puis choisissez Lancer des instances.

Pour lancer une instance sur un Hôte dédié à l'aide de l'assistant de lancement d'instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, Launch Instance (Lancer une instance).
3. Sélectionnez une AMI dans la liste. Les AMI SQL Server, SUSE et RHEL fournies par Amazon EC2 ne peuvent pas être utilisées avec des Hôtes dédiés.
4. Sélectionnez le type d'instance à lancer, puis choisissez Suivant : Configurer les détails de l'instance.
5. Sur la page Configurer les détails de l'instance, configurez les paramètres d'instance en fonction de vos besoins, puis configurez les paramètres suivants, qui sont spécifiques à l'Hôte dédié :

   - Location — Choisissez Hôte dédié - Lancez cette instance sur un Hôte dédié.
   - Hôte : choisissez Utiliser le placement automatique pour lancer l'instance sur n'importe quel Hôte dédié sur lequel le placement automatique est activé, ou sélectionnez un Hôte dédié spécifique dans la liste. La liste affiche uniquement les Hôtes dédiés qui prennent en charge le type d'instance sélectionné.
   - Affinité — Choisissez l'une des options suivantes :
     - Désactivé — L'instance est lancée sur l'hôte spécifié, mais il n'est pas garanti qu'elle redémarre sur celui-ci si elle est arrêtée.
     - Hôte — Si l'instance est arrêtée, elle redémarre toujours sur l'hôte spécifié.

   Pour plus d'informations, consultez Présentation du placement automatique et de l'affinité (p. 475).

Si vous ne pouvez pas voir ces paramètres, vérifiez que vous avez sélectionné un VPC dans le menu Network.

6. Choisissez Vérifier et lancer.
7. Sur la page Review Instance Launch, sélectionnez Launch.
8. Lorsque vous y êtes invité, sélectionnez une paire de clés existante ou créez-en une autre, puis choisissez Lancer des instances.

AWS CLI

Pour lancer une instance dans un Hôte dédié

Utilisez la commande run-instances de l'AWS CLI et spécifiez l'affinité de l'instance, la location et l'hôte dans le paramètre de demande `Placement`.

PowerShell

Pour lancer une instance dans un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Instance et spécifiez l'affinité de l'instance, la location et l'hôte dans le paramètre de demande `Placement`.

Lancement d'instances dans un groupe de ressources hôte

Lorsque vous lancez une instance dans un groupe de ressources hôte qui contient un Hôte dédié avec une capacité d'instance disponible, Amazon EC2 lance l'instance sur cet hôte. Si le groupe de ressources hôte ne contient pas d'hôte avec une capacité d'instance disponible, Amazon EC2 alloue automatiquement un nouvel hôte dans le groupe de ressources hôte, puis lance l'instance sur cet hôte. Pour de plus amples informations, veuillez consulter Groupes de ressources hôte dans le Guide de l'utilisateur AWS License Manager.

Exigences et limites

- Vous devez associer une configuration de licence basée sur le cœur/socket à l'AMI.
- Vous ne pouvez pas utiliser les AMI SQL Server, SUSE ou RHEL fournies par Amazon EC2 avec Hôtes dédiés.
- Vous ne pouvez pas cibler un hôte spécifique en choisissant un ID d'hôte et vous ne pouvez pas activer l'affinité d'instance lors du lancement d'une instance dans un groupe de ressources hôte.

Vous pouvez lancer une instance dans un groupe de ressources hôte à l'aide des méthodes suivantes.

Console

Pour lancer une instance dans un groupe de ressources hôte

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, Launch Instance (Lancer une instance).
3. Sélectionnez une AMI.
4. Sélectionnez le type d'instance à lancer, puis choisissez Suivant : Configurer les détails de l'instance.
5. Sur la page Configurer les détails de l'instance, configurez les paramètres d'instance en fonction de vos besoins, puis procédez comme suit :

   a. Pour Location, choisissez Hôte dédié.
   b. Pour Host resource group (Groupe de ressources hôte), choisissez Launch instance into a host resource group (Lancer une instance dans un groupe de ressources hôte).

   c. Pour Host resource group name (Nom du groupe de ressources hôte), choisissez le groupe de ressources hôte dans lequel lancer l'instance.
6. Choisissez Vérifier et lancer.
7. Sur la page Review Instance Launch, sélectionnez Launch.
8. Lorsque vous y êtes invité, sélectionnez une paire de clés existante ou créez-en une autre, puis choisissez Lancer des instances.

AWS CLI

Pour lancer une instance dans un groupe de ressources hôte

Utilisez la commande run-instances de l'AWS CLI et, dans le paramètre de demande `Placement`, omettez l'option Location et spécifiez l'ARN du groupe de ressources hôte.

PowerShell

Pour lancer une instance dans un groupe de ressources hôte

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Instance et, dans le paramètre de demande `Placement`, omettez l'option Location et spécifiez l'ARN du groupe de ressources hôte.

Présentation du placement automatique et de l'affinité

Le contrôle de placement pour l'Hôtes dédiés est effectué au niveau de l'instance et au niveau de l'hôte.

Placement automatique

Le placement automatique est configuré au niveau de l'hôte. Il vous permet de définir si les instances que vous lancez le sont sur un hôte spécifique ou sur n'importe quel hôte disponible doté de configurations correspondantes.

Lorsque le placement automatique d'un Hôte dédié est désactivé, il n'accepte que les lancements d'instance de location d'hôte qui spécifient son ID d'hôte unique. Il s'agit du paramètre par défaut pour un nouvel Hôtes dédiés.

Lorsque le placement automatique d'un Hôte dédié est activé, il accepte tous les lancements d'instances non ciblés qui correspondent à la configuration de son type d'instance.

Lors du lancement d'une instance, vous devez configurer sa location. Le lancement d'une instance sur un Hôte dédié sans indiquer un `HostId` spécifique permet de la lancer sur n'importe quel Hôte dédié sur lequel le placement automatique est activé et qui correspond à son type d'instance.

Affinité de l'hôte

L'affinité de l'hôte est configurée au niveau de l'instance. Elle établit une relation de lancement entre une instance et un Hôte dédié.

Lorsque l'affinité a pour valeur `Host`, une instance lancée sur un hôte spécifique redémarre toujours sur le même hôte si elle est arrêtée. Cela s'applique aussi bien aux lancements ciblés qu'aux lancements non-ciblés.

Lorsque l'affinité a pour valeur `Off` et que vous arrêtez et redémarrez l'instance, cette dernière peut être redémarrée sur tout hôte disponible. Toutefois, elle essaie de se relancer sur le dernier Hôte dédié sur lequel elle s'est exécutée (dans la mesure du possible).

Modification du placement automatique d'un Hôte dédié

Vous pouvez modifier les paramètres de placement automatique d'un Hôte dédié après l'avoir alloué à votre compte AWS, en utilisant l'une des méthodes suivantes.

Nouvelle console

Pour modifier le placement automatique d'un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sélectionnez un hôte et choisissez Actions, Modify host (Modifier l'hôte).
4. Pour Instance auto-placement (Placement automatique de l'instance), choisissez Activer pour activer le placement automatique ou Désactiver pour désactiver le placement automatique. Pour plus d'informations, consultez Présentation du placement automatique et de l'affinité (p. 475).
5. Choisissez Enregistrer.

Ancienne console

Pour modifier le placement automatique d'un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Hôtes dédiés dans le volet de navigation.
3. Sur la page Hôtes dédiés, sélectionnez un hôte, puis choisissez Actions et Modify Auto-Placement (Modifier le placement automatique).
4. Dans la fenêtre Modifier le placement automatique, pour Permettre le placement automatique des instances, choisissez Oui pour activer le placement automatique, ou Non pour le désactiver. Pour plus d'informations, consultez Présentation du placement automatique et de l'affinité (p. 475).
5. Choisissez Enregistrer.

AWS CLI

Pour modifier le placement automatique d'un Hôte dédié

Utilisez la commande modify-hosts de l'AWS CLI. Les exemples suivants activent le placement automatique pour l'Hôte dédié spécifié.

```
aws ec2 modify-hosts --auto-placement on --host-ids h-012a3456b7890cdef
```

PowerShell

Pour modifier le placement automatique d'un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell Edit-EC2Host. Les exemples suivants activent le placement automatique pour l'Hôte dédié spécifié.

```
PS C:\> Edit-EC2Host --AutoPlacement 1 --HostId h-012a3456b7890cdef
```

Modification des types d'instances pris en charge.

La prise en charge de plusieurs types d'instance sur le même hôte dédié est disponible pour les familles d'instances suivantes : `c5`, `m5`, `r5`, `c5n`, `r5n` et `m5n`. Les autres familles d'instances ne prennent en charge qu'un seul type d'instance sur le même Hôte dédié.

Vous pouvez allouer un Hôte dédié à l'aide des méthodes suivantes.

Vous pouvez modifier un Hôte dédié afin de modifier les types d'instances qu'il prend en charge. S'il prend actuellement en charge un seul type d'instance, vous pouvez le modifier afin qu'il en prenne en charge

plusieurs dans cette famille d'instances. De même, s'il prend en charge plusieurs types d'instances, vous pouvez le modifier afin qu'il n'en prenne plus qu'un seul.

Pour modifier un Hôte dédié afin qu'il prenne en charge plusieurs types d'instances, vous devez d'abord arrêter toutes les instances en cours d'exécution sur l'hôte. Cette modification prend effet au bout d'environ 10 minutes. L'Hôte dédié passe à l'état `pending` pendant que la modification est en cours. Vous ne pouvez pas démarrer les instances arrêtées ou lancer de nouvelles instances sur l'Hôte dédié lorsqu'il est à l'état `pending`.

Pour qu'il soit possible de modifier un Hôte dédié prenant en charge plusieurs types d'instances afin qu'il n'en prenne plus qu'un seul, l'hôte ne doit avoir aucune instance en cours d'exécution ou les instances en cours d'exécution doivent être du type qui devra être pris en charge par l'hôte. Par exemple, pour modifier un hôte prenant en charge plusieurs types d'instances dans la famille d'instances `m5` afin qu'il ne prenne plus en charge que les instances `m5.large`, il faut que l'Hôte dédié n'ait aucune instance en cours d'exécution ou que seules des instances `m5.large` soient en cours d'exécution sur l'hôte.

Vous pouvez modifier les types d'instance pris en charge à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour modifier les types d'instance pris en charge pour un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôte dédié.
3. Sélectionnez l'Hôte dédié à modifier et choisissez Actions, Modify host (Modifier l'hôte).
4. Selon la configuration actuelle de l'Hôte dédié, procédez comme indiqué ci-après.

   - Si l'Hôte dédié prend actuellement en charge un type d'instance spécifique, l'option Support multiple instance types (Prendre en charge plusieurs types d'instance) n'est pas activée et la liste Type d'instance répertorie le type d'instance pris en charge. Pour modifier l'hôte afin qu'il prenne en charge plusieurs types d'instances dans la famille d'instances actuelle, pour Support multiple instance types (Prendre en charge plusieurs types d'instances), choisissez Activer.

     Pour modifier un hôte afin qu'il prenne en charge plusieurs types d'instances, vous devez d'abord arrêter toutes les instances en cours d'exécution sur l'hôte.
   - Si l'Hôte dédié prend actuellement en charge plusieurs types d'instances d'une famille, Activé est sélectionné pour Support multiple instance types (Prendre en charge plusieurs types d'instances). Pour modifier l'hôte afin qu'il prenne en charge un type d'instance spécifique, pour Support multiple instance types (Prendre en charge plusieurs types d'instances), décochez Activer, puis pour Type d'instance, sélectionnez le type d'instance spécifique à prendre en charge.

     Vous ne pouvez pas modifier la famille d'instances prise en charge par l'Hôte dédié.
5. Choisissez Enregistrer.

Ancienne console

Pour modifier les types d'instance pris en charge pour un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôte dédié.
3. Sélectionnez l'Hôte dédié à modifier et choisissez Actions, Modify Supported Instance Types (Modifier les types d'instances pris en charge).
4. Selon la configuration actuelle de l'Hôte dédié, procédez comme indiqué ci-après.

   - Si l'Hôte dédié prend actuellement en charge un type d'instance spécifique, Non est sélectionné pour Support multiple instance types (Prendre en charge plusieurs types d'instances). Pour

modifier l'hôte afin qu'il prenne en charge plusieurs types d'instances dans la famille d'instances actuelle, pour Support multiple instance types (Prendre en charge plusieurs types d'instances), sélectionnez Oui.

Pour modifier un hôte afin qu'il prenne en charge plusieurs types d'instances, vous devez d'abord arrêter toutes les instances en cours d'exécution sur l'hôte.

- Si l'Hôte dédié prend actuellement en charge plusieurs types d'instances d'une famille, Oui est sélectionné pour Support multiple instance types (Prendre en charge plusieurs types d'instances) et Famille d'instances affiche la famille d'instances prise en charge. Pour modifier l'hôte afin qu'il prenne en charge un type d'instance spécifique, pour Support multiple instance types (Prendre en charge plusieurs types d'instances), sélectionnez Non, puis pour Type d'instance, sélectionnez le type d'instance spécifique à prendre en charge.

  Vous ne pouvez pas modifier la famille d'instances prise en charge par l'Hôte dédié.

5. Choisissez Enregistrer.

AWS CLI

Pour modifier les types d'instance pris en charge pour un Hôte dédié

Utilisez la commande modify-hosts de l'AWS CLI.

La commande suivante modifie un Hôte dédié afin qu'il prenne en charge plusieurs types d'instances au sein de la famille d'instances `m5`.

```
aws ec2 modify-hosts --instance-family m5 --host-ids h-012a3456b7890cdef
```

La commande suivante modifie un Hôte dédié afin qu'il prenne uniquement en charge les instances `m5.xlarge`.

```
aws ec2 modify-hosts --instance-type m5.xlarge --instance-family --host-
ids h-012a3456b7890cdef
```

PowerShell

Pour modifier les types d'instance pris en charge pour un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell Edit-EC2Host.

La commande suivante modifie un Hôte dédié afin qu'il prenne en charge plusieurs types d'instances au sein de la famille d'instances `m5`.

```
PS C:\> Edit-EC2Host --InstanceFamily m5 --HostId h-012a3456b7890cdef
```

La commande suivante modifie un Hôte dédié afin qu'il prenne uniquement en charge les instances `m5.xlarge`.

```
PS C:\> Edit-EC2Host --InstanceType m5.xlarge --HostId h-012a3456b7890cdef
```

Modification de l'affinité et de la location d'une instance

Vous pouvez modifier la location d'une instance en remplaçant `dedicated` par `host` ou `host` par `dedicated` après l'avoir lancée. Vous pouvez également modifier l'affinité entre l'instance et l'hôte. Pour qu'il soit possible de modifier l'affinité ou la location de l'instance, il faut que l'instance soit à l'état `stopped`.

Vous pouvez modifier la location et l'affinité d'une instance à l'aide des méthodes suivantes.

Console

Pour modifier la location d'instance ou l'affinité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instances, puis sélectionnez l'instance à modifier.
3. Choisissez État de l'instance, Arrêter.
4. Ouvrez le menu contextuel (clic droit) de l'instance et sélectionnez Instance Settings, puis Modify Instance Placement.
5. Sur la page Modifier le placement d'instance, configurez les éléments suivants :
   - Location — Choisissez l'une des options suivantes :
     - Exécuter une instance matérielle dédiée — Lance l'instance en tant qu'Instance dédiée. Pour plus d'informations, consultez Instances dédiées (p. 498).
     - Launch the instance on a Hôte dédié — Lance l'instance sur un Hôte dédié avec une affinité configurable.
   - Affinité — Choisissez l'une des options suivantes :
     - Cette instance peut être exécutée sur un de mes hôtes — L'instance est lancée sur n'importe quel Hôte dédié disponible de votre compte prenant en charge son type d'instance.
     - Cette instance ne peut être exécutée que sur l'hôte sélectionné — L'instance ne peut s'exécuter que sur l'Hôte dédié sélectionné pour Hôte cible.
   - Hôte cible — Sélectionnez l'Hôte dédié sur lequel l'instance doit s'exécuter. Si aucun hôte cible n'est répertorié, cela signifie que vous n'avez peut-être aucun Hôtes dédiés compatible disponible dans votre compte.

   Pour plus d'informations, consultez Présentation du placement automatique et de l'affinité (p. 475).
6. Choisissez Save.

AWS CLI

Pour modifier la location d'instance ou l'affinité

Utilisez la commande modify-instance-placement de l'AWS CLI. Les exemples suivants remplacent l'affinité de l'instance spécifiée `default` par `host`, et indiquent l'Hôte dédié avec lequel l'instance a une affinité.

```
aws ec2 modify-instance-placement --instance-id i-1234567890abcdef0 --affinity host --
host-id h-012a3456b7890cdef
```

PowerShell

Pour modifier la location d'instance ou l'affinité

Utilisez la commande Outils AWS pour Windows PowerShell Edit-EC2InstancePlacement. Les exemples suivants remplacent l'affinité de l'instance spécifiée `default` par `host`, et indiquent l'Hôte dédié avec lequel l'instance a une affinité.

```
PS C:\> Edit-EC2InstancePlacement -InstanceId i-1234567890abcdef0 -Affinity host -
HostId h-012a3456b7890cdef
```

Affichage des Hôtes dédiés

Vous pouvez afficher des détails sur un Hôte dédié et les instances individuelles qui s'y trouvent à l'aide des méthodes suivantes.

Nouvelle console

Pour afficher les détails d'un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sur la page Hôtes dédiés, sélectionnez un hôte.
4. Pour de plus amples informations sur l'hôte, choisissez Détails.

   Available vCPUs (UC virtuelles disponibles) indique les UC virtuelles (vCPU) qui sont disponibles sur l'Hôte dédié pour les nouveaux lancements d'instances. Par exemple, un Hôte dédié prenant en charge plusieurs types d'instances dans la famille d'instances `c5` et ne possédant aucune instance en cours d'exécution possède 72 UC virtuelles disponibles. Cela signifie que vous pouvez lancer différentes combinaisons de types d'instances sur l'Hôte dédié pour consommer les 72 UC virtuelles disponibles.

   Pour obtenir des informations sur les instances en cours d'exécution sur l'hôte, choisissez Instances en cours d'exécution.

Ancienne console

Pour afficher les détails d'un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sur la page Hôtes dédiés, sélectionnez un hôte.
4. Pour plus d'informations sur l'hôte, choisissez Description. Available vCPUs (UC virtuelles disponibles) indique les UC virtuelles (vCPU) qui sont disponibles sur l'Hôte dédié pour les nouveaux lancements d'instances. Par exemple, un Hôte dédié prenant en charge plusieurs types d'instances dans la famille d'instances `c5` et ne possédant aucune instance en cours d'exécution possède 72 UC virtuelles disponibles. Cela signifie que vous pouvez lancer différentes combinaisons de types d'instances sur l'Hôte dédié pour consommer les 72 UC virtuelles disponibles.

   Pour obtenir des informations sur les instances en cours d'exécution sur l'hôte, choisissez Instances.

AWS CLI

Pour afficher la capacité d'un Hôte dédié

Utilisez la commande describe-hosts de l'AWS CLI.

L'exemple suivant utilise la commande describe-hosts (AWS CLI) pour afficher la capacité d'instance disponible d'un Hôte dédié prenant en charge plusieurs types d'instances au sein de la famille d'instances `c5`. L'Hôte dédié possède déjà deux instances `c5.4xlarge` et quatre instances `c5.2xlarge` en cours d'exécution.

```
$  aws ec2 describe-hosts --host-id h-012a3456b7890cdef
```

```
"AvailableInstanceCapacity": [
    { "AvailableCapacity": 2,
      "InstanceType": "c5.xlarge",
      "TotalCapacity": 18 },
    { "AvailableCapacity": 4,
      "InstanceType": "c5.large",
      "TotalCapacity": 36 }
  ],
"AvailableVCpus": 8
```

PowerShell

Pour afficher la capacité d'instance d'un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell Get-EC2Host.

```
PS C:\> Get-EC2Host -HostId h-012a3456b7890cdef
```

Balisage des entités Hôte dédié

Vous pouvez allouer des balises personnalisées à vos entités Hôte dédié existantes pour classer celles-ci de différentes façons, par exemple, par objectif, par propriétaire ou par environnement. Cela vous permet de trouver rapidement un Hôte dédié spécifique en fonction des balises personnalisées que vous avez assignées. Les balises d'Hôte dédié peuvent également être utilisées pour le suivi de la répartition des coûts.

Vous pouvez aussi appliquer des balises aux Hôtes dédiés lors de la création. Pour plus d'informations, consultez Allocation d'Hôtes dédiés (p. 470).

Vous pouvez attribuer des balises à un Hôte dédié à l'aide des méthodes suivantes.

Nouvelle console

Pour attribuer des balises à un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sélectionnez l'Hôte dédié auquel attribuer des balises, puis choisissez Actions, Gérer les balises.
4. Dans l'écran Gérer les balises, choisissez Ajouter la balise, puis spécifiez la clé et la valeur de la balise.
5. (Facultatif) Choisissez Ajouter la balise pour ajouter des balises supplémentaires à l'Hôte dédié.
6. Sélectionnez Save Changes.

Ancienne console

Pour attribuer des balises à un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sélectionnez l'Hôte dédié à baliser et choisissez Balises.
4. Sélectionnez Ajouter/Modifier des balises.
5. Dans la boîte de dialogue Ajouter/Modifier des balises, choisissez Créer une balise, puis spécifiez la clé et la valeur de la balise.

6. (Facultatif) Choisissez Create Tag (Créer une balise)pour ajouter des balises supplémentaires au Hôte dédié.
7. Choisissez Enregistrer.

AWS CLI

Pour attribuer des balises à un Hôte dédié

Utilisez la commande create-tags de l'AWS CLI.

La commande suivante ajoute la balise `Owner=TeamA` à l'Hôte dédié spécifié.

```
aws ec2 create-tags --resources h-abc12345678909876 --tags Key=Owner,Value=TeamA
```

PowerShell

Pour attribuer des balises à un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Tag.

La commande `New-EC2Tag` nécessite un objet `Tag`, qui spécifie la paire clé-valeur à utiliser pour la balise d'Hôte dédié. Les commandes suivantes créent un objet `Tag` nommé `$tag` avec une paire clé-valeur `Owner` et `TeamA` :

```
PS C:\> $tag = New-Object Amazon.EC2.Model.Tag
PS C:\> $tag.Key = "Owner"
PS C:\> $tag.Value = "TeamA"
```

La commande suivante balise l'Hôte dédié spécifié avec l'objet `$tag` :

```
PS C:\> New-EC2Tag -Resource h-abc12345678909876 -Tag $tag
```

Surveillance des Hôtes dédiés

Amazon EC2 surveille constamment l'état de vos Hôtes dédiés. Les mises à jour sont communiquées sur la console Amazon EC2. Vous pouvez afficher des informations sur un Hôte dédié à l'aide des méthodes suivantes.

Console

Pour afficher l'état d'un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Recherchez l'Hôte dédié dans la liste et consultez la valeur située dans la colonne État.

AWS CLI

Pour afficher l'état d'un Hôte dédié

Utilisez la commande describe-hosts de l'AWS CLI, puis passez en revue la propriété `state` dans l'élément de réponse `hostSet`.

```
aws ec2 describe-hosts --host-id h-012a3456b7890cdef
```

PowerShell

Pour afficher l'état d'un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell Get-EC2Host, puis passez en revue la propriété `state` dans l'élément de réponse `hostSet`.

```
PS C:\> Get-EC2Host -HostId h-012a3456b7890cdef
```

Le tableau suivant explique les états possibles pour l'Hôte dédié.

État	Description
`available`	AWS n'a détecté aucun problème avec l'Hôte dédié. Aucune maintenance ni réparation n'est programmée. Les instances peuvent être lancées sur cet hôte dédié.
`released`	L'Hôte dédié a été libéré. l'ID de l'hôte n'est plus utilisé. Les hôtes libérés ne peuvent pas être réutilisés.
`under-assessment`	AWS étudie un problème potentiel avec l'Hôte dédié. Si une action doit être mise en place, vous recevez une notification via AWS Management Console ou par e-mail. Aucune instance ne peut être lancée sur un Hôte dédié dans cet état.
`pending`	L'Hôte dédié ne peut pas être utilisé le lancement de nouvelles instances. Soit il est en cours de modification afin de prendre en charge plusieurs types d'instances (p. 476), soit une récupération d'hôte (p. 492) est en cours.
`permanent-failure`	Une défaillance irrécupérable a été détectée. Vous recevez une notice d'expulsion par l'intermédiaire de vos instances et par e-mail. Vos instances peuvent continuer à s'exécuter. Si vous arrêtez ou mettez fin à toutes les instances sur un Hôte dédié dans cet état, AWS supprimera l'hôte. AWS ne redémarre pas les instances dans cet état. Aucune instance ne peut être lancée sur un Hôtes dédiés dans cet état.
`released-permanent-failure`	AWS libère en permanence chaque Hôtes dédiés défaillant et sur lequel il n'y a plus d'instance en cours d'exécution. L'ID de l'Hôte dédié ne peut plus être utilisé.

Libération d'Hôtes dédiés

Pour pouvoir libérer l'Hôte dédié, vous devez arrêter toutes les instances exécutées sur ce dernier. Ces instances peuvent être migrées vers un autre Hôtes dédiés de votre compte afin que vous puissiez continuer à les utiliser. Ces étapes ne concernent que les Hôtes dédiés à la demande.

Vous pouvez libérer un Hôte dédié à l'aide des méthodes suivantes.

Nouvelle console

Pour libérer un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sur la page Hôtes dédiés, sélectionnez le Hôte dédié à libérer.

4. Sélectionnez Actions, puis Libérer des hôtes.
5. Choisissez Libérer pour confirmer.

Ancienne console

Pour libérer un Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Hôtes dédiés dans le volet de navigation.
3. Sur la page Hôtes dédiés, sélectionnez le Hôte dédié à libérer.
4. Sélectionnez Actions, puis Release Hosts.
5. Choisissez Libérer pour confirmer.

AWS CLI

Pour libérer un Hôte dédié

Utilisez la commande release-hosts de l'AWS CLI.

```
aws ec2 release-hosts --host-ids h-012a3456b7890cdef
```

PowerShell

Pour libérer un Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell Remove-EC2hosts.

```
PS C:\> Remove-EC2Hosts -HostId h-012a3456b7890cdef
```

Une fois que vous avez libéré un Hôte dédié, vous ne pouvez plus réutiliser le même hôte ou ID d'hôte et la facturation à la demande pour cet hôte cesse. L'état de l'Hôte dédié devient `released` et vous ne pouvez plus lancer aucune instance sur cet hôte.

Note

Si vous avez récemment libéré des Hôtes dédiés, il peut s'écouler un peu de temps avant qu'ils cessent d'être comptabilisés dans le cadre de votre limite. Pendant ce temps, vous pouvez recevoir des erreurs `LimitExceeded` lorsque vous essayez d'allouer de nouveaux Hôtes dédiés. Dans ce cas, réessayez d'allouer ces nouveaux hôtes après quelques minutes.

Les instances qui ont été arrêtées peuvent toujours être utilisées et sont répertoriées à la page Instances. Elles conservent leur paramètre de location `host`.

Achat d'Réservations d'hôtes dédiés

Vous pouvez acheter des réservations en utilisant les méthodes suivantes :

Console

Pour acheter des réservations

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Hôtes dédiés, Réservations d'hôtes dédiés, Purchase Réservation d'hôtes dédiés (Acheter un hôte dédié).

3. Sur l'écran Purchase Réservation d'hôtes dédiés (Acheter un Réservation d'hôtes dédiés), vous pouvez rechercher les offres disponibles utilisant les paramètres par défaut ou spécifier des valeurs personnalisées pour les options suivantes :
   - Famille d'instances d'hôte — Les options répertoriées correspondent aux Hôtes dédiés de votre compte qui ne sont pas encore affectés à une réservation.
   - Zone de disponibilité — Zone de disponibilité des Hôtes dédiés de votre compte qui ne sont pas encore affectés à une réservation.
   - Option de paiement — Option de paiement de l'offre.
   - Durée — Durée de la réservation, qui peut être d'un ou trois ans.
4. Choisissez Trouver une offre et sélectionnez une offre correspondant à vos exigences.
5. Sélectionnez les Hôtes dédiés à associer à la réservation et choisissez Vérification.
6. Passez votre commande en revue et choisissez Order (Commander).

AWS CLI

Pour acheter des réservations

1. Utilisez la commande describe-host-reservation-offerings de l'AWS CLI pour répertorier les offres disponibles qui correspondent à vos besoins. L'exemple suivant répertorie les offres qui prennent en charge des instances dans la famille d'instances `m4` et ont une durée d'un an.

   Note

   La durée est indiquée en secondes. Une période d'un an comporte 31 536 000 secondes, tandis qu'une période de trois ans comporte 94 608 000 secondes.

```
aws ec2 describe-host-reservation-offerings --filter Name=instance-family,Values=m4
 --max-duration 31536000
```

   Les deux commandes renvoient une liste d'offres qui correspondent à vos critères de recherche. Notez l'`offeringId` de l'offre à acheter.

2. Utilisez la commande purchase-host-reservation de l'AWS CLI pour acheter l'offre et fournir la valeur de `offeringId` notée à l'étape précédente. L'exemple suivant achète la réservation spécifiée et l'associe à un Hôte dédié spécifique déjà alloué dans le compte AWS, puis applique une balise avec une clé `purpose` et une valeur `production`.

```
aws ec2 purchase-host-reservation --offering-id hro-03f707bf363b6b324 --
host-id-set h-013abcd2a00cbd123 --tag-specifications 'ResourceType=host-
reservation,Tags={Key=purpose,Value=production}'
```

PowerShell

Pour acheter des réservations

1. Utilisez la commande Outils AWS pour Windows PowerShell Get-EC2HostReservationOfferant pour répertorier les offres disponibles qui correspondent à vos besoins. Les exemples suivants répertorient les offres qui prennent en charge des instances dans la famille d'instances `m4` et ont une durée d'un an.

   Note

   La durée est indiquée en secondes. Une période d'un an comporte 31 536 000 secondes, tandis qu'une période de trois ans comporte 94 608 000 secondes.

```
PS C:\> $filter = @{Name="instance-family"; Value="m4"}
```

```
PS C:\> Get-EC2HostReservationOffering -filter $filter -MaxDuration 31536000
```

Les deux commandes renvoient une liste d'offres qui correspondent à vos critères de recherche. Notez l'`offeringId` de l'offre à acheter.

2. Utilisez la commande Outils AWS pour Windows PowerShell New-EC2HostReservation pour acheter l'offre et fournir la valeur de `offeringId` notée à l'étape précédente. Les exemples suivants permettent d'acheter la réservation spécifiée et l'associent à un Hôte dédié spécifique déjà alloué dans le compte AWS.

```
PS C:\> New-EC2HostReservation -OfferingId hro-03f707bf363b6b324 -
HostIdSet h-013abcd2a00cbd123
```

Affichage de réservations d'Hôte dédié.

Vous pouvez afficher des informations sur les Hôtes dédiés associés à votre réservation, en particulier :

- Durée de la réservation
- Options de paiement
- Dates de début et de fin

Vous pouvez consulter les détails de vos réservations d'Hôte dédié en utilisant les méthodes suivantes.

Console

Pour voir les détails d'une réservation d'Hôte dédié

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Hôtes dédiés dans le volet de navigation.
3. Sur la page Hôtes dédiés, choisissez Réservations d'Hôte dédié et sélectionnez la réservation dans la liste fournie.
4. Sélectionnez Détails pour en savoir plus sur la réservation.
5. Choisissez Hôtes pour en savoir plus sur les Hôtes dédiés auquel la réservation est associée.

AWS CLI

Pour voir les détails d'une réservation d'Hôte dédié

Utilisez la commande describe-host-reservations de l'AWS CLI.

```
aws ec2 describe-host-reservations
```

PowerShell

Pour voir les détails d'une réservation d'Hôte dédié

Utilisez la commande Outils AWS pour Windows PowerShell Get-EC2HostReservation.

```
PS C:\> Get-EC2HostReservation
```

Balisage d'Réservations d'hôtes dédiés

Vous pouvez allouer des balises personnalisées à vos Réservations d'hôtes dédiés pour classer celles-ci de différentes façons, par exemple, par objectif, par propriétaire ou par environnement. Cela vous aide à trouver rapidement un Réservation d'hôtes dédiés spécifique en fonction des balises personnalisées que vous lui avez attribuées.

Vous pouvez attribuer des balises à un Réservation d'hôtes dédiés uniquement à l'aide des outils de ligne de commande.

AWS CLI

Pour attribuer des balises à un Réservation d'hôtes dédiés

Utilisez la commande create-tags de l'AWS CLI.

```
aws ec2 create-tags --resources hr-1234563a4ffc669ae --tags Key=Owner,Value=TeamA
```

PowerShell

Pour attribuer des balises à un Réservation d'hôtes dédiés

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Tag.

La commande `New-EC2Tag` nécessite un paramètre `Tag`, qui spécifie la paire clé-valeur à utiliser pour la balise d'Réservation d'hôtes dédiés. Les commandes suivantes créent le paramètre `Tag` :

```
PS C:\> $tag = New-Object Amazon.EC2.Model.Tag
PS C:\> $tag.Key = "Owner"
PS C:\> $tag.Value = "TeamA"
```

```
PS C:\> New-EC2Tag -Resource hr-1234563a4ffc669ae -Tag $tag
```

Utilisation d'Hôtes dédiés partagés

Le partage d'Hôte dédié permet aux propriétaires d'un Hôte dédié de partager leurs Hôtes dédiés avec d'autres comptes AWS ou au sein d'une organisation AWS. Cela vous permet de créer et de gérer des Hôtes dédiés de manière centralisée, et de partager l'Hôte dédié entre plusieurs comptes AWS ou au sein de votre organisation AWS.

Dans ce modèle, le compte AWS auquel appartient l'Hôte dédié (le propriétaire) le partage avec d'autres comptes AWS (les consommateurs). Les consommateurs peuvent lancer des instances sur des Hôtes dédiés partagés avec eux comme ils le feraient sur des Hôtes dédiés qu'ils alloueraient dans leur propre compte. Le propriétaire est responsable de la gestion de l'Hôte dédié et des instances lancées sur celui-ci. Les propriétaires ne peuvent pas modifier les instances que les consommateurs lancent sur les Hôtes dédiés partagés. Les consommateurs sont responsables de la gestion des instances qu'ils lancent sur les Hôtes dédiés partagés avec eux. Les consommateurs ne peuvent ni afficher ni modifier les instances détenues par d'autres consommateurs ou par le propriétaire de l'Hôte dédié, et ils ne peuvent pas modifier les Hôtes dédiés qui sont partagés avec eux.

Un propriétaire d'Hôte dédié peut partager un Hôte dédié avec :

- Des comptes AWS spécifiques dans ou hors de son organisation AWS
- Une unité d'organisation dans son organisation AWS

- L'ensemble de son organisation AWS

Sommaire

- Conditions préalables au partage d'Hôtes dédiés (p. 488)
- Limites pour le partage des Hôte dédiés (p. 488)
- Services connexes (p. 488)
- Partage sur plusieurs zones de disponibilité (p. 488)
- Partage d'un Hôte dédié (p. 489)
- Annulation du partage d'un Hôte dédié partagé (p. 490)
- Identification d'un Hôte dédié partagé (p. 491)
- Affichage des instances en cours d'exécution sur un Hôte dédié partagé (p. 491)
- Autorisations relatives à un Hôte dédié partagé (p. 491)
- Facturation et mesures (p. 492)
- Limites Hôte dédié (p. 492)
- Récupération d'hôte et partage d'Hôte dédié (p. 492)

Conditions préalables au partage d'Hôtes dédiés

- Pour partager un Hôte dédié, vous devez le détenir dans votre compte AWS. Vous ne pouvez pas partager un Hôte dédié qui a été partagé avec vous.
- Pour partager un Hôte dédié avec votre organisation AWS ou avec une unité organisationnelle au sein de votre organisation AWS, vous devez activer le partage avec AWS Organizations. Pour plus d'informations, consultez la section relative à l'activation du partage avec des AWS Organizations dans le Guide de l'utilisateur AWS RAM.

Limites pour le partage des Hôte dédiés

Vous ne pouvez pas partager les Hôtes dédiés qui ont été alloués pour les types d'instance suivants : `u-6tb1.metal`, `u-9tb1.metal`, `u-12tb1.metal`, `u-18tb1.metal` et `u-24tb1.metal`.

Services connexes

AWS Resource Access Manager

Le partage d'un Hôte dédié s'intègre à AWS Resource Access Manager (AWS RAM). AWS RAM est un service qui vous permet de partager vos ressources AWS avec n'importe quel compte AWS ou via AWS Organizations. Avec AWS RAM, vous pouvez partager des ressources dont vous êtes propriétaire en créant un partage de ressources. Un partage de ressources spécifie les ressources à partager, ainsi que les consommateurs avec qui elles seront partagées. Les consommateurs peuvent être des comptes AWS individuels, des unités d'une organisation ou l'ensemble d'une organisation des AWS Organizations.

Pour plus d'informations sur AWS RAM, consultez le Guide de l'utilisateur AWS RAM.

Partage sur plusieurs zones de disponibilité

Pour garantir que les ressources sont réparties entre les zones de disponibilité d'une région, nous mappons indépendamment les zones de disponibilité aux noms de chaque compte. Cela peut entraîner des différences de nom de zone de disponibilité entre les comptes. Par exemple, la zone de disponibilité `us-east-1a` pour votre compte AWS peut avoir un emplacement autre que `us-east-1a` pour un autre compte AWS.

Pour identifier l'emplacement de vos Hôtes dédiés par rapport à vos comptes, vous devez utiliser l'ID de zone de disponibilité. L'ID de zone de disponibilité est un identifiant unique et cohérent pour une zone de disponibilité entre tous les comptes AWS. Par exemple, `use1-az1` est un ID de zone de disponibilité pour la région `us-east-1`, qui correspond au même emplacement dans chaque compte AWS.

Pour afficher les ID des zones de disponibilité dans votre compte

1. Ouvrez la console AWS RAM à l'adresse https://console.aws.amazon.com/ram.
2. Les ID de zone de disponibilité pour la région actuelle sont affichés dans le volet Your AZ ID (Votre ID de zone de disponibilité) dans la partie droite de l'écran.

Partage d'un Hôte dédié

Lorsqu'un propriétaire partage un Hôte dédié, il permet aux consommateurs de lancer des instances sur l'hôte. Les consommateurs peuvent lancer autant d'instances sur l'hôte partagé que sa capacité disponible le permet.

> **Important**
>
> Notez que vous êtes responsable de vous assurer que vous disposez des droits de licence appropriés pour partager les licences BYOL sur votre Hôtes dédiés.

Si vous partagez un Hôte dédié en ayant activé le placement automatique, gardez ce qui suit à l'esprit car cela pourrait conduire à une utilisation involontaire de l'Hôte dédié :

- Si les consommateurs lancent des instances avec location d'Hôte dédié et qu'ils n'ont pas de capacité sur un Hôte dédié qu'ils possèdent dans leur compte, l'instance est automatiquement lancée sur l'Hôte dédié partagé.

Pour partager un Hôte dédié, vous devez l'ajouter à un partage de ressources. Un partage de ressources est une ressource AWS RAM qui vous permet de partager vos ressources entre des comptes AWS. Un partage de ressources spécifie les ressources à partager, ainsi que les consommateurs avec qui elles seront partagées. Vous pouvez ajouter l'Hôte dédié à une ressource existante ou l'ajouter à un nouveau partage de ressources.

Si vous faites partie d'une organisation dans AWS Organizations et que le partage au sein de votre organisation est activé, l'accès à l'Hôte dédié partagé est automatiquement accordé aux consommateurs de votre organisation. Dans le cas contraire, les consommateurs reçoivent une invitation à rejoindre le partage de ressources et bénéficient d'un accès à l'Hôte dédié partagé après avoir accepté l'invitation.

> **Note**
>
> Après avoir partagé un Hôte dédié, les consommateurs peuvent y avoir accès en quelques minutes.

Vous pouvez partager un Hôte dédié que vous possédez à l'aide de l'une des méthodes suivantes.

Amazon EC2 console

Pour partager un Hôte dédié qui vous appartient à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Choisissez le Hôte dédié à partager, puis choisissez Actions, Partager l’hôte.
4. Sélectionnez le partage de ressources auquel vous souhaitez ajouter le Hôte dédié, puis choisissez Partager l’hôte.

Les consommateurs peuvent avoir accès à l'hôte partagé en quelques minutes.

AWS RAM console

Pour partager un Hôte dédié qui vous appartient à l'aide de la console AWS RAM

Consultez Création d'un partage de ressources dans le Guide de l'utilisateur AWS RAM.

AWS CLI

Pour partager un Hôte dédié qui vous appartient à l'aide d'AWS CLI

Utilisez la commande create-resource-share.

Annulation du partage d'un Hôte dédié partagé

Le propriétaire d'un Hôte dédié peut annuler le partage d'un Hôte dédié partagé à tout moment. Lorsque vous annulez le partage d'un Hôte dédié partagé, les règles suivantes s'appliquent :

- Les consommateurs avec qui l'Hôte dédié a été partagé ne peuvent plus lancer de nouvelles instances sur celui-ci.
- Les instances appartenant à des consommateurs qui s'exécutaient sur l'Hôte dédié au moment de l'annulation du partage continuent de s'exécuter, mais sont programmées pour être mises hors service. Les consommateurs reçoivent des notifications de mise hors service pour les instances, et disposent de deux semaines pour prendre les mesures nécessaires. Toutefois, si l'Hôte dédié est à nouveau partagé avec le consommateur au cours de la période de préavis de mise hors service, les mises hors service d'instance sont annulées.

Pour annuler le partage d'un Hôte dédié partagé qui vous appartient, vous devez le supprimer du partage de ressources. Pour ce faire, utilisez l'une des méthodes suivantes :

Amazon EC2 console

Pour annuler le partage d'un Hôte dédié partagé qui vous appartient à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Choisissez le Hôte dédié dont vous voulez annuler le partage et choisissez l'onglet Partage.
4. L'onglet Partage affiche la liste des partages de ressources auxquels le Hôte dédié a été ajouté. Sélectionnez le partage de ressources duquel vous souhaitez supprimer le Hôte dédié, puis choisissez Supprimer l'hôte du partage de ressources.

AWS RAM console

Pour annuler le partage d'un Hôte dédié partagé qui vous appartient à l'aide de la console AWS RAM

Consultez Mise à jour d'un partage de ressources dans le Guide de l'utilisateur AWS RAM.

Command line

Pour annuler le partage d'un Hôte dédié partagé qui vous appartient à l'aide d'AWS CLI

Utilisez la commande disassociate-resource-share.

Identification d'un Hôte dédié partagé

Les propriétaires et les consommateurs peuvent identifier les Hôtes dédiés partagés à l'aide de l'une des méthodes suivantes.

Amazon EC2 console

Pour identifier un Hôte dédié partagé à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés. L'écran affiche la liste des Hôtes dédiés qui vous appartiennent et des Hôtes dédiés qui sont partagés avec vous. La colonne Propriétaire affiche l'ID de compte AWS du propriétaire de l'Hôte dédié.

Command line

Pour identifier un Hôte dédié partagé à l'aide d'AWS CLI

Utilisez la commande describe-hosts. La commande renvoie les Hôtes dédiés qui vous appartiennent et les Hôtes dédiés qui sont partagés avec vous.

Affichage des instances en cours d'exécution sur un Hôte dédié partagé

Les propriétaires et les consommateurs peuvent afficher les instances s'exécutant sur un Hôte dédié partagé à tout moment à l'aide de l'une des méthodes suivantes.

Amazon EC2 console

Pour afficher les instances s'exécutant sur un Hôte dédié partagé à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sélectionnez l'Hôte dédié pour lequel afficher les instances et choisissez Instances. L'onglet répertorie les instances qui s'exécutent sur l'hôte. Les propriétaires voient toutes les instances s'exécutant sur l'hôte, y compris les instances lancées par les consommateurs. Les consommateurs ne voient que les instances en cours d'exécution qu'ils ont lancées sur l'hôte. La colonne Propriétaire affiche l'ID de compte AWS du compte qui a lancé l'instance.

Command line

Pour afficher les instances s'exécutant sur un Hôte dédié partagé à l'aide d'AWS CLI

Utilisez la commande describe-hosts. La commande renvoie les instances s'exécutant sur chaque Hôte dédié. Les propriétaires voient toutes les instances s'exécutant sur l'hôte. Les consommateurs ne voient que les instances en cours d'exécution qu'ils ont lancées sur les hôtes partagés. `InstanceOwnerId` affiche l'ID de compte AWS du propriétaire de l'instance.

Autorisations relatives à un Hôte dédié partagé

Autorisations accordées aux propriétaires

Les propriétaires sont responsables de la gestion de leurs Hôtes dédiés partagés et des instances qu'ils lancent sur eux. Les propriétaires peuvent afficher toutes les instances s'exécutant sur l'Hôte dédié

partagé, y compris celles lancées par les consommateurs. Toutefois, les propriétaires ne peuvent effectuer aucune action sur les instances en cours d'exécution lancées par les consommateurs.

Autorisations accordées aux consommateurs

Les consommateurs sont responsables de la gestion des instances qu'ils lancent sur un Hôte dédié partagé. Les consommateurs ne peuvent en aucun cas modifier l'Hôte dédié partagé. Ils ne peuvent pas non plus afficher ou modifier les instances qui ont été lancées par d'autres consommateurs ou par le propriétaire de l'Hôte dédié.

Facturation et mesures

Le partage d'Hôtes dédiés n'entraîne pas de frais supplémentaires.

Les propriétaires sont facturés pour les Hôtes dédiés qu'ils partagent. Les consommateurs ne sont pas facturés pour les instances qu'ils lancent sur des Hôtes dédiés partagés.

Les Réservations d'hôtes dédiés continuent à fournir des remises de facturation pour les Hôtes dédiés partagés. Seuls les propriétaires d'Hôte dédié peuvent acheter des Réservations d'hôtes dédiés pour les Hôtes dédiés partagés qu'ils possèdent.

Limites Hôte dédié

Les Hôtes dédiés partagés sont uniquement pris en compte dans les limites d'Hôtes dédiés du propriétaire. Les limites d'Hôtes dédiés du consommateur ne sont pas affectées par les Hôtes dédiés qui ont été partagés avec lui. De même, les instances que les consommateurs lancent sur les Hôtes dédiés partagés ne sont pas pris en compte dans leurs limites d'instance.

Récupération d'hôte et partage d'Hôte dédié

La récupération d'hôte permet de récupérer les instances lancées par le propriétaire d'un Hôte dédié et par les consommateurs avec qui ce dernier a été partagé. L'Hôte dédié de remplacement est alloué au compte du propriétaire. Il est ajouté aux mêmes partages de ressources que l'Hôte dédié d'origine, et il est partagé avec les mêmes consommateurs.

Pour plus d'informations, consultez Récupération de l'hôte (p. 492).

Récupération de l'hôte

La fonction de récupération de l'hôte redémarre automatiquement vos instances sur un nouvel hôte de remplacement si des incidents sont détectés sur votre Hôte dédié. La fonction de récupération de l'hôte permet de réduire les interventions manuelles et de diminuer la charge de travail opérationnelle en cas d'incident inattendu sur un Hôte dédié.

En outre, l'intégration standard d'AWS License Manager automatise le suivi et la gestion de vos licences en cas de récupération de l'hôte.

> Note
>
> L'intégration d'AWS License Manager est uniquement prise en charge dans les régions dans lesquelles AWS License Manager est disponible.

Sommaire

- Notions de base de la récupération de l'hôte (p. 493)
- Types d'instance pris en charge (p. 494)
- Configuration de la récupération de l'hôte (p. 494)

- États de la récupération de l'hôte (p. 495)
- Récupération manuelle des instances non prises en charge (p. 496)
- Services connexes (p. 496)
- Tarification (p. 496)

Notions de base de la récupération de l'hôte

La fonction de récupération de l'hôte fait intervenir des vérifications de l'état au niveau de l'hôte pour évaluer la disponibilité de l'hôte dédié et détecter les pannes système sous-jacentes. Voici quelques exemples de problèmes pouvant entraîner l'échec des vérifications de l'état au niveau de l'hôte :

- Perte de connectivité réseau
- Perte d'alimentation système
- Problèmes logiciels ou matériels sur l'hôte physique

Lorsqu'une panne système est détectée sur votre Hôte dédié, la fonction de récupération de l'hôte est lancée et Amazon EC2 alloue automatiquement un Hôte dédié de remplacement. L'Hôte dédié de remplacement reçoit un nouvel ID d'hôte, mais conserve les mêmes attributs que l'Hôte dédié d'origine, en particulier :

- Zone de disponibilité
- Type d'instance
- Balises
- Paramètres de placement automatique

Une fois l'Hôte dédié de remplacement alloué, les instances sont récupérées sur l'Hôte dédié de remplacement. Les instances récupérées conservent les mêmes attributs que les instances d'origine, en particulier :

- ID d'instance
- Adresses IP privées
- Adresses IP Elastic
- Pièces jointes de volume EBS
- Toutes les métadonnées d'instance

Si des instances ont des relations d'affinité avec l'Hôte dédié déficient, les instances récupérées établissent une relation d'affinité avec l'Hôte dédié de remplacement.

Une fois que toutes les instances ont été récupérées sur l'Hôte dédié de remplacement, l'Hôte dédié déficient est libéré et l'Hôte dédié de remplacement devient disponible.

Lors du lancement du processus de récupération de l'hôte, le propriétaire du compte AWS est averti par e-mail et par un événement AWS Personal Health Dashboard. Une seconde notification est envoyée une fois la récupération de l'hôte réalisée avec succès.

Les instances arrêtées ne sont pas récupérées sur l'Hôte dédié de remplacement. Si vous tentez de démarrer une instance arrêtée qui cible l'Hôte dédié déficient, son démarrage échoue. Nous vous recommandons de modifier l'instance arrêtée afin qu'elle cible un autre Hôte dédié ou de la lancer sur tout Hôte dédié disponible ayant des caractéristiques de configuration et de remplacement automatique correspondantes.

Les instances avec stockage d'instance ne sont pas récupérées sur l'Hôte dédié de remplacement. Afin de remédier à ce problème, l'Hôte dédié déficient est marqué comme devant être mis hors service et vous recevez une notification de mise hors service une fois la récupération de l'hôte terminée. Suivez les étapes correctives décrites dans la notification de mise hors service dans le temps imparti pour récupérer manuellement les instances restantes sur l'Hôte dédié déficient.

Si vous utilisez AWS License Manager pour effectuer le suivi de vos licences, AWS License Manager alloue de nouvelles licences pour l'Hôte dédié de remplacement en fonction des limites de configuration de licence. Si la configuration de licence définit des limites strictes qui seront outrepassées à la suite de la récupération de l'hôte, le processus de récupération n'est pas autorisé et vous êtes averti de l'échec de la récupération de l'hôte via une notification Amazon SNS. Si la configuration de licence définit des limites flexibles qui seront outrepassées à la suite de la récupération de l'hôte, le processus de récupération est autorisé et vous êtes averti du dépassement de la limite via une notification Amazon SNS. Pour plus d'informations, consultez Utilisation de configurations de licence dans le Guide de l'utilisateur AWS License Manager.

Types d'instance pris en charge

La récupération de l'hôte est prise en charge pour les familles d'instances suivantes : A1, C3, C4, C5, C5n, C6g, Inf1, M3, M4, M5, M5n, M6g, P3, R3, R4, R5, R5n, R6g, X1, X1e, u-6tb1, u-9tb1, u-12tb1, u-18tb1 et u-24tb1.

Pour récupérer des instances qui ne sont pas prises en charge, consultez Récupération manuelle des instances non prises en charge (p. 496).

Configuration de la récupération de l'hôte

Vous pouvez configurer la récupération de l'hôte au moment de l'allocation de l'Hôte dédié ou après l'allocation à l'aide de la console Amazon EC2 ou de l'AWS Command Line Interface (CLI).

Sommaire

- Activation de la récupération de l'hôte (p. 494)
- Désactivation de la récupération de l'hôte (p. 495)
- Affichage de la configuration de récupération de l'hôte (p. 495)

Activation de la récupération de l'hôte

Vous pouvez activer la récupération de l'hôte au moment de l'allocation de l'Hôte dédié ou après l'allocation.

Pour de plus amples informations sur l'activation de la récupération de l'hôte au moment de l'allocation de l'Hôte dédié, veuillez consulter Allocation d'Hôtes dédiés (p. 470).

Pour activer la récupération de l'hôte après l'allocation à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sélectionnez l'Hôte dédié pour lequel vous souhaitez activer la fonction de récupération de l'hôte, puis choisissez Actions, Modify Host Recovery (Modifier la récupération de l'hôte).
4. Pour Host recovery (Récupération de l'hôte), choisissez Enable (Activer), puis Save (Enregistrer).

Pour activer la récupération de l'hôte après l'allocation à l'aide de l'AWS CLI

Utilisez la commande modify-hosts et spécifiez le paramètre `host-recovery`.

```
$ aws ec2 modify-hosts --host-recovery on --host-ids h-012a3456b7890cdef
```

Désactivation de la récupération de l'hôte

Vous pouvez désactiver la récupération de l'hôte à tout moment après l'allocation de l'Hôte dédié.

Pour désactiver la récupération de l'hôte après l'allocation à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sélectionnez l'Hôte dédié pour lequel vous souhaitez désactiver la fonction de récupération de l'hôte, puis choisissez Actions, Modify Host Recovery (Modifier la récupération de l'hôte).
4. Pour Host recovery (Récupération de l'hôte), choisissez Disable (Désactiver), puis Save (Enregistrer).

Pour désactiver la récupération de l'hôte après l'allocation à l'aide de l'AWS CLI

Utilisez la commande modify-hosts et spécifiez le paramètre `host-recovery`.

```
$ aws ec2 modify-hosts --host-recovery off --host-ids h-012a3456b7890cdef
```

Affichage de la configuration de récupération de l'hôte

Vous pouvez afficher la configuration de récupération de l'hôte d'un Hôte dédié à tout moment.

Pour afficher la configuration de récupération de l'hôte d'un Hôte dédié à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Hôtes dédiés.
3. Sélectionnez l'Hôte dédié, puis, dans l'onglet Description, passez en revue le champ Host Recovery (Récupération de l'hôte).

Pour afficher la configuration de récupération de l'hôte d'un Hôte dédié à l'aide de l'AWS CLI

Utilisez la commande describe-hosts.

```
$ aws ec2 describe-hosts --host-ids h-012a3456b7890cdef
```

L'élément de réponse `HostRecovery` indique si la récupération de l'hôte est activée ou désactivée.

États de la récupération de l'hôte

Lorsqu'une déficience d'Hôte dédié est détectée, l'Hôte dédié déficient passe à l'état `under-assessment` et toutes les instances passent à l'état `impaired`. Vous ne pouvez pas lancer des instances sur l'Hôte dédié déficient tant qu'il est à l'état `under-assessment`.

Une fois l'Hôte dédié de remplacement alloué, il passe à l'état `pending`. Il reste dans cet état jusqu'à ce que le processus de récupération de l'hôte soit terminé. Vous ne pouvez pas lancer des instances sur l' Hôte dédié de remplacement tant qu'il est à l'état `pending`. Les instances récupérées situées sur l'Hôte dédié de remplacement restent à l'état `impaired` durant le processus de récupération.

Une fois la récupération de l'hôte terminée, l'Hôte dédié de remplacement passe à l'état `available` et les instances récupérées repassent à l'état `running`. Vous pouvez lancer des instances sur l'Hôte dédié

de remplacement une fois qu'il est à l'état `available`. L'Hôte dédié déficient d'origine est libéré de façon permanente et il passe à l'état `released-permanent-failure`.

Si l'Hôte dédié déficient possède des instances qui ne prennent pas en charge la récupération de l'hôte, telles que les instances comportant des volumes basés sur le stockage d'instance, l'Hôte dédié n'est pas libéré. Il est marqué comme devant être mis hors service et passe à l'état `permanent-failure`.

Récupération manuelle des instances non prises en charge

La fonction de récupération de l'hôte ne prend pas en charge la récupération des instances qui utilisent des volumes de stockage d'instance. Suivez les instructions ci-après pour récupérer manuellement les instances qui n'ont pas pu être récupérées automatiquement.

> Warning
>
> Les données stockées sur des volumes de stockage d'instance sont perdues lorsqu'une instance est arrêtée, mise en veille prolongée ou résiliée. Ceci inclut les volumes de stockage d'instance attachés à une instance ayant un volume EBS comme périphérique racine. Pour protéger les données provenant des volumes de stockage d'instance, sauvegardez-les sur un stockage permanent avant l'arrêt ou la résiliation de l'instance.

Récupération manuelle des instances basées sur des volumes EBS

Pour les instances basées sur des volumes EBS qui n'ont pas pu être récupérées automatiquement, nous vous recommandons de les arrêter puis de les redémarrer manuellement afin de les récupérer sur un nouvel Hôte dédié. Pour de plus amples informations sur l'arrêt de votre instance, ainsi que sur les changements apportés à la configuration de votre instance lorsque celle-ci est arrêtée, veuillez consulter Arrêt et démarrage de votre instance (p. 625).

Récupération manuelle des instances basées sur le stockage d'instance

Pour les instances basées sur le stockage d'instance qui n'ont pas pu être récupérées automatiquement, nous vous recommandons de procéder comme suit :

1. Lancez une instance de remplacement sur un nouvel Hôte dédié à partir de votre AMI la plus récente.
2. Migrez toutes les données nécessaires vers l'instance de remplacement.
3. Résiliez l'instance d'origine sur l'Hôte dédié déficient.

Services connexes

L'Hôte dédié intègre les services AWS suivants :

- AWS License Manager — Assure le suivi des licences sur vos Hôtes dédiés Amazon EC2 (pris en charge uniquement dans les régions dans lesquelles AWS License Manager est disponible). Pour de plus amples informations, veuillez consulter le Guide de l'utilisateur AWS License Manager.

Tarification

Il n'y a pas de facturation supplémentaire pour l'utilisation de la fonction de récupération de l'hôte, mais les frais habituellement appliqués pour l'Hôte dédié vous seront facturés. Pour de plus amples informations, veuillez consulter Tarification des hôtes dédiés Amazon EC2.

Dès que la fonction de récupération de l'hôte est lancée, vous n'êtes plus facturé pour l'Hôte dédié déficient. La facturation relative à l'hôte dédié de remplacement commence uniquement une fois qu'il est passé à l'état `available`.

Si l'Hôte dédié déficient était facturé au tarif à la demande, l'Hôte dédié de remplacement est également facturé au tarif à la demande. Si l'Hôte dédié déficient possédait une Réservation d'hôtes dédiés, elle est transférée à l'Hôte dédié de remplacement.

Suivi des changements de configuration

Vous pouvez utiliser AWS Config pour enregistrer les changements de configuration des Hôtes dédiés et des instances qui sont lancées, arrêtées ou résiliées sur ces hôtes. Vous pouvez utiliser les informations capturées par AWS Config comme source de données pour les rapports d'utilisation des licences.

AWS Config enregistre individuellement les informations de configuration des Hôtes dédiés et des instances, et associe ces informations via des relations. Il y a trois conditions pour la création de rapports :

- Statut de l'enregistrement AWS Config : lorsque ce paramètre est défini sur Activé, AWS Config enregistre un ou plusieurs types de ressource AWS, parmi lesquels peuvent figurer des Hôtes dédiés et des Instances dédiées. Pour capturer les informations requises pour les rapports d'utilisation des licences, vérifiez que les hôtes et les instances sont enregistrés avec les champs suivants.
- Statut de l'enregistrement de l'hôte — Lorsque ce paramètre a la valeur Activé, les informations de configuration concernant les Hôtes dédiés sont enregistrées.
- Statut de l'enregistrement de l'instance : lorsque ce paramètre est défini sur Activé, les informations de configuration concernant les Instances dédiées sont enregistrées.

Si l'une de ces trois conditions est désactivée, l'icône du bouton Edit Config Recording est rouge. Afin de tirer pleinement profit de cet outil, assurez-vous que les trois méthodes d'enregistrement soient activées. Lorsqu'elles sont toutes les trois activées, l'icône est verte. Pour modifier les paramètres, choisissez Edit Config Recording. Vous êtes alors dirigé vers la page Set up AWS Config de la console AWS Config, où vous pouvez configurer AWS Config et commencer l'enregistrement pour vos hôtes, instances et autres types de ressources pris en charge. Pour plus d'informations, consultez Configuration de AWS Config à l'aide de la console dans le Manuel du développeur AWS Config.

> **Note**
>
> AWS Config enregistre vos ressources après les avoir trouvées, ce qui peut prendre quelques minutes.

Une fois que AWS Config a commencé à enregistrer les changements de configuration apportés à vos hôtes et instances, vous pouvez obtenir l'historique de configuration de n'importe quel hôte que vous avez alloué ou libéré, et de n'importe quelle instance que vous avez lancée, arrêtée ou terminée. Par exemple, à tout moment dans l'historique de configuration d'un Hôte dédié, vous pouvez rechercher combien d'instances sont lancées sur cet hôte, ainsi que le nombre de sockets et de cœurs sur l'hôte. Pour n'importe laquelle de ces instances, vous pouvez également rechercher l'ID de son Amazon Machine Image (AMI). Vous pouvez utiliser ces informations pour les rapports de licences portant sur vos propres licences logicielles liées au serveur par socket ou par cœur.

Vous pouvez accéder aux historiques de configuration de l'une des façons suivantes :

- En utilisant la console AWS Config. Pour chaque ressource enregistrée, vous pouvez visualiser une page chronologique fournissant une historique des détails de configuration. Pour visualiser cette page, choisissez l'icône grise dans la colonne Chronologie de configuration de la page Hôtes dédiés. Pour plus d'informations, consultez Affichage des détails de configuration dans la console AWS Config dans le Manuel du développeur AWS Config.
- En exécutant des commandes d'AWS CLI. Tout d'abord, vous pouvez utiliser la commande list-discovered-resources pour obtenir une liste des hôtes et des instances. Vous pouvez ensuite utiliser la commande get-resource-config-history pour obtenir les détails de configuration d'un hôte ou d'une instance pour un intervalle de temps donné. Pour plus d'informations, veuillez consultez Afficher les détails de configuration à l'aide de la CLI dans le AWS Config Manuel du développeur.

- En utilisant l'API AWS Config dans vos applications. Tout d'abord, vous pouvez utiliser l'action ListDiscoveredResources pour obtenir une liste des hôtes et des instances. Vous pouvez ensuite utiliser l'action GetResourceConfigHistory pour obtenir les détails de configuration d'un hôte ou d'une instance pour un intervalle de temps donné.

Par exemple, pour obtenir la liste de la totalité de vos Hôtes dédiés à partir d'AWS Config, exécutez une commande d'interface de ligne de commande semblable à la suivante.

```
aws configservice list-discovered-resources --resource-type AWS::EC2::Host
```

Pour obtenir l'historique de configuration d'un Hôte dédié à partir d'AWS Config, exécutez une commande d'interface de ligne de commande semblable à la suivante.

```
aws configservice get-resource-config-history --resource-type AWS::EC2::Instance --
resource-id i-1234567890abcdef0
```

Pour gérer les paramètres AWS Config à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sur la page Hôtes dédiés, sélectionnez Modifier l'enregistrement de la configuration.
3. Dans la console AWS Config, suivez les étapes mentionnées pour activer l'enregistrement. Pour plus d'informations, veuillez consultez Configuration de AWS Config à l'aide de la console.

Pour plus d'informations, consultez Affichage des détails de configuration dans la console AWS Config.

Pour activer AWS Config à l'aide de la ligne de commande ou de l'API

- AWS CLI : Affichage des détails de configuration (AWS CLI) dans le Guide du développeur AWS Config.
- API Amazon EC2 : GetResourceConfigHistory.

Instances dédiées

Les Instances dédiées sont des instances Amazon EC2 qui s'exécutent dans un cloud privé virtuel (VPC) sur du matériel dédié à un seul client. Les Instances dédiées qui appartiennent à différents comptes AWS sont physiquement isolées au niveau matériel, même si ces comptes sont liés à un compte payeur unique. Cependant, les Instances dédiées peuvent partager du matériel avec d'autres instances du même compte AWS qui ne sont pas des Instances dédiées.

> Note
>
> Un Hôte dédié est également un serveur physique qui vous est dédié. Un Hôte dédié vous assure la visibilité et le contrôle sur la façon dont les instances sont placées sur le serveur. Pour plus d'informations, consultez Hôtes dédiés (p. 465).

Bases d'une Instance dédiée

Chaque instance que vous lancez dans un VPC a un attribut de location. Cet attribut a les valeurs suivantes.

Valeur de location	Description
`default`	Votre instance s'exécute sur un matériel partagé.

Valeur de location	Description
`dedicated`	Votre instance s'exécute sur un matériel à client unique.
`host`	Votre instance s'exécute sur un Hôte dédié, qui est un serveur isolé avec des configurations que vous pouvez contrôler.

Après le lancement d'une instance, certaines restrictions s'appliquent à la modification de sa location.

- Vous ne pouvez pas modifier la location d'une instance en remplaçant `default` par `dedicated` ou par `host` après l'avoir lancée.
- Vous ne pouvez pas modifier la location d'une instance en remplaçant `dedicated` ou `host` par `default` après l'avoir lancée.

Vous pouvez modifier la location d'une instance en remplaçant `dedicated` par `host` ou `host` par `dedicated` après l'avoir lancée. Pour plus d'informations, consultez Modification de la location d'une instance (p. 503).

Chaque VPC a un attribut de location d'instance qui lui est associé. Cet attribut a les valeurs suivantes.

Valeur de location	Description
`default`	Une instance lancée dans le VPC s'exécute sur un matériel partagé par défaut, sauf si vous spécifiez explicitement une autre location lors du lancement de l'instance.
`dedicated`	Une instance lancée dans le VPC est une Instance dédiée par défaut, sauf si vous spécifiez explicitement une location `host` lors de son lancement. Vous ne pouvez pas spécifier de location `default` lors du lancement de l'instance.

Vous pouvez modifier la location d'instance d'un VPC en remplaçant `dedicated` par `default` après l'avoir créé. Vous ne pouvez pas modifier la location d'instance d'un VPC en remplaçant `default` par `dedicated` après l'avoir créé.

Pour créer des Instances dédiées, vous pouvez effectuer les opérations suivantes :

- Créez le VPC avec la location d'instance `dedicated` (toutes les instances lancées dans ce VPC sont des Instances dédiées).
- Créez un VPC avec la location de l'instance définie sur `default` et spécifiez une location `dedicated` pour toutes les instances lors de leur lancement.

Restrictions des Instances dédiées

Certains services AWS ou leurs fonctions ne fonctionnent pas avec un VPC dont la location d'instance est définie comme `dedicated`. Vérifiez la documentation du service pour confirmer s'il existe des restrictions.

Certains types d'instance ne peuvent pas être lancés dans un VPC dont la location d'instance est définie comme `dedicated`. Pour plus d'informations sur les types d'instance pris en charge, consultez Amazon EC2 Instances dédiées.

Amazon EBS avec Instances dédiées

Quand vous lancez une Instance dédiée sauvegardée sur Amazon EBS, le volume EBS ne s'exécute pas sur un matériel dédié à client unique.

Instances réservées avec location dédiée

Pour garantir qu'une capacité suffisante est disponible pour le lancement d'Instances dédiées, vous pouvez acheter des Instances réservées dédiées. Pour plus d'informations, consultez Instances réservées (p. 329).

Quand vous achetez une Instance réservée dédiée, vous achetez la capacité nécessaire pour lancer une Instance dédiée dans un VPC à un coût d'utilisation grandement réduit ; la rupture de prix sur le coût d'utilisation s'applique uniquement si vous lancez une instance avec une location dédiée. Lorsque vous achetez une Instance réservée avec une location par défaut, celle-ci s'applique uniquement à une instance en cours d'exécution dotée d'un location `default`. Elle ne s'appliquerait pas à une instance en cours d'exécution dotée d'une location `dedicated`.

Vous ne pouvez pas utiliser le processus de modification pour modifier la location d'une Instance réservée après l'avoir achetée. Par contre, vous pouvez échanger une Instance réservée convertible contre une nouvelle Instance réservée convertible avec une autre location.

Scalabilité automatique des Instances dédiées

Vous pouvez utiliser Amazon EC2 Auto Scaling pour lancer des Instances dédiées. Pour plus d'informations, consultez Lancement d'instances Auto Scaling dans un VPC dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Récupération automatique des Instances dédiées

Vous pouvez configurer la récupération automatique de Instances dédiées si les instances sont dégradées suite à une défaillance du matériel sous-jacent ou à un problème nécessitant une intervention d'AWS pour sa résolution. Pour plus d'informations, consultez Récupération de votre instance (p. 652).

Instances Spot dédiées

Vous pouvez exécuter une instance Spot dédiée en spécifiant une location `dedicated` lorsque vous créez une demande d'instance Spot. Pour plus d'informations, consultez Définition d'une location pour vos Instances Spot (p. 393).

Tarification des Instances dédiées

La tarification des Instances dédiées est différente de celle des instances à la demande. Pour plus d'informations, consultez la page produit des Instances dédiées Amazon EC2.

Instances à capacité extensible avec Instances dédiées

Vous pouvez tirer parti des avantages liés à une exécution sur du matériel à location dédiée avec the section called “Instances à capacité extensible” (p. 233). Par défaut, les instances dédiées T3 sont lancées en mode illimité. Leur niveau de performances d'UC de base peut être étendu à un niveau supérieur lorsque la charge de travail l'exige. Les performances de base T3 et la possibilité d'émettre en rafale sont régies par les crédits UC. Compte tenu de la nature extensible des types d'instance T3, pour des performances optimales, nous vous recommandons de surveiller la façon dont vos instances T3 utilisent les ressources d'UC du matériel dédié. Les instances dédiées T3 s'adressent à des clients dont les charges de travail variées présentent un comportement d'UC aléatoire, mais dont le niveau d'utilisation d'UC est de préférence moyen ou inférieur aux niveaux d'utilisation de base. Pour de plus amples informations, veuillez consulter the section called “Crédits UC et utilisation de référence” (p. 234).

Amazon EC2 dispose de systèmes qui permettent d'identifier et de corriger les fluctuations au niveau des performances. Cependant, il est toujours possible d'observer des fluctuations à court terme si vous lancez plusieurs instances dédiées T3 dont les modèles d'utilisation d'UC sont corrélés. Pour les charges de travail plus exigeantes ou corrélées, nous recommandons d'utiliser des instances dédiées M5 ou M5 plutôt que des instances dédiées T3.

Utilisation des Instances dédiées

Vous pouvez créer un VPC avec une location d'instance `dedicated` afin de veiller à ce que toutes les instances lancées dans le VPC soient des Instances dédiées. Vous pouvez aussi spécifier la location de l'instance lors du lancement.

Rubriques

- Création d'un VPC avec une location d'instance dédiée (p. 501)
- Lancement d'Instances dédiées dans un VPC (p. 501)
- Affichage des informations de location (p. 502)
- Modification de la location d'une instance (p. 503)
- Modification de la location d'un VPC (p. 503)

Création d'un VPC avec une location d'instance dédiée

Quand vous créez un VPC, vous avez la possibilité de spécifier sa location d'instance. Si vous utilisez la console Amazon VPC, vous pouvez créer un VPC à l'aide de l'assistant VPC ou de la page Vos VPC.

Créer un VPC avec une location d'instance dédiée (assistant VPC)

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le tableau de bord, choisissez Démarrer l'assistant VPC.
3. Sélectionnez une configuration de VPC, puis choisissez Select.
4. Sur la page suivante de l'assistant, choisissez Dedicated dans la liste Hardware tenancy.
5. Sélectionnez Create VPC (Créer un VPC).

Créer un VPC avec une location d'instance dédiée (créer une boîte de dialogue VPC)

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, choisissez Your VPCs, puis Create VPC.
3. Pour Location, choisissez Dédiée. Spécifiez le bloc d'adresse CIDR, puis choisissez Oui, créer.

Pour définir l'option de location lorsque vous créez un VPC à l'aide de la ligne de commande

- create-vpc (AWS CLI)
- New-EC2Vpc (Outils AWS pour Windows PowerShell)

Si vous lancez une instance dans un VPC qui a une location d'instance `dedicated`, votre instance est automatiquement une Instance dédiée, quelle que soit la location de l'instance.

Lancement d'Instances dédiées dans un VPC

Vous pouvez lancer une Instance dédiée à l'aide de l'assistant de lancement d'instance Amazon EC2.

Pour lancer une Instance dédiée dans un VPC de location par défaut à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Sur la page Choose an Amazon Machine Image (AMI), sélectionnez une AMI puis choisissez Select.
4. Sur la page Choisir un type d'instance, sélectionnez le type d'instance, puis choisissez Suivant : Configurer les détails de l'instance.

Note

Veillez à choisir un type d'instance pris en charge en tant qu'Instance dédiée. Pour plus d'informations, consultez Instances dédiées Amazon EC2.

5. Sur la page Configure Instance Details, sélectionnez un VPC et un sous-réseau. Choisissez Dedicated - Run a dedicated instance dans la liste Tenancy, puis Next: Add Storage.
6. Continuez comme indiqué par l'assistant. Lorsque vous avez terminé de vérifier vos options sur la page Review Instance Launch, choisissez Launch pour choisir une paire de clés et lancer l'Instance dédiée.

Pour plus d'informations sur le lancement d'une instance avec une location `host`, consultez Lancement d'instances sur un Hôte dédié (p. 472).

Pour définir l'option de location pour une instance lors du lancement à l'aide de la ligne de commande

- run-instances (AWS CLI)
- New-EC2Instance

Affichage des informations de location

Pour afficher les informations de location pour votre VPC à l'aide de la console

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, sélectionnez Vos VPC.
3. Vérifiez la location d'instance de votre VPC dans la colonne Tenancy.
4. Si la colonne Tenancy n'est pas affichée, choisissez Edit Table Columns (icône en forme d'engrenage), Tenancy dans la boîte de dialogue Show/Hide Columns, puis Close.

Pour afficher les informations de location pour votre instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Vérifiez la location de votre instance dans la colonne Tenancy.
4. Si la colonne Tenancy n'est pas affichée, effectuez l'une des actions suivantes :
   - Choisissez Afficher / Masquer les colonnes (icône en forme d'engrenage), Location dans la boîte de dialogue Afficher / Masquer les colonnes, puis Fermer.
   - Sélectionnez l'instance. L'onglet Description dans le volet des détails affiche des informations sur l'instance, dont sa location.

Pour décrire la location de votre VPC à l'aide de la ligne de commande

- describe-vpcs (AWS CLI)
- Get-EC2Vpc (Outils AWS pour Windows PowerShell)

Pour décrire la location de votre instance à l'aide de la ligne de commande

- describe-instances (AWS CLI)
- Get-EC2Instance

Pour décrire la valeur de location d'une Instance réservée à l'aide de la ligne de commande

- describe-reserved-instances (AWS CLI)
- Get-EC2ReservedInstance (Outils AWS pour Windows PowerShell)

Pour décrire la valeur de location d'une offre d'Instance réservée à l'aide de la ligne de commande

- describe-reserved-instances-offerings (AWS CLI)
- Get-EC2ReservedInstancesOffering (Outils AWS pour Windows PowerShell)

Modification de la location d'une instance

Selon votre type d'instance et votre plateforme, vous pouvez remplacer la location d'une Instance dédiée arrêtée par `host` après l'avoir lancée. Au prochain démarrage de l'instance, elle démarre sur un Hôte dédié alloué à votre compte. Pour plus d'informations sur l'allocation et l'utilisation d'Hôtes dédiés, et les types d'instance qui peuvent être utilisés avec des Hôtes dédiés, consultez Utilisation d'Hôtes dédiés (p. 469). De même, vous pouvez remplacer la location d'une instance d'Hôte dédié arrêtée par `dedicated` après l'avoir lancée. La prochaine fois que l'instance démarre, elle démarre sur un matériel à client unique que nous contrôlons.

Pour modifier la location d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez votre instance.
3. Choisissez État de l'instance, Arrêter l'instance.
4. Choisissez Actions, Paramètres de l'instance, puis Modifier le placement d'instance.
5. Dans la liste Location, choisissez d'exécuter votre instance sur un matériel dédié ou sur un Hôte dédié. Choisissez Enregistrer.

Pour modifier la valeur de location d'une instance à l'aide de la ligne de commande

- modify-instance-placement (AWS CLI)
- Edit-EC2InstancePlacement (Outils AWS pour Windows PowerShell)

Modification de la location d'un VPC

Vous pouvez modifier l'attribut de location d'instance d'un VPC en remplaçant `dedicated` par `default`. La modification de la location d'instance du VPC n'affecte pas la location des instances figurant dans le VPC. La prochaine fois que vous lancerez une instance dans ce VPC, elle aura une location `default`, à moins que vous ne spécifiiez une autre option lors du lancement.

Vous ne pouvez pas modifier l'attribut de location d'un VPC en remplaçant `default` par `dedicated` après l'avoir créé.

Vous pouvez modifier l'attribut de location d'instance d'un VPC à l'aide de l'AWS CLI, d'un kit SDK AWS ou de l'API Amazon EC2 uniquement.

Pour modifier l'attribut de location d'instance d'un VPC à l'aide de l'AWS CLI

- Utilisez la commande modify-vpc-tenancy pour spécifier l'ID du VPC et la valeur de location d'instance. La seule valeur prise en charge est `default`.

```
aws ec2 modify-vpc-tenancy --vpc-id vpc-1a2b3c4d --instance-tenancy default
```

Réservations de capacité à la demande

Réservations de capacité à la demande vous permet de réserver de la capacité pour vos instances Amazon EC2 dans une zone de disponibilité spécifique pour la durée de votre choix. Cela vous permet de créer et de gérer les Réservations de capacité indépendamment des remises de facturation offertes par les Savings Plans ou les Instances réservées régionales. En créant Réservations de capacité, vous veillez à toujours avoir accès à la capacité EC2 lorsque vous en avez besoin, pendant aussi longtemps que nécessaire. Vous pouvez créer Réservations de capacité à tout moment, sans avoir à vous engager pour une durée de 1 à 3 ans, et la capacité est disponible immédiatement. Lorsque vous n'en avez plus besoin, annulez le Réservation de capacité pour ne plus encourir de frais.

Lorsque vous créez un Réservation de capacité, vous spécifiez :

- Zone de disponibilité dans laquelle la capacité est réservée
- Nombre d'instances pour lesquelles vous souhaitez réserver la capacité
- Attributs d'instances, dont le type d'instance, la location et la plateforme/le SO

Réservations de capacité peut uniquement être utilisé par les instances correspondant aux attributs. Par défaut, elles sont automatiquement utilisées par les instances en cours d'exécution dont les attributs correspondent. Si vous n'avez aucune instance en cours d'exécution dont les attributs correspondent à ceux de la Réservation de capacité, celle-ci reste inutilisée jusqu'à ce que vous lanciez une instance dont les attributs correspondent.

De plus, vous pouvez utiliser les Savings Plans et les Instances réservées régionales avec vos Réservations de capacité pour obtenir des remises de facturation. AWS applique automatiquement votre remise lorsque les attributs d'un Réservation de capacité correspondent à ceux d'un Savings Plan ou d'une Instance réservée régionale. Pour de plus amples informations, veuillez consulter Remises de facturation (p. 507).

Sommaire

- Différences entre les Réservations de capacité, les Instances réservées et les Savings Plans (p. 504)
- Plateformes prises en charge (p. 505)
- Limites de Réservation de capacité (p. 506)
- Limites et restrictions d'une Réservation de capacité (p. 506)
- Tarification et facturation d'une Réservation de capacité (p. 506)
- Utilisation de Réservations de capacité (p. 507)
- Réservations de capacité dans Zones locales (p. 517)
- Utilisation de Réservations de capacité partagées (p. 518)
- Métriques CloudWatch pour Réservations de capacité à la demande (p. 522)

Différences entre les Réservations de capacité, les Instances réservées et les Savings Plans

Le tableau suivant met en évidence les principales différences entre les Réservations de capacité, les Instances réservées et les Savings Plans :

	Réservations de capacité	Instances réservées zonales	Instances réservées régionales	Savings Plans
Durée	Aucun engagement requis. Peuvent être	Exigent un engagement d'un an ou de trois ans.		

	Réservations de capacité	Instances réservées zonales	Instances réservées régionales	Savings Plans
	créées et annulées selon les besoins.			
Avantage de capacité	Capacité réservée dans une zone de disponibilité spécifique.		Pas de réservation de capacité dans une zone de disponibilité.	
Remise de facturation	Pas de remise de facturation. Les instances lancées dans une Réservation de capacité sont facturées selon leurs tarifs à la demande standard. Toutefois, vous pouvez utiliser les Savings Plans ou les Instances réservées régionales avec les Réservations de capacité pour obtenir une remise de facturation. Les Instances réservées par zone ne s'appliquent pas à Réservations de capacité.	Permettent d'obtenir des remises de facturation		
Limites d'instance	Restreintes à vos limites d'instance à la demande par région.	Restreintes à 20 par zone de disponibilité. Une augmentation de limite peut être demandée.	Restreintes à 20 par région. Une augmentation de limite peut être demandée.	Aucune limite.

Pour plus d'informations, consultez les ressources suivantes :

- Instances réservées (p. 329)
- Guide de l'utilisateur Savings Plans

Plateformes prises en charge

Vous devez créer la Réservation de capacité avec la plate-forme appropriée pour vous assurer qu'elle correspond à vos instances. Les Réservations de capacité prennent en charge les plates-formes suivantes :

- Linux/UNIX
- Linux avec SQL Server Standard
- Linux avec SQL Server Web
- Linux avec SQL Server Enterprise

- Red Hat Enterprise Linux
- SUSE Linux

Pour plus d'informations sur les plates-formes Windows prises en charge, consultez Plates-formes prises en charge dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Limites de Réservation de capacité

Le nombre d'instances pour lesquelles vous êtes autorisé à réserver de la capacité est basé sur la limite d'instance à la demande de votre compte. Vous pouvez réserver de la capacité pour autant d'instances que cette limite le permet, moins le nombre d'instances que vous exécutez déjà.

Limites et restrictions d'une Réservation de capacité

Avant de créer des Réservations de capacité, prenez note des limitations et restrictions suivantes.

- Les Réservations de capacité actives et non utilisées sont prises en compte dans vos limites d'instance à la demande.
- Les Réservations de capacité ne sont pas transférables d'un compte AWS à un autre Toutefois, vous pouvez partager les Réservations de capacité avec d'autres comptes AWS. Pour plus d'informations, consultez Utilisation de Réservations de capacité partagées (p. 518).
- Les remises de facturation sur les Instance réservée par zone ne s'appliquent pas à Réservations de capacité
- Les Réservations de capacité ne peuvent pas être créées dans des groupes de placement
- Les Réservations de capacité ne peuvent pas être utilisées avec des Hôtes dédiés

Tarification et facturation d'une Réservation de capacité

Le prix d'un Réservation de capacité varie selon l'option de paiement.

Tarification

Lorsque la Réservation de capacité est active, vous êtes facturé selon l'équivalent du tarif à la demande, que vous exécutiez ou non les instances. Si vous n'utilisez pas la réservation, celle-ci apparaîtra en tant que réservation non utilisée sur votre facture EC2. Lorsque vous exécutez une instance qui correspond aux attributs d'une réservation, vous payez seulement pour l'instance, vous ne payez rien pour la réservation. Il n'y a aucun frais anticipé ou additionnel.

Par exemple, si vous créez une Réservation de capacité pour 20 instances Linux `m4.large` et que vous exécutez 15 instances Linux `m4.large` dans la même zone de disponibilité, vous serez facturé pour 15 instances actives et pour 5 instances non utilisées dans la réservation.

Les remises de facturation pour les Savings Plans et les Instances réservées régionales s'appliquent aux Réservations de capacité. Pour de plus amples informations, veuillez consulter Remises de facturation (p. 507).

Pour plus d'informations, consultez Tarification Amazon EC2.

Facturation

Les Réservations de capacité sont facturées à la seconde. Cela signifie que vous êtes facturé pour les heures partielles. Par exemple, si une réservation reste active dans votre compte pendant 24 heures et 15 minutes, vous serez facturé pour 24,25 heures de réservation.

L'exemple suivant présente la manière dont une Réservation de capacité est facturée. La Réservation de capacité est créée pour une instance Linux `m4.large`, dont le tarif à la demande est de 0,10 USD par heure d'utilisation. Dans cet exemple, la Réservation de capacité est active dans le compte pendant cinq heures. La Réservation de capacité n'étant pas utilisée la première heure, elle est facturée en tant qu'heure non utilisée au tarif à la demande standard du type d'instance `m4.large`. De la deuxième à la cinquième heure, la Réservation de capacité est occupée par une instance `m4.large`. Pendant ce laps de temps, la Réservation de capacité n'engendre pas de frais, et le compte est facturé pour l'instance `m4.large` qui l'occupe. Pour la sixième heure, la Réservation de capacité est annulée et l'instance `m4.large` s'exécute normalement en dehors de la capacité réservée. Cette heure est facturée selon le tarif à la demande du type d'instance `m4.large`.

Hour	1	2	3	
Unused Capacity Reservation	$0.10	$0.00	$0.00	$
On-demand Instance Usage	$0.00	$0.10	$0.10	$
Hourly cost	$0.10	$0.10	$0.10	$

Remises de facturation

Les remises de facturation pour les Savings Plans et les Instances réservées régionales s'appliquent aux Réservations de capacité. AWS applique automatiquement ces remises aux Réservations de capacité ayant des attributs correspondants. Lorsqu'un Réservation de capacité est utilisé par une instance, la remise est appliquée à cette instance. Les remises sont prioritairement appliquées à des instances en cours d'exécution avant de couvrir les Réservations de capacité inutilisées.

Les remises de facturation sur les Instances réservées zonales ne s'appliquent pas aux Réservations de capacité.

Pour plus d'informations, consultez les ressources suivantes :

- Instances réservées (p. 329)
- Guide de l'utilisateur Savings Plans

Affichage d'une facture

Vous pouvez vérifier les frais et les honoraires relatifs à votre compte sur la console AWS Billing and Cost Management.

- Le Tableau de bord affiche un récapitulatif des dépenses de votre compte.
- Sur la page Factures, sous Détails, développez la section Elastic Compute Cloud et la région pour obtenir des informations de facturation sur vos Réservations de capacité.

Vous pouvez consulter les frais en ligne ou télécharger un fichier CSV. Pour plus d'informations, consultez Postes de Réservation de capacité dans le Guide de l'utilisateur AWS Billing and Cost Management.

Utilisation de Réservations de capacité

Pour commencer à utiliser des Réservations de capacité, vous créez la réservation de capacité dans la zone de disponibilité requise. Vous pouvez ensuite lancer des instances dans la capacité réservée, afficher son utilisation de capacité en temps réel, et augmenter ou diminuer ses capacités en fonction de vos besoins.

Par défaut, les Réservations de capacité correspondent automatiquement aux nouvelles instances et aux instances en cours d'exécution dont les attributs correspondent (type d'instance, plateforme et zone de disponibilité). Cela signifie que toute instance avec des attributs correspondants est exécutée automatiquement dans la Réservation de capacité. Cependant, vous pouvez également cibler une Réservation de capacité pour des charges de travail spécifiques. Cela vous permet de contrôler explicitement les instances autorisées à s'exécuter dans cette capacité réservée.

Vous pouvez spécifiez comment votre réservation prend fin. Vous pouvez choisir d'annuler manuellement une Réservation de capacité ou de la terminer automatiquement à une date et une heure spécifiées. Si vous spécifiez une date et une heure de fin, la Réservation de capacité est annulée dans l'heure du moment spécifié. Par exemple, si vous spécifiez la date du 31/5/2019 à 13:30:55, la Réservation de capacité est assurée de prendre fin le 31/5/2019, entre 13:30:55 et 14:30:55. Lorsque la réservation prend fin, vous ne pouvez plus cibler d'instances sur la Réservation de capacité. Les instances en cours d'exécution dans la capacité réservée continuent à s'exécuter sans interruption. Si des instances ciblant une Réservation de capacité sont arrêtées, vous ne pouvez pas les redémarrer avant de supprimer leur préférence de ciblage de Réservation de capacité ou de les configurer de manière à cibler une Réservation de capacité différente.

Table des matières

- Création d'une Réservation de capacité (p. 508)
- Utilisation des groupes de Réservation de capacité (p. 510)
- Lancement d'instances dans une Réservation de capacité existante (p. 513)
- Modification d'une Réservation de capacité (p. 514)
- Modification des paramètres de Réservation de capacité d'une instance (p. 515)
- Affichage d'une Réservation de capacité (p. 516)
- Annulation d'une Réservation de capacité (p. 516)

Création d'une Réservation de capacité

Lorsque vous avez créé la Réservation de capacité, la capacité est disponible immédiatement. La capacité demeure réservée pour votre utilisation tant que la Réservation de capacité est active. Vous pouvez y lancer des instances à tout moment. Si la Réservation de capacité est ouverte, les nouvelles instances et les instances existantes dont les attributs correspondent s'exécutent automatiquement dans la capacité de la Réservation de capacité. Si la Réservation de capacité est `targeted`, les instances doivent la cibler spécifiquement pour s'exécuter dans la capacité réservée.

Votre demande de création d'une Réservation de capacité peut échouer si l'une des situations suivantes se produit :

- Amazon EC2 n'a pas une capacité suffisante pour répondre à la demande. Réessayez ultérieurement, essayez une zone de disponibilité différente ou essayez une capacité moins importante. Si votre application tolère plusieurs types et tailles d'instance, essayez des attributs d'instance différents.
- La quantité demandée dépasse votre limite d'instance à la demande pour la famille d'instance sélectionnée. Augmentez votre limite d'instance à la demande pour la famille d'instance requise et réessayez. Pour plus d'informations, consultez Limites instance à la demande (p. 326).

Pour créer une Réservation de capacité à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Réservations de capacité, puis Créer Réservation de capacité.
3. Sur la page Create a Réservation de capacité (Créer une capacité de réservation), configurez les paramètres suivants dans la section Instance details (Détails de l'instance). Le type d'instance, la plateforme et la zone de disponibilité des instances que vous lancez doivent correspondre au type d'instance, à la plateforme et à la zone de disponibilité que vous spécifiez ici ou la Réservation de

capacité ne s'applique pas. Par exemple, si un Réservation de capacité ouvert ne correspond pas, un lancement d'instance ciblant ce Réservation de capacité explicitement échouera.

a. Type d'instance : type d'instance à lancer dans la capacité réservée.
b. Launch EBS-optimized instances (Lancer des instances optimisées pour EBS) : spécifiez si vous souhaitez réserver la capacité pour des instances optimisées pour EBS. Cette option est sélectionnée par défaut pour certains types d'instances. Pour plus d'informations sur les instances optimisées pour EBS, consultez Amazon Elastic Block Store (p. 1084).
c. Attacher le stockage d'instance au lancement : spécifiez si les instances lancées dans la Réservation de capacité utilisent un stockage temporaire de niveau bloc. Les données figurant dans un volume de stockage d'instance ne sont maintenues que pendant la durée de vie de l'instance associée.
d. Plateforme : système d'exploitation pour vos instances. Pour de plus amples informations, veuillez consulter Plateformes prises en charge (p. 505). Pour plus d'informations sur les plates-formes Windows prises en charge, consultez Plates-formes prises en charge dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.
e. Zone de disponibilité : zone de disponibilité dans laquelle réserver la capacité.
f. Emplacement : spécifiez si vous voulez exécuter sur un matériel partagé (par défaut) ou une instance dédiée.
g. Quantité : nombre d'instances pour lesquelles vous souhaitez réserver la capacité. Si vous spécifiez une quantité qui dépasse votre limite d'instance à la demande restante pour le type d'instance sélectionné, la demande est refusée.

4. Configurez les paramètres suivants dans la section Reservation details (Détails de la réservation) :

   a. Reservation Ends (Fins de réservation) : choisissez une des options suivantes :

      - Manually (Manuellement) : réservez la capacité jusqu'à ce que vous l'annuliez de manière explicite.
      - Specific time (Date et heure spécifiques) : annule la réservation de capacité automatiquement à la date et à l'heure spécifiées.

   b. Instance eligibility (Éligibilité de l'instance) : choisissez une des options suivantes :

      - open (valeur par défaut) : la Réservation de capacité correspond à toute instance dont les attributs correspondent (type d'instance, plateforme, zone de disponibilité). Si vous lancez une instance avec les attributs correspondants, celle-ci est placée automatiquement dans la capacité réservée.
      - targeted : la Réservation de capacité accepte uniquement les instances dont les attributs correspondent (type d'instance, plateforme, zone de disponibilité) et qui ciblent explicitement la réservation.

5. Choisissez Request reservation (Demander une réservation).

Pour créer une Réservation de capacité avec l'AWS CLI

Utilisez la commande create-capacity-reservation. Pour de plus amples informations, veuillez consulter Plateformes prises en charge (p. 505). Pour plus d'informations sur les plates-formes Windows prises en charge, consultez Plates-formes prises en charge dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Par exemple, la commande suivante crée une Réservation de capacité qui réserve de la capacité pour trois instances `m5.2xlarge` exécutant des AMI Red Hat Enterprise Linux dans la zone de disponibilité `us-east-1a`.

```
aws ec2 create-capacity-reservation --instance-type m5.2xlarge --instance-platform Red Hat
 Enterprise Linux --availability-zone us-east-1a --instance-count 3
```

Utilisation des groupes de Réservation de capacité

Vous pouvez utiliser Groupes de ressources AWS pour créer des collections logiques de Réservations de capacité, appelées groupes de ressources. Un groupe de ressources est un regroupement logique de ressources AWS qui se trouvent toutes dans la même région AWS. Vous pouvez inclure plusieurs Réservations de capacité qui ont différents attributs (type d'instance, plate-forme et zone de disponibilité) dans un seul groupe de ressources.

Lorsque vous créez des groupes de ressources pour votre Réservations de capacité, vous pouvez cibler des instances vers un groupe de Réservations de capacité au lieu d'une Réservation de capacité seule. Les instances qui ciblent un groupe de Réservations de capacité correspondent à n'importe quelle Réservation de capacité du groupe disposant des attributs correspondants (type d'instance, plate-forme et zone de disponibilité) et de la capacité disponible. Si le groupe ne dispose pas d'une Réservation de capacité avec les attributs correspondants et de la capacité disponible, les instances s'exécutent à l'aide de la capacité à la demande. Si une Réservation de capacité adéquate est ajoutée au groupe cible à un stade ultérieur, l'instance est automatiquement mise en correspondance et déplacée vers sa capacité réservée.

Pour empêcher une utilisation non prévue des Réservations de capacité dans un groupe, configurez la Réservations de capacité dans le groupe pour accepter uniquement les instances qui ciblent explicitement la réserve de capacité. Pour ce faire, définissez l'éligibilité de l'instance sur ciblées (ancienne console) ou Uniquement des instances qui spécifient cette réservation (nouvelle console) lors de la création de la Réservation de capacité à l'aide de la console Amazon EC2. Lors de l'utilisation de l'AWS CLI, spécifiez `--instance-match-criteria targeted` lors de la création de la Réservation de capacité. On s'assure ainsi que seules les instances qui ciblent explicitement le groupe, ou une Réservation de capacité dans le groupe, peuvent s'exécuter dans le groupe.

Si une Réservation de capacité dans un groupe est annulée ou expire alors qu'elle a des instances en cours d'exécution, des dernières sont automatiquement déplacées vers une autre Réservation de capacité dans le groupe qui a des attributs correspondants et la capacité disponible. S'il ne reste pas de Réservations de capacité dans le groupe avec les attributs et la capacité disponible correspondants, les instances s'exécutent à l'aide de la capacité à la demande. Si une Réservation de capacité adéquate est ajoutée au groupe cible à un stade ultérieur, l'instance est automatiquement déplacée dans sa capacité réservée.

Pour créer un groupe pour vos Réservations de capacité

Utilisez la commande d'AWS CLI create-group. Pour `name`, indiquez un nom descriptif pour le groupe et pour `configuration`, spécifiez deux paramètres de `Type` demande :

- `AWS::EC2::CapacityReservationPool` pour s'assurer que le groupe de ressources peut être ciblé pour les lancements d'instances
- `AWS::ResourceGroups::Generic` avec `allowed-resource-types` définie sur `AWS::EC2::CapacityReservation` pour s'assurer que le groupe de ressources accepte uniquement les réserves de capacité

Par exemple, la commande suivante crée un groupe nommé `MyCRGroup`.

```
$ aws resource-groups create-group --name MyCRGroup --configuration
 '{"Type":"AWS::EC2::CapacityReservationPool"}' '{"Type":"AWS::ResourceGroups::Generic",
 "Parameters": [{"Name": "allowed-resource-types", "Values":
 ["AWS::EC2::CapacityReservation"]}]}'
```

Voici un exemple de sortie.

```
{
    "GroupConfiguration": {
        "Status": "UPDATE_COMPLETE",
```

```
        "Configuration": [
            {
                "Type": "AWS::EC2::CapacityReservationPool"
            },
            {
                "Type": "AWS::ResourceGroups::Generic",
                "Parameters": [
                    {
                        "Values": [
                            "AWS::EC2::CapacityReservation"
                        ],
                        "Name": "allowed-resource-types"
                    }
                ]
            }
        ]
    },
    "Group": {
        "GroupArn": "arn:aws:resource-groups:sa-east-1:123456789012:group/MyCRGroup",
        "Name": "MyCRGroup"
    }
}
```

Pour ajouter une Réservation de capacité à un groupe

Utilisez la commande AWS CLI group-resources. Pour `group`, spécifiez le nom du groupe auquel ajouter les Réservations de capacité, et pour `resources`, spécifiez les ARN des Réservations de capacité à ajouter. Pour ajouter plusieurs Réservations de capacité, séparez les ARN par un espace. Pour obtenir les ARN des Réservations de capacité à ajouter, utilisez la commande AWS CLI describe-capacity-reservations et spécifiez les ID des Réservations de capacité.

Par exemple, la commande suivante ajoute deux Réservations de capacité à un groupe nommé `MyCRGroup`.

```
$ aws resource-groups group-resources --group MyCRGroup --resource-arns arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/cr-1234567890abcdef1 arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/cr-54321abcdef567890
```

Voici un exemple de sortie.

```
{
    "Failed": [],
    "Succeeded": [
        "arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/cr-1234567890abcdef1",
        "arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/cr-54321abcdef567890"
    ]
}
```

Pour afficher les Réservations de capacité dans un groupe spécifique

Utilisez la commande d'AWS CLI list-group-resources. Pour `group`, spécifiez le nom du groupe.

Par exemple, la commande suivante répertorie les Réservations de capacité dans un groupe nommé `MyCRGroup`.

```
$ aws resource-groups list-group-resources --group MyCRGroup
```

Voici un exemple de sortie.

```
{
```

```
    "QueryErrors": [],
    "ResourceIdentifiers": [
        {
            "ResourceType": "AWS::EC2::CapacityReservation",
            "ResourceArn": "arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/
cr-1234567890abcdef1"
        },
        {
            "ResourceType": "AWS::EC2::CapacityReservation",
            "ResourceArn": "arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/
cr-54321abcdef567890"
        }
    ]
}
```

Pour afficher les groupes auxquels une Réservation de capacité spécifique a été ajoutée (AWS CLI)

Utilisez la commande d'AWS CLI get-groups-for-capacity-reservation.

Par exemple, la commande suivante répertorie les groupes auxquels une Réservation de capacité `cr-1234567890abcdef1` a été ajoutée.

```
$ aws ec2 get-groups-for-capacity-reservation --capacity-reservation-
id cr-1234567890abcdef1
```

Voici un exemple de sortie.

```
{
    "CapacityReservationGroups": [
        {
            "OwnerId": "123456789012",
            "GroupArn": "arn:aws:resource-groups:sa-east-1:123456789012:group/MyCRGroup"
        }
    ]
}
```

Pour afficher les groupes auxquels une Réservation de capacité spécifique a été ajoutée (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation Réservations de capacité, choisissez la Réservation de capacité à afficher, puis Afficher.

   Les groupes auxquels la Réservation de capacité a été ajoutée sont répertoriés dans la fiche Groupes.

Pour supprimer une Réservation de capacité d'un groupe

Utilisez la commande d'AWS CLI ungroup-resources. Pour `group`, spécifiez l'ARN du groupe duquel supprimer la Réservation de capacité, et pour `resources` spécifier les ARN des Réservations de capacité à supprimer. Pour supprimer plusieurs Réservations de capacité, séparez les ARN par un espace.

L'exemple suivant montre comment supprimer deux Réservations de capacité d'un groupe nommé `MyCRGroup`.

```
$ aws resource-groups ungroup-resources --group MyCRGroup --resource-arns arn:aws:ec2:sa-
east-1:123456789012:capacity-reservation/cr-0e154d26a16094dd arn:aws:ec2:sa-
east-1:123456789012:capacity-reservation/cr-54321abcdef567890
```

Voici un exemple de sortie.

```
{
    "Failed": [],
    "Succeeded": [
        "arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/cr-0e154d26a16094dd",
        "arn:aws:ec2:sa-east-1:123456789012:capacity-reservation/cr-54321abcdef567890"
    ]
}
```

Pour supprimer un groupe

Utilisez la commande d'AWS CLI delete-group. Pour `group` fournissez le nom du groupe à supprimer.

Par exemple, la commande suivante supprime un groupe appelé `MyCRGroup`.

```
$ aws resource-groups delete-group --group MyCRGroup
```

Voici un exemple de sortie.

```
{
    "Group": {
        "GroupArn": "arn:aws:resource-groups:sa-east-1:123456789012:group/MyCRGroup",
        "Name": "MyCRGroup"
    }
}
```

Lancement d'instances dans une Réservation de capacité existante

Lorsque vous lancez une instance, vous pouvez spécifier si elle doit être lancée dans n'importe quel Réservation de capacité `open`, dans une Réservation de capacité spécifique, ou dans un groupe de Réservations de capacité. Vous pouvez lancer une instance dans une Réservation de capacité si elle dispose des attributs correspondants (type d'instance, plate-forme et zone de disponibilité) et d'une capacité suffisante. Vous pouvez également configurer l'instance pour éviter qu'elle s'exécute dans une Réservation de capacité, même si vous avez une Réservation de capacité `open` qui a des attributs correspondants et la capacité disponible.

Le lancement d'une instance dans une Réservation de capacité réduit sa capacité disponible du nombre d'instances lancées. Par exemple, si vous lancez trois instances, la capacité disponible de la Réservation de capacité est réduite de trois.

Pour lancer des instances dans une Réservation de capacité existante à l'aide de la console

1. Ouvrez l'assistant de lancement d'instance en choisissant Launch Instances (Lancer des instances) depuis Dashboard (Tableau de bord) ou Instances.
2. Sélectionnez une Amazon Machine Image (AMI) et un type d'instance.
3. Remplissez la page Configurer les détails de l'instance. Pour Réservation de capacité, choisissez l'une des options suivantes :
   - None (Aucune) : empêche les instances de se lancer dans une Réservation de capacité. Les instances s'exécutent dans une capacité à la demande.
   - Open (Ouvrir) : lance les instances dans toute Réservation de capacité comportant des attributs correspondants et une capacité suffisante pour le nombre d'instances que vous avez sélectionnées. Si vous n'avez pas de Réservation de capacité correspondante avec une capacité suffisante, l'instance utilise une capacité à la demande.
   - Cible par ID — Lance les instances dans la Réservation de capacité sélectionnée. Si la Réservation de capacité sélectionnée ne dispose pas d'une capacité suffisante pour le nombre d'instances que vous avez sélectionnées, le lancement de l'instance échoue.

- Cible par groupe — Lance les instances dans n'importe quelle Réservation de capacité avec les attributs correspondants et la capacité disponible dans le groupe Réservation de capacité sélectionné. Si le groupe sélectionné ne dispose pas d'une Réservation de capacité avec les attributs correspondants et de la capacité disponible, les instances s'exécutent à l'aide de la capacité à la demande.

4. Complétez les étapes suivantes pour lancer les instances.

Pour lancer une instance dans une Réservation de capacité existante à l'aide de l'AWS CLI

Utilisez la commande run-instances et spécifiez le paramètre `--capacity-reservation-specification`.

L'exemple suivant lance une instance `t2.micro` dans toute Réservation de capacité ouverte disposant des attributs correspondants et de la capacité disponible :

```
aws ec2 run-instances --image-id ami-abc12345 --count 1 --instance-type t2.micro --key-
name MyKeyPair --subnet-id subnet-1234567890abcdef1 --capacity-reservation-specification
 CapacityReservationPreference=open
```

L'exemple suivant lance une instance `t2.micro` dans un `targeted` Réservation de capacité :

```
aws ec2 run-instances --image-id ami-abc12345 --count 1 --instance-type t2.micro --key-
name MyKeyPair --subnet-id subnet-1234567890abcdef1 --capacity-reservation-specification
 CapacityReservationTarget={CapacityReservationId=cr-a1234567}
```

L'exemple suivant lance une instance `t2.micro` dans un groupe Réservation de capacité :

```
aws ec2 run-instances --image-id ami-abc12345 --count 1 --instance-type t2.micro --key-
name MyKeyPair --subnet-id subnet-1234567890abcdef1 --capacity-reservation-specification
 CapacityReservationTarget={CapacityReservationResourceGroupArn=arn:aws:resource-groups:us-
west-1:123456789012:group/my-cr-group}
```

Modification d'une Réservation de capacité

Vous pouvez modifier les attributs d'un Réservation de capacité actif après l'avoir créé. Vous ne pouvez pas modifier une Réservation de capacité après son expiration ou après l'avoir explicitement annulée.

Lors de la modification d'une Réservation de capacité, vous pouvez uniquement augmenter ou diminuer la quantité et modifier la manière dont elle est libérée. Vous ne pouvez pas modifier le type d'instance, l'optimisation EBS, les paramètres de stockage d'instance, la plateforme, la zone de disponibilité ou l'éligibilité d'instance d'une Réservation de capacité. Si vous devez modifier un de ces attributs, nous vous recommandons d'annuler la réservation, puis d'en créer une nouvelle avec les attributs requis.

Si vous spécifiez une nouvelle quantité qui dépasse votre limite d'instance à la demande restante pour le type d'instance sélectionné, la mise à jour échoue.

Pour modifier une Réservation de capacité à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Réservations de capacité, sélectionnez la Réservation de capacité à modifier, puis choisissez Modifier.
3. Modifiez les options Quantité ou Reservation ends (Fins de réservation) selon vos besoins, puis choisissez Enregistrer les modifications.

Pour modifier une Réservation de capacité à l'aide de l'AWS CLI

Utilisez la commande modify-capacity-reservations :

Par exemple, la commande suivante modifie une Réservation de capacité pour réserver la capacité pour huit instances.

```
aws ec2 modify-capacity-reservation --capacity-reservation-id cr-1234567890abcdef0 --
instance-count 8
```

Modification des paramètres de Réservation de capacité d'une instance

Vous pouvez modifier les paramètres de Réservation de capacité pour une instance arrêtée à tout moment :

- Procédez au démarrage sur n'importe quelle Réservation de capacité disposant des attributs correspondants (type d'instance, plateforme et zone de disponibilité) et de la capacité disponible.
- Démarrez l'instance dans une Réservation de capacité spécifique.
- Démarrez l'instance dans n'importe quelle Réservation de capacité qui dispose des attributs correspondants et de la capacité disponible dans un groupe Réservation de capacité
- Empêchez l'instance de démarrer dans une Réservation de capacité.

Pour modifier les paramètres de la Réservation de capacité d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instances, puis sélectionnez l'instance à modifier. Arrêtez l'instance, si elle ne l'est pas déjà.
3. Choisissez Actions, Modify Réservation de capacité Settings (Modifier les paramètres de Réservation de capacité).
4. Pour Réservation de capacité, choisissez l'une des options suivantes :
   - Open (Ouvrir) : lance les instances dans toute Réservation de capacité comportant des attributs correspondants et une capacité suffisante pour le nombre d'instances que vous avez sélectionnées. Si vous n'avez pas de Réservation de capacité correspondante avec une capacité suffisante, l'instance utilise une capacité à la demande.
   - None (Aucune) : empêche les instances de se lancer dans une Réservation de capacité. Les instances s'exécutent dans une capacité à la demande.
   - Spécifier la réservation de capacité — Lance les instances dans la Réservation de capacité sélectionnée. Si la Réservation de capacité sélectionnée ne dispose pas d'une capacité suffisante pour le nombre d'instances que vous avez sélectionnées, le lancement de l'instance échoue.
   - Spécifier le groupe de réservation de capacité — Lance les instances dans n'importe quelle Réservation de capacité avec les attributs correspondants et la capacité disponible dans le groupe Réservation de capacité sélectionné. Si le groupe sélectionné ne dispose pas d'une Réservation de capacité avec les attributs correspondants et de la capacité disponible, les instances s'exécutent à l'aide de la capacité à la demande.

Pour modifier les paramètres de la Réservation de capacité d'une instance à l'aide de l'AWS CLI

Utilisez la commande modify-instance-capacity-reservation-attributes.

Par exemple, la commande suivante change le paramètre Réservation de capacité d'une instance pour `open` ou `none`.

```
aws ec2 modify-instance-capacity-reservation-attributes --instance-id i-1234567890abcdef0
 --capacity-reservation-specification CapacityReservationPreference=none|open
```

Par exemple, la commande suivante modifie une instance pour cibler une Réservation de capacité spécifique.

```
aws ec2 modify-instance-capacity-reservation-attributes --instance-
id i-1234567890abcdef0 --capacity-reservation-specification
 CapacityReservationTarget={CapacityReservationId=cr-1234567890abcdef0}
```

Par exemple, la commande suivante modifie une instance pour cibler un groupe Réservation de capacité spécifique.

```
aws ec2 modify-instance-capacity-reservation-attributes --instance-
id i-1234567890abcdef0 --capacity-reservation-specification
 CapacityReservationTarget={CapacityReservationResourceGroupArn=arn:aws:resource-groups:us-
west-1:123456789012:group/my-cr-group}
```

Affichage d'une Réservation de capacité

Réservations de capacité peut avoir les états suivants :

- `active` : la capacité peut être utilisée.
- `expired` : la Réservation de capacité a expiré automatiquement à la date et à l'heure spécifiées dans votre demande de réservation. La capacité réservée n'est plus disponible pour utilisation.
- `cancelled` : la Réservation de capacité a été annulée manuellement. La capacité réservée n'est plus disponible pour utilisation.
- `pending` : la demande de Réservation de capacité a abouti, mais la mise en service de la capacité est toujours en attente.
- `failed` : la demande de Réservation de capacité a échoué. Une demande peut échouer à cause de paramètres de demande non valides, de contraintes de capacité ou de contraintes de limite d'instance. Vous pouvez afficher une demande qui a échoué pendant 60 minutes.

Pour afficher vos Réservations de capacité à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Réservations de capacité puis sélectionnez une Réservation de capacité à afficher.
3. Choisissez View launched instances for this reservation (Afficher les instances lancées pour cette réservation).

Pour afficher vos Réservations de capacité à l'aide de l'AWS CLI

Utilisez la commande describe-capacity-reservations :

Par exemple, la commande suivante décrit toutes les Réservations de capacité.

```
aws ec2 describe-capacity-reservations
```

Annulation d'une Réservation de capacité

Vous pouvez annuler une Réservation de capacité à tout moment si vous n'avez plus besoin de la capacité réservée. Lorsque vous annulez une Réservation de capacité, la capacité est immédiatement libérée et n'est plus réservée pour votre utilisation.

Vous pouvez annuler des Réservations de capacité vides et des Réservations de capacité ayant des instances en cours d'exécution. Si vous annulez une Réservation de capacité avec des instances en cours d'exécution, les instances continuent leur exécution normale en dehors de la réservation de capacité aux

tarifs standard instance à la demande ou à un tarif réduit si vous avez un Savings Plan ou une Instance réservée régionale correspondant.

Une fois que vous avez annulé une Réservation de capacité, les instances la ciblant ne peuvent plus être lancées. Modifiez ces instances de sorte qu'elles ciblent une autre Réservation de capacité, lancez-les dans une Réservation de capacité « open » disposant des attributs correspondants et d'une capacité suffisante ou évitez de les lancer dans une Réservation de capacité. Pour plus d'informations, consultez Modification des paramètres de Réservation de capacité d'une instance (p. 515).

Pour annuler une Réservation de capacité à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Réservations de capacité et sélectionnez la Réservation de capacité à annuler.
3. Choisissez Cancel réservation (Annuler la réservation), Cancel réservation (Annuler la réservation).

Pour annuler une Réservation de capacité à l'aide de l'AWS CLI

Utilisez la commande cancel-capacity-reservation :

Par exemple, la commande suivante annule une Réservation de capacité avec un ID de `cr-1234567890abcdef0`.

```
aws ec2 cancel-capacity-reservation --capacity-reservation-id cr-1234567890abcdef0
```

Réservations de capacité dans Zones locales

Une Zone locale est une extension d'une région AWS qui est géographiquement proche de vos utilisateurs. Les ressources créées dans une Zone locale peuvent servir les utilisateurs locaux avec des communications à très faible latence. Pour de plus amples informations, veuillez consulter Zones locales AWS.

Vous pouvez étendre un VPC à partir de sa région AWS parente vers une Zone locale en créant un nouveau sous-réseau dans cette Zone locale. Lorsque vous créez un sous-réseau dans une Zone locale, votre VPC est étendu à cette Zone locale. Le sous-réseau de la Zone locale fonctionne de la même manière que les autres sous-réseaux de votre VPC.

En utilisant des Zones locales, vous pouvez placer des Réservations de capacité dans plusieurs emplacements qui sont plus proches de vos utilisateurs. Vous créez et utilisez des Réservations de capacité dans Zones locales de la même manière que vous créez et utilisez des Réservations de capacité dans les zones de disponibilité standard. Les fonctionnalités et le comportement de correspondance d'instance sont les mêmes. Pour plus d'informations sur les modèles de tarification pris en charge dans Zones locales, consultez les FAQ sur les zones locales AWS.

Limites

- Vous ne pouvez pas partager des Réservations de capacité qui sont créées dans une Zone locale.
- Vous ne pouvez pas utiliser de groupes Réservation de capacité dans une Zone locale.

Pour utiliser une Réservation de capacité dans une Zone locale

1. Activez la Zone locale pour une utilisation dans votre compte AWS. Pour de plus amples informations, veuillez consulter Activation des zones locales dans le Amazon EC2 Guide de l'utilisateur pour les instances Linux.
2. Créez une Réservation de capacité dans la Zone locale. Pour Zone de disponibilité, choisissez une Zone locale. Une Zone locale est représentée par un code de région AWS suivi d'un identifiant qui

indique l'emplacement ; par exemple, `us-west-2-lax-1a`. Pour de plus amples informations, veuillez consulter Création d'une Réservation de capacité (p. 508).

3. Créez un sous-réseau dans la Zone locale. Pour Zone de disponibilité, choisissez une Zone locale. Pour de plus amples informations, veuillez consulter Création d'un sous-réseau dans votre VPC dans le Amazon VPC Guide de l'utilisateur.
4. Lancez une instance. Pour Sous-réseau, choisissez le sous-réseau dans la Zone locale (par exemple `subnet-123abc | us-west-2-lax-1a`) et, pour Réservation de capacité, choisissez la spécification (`open` soit ciblez-la par ID) requise pour la Réservation de capacité que vous avez créée dans la Zone locale. Pour de plus amples informations, veuillez consulter Lancement d'instances dans une Réservation de capacité existante (p. 513).

Utilisation de Réservations de capacité partagées

Le partage de Réservation de capacité permet aux propriétaires d'une Réservation de capacité de partager leur capacité réservée avec d'autres comptes AWS ou au sein d'une organisation AWS. Cela vous permet de créer et de gérer des Réservations de capacité de manière centralisée, et de partager la capacité réservée entre plusieurs comptes AWS ou au sein de votre organisation AWS.

Dans ce modèle, le compte AWS auquel appartient la Réservation de capacité (propriétaire) la partage avec d'autres comptes AWS (consommateurs). Les consommateurs peuvent lancer des instances dans des Réservations de capacité partagées avec eux comme ils le feraient avec des Réservations de capacité qu'ils possèderaient dans leur propre compte. Le propriétaire d'une Réservation de capacité est responsable de la gestion de la Réservation de capacité et des instances lancées dans celle-ci. Les propriétaires ne peuvent pas modifier les instances lancées par les consommateurs dans des Réservations de capacité qu'ils ont partagées. Les consommateurs sont responsables de la gestion des instances qu'ils lancent dans des Réservations de capacité partagées avec eux. Les consommateurs ne peuvent pas voir ou modifier les instances appartenant à d'autres consommateurs ou au propriétaire de la Réservation de capacité.

Un propriétaire de Réservation de capacité peut partager une Réservation de capacité avec :

- Des comptes AWS spécifiques dans ou hors de son organisation AWS
- Une unité d'organisation dans son organisation AWS
- L'ensemble de son organisation AWS

Sommaire

- Conditions préalables au partage de Réservations de capacité (p. 518)
- Services connexes (p. 519)
- Partage sur plusieurs zones de disponibilité (p. 519)
- Partage d'une Réservation de capacité (p. 519)
- Arrêter de partager une Réservation de capacité (p. 520)
- Identification d'une Réservation de capacité partagée (p. 521)
- Affichage de l'utilisation d'une Réservation de capacité partagée (p. 521)
- Autorisations relatives à une Réservation de capacité partagée (p. 522)
- Facturation et mesures (p. 522)
- Limites d'instance (p. 522)

Conditions préalables au partage de Réservations de capacité

- Pour partager une Réservation de capacité, vous devez en disposer dans votre propre compte AWS. Vous ne pouvez pas partager une Réservation de capacité qui a été partagée avec vous.

- Vous pouvez uniquement partager des Réservations de capacité pour les instances de locations partagées. Vous ne pouvez pas partager de Réservations de capacité pour les instances de locations dédiées.
- Le partage de Réservation de capacité n'est pas disponible pour les nouveaux comptes AWS ou pour les comptes AWS qui présentent un historique de facturation limité. Les nouveaux comptes liés à un compte payeur qualifié ou à une organisation AWS sont dispensés de cette restriction.
- Pour partager une Réservation de capacité avec votre organisation AWS ou avec une unité d'organisation au sein de votre organisation AWS, vous devez activer le partage avec AWS Organizations. Pour plus d'informations, consultez la section relative à l'activation du partage avec des AWS Organizations dans le Guide de l'utilisateur AWS RAM.

Services connexes

Le partage d'une Réservation de capacité s'intègre à AWS Resource Access Manager (AWS RAM). AWS RAM est un service qui vous permet de partager vos ressources AWS avec n'importe quel compte AWS ou via des AWS Organizations. Avec AWS RAM, vous pouvez partager des ressources dont vous êtes propriétaire en créant un partage de ressources. Un partage de ressources spécifie les ressources à partager, ainsi que les consommateurs avec qui elles seront partagées. Les consommateurs peuvent être des comptes AWS individuels, des unités d'une organisation ou l'ensemble d'une organisation des AWS Organizations.

Pour plus d'informations sur AWS RAM, consultez le Guide de l'utilisateur AWS RAM.

Partage sur plusieurs zones de disponibilité

Pour garantir que les ressources sont réparties entre les zones de disponibilité d'une région, nous mappons indépendamment les zones de disponibilité aux noms de chaque compte. Cela peut entraîner des différences de nom de zone de disponibilité entre les comptes. Par exemple, la zone de disponibilité `us-east-1a` pour votre compte AWS peut avoir un emplacement autre que `us-east-1a` pour un autre compte AWS.

Pour identifier l'emplacement de vos Réservations de capacité par rapport à vos comptes, vous devez utiliser l'ID de zone de disponibilité. L'ID de zone de disponibilité est un identifiant unique et cohérent pour une zone de disponibilité entre tous les comptes AWS. Par exemple, `use1-az1` est un ID de zone de disponibilité pour la région `us-east-1`, qui correspond au même emplacement dans chaque compte AWS.

Pour afficher les ID de zone de disponibilité pour votre compte

1. Ouvrez la console AWS RAM à l'adresse https://console.aws.amazon.com/ram.
2. Les ID de zone de disponibilité pour la région actuelle sont affichés dans le volet Your AZ ID (Votre ID de zone de disponibilité) dans la partie droite de l'écran.

Partage d'une Réservation de capacité

Lorsque vous partagez une Réservation de capacité que vous possédez avec d'autres comptes AWS, vous autorisez ceux-ci à lancer des instances dans votre capacité réservée. Si vous partagez une Réservation de capacité ouverte, gardez présent à l'esprit les points suivants, car cela pourrait entraîner une utilisation indésirable de la Réservation de capacité :

- Si des consommateurs disposent d'instances en cours d'exécution correspondant aux attributs de la Réservation de capacité, du paramètre `CapacityReservationPreference` défini sur `open` et qu'ils ne procèdent pas à l'exécution dans une capacité réservée, ils utilisent automatiquement la Réservation de capacité partagée.
- Si des consommateurs lancent des instances disposant d'attributs correspondant (type d'instance, plateforme et zone de disponibilité) et du paramètre `CapacityReservationPreference` défini sur `open`, ils utilisent automatiquement la Réservation de capacité partagée.

Pour partager une Réservation de capacité, vous devez l'ajouter à un partage de ressources. Un partage de ressources est une ressource AWS RAM qui vous permet de partager vos ressources entre des comptes AWS. Un partage de ressources spécifie les ressources à partager, ainsi que les consommateurs avec qui elles seront partagées. Lorsque vous partagez une Réservation de capacité avec la console Amazon EC2, vous l'ajoutez à un partage de ressources existant. Pour ajouter une Réservation de capacité à un nouveau partage de ressources, vous devez créer le partage de ressources avec la console AWS RAM.

Si vous faites partie d'une organisation dans AWS Organizations et que le partage au sein de votre organisation est activé, l'accès à toute Réservation de capacité partagée est automatiquement accordé aux consommateurs de votre organisation. Dans le cas contraire, les consommateurs reçoivent une invitation à rejoindre le partage de ressources et bénéficient d'un accès à la Réservation de capacité partagée après avoir accepté l'invitation.

Vous pouvez partager une Réservation de capacité qui vous appartient avec la console Amazon EC2, la console AWS RAM ou l'AWS CLI.

Pour partager une Réservation de capacité qui vous appartient avec la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Réservations de capacité.
3. Choisissez la Réservation de capacité à partager, puis choisissez Actions, Share reservation (Partager une réservation).
4. Sélectionnez le partage de ressources auquel vous souhaitez ajouter la Réservation de capacité, puis choisissez Share Réservation de capacité (Partager la réservation de capacité).

   Les consommateurs peuvent avoir accès à la Réservation de capacité partagée en quelques minutes.

Pour partager une Réservation de capacité qui vous appartient avec la console AWS RAM

Consultez Création d'un partage de ressources dans le Guide de l'utilisateur AWS RAM.

Pour partager une Réservation de capacité qui vous appartient avec l'AWS CLI

Utilisez la commande create-resource-share.

Arrêter de partager une Réservation de capacité

Le propriétaire d'une Réservation de capacité peut cesser de partager une Réservation de capacité à tout moment. Les règles suivantes s'appliquent :

- Les instances appartenant aux consommateurs qui étaient en cours d'exécution dans la capacité partagée au moment où le partage s'arrête continuent de s'exécuter normalement en dehors de la capacité réservée, et la capacité est restaurée dans la Réservation de capacité soumise à la disponibilité de capacité Amazon EC2.
- Les consommateurs avec lesquels la Réservation de capacité était partagée ne peuvent plus lancer de nouvelles instances dans la capacité réservée.

Pour arrêter de partager une Réservation de capacité que vous possédez, vous devez la supprimer du partage de ressources. Pour ce faire, vous pouvez utiliser la console Amazon EC2, la console AWS RAM ou l'AWS CLI.

Pour arrêter le partage d'une Réservation de capacité que vous possédez à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le volet de navigation, choisissez Réservations de capacité.
3. Sélectionnez la Réservation de capacité et choisissez l'onglet Sharing (Partage).
4. L'onglet Sharing (Partage) affiche la liste des partages de ressources auxquels la Réservation de capacité a été ajoutée. Sélectionnez le partage de ressources duquel vous souhaitez supprimer la Réservation de capacité, puis choisissez Remove from resource share (Supprimer du partage de ressources).

Pour arrêter le partage d'une Réservation de capacité que vous possédez à l'aide de la console AWS RAM

Consultez Mise à jour d'un partage de ressources dans le Guide de l'utilisateur AWS RAM.

Pour arrêter de partager une Réservation de capacité que vous possédez à l'aide de l'AWS CLI

Utilisez la commande disassociate-resource-share.

Identification d'une Réservation de capacité partagée

Les propriétaires et les consommateurs peuvent identifier des Réservations de capacité partagées avec la console Amazon EC2 et l'AWS CLI

Pour identifier une Réservation de capacité partagée avec la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Réservations de capacité. L'écran affiche la liste des Réservations de capacité qui vous appartiennent et des Réservations de capacité qui sont partagées avec vous. La colonne Owner (Propriétaire) contient l'ID de compte AWS du propriétaire de la Réservation de capacité. `(me)` en regard de l'ID de compte AWS indique que vous êtes le propriétaire.

Pour identifier une Réservation de capacité partagée avec l'AWS CLI

Utilisez la commande describe-capacity-reservations : La commande renvoie les Réservations de capacité qui vous appartiennent et les Réservations de capacité qui sont partagées avec vous. `OwnerId` indique l'ID de compte AWS du propriétaire de la Réservation de capacité.

Affichage de l'utilisation d'une Réservation de capacité partagée

Le propriétaire d'une Réservation de capacité partagée peut afficher à tout moment son utilisation avec la console Amazon EC2 et l'AWS CLI.

Pour afficher l'utilisation d'une Réservation de capacité à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Réservations de capacité.
3. Sélectionnez la Réservation de capacité dont vous souhaitez voir l'utilisation et choisissez l'onglet Usage (Utilisation).

   La colonne AWS account ID (ID de compte AWS) indique l'ID de compte des consommateurs qui utilisent actuellement la Réservation de capacité. La colonne Launched instances (Instances lancées) indique le nombre d'instances en cours d'exécution pour chaque consommateur dans la capacité réservée.

Pour afficher l'utilisation d'une Réservation de capacité à l'aide de l'AWS CLI

Utilisez la commande get-capacity-reservation-usage. `AccountId` indique l'ID du compte utilisant la Réservation de capacité. `UsedInstanceCount` indique le nombre d'instances en cours d'exécution pour le consommateur dans la capacité réservée.

Autorisations relatives à une Réservation de capacité partagée

Autorisations accordées aux propriétaires

Les propriétaires sont responsables de la gestion et de l'annulation de leurs Réservations de capacité partagées. Les propriétaires ne peuvent pas modifier des instances appartenant à d'autres comptes et en cours d'exécution dans la Réservation de capacité. Les propriétaires sont responsables de la gestion des instances qu'ils lancent dans la Réservation de capacité partagée.

Autorisations accordées aux consommateurs

Les consommateurs sont responsables de la gestion de leurs instances exécutées dans la Réservation de capacité partagée. Les consommateurs ne peuvent pas modifier la Réservation de capacité partagée. Ils ne peuvent pas non plus afficher ou modifier des instances qui appartiennent à d'autres consommateurs ou au propriétaire de la Réservation de capacité.

Facturation et mesures

Le partage de Réservations de capacité n'entraîne pas de frais supplémentaires.

Le propriétaire de la Réservation de capacité est facturé pour les instances qu'il exécute dans la Réservation de capacité et pour la capacité réservée non utilisée. Les consommateurs sont facturés pour les instances qu'ils exécutent dans la Réservation de capacité partagée.

Limites d'instance

Toute utilisation d'une Réservation de capacité est prise en compte par rapport aux limites instance à la demande du propriétaire de la Réservation de capacité. Cela comprend :

- La capacité réservée non utilisée
- L'utilisation par des instances qui appartiennent au propriétaire de la Réservation de capacité
- L'utilisation par des instances qui appartiennent aux consommateurs

Les instances lancées dans la capacité partagée par des consommateurs sont prises en compte par rapport à la limite instance à la demande du propriétaire de la Réservation de capacité. Les limites d'instance des consommateurs sont égales à la somme de leurs propres limites instance à la demande et de la capacité disponible dans les Réservations de capacité partagées auxquelles ils ont accès.

Métriques CloudWatch pour Réservations de capacité à la demande

Grâce aux métriques CloudWatch, vous pouvez surveiller efficacement votre Réservations de capacité et identifier la capacité inutilisée en définissant des alarmes CloudWatch pour vous avertir lorsque les seuils d'utilisation sont atteints. Cela peut vous aider à maintenir un volume de Réservation de capacité constant et à atteindre un niveau d'utilisation plus élevé.

Les Réservations de capacité à la demande envoient toutes les cinq minutes des données métriques à CloudWatch. Les métriques ne sont pas prises en charge pour des Réservations de capacité qui sont actives pendant moins de cinq minutes.

Pour plus d'informations sur l'affichage des métriques dans la console CloudWatch, consultez Utilisation des métriques Amazon CloudWatch. Pour plus d'informations sur la création d’alarmes, consultez Création d'alarmes Amazon CloudWatch.

Sommaire

- Métriques d'utilisation Réservation de capacité (p. 523)
- Dimensions de métriques Réservation de capacité (p. 523)
- Affichage des métriques CloudWatch pour Réservations de capacité (p. 523)

Métriques d'utilisation Réservation de capacité

L'espace de nom `AWS/EC2CapacityReservations` inclut les mesures d'utilisation suivantes que vous pouvez employer pour surveiller et maintenir la capacité à la demande à l'intérieur des seuils que vous spécifiez pour votre réservation.

Métrique	Description
`UsedInstanceCount`	Nombre d'instances actuellement utilisées. Unité : nombre
`AvailableInstanceCount`	Nombre d'instances qui sont disponibles. Unité : nombre
`TotalInstanceCount`	Nombre total d'instances que vous avez réservées. Unité : nombre
`InstanceUtilization`	Pourcentage d'instances de capacité réservées qui sont actuellement utilisées. Unité : pourcentage

Dimensions de métriques Réservation de capacité

Vous pouvez utiliser les dimensions suivantes pour affiner les métriques répertoriées dans les tableaux précédents.

Dimension	Description
`CapacityReservationId`	Cette dimension globalement unique filtre uniquement les données que vous demandez pour la réservation de capacité identifiée.

Affichage des métriques CloudWatch pour Réservations de capacité

Les métriques sont d'abord regroupées par espaces de noms de service, puis par dimensions prises en charge. Vous pouvez utiliser les procédures ci-dessous pour afficher les métriques pour vos Réservations de capacité.

Pour afficher des métriques Réservation de capacité à l'aide de la console CloudWatch

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Si nécessaire, changez la région. Dans la barre de navigation, sélectionnez la région où réside la Réservation de capacité. Pour plus d'informations, consultez Régions et points de terminaison.
3. Dans le volet de navigation, choisissez Métriques.
4. Pour Toutes les mesures, choisissez Réservations de capacité EC2.

5. Choisissez la dimension de métrique Par réservation de capacité. Les métriques seront regroupées par `CapacityReservationId`.
6. Pour trier les métriques, utilisez l'en-tête de colonne. Pour représenter graphiquement une métrique, cochez la case en regard de la métrique.

Pour afficher des métriques Réservation de capacité (AWS CLI)

Utilisez la commande list-metrics suivante :

```
aws cloudwatch list-metrics --namespace "AWS/EC2CapacityReservations"
```

Cycle de vie d'une instance

Une instance Amazon EC2 passe par différents états entre le moment où vous la lancez et où vous la désactivez.

L'illustration suivante représente les transitions entre les états de l'instance. Notez que vous ne pouvez pas arrêter et démarrer une instance basée sur le stockage d'instance. Pour plus d'informations sur les instances basées sur le stockage d'instance, consultez Stockage pour le périphérique racine (p. 106).

Le tableau suivant décrit brièvement chaque état d'instance et indique s'il fait l'objet d'une facturation ou non.

Note

Le tableau indique la facturation pour l'utilisation de l'instance uniquement. Certaines ressources AWS, comme les volumes Amazon EBS et les adresses IP Elastic, entraînent des frais quel que

soit l'état de l'instance. Pour plus d'informations, consultez Éviter les frais inattendus dans le AWS Billing and Cost Management Guide de l'utilisateur.

État de l'instance	Description	Facturation de l'utilisation de l'instance
`pending`	L'instance se prépare à passer à l'état `running`. Une instance passe à l'état `pending` lorsqu'elle est lancée pour la première fois, ou lorsqu'elle est démarrée après avoir été à l'état `stopped`.	Non facturé
`running`	L'instance est en cours d'exécution et prête à être utilisée.	Facturé
`stopping`	L'instance se prépare à être arrêtée ou à entrer dans l'état arrêt-veille prolongée.	Non facturé en cas de préparation à l'arrêt Facturé en cas de préparation à la mise en veille prolongée
`stopped`	L'instance est arrêtée et ne peut pas être utilisée. L'instance peut être démarrée à tout moment.	Non facturé
`shutting down`	L'instance se prépare à être supprimée.	Non facturé
`terminated`	L'instance a été définitivement supprimée et ne peut pas être démarrée.	Non facturé Note Les instances réservées appliquées aux instances résiliées sont facturées jusqu'à la fin de leur période de validité, selon l'option de paiement. Pour plus d'informations, consultez Instances réservées (p. 329).

Note

Le redémarrage d'une instance ne commence pas une période de facturation d'une nouvelle instance, car l'instance reste à l'état `running`.

Lancement d'une instance

Lorsque vous lancez une instance, elle entre dans l'état `pending`. Le type d'instance que vous avez spécifié au lancement détermine les capacités matérielles de l'ordinateur hôte de votre instance. Nous utilisons l'Amazon Machine Image (AMI) que vous avez spécifié au lancement pour démarrer l'instance. Une fois que l'instance est prête, elle entre dans l'état `running`. Vous pouvez vous connecter à votre instance en cours d'exécution et l'utiliser comme vous le feriez d'un ordinateur devant lequel vous êtes assis.

Dès que votre instance passe à l'état `running`, chaque ou seconde avec un minimum d'une minute pendant laquelle votre instance s'exécute vous est facturée, même si l'instance demeure inactive et que vous ne vous y connectez pas.

Pour plus d'informations, consultez Lancer votre instance (p. 529) et Connectez-vous à votre instance Linux (p. 598).

Arrêt et démarrage d'une instance (instances basées sur les volumes Amazon EBS uniquement)

Si votre instance ne passe pas avec succès un contrôle de statut ou n'exécute pas vos applications comme escompté, et que le volume racine de votre instance est un volume Amazon EBS, vous pouvez arrêter et démarrer votre instance pour tenter de corriger le problème.

Lorsque vous arrêtez votre instance, elle entre dans l'état `stopping`, puis dans l'état `stopped`. Nous ne vous facturons ni l'utilisation ni les frais de transfert de données après que vous avez arrêté l'instance. En revanche, nous facturons le stockage des volumes Amazon EBS. Lorsque votre instance se trouve dans l'état `stopped`, vous pouvez modifier certains attributs de l'instance, y compris le type d'instance.

Lorsque vous démarrez votre instance, elle passe à l'état `pending` et nous déplaçons l'instance vers un nouvel ordinateur hôte (même si, dans la plupart des cas, elle reste sur l'hôte actuel). Lorsque vous arrêtez et démarrez votre instance, vous perdez toutes les données des volumes de stockage d'instance de l'ordinateur hôte précédent.

Votre instance conserve son adresse IPv4 privée, ce qui signifie qu'une adresse IP Elastic associée à l'adresse IPv4 privée ou à l'interface réseau continue d'être associée à votre instance. Si votre instance a une adresse IPv6, elle conserve cette dernière.

Chaque fois qu'une instance passe de l'état `stopped` à l'état `running`, nous facturons par seconde lorsque l'instance est en cours d'exécution, avec un minimum d'une minute chaque fois que vous démarrez votre instance.

Pour plus d'informations, consultez Arrêt et démarrage de votre instance (p. 625).

Mise en veille prolongée d'une instance (instances basées sur Amazon EBS uniquement)

Lorsque vous mettez une instance en veille prolongée, nous demandons au système d'exploitation d'exécuter l'opération correspondante (suspend-to-disk), ce qui enregistre le contenu de la mémoire de l'instance (RAM) sur votre volume racine Amazon EBS. Nous conservons le volume racine Amazon EBS de l'instance et les volumes de données Amazon EBS attachés. Lorsque vous démarrez votre instance, le volume racine Amazon EBS est restauré à son état précédent et le contenu de la mémoire RAM est rechargé. Les volumes de données précédemment attachés sont attachés à nouveau et l'instance conserve son ID d'instance.

Lorsque vous mettez votre instance en veille prolongée, elle entre dans l'état `stopping`, puis dans l'état `stopped`. Nous ne facturons pas l'utilisation de l'instance en veille prolongée pendant que cette dernière est à l'état `stopped`, mais nous la facturons quand elle est à l'état `stopping` (contrairement à ce qui se passe lorsque vous arrêtez une instance (p. 526) sans la mettre en veille prolongée). Nous ne facturons pas de frais de transfert de données pour l'utilisation. En revanche, nous facturons le stockage des volumes Amazon EBS, y compris le stockage des données de la mémoire RAM.

Lorsque vous démarrez votre instance mise en veille prolongée, elle passe à l'état `pending` et nous déplaçons l'instance vers un nouvel ordinateur hôte (même si, dans la plupart des cas, elle reste sur l'hôte actuel).

Votre instance conserve son adresse IPv4 privée, ce qui signifie qu'une adresse IP Elastic associée à l'adresse IPv4 privée ou à l'interface réseau continue d'être associée à votre instance. Si votre instance a une adresse IPv6, elle conserve cette dernière.

Pour plus d'informations, consultez Mettre votre instance Linux en veille prolongée (p. 628).

Redémarrage d'instance

Vous pouvez redémarrer votre instance à l'aide de la console Amazon EC2, d'un outil de ligne de commande et de l'API Amazon EC2. Nous vous recommandons d'utiliser Amazon EC2 pour redémarrer votre instance au lieu d'exécuter la commande de redémarrage du système d'exploitation à partir de votre instance.

Le redémarrage d'une instance est similaire à celui d'un système d'exploitation. L'instance demeure sur le même ordinateur hôte et conserve son nom DNS public, son adresse IP privée et les données de ses volumes de stockage d'instance. Le redémarrage nécessite généralement quelques minutes pour s'exécuter, mais le temps réel dépend de la configuration de l'instance.

Le redémarrage d'une instance ne déclenche pas de nouvelle période ; la facturation par seconde continue sans frais d'une minute minimum supplémentaire.

Pour plus d'informations, consultez Redémarrer votre instance (p. 641).

Mise hors service d'instance

Une instance est planifiée pour être mise hors service quand AWS détecte une défaillance irréparable du matériel sous-jacent hébergeant l'instance. Quand une instance atteint sa date de mise hors service planifiée, elle est arrêtée ou terminée par AWS. Si le périphérique racine de votre instance est un volume Amazon EBS, l'instance est arrêtée et vous pouvez la redémarrer à tout moment. Si le périphérique racine de votre instance est un volume de stockage d'instance, l'instance est terminée et ne peut pas être utilisée à nouveau.

Pour plus d'informations, consultez Mise hors service d'instance (p. 642).

Terminaison d'instance

Si vous jugez que vous n'avez plus besoin d'une instance, vous pouvez la mettre hors service. Dès que l'état d'une instance passe à `shutting-down` ou `terminated`, l'instance ne vous est plus facturée.

Si vous activez la protection de la résiliation, il ne vous est pas possible de résilier l'instance à l'aide de la console, de l'interface de ligne de commande ou de l'API.

Une fois que vous avez mis une instance hors service, elle demeure visible sur la console pendant un court instant, puis l'entrée est supprimée automatiquement. Vous pouvez aussi décrire une instance terminée à l'aide de l'interface ligne de commande ou de l'API. Les ressources (telles que les balises) sont progressivement dissociées de l'instance résiliées. Par conséquent, elles ne seront plus visibles dans l'instance terminée après un certain temps. Vous ne pouvez pas vous connecter à une instance terminée, ni la récupérer.

Chaque instance basée sur les volumes Amazon EBS prend en charge l'attribut `InstanceInitiatedShutdownBehavior`, qui contrôle si l'instance s'arrête ou se termine quand vous déclenchez un arrêt à partir de l'instance elle-même (par exemple, avec la commande shutdown sur Linux). Le comportement par défaut est celui de l'arrêt de l'instance. Vous pouvez modifier la valeur de cet attribut tandis que l'instance est en cours d'exécution ou arrêtée.

Chaque volume Amazon EBS prend en charge l'attribut `DeleteOnTermination`, qui contrôle si le volume est supprimé ou conservé lorsque vous terminez l'instance à laquelle il est attaché. Par défaut, le volume du périphérique racine est supprimé et les autres volumes EBS sont conservés.

Pour plus d'informations, consultez Résilier une instance (p. 645).

Différences entre redémarrage, arrêt, mise en veille prolongée et résiliation

Le tableau suivant résume les principales différences entre le redémarrage, l'arrêt, la mise en veille prolongée et la résiliation d'une instance.

Caractéristiqu	Redémarrer	Arrêt/démarrage (instances basées sur les volumes Amazon EBS uniquement)	Mise en veille prolongée (instances basées sur Amazon EBS uniquement)	Terminer
Ordinateur hôte	L'instance demeure sur le même ordinateur hôte.	Nous déplaçons l'instance vers un nouvel ordinateur hôte (même si, dans certains cas, elle reste sur l'hôte actuel).	Nous déplaçons l'instance vers un nouvel ordinateur hôte (même si, dans certains cas, elle reste sur l'hôte actuel).	Aucune
Adresses IPv4 publiques et privées	Ces adresses demeurent identiques.	L'instance conserve son adresse IPv4 privée. L'instance obtient une nouvelle adresse IPv4 publique, à moins qu'elle ne possède une adresse IP Elastic, laquelle ne change pas lors d'un arrêt/démarrage.	L'instance conserve son adresse IPv4 privée. L'instance obtient une nouvelle adresse IPv4 publique, à moins qu'elle ne possède une adresse IP Elastic, laquelle ne change pas lors d'un arrêt/démarrage.	Aucune
Adresses IP Elastic (IPv4)	L'adresse IP Elastic reste associée à l'instance	L'adresse IP Elastic reste associée à l'instance	L'adresse IP Elastic reste associée à l'instance	L'adresse IP Elastic est dissociée de l'instance.
Adresse IPv6	L'adresse reste la même.	L'instance conserve son adresse IPv6.	L'instance conserve son adresse IPv6.	Aucune
Volumes de stockage d'instance	Les données sont conservées.	Les données sont effacées.	Les données sont effacées.	Les données sont effacées.
volume du périphérique racine	Le volume est conservé	Le volume est conservé	Le volume est conservé	Le volume est supprimé par défaut.
RAM (contenu de la mémoire)	Les données de la mémoire RAM sont effacées.	Les données de la mémoire RAM sont effacées.	La mémoire RAM est enregistrée dans un fichier sur le volume racine.	Les données de la mémoire RAM sont effacées.
Facturation	L'heure de facturation de l'instance ne change pas.	Vous cessez d'être facturé aussitôt que l'état d'une instance devient `stopping`. Chaque fois qu'une instance passe de	Des frais vous sont facturés lorsque l'instance est à l'état `stopping`, mais ne le sont plus lorsque l'instance passe à l'état	Vous cessez d'être facturé aussitôt que l'état d'une instance devient `shutting-down`.

Caractéristiqu	Redémarrer	Arrêt/démarrage (instances basées sur les volumes Amazon EBS uniquement)	Mise en veille prolongée (instances basées sur Amazon EBS uniquement)	Terminer
		l'état `stopped` à `running`, nous démarrons une nouvelle période de facturation d'instance, en facturant un minimum d'une minute chaque fois que vous démarrez votre instance.	`stopped`. Chaque fois qu'une instance passe de l'état `stopped` à `running`, nous démarrons une nouvelle période de facturation d'instance, en facturant un minimum d'une minute chaque fois que vous démarrez votre instance.	

Les commandes d'arrêt du système d'exploitation terminent toujours une instance basée sur le stockage d'instance. Vous pouvez contrôler si les commandes d'arrêt du système d'exploitation arrêtent ou terminent une instance basée sur les volumes Amazon EBS. Pour plus d'informations, consultez Modification du comportement d'arrêt lancé de l'instance (p. 649).

Lancer votre instance

Une instance est un serveur virtuel dans le cloud AWS. Vous lancez une instance à partir d'une Amazon Machine Image (AMI). L'AMI fournit le système d'exploitation, le serveur d'applications, ainsi que les applications de votre instance.

Lorsque vous vous inscrivez à AWS, vous pouvez démarrer gratuitement avec Amazon EC2 grâce à l'offre gratuite AWS. Vous pouvez utiliser l'offre gratuite pour lancer et utiliser une instance `t2.micro` gratuitement pendant 12 mois (dans les régions où `t2.micro` n'est pas disponible, vous pouvez utiliser une instance `t3.micro` avec l'offre gratuite). Si vous lancez une instance qui ne fait pas partie de l'offre gratuite, les frais d'utilisation standard d'Amazon EC2 vous seront facturés pour l'instance. Pour plus d'informations, consultez Tarification Amazon EC2.

Vous pouvez lancer une instance à l'aide des méthodes suivantes.

Méthode	Documentation
[Console Amazon EC2] Utilisation de l'assistant de lancement d'instance pour spécifier les paramètres de lancement.	Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530)
[Console Amazon EC2] Création d'un modèle de lancement et lancement de l'instance à partir de celui-ci.	Lancement d'une instance à partir d'un modèle de lancement (p. 537)
[Console Amazon EC2] Utilisation d'une instance existante comme base.	Lancement d'une instance à l'aide des paramètres d'une instance existante (p. 555)
[Console Amazon EC2] Utilisation d'une AMI que vous avez achetée sur l'AWS Marketplace.	Lancement d'une instance AWS Marketplace (p. 557)
[AWS CLI] Utilisation d'une AMI que vous sélectionnez.	Utilisation de Amazon EC2 via l'interface de ligne de commande AWS

Méthode	Documentation
[Outils AWS pour Windows PowerShell] Utilisation d'une AMI que vous sélectionnez.	Amazon EC2 à partir des Outils AWS pour Windows PowerShell
[AWS CLI] Utilisez le parc d'instances EC2 pour allouer la capacité entre différents types d'instance EC2 et zones de disponibilité, et entre les modèles d'achat instance à la demande, Instance réservée et instance Spot.	Lancement d'instances à l'aide d'un Flotte EC2 (p. 558)

Lorsque vous lancez votre instance, vous pouvez le faire dans un sous-réseau associé à l'une des ressources suivantes :

- Une zone de disponibilité - Il s'agit de l'option par défaut.
- Une zone locale - Pour lancer une instance dans une zone locale, vous devez vous inscrire à la zone locale, puis créer un sous-réseau dans la zone. Pour de plus amples informations, veuillez consulter Zones locales.
- Une zone de longueur d'onde - Pour lancer une instance dans une zone de longueur d'onde, vous devez choisir la zone de longueur d'onde, puis créer un sous-réseau dans la zone. Pour de plus amples informations sur le lancement d'une instance dans une zone de longueur d'onde, veuillez consulter Démarrer avec AWS Wavelength dans le Manuel du développeur AWS Wavelength.
- Un Outpost - Pour lancer une instance dans un Outpost, vous devez créer un Outpost. Pour de plus amples informations sur la création d'un Outpost, veuillez consulter Démarrage avec AWS Outposts dans le Guide de l'utilisateur AWS Outposts.

Une fois que vous avez lancé votre instance, vous pouvez la connecter et l'utiliser. Au début, l'état de l'instance est `pending`. Lorsque l'état de l'instance indique `running`, cela signifie que le démarrage de l'instance a commencé. Il peut y avoir un bref délai avant que vous puissiez vous connecter à l'instance. Notez que le lancement de types d'instances nues peut prendre plus de temps. Pour de plus amples informations sur les instances nues, veuillez consulter Instances reposant sur le système Nitro (p. 218).

L'instance reçoit un nom DNS public que vous pouvez utiliser pour la contacter depuis Internet. L'instance reçoit également un nom DNS privé que d'autres instances au sein du même VPC peuvent utiliser pour la contacter. Pour plus d'informations sur la connexion à votre instance, consultez Connectez-vous à votre instance Linux (p. 598).

Lorsque vous n'avez plus besoin d'utiliser une instance, veillez à la mettre hors service. Pour plus d'informations, consultez Résilier une instance (p. 645).

Lancement d'une instance à l'aide de l'assistant de lancement d'instance

Vous pouvez lancer une instance à l'aide de l'assistant de lancement d'instance. L'assistant de lancement d'instance spécifie tous les paramètres de lancement requis pour lancer une instance. Lorsque l'assistant de lancement d'instance fournit une valeur par défaut, vous pouvez accepter la valeur par défaut ou spécifier votre propre valeur. À tout le moins, vous devez sélectionner une AMI et une paire de clés pour lancer une instance.

Avant de lancer l'instance, vérifiez que tout est prêt. Pour plus d'informations, consultez Configuration avec Amazon EC2 (p. 28).

> **Important**
>
> Lorsque vous lancez une instance qui ne fait pas partie de l'offre gratuite AWS, la durée d'exécution de l'instance vous est facturée, même si celle-ci reste inactive.

Étapes de lancement d'une instance :

- Commencer le lancement de l'instance (p. 531)
- Étape 1 : Sélection d'une Amazon Machine Image (AMI) (p. 531)
- Étape 2 : Choisir un type d'instance (p. 532)
- Étape 3 : Configurer les détails de l'instance (p. 532)
- Étape 4 : Ajouter du stockage (p. 535)
- Étape 5 : Ajouter des balises (p. 536)
- Étape 6 : Configurer un groupe de sécurité (p. 536)
- Étape 7 : Vérifier le lancement de l'instance et sélectionner une paire de clés (p. 536)

Commencer le lancement de l'instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation en haut de l'écran, la région actuelle est affichée (par exemple, USA Est (Ohio)). Sélectionnez une région pour l'instance répondant à vos besoins. Ce choix est important car certaines ressources Amazon EC2 peuvent être partagées entre plusieurs régions, contrairement à d'autres ressources. Pour plus d'informations, consultez Emplacements des ressources (p. 1301).
3. Sur le tableau de bord de la console Amazon EC2, sélectionnez Launch instance (Lancer une instance).

Étape 1 : Sélection d'une Amazon Machine Image (AMI)

Lorsque vous lancez une instance, vous devez sélectionner une configuration connue sous le nom d'Amazon Machine Image (AMI). Une AMI contient les informations nécessaires à la création d'une instance. Par exemple, une AMI peut contenir le logiciel nécessaire pour fonctionner en tant que serveur web : par exemple, Linux, Apache et votre site web.

Lorsque vous lancez une instance, vous pouvez sélectionner une AMI dans la liste ou sélectionner un paramètre Systems Manager pointant vers un ID AMI. Pour de plus amples informations, veuillez consulter Utilisation d'un paramètre Systems Manager pour rechercher une AMI.

Sur la page Choisir une image de machine Amazon (AMI) utilisez l'une des deux options pour choisir une AMI. Recherchez la liste des AMI (p. 531), ou effectuez une recherche par paramètre Systems Manager (p. 532).

En recherchant la liste des AMI

1. Sélectionnez le type d'AMI à utiliser dans le volet gauche :

   Démarrage rapide

   Quelques AMI couramment utilisées pour vous aider à démarrer rapidement. Pour sélectionner une AMI éligible pour l'offre gratuite, choisissez Offre gratuite uniquement dans le volet gauche. Ces AMI sont indiquées comme Admissible à l'offre gratuite.

   Mes AMI

   Les AMI privées que vous possédez ou qui ont été partagées avec vous. Pour voir les AMI partagées avec vous, choisissez Shared with me (Partagé avec moi) dans le volet de gauche.

   AWS Marketplace

   Boutique en ligne où vous pouvez acheter des logiciels exécutés sur AWS, et notamment les AMI. Pour plus d'informations sur le lancement d'une instance depuis l'AWS Marketplace, consultez Lancement d'une instance AWS Marketplace (p. 557).

AMI de la communauté

Les AMI qui ont été mises à la disposition des utilisateurs par des membres de la communauté AWS. Pour filtrer la liste des AMI par système d'exploitation, activez la case à cocher appropriée sous Système d'exploitation. Vous pouvez également filtrer par architecture et par type d'appareil racine.

2. Vérifiez le Type de périphérique racine spécifié pour chaque AMI. Déterminez les types d'AMI dont vous avez besoin, soit `ebs` (basé sur Amazon EBS) soit `instance-store` (basé sur le stockage d'instance). Pour plus d'informations, consultez Stockage pour le périphérique racine (p. 106).
3. Vérifiez le Type de virtualisation spécifié pour chaque AMI. Déterminez le type d'AMI dont vous avez besoin, `hvm` ou `paravirtual`. Par exemple, certains types d'instance requièrent HVM. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108).
4. Choisissez une AMI correspondant à vos besoins, puis choisissez Sélectionner.

Par paramètre Systems Manager

1. Choisissez Rechercher par paramètre Systems Manager (en haut à droite).
2. Pour Paramètre Systems Manager, sélectionnez un paramètre. L'ID AMI correspondant apparaît à côté de Currently resolves to (Se résout en).
3. Choisissez Search (Rechercher). Les AMI correspondant à l'ID AMI apparaissent dans la liste.
4. Sélectionnez l'AMI dans la liste, puis choisissez Select (Sélectionner).

Étape 2 : Choisir un type d'instance

Sur la page Choisir un type d'instance, sélectionnez la configuration matérielle et la taille de l'instance à lancer. Les types d'instance plus importants disposent de plus d'UC et de mémoire. Pour plus d'informations, consultez Types d'instance (p. 213).

Pour rester éligible pour l'offre gratuite, choisissez le type d'instance t2.micro (ou le type d'instance t3.micro dans des régions où t2.micro n'est pas disponible). Pour de plus amples informations, veuillez consulter Instances à capacité extensible (p. 233).

Par défaut, l'assistant affiche les types d'instance de la génération actuelle et sélectionne le premier type d'instance disponible en fonction de l'AMI que vous avez sélectionnée. Pour afficher les types d'instances de la génération précédente, choisissez Toutes les générations dans la liste de filtres.

Note

Pour configurer une instance rapidement à des fins de test, choisissez Vérifier et lancer afin d'accepter les paramètres de configuration par défaut, puis lancer votre instance. Sinon, pour configurer votre instance plus en détails, choisissez Suivant : Configurer les détails de l'instance.

Étape 3 : Configurer les détails de l'instance

Sur la page Configurer les détails de l'instance, modifiez les paramètres suivants en fonction de vos besoins (développez Détails avancés pour afficher tous les paramètres), puis choisissez Suivant : Ajouter le stockage :

- Nombre d'instances : entrez le nombre d'instances à lancer.

  Tip

  Pour accélérer les lancements d'instances, divisez les demandes volumineuses en lots plus petits. Par exemple, créez cinq demandes de lancement distinctes pour 100 instances au lieu d'un lancement pour 500 instances.
- (Facultatif) Pour vous assurer de conserver le nombre correct d'instances pour gérer la demande de votre application, vous pouvez choisir Lancer dans le groupe Auto Scaling pour créer une configuration

de lancement et un groupe Auto Scaling. Auto Scaling met à l'échelle le nombre d'instances du groupe en fonction de vos spécifications. Pour plus d'informations, consultez l'Amazon EC2 Auto Scaling Guide de l'utilisateur.

> **Note**
>
> Si Amazon EC2 Auto Scaling marque une instance qui se trouve dans un groupe Auto Scaling comme non saine, elle est automatiquement planifiée pour le remplacement lorsqu'elle est terminée et qu'une autre est lancée, et vous perdez vos données sur l'instance d'origine. Une instance est marquée comme non saine si vous arrêtez ou redémarrez l'instance, ou si un autre événement marque l'instance comme non saine. Pour de plus amples informations, veuillez consulter Vérification de l'état des instances Auto Scaling dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

- Option d'achat : sélectionnez Demander des instances Spot pour lancer une instance Spot. Cet utilitaire ajoute et supprime des options de cette page. Définissez le prix maximum et mettez éventuellement à jour le type de demande, le comportement d'interruption et la validité de la demande. Pour de plus amples informations, veuillez consulter Création d'une demande d'instance Spot (p. 395).
- Réseau : sélectionnez le VPC ou, pour créer un nouveau VPC, choisissez Créer un nouveau VPC afin d'accéder à la console Amazon VPC. Une fois que vous avez terminé, revenez dans l'assistant et choisissez Actualiser pour charger votre VPC dans la liste.
- Sous-réseau : vous pouvez lancer une instance dans un sous-réseau associé à une zone de disponibilité, une zone locale, une zone de longueur d'onde ou un Outpost.

  Pour lancer l'instance dans une zone de disponibilité, sélectionnez le sous-réseau dans lequel lancer votre instance. Vous pouvez sélectionner Aucune préférence pour laisser AWS choisir un sous-réseau par défaut dans une zone de disponibilité. Pour créer un sous-réseau, choisissez Créer un nouveau sous-réseau afin d'accéder à la console Amazon VPC. Une fois que vous avez terminé, revenez dans l'assistant et choisissez Actualiser afin de charger votre sous-réseau dans la liste.

  Pour lancer l'instance dans une Zone locale, sélectionnez un sous-réseau que vous avez créé dans la Zone locale.

  Pour lancer une instance dans un Outpost, sélectionnez un sous-réseau dans un VPC que vous avez associé à un Outpost.
- Attribuer automatiquement l'adresse IP publique : indiquez si l'instance est associée à une adresse IPv4 publique. Par défaut, les instances d'un sous-réseau par défaut se voient attribuer une adresse IPv4 publique, contrairement aux instances d'un sous-réseau personnalisé. Vous pouvez sélectionner Activer ou Désactiver pour remplacer la configuration par défaut du sous-réseau. Pour plus d'informations, consultez Adresses IPv4 publiques et noms d'hôte DNS externes (p. 809).
- Attribuer automatiquement l'adresse IP IPv6 : spécifiez si votre instance reçoit une adresse IPv6 de la plage du sous-réseau. Sélectionnez Activer ou Désactiver pour remplacer le paramètre par défaut du sous-réseau. Cette option est uniquement disponible si vous avez associé un bloc d'adresses CIDR IPv6 à votre VPC et à votre sous-réseau. Pour plus d'informations, consultez Vos VPC et sous-réseaux dans le Amazon VPC Guide de l'utilisateur.
- Répertoire de jonction de domaines : sélectionnez le répertoire AWS Directory Service (domaine) auquel votre instance Linux est jointe après le lancement. Si vous sélectionnez un domaine, vous devez sélectionner un rôle IAM avec les autorisations requises. Pour de plus amples informations, veuillez consulter Jonction en toute transparence d'une instance Linux EC2 à votre répertoire AWS Managed Microsoft AD.
- Groupe de placement : un groupe de placement détermine la stratégie de placement de vos instances. Sélectionnez un groupe de placement existant ou créez-en un nouveau. Cette option est disponible uniquement si vous avez sélectionné un type d'instance qui prend en charge les groupes de placement. Pour plus d'informations, consultez Groupes de placement (p. 924).
- Réservation de capacité : indiquez s'il convient de lancer l'instance dans la capacité partagée, dans une Réservation de capacité `open`, une Réservation de capacité spécifique ou un groupe Réservation de

capacité. Pour plus d'informations, consultez Lancement d'instances dans une Réservation de capacité existante (p. 513)..

- Rôle IAM : sélectionnez un rôle AWS Identity and Access Management (IAM) à associer à l'instance. Pour plus d'informations, consultez Rôles IAM pour Amazon EC2 (p. 1034).
- CPU options (Options d'UC) : choisissez Specify CPU options (Spécifier les options d'UC) pour spécifier un nombre personnalisé de vCPU lors du lancement. Définissez le nombre de cœurs d'UC et de threads par cœur. Pour plus d'informations, consultez Optimisations des options d'UC (p. 673).
- Comportement d'arrêt : indiquez si l'instance doit s'arrêter ou être résiliée lorsque vous arrêtez l'ordinateur. Pour plus d'informations, consultez Modification du comportement d'arrêt lancé de l'instance (p. 649).
- Comportement d'arrêt - mise en veille prolongée : pour activer la mise en veille prolongée, sélectionnez cette case à cocher. Cette option est uniquement disponible si votre instance satisfait les conditions préalables à la mise en veille prolongée. Pour plus d'informations, consultez Mettre votre instance Linux en veille prolongée (p. 628).
- Activer la protection de la résiliation : activez cette case à cocher pour éviter toute mise hors service accidentelle. Pour plus d'informations, consultez Activation de la protection contre la résiliation (p. 648).
- Monitoring (Surveillance) : activez cette case à cocher pour mettre en place une surveillance détaillée de votre instance à l'aide d'Amazon CloudWatch. Des frais supplémentaires seront facturés. Pour plus d'informations, consultez Surveillance de vos instances avec CloudWatch (p. 761).
- Instance optimisée pour EBS : une instance optimisée pour Amazon EBS a recours à une pile de configuration optimisée et fournit une capacité supplémentaire dédiée pour les I/O Amazon EBS. Si le type d'instance prend en charge cette fonction, coche cette case afin de l'activer. Des frais supplémentaires seront facturés. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).
- Location : si vous lancez votre instance sur un VPC, vous pouvez choisir de l'exécuter sur un matériel isolé dédié (Dédié) ou sur un hôte dédié (Hôte dédié). Des frais supplémentaires peuvent être facturés. Pour plus d'informations, consultez Instances dédiées (p. 498) et Hôtes dédiés (p. 465).
- T2/T3 illimité) : activez cette case à cocher pour permettre aux applications de s'exécuter au-delà du niveau de référence aussi longtemps que nécessaire. Des frais supplémentaires peuvent être facturés. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).
- Systèmes de fichiers : choisissez Add file system (Ajouter un système de fichiers) pour monter un ou plusieurs systèmes de fichiers Amazon EFS sur votre instance. Pour plus d'informations, consultez Utilisation de Amazon EFS avec Amazon EC2 (p. 1283).
- Interfaces réseau : si vous avez sélectionné un sous-réseau spécifique, vous pouvez spécifier jusqu'à deux interfaces réseau pour votre instance :
  - Pour Interfaces réseau, sélectionnez Nouvelle interface réseau pour laisser AWS créer une nouvelle interface, ou choisissez une interface réseau disponible existante.
  - Pour IP principale, entrez une adresse IPv4 privée de la plage de votre sous-réseau ou conservez Attribution automatique pour permettre à AWS de choisir une adresse IPv4 pour vous.
  - Pour Adresses IP secondaires, choisissez Ajouter l'IP pour affecter plusieurs adresses IPv4 privées à l'interface réseau sélectionnée.
  - (IPv6 uniquement) Pour Adresses IP IPv6, choisissez Ajouter l'IP et entrez une adresse IPv6 de la plage du sous-réseau, ou conservez Attribution automatique pour permettre à AWS de choisir une adresse pour vous.
  - Choisissez Ajouter périphérique pour ajouter une interface réseau secondaire. Une interface réseau secondaire peut résider dans un autre sous-réseau du VPC, à condition que celui-ci figure dans la même zone de disponibilité que votre instance.

  Pour plus d'informations, consultez Interfaces réseau Elastic (p. 840). Si vous spécifiez plusieurs interfaces réseau, votre instance ne peut recevoir aucune adresse IPv4 publique. En outre, si vous spécifiez une interface réseau existante pour eth0, vous ne pouvez pas remplacer le paramètre d'adresse IPv4 publique du sous-réseau à l'aide de Attribuer automatiquement l'adresse IP publique.

Pour plus d'informations, consultez Attribution d'une adresse IPv4 publique lors du lancement d'une instance (p. 813).

- ID du noyau : (valide uniquement pour les AMIs paravirtuelles, PV) sélectionnez Utiliser la valeur par défaut sauf si vous souhaitez utiliser un noyau spécifique.
- ID de disque RAM : (valide uniquement pour les AMIs paravirtuelles, PV) sélectionnez Utiliser la valeur par défaut sauf si vous souhaitez utiliser un disque RAM spécifique. Si vous avez sélectionné un noyau, vous devrez peut-être sélectionner un disque RAM spécifique avec les pilotes qui l'accompagnent.
- Metadata accessible (Métadonnées accessibles) : vous pouvez activer ou désactiver l'accès aux métadonnées de l'instance. Pour de plus amples informations, veuillez consulter Configuration du service des métadonnées d'instance (p. 702).
- Metadata version (Version des métadonnées) : si vous activez l'accès aux métadonnées de l'instance, vous pouvez choisir d'exiger l'utilisation de Service des métadonnées d'instance Version 2 lors de la demande de métadonnées d'instance. Pour de plus amples informations, veuillez consulter Configuration des options de métadonnées d'instance pour les nouvelles instances (p. 706).
- Metadata token response hop limit (Durée de vie de réponse du jeton de métadonnées) : si vous activez les métadonnées d'instance, vous pouvez définir le nombre autorisé de sauts réseau pour le jeton de métadonnées. Pour de plus amples informations, veuillez consulter Configuration du service des métadonnées d'instance (p. 702).
- Données utilisateur : vous pouvez spécifier les données utilisateur pour configurer une instance lors du lancement ou pour exécuter un script de configuration. Pour attacher un fichier, sélectionnez l'option Sous forme de fichier et parcourez la liste jusqu'à ce que vous trouviez le fichier à attacher.

Étape 4 : Ajouter du stockage

L'AMI sélectionnée inclut un ou plusieurs volumes de stockage, notamment le volume du périphérique racine. Sur la page Ajouter le stockage, vous pouvez spécifier des volumes supplémentaires à attacher à l'instance en choisissant Ajouter un nouveau volume. Configurez chaque volume comme suit, puis choisissez Suivant : Ajouter des balises.

- Type : sélectionnez le stockage d'instance ou les volumes Amazon EBS à associer à votre instance. Les types de volumes disponibles dans la liste dépendent du type d'instance que vous avez sélectionné. Pour plus d'informations, consultez Stockage d'instance Amazon EC2 (p. 1264) et Volumes Amazon EBS (p. 1086).
- Dispositif : sélectionnez l'appareil dans la liste des noms d'appareils disponibles pour le volume.
- Instantané : entrez le nom ou l'ID de l'instantané à partir duquel vous souhaitez restaurer un volume. Vous pouvez également rechercher les instantanés partagés et publics disponibles en saisissant un texte dans le champ Instantané. Les descriptions d'instantané sont sensibles à la casse.
- Taille : pour les volumes EBS, vous pouvez spécifier une taille de stockage. Même si vous avez sélectionné une AMI et une instance éligibles pour l'offre gratuite, afin de ne pas dépasser les limites de celle-ci, vous devez veiller à ne pas dépasser 30 GiO de stockage au total. Pour plus d'informations, consultez Contraintes sur la taille et la configuration d'un volume EBS (p. 1104).
- Type de volume : pour les volumes EBS, sélectionnez un type de volume. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).
- IOPS : si vous avez sélectionné le type de volume IOPS provisionnés SSD, vous pouvez entrer le nombre d'opérations d'I/O par seconde (IOPS) que le volume peut prendre en charge.
- Supprimer à la résiliation : pour les volumes Amazon EBS, activez cette case à cocher afin de supprimer le volume une fois l'instance résiliée. Pour plus d'informations, consultez Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649).
- Chiffré : si le type d'instance prend en charge le chiffrement EBS, vous pouvez spécifier l'état de chiffrement du volume. Si vous avez activé le chiffrement par défaut dans cette région, la clé CMK par défaut est sélectionnée automatiquement. Vous pouvez sélectionner une autre clé ou désactiver le chiffrement. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Étape 5 : Ajouter des balises

Sur la page Ajouter des balises, spécifiez les balises (p. 1309) en fournissant les combinaisons clé et valeur. Vous pouvez attribuer une balise à l'instance, aux volumes ou aux deux. Pour les Instances Spot, vous pouvez baliser la demande d'instance Spot uniquement. Choisissez Ajouter une autre balise pour ajouter plusieurs balises à vos ressources. Choisissez Suivant : Configurer le groupe de sécurité une fois que vous avez terminé.

Étape 6 : Configurer un groupe de sécurité

Sur la page Configurer le groupe de sécurité, utilisez un groupe de sécurité afin de définir les règles de pare-feu de votre instance. Ces règles déterminent le trafic réseau entrant acheminé vers votre instance. Le reste du trafic est ignoré. Pour plus d'informations sur les groupes de sécurité, consultez Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062). Sélectionnez ou créez un groupe de sécurité de la façon suivante, puis choisissez Vérifier et lancer.

- Pour sélectionner un groupe de sécurité existant, choisissez Select an existing security group (Sélectionner un groupe de sécurité existant), puis sélectionnez votre groupe de sécurité. Vous ne pouvez pas modifier les règles d'un groupe de sécurité existant, mais vous pouvez les copier dans un nouveau groupe en sélectionnant Copier vers le nouveau. Vous pouvez ensuite ajouter des règles, comme indiqué à l'étape suivante.
- Pour créer un nouveau groupe de sécurité, sélectionnez Create a new security group (Créer un nouveau groupe de sécurité). L'assistant définit automatiquement le groupe de sécurité launch-wizard-x et crée une règle entrante afin de vous permettre de vous connecter à l'instance via SSH (port 22).
- Vous pouvez ajouter des règles en fonction de vos besoins. Par exemple, si votre instance est un serveur web, ouvrez les ports 80 (HTTP) et 443 (HTTPS) afin d'autoriser le trafic Internet.

  Pour ajouter une règle, choisissez Ajouter une règle, sélectionnez le protocole à ouvrir au trafic réseau, puis spécifiez la source. Sélectionnez Mon IP dans la liste Source afin de laisser l'assistant ajouter votre adresse IP publique à l'ordinateur. Toutefois, si votre connexion s'effectue via un ISP ou derrière un pare-feu sans adresse IP statique, vous devez déterminer la plage d'adresses IP utilisée par les ordinateurs clients.

  > **Warning**
  >
  > Les règles qui permettent à toutes les adresses IP (`0.0.0.0/0`) d'accéder à votre instance via SSH ou RDP sont acceptables dans le cadre de cet court exercice, mais pas assez sécurisées pour un environnement de production. Veillez à autoriser une seule adresse IP ou plage d'adresses à accéder à votre instance.

Étape 7 : Vérifier le lancement de l'instance et sélectionner une paire de clés

Sur la page Examiner le lancement de l'instance, vérifiez les détails de votre instance, puis effectuez les modifications nécessaires en sélectionnant le lien Modifier approprié.

Une fois que vous êtes prêt, choisissez Lancer.

Dans la boîte de dialogue Select an existing key pair or create a new key pair (Sélectionner une paire de clés existante ou créer une nouvelle paire de clés), vous pouvez choisir une paire de clés existante ou en créer une nouvelle. Par exemple, sélectionnez Choisir une paire de clés existante, puis choisissez la paire de clés que vous avez créée lors de la configuration. Pour de plus amples informations, veuillez consulter Paires de clés Amazon EC2 et instances Linux (p. 1047).

> **Important**
>
> Si vous sélectionnez l'option Proceed without key pair (Continuer sans paire de clé), vous ne pourrez pas vous connecter à l'instance à moins de choisir une AMI configurée de façon à autoriser les utilisateurs à se connecter d'une autre façon.

Pour lancer votre instance, activez la case à cocher de confirmation, puis choisissez Lancer des instances.

(Facultatif) Vous pouvez créer une alarme de contrôle de statut pour l'instance (des frais supplémentaires peuvent être appliqués). En cas de doute, vous pouvez toujours en ajouter une par la suite. Sur l'écran de confirmation, choisissez Créer des alarmes de contrôle de statut et suivez les instructions. Pour plus d'informations, consultez Création et modification d'alarmes de contrôle de statut (p. 747).

Si l'instance ne peut pas être lancée ou que l'état passe immédiatement à `terminated` au lieu de `running`, consultez Résolution des problèmes de lancement d'instance (p. 1325).

Lancement d'une instance à partir d'un modèle de lancement

Vous pouvez créer un modèle de lancement contenant les informations de configuration pour lancer une instance. Les modèles de lancement vous permettent de stocker des paramètres de lancement, ce qui vous évite de devoir les spécifier à chaque lancement d'une instance. Par exemple, un modèle de lancement peut contenir l'ID d'AMI, le type d'instance et les paramètres réseau que vous utilisez généralement pour lancer des instances. Lorsque vous lancez une instance à l'aide de la console Amazon EC2, d'un kit SDK AWS ou d'un outil de ligne de commande, vous pouvez spécifier le modèle de lancement à utiliser.

Pour chaque modèle de lancement, vous pouvez créer une ou plusieurs versions de modèle de lancement numérotées. Chaque version peut comporter différents paramètres de lancement. Lorsque vous lancez une instance à partir d'un modèle de lancement, vous pouvez utiliser une version quelconque du modèle de lancement. Si vous ne spécifiez pas une version, c'est la version par défaut qui est utilisée. Vous pouvez définir n'importe quelle version du modèle de lancement comme version par défaut. Par défaut, il s'agit de la première version du modèle de lancement.

Le schéma suivant présente trois versions d'un modèle de lancement. La première version spécifie le type d'instance, l'ID d'AMI, le sous-réseau et la paire de clés à utiliser pour lancer l'instance. La deuxième version est basée sur la première et spécifie également un groupe de sécurité pour l'instance. La troisième version utilise différentes valeurs pour certains des paramètres. La version 2 est définie comme version par défaut. Si vous avez lancé une instance à partir de ce modèle de lancement, les paramètres de lancement de la version 2 sont utilisés si aucune autre version n'a été spécifiée.

Sommaire

- Restrictions du modèle de lancement (p. 538)
- Utilisation des modèles de lancement pour contrôler les paramètres de lancement (p. 538)
- Contrôle de l'utilisation des modèles de lancement (p. 538)
- Création d'un modèle de lancement (p. 539)
- Gestion des versions du modèle de lancement (p. 548)
- Lancement d'une instance à partir d'un modèle de lancement (p. 552)
- Utilisation de modèles de lancement avec Amazon EC2 Auto Scaling (p. 553)

- Utilisation de modèles de lancement avec Flotte EC2 (p. 554)
- Utilisation de modèles de lancement avec Parc d'instances Spot (p. 554)
- Suppression d'un modèle de lancement (p. 555)

Restrictions du modèle de lancement

Les règles suivantes s'appliquent aux modèles de lancement et à leurs versions :

- Vous ne pouvez pas créer plus de 5 000 modèles de lancement par région et plus de 10 000 versions par modèle de lancement.
- Les paramètres du modèle de lancement sont facultatifs. Néanmoins, vous devez vous assurer que votre demande de lancement d'une instance inclut tous les paramètres obligatoires. Par exemple, si votre modèle de lancement n'inclut pas un ID d'AMI, vous devez le spécifier avec un ID d'AMI lorsque vous lancez une instance.
- Les paramètres du modèle de lancement ne sont pas entièrement validés lorsque vous créez le modèle de lancement. Si vous spécifiez des valeurs incorrectes pour les paramètres ou si vous n'utilisez pas de combinaisons de paramètres prises en charge, aucune instance ne peut se lancer à l'aide de ce modèle de lancement. Veillez à spécifier les valeurs correctes des paramètres et à utiliser des combinaisons de paramètres prises en charge. Par exemple, pour lancer une instance dans un groupe de placement, vous devez spécifier un type d'instance pris en charge.
- Vous pouvez baliser un modèle de lancement, mais pas une version de modèle de lancement.
- Les versions de modèle de lancement sont numérotées dans l'ordre de leur création. Après avoir créé une version de modèle de lancement, vous ne pouvez pas spécifier vous-même le numéro de version.

Utilisation des modèles de lancement pour contrôler les paramètres de lancement

Un modèle de lancement peut contenir tout ou partie des paramètres permettant de lancer une instance. Lorsque vous lancez une instance à l'aide d'un modèle de lancement, vous pouvez remplacer les paramètres spécifiés dans le modèle de lancement. Vous pouvez également spécifier d'autres paramètres qui ne figurent pas dans le modèle de lancement.

> Note
>
> Vous ne pouvez pas supprimer les paramètres du modèle de lancement au cours du lancement (par exemple, vous ne pouvez pas spécifier une valeur null pour le paramètre). Pour supprimer un paramètre, créez une nouvelle version du modèle de lancement sans ce paramètre, puis utilisez cette version pour lancer l'instance.

Pour lancer des instances, les utilisateurs d'IAM doivent avoir l'autorisation d'utiliser l'action `ec2:RunInstances`. Vous devez également être autorisés à créer ou à utiliser les ressources créées ou associées à l'instance. Vous pouvez utiliser des autorisations au niveau des ressources pour l'action `ec2:RunInstances` afin de contrôler les paramètres de lancement pouvant être spécifiés par les utilisateurs. Vous pouvez également autoriser les utilisateurs à lancer une instance à l'aide d'un modèle de lancement. Cela vous permet de gérer les paramètres de lancement dans un modèle de lancement plutôt que dans une stratégie IAM et d'utiliser un modèle de lancement comment moyen d'autoriser le lancement d'instances. Par exemple, vous pouvez préciser que les utilisateurs peuvent uniquement lancer des instances à l'aide d'un modèle de lancement et qu'ils ne peuvent utiliser qu'un modèle de lancement particulier. Vous pouvez également contrôler les paramètres de lancement que les utilisateurs peuvent remplacer dans le modèle de lancement. Pour obtenir des exemples de stratégies, consultez Modèles de lancement (p. 1010).

Contrôle de l'utilisation des modèles de lancement

Par défaut, les utilisateurs d'IAM ne sont pas autorisés à utiliser des modèles de lancement. Vous pouvez créer une stratégie utilisateur IAM qui autorise les utilisateurs à créer, modifier, décrire et supprimer des

modèles de lancement et leurs versions. Vous pouvez également appliquer des autorisations au niveau des ressources à certaines actions de modèle de lancement pour contrôler la capacité d'un utilisateur à utiliser des ressources spécifiques pour ces actions. Pour plus d'informations, consultez Exemple : Utilisation des modèles de lancement (p. 1022).

Soyez vigilant lorsque vous autorisez des utilisateurs à effectuer les actions `ec2:CreateLaunchTemplate` et `ec2:CreateLaunchTemplateVersion`. Vous ne pouvez pas utiliser les autorisations au niveau des ressources pour contrôler les ressources que les utilisateurs peuvent spécifier dans le modèle de lancement. Pour limiter les ressources utilisées pour lancer une instance, veillez à autoriser uniquement les administrateurs appropriés à créer des modèles de lancement et des versions de modèles de lancement.

Création d'un modèle de lancement

Créez un modèle de lancement à l'aide des paramètres que vous définissez ou utilisez un modèle de lancement ou une instance existant(e) comme base d'un nouveau modèle de lancement.

Tâches

- Création d'un nouveau modèle de lancement à l'aide des paramètres que vous définissez (p. 539)
- Création d'un modèle de lancement à partir d'un modèle de lancement existant (p. 546)
- Création d'un modèle de lancement à partir d'une instance (p. 546)

Création d'un nouveau modèle de lancement à l'aide des paramètres que vous définissez

Nouvelle console

Pour créer un modèle de lancement à l'aide des paramètres définis dans la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement, puis Créer un modèle de lancement.
3. Pour Nom du modèle de lancement, entrez un nom descriptif pour le modèle.
4. Pour Description de la version du modèle, fournissez une brève description pour la version du modèle de lancement.
5. Pour baliser le modèle de lancement à la création, développez Template tags (Balises du modèle), Ajouter la balise, puis entrez une paire de clé et de valeur de balise.
6. Pour Contenu du modèle de lancement, indiquez les informations qui suivent :
   - AMI : AMI à partir de laquelle lancer l'instance. Pour effectuer une recherche parmi toutes les AMI disponibles, choisissez Search for AMI (Rechercher l'AMI). Pour sélectionner une AMI couramment utilisée, choisissez Lancement rapide. Vous pouvez aussi choisir AWS Marketplace ou AMI de la communauté. Vous pouvez utiliser une AMI qui vous appartient ou rechercher une AMI appropriée.
   - Type d'instance : assurez-vous que le type d'instance est compatible avec l'AMI spécifiée. Pour plus d'informations, consultez Types d'instance (p. 213).
   - Nom de la paire de clés : indiquez la paire de clés correspondant à l'instance. Pour plus d'informations, consultez Paires de clés Amazon EC2 et instances Linux (p. 1047).
   - Type de plateforme : le cas échéant, choisissez de lancer l'instance dans VPC ou dans EC2-Classic. Si vous choisissez VPC, indiquez le sous-réseau dans la section Interfaces réseau. Si vous choisissez Classic, vérifiez que le type d'instance spécifié est pris en charge dans EC2-Classic et spécifiez la zone de disponibilité de l'instance.
   - Groupes de sécurité : un ou plusieurs groupes de sécurité à associer à l'instance. Si vous ajoutez une interface réseau au modèle de lancement, omettez ce paramètre et spécifiez les groupes de sécurité dans le cadre de la spécification de l'interface réseau. Vous ne pouvez pas

lancer une instance à partir d'un modèle de lancement qui spécifie des groupes de sécurité et une interface réseau. Pour de plus amples informations, veuillez consulter Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062).

7. Pour Stockage (volumes), indiquez les volumes à attacher à l'instance, outre ceux qui sont spécifiés par l'AMI. (Volume 1 (AMI Root) (Volume 1 (racine AMI))) Pour ajouter un nouveau volume, choisissez Ajouter un volume.

   - Type de volume : volumes du stockage d'instance ou Amazon EBS auxquels associer votre instance. Le type de volume dépend du type d'instance que vous avez sélectionné. Pour plus d'informations, consultez Stockage d'instance Amazon EC2 (p. 1264) et Volumes Amazon EBS (p. 1086).
   - Nom du périphérique : nom de périphérique pour le volume.
   - Instantané : ID de l'instantané à partir duquel créer le volume.
   - Taille : taille du stockage des volumes Amazon EBS.
   - Type de volume : type de volume pour les volumes Amazon EBS. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).
   - IOPS : nombre d'opérations d'E/S par seconde que les volumes de type IOPS provisionnés SSD peuvent prendre en charge.
   - Supprimer à la résiliation : pour les volumes Amazon EBS, choisissez si le volume doit être supprimé une fois l'instance associée résiliée. Pour plus d'informations, consultez Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649).
   - Chiffré : si le type d'instance prend en charge le chiffrement EBS, vous pouvez activer le chiffrement pour le volume. Si vous avez activé le chiffrement par défaut dans cette région, le chiffrement est activé automatiquement. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).
   - Clé : clé CMK à utiliser pour le chiffrement EBS. Vous pouvez spécifier l'ARN de n'importe quelle clé principale client (clé CMK) que vous avez créée avec AWS Key Management Service. Si vous spécifiez une clé CMK, vous devez également utiliser Chiffré pour activer le chiffrement.
8. Pour le champ Resource tags (Balises de ressource), spécifiez les balises (p. 1309) en fournissant les combinaisons clé et valeur. Vous pouvez baliser l'instance, les volumes, les demandes instance Spot ou les trois.
9. Pour Interfaces réseau, vous pouvez indiquer jusqu'à deux interfaces réseau (p. 840) pour l'instance.

   - Index du périphérique : numéro du périphérique correspondant à l'interface réseau, par exemple `eth0` pour l'interface réseau principale. Si vous ne renseignez pas ce champ, AWS crée l'interface réseau principale.
   - Interface réseau : ID de l'interface réseau ; laissez ce champ vide pour permettre à AWS de créer une interface réseau.
   - Description : (facultatif) description de la nouvelle interface réseau.
   - Sous-réseau : sous-réseau dans lequel créer une interface réseau. Pour l'interface réseau principale (`eth0`), il s'agit du sous-réseau dans lequel l'instance est lancée. Si vous avez indiqué une interface réseau existante pour `eth0`, l'instance est lancée dans le sous-réseau dans lequel l'interface réseau est située.
   - Attribuer automatiquement l'adresse IP publique : spécifiez s'il convient d'attribuer automatiquement une adresse IP publique à l'interface réseau avec l'index de périphérique `eth0`. Ce paramètre ne peut être activé que pour une seule nouvelle interface réseau.
   - IP principale : une adresse IPv4 privée de la plage d'adresses de votre sous-réseau. Laissez ce champ vide pour qu'AWS choisisse une adresse IPv4 privée à votre place.
   - IP secondaire : adresse IPv4 privée secondaire de la plage d'adresses de votre sous-réseau. Laissez ce champ vide pour qu'AWS en choisisse une à votre place.
   - (IPv6 uniquement) Adresses IP IPv6 : adresse IPv6 comprise dans la plage du sous-réseau.

- Groupes de sécurité : un ou plusieurs groupes de sécurité de votre VPC auxquels associer l'interface réseau.
- Supprimer à la résiliation : indiquez s'il convient de supprimer l'interface réseau à la suppression de l'instance.
- Elastic Fabric Adapter (EFA) : Indique si l'interface réseau est une Elastic Fabric Adapter (EFA). Pour plus d'informations, consultez Elastic Fabric Adapter (EFA).

10. Développez la section Détails avancés pour afficher les champs et spécifier des paramètres supplémentaires pour l'instance.

    - Option d'achat : modèle d'achat. Choisissez Request Spot Instances (Demander des instances Spot) pour demander des Instances Spot au prix Spot, plafonné au prix des instances à la demande, et choisissez Personnaliser pour modifier les paramètres par défaut de l'instance Spot. Si vous ne demandez pas une instance Spot, EC2 lance une instance à la demande par défaut. Pour plus d'informations, consultez Instances Spot (p. 373).
    - Profil d'instance IAM : profil d'instance AWS Identity and Access Management (IAM) à associer à l'instance. Pour plus d'informations, consultez Rôles IAM pour Amazon EC2 (p. 1034).
    - Comportement d'arrêt : indiquez si l'instance doit s'arrêter ou être mise hors service lorsque vous arrêtez l'ordinateur. Pour plus d'informations, consultez Modification du comportement d'arrêt lancé de l'instance (p. 649).
    - Stop - Hibernate behavior (Arrêt - Veille prolongée) : indiquez si la mise en veille prolongée est activée pour l'instance. Ce champ est valide uniquement pour les instances pour lesquelles les prérequis de mise en veille prolongée sont satisfaits. Pour plus d'informations, consultez Mettre votre instance Linux en veille prolongée (p. 628).
    - Protection de la résiliation : indiquez s'il convient d'éviter toute résiliation accidentelle. Pour plus d'informations, consultez Activation de la protection contre la résiliation (p. 648).
    - Detailed CloudWatch monitoring (Surveillance CloudWatch détaillée) : indiquez s'il convient d'activer la surveillance détaillée de l'instance à l'aide de Amazon CloudWatch. Des frais supplémentaires seront facturés. Pour plus d'informations, consultez Surveillance de vos instances avec CloudWatch (p. 761).
    - Elastic inference (Inférence Elastic) : accélérateur d'inférence Elastic à attacher à votre instance de CPU EC2. Pour de plus amples informations, veuillez consulter Utilisation d'Amazon Elastic Inference dans le Guide du développeur Amazon Elastic Inference.
    - T2/T3 Unlimited (T2/T3 illimité) : indiquez s'il convient de permettre aux applications de s'exécuter au-delà du niveau de référence aussi longtemps que nécessaire. Ce champ n'est valide que pour les instances T2, T3 et T3a. Des frais supplémentaires peuvent être facturés. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).
    - Nom du groupe de placement : indiquez un groupe de placement dans lequel lancer l'instance. Le lancement dans un groupe de placement n'est pas possible pour tous les types d'instance. Pour plus d'informations, consultez Groupes de placement (p. 924).
    - Instance optimisée pour EBS : fournit une capacité supplémentaire dédiée pour les I/O Amazon EBS. Cette fonction n'est pas prise en charge par tous les types d'instance et elle implique des frais supplémentaires. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).
    - Réservation de capacité : indiquez s'il convient de lancer l'instance dans la capacité partagée, dans une Réservation de capacité `open`, une Réservation de capacité spécifique ou un groupe Réservation de capacité. Pour de plus amples informations, veuillez consulter Lancement d'instances dans une Réservation de capacité existante (p. 513).
    - Location : indiquez s'il convient d'exécuter votre instance sur un matériel partagé (Partagé), isolé, dédié (Dédié) ou sur un Hôte dédié (Hôte dédié). Si vous choisissez de lancer l'instance sur un Hôte dédié, vous pouvez spécifier si l'instance doit être lancée dans un groupe de ressources hôte ou vous pouvez cibler un Hôte dédié spécifique. Des frais supplémentaires peuvent être facturés. Pour plus d'informations, consultez Instances dédiées (p. 498) et Hôtes dédiés (p. 465).

- ID de disque RAM : (Valide uniquement pour les AMIs paravirtuelles (PV)) Disque RAM pour l'instance. Si vous avez sélectionné un noyau, vous devrez peut-être spécifier un disque RAM spécifique avec les pilotes qui l'accompagnent.
- ID du noyau : (Valide uniquement pour les AMIs paravirtuelles (PV)) Noyau pour l'instance.
- Configurations de licence : vous pouvez lancer des instances sur la configuration de licence spécifiée pour suivre l'utilisation de votre licence. Pour de plus amples informations, veuillez consulter Création d'une configuration de licence dans le Guide de l'utilisateur License Manager AWS.
- Metadata accessible (Métadonnées accessibles) : pour activer ou désactiver l'accès aux métadonnées de l'instance. Pour de plus amples informations, veuillez consulter Configuration du service des métadonnées d'instance (p. 702).
- Metadata version (Version des métadonnées) : si vous activez l'accès aux métadonnées de l'instance, vous pouvez choisir d'exiger l'utilisation de Service des métadonnées d'instance Version 2 lors de la demande de métadonnées d'instance. Pour de plus amples informations, veuillez consulter Configuration des options de métadonnées d'instance pour les nouvelles instances (p. 706).
- Limite de sauts de réponse de jeton de métadonnées : si vous activez les métadonnées d'instance, vous pouvez définir le nombre autorisé de sauts réseau pour le jeton de métadonnées. Pour de plus amples informations, veuillez consulter Configuration du service des métadonnées d'instance (p. 702).
- Données utilisateur : vous pouvez spécifier les données utilisateur pour configurer une instance lors du lancement ou pour exécuter un script de configuration. Pour plus d'informations, consultez Exécution de commandes sur votre instance Linux lors du lancement (p. 694).

11. Choisissez Créer un modèle de lancement.

Ancienne console

Pour créer un modèle de lancement à l'aide des paramètres définis dans la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement, puis Créer un modèle de lancement.
3. Pour Nom du modèle de lancement, entrez un nom descriptif pour le modèle. Pour baliser le modèle de lancement à la création, choisissez Afficher les balises, Ajouter une balise, puis entrez une paire de clé et de valeur de balise.
4. Pour Description de la version du modèle, fournissez une brève description pour la version du modèle de lancement.
5. Pour Contenu du modèle de lancement, indiquez les informations qui suivent :

   - ID d'AMI : AMI à partir de laquelle lancer l'instance. Pour effectuer une recherche parmi toutes les AMI disponibles, choisissez Search for AMI (Rechercher l'AMI). Pour sélectionner une AMI couramment utilisée, choisissez Lancement rapide. Vous pouvez aussi choisir AWS Marketplace ou AMI de la communauté. Vous pouvez utiliser une AMI qui vous appartient ou rechercher une AMI appropriée.
   - Type d'instance : assurez-vous que le type d'instance est compatible avec l'AMI spécifiée. Pour plus d'informations, consultez Types d'instance (p. 213).
   - Nom de la paire de clés : indiquez la paire de clés correspondant à l'instance. Pour plus d'informations, consultez Paires de clés Amazon EC2 et instances Linux (p. 1047).
   - Type de réseau : le cas échéant, choisissez de lancer l'instance dans VPC ou EC2-Classic. Si vous choisissez VPC, indiquez le sous-réseau dans la section Interfaces réseau. Si vous choisissez Classic, vérifiez que le type d'instance spécifié est pris en charge dans EC2-Classic et spécifiez la zone de disponibilité de l'instance.

- Groupes de sécurité : un ou plusieurs groupes de sécurité à associer à l'instance. Pour de plus amples informations, veuillez consulter Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062).

6. Pour Interfaces réseau, vous pouvez indiquer jusqu'à deux interfaces réseau (p. 840) pour l'instance.

   - Périphérique : numéro du périphérique correspondant à l'interface réseau, par exemple `eth0` pour l'interface réseau principale. Si vous ne renseignez pas ce champ, AWS crée l'interface réseau principale.
   - Interface réseau : ID de l'interface réseau ; laissez ce champ vide pour permettre à AWS de créer une interface réseau.
   - Description : (facultatif) description de la nouvelle interface réseau.
   - Sous-réseau : sous-réseau dans lequel créer une interface réseau. Pour l'interface réseau principale (`eth0`), il s'agit du sous-réseau dans lequel l'instance est lancée. Si vous avez indiqué une interface réseau existante pour `eth0`, l'instance est lancée dans le sous-réseau dans lequel l'interface réseau est située.
   - Attribuer automatiquement l'adresse IP publique : spécifiez s'il convient d'attribuer automatiquement une adresse IP publique à l'interface réseau avec l'index de périphérique `eth0`. Ce paramètre ne peut être activé que pour une seule nouvelle interface réseau.
   - IP principale : une adresse IPv4 privée de la plage d'adresses de votre sous-réseau. Laissez ce champ vide pour qu'AWS choisisse une adresse IPv4 privée à votre place.
   - IP secondaire : adresse IPv4 privée secondaire de la plage d'adresses de votre sous-réseau. Laissez ce champ vide pour qu'AWS en choisisse une à votre place.
   - (IPv6 uniquement) Adresses IP IPv6 : adresse IPv6 comprise dans la plage du sous-réseau.
   - ID de groupe de sécurité : ID d'un groupe de sécurité de votre VPC auquel associer l'interface réseau.
   - Supprimer à la résiliation : indiquez s'il convient de supprimer l'interface réseau à la suppression de l'instance.
   - Elastic Fabric Adapter (EFA) : Indique si l'interface réseau est une Elastic Fabric Adapter (EFA). Pour plus d'informations, consultez Elastic Fabric Adapter (EFA).

7. Pour Stockage (Volumes), indiquez les volumes à attacher à l'instance, outre ceux qui sont spécifiés par l'AMI.

   - Type de volume : volumes du stockage d'instance ou Amazon EBS auxquels associer votre instance. Le type de volume dépend du type d'instance que vous avez sélectionné. Pour plus d'informations, consultez Stockage d'instance Amazon EC2 (p. 1264) et Volumes Amazon EBS (p. 1086).
   - Nom du périphérique : nom de périphérique pour le volume.
   - Instantané : ID de l'instantané à partir duquel créer le volume.
   - Taille : taille du stockage des volumes Amazon EBS.
   - Type de volume : type de volume pour les volumes Amazon EBS. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).
   - IOPS : nombre d'opérations d'E/S par seconde que les volumes de type IOPS provisionnés SSD peuvent prendre en charge.
   - Supprimer à la résiliation : pour les volumes Amazon EBS, choisissez si le volume doit être supprimé une fois l'instance associée résiliée. Pour plus d'informations, consultez Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649).
   - Chiffré : si le type d'instance prend en charge le chiffrement EBS, vous pouvez activer le chiffrement pour le volume. Si vous avez activé le chiffrement par défaut dans cette région, le chiffrement est activé automatiquement. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

- Clé : clé CMK à utiliser pour le chiffrement EBS. Vous pouvez spécifier l'ARN de n'importe quelle clé principale client (clé CMK) que vous avez créée avec AWS Key Management Service. Si vous spécifiez une clé CMK, vous devez également utiliser Chiffré pour activer le chiffrement.

8. Pour le champ Instance tags (Balises d'instance), spécifiez les balises (p. 1309) en fournissant les combinaisons clé et valeur. Vous pouvez attribuer une balise à l'instance, aux volumes ou aux deux.
9. Développez la section Détails avancés pour afficher les champs et spécifier des paramètres supplémentaires pour l'instance.
   - Option d'achat : modèle d'achat. Choisissez Demander des instances Spot pour demander des Instances Spot au prix Spot, plafonné au prix des instances à la demande, et choisissez Customize Spot parameters (Personnaliser les paramètres Spot) pour modifier les paramètres instance Spot par défaut. Si vous ne demandez pas une instance Spot, EC2 lance une instance à la demande par défaut. Pour plus d'informations, consultez Instances Spot (p. 373).
   - Profil d'instance IAM : profil d'instance AWS Identity and Access Management (IAM) à associer à l'instance. Pour plus d'informations, consultez Rôles IAM pour Amazon EC2 (p. 1034).
   - Comportement d'arrêt : indiquez si l'instance doit s'arrêter ou être mise hors service lorsque vous arrêtez l'ordinateur. Pour plus d'informations, consultez Modification du comportement d'arrêt lancé de l'instance (p. 649).
   - Stop - Hibernate behavior (Arrêt - Veille prolongée) : indiquez si la mise en veille prolongée est activée pour l'instance. Ce champ est valide uniquement pour les instances pour lesquelles les prérequis de mise en veille prolongée sont satisfaits. Pour plus d'informations, consultez Mettre votre instance Linux en veille prolongée (p. 628).
   - Protection de la résiliation : indiquez s'il convient d'éviter toute résiliation accidentelle. Pour plus d'informations, consultez Activation de la protection contre la résiliation (p. 648).
   - Surveillance : indiquez s'il convient d'activer la surveillance détaillée de l'instance via Amazon CloudWatch. Des frais supplémentaires seront facturés. Pour plus d'informations, consultez Surveillance de vos instances avec CloudWatch (p. 761).
   - T2/T3 Unlimited (T2/T3 illimité) : indiquez s'il convient de permettre aux applications de s'exécuter au-delà du niveau de référence aussi longtemps que nécessaire. Ce champ est valable uniquement pour les instances T2 et T3. Des frais supplémentaires peuvent être facturés. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).
   - Nom du groupe de placement : indiquez un groupe de placement dans lequel lancer l'instance. Le lancement dans un groupe de placement n'est pas possible pour tous les types d'instance. Pour plus d'informations, consultez Groupes de placement (p. 924).
   - Instance optimisée pour EBS : fournit une capacité supplémentaire dédiée pour les I/O Amazon EBS. Cette fonction n'est pas prise en charge par tous les types d'instance et elle implique des frais supplémentaires. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).
   - Location : indiquez s'il convient d'exécuter votre instance sur un matériel partagé (Partagé), isolé, dédié (Dédié) ou sur un Hôte dédié (Hôte dédié). Si vous choisissez de lancer l'instance sur un Hôte dédié, vous pouvez spécifier si l'instance doit être lancée dans un groupe de ressources hôte ou vous pouvez cibler un Hôte dédié spécifique. Des frais supplémentaires peuvent être facturés. Pour plus d'informations, consultez Instances dédiées (p. 498) et Hôtes dédiés (p. 465).
   - ID de disque RAM : disque RAM de l'instance. Si vous avez sélectionné un noyau, vous devrez peut-être spécifier un disque RAM spécifique avec les pilotes qui l'accompagnent. Valide uniquement pour les AMIs paravirtuelles, (PV).
   - ID du noyau : noyau de l'instance. Valide uniquement pour les AMIs paravirtuelles, (PV).
   - Données utilisateur : vous pouvez spécifier les données utilisateur pour configurer une instance lors du lancement ou pour exécuter un script de configuration. Pour plus d'informations, consultez Exécution de commandes sur votre instance Linux lors du lancement (p. 694).
10. Choisissez Créer un modèle de lancement.

AWS CLI

Pour créer un modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande create-launch-template. L'exemple suivant crée un modèle de lancement qui spécifie ce qui suit :
  - Une balise pour le modèle de lancement (`purpose=production`)
  - Type d'instance (`r4.4xlarge`) et AMI (`ami-8c1be5f6`) à lancer
  - Nombre de cœurs (`4`) et threads par cœur (`2`) pour un total de 8 vCPU (4 cœurs x 2 threads)
  - Sous-réseau dans lequel lancer l'instance (`subnet-7b16de0c`)

  Le modèle attribue une adresse IP publique et une adresse IPv6 à l'instance et crée une balise pour l'instance (`Name=webserver`).

```
aws ec2 create-launch-template \
    --launch-template-name TemplateForWebServer \
    --version-description WebVersion1 \
    --tag-specifications 'ResourceType=launch-
template,Tags=[{Key=purpose,Value=production}]' \
    --launch-template-data file://template-data.json
```

  Voici un exemple de fichier `template-data.json`.

```
{
    "NetworkInterfaces": [{
        "AssociatePublicIpAddress": true,
        "DeviceIndex": 0,
        "Ipv6AddressCount": 1,
        "SubnetId": "subnet-7b16de0c"
    }],
    "ImageId": "ami-8c1be5f6",
    "InstanceType": "r4.4xlarge",
    "TagSpecifications": [{
        "ResourceType": "instance",
        "Tags": [{
            "Key":"Name",
            "Value":"webserver"
        }]
    }],
    "CpuOptions": {
        "CoreCount":4,
        "ThreadsPerCore":2
    }
}
```

  Voici un exemple de sortie.

```
{
    "LaunchTemplate": {
        "LatestVersionNumber": 1,
        "LaunchTemplateId": "lt-01238c059e3466abc",
        "LaunchTemplateName": "TemplateForWebServer",
        "DefaultVersionNumber": 1,
        "CreatedBy": "arn:aws:iam::123456789012:root",
        "CreateTime": "2017-11-27T09:13:24.000Z"
    }
}
```

Création d'un modèle de lancement à partir d'un modèle de lancement existant

Nouvelle console

Pour créer un modèle de lancement à partir d'un modèle de lancement existant à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement, puis Créer un modèle de lancement.
3. Pour Nom du modèle de lancement, entrez un nom descriptif pour le modèle.
4. Pour Description de la version du modèle, fournissez une brève description pour la version du modèle de lancement.
5. Pour baliser le modèle de lancement à la création, développez Template tags (Balises du modèle), Ajouter la balise, puis entrez une paire de clé et de valeur de balise.
6. Développez Modèle source, et, pour Nom du modèle de lancement, choisissez un modèle de lancement sur lequel baser le nouveau modèle.
7. Pour Version du modèle source, choisissez la version du modèle de lancement sur laquelle baser le nouveau modèle de lancement.
8. Ajustez les paramètres de lancement si nécessaire, puis choisissez Créer un modèle de lancement.

Ancienne console

Pour créer un modèle de lancement à partir d'un modèle de lancement existant à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Choisissez Créer un modèle de lancement. Indiquez un nom, une description et des balises pour le modèle de lancement.
4. Pour Modèle source, choisissez un modèle de lancement sur lequel baser le nouveau.
5. Pour Version du modèle source, choisissez la version du modèle de lancement sur laquelle baser le nouveau modèle de lancement.
6. Ajustez les paramètres de lancement si nécessaire, puis choisissez Créer un modèle de lancement.

Création d'un modèle de lancement à partir d'une instance

Nouvelle console

Pour créer un modèle de lancement à partir d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Créer un modèle à partir d'une instance.
4. Indiquez un nom, une description eet des balises, puis ajustez les paramètres de lancement si nécessaire.

   Note

   Lorsque vous créez un modèle de lancement à partir d'une instance, les ID et adresses IP de l'interface réseau de l'instance ne sont pas inclus dans le modèle.

5. Choisissez Créer un modèle de lancement.

Ancienne console

Pour créer un modèle de lancement à partir d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Create Template From Instance (Créer un modèle à partir d'une instance).
4. Indiquez un nom, une description eet des balises, puis ajustez les paramètres de lancement si nécessaire.

   Note

   Lorsque vous créez un modèle de lancement à partir d'une instance, les ID et adresses IP de l'interface réseau de l'instance ne sont pas inclus dans le modèle.
5. Choisissez Create Template From Instance (Créer un modèle à partir d'une instance).

AWS CLI

Vous pouvez utiliser le AWS CLI pour créer un modèle de lancement à partir d'une instance existante en obtenant d'abord les données du modèle de lancement à partir d'une instance, puis en créant un modèle de lancement à l'aide des données du modèle de lancement.

Pour obtenir des données de modèle de lancement à partir d'une instance à l'aide de l'AWS CLI

- Utilisez la commande get-launch-template-data et spécifiez l'ID d'instance. Vous pouvez utiliser la sortie comme base pour créer un modèle de lancement ou une version de modèle de lancement. Par défaut, la sortie contient un objet `LaunchTemplateData` de niveau supérieur qui ne peut pas être spécifié dans les données de modèle de lancement. Excluez cet objet à l'aide de l'option `--query`.

```
aws ec2 get-launch-template-data \
    --instance-id i-0123d646e8048babc \
    --query "LaunchTemplateData"
```

Voici un exemple de sortie.

```
    {
        "Monitoring": {},
        "ImageId": "ami-8c1be5f6",
        "BlockDeviceMappings": [
            {
                "DeviceName": "/dev/xvda",
                "Ebs": {
                    "DeleteOnTermination": true
                }
            }
        ],
        "EbsOptimized": false,
        "Placement": {
            "Tenancy": "default",
            "GroupName": "",
            "AvailabilityZone": "us-east-1a"
        },
```

```
        "InstanceType": "t2.micro",
        "NetworkInterfaces": [
            {
                "Description": "",
                "NetworkInterfaceId": "eni-35306abc",
                "PrivateIpAddresses": [
                    {
                        "Primary": true,
                        "PrivateIpAddress": "10.0.0.72"
                    }
                ],
                "SubnetId": "subnet-7b16de0c",
                "Groups": [
                    "sg-7c227019"
                ],
                "Ipv6Addresses": [
                    {
                        "Ipv6Address": "2001:db8:1234:1a00::123"
                    }
                ],
                "PrivateIpAddress": "10.0.0.72"
            }
        ]
    }
```

Par exemple, vous pouvez écrire directement la sortie dans un fichier :

```
aws ec2 get-launch-template-data \
    --instance-id i-0123d646e8048babc \
    --query "LaunchTemplateData" >> instance-data.json
```

Pour créer un modèle de lancement à l'aide des données du modèle de lancement

Utilisez la commande create-launch-template pour créer un modèle de lancement à l'aide de la sortie de la procédure précédente. Pour de plus amples informations sur la création d'un modèle de lancement à l'aide de AWS CLI, veuillez consulter Création d'un nouveau modèle de lancement à l'aide des paramètres que vous définissez (p. 539).

Gestion des versions du modèle de lancement

Vous pouvez créer des versions d'un modèle de lancement spécifique, définir la version par défaut, décrire une version du modèle de lancement et supprimer les versions dont vous n'avez plus besoin.

Tâches

- Création d'une version d'un modèle de lancement (p. 548)
- Définition de la version par défaut du modèle de lancement (p. 550)
- Description d'une version du modèle de lancement (p. 550)
- Suppression d'une version du modèle de lancement (p. 551)

Création d'une version d'un modèle de lancement

Lorsque vous créez une version d'un modèle de lancement, vous pouvez spécifier de nouveaux paramètres de lancement ou utiliser une version existante comme base de la nouvelle version. Pour plus d'informations sur les paramètres de lancement, consultez Création d'un modèle de lancement (p. 539).

Nouvelle console

Pour créer une version d'un modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez un modèle de lancement, puis choisissez Actions, Modify template (Create new version) (Modifier le modèle (Créer une nouvelle version)).
4. Pour Description de la version du modèle, saisissez une description pour la version du modèle de lancement.
5. (Facultatif) Développez Modèle source et sélectionnez une version du modèle de lancement à utiliser comme base pour la nouvelle version du modèle. La nouvelle version de modèle de lancement hérite des paramètres de lancement de cette version.
6. Modifiez les paramètres de lancement si nécessaire, puis choisissez Créer un modèle de lancement.

Ancienne console

Pour créer une version d'un modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Choisissez Créer un modèle de lancement.
4. Pour Que souhaitez-vous faire ?, choisissez Créer une nouvelle version de modèle.
5. Pour Nom du modèle de lancement, sélectionnez le nom du modèle de lancement existant dans la liste.
6. Pour Description de la version du modèle, saisissez une description pour la version du modèle de lancement.
7. (Facultatif) Sélectionnez une version du modèle de lancement ou une version d'un autre modèle de lancement à utiliser comme base pour la nouvelle version de modèle de lancement. La nouvelle version de modèle de lancement hérite des paramètres de lancement de cette version.
8. Modifiez les paramètres de lancement si nécessaire, puis choisissez Créer un modèle de lancement.

AWS CLI

Pour créer une version d'un modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande create-launch-template-version. Vous pouvez spécifier une version source sur laquelle baser la nouvelle version. La nouvelle version hérite des paramètres de lancement de cette version et vous pouvez les remplacer en utilisant `--launch-template-data`. L'exemple suivant illustre la création d'une nouvelle version basée sur la version 1 du modèle de lancement et la spécification d'un autre ID d'AMI.

```
aws ec2 create-launch-template-version \
    --launch-template-id lt-0abcd290751193123 \
    --version-description WebVersion2 \
    --source-version 1 \
    --launch-template-data "ImageId=ami-c998b6b2"
```

Définition de la version par défaut du modèle de lancement

Vous pouvez définir la version par défaut du modèle de lancement. Si vous lancez une instance à partir d'un modèle de lancement sans spécifier de version, le lancement est effectué à l'aide des paramètres de la version par défaut.

Nouvelle console

Pour définir la version par défaut du modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Définir la version par défaut.
4. Pour Version du modèle, sélectionnez le numéro de la version à définir par défaut et choisissez Définir comme version par défaut.

Ancienne console

Pour définir la version par défaut du modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Définir la version par défaut.
4. Pour Version par défaut, sélectionnez le numéro de version et choisissez Définir comme version par défaut.

AWS CLI

Pour définir la version par défaut du modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande modify-launch-template et spécifiez la version que vous souhaitez définir comme version par défaut.

```
aws ec2 modify-launch-template \
    --launch-template-id lt-0abcd290751193123 \
    --default-version 2
```

Description d'une version du modèle de lancement

À l'aide de la console, vous pouvez afficher toutes les versions du modèle de lancement sélectionné ou obtenir une liste des modèles de lancement dont la version la plus récente ou par défaut correspond à un numéro de version spécifique. À l'aide de l'AWS CLI, vous pouvez décrire toutes les versions, des versions individuelles ou une plage de versions d'un modèle de lancement spécifié. Vous pouvez également décrire toutes les dernières versions ou toutes les versions par défaut de tous les modèles de lancement de votre compte.

Nouvelle console

Pour décrire une version d'un modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Vous pouvez afficher une version d'un modèle de lancement spécifique ou obtenir une liste des modèles de lancement dont la version la plus récente ou par défaut correspond à un numéro de version spécifique.

- Pour afficher une version d'un modèle de lancement : sélectionnez le modèle de lancement. Sous l'onglet Versions dans Version, sélectionnez une version pour afficher ses détails.
- Pour obtenir une liste de tous les modèles de lancement dont la dernière version correspond à un numéro de version spécifique : dans la barre de recherche, choisissez Dernière version, puis sélectionnez un numéro de version.
- Pour obtenir la liste de tous les modèles de lancement dont la version par défaut correspond à un numéro de version spécifique : dans la barre de recherche, choisissez Version par défaut, puis sélectionnez un numéro de version.

AWS CLI

Pour supprimer une version du modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande describe-launch-template-versions et spécifiez les numéros de version. Dans l'exemple suivant, les versions 1 et 3 sont spécifiées.

```
aws ec2 describe-launch-template-versions \
    --launch-template-id lt-0abcd290751193123 \
    --versions 1 3
```

Pour décrire toutes les versions de modèles de lancement les plus récentes et par défaut de votre compte à l'aide de AWS CLI

- Utilisez la commande describe-launch-template-versions et spécifiez `$Latest`, `$Default` ou les deux. Vous devez omettre l'ID et le nom du modèle de lancement dans l'appel. Vous ne pouvez pas spécifier de numéros de version.

```
aws ec2 describe-launch-template-versions \
    --versions "$Latest,$Default"
```

Suppression d'une version du modèle de lancement

Si vous n'avez plus besoin d'une version du modèle de lancement, vous pouvez la supprimer. Vous ne pouvez pas remplacer le numéro d'une version après l'avoir supprimée. Vous ne pouvez pas supprimer la version par défaut du modèle de lancement et devez attribuer une autre version comme version par défaut.

Nouvelle console

Pour supprimer une version d'un modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Supprimer la version du modèle.
4. Sélectionnez la version à supprimer et choisissez Supprimer.

Ancienne console

Pour supprimer une version d'un modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Supprimer la version du modèle.

4. Sélectionnez la version à supprimer, puis Supprimer la version du modèle de lancement.

AWS CLI

Pour supprimer une version du modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande delete-launch-template-versions et spécifiez les numéros de version à supprimer.

```
aws ec2 delete-launch-template-versions \
    --launch-template-id lt-0abcd290751193123 \
    --versions 1
```

Lancement d'une instance à partir d'un modèle de lancement

Vous pouvez utiliser les paramètres contenus dans un modèle de lancement pour lancer une instance. Avant de lancer l'instance, vous pouvez remplacer ou ajouter des paramètres de lancement.

Deux balises accompagnées des clés `aws:ec2launchtemplate:id` et `aws:ec2launchtemplate:version` sont attribuées automatiquement aux instances lancées à l'aide d'un modèle de lancement. Vous ne pouvez ni supprimer ni modifier ces balises.

Nouvelle console

Pour lancer une instance à partir d'un modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Lancement d'une instance à partir d'un modèle.
4. Pour Version du modèle source, sélectionnez la version du modèle de lancement à utiliser.
5. Pour Nombre d'instances, spécifiez le nombre d'instances à lancer.
6. (Facultatif) Vous pouvez remplacer ou ajouter des paramètres du modèle de lancement dans la section Détails de l'instance.
7. Choisissez Lancer une instance à partir d'un modèle.

Ancienne console

Pour lancer une instance à partir d'un modèle de lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Lancement d'une instance à partir d'un modèle.
4. Sélectionnez la version du modèle de lancement à utiliser.
5. (Facultatif) Vous pouvez remplacer ou ajouter des paramètres du modèle de lancement dans la section Détails de l'instance.
6. Choisissez Lancer une instance à partir d'un modèle.

AWS CLI

Pour lancer une instance à partir d'un modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande run-instances et spécifiez le paramètre `--launch-template`. Spécifiez éventuellement la version du modèle de lancement à utiliser. Si vous ne la spécifiez pas, c'est la version par défaut qui est utilisée.

```
aws ec2 run-instances \
    --launch-template LaunchTemplateId=lt-0abcd290751193123,Version=1
```

- Pour remplacer un paramètre du modèle de lancement, spécifiez-le dans la commande run-instances. Dans l'exemple suivant, le type d'instance spécifié dans le modèle de lancement (le cas échéant) est remplacé.

```
aws ec2 run-instances \
    --launch-template LaunchTemplateId=lt-0abcd290751193123 \
    --instance-type t2.small
```

- Si vous spécifiez un paramètre imbriqué faisant partie d'une structure complexe, l'instance est lancée à l'aide de la structure complexe spécifiée dans le modèle de lancement et des éventuels paramètres imbriqués supplémentaires définis.

  Dans l'exemple suivant, l'instance est lancée avec la balise `Owner=TeamA` et toute autre balise spécifiée dans le modèle de lancement. Si le modèle de lancement comporte une balise avec une clé `Owner`, la valeur est remplacée par `TeamA`.

```
aws ec2 run-instances \
    --launch-template LaunchTemplateId=lt-0abcd290751193123 \
    --tag-specifications "ResourceType=instance,Tags=[{Key=Owner,Value=TeamA}]"
```

  Dans l'exemple suivant, l'instance est lancée avec un volume pourvu du nom de périphérique `/dev/xvdb` et d'autres mappages de périphérique de stockage en mode bloc spécifiés dans le modèle de lancement. Si le modèle de lancement possède un volume existant défini pour `/dev/xvdb`, ses valeurs sont remplacées par celles qui sont spécifiées.

```
aws ec2 run-instances \
    --launch-template LaunchTemplateId=lt-0abcd290751193123 \
    --block-device-mappings "DeviceName=/dev/xvdb,Ebs={VolumeSize=20,VolumeType=gp2}"
```

Si l'instance ne peut pas être lancée ou que l'état passe immédiatement à `terminated` au lieu de `running`, consultez Résolution des problèmes de lancement d'instance (p. 1325).

Utilisation de modèles de lancement avec Amazon EC2 Auto Scaling

Vous pouvez créer un groupe Auto Scaling et spécifier un modèle de lancement à utiliser pour le groupe. Lorsque Amazon EC2 Auto Scaling lance des instances dans le groupe Auto Scaling, il utilise les paramètres de lancement définis dans le modèle de lancement associé. Pour plus d'informations, consultez Création d'un groupe Auto Scaling à l'aide d'un modèle de lancement dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Avant de pouvoir créer un groupe Auto Scaling à l'aide d'un modèle de lancement, vous devez créer un modèle de lancement qui comprend les paramètres requis pour lancer une instance dans un groupe Auto Scaling, comme l'ID de l'AMI. La nouvelle console fournit des conseils pour vous aider à créer un modèle que vous pouvez utiliser avec Auto Scaling.

Pour créer un modèle de lancement à utiliser avec Auto Scaling à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement, puis Créer un modèle de lancement.
3. Pour Nom du modèle de lancement, entrez un nom descriptif pour le modèle.
4. Pour Description de la version du modèle, fournissez une brève description pour la version du modèle de lancement.
5. Sous Auto Scaling guidance (Conseils Auto Scaling), cochez la case pour obtenir des conseils d'Amazon EC2 pour créer un modèle à utiliser avec Auto Scaling.
6. Modifiez les paramètres de lancement selon vos besoins. Étant donné que vous avez sélectionné des conseils Auto Scaling, certains champs sont obligatoires et certains ne sont pas disponibles. Pour connaître les considérations à prendre en compte lors de la création d'un modèle de lancement et pour obtenir des informations sur la configuration des paramètres de lancement pour Auto Scaling, veuillez consulter Création d'un modèle de lancement pour un groupe Auto Scaling dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.
7. Choisissez Créer un modèle de lancement.
8. (Facultatif) Pour créer un groupe Auto Scaling à l'aide de ce modèle de lancement, dans la page Étapes suivantes choisissez Créer un groupe Auto Scaling.

Pour créer ou mettre à jour un groupe Amazon EC2 Auto Scaling avec un modèle de lancement à l'aide de l'AWS CLI

- Utilisez les commandes create-auto-scaling-group ou update-auto-scaling-group et spécifiez le paramètre `--launch-template`.

Utilisation de modèles de lancement avec Flotte EC2

Vous pouvez créer une demande Flotte EC2 et spécifier un modèle de lancement dans la configuration d'instance. Si Amazon EC2 satisfait à la demande Flotte EC2, il utilise les paramètres de lancement définis dans le modèle de lancement associé. Vous pouvez remplacer certains des paramètres spécifiés dans le modèle de lancement.

Pour plus d'informations, consultez Création d'un Flotte EC2 (p. 579).

Pour créer une Flotte EC2 avec un modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande create-fleet. Utilisez le paramètre `--launch-template-configs` pour spécifier le modèle de lancement et tous les remplacements de celui-ci.

Utilisation de modèles de lancement avec Parc d'instances Spot

Vous pouvez créer une demande Parc d'instances Spot et spécifier un modèle de lancement dans la configuration d'instance. Si Amazon EC2 satisfait à la demande Parc d'instances Spot, il utilise les paramètres de lancement définis dans le modèle de lancement associé. Vous pouvez remplacer certains des paramètres spécifiés dans le modèle de lancement.

Pour plus d'informations, consultez Parc d'instances Spot requêtes (p. 409).

Pour créer une demande Parc d'instances Spot avec un modèle de lancement à l'aide de l'AWS CLI

- Utilisez la commande request-spot-fleet. Utilisez le paramètre `LaunchTemplateConfigs` pour spécifier le modèle de lancement et tous les remplacements de celui-ci.

Suppression d'un modèle de lancement

Si vous n'avez plus besoin d'un modèle de lancement, vous pouvez le supprimer. La suppression d'un modèle de lancement entraîne celle de toutes ses versions.

Nouvelle console

Pour supprimer un modèle de lancement (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Supprimer le modèle.
4. Entrez **`Delete`** pour confirmer la suppression, puis choisissez Supprimer.

Ancienne console

Pour supprimer un modèle de lancement (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Modèles de lancement.
3. Sélectionnez le modèle de lancement et choisissez Actions, Supprimer le modèle.
4. Choisissez Supprimer le modèle de lancement.

AWS CLI

Pour supprimer un modèle de lancement (AWS CLI)

- Utilisez la commande delete-launch-template (AWS CLI) et spécifiez le modèle de lancement.

```
aws ec2 delete-launch-template --launch-template-id lt-01238c059e3466abc
```

Lancement d'une instance à l'aide des paramètres d'une instance existante

La console Amazon EC2 fournit l'option d'assistant En lancer plus comme ceci qui vous permet d'utiliser une instance actuelle comme base afin de lancer d'autres instances. Cette option remplit automatiquement l'assistant de lancement Amazon EC2 avec des détails de configuration issus de l'instance sélectionnée.

> **Note**
>
> L'option d'assistant En lancer plus comme ceci ne clone pas l'instance sélectionnée, mais elle duplique certains détails de configuration. Pour créer une copie de l'instance, commencez par créer une AMI sur la base de cette instance, puis lancez des instances supplémentaires à partir de l'AMI.
> Vous pouvez également créer un modèle de lancement (p. 537) pour stocker les paramètres de lancement de vos instances.

Les détails de configuration suivants sont copiés de l'instance sélectionnée vers l'assistant de lancement :

- ID d'AMI
- Type d'instance
- zone de disponibilité ou le VPC et le sous-réseau où se trouve l'instance sélectionnée

- Adresse IPv4 publique. Si l'instance sélectionnée a une adresse IPv4 publique, la nouvelle instance en reçoit une aussi, quel que soit le paramètre d'adresse IPv4 public par défaut de l'instance sélectionnée. Pour plus d'informations sur les adresses IPv4 publiques, consultez Adresses IPv4 publiques et noms d'hôte DNS externes (p. 809).
- Groupe de placement, le cas échéant
- Rôle IAM associé à l'instance, le cas échéant
- Paramètre du comportement lors de la mise hors tension (arrêt ou mise hors service)
- Paramètre de protection de mise hors service de l'instance (vrai ou faux)
- Surveillance CloudWatch (activée ou désactivée)
- Paramètre d'optimisation Amazon EBS (vrai ou faux)
- Paramètre de location, en cas de lancement sur un VPC (partagé ou dédié)
- ID du noyau et ID du disque RAM, le cas échéant
- Données utilisateur, le cas échéant
- Balises associées à l'instance, le cas échéant
- Groupes de sécurité associés à l'instance

Les détails de configuration suivants ne sont pas copiés à partir de l'instance sélectionnée. Au lieu de cela, l'assistant applique leurs paramètres ou leur comportement par défaut :

- Nombre d'interfaces réseau : par défaut, il y a une interface réseau, qui est l'interface réseau principale (eth0).
- Stockage : la configuration de stockage par défaut est déterminée par l'AMI et le type d'instance.

Nouvelle console

Pour utiliser l'instance actuelle comme modèle

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance que vous souhaitez utiliser, puis choisissez Actions, Images and templates (Images et modèles), puis En lancer plus comme ceci.
4. L'assistant de lancement s'ouvre sur la page Examiner le lancement de l'instance. Vous pouvez apporter les modifications nécessaires en choisissant le lien Modifier approprié.

   Une fois que vous êtes prêt, choisissez Lancer afin de sélectionner une paire de clés et de lancer votre instance.
5. Si l'instance ne peut pas être lancée ou que l'état passe immédiatement à `terminated` au lieu de `running`, consultez Résolution des problèmes de lancement d'instance (p. 1325).

Ancienne console

Pour utiliser l'instance actuelle comme modèle

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance que vous souhaitez utiliser, puis choisissez Actions, En lancer plus comme ceci.
4. L'assistant de lancement s'ouvre sur la page Examiner le lancement de l'instance. Vous pouvez apporter les modifications nécessaires en choisissant le lien Modifier approprié.

Une fois que vous êtes prêt, choisissez Lancer afin de sélectionner une paire de clés et de lancer votre instance.

5. Si l'instance ne peut pas être lancée ou que l'état passe immédiatement à `terminated` au lieu de `running`, consultez Résolution des problèmes de lancement d'instance (p. 1325).

Lancement d'une instance AWS Marketplace

Vous pouvez vous abonner à un produit AWS Marketplace et lancer une instance depuis l'AMI du produit à l'aide de l'Assistant de lancement Amazon EC2. Pour plus d'informations sur les AMI payantes, consultez AMI payantes (p. 128). Pour annuler votre abonnement après le lancement, vous devez d'abord mettre fin à toutes les instances qui s'exécutent à partir de l'abonnement. Pour plus d'informations, consultez Gérer vos abonnements AWS Marketplace (p. 131).

Pour lancer une instance à partir de AWS Marketplace à l'aide de l'Assistant de lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le tableau de bord Amazon EC2, choisissez Lancer une instance.
3. Sur la page Choose an Amazon Machine Image (AMI), sélectionnez la catégorie AWS Marketplace sur la gauche. Recherchez une AMI appropriée en parcourant les catégories ou à l'aide de la fonctionnalité de recherche. Sélectionnez Select pour choisir votre produit.
4. Une présentation du produit sélectionné s'affiche dans une boîte de dialogue. Vous pouvez afficher les informations de tarification, ainsi que toute autre information communiquée par le fournisseur. Lorsque vous êtes prêt, sélectionnez Continue.

   Note

   L'utilisation du produit ne vous est plus facturée jusqu'à ce que vous ayez lancé une instance avec l'AMI. Notez la tarification de chaque type d'instance pris en charge, car vous allez être invité à sélectionner un type d'instance sur la page suivante de l'Assistant. Des taxes supplémentaires peuvent également être appliquées au produit.
5. Sur la page Choisir un type d'instance, sélectionnez la configuration matérielle et la taille de l'instance à lancer. Lorsque vous avez terminé, sélectionnez Next: Configure Instance Details.
6. Sur les pages suivantes de l'Assistant, vous pouvez configurer votre instance et ajouter du stockage, ainsi que des balises. Pour plus d'informations sur les différentes options que vous pouvez configurer, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530). Choisissez Next jusqu'à la page Configure Security Group.

   L'Assistant crée un groupe de sécurité conforme aux spécifications du fournisseur pour le produit. Le groupe de sécurité peut inclure des règles autorisant l'accès à toutes les adresses IPv4 (`0.0.0.0/0`) sur SSH (port 22) sur Linux ou sur RDP (port 3389) sur Windows. Il est recommandé d'ajuster ces règles pour n'autoriser qu'une adresse ou plage d'adresses spécifiques à accéder à votre instance sur ces ports.

   Lorsque vous êtes prêt, sélectionnez Review and Launch.
7. Sur la page Review Instance Launch, vérifiez les détails de l'AMI à partir de laquelle vous vous apprêtez à lancer l'instance, ainsi que les autres détails de configuration que vous avez définis dans l'Assistant. Lorsque vous êtes prêt, sélectionnez Launch pour choisir ou créer une paire de clés, et démarrez votre instance.
8. Selon le produit auquel vous êtes abonné, le lancement de l'instance peut prendre quelques minutes, voire plus. Vous devez vous abonner au produit avant de pouvoir lancer une instance. En cas de problème avec les informations de votre carte bancaire, vous serez invité à mettre à jour les coordonnées de votre compte. Lorsque la page de confirmation de lancement s'affiche, sélectionnez View Instances pour accéder à la page Instances.

> **Note**
>
> Vous êtes facturé pour le prix de l'abonnement aussi longtemps que votre instance s'exécute, même si elle est inactive. Si votre instance est arrêtée, il se peut que vous continuiez à être facturé pour le stockage.

9. Lorsque votre instance est à l'état `running`, vous pouvez vous y connecter. Pour ce faire, sélectionnez votre instance dans la liste et choisissez Connect. Suivez les instructions de la boîte de dialogue. Pour plus d'informations sur la connexion à votre instance, consultez Connectez-vous à votre instance Linux (p. 598).

   > **Important**
   >
   > Lisez attentivement les instructions d'utilisation du fournisseur, car il se peut que vous deviez choisir un nom utilisateur spécifique pour vous connecter à l'instance. Pour plus d'informations sur l'accès aux détails de votre abonnement, consultez Gérer vos abonnements AWS Marketplace (p. 131).

10. Si l'instance ne peut pas être lancée ou que l'état passe immédiatement à `terminated` au lieu de `running`, consultez Résolution des problèmes de lancement d'instance (p. 1325).

Lancement d'une instance AMI AWS Marketplace à l'aide de l'API et de l'interface ligne de commande

Pour lancer les instances depuis les produits AWS Marketplace à l'aide de l'API ou des outils de ligne de commande, vérifiez d'abord que vous êtes abonné au produit. Vous pouvez alors lancer une instance avec l'ID d'AMI du produit en utilisant les méthodes suivantes :

Method	Documentation
AWS CLI	Utilisez la commande run-instances ou consultez la rubrique suivante pour plus d'informations : Lancement d'une instance.
Outils AWS pour Windows PowerShell	Utilisez la commande New-EC2Instance ou consultez la rubrique suivante pour plus d'informations : Lancer une instance Amazon EC2 à l'aide de Windows PowerShell
API de requête	Utilisez la demande RunInstances request.

Lancement d'instances à l'aide d'un Flotte EC2

Un Flotte EC2 contient les informations de configuration permettant de lancer un parc – ou un groupe – d'instances. Dans un seul appel d'API, un parc d'instances peut lancer plusieurs types d'instance dans plusieurs zones de disponibilité, en utilisant conjointement les options d'achat instance à la demande, Instance réservée et instance Spot. Flotte EC2 vous permet d'effectuer les opérations suivantes :

- Définir des cibles de capacité à la demande et Spot distinctes et le montant maximum que vous être prêt à payer par heure.
- Spécifier les types d'instance qui fonctionnent le mieux pour vos applications.
- Spécifiez comment Amazon EC2 doit répartir votre capacité de flotte au sein de chaque option d'achat

Vous pouvez également définir le montant maximum que vous être prêt à payer par heure pour votre parc et Flotte EC2 lance les instances jusqu'à ce que le montant maximum soit atteint. Une fois le montant maximum que vous être prêt à payer atteint, le parc arrête de lancer des instances même si la capacité cible n'a pas été atteinte.

Le Flotte EC2 tente de lancer le nombre d'instances nécessaires conformément à la capacité cible que vous avez spécifiée dans votre demande. Si vous avez spécifié un prix maximum total par heure, il remplit la capacité jusqu'à ce que le montant maximum que vous êtes prêt à payer soit atteint. Le parc tente également de préserver le parc de capacité cible si Instances Spot est interrompu. Pour de plus amples informations, veuillez consulter Fonctionnement des Instances Spot (p. 379).

Vous pouvez spécifier un nombre illimité de types d'instance par Flotte EC2. Ces types d'instance peuvent être alloués à l'aide des options d'achat À la demande et Spot. Vous pouvez également spécifier plusieurs zones de disponibilité, spécifier différents prix spot maximum pour chaque instance et choisir des options Spot supplémentaires pour chaque parc d'instances. Amazon EC2 utilise les options spécifiées pour allouer de la capacité lors du lancement du parc d'instances.

Lorsque le parc d'instances est en cours d'exécution, si Amazon EC2 récupère une instance Spot en raison d'une augmentation de prix ou de la défaillance d'une instance, le Flotte EC2 peut essayer de remplacer les instances par l'un des types d'instance que vous spécifiez. Il est ainsi plus facile de récupérer de la capacité de lors d'un pic de tarification Spot. Vous pouvez développer une stratégie de ressources flexible et élastique pour chaque parc d'instances. Par exemple, au sein de parcs spécifiques, votre capacité principale peut être à la demande, complétée par une capacité spot moins onéreuse si possible.

Si vous avez des Instances réservées et que vous spécifiez des Instances à la demande dans votre parc d'instances, le Flotte EC2 utilise vos Instances réservées. Par exemple, si votre parc spécifie une instance à la demande en tant que `c4.large`, et que vous avez des Instances réservées pour `c4.large`, vous recevez la tarification pour l'Instance réservée.

L'utilisation d'un Flotte EC2 n'entraîne pas de frais supplémentaires. Vous payez uniquement les instances EC2 que le parc lance pour vous.

Sommaire

- Limites Flotte EC2 (p. 559)
- Limites Flotte EC2 (p. 560)
- Instances à capacité extensible (p. 560)
- Stratégies de configuration d'un Flotte EC2 (p. 560)
- Gestion d'un Flotte EC2 (p. 570)

Limites Flotte EC2

Les limites suivantes s'appliquent à un Flotte EC2 :

- Un Flotte EC2 est disponible uniquement via l'API ou l'AWS CLI.
- Une demande de Flotte EC2 ne peut pas couvrir plusieurs régions AWS. Vous devez créer un Flotte EC2 distinct pour chaque région.
- Une demande de Flotte EC2 ne peut pas couvrir différents sous-réseaux d'une même zone de disponibilité.

Limites Flotte EC2

Les limites Amazon EC2 habituelles s'appliquent aux instances lancées par une Flotte EC2, par exemple les limites de prix d'une demande Spot, les limites d'instance et les limites de volume. En outre, les limites suivantes s'appliquent :

- Nombre de Flottes EC2 actives par région AWS : 1 000 * †
- Nombre de groupes instance Spot (combinaison unique de type d'instance et de sous-réseau) : 300* ‡
- Taille des données utilisateur dans une spécification de lancement : 16 Ko †
- Capacité cible par Flotte EC2 : 10 000
- Capacité cible pour toutes les Flottes EC2 d'une région : 100 000 *
- Une demande de Flotte EC2 ne peut pas couvrir plusieurs régions.
- Une demande de Flotte EC2 ne peut pas couvrir différents sous-réseaux d'une même zone de disponibilité.

S'il vous faut une capacité cible supérieure aux limites par défaut, remplissez le formulaire Créer une demande du centre AWS Support pour demander l'augmentation de la limite. Pour Limit type (Type de limite), choisissez EC2 Fleet (Flotte EC2), choisissez une région, puis choisissez Target Fleet Capacity per Fleet (in units) (Capacité de flotte cible par flotte (en unités)), Target Fleet Capacity per Region (in units) (Capacité de flotte cible par région (en unités)) ou les deux.

* Ces limites s'appliquent à vos Flottes EC2 et à vos Parcs d'instances Spot.

† Ces limites sont finales. Vous ne pouvez pas demander une augmentation de ces limites.

‡ Cette limite ne s'applique qu'aux flottes de type `request` ou `maintain`. Cette limite ne s'applique pas aux flottes `instant`.

Instances à capacité extensible

Si vous lancez vos Instances Spot à l'aide d'un type d'instance à capacité extensible (p. 233), et si vous prévoyez d'utiliser vos Instances Spot à capacité extensible immédiatement et pour une courte durée, sans temps d'inactivité pour accumuler des crédits UC, nous vous recommandons de les lancer en mode standard (p. 246) pour éviter de payer des coûts plus élevés. Si vous lancez vos Instances Spot à capacité extensible en mode Illimité (p. 237) et que vous étendez immédiatement l'utilisation de l'UC, vous dépensez des crédits excédentaires pour cette extension d'utilisation. Si vous utilisez l'instance pour une courte durée, elle n'a pas le temps d'accumuler des crédits UC pour rembourser les crédits excédentaires, et ces derniers vous sont facturés lorsque vous résiliez l'instance.

Le mode Illimité convient aux Instances Spot à capacité extensible uniquement si l'instance s'exécute suffisamment longtemps pour accumuler des crédits UC pour l'extension d'utilisation. Sinon, payer des crédits excédentaires rend les Instances Spot à capacité extensible plus coûteuses que les autres instances. Pour de plus amples informations, veuillez consulter Quand utiliser le mode illimité/mode d'UC fixe ? (p. 238).

Les crédits de lancement visent à optimiser la productivité du lancement initial des instances T2 en leur fournissant suffisamment de ressources de calcul pour pouvoir configurer l'instance. Il est interdit de procéder à des lancements répétés d'instances T2 pour bénéficier de nouveaux crédits de lancement. Si vous avez besoin de performances soutenues de l'UC, vous pouvez obtenir des crédits (en restant inactif pendant un certain temps) : utilisez le mode Illimité (p. 237) pour les Instances Spot T2 ou un type d'instance avec UC dédiée.

Stratégies de configuration d'un Flotte EC2

Un Flotte EC2 est un groupe d'Instances à la demande et d'Instances Spot.

Le Flotte EC2 tente de lancer le nombre d'instances nécessaires pour répondre à la capacité cible que vous spécifiez dans votre demande. Le parc peut comprendre uniquement des Instances à la demande uniquement des Instances Spot, ou une combinaison d'Instances à la demande et d'Instances Spot. La demande des Instances Spot est satisfaite si la capacité disponible et le prix maximum par heure que vous avez spécifié pour la demande dépassent le prix spot actuel. Le parc tente également de préserver le parc de capacité cible si Instances Spot est interrompu.

Vous pouvez également définir le montant maximum que vous être prêt à payer par heure pour votre parc et Flotte EC2 lance les instances jusqu'à ce que le montant maximum soit atteint. Une fois le montant maximum que vous être prêt à payer atteint, le parc arrête de lancer des instances même si la capacité cible n'a pas été atteinte.

Un groupe d'instance Spot est un ensemble d'instances EC2 inutilisées ayant les mêmes type d'instance, système d'exploitation, zone de disponibilité et plateforme réseau. Lorsque vous créez un Flotte EC2, vous pouvez inclure plusieurs spécifications de lancement qui varient en terme de type d'instance, zone de disponibilité, sous-réseau ou prix maximum. Le parc d'instances sélectionne les groupes d'instance Spot servant à satisfaire la demande, selon les spécifications de lancement incluses dans votre demande et la configuration de cette demande. Les Instances Spot proviennent des groupes sélectionnés.

Un Flotte EC2 vous permet d'allouer un gros volume de capacité EC2 logique pour votre application sur la base du nombre de cœurs ou d'instances, ou de la quantité de mémoire. Par exemple, vous pouvez spécifier un Flotte EC2 pour lancer une capacité cible de 200 instances, dont 130 sont des Instances à la demande et le reste des Instances Spot.

Utilisez les stratégies de configuration appropriées pour créer un Flotte EC2 qui réponde à vos besoins.

Table des matières

- Planification d'un Flotte EC2 (p. 561)
- Types de demande Flotte EC2 (p. 562)
- Stratégies d'allocation pour Instances Spot (p. 562)
- Configuration du Flotte EC2 pour la sauvegarde à la demande (p. 564)
- Remplacements du prix maximum (p. 565)
- Contrôle des dépenses (p. 565)
- Pondération d'instance Flotte EC2 (p. 566)
- Didacticiel : Utilisation d'un Flotte EC2 avec pondération des instances (p. 567)
- Didacticiel : Utilisation d'un Flotte EC2 avec la capacité à la demande comme capacité principale (p. 569)

Planification d'un Flotte EC2

Lors de la planification de votre Flotte EC2, nous vous recommandons de procéder comme suit :

- Déterminez si vous souhaitez créer un Flotte EC2 qui envoie une demande unique synchrone ou asynchrone pour la capacité cible souhaitée ou qui conserve une capacité cible au fil du temps. Pour plus d'informations, consultez Types de demande Flotte EC2 (p. 562).
- Déterminez les types d'instance qui correspondent aux exigences de votre application.
- Si vous envisagez d'inclure des Instances Spot dans votre Flotte EC2, passez en revue les bonnes pratiques en matière d'instances Spot avant de créer le parc d'instances. Appuyez-vous sur ces bonnes pratiques lorsque vous planifiez votre parc d'instances afin de pouvoir mettre en service ces instances au prix le plus bas possible.
- Déterminez la capacité cible de votre Flotte EC2. Vous pouvez définir la capacité cible en instances ou en unités personnalisées. Pour plus d'informations, consultez Pondération d'instance Flotte EC2 (p. 566).

- Déterminez quelle portion de la capacité cible du Flotte EC2 doit correspondre à la capacité à la demande et à la capacité des instances spot. Vous pouvez spécifier 0 pour la capacité à la demande, pour la capacité des instances spot, ou pour les deux.
- Déterminez le prix par unité si vous avez recours à la pondération d'instance. Pour calculer le prix par unité, divisez le prix pour une heure d'instance par le nombre d'unités (ou pondération) que cette instance représente. Si vous n'utilisez pas la pondération d'instance, le prix par unité défini par défaut est le prix par heure d'instance.
- Déterminez le montant maximum par heure que vous êtes prêt à payer pour votre parc. Pour de plus amples informations, veuillez consulter Contrôle des dépenses (p. 565).
- Passez en revue les options possibles pour votre Flotte EC2. Pour plus d'informations, consultez le Référence du fichier de configuration JSON de Flotte EC2 (p. 576). Pour accéder à des exemples de configuration de Flotte EC2, consultez Exemples de configuration d'un Flotte EC2 (p. 587).

Types de demande Flotte EC2

Il existe trois types de demandes Flotte EC2 :

`instant`

Si vous configurez le type de demande comme `instant`, Flotte EC2 passe une demande unique synchrone de la capacité souhaitée. Dans la réponse de l'API, il renvoie les instances qui ont été lancées, ainsi que les erreurs liées aux instances qui n'ont pas pu être lancées.

`request`

Si vous configurez le type de demande comme `request`, Flotte EC2 passe une demande unique asynchrone de la capacité souhaitée. Ensuite, si la capacité est réduite en raison d'interruptions Spot, le parc d'instances n'essaie pas de réapprovisionner les Instances Spot et il ne soumet pas les demandes dans d'autres groupes instance Spot si la capacité n'est pas disponible.

`maintain`

(Par défaut) Si vous configurez le type de demande comme `maintain`, Flotte EC2 passe une demande asynchrone de la capacité souhaitée et maintient la capacité en réapprovisionnant automatiquement les Instances Spot interrompues.

Les trois types de demandes bénéficient d'une stratégie d'allocation. Pour plus d'informations, consultez Stratégies d'allocation pour Instances Spot (p. 562).

Stratégies d'allocation pour Instances Spot

La stratégie d'allocation de votre Flotte EC2 détermine la façon dont la demande d'Instances Spot est satisfaite à partir des groupes d'instance Spot possibles représentés dans ses spécifications de lancement. Les stratégies d'allocation que vous pouvez spécifier dans votre parc d'instances sont indiquées ci-après :

`lowest-price`

Les Instances Spot proviennent du groupe offrant le prix le plus bas. Il s'agit de la stratégie par défaut.

`diversified`

Les Instances Spot sont réparties entre tous les groupes.

`capacity-optimized`

Les Instances Spot proviennent du pool d'instances Spot avec une capacité optimale pour le nombre d'instances qui sont lancées.

`InstancePoolsToUseCount`

Les Instances Spot sont réparties entre le nombre de groupes d'instances Spot que vous spécifiez. Ce paramètre n'est valide que s'il est utilisé conjointement avec `lowest-price`.

Maintien de la capacité cible

Une fois les Instances Spot résiliées en raison d'un changement de prix spot ou de la modification de la capacité disponible d'un groupe d'instance Spot, un Flotte EC2 de type `maintain` lance des Instances Spot de remplacement. Si la stratégie d'allocation `lowest-price` est sélectionnée, le parc d'instances lance les instances de remplacement dans le pool où le prix Spot est actuellement le plus faible. Si la stratégie d'allocation est `lowest-price` conjointement à `InstancePoolsToUseCount`, le parc sélectionne les groupes d'instances Spot ayant le prix le plus bas et lance les Instances Spot sur le nombre de groupes d'instances Spot que vous spécifiez. Si la stratégie d'allocation `capacity-optimized` est sélectionnée, le parc d'instances lance les instances de remplacement dans le pool avec le plus de capacité de instance Spot disponible. Si la stratégie d'allocation est `diversified`, le parc d'instances répartit les Instances Spot de remplacement entre les groupes restants.

Configuration du Flotte EC2 pour l'optimisation des coûts

Pour optimiser les coûts relatifs à votre utilisation des Instances Spot, spécifiez la stratégie d'allocation `lowest-price` pour que le Flotte EC2 déploie automatiquement la combinaison la plus économique de types d'instance et de zones de disponibilité en fonction du prix Spot actuel.

Pour la capacité cible des instance à la demande, le Flotte EC2 sélectionne toujours le type d'instance le moins cher en fonction du prix à la demande public, tout en continuant à suivre la stratégie d'allocation (`lowest-price` ou `capacity-optimized`) pour les `diversified` pour les Instances Spot.

Configuration du Flotte EC2 pour l'optimisation des coûts et la diversification

Pour créer un parc d'Instances Spot à la fois peu onéreux et diversifié, utilisez la stratégie d'allocation `lowest-price` conjointement avec `InstancePoolsToUseCount`. Flotte EC2 déploie automatiquement la combinaison la plus économique de types d'instance et de zones de disponibilité en fonction du prix spot actuel sur le nombre de groupes d'instances Spot que vous spécifiez. Il est possible d'utiliser cette combinaison pour éviter les Instances Spot les plus onéreuses.

Configuration du Flotte EC2 pour l'optimisation de la capacité

Avec les Instances Spot, la tarification change lentement au fil du temps en fonction des tendances à long terme en matière d'offre et de demande, mais la capacité fluctue en temps réel. La stratégie `capacity-optimized` lance automatiquement des Instances Spot dans les pools les plus disponibles en examinant les données de capacité en temps réel et en prédisant les instances les plus disponibles. Cela convient parfaitement aux charges de travail telles que le Big Data et l'analyse, le rendu d'images et de médias, l'apprentissage automatique et le calcul haute performance qui peuvent avoir un coût d'interruption plus élevé, associé au redémarrage des tâches et aux points de contrôle. En offrant la possibilité de moins d'interruptions, la stratégie `capacity-optimized` peut réduire le coût global de votre charge de travail.

Choix de la stratégie d'allocation appropriée

Vous pouvez optimiser votre parc d'instances en fonction des cas d'utilisation.

Si votre parc est de petite taille ou qu'il s'exécute pendant un temps limité, il y a peu de risques que vos Instances Spot soient interrompues, même si toutes les instances se trouvent dans un même groupe d'instance Spot. C'est pourquoi la stratégie `lowest-price` répondra certainement à vos besoins pour le prix le plus faible.

Si votre parc est important ou qu'il s'exécute pendant une longue durée, vous pouvez améliorer sa disponibilité en répartissant les Instances Spot entre plusieurs groupes. Par exemple, si votre Flotte EC2 spécifie 10 groupes et une capacité cible de 100 instances, le parc lance 10 Instances Spot dans chaque groupe. Si le prix Spot d'un pool dépasse le prix maximum de ce pool, seul 10 % de votre parc est touché. Avec cette stratégie, votre parc est également moins affecté par les augmentations du prix Spot dans un pool au fil du temps.

Avec la stratégie `diversified`, le Flotte EC2 ne lance pas d'Instances Spot dans des groupes dont le prix spot est supérieur ou égal au prix à la demande.

Pour créer un parc peu onéreux et diversifié, utilisez la stratégie `lowest-price` conjointement avec `InstancePoolsToUseCount`. Vous pouvez utiliser un nombre plus ou moins élevé de groupes d'instances Spot auxquels allouer vos Instances Spot. Par exemple, si vous exécutez un traitement par lots, nous vous recommandons de spécifier un petit nombre de groupes d'instances Spot (par exemple, `InstancePoolsToUseCount=2`) pour garantir que votre file d'attente aura toujours une capacité de calcul suffisante tout en optimisant les économies. Si vous exécutez un service web, nous vous recommandons de spécifier un nombre élevé de groupes d'instances Spot (par exemple, `InstancePoolsToUseCount=10`) pour minimiser l'impact si un groupe d'instance Spot devient temporairement indisponible.

Si votre parc exécute des charges de travail dont l'interruption entraîne des coûts plus élevés associés au redémarrage du travail et aux points de contrôle, utilisez la stratégie `capacity-optimized`. Cette stratégie offre la possibilité de moins d'interruptions, ce qui peut réduire le coût global de votre charge de travail.

Configuration du Flotte EC2 pour la sauvegarde à la demande

En cas de besoin d'une mise à l'échelle imprévisible et urgente, par exemple pour un site web d'actualité qui doit être dimensionné à la hausse lors d'un événement d'actualité ou de la sortie d'un jeu majeur, nous vous recommandons de spécifier des types d'instance alternatifs pour vos Instances à la demande, au cas où votre option préférée n'aurait pas une capacité disponible suffisante. Par exemple, vous pouvez préférer des Instances à la demande `c5.2xlarge`, mais si la capacité disponible est insuffisante, vous pouvez utiliser certaines instances `c4.2xlarge` lors des pics de charges. Dans ce cas, le Flotte EC2 tente d'assurer toute votre capacité cible en utilisant des instances `c5.2xlarge`, mais si la capacité est insuffisante, il lance automatiquement des instances `c4.2xlarge` pour assurer la capacité cible.

Hiérarchisation des types d'instance pour la capacité à la demande

Lorsque le Flotte EC2 essaie de traiter l'affectation de capacité à la demande, il lance par défaut le type d'instance dont le prix est le plus bas en premier. Si `AllocationStrategy` a pour valeur `prioritized`, le Flotte EC2 utilise la priorité pour déterminer quel type d'instance utiliser en premier afin de traiter l'affectation de capacité à la demande. La priorité est affectée au remplacement du modèle de lancement, et la priorité la plus élevée est lancée en premier.

Par exemple, vous avez configuré trois remplacements du modèle de lancement, ayant chacun un type d'instance différent : `c3.large`, `c4.large` et `c5.large`. Le prix à la demande d'une instance `c5.large` est inférieur à celui d'une instance `c4.large`. L'instance `c3.large` est la moins chère. Si vous n'utilisez pas la priorité pour déterminer l'ordre, le parc traite l'affectation de capacité à la demande en commençant par `c3.large`, puis `c5.large`. Étant donné que vous avez souvent des Instances réservées non utilisées pour `c4.large`, vous pouvez définir la priorité du remplacement du modèle de lancement afin que l'ordre soit `c4.large`, `c3.large`, puis `c5.large`.

Utilisation de Réservations de capacité pour Instances à la demande

Vous pouvez configurer une flotte pour qu'elle utilise d'abord Réservations de capacité à la demande lors du lancement d'Instances à la demande en définissant la stratégie d'utilisation des réservations de capacité sur `use-capacity-reservations-first`. Vous pouvez utiliser ce paramètre conjointement avec la stratégie d'allocation pour l'Instances à la demande (`lowest-price` ou `prioritized`).

Lorsque les réservations de capacité inutilisées sont employées pour satisfaire la capacité à la demande :

- La flotte utilise les réservations de capacité inutilisées pour satisfaire la capacité à la demande jusqu'à la capacité à la demande cible.
- Si plusieurs groupes d'instances ont des réservations de capacité inutilisées, la stratégie d'allocation à la demande (`lowest-price` ou `prioritized`) est appliquée.
- Si le nombre de réservations de capacité inutilisées est inférieur à la capacité à la demande cible, la capacité à la demande cible restante est lancée conformément à la stratégie d'allocation à la demande (`lowest-price` ou `prioritized`).

Vous ne pouvez uniquement utiliser Réservations de capacité à la demande non utilisé pour les flottes de type `instant`.

Pour obtenir des exemples de la façon de configurer une flotte pour qu'elle utilise Réservations de capacité pour satisfaire la capacité à la demande, consultez Exemples de configuration d'un Flotte EC2 (p. 587). Pour plus d'informations, consultez Réservations de capacité à la demande (p. 504) et la FAQ sur la réservation de capacité à la demande.

Remplacements du prix maximum

Chaque Flotte EC2 peut inclure un prix maximum global ou utiliser la valeur par défaut (prix à la demande). Le parc d'instances utilise ce prix comme prix maximum par défaut pour chacune de ses spécifications de lancement.

Si vous le souhaitez, vous pouvez également spécifier un prix maximum dans une ou plusieurs spécifications de lancement. Ce prix est propre à la spécification de lancement. Si une spécification de lancement comprend un prix spécifique, le Flotte EC2 utilise ce prix maximum à la place du prix maximum global. Toute autre spécification de lancement qui ne comprend pas de prix maximum spécifique continue à utiliser le prix maximum global.

Contrôle des dépenses

Flotte EC2 arrête le lancement des instances une fois l'un des paramètres suivants atteints : le `TotalTargetCapacity` ou le `MaxTotalPrice` (montant maximum que vous êtes prêt à payer). Pour contrôler le montant payé par heure pour votre parc, vous pouvez spécifier `MaxTotalPrice`. Une fois le prix total atteint, Flotte EC2 arrête de lancer des instances même si la capacité cible n'a pas été atteinte.

Les exemples suivants montrent deux manières de le faire. Dans le premier, Flotte EC2 arrête de lancer des instances une fois la capacité cible atteinte. Dans le deuxième, Flotte EC2 arrête le lancement des instances une fois le montant maximum que vous êtes prêt à payer atteint (`MaxTotalPrice`).

Exemple : Arrêt du lancement des instances lorsque la capacité cible est atteinte

Prenons l'exemple d'une demande pour `m4.large` Instances à la demande, avec :

- Prix à la demande : 0,10 USD par heure
- `OnDemandTargetCapacity` : 10
- `MaxTotalPrice` : 1,50 USD

Flotte EC2 lance 10 Instances à la demande car le total de 1 USD (10 instances x 0,10 USD) ne dépasse pas le `MaxTotalPrice` de 1,50 USD pour Instances à la demande.

Exemple : Arrêt du lancement des instances lorsque le prix total maximum est atteint

Prenons l'exemple d'une demande pour `m4.large` Instances à la demande, avec :

- Prix à la demande : 0,10 USD par heure
- `OnDemandTargetCapacity` : 10
- `MaxTotalPrice` : 0,80 USD

Si Flotte EC2 lance la capacité cible à la demande (10 Instances à la demande), le coût total par heure est de 1 USD. Ce montant est supérieur à celui (0,80 USD) spécifié pour `MaxTotalPrice` pour Instances à la demande. Afin d'éviter de dépenser plus que vous le souhaitez, Flotte EC2 lance uniquement 8 Instances à la demande (ce qui est inférieur à la capacité cible à la demande) car le lancement d'instances supplémentaires dépasserait `MaxTotalPrice` pour Instances à la demande.

Pondération d'instance Flotte EC2

Lorsque vous créez un Flotte EC2, vous pouvez définir les unités de capacité que chaque type d'instance apporterait aux performances de votre application. Vous pouvez ensuite ajuster votre prix maximum pour chaque spécification de lancement en utilisant la pondération d'instance.

Par défaut, le prix que vous spécifiez représente le prix par heure d'instance. Lorsque vous utilisez la fonction de pondération d'instance, le prix que vous spécifiez correspond au prix par heure d'unité. Vous pouvez calculer le prix de l'offre par heure d'unité en divisant le prix de l'offre pour un type d'instance par le nombre d'unités qu'il représente. Le Flotte EC2 calcule le nombre d'instances à lancer en divisant la capacité cible par la pondération de l'instance. Si le résultat n'est pas un entier, le parc d'instances l'arrondit à l'entier suivant afin que la taille de votre parc ne soit pas inférieure à sa capacité cible. Le parc d'instances peut sélectionner n'importe quel groupe indiqué dans votre spécification de lancement, même si la capacité des instances lancées dépasse la capacité cible demandée.

Le tableau suivant inclut des exemples de calcul afin de déterminer le prix par unité pour un Flotte EC2 ayant une capacité cible de 10.

Type d'instance	Pondération de l'instance	Capacité cible	Nombre d'instances lancées	Prix par heure d'instance	Prix par heure d'unité
`r3.xlarge`	2	10	5 (10 divisé par 2)	0,05 USD	0,025 USD (0,05 divisé par 2)
`r3.8xlarge`	8	10	2 (10 divisé par 8, résultat arrondi)	0,10 USD	0,0125 USD (0,10 divisé par 8)

Utilisez la pondération d'instance de Flotte EC2 comme suit, afin de mettre en service la capacité cible que vous voulez dans les groupes selon le prix par unité le plus bas au moment de l'exécution :

1. Définissez la capacité cible de votre Flotte EC2 en instances (valeur par défaut) ou dans les unités de votre choix, par exemple les UC virtuelles, la mémoire, le stockage ou le débit.
2. Définissez le prix par unité.
3. Pour chaque spécification de lancement, spécifiez la pondération, à savoir le nombre d'unités que représente ce type d'instance par rapport à la capacité cible.

Exemple de pondération d'instance

Prenons l'exemple d'une demande de Flotte EC2 avec la configuration suivante :

- Capacité cible de 24
- Spécification de lancement avec le type d'instance `r3.2xlarge` et une pondération de 6
- Spécification de lancement avec le type d'instance `c3.xlarge` et une pondération de 5

La pondération correspond au nombre d'unités du type d'instance par rapport à la capacité cible. Si la première spécification de lancement fournit le prix par unité le plus faible (prix pour `r3.2xlarge` par heure d'instance divisé par 6), le Flotte EC2 lance quatre de ces instances (24 divisé par 6).

Si la deuxième spécification de lancement fournit le prix par unité le plus bas (prix pour `c3.xlarge` par heure d'instance divisé par 5), le Flotte EC2 lance cinq de ces instances (24 divisé par 5, résultat arrondi).

Pondération d'instance et stratégie d'attribution

Prenons l'exemple d'une demande de Flotte EC2 avec la configuration suivante :

- Capacité cible de 30 Instances Spot
- Spécification de lancement avec le type d'instance `c3.2xlarge` et une pondération de 8
- Spécification de lancement avec le type d'instance `m3.xlarge` et une pondération de 8
- Spécification de lancement avec le type d'instance `r3.xlarge` et une pondération de 8

Le Flotte EC2 lancerait quatre instances (30 divisé par 8, résultat arrondi). Avec la stratégie `lowest-price`, les quatre instances sont issues du pool d'instances Spot qui fournit le prix par unité le plus bas. Avec la stratégie `diversified`, le parc d'instances lance une instance dans chacun des trois groupes, et la quatrième instance dans l'un des trois groupes fournit le prix par unité le plus bas.

Didacticiel : Utilisation d'un Flotte EC2 avec pondération des instances

Cette procédure utilise une société fictive nommée Example Corp pour illustrer le processus de demande d'un Flotte EC2 utilisant la pondération des instances.

Objectif

Example Corp est une entreprise pharmaceutique qui souhaite utiliser la puissance de calcul d'Amazon EC2 pour analyser les composants chimiques susceptibles d'être utilisés dans la lutte contre le cancer.

Planification

Example Corp commence par examiner les bonnes pratiques en matière d'instances Spot. Ensuite, Example Corp détermine les exigences suivantes pour son Flotte EC2.

Types d'instance

Example Corp a une application qui exige beaucoup de calculs et de mémoire. Pour un fonctionnement optimal, cette application a besoin d'au moins 60 Go de mémoire et de huit UC virtuelles (vCPU). L'entreprise souhaite optimiser ces ressources pour l'application au prix le plus bas possible. Example Corp décide que l'un des types d'instance EC2 suivants est capable de répondre à ses besoins :

Type d'instance	Mémoire (Gio)	vCPU
r3.2xlarge	61	8
r3.4xlarge	122	16
r3.8xlarge	244	32

Capacité cible en unités

Avec la pondération d'instance, la capacité cible peut être égale à un certain nombre d'instances (valeur par défaut) ou à une combinaison de critères, par exemple les noyaux (vCPU), la mémoire (GiO) et le stockage (Go). En considérant que la base de son application (60 Go de RAM et huit vCPU) compte pour une unité, Example Corp décide que 20 fois cette quantité suffirait pour répondre à ses besoins. L'entreprise définit donc la capacité cible de sa demande de Flotte EC2 sur 20.

Pondérations d'instance

Après avoir déterminé sa capacité cible, Example Corp calcule ses pondérations d'instance. Pour calculer la pondération de chaque type d'instance, l'entreprise détermine les unités de chaque type d'instance nécessaires pour atteindre la capacité cible de la façon suivante :

- r3.2xlarge (61 Go, 8 vCPU) = 1 unité de 20
- r3.4xlarge (122 Go, 16 vCPU) = 2 unités de 20
- r3.8xlarge (244 Go, 32 vCPU) = 4 unités de 20

Par conséquent, Example Corp assigne des pondérations d'instance de 1, 2 et 4 aux configurations de lancement respectives dans sa demande de Flotte EC2.

Prix par heure d'unité

Example Corp utilise le prix à la Demande par heure d'instance comme point de départ de son prix. Elle peut également utiliser les prix Spot récents ou une combinaison des deux. Pour calculer le prix par heure d'unité, elle divise le prix de départ basé sur l'heure d'instance par la pondération. Exemples :

Type d'instance	Prix à la Demande	Pondération de l'instance	Prix par heure d'unité
r3.2xLarge	0,7 USD	1	0,7 USD
r3.4xLarge	1,4 USD	2	0,7 USD
r3.8xLarge	$2.8	4	0,7 USD

Example Corp peut utiliser un prix global par heure d'unité s'élevant à 0,7 USD et rester concurrentielle pour les trois types d'instance. Elle peut également utiliser un prix global par heure d'unité s'élevant à 0,7 USD et un prix spécifique par heure d'unité de 0,9 USD dans la spécification de lancement du type d'instance `r3.8xlarge`.

Vérification des autorisations

Avant de créer un Flotte EC2, la société Example Corp vérifie qu'elle dispose d'un rôle IAM avec les autorisations requises. Pour plus d'informations, consultez Conditions préalables requises Flotte EC2 (p. 572).

Création d'un modèle de lancement

Ensuite, Example Corp crée un modèle de lancement. L'ID de modèle de lancement est utilisé à l'étape suivante. Pour plus d'informations, consultez Création d'un modèle de lancement (p. 539).

Création du Flotte EC2

Example Corp crée un fichier, `config.json`, avec la configuration suivante pour son Flotte EC2. Dans l'exemple suivant, remplacez les identificateurs de ressources par vos propres identificateurs de ressources.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-07b3bc7625cdab851",
                "Version": "1"
            },
            "Overrides": [
                {
```

```
                    "InstanceType": "r3.2xlarge",
                    "SubnetId": "subnet-482e4972",
                    "WeightedCapacity": 1
                },
                {
                    "InstanceType": "r3.4xlarge",
                    "SubnetId": "subnet-482e4972",
                    "WeightedCapacity": 2
                },
                {
                    "InstanceType": "r3.8xlarge",
                    "MaxPrice": "0.90",
                    "SubnetId": "subnet-482e4972",
                    "WeightedCapacity": 4
                }
            ]
        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 20,
        "DefaultTargetCapacityType": "spot"
    }
}
```

Example Corp crée le Flotte EC2 à l'aide de la commande create-fleet suivante.

```
aws ec2 create-fleet \
    --cli-input-json file://config.json
```

Pour de plus amples informations, veuillez consulter Création d'un Flotte EC2 (p. 579).

Exécution

La stratégie d'allocation détermine de quels groupes d'instance Spot sont issues vos Instances Spot.

Avec la stratégie `lowest-price` (qui est la stratégie par défaut), les Instances Spot sont issues du groupe ayant le prix par unité le plus bas au moment de l'exécution. Pour fournir 20 unités de capacité, le Flotte EC2 lance 20 instances `r3.2xlarge` (20 divisé par 1), 10 instances `r3.4xlarge` (20 divisé par 2) ou 5 instances `r3.8xlarge` (20 divisé par 4).

Si Example Corp avait utilisé la stratégie `diversified`, les Instances Spot auraient été issues des trois groupes. Le Flotte EC2 aurait lancé 6 instances `r3.2xlarge` (soit 6 unités), 3 instances `r3.4xlarge` (soit 6 unités) et 2 instances `r3.8xlarge` (soit 8 unités), pour un total de 20 unités.

Didacticiel : Utilisation d'un Flotte EC2 avec la capacité à la demande comme capacité principale

Ce didacticiel utilise une société fictive nommée ABC Online pour illustrer le processus de demande d'un Flotte EC2 avec la capacité à la demande comme capacité principale, et la capacité des instances spot si elle est disponible.

Objectif

ABC Online est une compagnie de livraison de restaurants qui veut être capable d'allouer une capacité d'Amazon EC2 entre les types d'instance EC2 et les options d'achat pour atteindre l'échelle, la performance et le coût qu'elle s'est fixés.

Planification

ABC Online nécessite une capacité fixe pour faire face aux périodes de pic, mais souhaiterait bénéficier d'une capacité augmentée pour un prix inférieur. ABC Online détermine les exigences suivantes pour son Flotte EC2 :

- Capacité des instance à la demande – ABC Online nécessite 15 Instances à la demande pour s'assurer de pouvoir prendre en charge le trafic dans les périodes de pic.
- Capacité des instance Spot – ABC Online souhaite améliorer la performance, mais pour un prix inférieur, en mettant en service 5 Instances Spot.

Vérification des autorisations

Avant de créer un Flotte EC2, la société ABC Online vérifie qu'elle dispose d'un rôle IAM avec les autorisations requises. Pour plus d'informations, consultez Conditions préalables requises Flotte EC2 (p. 572).

Création d'un modèle de lancement

Ensuite, ABC Online crée un modèle de lancement. L'ID de modèle de lancement est utilisé à l'étape suivante. Pour plus d'informations, consultez Création d'un modèle de lancement (p. 539).

Création du Flotte EC2

ABC Online crée un fichier, `config.json`, avec la configuration suivante pour son Flotte EC2. Dans l'exemple suivant, remplacez les identificateurs de ressources par vos propres identificateurs de ressources.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-07b3bc7625cdab851",
                "Version": "2"
            }

        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 20,
        "OnDemandTargetCapacity":15,
        "DefaultTargetCapacityType": "spot"
    }
}
```

ABC Online crée le Flotte EC2 à l'aide de la commande create-fleet suivante.

```
aws ec2 create-fleet \
    --cli-input-json file://config.json
```

Pour de plus amples informations, veuillez consulter Création d'un Flotte EC2 (p. 579).

Exécution

La stratégie d'allocation détermine que la capacité à la demande est toujours assurée, tandis que l'équilibre de la capacité cible est assuré sous la forme d'instances spot si la capacité et la disponibilité nécessaires sont assurées.

Gestion d'un Flotte EC2

Pour utiliser un Flotte EC2, vous créez une demande comprenant la capacité cible totale, une capacité à la demande, une capacité d'instances Spot, une ou plusieurs spécifications de lancement pour les instances

et le prix maximum que vous êtes prêt à payer. La demande de parc d'instances doit inclure un modèle de lancement qui définit les informations dont le parc d'instances a besoin pour lancer une instance, par exemple une AMI, un type d'instance, un sous-réseau ou une zone de disponibilité, et un ou plusieurs groupes de sécurité. Vous pouvez spécifier des remplacements de spécification de lancement pour le type d'instance, le sous-réseau, la zone de disponibilité et le prix maximum que vous êtes prêt à payer, et vous pouvez affecter une capacité pondérée à chaque remplacement de spécification de lancement.

Si votre parc d'instances inclut des Instances Spot, Amazon EC2 tente de maintenir la capacité cible de votre parc d'instances au fur et à mesure de l'évolution des prix Spot.

Une demande de Flotte EC2 reste active jusqu'à ce qu'elle arrive à expiration ou que vous la supprimiez. Lorsque vous supprimez un parc d'instances, vous pouvez spécifier si la suppression met hors service les instances dans le parc d'instances.

Sommaire

- États des demandes Flotte EC2 (p. 571)
- Conditions préalables requises Flotte EC2 (p. 572)
- Vérifications de l'état par Flotte EC2 (p. 574)
- Génération d'un fichier de configuration JSON de Flotte EC2 (p. 575)
- Création d'un Flotte EC2 (p. 579)
- Balisage d'un Flotte EC2 (p. 582)
- Surveillance de votre Flotte EC2 (p. 583)
- Modification d'un Flotte EC2 (p. 585)
- Suppression d'un Flotte EC2 (p. 586)
- Exemples de configuration d'un Flotte EC2 (p. 587)

États des demandes Flotte EC2

Une demande de Flotte EC2 peut avoir l'un des états suivants :

- `submitted` – La demande de Flotte EC2 est en cours d'évaluation et Amazon EC2 se prépare à lancer le nombre cible d'instances, qui peuvent inclure des Instances à la demande, des Instances Spot ou les deux.
- `active` – La demande de Flotte EC2 a été validée et Amazon EC2 tente de conserver le nombre cible d'instances en cours d'exécution. La demande conserve cet état jusqu'à ce qu'elle soit modifiée ou supprimée.
- `modifying` – La demande de Flotte EC2 est en cours de modification. La demande conserve cet état jusqu'à ce que la modification soit totalement traitée ou que la demande soit supprimée. Seule une demande de type `maintain` peut être modifiée. Cet état ne s'applique pas aux autres types de demandes.
- `deleted_running` – La demande de Flotte EC2 est supprimée et ne lance pas d'instances supplémentaires. Ses instances existantes continuent de s'exécuter jusqu'à ce qu'elles soient interrompues ou mises hors service. La demande conserve cet état jusqu'à ce que toutes les instances soient interrompues ou mises hors service.
- `deleted_terminating` – La demande de Flotte EC2 est supprimée et ses instances sont résiliées. La demande conserve cet état jusqu'à ce que toutes les instances soient mises hors service.
- `deleted` – Le Flotte EC2 est supprimé et n'a aucune instance en cours d'exécution. La demande est supprimée deux jours après la mise hors service de ses instances.

L'illustration suivante représente les transitions entre les états de la demande de Flotte EC2. Si vous dépassez les limites de votre parc d'instances, la demande est immédiatement supprimée.

Conditions préalables requises Flotte EC2

Pour créer un Flotte EC2, les prérequis suivants doivent être en place.

Modèle de lancement

Un modèle de lancement inclut des informations sur les instances à lancer, telles que le type d'instance, la zone de disponibilité et le prix maximum que vous êtes disposé à payer. Pour plus d'informations, consultez Lancement d'une instance à partir d'un modèle de lancement (p. 537).

Rôle lié à un service pour Flotte EC2

Le rôle `AWSServiceRoleForEC2Fleet` accorde au Flotte EC2 l'autorisation de demander, lancer, résilier et baliser des instances en votre nom. Amazon EC2 utilise ce rôle lié à un service pour exécuter les actions suivantes :

- `ec2:RunInstances` – Lancer des instances
- `ec2:RequestSpotInstances` – Demander des Instances Spot.
- `ec2:TerminateInstances` – Résilier des instances
- `ec2:DescribeImages` – Décrire des AMI (Amazon Machine Image) pour les Instances Spot
- `ec2:DescribeInstanceStatus` – Décrire le statut des Instances Spot.
- `ec2:DescribeSubnets` – Décrire les sous-réseaux pour les Instances Spot.
- `ec2:CreateTags` – Ajoutez des balises aux Flotte EC2, aux instances et aux volumes.

Assurez-vous que ce rôle existe avant d'utiliser l'AWS CLI ou une API pour créer un Flotte EC2.

> **Note**
>
> Un `instant` Flotte EC2 ne requiert pas ce rôle.

Pour créer le rôle, utilisez la console IAM comme suit.

Pour créer le rôle AWSServiceRoleForEC2Fleet pour Flotte EC2

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le volet de navigation, choisissez Rôles, puis Créer un rôle.

3. Pour Select type of trusted entity (Sélectionner le type d'entité de confiance), choisissez Service AWS.
4. Pour Choisir le service qui utilisera ce rôle, choisissez EC2 - Fleet (EC2 - Flotte), puis choisissez Suivant : Autorisations, Suivant : Balises et Suivant : Vérification.
5. Sur la page Vérification, choisissez Create Role (Créer un rôle).

Si vous n'avez plus besoin d'utiliser le Flotte EC2, nous vous recommandons de supprimer le rôle AWSServiceRoleForEC2Fleet. Après la suppression de ce rôle de votre compte, vous pouvez créer de nouveau le rôle si vous créez un autre parc d'instances.

Pour plus d'informations, consultez Utilisation des rôles liés à un service dans le IAM Guide de l'utilisateur.

Octroi d'un accès aux CMK en vue de leur utilisation avec les AMI chiffrées et les instantanés EBS

Si vous spécifiez une AMI chiffrée (p. 168) ou un instantané Amazon EBS chiffré (p. 1194) dans votre Flotte EC2 et que vous utilisez la clé principale client gérée (CMK) pour le chiffrement, vous devez autoriser le rôle AWSServiceRoleForEC2Fleet à utiliser la CMK afin que Amazon EC2 puisse lancer les instances en votre nom. Pour cela, vous devez ajouter une autorisation à la CMK, comme indiqué dans la procédure suivante.

Lorsque vous définissez les autorisations, les octrois constituent une alternative aux stratégies de clé. Pour de plus amples informations, veuillez consulter Utilisation des autorisations et Utilisation des stratégies de clé dans AWS KMS dans le AWS Key Management Service Developer Guide.

Pour autoriser le rôle AWSServiceRoleForEC2Fleet à utiliser la CMK

- Utilisez la commande create-grant pour ajouter un octroi à la CMK et spécifier le mandataire (le rôle lié à un service AWSServiceRoleForEC2Fleet) qui reçoit l'autorisation d'effectuer les opérations autorisées par l'octroi. La CMK est spécifiée par le paramètre `key-id` et l'ARN de la CMK. Le mandataire est spécifié par le paramètre `grantee-principal` et l'ARN du rôle lié à un service AWSServiceRoleForEC2Fleet.

```
aws kms create-grant \
    --region us-east-1 \
    --key-id arn:aws:kms:us-
east-1:444455556666:key/1234abcd-12ab-34cd-56ef-1234567890ab \
    --grantee-principal arn:aws:iam::111122223333:role/AWSServiceRoleForEC2Fleet \
    --operations "Decrypt" "Encrypt" "GenerateDataKey"
 "GenerateDataKeyWithoutPlaintext" "CreateGrant" "DescribeKey" "ReEncryptFrom"
 "ReEncryptTo"
```

Utilisateurs Flotte EC2 et IAM

Si vos utilisateurs IAM sont appelés à créer ou à gérer un Flotte EC2, veillez à leur accorder les autorisations nécessaires comme suit.

Pour accorder à un utilisateur IAM des autorisations pour un Flotte EC2

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le panneau de navigation, choisissez Policies.
3. Choisissez Créer une stratégie.
4. Sur la page Créer une stratégie, choisissez l'onglet JSON, remplacez le texte par le suivant, puis choisissez Examiner une stratégie.

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Effect": "Allow",
            "Action": [
                "ec2:*"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
              "iam:ListRoles",
              "iam:PassRole",
              "iam:ListInstanceProfiles"
            ],
            "Resource": "*"
        }
    ]
}
```

La fonction `ec2:*` accorde à un utilisateur IAM l'autorisation d'appeler toutes les actions d'API Amazon EC2. Pour limiter les actions API Amazon EC2 susceptibles d'être effectuées par l'utilisateur, spécifiez celles qui sont autorisées.

Un utilisateur IAM doit être autorisé à appeler l'action `iam:ListRoles` pour énumérer les rôles IAM existants, l'action `iam:PassRole` pour spécifier le rôle de Flotte EC2 et l'action `iam:ListInstanceProfiles` pour énumérer les profils d'instance existants.

(Facultatif) Pour autoriser un utilisateur IAM à créer des rôles ou des profils d'instances à l'aide de la console IAM, vous devez aussi ajouter les actions suivantes à la stratégie :

- `iam:AddRoleToInstanceProfile`
- `iam:AttachRolePolicy`
- `iam:CreateInstanceProfile`
- `iam:CreateRole`
- `iam:GetRole`
- `iam:ListPolicies`

5. Sur la page Review Policy (Vérifier la stratégie), saisissez un nom et une description pour la stratégie, puis choisissez Create policy (Créer une stratégie).
6. Dans le panneau de navigation, choisissez Utilisateurs et sélectionnez l'utilisateur.
7. Sous l'onglet Autorisations, choisissez Ajouter des autorisations.
8. Choisissez Attacher directement les stratégies existantes. Sélectionnez la stratégie que vous avez créée précédemment, puis choisissez Suivant : Vérification.
9. Choisissez Add permissions.

Vérifications de l'état par Flotte EC2

Le Flotte EC2 vérifie l'état de santé des instances du parc d'instances toutes les deux minutes. Le statut de l'état d'une instance est `healthy` ou `unhealthy`. Le parc d'instances détermine l'état de santé d'une instance à l'aide des contrôles d'état fournis par Amazon EC2. Si le contrôle du statut de l'instance ou de celui du système est `impaired` pendant trois vérifications de l'état consécutives, le statut de l'état de l'instance est `unhealthy`. Sinon, le statut d'état est `healthy`. Pour plus d'informations, consultez Contrôles de statut pour vos instances (p. 742).

Vous pouvez configurer votre Flotte EC2 pour qu'il remplace les instances non saines. Après activation du remplacement lié à la vérification de l'état, une instance est remplacée lorsque son statut d'état est `unhealthy`. Notez que la taille du parc d'instances peut être inférieure à sa capacité cible pendant quelques minutes pendant le remplacement d'une instance non saine.

Prérequis

- Le remplacement lié à la vérification de l'état de santé est pris en charge uniquement avec les Flottes EC2 qui maintiennent une capacité cible, pas avec les parcs d'instances uniques.
- Vous pouvez configurer votre Flotte EC2 pour qu'il remplace les instances non saines au moment de sa création uniquement.
- Les utilisateurs IAM peuvent utiliser le remplacement lié à la vérification de l'état seulement s'ils sont autorisés à appeler l'action `ec2:DescribeInstanceStatus`.

Génération d'un fichier de configuration JSON de Flotte EC2

Pour créer une Flotte EC2, il vous suffit de spécifier le modèle de lancement, la capacité cible totale et si l'option d'achat par défaut est à la demande ou Spot. Si vous ne spécifiez pas un paramètre, le parc d'instances utilise la valeur par défaut. Pour afficher la liste complète des paramètres de configuration du parc d'instances, vous pouvez générer un fichier JSON comme suit.

Pour générer un fichier JSON avec tous les paramètres de Flotte EC2 possibles à l'aide de la ligne de commande

- Utilisez la commande create-fleet (AWS CLI) et le paramètre `--generate-cli-skeleton` pour générer un fichier JSON de Flotte EC2 :

```
aws ec2 create-fleet \
    --generate-cli-skeleton
```

Les paramètres de Flotte EC2 suivants sont disponibles :

```
{
    "DryRun": true,
    "ClientToken": "",
    "SpotOptions": {
        "AllocationStrategy": "lowest-price",
        "InstanceInterruptionBehavior": "hibernate",
        "InstancePoolsToUseCount": 0,
        "SingleInstanceType": true,
        "SingleAvailabilityZone": true,
        "MaxTotalPrice": 0,
        "MinTargetCapacity": 0
    },
    "OnDemandOptions": {
        "AllocationStrategy": "prioritized",
        "SingleInstanceType": true,
        "SingleAvailabilityZone": true,
        "MaxTotalPrice": 0,
        "MinTargetCapacity": 0
    },
    "ExcessCapacityTerminationPolicy": "termination",
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "",
                "LaunchTemplateName": "",
                "Version": ""
            },
            "Overrides": [
                {
                    "InstanceType": "t2.micro",
                    "MaxPrice": "",
                    "SubnetId": "",
                    "AvailabilityZone": "",
```

```
                    "WeightedCapacity": null,
                    "Priority": null,
                    "Placement": {
                        "AvailabilityZone": "",
                        "Affinity": "",
                        "GroupName": "",
                        "PartitionNumber": 0,
                        "HostId": "",
                        "Tenancy": "dedicated",
                        "SpreadDomain": ""
                }
            ]
        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 0,
        "OnDemandTargetCapacity": 0,
        "SpotTargetCapacity": 0,
        "DefaultTargetCapacityType": "spot"
    },
    "TerminateInstancesWithExpiration": true,
    "Type": "maintain",
    "ValidFrom": "1970-01-01T00:00:00",
    "ValidUntil": "1970-01-01T00:00:00",
    "ReplaceUnhealthyInstances": true,
    "TagSpecifications": [
        {
            "ResourceType": "fleet",
            "Tags": [
                {
                    "Key": "",
                    "Value": ""
                }
            ]
        }
    ]
}
```

Référence du fichier de configuration JSON de Flotte EC2

Note

Utilisez des minuscules pour toutes les valeurs de paramètres. Sinon, vous obtenez une erreur quand Amazon EC2 utilise le fichier JSON pour lancer le Flotte EC2.

AllocationStrategy (pour SpotOptions)

(Facultatif) Indique comment allouer la capacité cible des instance Spot sur l'ensemble des groupes d'instance Spot spécifiés par le Flotte EC2. Les valeurs valides sont `lowest-price`, `diversified` et `capacity-optimized`. La valeur par défaut est `lowest-price`. Spécifiez la stratégie d'allocation qui répond à vos besoins. Pour de plus amples informations, veuillez consulter Stratégies d'allocation pour Instances Spot (p. 562).

InstanceInterruptionBehavior

(Facultatif) Comportement lorsqu'une instance Spot est interrompue. Les valeurs valides sont `hibernate`, `stop` et `terminate`. Par défaut, le service Spot met hors service les Instances Spot lorsqu'elles sont interrompues. Si le type de parc est `maintain`, vous pouvez demander que le service Spot mette en veille prolongée ou arrête les Instances Spot lorsqu'elles sont interrompues.

InstancePoolsToUseCount

Nombre de groupes d'instances Spot auxquels allouer votre capacité Spot cible. Valide uniquement lorsque le paramètre Spot AllocationStrategy a pour valeur `lowest-price`. Le Flotte EC2 sélectionne

les groupes d'instances Spot les moins chers et alloue votre capacité Spot cible en la répartissant équitablement entre le nombre de groupes d'instances Spot que vous spécifiez.

SingleInstanceType

Indique que le parc utilise un type d'instance unique pour lancer toutes les Instances Spot dans le parc.

SingleAvailabilityZone

Indique que le parc lance toutes les Instances Spot dans une seule zone de disponibilité.

MaxTotalPrice

Montant maximum par heure pour Instances Spot que vous êtes prêt à payer.

MinTargetCapacity

Capacité cible minimum pour Instances Spot dans le parc. Si la capacité cible minimum n'est pas atteinte, le parc ne lance aucune instance.

AllocationStrategy (pour OnDemandOptions)

Ordre des remplacements du modèle de lancement à utiliser lors du traitement de l'affectation de capacité à la demande. Si vous spécifiez `lowest-price`, le Flotte EC2 utilise le prix pour déterminer l'ordre, en lançant le prix le plus bas en premier. Si vous spécifiez la valeur prioritized, le Flotte EC2 utilise la priorité que vous avez affectée à chaque remplacement du modèle de lancement, en lançant la priorité la plus élevée en premier. Si vous ne spécifiez aucune valeur, le Flotte EC2 utilise la valeur `lowest-price` par défaut.

SingleInstanceType

Indique que le parc utilise un type d'instance unique pour lancer toutes les instances à la demande dans le parc.

SingleAvailabilityZone

Indique que le parc lance toutes les instances à la demande dans une seule zone de disponibilité.

MaxTotalPrice

Montant maximum par heure pour les instances à la demande que vous êtes prêt à payer.

MinTargetCapacity

Capacité cible minimum pour les instances à la demande dans le parc. Si la capacité cible minimum n'est pas atteinte, le parc ne lance aucune instance.

ExcessCapacityTerminationPolicy

(Facultatif) Indique si les instances en cours d'exécution doivent être résiliées si la capacité cible totale du Flotte EC2 est définie sous la taille actuelle du Flotte EC2. Les valeurs valides sont `no-termination` et `termination`.

LaunchTemplateId

ID du modèle de lancement à utiliser. Vous devez spécifier l'ID ou le nom du modèle de lancement. Le modèle de lancement doit spécifier une Amazon Machine Image (AMI). Pour de plus amples informations sur la création de modèles de lancement, veuillez consulter Lancement d'une instance à partir d'un modèle de lancement (p. 537).

LaunchTemplateName

Nom du modèle de lancement à utiliser. Vous devez spécifier l'ID ou le nom du modèle de lancement. Le modèle de lancement doit spécifier une Amazon Machine Image (AMI). Pour plus d'informations, consultez Lancement d'une instance à partir d'un modèle de lancement (p. 537).

Version

Numéro de version du modèle de lancement, `$Latest` ou `$Default`. Vous devez spécifier une valeur, sinon la demande échoue. Si la valeur est `$Latest`, Amazon EC2 utilise la dernière version du

modèle de lancement. Si la valeur est `$Default`, Amazon EC2 utilise la version par défaut du modèle de lancement. Pour de plus amples informations, veuillez consulter Gestion des versions du modèle de lancement (p. 548).

InstanceType

(Facultatif) Type d'instance. Cette valeur, si elle est entrée, remplace le modèle de lancement. Les types d'instance doivent avoir les spécifications matérielles minimum dont vous avez besoin (vCPU, mémoire ou stockage).

MaxPrice

(Facultatif) Prix maximum par unité-heure que vous êtes disposé à payer pour une instance Spot. Cette valeur, si elle est entrée, remplace le modèle de lancement. Vous pouvez utiliser le prix maximum par défaut (prix à la demande) ou indiquer le prix maximum que vous êtes disposé à payer. Vos Instances Spot ne sont pas lancées si votre prix maximum est inférieur au prix spot pour les types d'instance que vous avez spécifiés.

SubnetId

(Facultatif) ID du sous-réseau dans lequel lancer les instances. Cette valeur, si elle est entrée, remplace le modèle de lancement.

Pour créer un nouveau VPC, accédez à la console Amazon VPC. Lorsque vous avez terminé, revenez dans le fichier JSON et entrez le nouvel ID de sous-réseau.

AvailabilityZone

(Facultatif) Zone de disponibilité dans laquelle lancer les instances. Le réglage par défaut consiste à laisser AWS choisir les zones pour vos instances. Si vous préférez, vous pouvez spécifier des zones spécifiques. Cette valeur, si elle est entrée, remplace le modèle de lancement.

Spécifiez une ou plusieurs zones de disponibilité. Si vous avez plusieurs sous-réseaux dans une zone, spécifiez le sous-réseau approprié. Pour ajouter des sous-réseaux, accédez à la console Amazon VPC. Lorsque vous avez terminé, revenez dans le fichier JSON et entrez le nouvel ID de sous-réseau.

WeightedCapacity

(Facultatif) Nombre d'unités fournies par le type d'instance spécifié. Cette valeur, si elle est entrée, remplace le modèle de lancement.

Priority

Priorité du remplacement du modèle de lancement. Si AllocationStrategy a pour valeur `prioritized`, le Flotte EC2 utilise la priorité pour déterminer quel remplacement du modèle de lancement utiliser en premier afin de traiter l'affectation de capacité à la demande. La priorité la plus élevée est lancée en premier. Les valeurs valides sont les nombres entiers à partir de `0`. Plus le nombre est bas, plus la priorité est élevée. Si aucun nombre n'est défini, le remplacement a la priorité la plus faible.

TotalTargetCapacity

Nombre d'instances à lancer. Vous pouvez choisir des instances ou des caractéristiques de performances importantes pour la charge de travail de votre application, par exemple les vCPU, la mémoire ou le stockage. Si le type de demande est `maintain`, vous pouvez spécifier une capacité cible de 0 et ajouter une capacité ultérieurement.

OnDemandTargetCapacity

(Facultatif) Nombre d'Instances à la demande à lancer. Ce nombre doit être inférieur à `TotalTargetCapacity`.

SpotTargetCapacity

(Facultatif) Nombre d'Instances Spot à lancer. Ce nombre doit être inférieur à `TotalTargetCapacity`.

DefaultTargetCapacityType

Si la valeur de `TotalTargetCapacity` est supérieure aux valeurs combinées pour `OnDemandTargetCapacity` et `SpotTargetCapacity`, la différence est lancée sous forme d'option d'achat d'instance spécifié ici. Les valeurs valides sont `on-demand` ou `spot`.

TerminateInstancesWithExpiration

(Facultatif) Par défaut, Amazon EC2 résilie vos instances à l'expiration de la demande de Flotte EC2. La valeur par défaut est `true`. Pour les maintenir actives après l'expiration de votre demande, n'entrez pas de valeur pour ce paramètre.

Type

(Facultatif) Indique si le Flotte EC2 envoie une demande unique synchrone pour la capacité souhaitée (`instant`) ou une demande unique asynchrone pour la capacité souhaitée, mais sans tenter de maintenir la capacité ni d'envoyer des demandes dans d'autres groupes de capacité si la capacité n'est pas disponible (`request`), ou envoie une demande asynchrone pour la capacité souhaitée et continue de maintenir la capacité souhaitée en réapprovisionnant les Instances Spot interrompues (`maintain`). Les valeurs valides sont `instant`, `request` et `maintain`. La valeur par défaut est `maintain`. Pour plus d'informations, consultez Types de demande Flotte EC2 (p. 562).

ValidFrom

(Facultatif) Pour créer une demande valide uniquement au cours d'une période spécifique, entrez une date de début.

ValidUntil

(Facultatif) Pour créer une demande valide uniquement au cours d'une période spécifique, entrez une date de fin.

ReplaceUnhealthyInstances

(Facultatif) Pour remplacer les instances non saines dans un Flotte EC2 configuré sur `maintain`, le parc d'instances, entrez `true`. Sinon, laissez ce paramètre vide.

TagSpecifications

(Facultatif) Paire clé-valeur pour le balisage de la demande Flotte EC2 à sa création. La valeur pour `ResourceType` doit être `fleet` ; sinon, la demande de parc d'instances échoue. Pour baliser les instances au moment du lancement, spécifiez les balises dans le modèle de lancement (p. 539). Pour obtenir des informations sur le balisage après lancement, consultez Balisage de vos ressources (p. 1310).

Création d'un Flotte EC2

Lorsque vous créez un Flotte EC2, vous devez spécifier un modèle de lancement qui inclut des informations sur les instances à lancer, telles que le type d'instance, la zone de disponibilité et le prix maximum que vous êtes disposé à payer.

Vous pouvez créer un Flotte EC2 incluant plusieurs spécifications de lancement qui remplacent le modèle de lancement. Les spécifications de lancement peuvent varier en terme de type d'instance, zone de disponibilité, sous-réseau ou prix maximum, et elles peuvent inclure une capacité pondérée différente.

Lorsque vous créez un Flotte EC2, utilisez un fichier JSON pour spécifier des informations sur les instances à lancer. Pour plus d'informations, consultez Référence du fichier de configuration JSON de Flotte EC2 (p. 576).

Flottes EC2 peut uniquement être créé à l'aide de l'AWS CLI.

Pour créer un Flotte EC2 (AWS CLI)

- Utilisez la commande create-fleet (AWS CLI) pour créer un Flotte EC2.

```
aws ec2 create-fleet \
    --cli-input-json file://file_name.json
```

Pour accéder à des exemples de fichiers de configuration, consultez Exemples de configuration d'un Flotte EC2 (p. 587).

Voici un exemple de sortie d'un parc d'instances du type `request` ou `maintain`.

```
{
    "FleetId": "fleet-12a34b55-67cd-8ef9-ba9b-9208dEXAMPLE"
}
```

Voici un exemple de sortie d'un parc d'instances du type `instant` qui a lancé la capacité cible.

```
{
  "FleetId": "fleet-12a34b55-67cd-8ef9-ba9b-9208dEXAMPLE",
  "Errors": [],
  "Instances": [
    {
      "LaunchTemplateAndOverrides": {
        "LaunchTemplateSpecification": {
          "LaunchTemplateId": "lt-01234a567b8910abcEXAMPLE",
          "Version": "1"
        },
        "Overrides": {
          "InstanceType": "c5.large",
          "AvailabilityZone": "us-east-1a"
        }
      },
      "Lifecycle": "on-demand",
      "InstanceIds": [
        "i-1234567890abcdef0",
        "i-9876543210abcdef9"
      ],
      "InstanceType": "c5.large",
      "Platform": null
    },
    {
      "LaunchTemplateAndOverrides": {
        "LaunchTemplateSpecification": {
          "LaunchTemplateId": "lt-01234a567b8910abcEXAMPLE",
          "Version": "1"
        },
        "Overrides": {
          "InstanceType": "c4.large",
          "AvailabilityZone": "us-east-1a"
        }
      },
      "Lifecycle": "on-demand",
      "InstanceIds": [
        "i-5678901234abcdef0",
        "i-5432109876abcdef9"
      ],
      "InstanceType": "c4.large",
      "Platform": null
    },
  ]
}
```

Voici un exemple de sortie d'un parc d'instances du type `instant` qui a lancé une partie de la capacité cible avec les erreurs liées aux instances qui n'ont pas été lancées.

```
{
  "FleetId": "fleet-12a34b55-67cd-8ef9-ba9b-9208dEXAMPLE",
  "Errors": [
    {
      "LaunchTemplateAndOverrides": {
        "LaunchTemplateSpecification": {
          "LaunchTemplateId": "lt-01234a567b8910abcEXAMPLE",
          "Version": "1"
        },
        "Overrides": {
          "InstanceType": "c4.xlarge",
          "AvailabilityZone": "us-east-1a",
        }
      },
      "Lifecycle": "on-demand",
      "ErrorCode": "InsufficientInstanceCapacity",
      "ErrorMessage": "",
      "InstanceType": "c4.xlarge",
      "Platform": null
    },
  ],
  "Instances": [
    {
      "LaunchTemplateAndOverrides": {
        "LaunchTemplateSpecification": {
          "LaunchTemplateId": "lt-01234a567b8910abcEXAMPLE",
          "Version": "1"
        },
        "Overrides": {
          "InstanceType": "c5.large",
          "AvailabilityZone": "us-east-1a"
        }
      },
      "Lifecycle": "on-demand",
      "InstanceIds": [
        "i-1234567890abcdef0",
        "i-9876543210abcdef9"
      ],
      "InstanceType": "c5.large",
      "Platform": null
    },
  ]
}
```

Voici un exemple de sortie d'un parc d'instances du type `instant` qui n'a lancé aucune instance.

```
{
  "FleetId": "fleet-12a34b55-67cd-8ef9-ba9b-9208dEXAMPLE",
  "Errors": [
    {
      "LaunchTemplateAndOverrides": {
        "LaunchTemplateSpecification": {
          "LaunchTemplateId": "lt-01234a567b8910abcEXAMPLE",
          "Version": "1"
        },
        "Overrides": {
          "InstanceType": "c4.xlarge",
          "AvailabilityZone": "us-east-1a",
        }
      },
      "Lifecycle": "on-demand",
      "ErrorCode": "InsufficientCapacity",
      "ErrorMessage": "",
      "InstanceType": "c4.xlarge",
```

```
      "Platform": null
    },
    {
      "LaunchTemplateAndOverrides": {
        "LaunchTemplateSpecification": {
          "LaunchTemplateId": "lt-01234a567b8910abcEXAMPLE",
          "Version": "1"
        },
        "Overrides": {
          "InstanceType": "c5.large",
          "AvailabilityZone": "us-east-1a",
        }
      },
      "Lifecycle": "on-demand",
      "ErrorCode": "InsufficientCapacity",
      "ErrorMessage": "",
      "InstanceType": "c5.large",
      "Platform": null
    },
  ],
  "Instances": []
}
```

Balisage d'un Flotte EC2

Pour vous aider à classer et à gérer vos demandes de Flotte EC2, vous pouvez les baliser avec des métadonnées personnalisées. Vous pouvez affecter une balise à une demande de Flotte EC2 lorsque vous la créez, ou après.

Lorsque vous balisez une demande de parc, les instances et les volumes lancés par le parc ne sont pas balisés automatiquement. Vous devez baliser explicitement les instances et les volumes lancés par le parc. Vous pouvez choisir d'affecter des balises uniquement à la demande de parc, ou uniquement aux instances lancées par le parc, ou uniquement aux volumes attachés aux instances lancées par le parc, ou aux trois.

> **Note**
>
> Pour les types de parc `instant`, vous pouvez baliser les volumes attachés à Instances à la demande et Instances Spot. Pour les types de parc `request` ou `maintain`, vous pouvez uniquement baliser les volumes attachés à Instances à la demande.

Pour plus d'informations sur le fonctionnement des balises, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Prérequis

Octroyez à l'utilisateur IAM l'autorisation de baliser les ressources. Pour plus d'informations, consultez Exemple : Balisage de vos ressources (p. 1018).

Pour accorder à un utilisateur IAM l'autorisation de baliser les ressources

Créez une stratégie IAM qui inclut les éléments suivants :

- L'action `ec2:CreateTags`. Celle-ci accorde à l'utilisateur IAM l'autorisation de créer des balises.
- L'action `ec2:CreateFleet`. Celle-ci accorde à l'utilisateur IAM l'autorisation de créer une demande Flotte EC2.
- Pour `Resource`, nous vous recommandons de spécifier `"*"`. Cela permet aux utilisateurs de baliser tous les types de ressources.

```
{
    "Version": "2012-10-17",
```

```
    "Statement": [
        {
            "Sid": "TagEC2FleetRequest",
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags",
                "ec2:CreateFleet"
            ],
            "Resource": "*"
}
```

> **Important**
>
> Actuellement, nous ne prenons pas en charge les autorisations de niveau ressource pour la ressource `create-fleet`. Si vous spécifiez `create-fleet` en tant que ressource, vous recevrez une exception de non-autorisation lorsque vous tenterez de baliser le parc. L'exemple suivant illustre comme ne pas définir la stratégie.

```
{
    "Effect": "Allow",
    "Action": [
        "ec2:CreateTags",
        "ec2:CreateFleet"
    ],
    "Resource": "arn:aws:ec2:us-east-1:111122223333:create-fleet/*"
}
```

Pour baliser une nouvelle demande Flotte EC2

Pour baliser une demande de Flotte EC2 lorsque vous la créez, spécifiez la paire clé-valeur dans le fichier JSON (p. 575) utilisé pour créer le parc d'instances. La valeur pour `ResourceType` doit être `fleet`. Si vous spécifiez une autre valeur, la demande de parc d'instances échoue.

Pour baliser des instances et des volumes lancés par un Flotte EC2

Pour baliser des instances et des volumes lorsqu'ils sont lancés par le parc d'instances, spécifiez les balises dans le modèle de lancement (p. 539) référencé dans la demande Flotte EC2.

> **Note**
>
> Vous ne pouvez pas baliser les volumes attachés à Instances Spot qui sont lancés par un type de parc `request` ou `maintain`.

Pour baliser une demande Flotte EC2, une instance et un volume existants (AWS CLI)

Utilisez la commande create-tags pour baliser les ressources existantes.

```
aws ec2 create-tags \
    --resources fleet-12a34b55-67cd-8ef9-
ba9b-9208dEXAMPLE i-1234567890abcdef0 vol-1234567890EXAMPLE \
    --tags Key=purpose,Value=test
```

Surveillance de votre Flotte EC2

Le Flotte EC2 lance des Instances à la demande lorsque la capacité requise est disponible, et il lance des Instances Spot lorsque votre prix maximum dépasse le prix spot et que la capacité est disponible. Les Instances à la demande s'exécutent jusqu'à ce que vous les résiliiez, et les Instances Spot s'exécutent jusqu'à ce qu'elles soient interrompues ou que vous les résiliiez.

La liste renvoyée des instances en cours d'exécution est actualisée périodiquement et peut ne pas être à jour.

Pour surveiller votre Flotte EC2 (AWS CLI)

Utilisez la commande describe-fleets suivante pour décrire vos Flottes EC2 :

```
aws ec2 describe-fleets
```

Voici un exemple de sortie.

```
{
    "Fleets": [
        {
            "Type": "maintain",
            "FulfilledCapacity": 2.0,
            "LaunchTemplateConfigs": [
                {
                    "LaunchTemplateSpecification": {
                        "Version": "2",
                        "LaunchTemplateId": "lt-07b3bc7625cdab851"
                    }
                }
            ],
            "TerminateInstancesWithExpiration": false,
            "TargetCapacitySpecification": {
                "OnDemandTargetCapacity": 0,
                "SpotTargetCapacity": 2,
                "TotalTargetCapacity": 2,
                "DefaultTargetCapacityType": "spot"
            },
            "FulfilledOnDemandCapacity": 0.0,
            "ActivityStatus": "fulfilled",
            "FleetId": "fleet-76e13e99-01ef-4bd6-ba9b-9208de883e7f",
            "ReplaceUnhealthyInstances": false,
            "SpotOptions": {
                "InstanceInterruptionBehavior": "terminate",
                "InstancePoolsToUseCount": 1,
                "AllocationStrategy": "lowest-price"
            },
            "FleetState": "active",
            "ExcessCapacityTerminationPolicy": "termination",
            "CreateTime": "2018-04-10T16:46:03.000Z"
        }
    ]
}
```

Utilisez la commande describe-fleet-instances suivante afin de décrire les instances pour le Flotte EC2 spécifié.

```
aws ec2 describe-fleet-instances \
    --fleet-id fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE
```

```
{
    "ActiveInstances": [
        {
            "InstanceId": "i-09cd595998cb3765e",
            "InstanceHealth": "healthy",
            "InstanceType": "m4.large",
            "SpotInstanceRequestId": "sir-86k84j6p"
        },
        {
            "InstanceId": "i-09cf95167ca219f17",
            "InstanceHealth": "healthy",
            "InstanceType": "m4.large",
```

```
            "SpotInstanceRequestId": "sir-dvxi7fsm"
        }
    ],
    "FleetId": "fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE"
}
```

Utilisez la commande describe-fleet-history suivante afin de décrire l'historique du Flotte EC2 spécifié pour la période spécifiée.

```
aws ec2 describe-fleet-history --fleet-request-id fleet-73fbd2ce-
aa30-494c-8788-1cee4EXAMPLE --start-time 2018-04-10T00:00:00Z
```

```
{
    "HistoryRecords": [],
    "FleetId": "fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE",
    "LastEvaluatedTime": "1970-01-01T00:00:00.000Z",
    "StartTime": "2018-04-09T23:53:20.000Z"
}
```

Modification d'un Flotte EC2

Vous pouvez modifier un Flotte EC2 qui présente l'état `submitted` ou `active`. Lorsque vous modifiez un parc d'instances, il prend l'état `modifying`.

Vous pouvez uniquement modifier un Flotte EC2 qui a le type `maintain`. Vous ne pouvez pas modifier un Flotte EC2 ayant le type `request` ou `instant`.

Vous pouvez modifier les paramètres suivants d'un Flotte EC2 :

- `target-capacity-specification` – Augmentez ou diminuez la capacité cible pour `TotalTargetCapacity`, `OnDemandTargetCapacity` et `SpotTargetCapacity`.
- `excess-capacity-termination-policy` – Indiquez si les instances en cours d'exécution doivent être résiliées si la capacité cible totale du Flotte EC2 est définie sous la taille actuelle du parc d'instances. Les valeurs valides sont `no-termination` et `termination`.

Lorsque vous augmentez la capacité cible, la Flotte EC2 lance les instances supplémentaires en fonction de l'option d'achat d'instance spécifiée pour `DefaultTargetCapacityType`, qui correspond à des Instances à la demande ou à des Instances Spot.

Si le paramètre `DefaultTargetCapacityType` a pour valeur `spot`, le Flotte EC2 lance les Instances Spot supplémentaires en fonction de sa stratégie d'allocation. Si la stratégie d'allocation `lowest-price` est sélectionnée, le parc d'instances lance les instances du groupe d'instance Spot offrant le tarif le moins élevé de la demande. Si la stratégie d'allocation `diversified` est sélectionnée, le parc d'instances répartit les instances entre les groupes de la demande.

Lorsque vous diminuez la capacité cible, le Flotte EC2 supprime toutes les demandes ouvertes qui dépassent la nouvelle capacité cible. Vous pouvez demander à ce que le parc d'instances mette hors service les instances jusqu'à ce que la taille du parc atteigne la nouvelle capacité cible. Si la stratégie d'allocation est `lowest-price`, le parc d'instances met hors service les instances ayant le prix par unité le plus élevé. En revanche, si la stratégie d'allocation est `diversified`, le parc d'instances met hors service les instances des divers pools. Vous pouvez aussi demander à ce que le Flotte EC2 conserve sa taille actuelle, mais sans remplacer les Instances Spot interrompues ni les instances que vous résiliez manuellement.

Lorsqu'un Flotte EC2 résilie une instance Spot du fait de la diminution de la capacité cible, l'instance reçoit un avis d'interruption d'instance Spot.

Pour modifier un Flotte EC2 (AWS CLI)

Utilisez la commande modify-fleet suivante pour mettre à jour la capacité cible du Flotte EC2 spécifié.

```
aws ec2 modify-fleet \
    --fleet-id fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
    --target-capacity-specification TotalTargetCapacity=20
```

Si vous diminuez la capacité cible, mais que vous souhaitez conserver la taille actuelle du parc d'instances, vous pouvez modifier la commande précédente comme suit :

```
aws ec2 modify-fleet \
    --fleet-id fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
    --target-capacity-specification TotalTargetCapacity=10 \
    --excess-capacity-termination-policy no-termination
```

Suppression d'un Flotte EC2

Si vous n'avez plus besoin d'un Flotte EC2, vous pouvez le supprimer. Après avoir supprimé un parc d'instances, il ne lance pas de nouvelles instances.

Vous devez indiquer si le Flotte EC2 doit mettre hors service ses instances. Si vous spécifiez que les instances doivent être mises hors service lors de la suppression du parc d'instances, ce dernier prend l'état `deleted_terminating`. Sinon, il passe à l'état `deleted_running` et les instances continuent à s'exécuter jusqu'à ce qu'elles soient interrompues ou jusqu'à ce que vous les mettiez hors service manuellement.

Vous ne pouvez supprimer que des parcs de type `request` et `maintain`. Vous ne pouvez pas supprimer un `instant` Flotte EC2.

Pour supprimer un Flotte EC2 (AWS CLI)

Utilisez la commande delete-fleets et le paramètre `--terminate-instances` pour supprimer le Flotte EC2 spécifié et résilier les instances :

```
aws ec2 delete-fleets \
    --fleet-ids fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
    --terminate-instances
```

Voici un exemple de sortie.

```
{
    "UnsuccessfulFleetDeletions": [],
    "SuccessfulFleetDeletions": [
        {
            "CurrentFleetState": "deleted_terminating",
            "PreviousFleetState": "active",
            "FleetId": "fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE"
        }
    ]
}
```

Vous pouvez modifier la commande précédente avec le paramètre `--no-terminate-instances` pour supprimer le Flotte EC2 spécifié sans résilier les instances :

```
aws ec2 delete-fleets \
    --fleet-ids fleet-73fbd2ce-aa30-494c-8788-1cee4EXAMPLE \
    --no-terminate-instances
```

Voici un exemple de sortie.

```
{
    "UnsuccessfulFleetDeletions": [],
    "SuccessfulFleetDeletions": [
        {
            "CurrentFleetState": "deleted_running",
            "PreviousFleetState": "active",
            "FleetId": "fleet-4b8aaae8-dfb5-436d-a4c6-3dafa4c6b7dcEXAMPLE"
        }
    ]
}
```

Raisons de l'échec de la suppression

Si un Flotte EC2 ne peut pas être supprimé, `UnsuccessfulFleetDeletions` renvoie l'ID du Flotte EC2, un code d'erreur et un message d'erreur. Les codes d'erreur sont `fleetIdDoesNotExist`, `fleetIdMalformed`, `fleetNotInDeletableState` et `unexpectedError`.

Exemples de configuration d'un Flotte EC2

Les exemples suivants montrent les configurations de lancement que vous pouvez utiliser avec la commande create-fleet pour créer une Flotte EC2. Pour plus d'informations sur les paramètres create-fleet consultez Référence du fichier de configuration JSON de Flotte EC2 (p. 576).

Exemples

- Exemple 1 : Lancer Instances Spot en tant qu'option d'achat par défaut (p. 587)
- Exemple 2 : Lancer Instances à la demande en tant qu'option d'achat par défaut (p. 588)
- Exemple 3 : Lancer Instances à la demande en tant que capacité principale (p. 588)
- Exemple 4 : Lancer Instances Spot à l'aide de la stratégie d'attribution lowest-price (p. 588)
- Exemple 5 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution prioritized (p. 589)
- Exemple 6 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution prioritized lorsque la capacité cible totale est supérieure au nombre de Réservations de capacité inutilisés (p. 591)
- Exemple 7 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution lowest-price (p. 593)
- Exemple 8 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution lowest-price lorsque la capacité cible totale est supérieure au nombre de Réservations de capacité inutilisés (p. 596)

Exemple 1 : Lancer Instances Spot en tant qu'option d'achat par défaut

L'exemple suivant spécifie les paramètres minimum requis dans une Flotte EC2 : un modèle de lancement, une capacité cible et une option d'achat par défaut. Le modèle de lancement est identifié par son ID de modèle de lancement et son numéro de version. La capacité cible du parc d'instances est de 2 instances et l'option d'achat par défaut est `spot`, ce qui entraîne le lancement par le parc d'instances de 2 Instances Spot.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-0e8c754449b27161c",
                "Version": "1"
            }

        }
```

```
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 2,
        "DefaultTargetCapacityType": "spot"
    }
}
```

Exemple 2 : Lancer Instances à la demande en tant qu'option d'achat par défaut

L'exemple suivant spécifie les paramètres minimum requis dans une Flotte EC2 : un modèle de lancement, une capacité cible et une option d'achat par défaut. Le modèle de lancement est identifié par son ID de modèle de lancement et son numéro de version. La capacité cible du parc d'instances est de 2 instances et l'option d'achat par défaut est `on-demand`, ce qui entraîne le lancement par le parc d'instances de 2 Instances à la demande.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-0e8c754449b27161c",
                "Version": "1"
            }

        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 2,
        "DefaultTargetCapacityType": "on-demand"
    }
}
```

Exemple 3 : Lancer Instances à la demande en tant que capacité principale

L'exemple suivant spécifie la capacité cible totale de 2 instances pour le parc d'instances et une capacité cible de 1 instance à la demande. L'option d'achat par défaut est `spot`. Le parc d'instances lance 1 instance à la demande comme spécifié, mais a besoin de lancer une instance supplémentaire pour assurer la capacité cible totale. L'option d'achat pour la différence est calculé comme `TotalTargetCapacity` – `OnDemandTargetCapacity` = `DefaultTargetCapacityType`, ce qui entraîne le lancement par le parc d'instances de 1 instance Spot.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-0e8c754449b27161c",
                "Version": "1"
            }

        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 2,
        "OnDemandTargetCapacity": 1,
        "DefaultTargetCapacityType": "spot"
    }
}
```

Exemple 4 : Lancer Instances Spot à l'aide de la stratégie d'attribution `lowest-price`

Si la stratégie d'allocation pour les Instances Spot n'est pas spécifiée, la stratégie d'allocation par défaut, à savoir `lowest-price`, est utilisée. L'exemple suivant utilise la stratégie d'attribution `lowest-price`.

Les trois spécifications de lancement, qui remplacent le modèle de lancement, ont des types d'instance différents mais la même capacité pondérée et le même sous-réseau. La capacité cible totale est de 2 instances et l'option d'achat par défaut est `spot`. Le Flotte EC2 lance 2 Instances Spot en utilisant le type d'instance de la spécification de lancement au prix le plus bas.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-0e8c754449b27161c",
                "Version": "1"
            }
  "Overrides": [
        {
          "InstanceType": "c4.large",
          "WeightedCapacity": 1,
          "SubnetId": "subnet-a4f6c5d3"
        },
        {
          "InstanceType": "c3.large",
          "WeightedCapacity": 1,
          "SubnetId": "subnet-a4f6c5d3"
        },
        {
          "InstanceType": "c5.large",
          "WeightedCapacity": 1,
          "SubnetId": "subnet-a4f6c5d3"
        }
      ]

        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 2,
        "DefaultTargetCapacityType": "spot"
    }
}
```

Exemple 5 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution `prioritized`

Vous pouvez configurer une flotte pour qu'elle utilise d'abord Réservations de capacité à la demande lors du lancement d'Instances à la demande en définissant la stratégie d'utilisation pour Réservations de capacité sur `use-capacity-reservations-first`. Et si plusieurs groupes d'instances n'utilisent pas Réservations de capacité, la stratégie d'allocation à la demande choisie est appliquée. Dans cet exemple, la stratégie d'allocation à la demande est `prioritized`.

Dans cet exemple, il y a 15 Réservations de capacité inutilisés disponibles. Ce chiffre est supérieur à la capacité à la demande cible de la flotte de 12 Instances à la demande.

Le compte a les 15 Réservations de capacité suivants inutilisés dans 3 groupes différents. Le nombre de Réservations de capacité dans chaque pool est indiqué par `AvailableInstanceCount`.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "c4.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}
```

```
{
    "CapacityReservationId": "cr-222",
    "InstanceType": "c3.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}

{
    "CapacityReservationId": "cr-333",
    "InstanceType": "c5.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount":5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}
```

La configuration de flotte suivante affiche uniquement les configurations pertinentes pour cet exemple. La stratégie d'allocation à la demande est `prioritized`, et la stratégie d'utilisation pour Réservations de capacité est `use-capacity-reservations-first`. La capacité totale cible est 12 et le type de capacité cible par défaut est `on-demand`.

> **Note**
>
> Le type de flotte doit être `instant`. Les Réservations de capacité ne sont pas pris en charge pour les autres types de flotte.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-1234567890abcdefg",
                "Version": "1"
            }
            "Overrides": [
                {
                  "InstanceType": "c4.large",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1,
                  "Priority": 1.0
                },
                {
                  "InstanceType": "c3.large",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1,
                  "Priority": 2.0
                },
                {
                  "InstanceType": "c5.large",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1,
                  "Priority": 3.0
                }
              ]
        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 12,
        "DefaultTargetCapacityType": "on-demand"
    },
```

```
    "OnDemandOptions": {
        "AllocationStrategy": "prioritized"
        "CapacityReservationOptions": {
            "UsageStrategy": "use-capacity-reservations-first"
        }
    },
    "Type": "instant",
}
```

Après avoir créé la flotte `instant` à l'aide de la configuration précédente, les 12 instances suivantes sont lancées pour atteindre la capacité cible :

- 5 c4.large Instances à la demande dans us-east-1a – c4.large dans us-east-1a est priorisé en premier, et il y a 5 Réservations de capacité c4.large disponibles inutilisés
- 5 c3.large Instances à la demande dans us-east-1a – c3.large dans us-east-1a est priorisé en second, et il y a 5 Réservations de capacité c3.large disponibles inutilisés
- 2 c5.large Instances à la demande dans us-east-1a – c5.large dans us-east-1a est priorisé en troisième, et il y a 5 Réservations de capacité c5.large disponibles inutilisés, dont seulement 2 sont nécessaires pour atteindre la capacité cible

Une fois la flotte lancée, vous pouvez exécuter describe-capacity-reservations pour voir combien il reste de Réservations de capacité inutilisés. Dans cet exemple, vous devriez voir la réponse suivante, qui montre que tous les Réservations de capacité c4.large et c3.large ont été utilisés, avec 3 Réservations de capacité c5.large restants inutilisés.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "c4.large",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-222",
    "InstanceType": "c3.large",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-333",
    "InstanceType": "c5.large",
    "AvailableInstanceCount": 3
}
```

Exemple 6 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution `prioritized` lorsque la capacité cible totale est supérieure au nombre de Réservations de capacité inutilisés

Vous pouvez configurer une flotte pour qu'elle utilise d'abord Réservations de capacité à la demande lors du lancement d'Instances à la demande en définissant la stratégie d'utilisation pour Réservations de capacité sur `use-capacity-reservations-first`. Et si le nombre de Réservations de capacité inutilisés est inférieur à la capacité cible à la demande, la capacité cible à la demande restante est lancée en fonction de la stratégie d'allocation à la demande choisie. Dans cet exemple, la stratégie d'allocation à la demande est `prioritized`.

Dans cet exemple, il y a 15 Réservations de capacité inutilisés disponibles. Cette capacité est inférieure à la capacité cible à la demande de la flotte de 16 Instances à la demande.

Le compte a les 15 Réservations de capacité suivants inutilisés dans 3 groupes différents. Le nombre de Réservations de capacité dans chaque pool est indiqué par `AvailableInstanceCount`.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "c4.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}

{
    "CapacityReservationId": "cr-111",
    "InstanceType": "c3.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}

{
    "CapacityReservationId": "cr-111",
    "InstanceType": "c5.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount":5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}
```

La configuration de flotte suivante affiche uniquement les configurations pertinentes pour cet exemple. La stratégie d'allocation à la demande est `prioritized`, et la stratégie d'utilisation pour Réservations de capacité est `use-capacity-reservations-first`. La capacité cible totale est 16 et le type de capacité cible par défaut est `on-demand`.

> **Note**
>
> Le type de flotte doit être `instant`. Les Réservations de capacité ne sont pas pris en charge pour les autres types de flotte.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-0e8c754449b27161c",
                "Version": "1"
            }
            "Overrides": [
                {
                  "InstanceType": "c4.large",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1,
                  "Priority": 1.0
                },
                {
                  "InstanceType": "c3.large",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1,
                  "Priority": 2.0
                },
                {
                  "InstanceType": "c5.large",
                  "AvailabilityZone": "us-east-1a",
```

```
                  "WeightedCapacity": 1,
                  "Priority": 3.0
                }
              ]

        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 16,
        "DefaultTargetCapacityType": "on-demand"
    },
    "OnDemandOptions": {
        "AllocationStrategy": "prioritized"
        "CapacityReservationOptions": {
            "UsageStrategy": "use-capacity-reservations-first"
        }
    },
    "Type": "instant",
}
```

Après avoir créé la flotte `instant` à l'aide de la configuration précédente, les 16 instances suivantes sont lancées pour atteindre la capacité cible :

- 6 c4.large Instances à la demande dans us-east-1a – c4.large dans us-east-1a est priorisé en premier, et il y a 5 Réservations de capacité c4.large disponibles inutilisés. Les Réservations de capacité sont d'abord utilisés pour lancer 5 Instances à la demande, puis un instance à la demande supplémentaire est lancé selon la stratégie d'allocation à la demande, qui est `prioritized` dans cet exemple.
- 5 c3.large Instances à la demande dans us-east-1a – c3.large dans us-east-1a est priorisé en second, et il y a 5 Réservations de capacité c3.large disponibles inutilisés
- 5 c5.large Instances à la demande dans us-east-1a – c5.large dans us-east-1a est priorisé en troisième, et il y a 5 Réservations de capacité c5.large disponibles inutilisés

Une fois la flotte lancée, vous pouvez exécuter describe-capacity-reservations pour voir combien il reste de Réservations de capacité inutilisés. Dans cet exemple, vous devriez voir la réponse suivante, qui montre que tous les Réservations de capacité de tous les groupes ont été utilisés.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "c4.large",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-222",
    "InstanceType": "c3.large",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-333",
    "InstanceType": "c5.large",
    "AvailableInstanceCount": 0
}
```

Exemple 7 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution `lowest-price`

Vous pouvez configurer une flotte pour qu'elle utilise d'abord Réservations de capacité à la demande lors du lancement d'Instances à la demande en définissant la stratégie d'utilisation pour Réservations de

capacité sur `use-capacity-reservations-first`. Et si plusieurs groupes d'instances n'utilisent pas Réservations de capacité, la stratégie d'allocation à la demande choisie est appliquée. Dans cet exemple, la stratégie d'allocation à la demande est `lowest-price`.

Dans cet exemple, il y a 15 Réservations de capacité inutilisés disponibles. Ce chiffre est supérieur à la capacité à la demande cible de la flotte de 12 Instances à la demande.

Le compte a les 15 Réservations de capacité suivants inutilisés dans 3 groupes différents. Le nombre de Réservations de capacité dans chaque pool est indiqué par `AvailableInstanceCount`.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "m5.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}

{
    "CapacityReservationId": "cr-222",
    "InstanceType": "m4.xlarge",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}

{
    "CapacityReservationId": "cr-333",
    "InstanceType": "m4.2xlarge",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount":5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}
```

La configuration de flotte suivante affiche uniquement les configurations pertinentes pour cet exemple. La stratégie d'allocation à la demande est `lowest-price`, et la stratégie d'utilisation pour Réservations de capacité est `use-capacity-reservations-first`. La capacité totale cible est 12 et le type de capacité cible par défaut est `on-demand`.

Dans cet exemple, le prix des instance à la demande est :

- m5.large – 0,096 $par heure
- m4.xlarge – 0,20 $par heure
- m4.2xlarge – 0,40 $par heure

> **Note**
>
> Le type de flotte doit être `instant`. Les Réservations de capacité ne sont pas pris en charge pour les autres types de flotte.

```
{
    "LaunchTemplateConfigs": [
        {
```

```
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-0e8c754449b27161c",
                "Version": "1"
            }
            "Overrides": [
                {
                  "InstanceType": "m5.large",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1
                },
                {
                  "InstanceType": "m4.xlarge",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1
                },
                {
                  "InstanceType": "m4.2xlarge",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1
                }
              ]

        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 12,
        "DefaultTargetCapacityType": "on-demand"
    },
    "OnDemandOptions": {
        "AllocationStrategy": "lowest-price"
        "CapacityReservationOptions": {
            "UsageStrategy": "use-capacity-reservations-first"
        }
    },
    "Type": "instant",
}
```

Après avoir créé la flotte `instant` à l'aide de la configuration précédente, les 12 instances suivantes sont lancées pour atteindre la capacité cible :

- 5 m5.large Instances à la demande dans us-east-1a – m5.large dans us-east-1a est le prix le plus bas, et il y a 5 Réservations de capacité m5.large disponibles inutilisés
- 5 m4.xlarge Instances à la demande dans us-east-1a – m4.xlarge dans us-east-1a est le prix le plus bas suivant, et il y a 5 Réservations de capacité m4.xlarge disponibles inutilisés
- 2 m4.2xlarge Instances à la demande dans us-est-1a – m4.2xlarge en us-est-1a est le troisième prix le plus bas, et il y a 5 Réservations de capacité m4.2xlarge disponibles inutilisés dont seulement 2 sont nécessaires pour atteindre la capacité cible

Une fois la flotte lancée, vous pouvez exécuter describe-capacity-reservations pour voir combien il reste de Réservations de capacité inutilisés. Dans cet exemple, vous devriez voir la réponse suivante, qui montre que tous les Réservations de capacité m5.large et m4.xlarge ont été utilisés, avec 3 Réservations de capacité m4.2xlarge restants inutilisés.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "m5.large",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-222",
```

```
    "InstanceType": "m4.xlarge",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-333",
    "InstanceType": "m4.2xlarge",
    "AvailableInstanceCount": 3
}
```

Exemple 8 : Lancer Instances à la demande en utilisant Réservations de capacité et la stratégie d'attribution `lowest-price` lorsque la capacité cible totale est supérieure au nombre de Réservations de capacité inutilisés

Vous pouvez configurer une flotte pour qu'elle utilise d'abord Réservations de capacité à la demande lors du lancement d'Instances à la demande en définissant la stratégie d'utilisation pour Réservations de capacité sur `use-capacity-reservations-first`. Et si le nombre de Réservations de capacité inutilisés est inférieur à la capacité cible à la demande, la capacité cible à la demande restante est lancée en fonction de la stratégie d'allocation à la demande choisie. Dans cet exemple, la stratégie d'allocation à la demande est `lowest-price`.

Dans cet exemple, il y a 15 Réservations de capacité inutilisés disponibles. Cette capacité est inférieure à la capacité cible à la demande de la flotte de 16 Instances à la demande.

Le compte a les 15 Réservations de capacité suivants inutilisés dans 3 groupes différents. Le nombre de Réservations de capacité dans chaque pool est indiqué par `AvailableInstanceCount`.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "m5.large",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}

{
    "CapacityReservationId": "cr-222",
    "InstanceType": "m4.xlarge",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount": 5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}

{
    "CapacityReservationId": "cr-333",
    "InstanceType": "m4.2xlarge",
    "InstancePlatform": "Linux/UNIX",
    "AvailabilityZone": "us-east-1a",
    "AvailableInstanceCount":5,
    "InstanceMatchCriteria": "open",
    "State": "active"
}
```

La configuration de flotte suivante affiche uniquement les configurations pertinentes pour cet exemple. La stratégie d'allocation à la demande est `lowest-price`, et la stratégie d'utilisation pour Réservations de capacité est `use-capacity-reservations-first`. La capacité cible totale est 16 et le type de capacité cible par défaut est `on-demand`.

Dans cet exemple, le prix des instance à la demande est :

- m5.large – 0,096 $par heure
- m4.xlarge – 0,20 $par heure
- m4.2xlarge – 0,40 $par heure

> **Note**
>
> Le type de flotte doit être `instant`. Les Réservations de capacité ne sont pas pris en charge pour les autres types de flotte.

```
{
    "LaunchTemplateConfigs": [
        {
            "LaunchTemplateSpecification": {
                "LaunchTemplateId": "lt-0e8c754449b27161c",
                "Version": "1"
            }
            "Overrides": [
                {
                  "InstanceType": "m5.large",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1
                },
                {
                  "InstanceType": "m4.xlarge",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1
                },
                {
                  "InstanceType": "m4.2xlarge",
                  "AvailabilityZone": "us-east-1a",
                  "WeightedCapacity": 1
                }
              ]

        }
    ],
    "TargetCapacitySpecification": {
        "TotalTargetCapacity": 16,
        "DefaultTargetCapacityType": "on-demand"
    },
    "OnDemandOptions": {
        "AllocationStrategy": "lowest-price"
        "CapacityReservationOptions": {
            "UsageStrategy": "use-capacity-reservations-first"
        }
    },
    "Type": "instant",
}
```

Après avoir créé la flotte `instant` à l'aide de la configuration précédente, les 16 instances suivantes sont lancées pour atteindre la capacité cible :

- 6 m5.large Instances à la demande dans us-east-1a – m5.large dans us-east-1a est le prix le plus bas, et il y a 5 Réservations de capacité m5.large disponibles inutilisés. Les Réservations de capacité sont d'abord utilisés pour lancer 5 Instances à la demande, puis un instance à la demande supplémentaire est lancé selon la stratégie d'allocation à la demande, qui est `lowest-price` dans cet exemple.
- 5 m4.xlarge Instances à la demande dans us-east-1a – m4.xlarge dans us-east-1a est le prix le plus bas suivant, et il y a 5 Réservations de capacité m4.xlarge disponibles inutilisés

- 5 m4.2xlarge Instances à la demande dans us-east-1a – m4.2xlarge dans us-east-1a est le troisième prix le plus bas, et il y a 5 Réservations de capacité m4.2xlarge inutilisés disponibles

Une fois la flotte lancée, vous pouvez exécuter describe-capacity-reservations pour voir combien il reste de Réservations de capacité inutilisés. Dans cet exemple, vous devriez voir la réponse suivante, qui montre que tous les Réservations de capacité de tous les groupes ont été utilisés.

```
{
    "CapacityReservationId": "cr-111",
    "InstanceType": "m5.large",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-222",
    "InstanceType": "m4.xlarge",
    "AvailableInstanceCount": 0
}

{
    "CapacityReservationId": "cr-333",
    "InstanceType": "m4.2xlarge",
    "AvailableInstanceCount": 0
}
```

Connectez-vous à votre instance Linux

Connectez-vous aux instances Linux que vous avez lancées et transférez les fichiers de votre ordinateur local à votre instance.

Pour vous connecter à une instance Windows, consultez Connexion à votre instance Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Options de connexion

Le système d'exploitation de votre ordinateur local détermine les options dont vous disposez pour vous connecter à votre instance Linux à partir de votre ordinateur local.

Si le système d'exploitation de votre ordinateur local est Linux ou macOS X

- Client SSH (p. 601)
- EC2 Instance Connect (p. 604)
- Gestionnaire de session AWS Systems Manager

Si votre système d'exploitation de l'ordinateur local est Windows

- PuTTY (p. 614)
- Client SSH (p. 601)
- Gestionnaire de session AWS Systems Manager
- WSL (Windows Subsystem for Linux) (p. 620)

Prérequis généraux pour se connecter à votre instance

Avant de vous connecter à votre instance Linux, vérifiez les prérequis généraux suivants :

- Obtenez des informations sur votre instance (p. 599)
- Autorisez le trafic entrant vers votre instance (p. 600)
- Rechercher la clé privée (p. 600)

- (Facultatif) Obtenez l'empreinte digitale de l'instance (p. 600)

Obtenez des informations sur votre instance

- Obtenir l'ID de l'instance.

  Vous pouvez obtenir l'ID de votre instance à l'aide de la console Amazon EC2 (dans la colonne ID d'instance. Si vous préférez, vous pouvez utiliser la commande describe-instances (AWS CLI) ou Get-EC2Instance (Outils AWS pour Windows PowerShell).
- Obtenir le nom DNS public de l'instance.

  Vous pouvez obtenir le nom DNS public de votre instance à l'aide de la console Amazon EC2. Consultez la colonne DNS Public (IPv4). Si cette colonne est masquée, choisissez l'icône des paramètres

  ( ) dans le coin supérieur droit de l'écran et sélectionnez Public DNS (IPv4) (DNS public (IPv4)). Si vous préférez, vous pouvez utiliser la commande describe-instances (AWS CLI) ou Get-EC2Instance (Outils AWS pour Windows PowerShell).
- (IPv6 uniquement) Obtenir l'adresse IPv6 de l'instance.

  Si vous avez attribué une adresse IPv6 à votre instance, vous pouvez vous connecter à l'instance à l'aide de son adresse IPv6 au lieu d'une adresse IPv4 publique ou d'un nom d'hôte DNS IPv4 public. Votre ordinateur local doit avoir une adresse IPv6 et doit être configuré pour utiliser IPv6. Vous pouvez obtenir l'adresse IPv6 de votre instance à l'aide de la console Amazon EC2. Consultez le champ Adresses IP IPv6. Si vous préférez, vous pouvez utiliser la commande describe-instances (AWS CLI) ou Get-EC2Instance (Outils AWS pour Windows PowerShell). Pour plus d'informations sur IPv6, consultez Adresses IPv6 (p. 810).

- Obtenez le nom d'utilisateur de votre instance.

  Vous pouvez vous connecter à votre instance à l'aide du nom d'utilisateur de votre compte d'utilisateur ou du nom d'utilisateur par défaut de l'AMI que vous avez utilisée pour lancer votre instance.

  - Obtenez le nom d'utilisateur de votre compte d'utilisateur.

    Pour de plus amples informations sur la création d'un compte utilisateur, veuillez consulter Gestion des comptes d'utilisateur sur votre instance Amazon Linux (p. 660).
  - Obtenir le nom d'utilisateur par défaut pour l'AMI que vous avez utilisée pour lancer votre instance
    - Pour Amazon Linux 2 ou l'AMI Amazon Linux, le nom d'utilisateur est `ec2-user`.
    - Pour une AMI CentOS, le nom d'utilisateur est `centos`.
    - Pour une AMI Debian, le nom d'utilisateur est `admin`.
    - Pour une AMI Fedora, le nom d'utilisateur est `ec2-user` ou `fedora`.
    - Pour une AMI RHEL, le nom d'utilisateur est `ec2-user` ou `root`.
    - Pour une AMI SUSE, le nom d'utilisateur est `ec2-user` ou `root`.
    - Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.
    - Si `ec2-user` et `root` ne fonctionnent pas, renseignez-vous auprès de votre fournisseur AMI.

Autorisez le trafic entrant vers votre instance

- Activer le trafic SSH entrant à partir de votre adresse IP vers votre instance

  Vérifiez que le groupe de sécurité associé à votre instance autorise le trafic SSH entrant à partir de votre adresse IP. Le groupe de sécurité par défaut pour le VPC n'autorise pas le trafic SSH entrant par défaut. Le groupe de sécurité créé par l'assistant de lancement d'instance autorise le trafic SSH par défaut. Pour plus d'informations, consultez Autorisation du trafic entrant pour vos instances Linux (p. 1044).

Rechercher la clé privée

- Rechercher la clé privée

  Vous aurez besoin du chemin d'accès qualifié complet à l'emplacement sur votre ordinateur du fichier `.pem` pour la paire de clés que vous avez spécifiée lorsque vous avez lancé l'instance. Pour plus d'informations sur la création de votre paire de clés, consultez Création d'une paire de clés à l'aide d'Amazon EC2.
- Définir les autorisations de votre clé privée

  Si vous envisagez d'utiliser un client SSH sur un ordinateur macOS ou Linux pour vous connecter à votre instance Linux, utilisez la commande suivante pour définir les autorisations de votre fichier de clé privée afin d'être la seule personne autorisée à le lire.

```
chmod 400 my-key-pair.pem
```

  Si vous ne définissez pas ces autorisations, vous ne pouvez pas vous connecter à votre instance à l'aide de cette paire de clés. Pour plus d'informations, consultez Erreur : fichier de clé privée non protégé (p. 1334).

(Facultatif) Obtenez l'empreinte digitale de l'instance

Afin de vous protéger contre les attaques de l'homme au milieu, vous pouvez vérifier l'empreinte digitale de la clé RSA lorsque vous vous connectez à votre instance. La vérification de l'empreinte digitale est utile si vous avez lancé votre instance à partir d'une AMI publique d'un tiers.

Vous obtenez d'abord l'empreinte digitale de l'instance. Ensuite, lorsque vous vous connectez à l'instance, on vous demandera de vérifier l'empreinte. Vous pouvez comparer l'empreinte obtenue avec l'empreinte affichée pour la vérification. Si ces empreintes ne correspondent pas, quelqu'un essaie peut-être d'effectuer une attaque MITM. Si elles correspondent, vous pouvez vous connecter à votre instance en toute confiance.

Prérequis pour l'obtention d'une empreinte digitale d'instance :

- Pour obtenir l'empreinte digitale de l'instance, vous devez utiliser l'AWS CLI. Pour de plus amples informations sur l'installation de l'AWS CLI, veuillez consulter Installation de l'AWS Command Line Interface dans le AWS Command Line Interface Guide de l'utilisateur.
- L'instance ne doit pas être dans l'état `pending`. L'empreinte digitale n'est disponible qu'une fois le premier démarrage de l'instance terminé.

Pour obtenir l'empreinte digitale de l'instance

1. Sur votre ordinateur local (pas sur l'instance), utilisez la commande get-console-output (AWS CLI) comme suit pour obtenir l'empreinte digitale :

```
aws ec2 get-console-output --instance-id instance_id --output text
```

2. Voici un exemple de ce que vous devez rechercher dans la sortie. La sortie exacte peut varier selon le système d'exploitation, la version AMI et si AWS a créé la clé.

```
ec2: #############################################################
ec2: -----BEGIN SSH HOST KEY FINGERPRINTS-----
ec2: 1024 SHA256:7HItIgTONZ/b0CH9c5Dq1ijgqQ6kFn86uQhQ5E/F9pU root@ip-10-0-2-182 (DSA)
ec2: 256 SHA256:l4UB/neBad9tvkgJf1QZWxheQmR59WgrgzEimCG6kZY root@ip-10-0-2-182 (ECDSA)
ec2: 256 SHA256:kpEa+rw/Uq3zxaYZN8KT501iBtJOIdHG52dFi66EEfQ no comment (ED25519)
ec2: 2048 SHA256:L8l6pepcA7iqW/jBecQjVZClUrKY+o2cHLI0iHerbVc root@ip-10-0-2-182 (RSA)
ec2: -----END SSH HOST KEY FINGERPRINTS-----
ec2: #############################################################
```

Connexion à votre instance Linux à l'aide de SSH

Après avoir lancé votre instance, vous pouvez connecter à celle-ci et l'utiliser comme vous le feriez d'un ordinateur devant lequel vous seriez assis.

Les instructions suivantes expliquent comment vous connecter à votre instance à l'aide d'un client SSH. Si vous recevez une erreur lors d'une tentative de connexion à votre instance, consultez Résolution de la connexion à votre instance (p. 1327). Pour plus d'options de connexion, veuillez consulter Connectez-vous à votre instance Linux (p. 598).

Prérequis

Avant de vous connecter à votre instance Linux, remplissez les prérequis suivants.

Vérifiez que l'instance est prête.

Une fois l'instance lancée, il peut falloir quelques minutes pour qu'elle soit prête pour que vous puissiez vous y connecter. Vérifiez que votre instance a réussi les contrôles de statut. Vous pouvez afficher ces informations dans la colonne Vérification de statut de la page Instances.

Vérifiez les prérequis généraux pour la connexion à votre instance.

Pour rechercher le nom DNS public ou l'adresse IP de votre instance et le nom d'utilisateur que vous devez utiliser pour vous connecter à cette dernière, veuillez consulter Prérequis généraux pour se connecter à votre instance (p. 598).

Installez un client SSH sur votre ordinateur local, si besoin.

Un client SSH peut être installé par défaut sur votre ordinateur local. Vous pouvez vérifier cela en tapant ssh sur la ligne de commande. Si votre ordinateur ne reconnaît pas la commande, vous pouvez installer un client SSH.

- Les versions récentes de Windows Server 2019 et Windows 10 - OpenSSH sont incluses en tant que composant installable. Pour plus d'informations, consultez OpenSSH dans Windows.
- Versions antérieures de Windows - Téléchargez et installez OpenSSH. Pour plus d'informations, consultez Win32-OpenSSH.
- Linux et macOS X - Téléchargez et installez OpenSSH. Pour plus d'informations, consultez http://www.openssh.com.

Connexion à votre instance Linux à l'aide d'un client SSH

Utilisez la procédure suivante pour vous connecter à votre instance Linux à l'aide d'un client SSH. Si vous recevez une erreur lors d'une tentative de connexion à votre instance, consultez Résolution de la connexion à votre instance (p. 1327).

Pour vous connecter à votre instance à l'aide de SSH

1. Utilisez la commande ssh dans une fenêtre de terminal pour vous connecter à l'instance. Vous spécifiez le chemin et le nom de fichier de la clé privée (`.pem`), le nom d'utilisateur de votre AMI et le nom DNS public ou l'adresse IPv6 de votre instance. Pour savoir comment trouver la clé privée, le nom d'utilisateur d'une instance et le nom DNS ou l'adresse IPv6 d'une instance, veuillez consulter Rechercher la clé privée (p. 600) et Obtenez des informations sur votre instance (p. 599). Pour vous connecter à votre instance, utilisez l'une des commandes suivantes.
   - (DNS public) Pour vous connecter à l'aide du nom DNS public de votre instance, entrez la commande suivante.

     ```
     ssh -i /path/my-key-pair.pem my-instance-user-name@my-instance-public-dns-name
     ```

   - (IPv6) Sinon, si votre instance possède une adresse IPv6, pour vous connecter en utilisant l'adresse IPv6 de votre instance, entrez la commande suivante.

     ```
     ssh -i /path/my-key-pair.pem my-instance-user-name@my-instance-IPv6-address
     ```

   Vous verrez une réponse telle que celle ci-après:

   ```
   The authenticity of host 'ec2-198-51-100-1.compute-1.amazonaws.com (198-51-100-1)'
    can't be established.
   ECDSA key fingerprint is l4UB/neBad9tvkgJf1QZWxheQmR59WgrgzEimCG6kZY.
   Are you sure you want to continue connecting (yes/no)?
   ```

2. (Facultatif) Vérifiez que l'empreinte de l'alerte de sécurité correspond à l'empreinte que vous avez précédemment obtenue dans (Facultatif) Obtenez l'empreinte digitale de l'instance (p. 600). Si ces empreintes ne correspondent pas, quelqu'un essaie peut-être d'effectuer une attaque MITM. Si elles correspondent, passez à l'étape suivante.
3. Saisissez **`yes`**.

   Vous verrez une réponse telle que celle ci-après:

   ```
   Warning: Permanently added 'ec2-198-51-100-1.compute-1.amazonaws.com' (ECDSA) to the
    list of known hosts.
   ```

Transfert de fichiers des instances Linux à partir de Linux à l'aide de SCP

Le transfert de fichiers entre votre ordinateur local et une instance Linux peut se faire en le protocole de copie sécurisée (SCP). Cette section décrit comment transférer des fichiers avec SCP. La procédure est similaire à celle de la connexion à une instance avec SSH.

Prérequis

- Vérifiez les prérequis généraux pour le transfert de fichiers à votre instance.

  Les prérequis généraux pour le transfert de fichiers vers une instance sont les mêmes que les prérequis généraux pour se connecter à une instance. Pour plus d'informations, consultez Prérequis généraux pour se connecter à votre instance (p. 598).
- Installez un client SCP

  La plupart des ordinateurs Linux, Unix et Apple comporte un client SCP par défaut. Si ce n'est pas le cas pour le vôtre, le projet OpenSSH offre une implémentation gratuite de l'ensemble de la suite d'outils SSH, notamment un client SCP. Pour plus d'informations, consultez http://www.openssh.org.

La procédure suivante vous guide pour le transfert d'un fichier avec SCP. Si vous êtes déjà connecté à l'instance avec SSH et que vous avez vérifié ses empreintes, vous pouvez commencer l'étape qui contient la commande SCP (étape 4).

Pour utiliser SCP pour transférer un fichier

1. Transférez un fichier vers votre instance en utilisant le nom DNS public de l'instance ou l'adresse IPv6 si votre instance en a un. Par exemple, si le nom de votre fichier de clé privée est `my-key-pair`, le fichier à transférer est `SampleFile.txt`, le nom d'utilisateur de votre instance est `my-instance-user-name`, et le nom DNS public de l'instance est `my-instance-public-dns-name`, ou `my-instance-IPv6-address` si votre instance a une adresse IPv6, utilisez l'une des commandes suivantes pour copier le fichier dans le répertoire de base `my-instance-user-name`.
   - (DNS public) Pour transférer un fichier vers votre instance en utilisant le nom DNS public de votre instance, entrez la commande suivante.

     ```
     scp -i /path/my-key-pair.pem /path/SampleFile.txt my-instance-user-name@my-instance-public-dns-name:~
     ```
   - (IPv6) Sinon, si votre instance possède une adresse IPv6, pour vous connecter en utilisant l'adresse IPv6 de votre instance, entrez la commande suivante. L'adresse IPv6 doit être placée entre crochets (`[ ]`]), lesquels doivent être précédés d'un caractère d'échappement (`\`).

     ```
     scp -i /path/my-key-pair.pem /path/SampleFile.txt my-instance-user-name@\[my-instance-IPv6-address\]:~
     ```

   Vous verrez une réponse telle que celle ci-après:

   ```
   The authenticity of host 'ec2-198-51-100-1.compute-1.amazonaws.com (10.254.142.33)'
   can't be established.
   RSA key fingerprint is 1f:51:ae:28:bf:89:e9:d8:1f:25:5d:37:2d:7d:b8:ca:9f:f5:f1:6f.
   Are you sure you want to continue connecting (yes/no)?
   ```
2. (Facultatif) Vérifiez que l'empreinte de l'alerte de sécurité correspond à l'empreinte que vous avez précédemment obtenue dans (Facultatif) Obtenez l'empreinte digitale de l'instance (p. 600). Si ces empreintes ne correspondent pas, quelqu'un essaie peut-être d'effectuer une attaque MITM. Si elles correspondent, passez à l'étape suivante.
3. Saisissez **`yes`**.

Vous verrez une réponse telle que celle ci-après:

```
Warning: Permanently added 'ec2-198-51-100-1.compute-1.amazonaws.com' (RSA)
to the list of known hosts.
Sending file modes: C0644 20 SampleFile.txt
Sink: C0644 20 SampleFile.txt
SampleFile.txt                                100%   20     0.0KB/s   00:00
```

Si l'erreur « bash: scp: command not found » s'affiche, vous devez d'abord installer scp sur votre instance Linux. Pour certains systèmes d'exploitation, elle se trouve dans le package `openssh-clients`. Pour les variantes Amazon Linux comme l'AMI optimisée pour Amazon ECS, utilisez la commande suivante pour installer scp :

```
[ec2-user ~]$ sudo yum install -y openssh-clients
```

4. Pour transférer les fichiers dans l'autre direction (de votre instance Amazon EC2 à votre ordinateur local), inversez l'ordre des paramètres de l'hôte. Par exemple, pour transférer le fichier `SampleFile.txt` depuis votre instance EC2 vers le répertoire de base de votre ordinateur local en tant que `SampleFile2.txt`, utilisez la commande suivante sur votre ordinateur local.
   - (DNS public) Pour transférer un fichier de votre instance en utilisant le nom DNS public de votre instance, entrez la commande suivante.

     ```
     scp -i /path/my-key-pair.pem my-instance-user-name@my-instance-public-dns-name:~/
     SampleFile.txt ~/SampleFile2.txt
     ```

   - (IPv6) Sinon, si votre instance possède une adresse IPv6, pour vous connecter en utilisant l'adresse IPv6 de votre instance, entrez la commande suivante. L'adresse IPv6 doit être placée entre crochets (`[ ]`), lesquels doivent être précédés d'un caractère d'échappement (`\`).

     ```
     scp -i /path/my-key-pair.pem my-instance-user-name@\[my-instance-IPv6-address\]:~/
     SampleFile.txt ~/SampleFile2.txt
     ```

Connexion à votre instance Linux à l'aide de EC2 Instance Connect

Amazon EC2 Instance Connect vous fournit une solution simple et sécurisée pour vous connecter à vos instances à l'aide de Secure Shell (SSH). Avec EC2 Instance Connect, vous utilisez les mandataires et les stratégies AWS Identity and Access Management (IAM) pour contrôler l'accès SSH à vos instances, ce qui vous dispense de devoir gérer et partager les clés SSH. Toutes les demandes de connexion effectuées avec EC2 Instance Connect sont enregistrées sur AWS CloudTrail, de sorte que vous pouvez auditer les demandes de connexion (p. 806).

Vous pouvez utiliser Instance Connect pour vous connecter à vos instances Linux à l'aide d'un client basé sur un navigateur, de l'interface de ligne de commande Amazon EC2 Instance Connect ou du client SSH de votre choix.

Lorsque vous vous connectez à une instance à l'aide d'EC2 Instance Connect, l'API Instance Connect envoie en mode push une clé publique SSH à utilisation unique vers les métadonnées de l'instance (p. 701), où elle demeure pendant 60 secondes. La stratégie IAM attachée à votre utilisateur IAM autorise votre utilisateur IAM à envoyer en mode push la clé publique vers les métadonnées de l'instance. Le démon SSH utilise `AuthorizedKeysCommand` et `AuthorizedKeysCommandUser`, qui sont configurés lors de l'installation d'Instance Connect, pour rechercher la clé publique à partir des métadonnées de l'instance en vue de son authentification, et vous connecte à l'instance.

Conseil

Si vous vous connectez à une instance Linux à partir d'un ordinateur local sous Windows, consultez plutôt la documentation suivante :

- Connexion à votre instance Linux à partir de Windows à l'aide de PuTTY (p. 614)
- Connexion à votre instance Linux à l'aide de SSH (p. 601)
- Connexion de votre instance Linux à partir de Windows à l'aide de Windows Subsystem pour Linux (p. 620)

Sommaire

- Configurer EC2 Instance Connect (p. 605)
- Connexion à l'aide de EC2 Instance Connect (p. 611)
- Désinstallation d'EC2 Instance Connect (p. 613)

Configurer EC2 Instance Connect

Pour utiliser EC2 Instance Connect pour vous connecter à une instance, vous devez configurer chaque instance qui prendra en charge l'utilisation d'Instance Connect (vous n'aurez besoin d'effectuer cette opération qu'une seule fois pour chaque instance) et accorder l'autorisation nécessaire à chaque utilisateur IAM qui utilisera Instance Connect.

Tâches de configuration d'Instance Connect

- Tâche 1 : Configurer l'accès réseau à une instance (p. 606)
- Tâche 2 : (Conditionnel) Installer EC2 Instance Connect sur une instance (p. 606)
- Tâche 3 : (Facultatif) Installer l'interface de ligne de commande EC2 Instance Connect (p. 609)
- Tâche 4 : Configurer les autorisations IAM pour EC2 Instance Connect (p. 609)

Pour de plus amples informations sur la configuration de EC2 Instance Connect, veuillez consulter Sécurisation de vos hôtes bastion avec Amazon EC2 Instance Connect.

Limites

- Les distributions Linux suivantes sont prises en charge :
  - Amazon Linux 2 (toute version)
  - Ubuntu 16.04 ou version ultérieure
- Si vous avez configuré les paramètres `AuthorizedKeysCommand` et `AuthorizedKeysCommandUser` pour l'authentification SSH, l'installation d'EC2 Instance Connect ne les mettra pas à jour. En conséquence, vous ne pourrez pas utiliser Instance Connect.

Prérequis

- Vérifiez les conditions préalables requises pour la connexion à votre instance à l'aide de SSH.

  Pour plus d'informations, consultez Prérequis généraux pour se connecter à votre instance (p. 598).
- Installez un client SSH sur votre ordinateur local.

  Votre ordinateur local comporte probablement un client SSH par défaut. Vous pouvez vérifier un client SSH en tapant ssh dans la ligne de commande. Si votre ordinateur local ne reconnaît pas la commande, vous pouvez installer un client SSH. Pour plus d'informations sur l'installation d'un client SSH sur Linux ou maxOS X, consultez http://www.openssh.com. Pour de plus amples informations sur l'installation d'un client SSH sous Windows 10, veuillez consulter OpenSSH dans Windows.

- Installez AWS CLI sur votre ordinateur local.

  Pour configurer les autorisations IAM, vous devez utiliser AWS CLI. Pour plus d'informations sur l'installation d'AWS CLI, consultez Installation de l'interface de ligne de commande AWS dans le AWS Command Line Interface Guide de l'utilisateur.
- [Ubuntu] Installez AWS CLI sur votre instance.

  Pour installer EC2 Instance Connect sur une instance Ubuntu, vous devez utiliser l'interface de ligne de commande AWS sur l'instance. Pour plus d'informations sur l'installation d'AWS CLI, consultez Installation de l'interface de ligne de commande AWS dans le AWS Command Line Interface Guide de l'utilisateur.

Tâche 1 : Configurer l'accès réseau à une instance

Vous devez configurer l'accès réseau suivant à votre instance de façon à pouvoir installer EC2 Instance Connect et à permettre à vos utilisateurs de se connecter à votre instance :

- Vérifiez que le groupe de sécurité associé à votre instance autorise le trafic SSH entrant (p. 1045) sur le port 22 à partir de votre adresse IP. Le groupe de sécurité par défaut pour le VPC n'autorise pas le trafic SSH entrant par défaut. Le groupe de sécurité créé par l'assistant de lancement autorise le trafic SSH entrant par défaut. Pour plus d'informations, consultez Autorisation du trafic entrant pour vos instances Linux (p. 1044).
- (Client basé sur un navigateur) Il est recommandé que votre instance autorise le trafic SSH entrant à partir du bloc IP recommandé publié pour le service. Utilisez le filtre `EC2_INSTANCE_CONNECT` pour que le paramètre `service` obtienne les plages d'adresse IP du sous-ensemble EC2 Instance Connect. Pour plus d'informations, consultez Plages d'adresses IP AWS dans le Référence générale d'Amazon Web Services.

Tâche 2 : (Conditionnel) Installer EC2 Instance Connect sur une instance

Amazon Linux 2 2.0.20190618 ou version ultérieure et Ubuntu 20.04 ou version ultérieure est préconfiguré avec EC2 Instance Connect. Si vous avez lancé votre instance à l'aide de l'une de ces AMI, vous pouvez ignorer cette tâche. Pour les autres distributions Linux prises en charge, vous devez installer Instance Connect sur chaque instance prenant en charge la connexion à l'aide d'Instance Connect.

L'installation d'Instance Connect configure le démon SSH sur l'instance. La procédure d'installation d'Instance Connect est différente dans le cas des instances lancées avec Amazon Linux 2 et Ubuntu.

Amazon Linux 2

Pour installer EC2 Instance Connect sur une instance lancée avec Amazon Linux 2

1. Connectez-vous à votre instance à l'aide de SSH.

   Utilisez la paire de clés SSH attribuée à votre instance lors de son lancement, ainsi que le nom d'utilisateur par défaut de l'AMI utilisée pour lancer votre instance. Pour Amazon Linux 2, le nom d'utilisateur par défaut est `ec2-user`.

   Par exemple, si votre instance a été lancée avec Amazon Linux 2, que le nom DNS public de votre instance est `ec2-a-b-c-d.us-west-2.compute.amazonaws.com` et que la paire de clés est `my_ec2_private_key.pem`, utilisez la commande suivante pour établir une connexion SSH à votre instance :

```
$ ssh -i my_ec2_private_key.pem ec2-user@ec2-a-b-c-d.us-
west-2.compute.amazonaws.com
```

Pour plus d'informations sur la connexion à votre instance, consultez Connexion à votre instance Linux à l'aide de SSH (p. 601).

2. Installez le package EC2 Instance Connect sur votre instance.

   Pour Amazon Linux 2, utilisez la commande yum install.

```
[ec2-user ~]$ sudo yum install ec2-instance-connect
```

   Quatre nouveaux fichiers doivent apparaître dans le dossier `/opt/aws/bin/` :

```
eic_curl_authorized_keys
eic_harvest_hostkeys
eic_parse_authorized_keys
eic_run_authorized_keys
```

3. (Facultatif) Vérifiez qu'Instance Connect a été installé avec succès sur votre instance.

   Utilisez la commande sudo less pour vérifier que le fichier `/etc/ssh/sshd_config` a été correctement mis à jour comme suit :

```
[ec2-user ~]$ sudo less /etc/ssh/sshd_config
```

   Instance Connect a été correctement installé si les lignes `AuthorizedKeysCommand` et `AuthorizedKeysCommandUser` du fichier `/etc/ssh/sshd_config` contiennent les valeurs suivantes :

```
AuthorizedKeysCommand /opt/aws/bin/eic_run_authorized_keys %u %f
AuthorizedKeysCommandUser ec2-instance-connect
```

   - `AuthorizedKeysCommand` définit le fichier `eic_run_authorized_keys` pour rechercher les clés à partir des métadonnées de l'instance.
   - `AuthorizedKeysCommandUser` définit l'utilisateur système comme `ec2-instance-connect`.

   > **Note**
   >
   > Si vous avez déjà configuré `AuthorizedKeysCommand` et `AuthorizedKeysCommandUser`, l'installation d'Instance Connect ne modifie pas les valeurs et vous ne pouvez pas utiliser Instance Connect.

Ubuntu

Pour installer EC2 Instance Connect sur une instance lancée avec Ubuntu 16.04 ou version ultérieure

1. Connectez-vous à votre instance à l'aide de SSH.

   Utilisez la paire de clés SSH attribuée à votre instance lors de son lancement, ainsi que le nom d'utilisateur par défaut de l'AMI utilisée pour lancer votre instance. Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.

   Si votre instance a été lancée avec Ubuntu, que le nom DNS public de votre instance est `ec2-a-b-c-d.us-west-2.compute.amazonaws.com` et que la paire de clés est

`my_ec2_private_key.pem`, utilisez la commande suivante pour établir une connexion SSH à votre instance :

```
$ ssh -i my_ec2_private_key.pem ubuntu@ec2-a-b-c-d.us-west-2.compute.amazonaws.com
```

Pour plus d'informations sur la connexion à votre instance, consultez Connexion à votre instance Linux à l'aide de SSH (p. 601).

2. (Facultatif) Assurez-vous que votre instance possède l'AMI Ubuntu la plus récente.

   Pour Ubuntu, utilisez les commandes suivantes pour mettre à jour tous les packages de votre instance.

```
ubuntu:~$ sudo apt-get update
```

```
ubuntu:~$ sudo apt-get upgrade
```

3. Installez le package Instance Connect sur votre instance.

   Pour Ubuntu, utilisez la commande sudo apt-get pour installer le package `.deb`.

```
ubuntu:~$ sudo apt-get install ec2-instance-connect
```

   Quatre nouveaux fichiers doivent apparaître dans le dossier `/usr/share/ec2-instance-connect/` :

```
eic_curl_authorized_keys
eic_harvest_hostkeys
eic_parse_authorized_keys
eic_run_authorized_keys
```

4. (Facultatif) Vérifiez qu'Instance Connect a été installé avec succès sur votre instance.

   Utilisez la commande sudo less pour vérifier que le fichier `/lib/systemd/system/ssh.service.d/ec2-instance-connect.conf` a été correctement mis à jour comme suit :

```
ubuntu:~$ sudo less /lib/systemd/system/ssh.service.d/ec2-instance-connect.conf
```

   Instance Connect a été correctement installé si les lignes `AuthorizedKeysCommand` et `AuthorizedKeysCommandUser` du fichier `/lib/systemd/system/ssh.service.d/ec2-instance-connect.conf` contiennent les valeurs suivantes :

```
AuthorizedKeysCommand /usr/share/ec2-instance-connect/eic_run_authorized_keys %u %f
AuthorizedKeysCommandUser ec2-instance-connect
```

   - `AuthorizedKeysCommand` définit le fichier `eic_run_authorized_keys` pour rechercher les clés à partir des métadonnées de l'instance.
   - `AuthorizedKeysCommandUser` définit l'utilisateur système comme `ec2-instance-connect`.

   > **Note**
   >
   > Si vous avez déjà configuré `AuthorizedKeysCommand` et `AuthorizedKeysCommandUser`, l'installation d'Instance Connect ne modifie pas les valeurs et vous ne pouvez pas utiliser Instance Connect.

Pour de plus amples informations sur le package EC2 Instance Connect, consultez aws/aws-ec2-instance-connect-config sur le site web GitHub.

Tâche 3 : (Facultatif) Installer l'interface de ligne de commande EC2 Instance Connect

L'interface de ligne de commande EC2 Instance Connect fournit une interface similaire aux appels SSH standard, ce qui inclut l'interrogation des informations sur l'instance EC2, la génération et la publication de clés publiques éphémères et l'établissement d'une connexion SHH via une seule commande, `mssh instance_id`.

> **Note**
>
> Il n'est pas nécessaire d'installer l'interface de ligne de commande EC2 Instance Connect si les utilisateurs n'utilisent que le client basé sur un navigateur ou un client SSH pour se connecter à une instance.

Pour installer le package de l'interface de ligne de commande EC2 Instance Connect

Utilisez `pip` pour installer le package `ec2instanceconnectcli`. Pour plus d'informations, consultez aws/aws-ec2-instance-connect-cli sur le site web GitHub, et https://pypi.org/project/ec2instanceconnectcli/ sur le site web Python Package Index (PyPI).

```
$ pip install ec2instanceconnectcli
```

Tâche 4 : Configurer les autorisations IAM pour EC2 Instance Connect

Si vos utilisateurs IAM se connectent à une instance avec EC2 Instance Connect, vous devez leur accorder l'autorisation d'envoyer la clé publique en mode push à l'instance. Vous leur accordez l'autorisation en créant une stratégie IAM et en attachant la stratégie aux utilisateurs IAM qui en ont besoin. Pour de plus amples informations, veuillez consulter Actions, ressources et clés de condition pour Amazon EC2 Instance Connect dans le IAM Guide de l'utilisateur.

Les instructions suivantes expliquent comment créer la stratégie et l'attacher à l'aide de l'AWS CLI. Pour obtenir les instructions d'utilisation d'AWS Management Console, consultez Création de stratégies IAM (console) et Ajout d'autorisations par l'attachement direct de stratégies à l'utilisateur dans le IAM Guide de l'utilisateur.

Pour accorder à un utilisateur IAM l'autorisation d'utiliser EC2 Instance Connect (AWS CLI)

1. Créez un document de stratégie JSON incluant le contenu suivant :

   - L'action `ec2-instance-connect:SendSSHPublicKey`. Cette action accorde à un utilisateur IAM l'autorisation d'envoyer la clé publique en mode push à une instance. Avec `ec2-instance-connect:SendSSHPublicKey`, envisagez de limiter l'accès à des instances EC2 spécifiques. Sinon, tous les utilisateurs IAM disposant de cette autorisation peuvent se connecter à toutes les instances EC2.
   - La condition `ec2:osuser`. Elle spécifie le nom de l'utilisateur du système d'exploitation qui peut envoyer la clé publique en mode push à une instance. Utilisez le nom d'utilisateur par défaut pour l'AMI que vous avez utilisée pour lancer l'instance. Le nom d'utilisateur par défaut est `ec2-user` pour Amazon Linux 2 et `ubuntu` pour Ubuntu.
   - L'action `ec2:DescribeInstances`. Cette action est obligatoire lorsque vous utilisez l'interface de ligne de commande EC2 Instance Connect, car l'encapsuleur l'appelle. Les utilisateurs IAM disposent peut-être déjà de l'autorisation d'appeler cette action à partir d'une autre stratégie.

   Voici un exemple de document stratégie : Vous pouvez omettre la déclaration de l'action `ec2:DescribeInstances` si vos utilisateurs utilisent uniquement un client SSH pour se connecter à vos instances. Vous pouvez remplacer les instances spécifiées dans `Resource` par le caractère générique `*` pour accorder aux utilisateurs l'accès à toutes les instances EC2 avec EC2 Instance Connect.

```
{
    "Version": "2012-10-17",
    "Statement": [
      {
        "Effect": "Allow",
        "Action": "ec2-instance-connect:SendSSHPublicKey",
        "Resource": [
            "arn:aws:ec2:region:account-id:instance/i-1234567890abcdef0",
            "arn:aws:ec2:region:account-id:instance/i-0598c7d356eba48d7"
        ],
        "Condition": {
            "StringEquals": {
                "ec2:osuser": "ami-username"
            }
        }
      },
      {
        "Effect": "Allow",
        "Action": "ec2:DescribeInstances",
        "Resource": "*"
      }
    ]
}
```

La stratégie précédente autorise l'accès à des instances spécifiques, identifiées par leur ID d'instance. Vous pouvez également utiliser des balises de ressources pour contrôler l'accès à une instance. Le contrôle d'accès basé sur les attributs permet de définir des autorisations basées sur des balises et pouvant être associées à des utilisateurs et des ressources AWS. Par exemple, la stratégie suivante permet à un utilisateur IAM d'accéder à une instance uniquement si cette instance possède une balise de ressource avec clé=`tag-key` et valeur=`tag-value`. Pour de plus amples informations sur l'utilisation de balises pour contrôler l'accès à vos ressources AWS, veuillez consulter Contrôle de l'accès aux ressources AWS dans le IAM Guide de l'utilisateur.

```
{
   "Version":"2012-10-17",
   "Statement":[
      {
         "Effect":"Allow",
         "Action":"ec2-instance-connect:SendSSHPublicKey",
         "Resource": "arn:aws:ec2:region:account-id:instance/*",
         "Condition":{
            "StringEquals":{
               "aws:ResourceTag/tag-key":"tag-value"
            }
         }
      },
      {
        "Effect": "Allow",
        "Action": "ec2:DescribeInstances",
        "Resource": "*"
      }
   ]
}
```

2. Utilisez la commande create-policy pour créer une nouvelle stratégie gérée et spécifiez le document JSON que vous avez créé pour l'utiliser comme contenu de la nouvelle stratégie.

```
$ aws iam create-policy --policy-name my-policy --policy-document file://JSON-file-name
```

3. Utilisez la commande attach-user-policy pour attacher la stratégie gérée à l'utilisateur IAM spécifié. Pour le paramètre `--user-name`, spécifiez le nom convivial (pas l'ARN) de l'utilisateur IAM.

```
$ aws iam attach-user-policy --policy-arn arn:aws:iam::account-id:policy/my-policy --
user-name IAM-friendly-name
```

Connexion à l'aide de EC2 Instance Connect

Les instructions suivantes expliquent comment vous connecter à votre instance Linux avec EC2 Instance Connect.

Options de connexion à l'aide d'EC2 Instance Connect

- Connexion à l'aide du client basé sur un navigateur (p. 611)
- Connexion à l'aide de l'interface de ligne de commande EC2 Instance Connect (p. 612)
- Connexion à l'aide de votre propre clé et d'un client SSH (p. 612)

Limites

- Les distributions Linux suivantes sont prises en charge :
  - Amazon Linux 2 (toute version)
  - Ubuntu 16.04 ou version ultérieure
- Pour se connecter à l'aide du client basé sur le navigateur, l'instance doit avoir une adresse IPv4 publique.
- Si l'instance ne possède pas d'adresse IP publique, vous ne pouvez vous connecter à l'instance qu'à l'aide de l'interface de ligne de commande EC2 Instance Connect et uniquement à partir d'une machine au sein du même VPC.
- EC2 Instance Connect ne prend pas en charge la connexion à l'aide d'une adresse IPv6.

Prérequis

- Installez Instance Connect sur votre instance.

  Pour plus d'informations, consultez Configurer EC2 Instance Connect (p. 605).
- (Facultatif) Installez un client SSH sur votre ordinateur local.

  Il n'est pas nécessaire d'installer un client SSH si les utilisateurs utilisent uniquement la console ou l'interface de ligne de commande EC2 Instance Connect pour se connecter à une instance. Votre ordinateur local comporte probablement un client SSH par défaut. Vous pouvez vérifier un client SSH en tapant ssh dans la ligne de commande. Si votre ordinateur local ne reconnaît pas la commande, vous pouvez installer un client SSH. Pour plus d'informations sur l'installation d'un client SSH sur Linux ou maxOS X, consultez http://www.openssh.com. Pour de plus amples informations sur l'installation d'un client SSH sous Windows 10, veuillez consulter OpenSSH dans Windows.
- (Facultatif) Installez l'interface de ligne de commande EC2 Instance Connect sur votre ordinateur local.

  Il n'est pas nécessaire d'installer l'interface de ligne de commande EC2 Instance Connect si les utilisateurs n'utilisent que la console ou un client SSH pour se connecter à une instance. Pour plus d'informations, consultez Tâche 3 : (Facultatif) Installer l'interface de ligne de commande EC2 Instance Connect (p. 609).

Connexion à l'aide du client basé sur un navigateur

Vous pouvez vous connecter à une instance à l'aide du client basé sur un navigateur en sélectionnant l'instance à partir de la console Amazon EC2 et en choisissant de vous connecter avec EC2 Instance Connect. Instance Connect gère les autorisations et garantit une connexion réussie.

Pour vous connecter à votre instance à l'aide du client basé sur un navigateur à partir de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances.
3. Sélectionnez l'instance, puis choisissez Connect (Connexion).
4. Choisissez EC2 Instance Connect.
5. Vérifiez le nom d'utilisateur et choisissez Se connecter pour ouvrir une fenêtre de terminal.

Connexion à l'aide de l'interface de ligne de commande EC2 Instance Connect

Vous pouvez vous connecter à une instance avec l'interface de ligne de commande EC2 Instance Connect en fournissant uniquement l'ID d'instance, tandis que l'interface de ligne de commande Instance Connect exécute les trois actions suivantes en un seul appel ; elle génère une clé publique SSH à utilisation unique, envoie en mode push la clé à l'instance où elle demeure pendant 60 secondes, et connecte l'utilisateur à l'instance. Vous pouvez utiliser les commandes SSH/SFTP de base avec l'interface de ligne de commande Instance Connect.

> **Note**
>
> `-i` n' est pas pris en charge lors de l'utilisation de mssh. Lorsque vous utilisez la commande mssh pour vous connecter à votre instance, vous n'avez pas besoin de spécifier un type de fichier d'identité car Instance Connect gère la paire de clés.

Amazon Linux 2

Pour vous connecter à une instance avec l'interface de ligne de commande EC2 Instance Connect

Utilisez la commande mssh avec l'ID d'instance comme suit. Vous n'avez pas besoin de spécifier le nom d'utilisateur de l'AMI.

```
$ mssh i-001234a4bf70dec41EXAMPLE
```

Ubuntu

Pour vous connecter à une instance avec l'interface de ligne de commande EC2 Instance Connect

Utilisez la commande mssh avec l'ID d'instance et le nom d'utilisateur par défaut pour l'AMI Ubuntu comme suit. Vous devez spécifier le nom d'utilisateur de l'AMI ou vous obtiendrez l'erreur suivante : Échec de l'authentification.

```
$ mssh ubuntu@i-001234a4bf70dec41EXAMPLE
```

Connexion à l'aide de votre propre clé et d'un client SSH

Vous pouvez utiliser votre propre clé SSH et vous connecter à votre instance à partir du client SSH de votre choix en utilisant l'interface de ligne de commande EC2 Instance Connect. Cela vous permet de bénéficier de la capacité d'Instance Connect d'envoyer une clé publique en mode push à l'instance.

Exigence

Les types de clé RSA pris en charge sont OpenSSH et SSH2. Les longueurs prises en charge sont 2048 et 4096. Pour plus d'informations, consultez Option 2 : Importez votre propre clé publique dans Amazon EC2 (p. 1049).

Pour vous connecter à votre instance à l'aide de votre propre clé et d'un client SSH

1. (Facultatif) Générer de nouvelles clés SSH publiques et privées

   Vous pouvez générer de nouvelles clés SSH privées et publiques, `my_rsa_key` et `my_rsa_key.pub`, à l'aide de la commande suivante :

   ```
   $ ssh-keygen -t rsa -f my_rsa_key
   ```

2. Envoyer votre clé publique SSH en mode push à l'instance

   Utilisez la commande send-ssh-public-key pour pousser votre clé publique SSH vers l'instance. Si vous avez lancé votre instance avec Amazon Linux 2, le nom d'utilisateur par défaut de l'AMI est `ec2-user`. Si vous avez lancé votre instance avec Ubuntu, le nom d'utilisateur par défaut de l'AMI est `ubuntu`.

   L'exemple suivant pousse la clé publique vers l'instance spécifiée dans la zone de disponibilité spécifiée, afin d'authentifier `ec2-user` :

   ```
   $ aws ec2-instance-connect send-ssh-public-key \
       --instance-id i-001234a4bf70dec41EXAMPLE \
       --availability-zone us-west-2b \
       --instance-os-user ec2-user \
       --ssh-public-key file://my_rsa_key.pub
   ```

3. Connexion à l'instance avec votre clé privée

   Utilisez la commande ssh pour vous connecter à l'instance à l'aide de la clé privée avant que la clé publique ne soit supprimée des métadonnées de l'instance (vous disposez de 60 secondes avant qu'elle ne soit supprimée). Spécifiez la clé privée qui correspond à la clé publique, le nom d'utilisateur par défaut de l'AMI que vous avez utilisé pour lancer votre instance et le DNS public de l'instance. Ajoutez l'option `IdentitiesOnly=yes` pour vous assurer que seuls les fichiers de la configuration ssh et la clé spécifiée sont utilisés pour la connexion.

   ```
   $ ssh -o "IdentitiesOnly=yes" -i my_rsa_key ec2-
   user@ec2-198-51-100-1.compute-1.amazonaws.com
   ```

Désinstallation d'EC2 Instance Connect

Pour désactiver EC2 Instance Connect, connectez-vous à votre instance et désinstallez le package `ec2-instance-connect` que vous avez installé sur le système d'exploitation. Si la configuration `sshd` correspond à ce qui a été défini quand vous avez installé EC2 Instance Connect, la désinstallation du package `ec2-instance-connect` supprime aussi la configuration `sshd`. Si vous avez modifié la configuration `sshd` après l'installation d'EC2 Instance Connect, vous devez la mettre à jour manuellement.

Amazon Linux 2

Vous pouvez désinstaller EC2 Instance Connect sur Amazon Linux 2 version 2.0.20190618 ou ultérieure, où EC2 Instance Connect est préconfiguré.

Pour désinstaller EC2 Instance Connect sur une instance lancée avec Amazon Linux 2

1. Connectez-vous à votre instance à l'aide de SSH. Spécifiez la paire de clés SSH que vous avez utilisée pour votre instance lorsque vous l'avez lancée et le nom d'utilisateur par défaut pour l'AMI Amazon Linux 2, c'est-à-dire `ec2-user`.

Par exemple, la commande ssh suivante vous connecte à l'instance ayant le nom DNS public `ec2-a-b-c-d.us-west-2.compute.amazonaws.com` en utilisant la paire de clés `my_ec2_private_key.pem`.

```
$ ssh -i my_ec2_private_key.pem ec2-user@ec2-a-b-c-d.us-
west-2.compute.amazonaws.com
```

2. Désinstallez le package `ec2-instance-connect` à l'aide de la commande yum.

```
[ec2-user ~]$ sudo yum remove ec2-instance-connect
```

Ubuntu

Pour désinstaller EC2 Instance Connect sur une instance lancée avec une AMI Ubuntu

1. Connectez-vous à votre instance à l'aide de SSH. Spécifiez la paire de clés SSH que vous avez utilisée pour votre instance lorsque vous l'avez lancée et le nom d'utilisateur par défaut pour l'AMI Ubuntu, c'est-à-dire `ubuntu`.

   Par exemple, la commande ssh suivante vous connecte à l'instance ayant le nom DNS public `ec2-a-b-c-d.us-west-2.compute.amazonaws.com` en utilisant la paire de clés `my_ec2_private_key.pem`.

```
$ ssh -i my_ec2_private_key.pem ubuntu@ec2-a-b-c-d.us-west-2.compute.amazonaws.com
```

2. Désinstallez le package `ec2-instance-connect` à l'aide de la commande apt-get.

```
ubuntu:~$ sudo apt-get remove ec2-instance-connect
```

Connexion à votre instance Linux à partir de Windows à l'aide de PuTTY

Après avoir lancé votre instance, vous pouvez connecter à celle-ci et l'utiliser comme vous le feriez d'un ordinateur devant lequel vous seriez assis.

Les instructions suivantes expliquent comment vous connecter à votre instance à l'aide de PuTTY, un client SSH gratuit pour Windows. Si vous recevez une erreur en essayant de vous connecter à votre instance, consultez Résolution de la connexion à votre instance.

Conditions préalables requises

Avant de vous connecter à votre instance Linux à l'aide de PuTTY, remplissez les prérequis suivants.

Vérifiez que l'instance est prête.

Une fois l'instance lancée, il peut falloir quelques minutes pour qu'elle soit prête pour que vous puissiez vous y connecter. Vérifiez que votre instance a réussi les contrôles de statut. Vous pouvez afficher ces informations dans la colonne Vérification de statut de la page Instances.

Vérifiez les prérequis généraux pour la connexion à votre instance.

Pour rechercher le nom DNS public ou l'adresse IP de votre instance et le nom d'utilisateur que vous devez utiliser pour vous connecter à cette dernière, veuillez consulter Prérequis généraux pour se connecter à votre instance (p. 598).

Installez PuTTY sur votre ordinateur local.

Téléchargez et installez PuTTY à partir de la page de téléchargement PuTTY. Si une version antérieure de PuTTY est installée, nous vous recommandons de télécharger la dernière version. Assurez-vous d'installer toute la suite.

Convertir votre clé privée avec PuTTYgen

Recherchez la clé privée (fichier .pem) pour la paire de clés que vous avez spécifiée lorsque vous avez lancé l'instance. Convertissez le fichier .pem en fichier .ppk pour une utilisation avec PuTTY. Pour plus d'informations, suivez les étapes décrites dans la section suivante.

Convertir votre clé privée avec PuTTYgen

PuTTY ne prend pas en charge de manière native le format clé privée pour les clés SSH. PuTTY fournit un outil nommé PuTTYgen qui convertit les clés au format requis pour PuTTY. Vous devez convertir votre clé privée (fichier .pem) dans ce format (fichier .ppk) comme suit pour vous connecter à votre instance avec PuTTY.

Pour convertir votre clé privée

1. Depuis le menu Start (Démarrer), choisissez All Programs (Tous les programmes), PuTTY, PuTTYgen.
2. Sous Type of key to generate (Type de clé à générer), sélectionnez RSA. Si votre version de PuTTYgen n'inclut pas cette option, choisissez SSH-2 RSA.

3. Choisissez Charger. Par défaut, PuTTYgen affiche uniquement les fichiers ayant l'extension `.ppk`. Pour trouver votre fichier `.pem`, choisissez l'option permettant d'afficher tous les types de fichiers.

4. Sélectionnez votre fichier `.pem` pour la paire de clés que vous avez spécifiée lors du lancement de votre instance et choisissez Open (Ouvrir). PuTTYgen affiche un avis indiquant que le fichier `.pem` a été importé avec succès. Choisissez OK.
5. Pour enregistrer la clé dans le format utilisable par PuTTY, choisissez Save private key (Enregistrer la clé privée). PuTTYgen affiche un avertissement à propos de l'enregistrement de la clé sans phrase secrète. Choisissez Oui.

   Note

   La phrase secrète d'une clé privée constitue une couche supplémentaire de protection. Même si votre clé privée est découverte, elle ne peut pas être utilisée sans la phrase secrète. Le désavantage d'une phrase secrète est qu'elle rend l'automatisation plus difficile, car l'intervention humaine est nécessaire pour se connecter à une instance, ou copier des fichiers vers une instance.
6. Spécifiez le même nom pour la clé que celui que vous avez utilisé pour la paire de clés (par exemple, `my-key-pair`) et choisissez Save (Enregistrer). PuTTY ajoute automatiquement l'extension de fichier `.ppk`.

Votre clé privée est désormais dans bon format pour être utilisée avec PuTTY. Vous pouvez désormais vous connecter à votre instance en utilisant le client SSH de PuTTY.

Connexion à votre instance Linux

Utilisez la procédure suivante pour vous connecter à votre instance Linux à l'aide de PuTTY. Vous aurez besoin du fichier `.ppk` que vous avez créé pour votre clé privée. Pour plus d'informations, consultez Convertir votre clé privée avec PuTTYgen (p. 615) dans la section précédente. Si vous recevez une erreur en essayant de vous connecter à votre instance, consultez Résolution de la connexion à votre instance.

Pour vous connecter à votre instance à l'aide de PuTTY

1. Lancez PuTTY (dans le menu Démarrer, choisissez Tous les programmes, PuTTY, PuTTY).
2. Dans le volet Catégorie, Choisissez Session et complétez les champs suivants :

   a. Dans la zone Host Name (Nom d'hôte), effectuez l'une des opérations suivantes :
      - (DNS public) Pour vous connecter à l'aide du nom DNS public de votre instance, entrez *my-instance-user-name*@*my-instance-public-dns-name*.
      - (IPv6) Sinon, si votre instance possède une adresse IPv6, pour vous connecter en utilisant l'adresse IPv6 de votre instance, entrez *my-instance-user-name*@*my-instance-IPv6-address*.

      Pour de plus amples informations sur la façon d'obtenir le nom d'utilisateur de votre instance, ainsi que le nom DNS public ou l'adresse IPv6 de votre instance, veuillez consulter Obtenez des informations sur votre instance (p. 599).

   b. Vérifiez que Port a pour valeur 22.

   c. Sous Type de connexion, sélectionnez SSH.

3. (Facultatif) Vous pouvez configurer PuTTY pour envoyer automatiquement des données « keepalive » à intervalles réguliers afin de garder votre session active. Cela est particulièrement utile et vous évite de vous déconnecter de votre instance en raison de l'inactivité de la session. Dans le volet Catégorie, choisissez Connexion, puis entrez l'intervalle requis dans le champ Seconds between keepalives (Secondes écoulées entre les paquets keepalive). Par exemple, si votre session se déconnecte après

10 minutes d'inactivité, entrez 180 pour configurer PuTTY pour envoyer des données keepalive toutes les 3 minutes.

4. Dans le volet Catégorie, développez Connexion, développez SSH, puis choisissez Auth. Suivez les instructions suivantes :
   a. Choisissez Parcourir.
   b. Sélectionnez le fichier `.ppk` que vous avez généré pour votre paire de clés, puis choisissez Ouvrir.
   c. (Facultatif) Si vous comptez relancer cette session plus tard, vous pouvez enregistrer les informations correspondantes pour les utiliser à l'avenir. Sous Category (Catégorie), choisissez Session, entrez un nom pour la session dans Saved Sessions (Sessions enregistrées), puis choisissez Save (Enregistrer).
   d. Choisissez Open.
5. S'il s'agit de votre première connexion à cette instance, PuTTY affiche une boîte de dialogue d'alerte de sécurité qui vous demande si vous faites confiance à l'hôte auquel vous vous connectez.
   a. (Facultatif) Vérifiez que l'empreinte dans la boîte de dialogue d'alerte de sécurité correspond à l'empreinte que vous avez obtenue précédemment dans (Facultatif) Obtenez l'empreinte digitale de l'instance (p. 600). Si ces empreintes ne correspondent pas, quelqu'un essaie peut-être d'effectuer une attaque MITM. Si elles correspondent, passez à l'étape suivante.
   b. Choisissez Oui. Une fenêtre s'ouvre et vous êtes connecté à votre instance.

      Note

      Si vous avez spécifié une phrase passe lorsque vous avez converti votre clé privée au format PuTTY vous devez fournir cette phrase passe au moment de la connexion à l'instance.

Si vous recevez une erreur en essayant de vous connecter à votre instance, consultez Résolution de la connexion à votre instance.

Transfert de fichiers vers votre instance Linux à l'aide du client Secure Copy PuTTY

Le client Secure Copy PuTTY (PSCP) est un outil de ligne de commande que vous pouvez utiliser pour transférer les fichiers entre votre ordinateur Windows et votre instance Linux. Si vous préférez une interface utilisateur graphique (GUI), vous pouvez utiliser un outil GUI open source nommé WinSCP. Pour plus d'informations, consultez Transfert de fichiers vers votre instance Linux à l'aide de WinSCP (p. 618).

Pour utiliser PSCP, vous aurez besoin de la clé privée que vous avez générée dans Convertir votre clé privée avec PuTTYgen (p. 615). Vous avez également besoin du nom DNS public de votre instance Linux, ou de l'adresse IPv6 si votre instance en a une.

Dans l'exemple suivant, le fichier `Sample_file.txt` est transféré depuis le lecteur C:\ d'un ordinateur Windows vers le répertoire de base `my-instance-user-name` d'une instance Amazon Linux. Pour transférer un fichier, utilisez l'une des commandes suivantes.

- (DNS public) Pour transférer un fichier à l'aide du nom DNS public de votre instance, entrez la commande suivante.

```
pscp -i C:\path\my-key-pair.ppk C:\path\Sample_file.txt my-instance-user-name@my-
instance-public-dns-name:/home/my-instance-user-name/Sample_file.txt
```

- (IPv6) Sinon, si votre instance possède une adresse IPv6, pour vous connecter en utilisant l'adresse IPv6 de votre instance, entrez la commande suivante. L'adresse IPv6 doit être entre crochets (`[ ]`).

```
pscp -i C:\path\my-key-pair.ppk C:\path\Sample_file.txt my-instance-user-name@[my-
instance-IPv6-address]:/home/my-instance-user-name/Sample_file.txt
```

Transfert de fichiers vers votre instance Linux à l'aide de WinSCP

WinSCP est un gestionnaire de fichiers basé sur l'interface utilisateur graphique pour Windows qui vous permet de charger et de transférer des fichiers vers un ordinateur distant à l'aide des protocoles SFTP, SCP, FTP, et FTPS. WinSCP vous donne la possibilité de glisser et déposer les fichiers de votre ordinateur Windows vers votre instance Linux ou de synchroniser des structures entières de répertoires entre les deux systèmes.

Prérequis

- Vous devez disposer de la clé privée que vous avez générée dans Convertir votre clé privée avec PuTTYgen (p. 615).
- Vous avez également besoin du nom DNS public de votre instance Linux.
- Le package `scp` doit être installé sur votre instance Linux. Pour certains systèmes d'exploitation, vous installez le package `openssh-clients`. Pour d'autres, tels que l'AMI optimisé pour Amazon ECS, vous installez le `scp` package. Consultez la documentation de votre distribution Linux.

Pour vous connecter à votre instance à l'aide de WinSCP

1. Téléchargez et installez WinSCP à partir de http://winscp.net/eng/download.php. Pour la plupart des utilisateurs, les options d'installation par défaut sont correctes.
2. Démarrez WinSCP.
3. Dans l'écran Connexion WinSCP pour Nom d'hôte, entrez l'une des options suivantes :
   - (Adresse DNS publique ou IPv4) Pour vous connecter à l'aide du nom DNS public ou de l'adresse IPv4 publique de votre instance, entrez le nom DNS public ou l'adresse IPv4 publique de votre instance.
   - (IPv6) Sinon, si votre instance possède une adresse IPv6, pour vous connecter en utilisant l'adresse IPv6 de votre instance, entrez l'adresse IPv6 de votre instance.
4. Pour User name (Nom d'utilisateur), saisissez le nom utilisateur par défaut pour votre AMI.
   - Pour Amazon Linux 2 ou l'AMI Amazon Linux, le nom d'utilisateur est `ec2-user`.
   - Pour une AMI CentOS, le nom d'utilisateur est `centos`.
   - Pour une AMI Debian, le nom d'utilisateur est `admin`.
   - Pour une AMI Fedora, le nom d'utilisateur est `ec2-user` ou `fedora`.
   - Pour une AMI RHEL, le nom d'utilisateur est `ec2-user` ou `root`.
   - Pour une AMI SUSE, le nom d'utilisateur est `ec2-user` ou `root`.
   - Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.
   - Si `ec2-user` et `root` ne fonctionnent pas, renseignez-vous auprès de votre fournisseur AMI.
5. Spécifiez la clé privée pour votre instance. Pour Private key (Clé privée), entrez le chemin d'accès de votre clé privée, ou cliquez sur le bouton « ... » pour rechercher le fichier. Pour ouvrir les paramètres de site avancés, pour les versions les plus récentes de WinSCP, choisissez Advanced (Avancé). Pour rechercher le paramètre Private key file (Fichier de clé privée), sous SSH, choisissez Authentication (Authentification).

   Voici une capture d'écran de WinSCP version 5.9.4 :

WinSCP a besoin d'un fichier clé privé PuTTY (`.ppk`). Vous pouvez convertir un fichier de clé de sécurité `.pem` au format `.ppk` à l'aide de PuTTYgen. Pour plus d'informations, consultez Convertir votre clé privée avec PuTTYgen (p. 615).

6. (Facultatif) Dans le volet gauche, choisissez Directories (Répertoires). Pour Remote directory (Répertoire distant),entrez le chemin d'accès du répertoire auquel ajouter les fichiers. Pour ouvrir les paramètres de site avancés, pour les versions les plus récentes de WinSCP, choisissez Advanced (Avancé). Pour rechercher le paramètre Remote directory (Répertoire distant), sous Environment (Environnement), choisissez Directories (Répertoires).
7. Choisissez Login (Connexion). Pour ajouter l'empreinte hôte au cache hôte, choisissez Yes (Oui).

8. Une fois que la connexion est établie, dans la fenêtre de connexion, votre instance Linux est à droite et votre machine locale est à gauche. Vous pouvez glisser-déposer des fichiers entre le système de fichiers distant et votre ordinateur local. Pour obtenir plus d'information sur WinSCP, consultez la documentation du projet sur http://winscp.net/eng/docs/start.

   Si vous recevez une erreur « Cannot execute SCP to start transfer (Impossible d'exécuter SCP pour démarrer le transfert) », vérifiez que scp est installé sur l'instance Linux.

Connexion de votre instance Linux à partir de Windows à l'aide de Windows Subsystem pour Linux

Après avoir lancé votre instance, vous pouvez connecter à celle-ci et l'utiliser comme vous le feriez d'un ordinateur devant lequel vous seriez assis.

Les instructions suivantes expliquent comment vous connecter à votre instance à l'aide d'une distribution Linux sur Windows Subsystem pour Linux (WSL). WSL est en téléchargement gratuit et vous permet d'exécuter des outils de ligne de commande directement sous Windows, avec votre bureau Windows traditionnel, sans la surcharge d'une machine virtuelle.

En installant WSL, vous pouvez utiliser un environnement Linux natif pour vous connecter à vos instances Linux EC2 au lieu de faire appel à PuTTY ou PuTTYgen. L'environnement Linux vous permet de vous connecter à vos instances Linux plus facilement, car il comprend un client SSH natif que vous pouvez utiliser pour vous connecter à vos instances Linux et pour modifier les autorisations du fichier de clé .pem. La console Amazon EC2 fournit la commande SSH pour se connecter à l'instance Linux, et vous pouvez

obtenir des informations plus détaillées à partir de la commande SSH pour le dépannage. Pour plus d'informations, consultez la documentation Windows Subsystem pour Linux.

> **Note**
>
> Une fois que vous avez installé WSL, tous les prérequis et les étapes sont les mêmes que ceux décrits dans Connexion à votre instance Linux à l'aide de SSH (p. 601), et vous bénéficiez d'une expérience similaire à l'utilisation de Linux natif.

Si vous recevez une erreur en essayant de vous connecter à votre instance, consultez Résolution de la connexion à votre instance.

Table des matières

- Conditions préalables requises (p. 601)
- Se connecter à votre instance Linux avec WSL (p. 621)
- Transfert de fichiers des instances Linux à partir de Linux à l'aide de SCP (p. 622)
- Désinstallation de WSL (p. 624)

Conditions préalables requises

Avant de vous connecter à votre instance Linux, remplissez les prérequis suivants.

Vérifiez que l'instance est prête.

> Une fois l'instance lancée, il peut falloir quelques minutes pour qu'elle soit prête pour que vous puissiez vous y connecter. Vérifiez que votre instance a réussi les contrôles de statut. Vous pouvez afficher ces informations dans la colonne Vérification de statut de la page Instances.

Vérifiez les prérequis généraux pour la connexion à votre instance.

> Pour rechercher le nom DNS public ou l'adresse IP de votre instance et le nom d'utilisateur que vous devez utiliser pour vous connecter à cette dernière, veuillez consulter Prérequis généraux pour se connecter à votre instance (p. 598).

Installez Windows Subsystem for Linux (WSL) et une distribution Linux sur votre ordinateur local.

> Installez WSL et une distribution Linux à l'aide des instructions du Guide d'installation de Windows 10. L'exemple des instructions installe la distribution Ubuntu de Linux, mais vous pouvez installer n'importe quelle distribution. Vous êtes invité à redémarrer votre ordinateur pour que les modifications prennent effet.

Copiez la clé privée depuis Windows vers WSL

> Dans une fenêtre de terminal WSL, copiez le fichier `.pem` (pour la paire de clés que vous avez spécifiée lorsque vous avez lancé l'instance) de Windows vers WSL. Notez le chemin d'accès qualifié complet vers le fichier `.pem` sur WSL à utiliser lors de la connexion à votre instance. Pour plus d'informations sur la façon de spécifier le chemin vers votre disque dur Windows, consultez How do I access my C drive?. Pour de plus amples informations sur les paires de clés et les instances Windows, veuillez consulter Paires de clés Amazon EC2 et instances Windows.
>
> ```
> cp /mnt/<Windows drive letter>/path/my-key-pair.pem ~/WSL-path/my-key-pair.pem
> ```

Se connecter à votre instance Linux avec WSL

Utilisez la procédure suivante pour vous connecter à votre instance Linux à l'aide de Windows Subsystem pour Linux (WSL). Si vous recevez une erreur en essayant de vous connecter à votre instance, consultez Résolution de la connexion à votre instance.

Pour vous connecter à votre instance à l'aide de SSH

1. Utilisez la commande ssh dans une fenêtre de terminal pour vous connecter à l'instance. Vous spécifiez le chemin et le nom de fichier de la clé privée (`.pem`), le nom d'utilisateur de votre AMI et le nom DNS public ou l'adresse IPv6 de votre instance. Pour savoir comment trouver la clé privée, le nom d'utilisateur d'une instance et le nom DNS ou l'adresse IPv6 d'une instance, veuillez consulter Rechercher la clé privée (p. 600) et Obtenez des informations sur votre instance (p. 599). Pour vous connecter à votre instance, utilisez l'une des commandes suivantes.
   - (DNS public) Pour vous connecter à l'aide du nom DNS public de votre instance, entrez la commande suivante.

     ```
     sudo ssh -i /path/my-key-pair.pem my-instance-user-name@my-instance-public-dns-name
     ```

   - (IPv6) Sinon, si votre instance possède une adresse IPv6, vous pouvez vous connecter à l'instance à l'aide de son adresse IPv6. Spécifiez la commande ssh avec le chemin d'accès au fichier de clé privée (.pem), le nom d'utilisateur approprié et l'adresse IPv6.

     ```
     sudo ssh -i /path/my-key-pair.pem my-instance-user-name@my-instance-IPv6-address
     ```

   Vous verrez une réponse telle que celle ci-après:

   ```
   The authenticity of host 'ec2-198-51-100-1.compute-1.amazonaws.com (10.254.142.33)'
   can't be established.
   RSA key fingerprint is 1f:51:ae:28:bf:89:e9:d8:1f:25:5d:37:2d:7d:b8:ca:9f:f5:f1:6f.
   Are you sure you want to continue connecting (yes/no)?
   ```

2. (Facultatif) Vérifiez que l'empreinte de l'alerte de sécurité correspond à l'empreinte que vous avez précédemment obtenue dans (Facultatif) Obtenez l'empreinte digitale de l'instance (p. 600). Si ces empreintes ne correspondent pas, quelqu'un essaie peut-être d'effectuer une attaque MITM. Si elles correspondent, passez à l'étape suivante.
3. Saisissez `yes`.

   Vous verrez une réponse telle que celle ci-après:

   ```
   Warning: Permanently added 'ec2-198-51-100-1.compute-1.amazonaws.com' (RSA)
   to the list of known hosts.
   ```

Transfert de fichiers des instances Linux à partir de Linux à l'aide de SCP

Le transfert de fichiers entre votre ordinateur local et une instance Linux peut se faire en le protocole de copie sécurisée (SCP). Cette section décrit comment transférer des fichiers avec SCP. La procédure est similaire à celle de la connexion à une instance avec SSH.

Prérequis

- Vérifiez les prérequis généraux pour le transfert de fichiers à votre instance.

  Les prérequis généraux pour le transfert de fichiers vers une instance sont les mêmes que les prérequis généraux pour se connecter à une instance. Pour plus d'informations, consultez Prérequis généraux pour se connecter à votre instance (p. 598).
- Installez un client SCP

  La plupart des ordinateurs Linux, Unix et Apple comporte un client SCP par défaut. Si ce n'est pas le cas pour le vôtre, le projet OpenSSH offre une implémentation gratuite de l'ensemble de la suite d'outils SSH, notamment un client SCP. Pour plus d'informations, consultez http://www.openssh.org.

La procédure suivante vous guide pour le transfert d'un fichier avec SCP. Si vous êtes déjà connecté à l'instance avec SSH et que vous avez vérifié ses empreintes, vous pouvez commencer l'étape qui contient la commande SCP (étape 4).

Pour utiliser SCP pour transférer un fichier

1. Transférez un fichier vers votre instance à l'aide du nom DNS public de l'instance. Par exemple, si le nom du fichier clé privé est `my-key-pair`, le fichier à transférer est `SampleFile.txt`, le nom d'utilisateur est `my-instance-user-name` et le nom DNS public de l'instance est `my-instance-public-dns-name` ou l'adresse IPv6 est `my-instance-IPv6-address`, utilisez la commande suivante pour copier le fichier dans le répertoire de base `my-instance-user-name`.
   - (DNS public) Pour transférer un fichier à l'aide du nom DNS public de votre instance, entrez la commande suivante.

     ```
     scp -i /path/my-key-pair.pem /path/SampleFile.txt my-instance-user-name@my-instance-public-dns-name:~
     ```

   - (IPv6) Sinon, si votre instance possède une adresse IPv6, vous pouvez transférer un fichier à l'aide de l'adresse IPv6 de l'instance. L'adresse IPv6 doit être placée entre crochets (`[ ]`]), lesquels doivent être précédés d'un caractère d'échappement (`\`).

     ```
     scp -i /path/my-key-pair.pem /path/SampleFile.txt my-instance-user-name@\[my-instance-IPv6-address\]:~
     ```

   Vous verrez une réponse telle que celle ci-après:

   ```
   The authenticity of host 'ec2-198-51-100-1.compute-1.amazonaws.com (10.254.142.33)'
   can't be established.
   RSA key fingerprint is 1f:51:ae:28:bf:89:e9:d8:1f:25:5d:37:2d:7d:b8:ca:9f:f5:f1:6f.
   Are you sure you want to continue connecting (yes/no)?
   ```

2. (Facultatif) Vérifiez que l'empreinte de l'alerte de sécurité correspond à l'empreinte que vous avez précédemment obtenue dans (Facultatif) Obtenez l'empreinte digitale de l'instance (p. 600). Si ces empreintes ne correspondent pas, quelqu'un essaie peut-être d'effectuer une attaque MITM. Si elles correspondent, passez à l'étape suivante.
3. Saisissez **`yes`**.

   Vous verrez une réponse telle que celle ci-après:

   ```
   Warning: Permanently added 'ec2-198-51-100-1.compute-1.amazonaws.com' (RSA)
   to the list of known hosts.
   Sending file modes: C0644 20 SampleFile.txt
   Sink: C0644 20 SampleFile.txt
   SampleFile.txt                                100%   20     0.0KB/s   00:00
   ```

   Si l'erreur « bash: scp: command not found » s'affiche, vous devez d'abord installer scp sur votre instance Linux. Pour certains systèmes d'exploitation, elle se trouve dans le package `openssh-clients`. Pour les variantes Amazon Linux comme l'AMI optimisée pour Amazon ECS, utilisez la commande suivante pour installer scp :

   ```
   [ec2-user ~]$ sudo yum install -y openssh-clients
   ```

4. Pour transférer les fichiers dans l'autre direction (de votre instance Amazon EC2 à votre ordinateur local), inversez l'ordre des paramètres de l'hôte. Par exemple, pour retransférer le fichier `SampleFile.txt` depuis votre instance EC2 vers le répertoire de base de votre ordinateur local en tant que `SampleFile2.txt`, utilisez l'une des commandes suivantes sur votre ordinateur local.

- (DNS public) Pour transférer un fichier à l'aide du nom DNS public de votre instance, entrez la commande suivante.

```
scp -i /path/my-key-pair.pem my-instance-user-
name@ec2-198-51-100-1.compute-1.amazonaws.com:~/SampleFile.txt ~/SampleFile2.txt
```

- (IPv6) Sinon, si votre instance possède une adresse IPv6, pour transférer des fichiers dans l'autre sens à l'aide de l'adresse IPv6 de l'instance, entrez la commande suivante.

```
scp -i /path/my-key-pair.pem my-instance-user-name@
\[2001:db8:1234:1a00:9691:9503:25ad:1761\]:~/SampleFile.txt ~/SampleFile2.txt
```

Désinstallation de WSL

Pour plus d'informations sur la façon de désinstaller Windows Subsystem pour Linux, consultez How do I uninstall a WSL Distribution ?.

Connexion à votre instance Linux à l'aide du Gestionnaire de session

Le Gestionnaire de session est une fonctionnalité AWS Systems Manager entièrement gérée qui vous permet de gérer vos instances Amazon EC2 via un shell interactif basé sur un navigateur à un clic ou via l'AWS CLI. Vous pouvez utiliser le Gestionnaire de session pour démarrer une session avec une instance dans votre compte. Une fois la session démarrée, vous pouvez exécuter des commandes bash comme vous le feriez avec n'importe quel autre type de connexion. Pour plus d'informations sur le Gestionnaire de session, consultez Gestionnaire de sessions AWS Systems Manager dans le Guide de l'utilisateur AWS Systems Manager.

Avant d'essayer de vous connecter à une instance à l'aide du Gestionnaire de session, assurez-vous que les étapes d'installation nécessaires sont terminées. Pour plus d'informations et d'instructions, consultez Démarrer avec le Gestionnaire de session.

Pour se connecter à une instance Linux à l'aide du Gestionnaire de session avec la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances.
3. Sélectionnez l'instance, puis choisissez Connect (Connexion).
4. Pour Connection method (Méthode de connexion), choisissez Session Manager (Gestionnaire de session).
5. Choisissez Connect (Se connecter).

Dépannage

Si vous recevez une erreur indiquant que vous n'êtes pas autorisé à effectuer une ou plusieurs actions Systems Manager (`ssm:`*`command-name`*), vous devez mettre à jour vos stratégies qui vous permettront de démarrer des sessions à partir de la console Amazon EC2. Pour plus d'informations, consultez Stratégies IAM par défaut de démarrage rapide pour le Gestionnaire de session dans le Guide de l'utilisateur AWS Systems Manager.

Arrêt et démarrage de votre instance

Vous pouvez arrêter et démarrer votre instance si elle comporte un volume Amazon EBS comme périphérique racine. L'instance conserve son ID d'instance, mais vous pouvez le modifier comme expliqué dans la section Présentation (p. 625).

Lorsque vous arrêtez une instance, nous la fermons. Nous ne vous facturons pas l'utilisation d'une instance arrêtée, ni les frais de transfert de données. Par contre, nous facturons le stockage des volumes Amazon EBS. Chaque fois que vous démarrez une instance arrêtée, nous facturons au minimum une minute pour l'utilisation. Après une minute, seules les secondes que vous utilisez vous sont facturées. Si, par exemple, vous exécutez une instance pendant 20 secondes, puis que vous l'arrêtez, nous vous facturons une minute complète. Si vous exécutez une instance pendant 3 minutes et 40 secondes, nous vous facturons exactement 3 minutes et 40 secondes d'utilisation.

Pendant que l'instance est arrêtée, vous pouvez traiter son volume racine comme tout autre volume et le modifier (par exemple, résoudre des problèmes de système de fichiers ou mettre à jour le logiciel). Il suffit de détacher le volume de l'instance arrêtée, l'attacher à une instance en cours d'exécution, effectuer vos modifications, détacher le volume de l'instance en cours d'exécution, puis le ré-attacher à l'instance arrêtée. Veillez à le ré-attacher à l'aide du nom de périphérique de stockage que vous avez spécifié comme périphérique racine dans le mappage de périphérique de stockage en mode bloc pour l'instance.

Si vous jugez que vous n'avez plus besoin d'une instance, vous pouvez y mettre fin. Dès que l'état d'une instance passe à `shutting-down` ou `terminated`, nous arrêtons de facturer cette instance. Pour plus d'informations, consultez Résilier une instance (p. 645). Si vous préférez mettre l'instance en veille prolongée, consultez Mettre votre instance Linux en veille prolongée (p. 628). Pour plus d'informations, consultez Différences entre redémarrage, arrêt, mise en veille prolongée et résiliation (p. 528).

Sommaire

- Présentation (p. 625)
- Ce qui se passe lorsque vous arrêtez une instance (p. 626)
- Arrêt et démarrage de vos instances (p. 627)
- Modification d'une instance arrêtée (p. 628)
- Dépannage (p. 628)

Présentation

Vous pouvez uniquement arrêter une instance basée sur Amazon EBS. Pour vérifier le type de périphérique racine de votre instance, décrivez l'instance et déterminez si le type de périphérique de son volume racine est `ebs` (instance basée sur Amazon EBS) ou `instance store` (instance basée sur le stockage d'instance). Pour plus d'informations, consultez Sélection du type de périphérique racine de votre AMI (p. 107).

Lorsque vous arrêtez une instance en cours d'exécution, ce qui suit se produit :

- L'instance effectue une fermeture normale et arrête de s'exécuter ; son état passe à `stopping`, puis à `stopped`.
- Les volumes Amazon EBS restent attachées à l'instance et leurs données persistent.
- Les données stockées dans la RAM de l'ordinateur hôte ou les volumes de stockage d'instance de l'ordinateur hôte sont perdues.
- L'instance est généralement migrée vers un nouvel ordinateur hôte sous-jacent lorsqu'elle est lancée (même si, dans certains cas, elle reste sur l'hôte actuel).
- L'instance conserve ses adresses IPv4 privées et les éventuelles adresses IPv6 lorsqu'elle est arrêtée et démarrée. Nous libérons l'adresse IPv4 publique et nous en affectons une nouvelle quand vous la démarrez.

- L'instance conserve les adresses IP Elastic qui lui sont associées. Les adresses IP Elastic qui sont associées à une instance arrêtée vous sont facturées. Avec EC2-Classic, une adresse IP Elastic est dissociée de votre instance lorsque vous l'arrêtez. Pour plus d'informations, consultez EC2-Classic (p. 941).
- Lorsque vous arrêtez et démarrez une instance Windows, le service EC2Config exécute des tâches sur l'instance, telles que la modification des lettres de lecteur pour les volumes Amazon EBS attachés. Pour de plus amples informations sur ces comportements par défaut et pour savoir comment les modifier, veuillez consulter Configuration d'une instance Windows à l'aide du service EC2Config dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.
- Si votre instance est dans un groupe Auto Scaling, le service Amazon EC2 Auto Scaling marque l'instance arrêtée comme étant défectueuse, et peut y mettre fin et lancer une instance de remplacement. Pour plus d'informations, consultez Vérification de l'état des instances Auto Scaling dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.
- Lorsque vous arrêtez une instance ClassicLink, celle-ci est détachée du VPC auquel elle était liée. Vous devez lier à nouveau l'instance au VPC après l'avoir démarrée. Pour plus d'informations sur ClassicLink, consultez ClassicLink (p. 949).

Pour plus d'informations, consultez Différences entre redémarrage, arrêt, mise en veille prolongée et résiliation (p. 528).

Vous pouvez modifier les attributs suivants d'une instance uniquement quand celle-ci est arrêtée :

- Type d'instance
- Données utilisateur
- Kernel
- Disque RAM

Si vous essayez de modifier ces attributs lorsque l'instance est en cours d'exécution, Amazon EC2 renvoie le code `IncorrectInstanceState`.

Ce qui se passe lorsque vous arrêtez une instance

Lorsqu'une instance EC2 est arrêtée à l'aide de la commande `stop-instances`, les éléments suivants sont enregistrés au niveau du système d'exploitation :

- La demande d'API envoie un événement d'appui sur un bouton à l'invité.
- Divers services système sont arrêtés à la suite de l'événement d'appui sur le bouton. L'arrêt normal est déclenché par l'événement d'appui sur un bouton d'arrêt ACPI à partir de l'hyperviseur.
- L'arrêt ACPI est lancé.
- L'instance s'arrête lorsque le processus d'arrêt normal se termine. L'heure d'arrêt du système d'exploitation n'est pas configurable.
- Si le système d'exploitation d'instance ne s'arrête pas proprement en quelques minutes, un arrêt dur est effectué.

Par défaut, lorsque vous initiez une fermeture à partir d'une instance basée sur Amazon EBS (par exemple, à l'aide de la commande shutdown ou poweroff), l'instance s'arrête. Vous pouvez modifier ce comportement pour que l'instance prenne fin. Pour de plus amples informations, veuillez consulter Modification du comportement d'arrêt lancé de l'instance (p. 649).

L'utilisation de la commande halt d'une instance ne déclenche pas un arrêt. Si elle est utilisée, l'instance n'est pas résiliée. Au lieu de cela, elle place l'UC à l'état `HLT` et l'instance continue de s'exécuter.

Arrêt et démarrage de vos instances

Vous pouvez arrêter et démarrer votre instance basée sur Amazon EBS à l'aide de la console ou de la ligne de commande.

Nouvelle console

Pour arrêter et démarrer une instance basée sur Amazon EBS à l'aide de la console

1. Lorsque vous arrêtez une instance, les données contenues sur les volumes de stockage d'instance sont effacées. Avant d'arrêter une instance, vérifiez que vous avez copié toutes les données dont vous avez besoin à partir de vos volumes de stockage d'instance vers un stockage persistant, tel que Amazon EBS ou Amazon S3.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez l'instance.
3. Choisissez État de l'instance, Arrêter l'instance. Si cette option est désactivée, l'instance est déjà arrêtée ou son périphérique racine est un volume de stockage d'instance.
4. Lorsque vous êtes invité à confirmer l'opération, choisissez Arrêter. L'arrêt de l'instance peut prendre quelques minutes.
5. (Facultatif) Pendant que l'instance est en cours d'arrêt, vous pouvez modifier certains de ses attributs. Pour plus d'informations, consultez Modification d'une instance arrêtée (p. 628).
6. Pour démarrer l'instance arrêtée, sélectionnez l'instance et choisissez État de l'instance, Démarrer l'instance.
7. Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`.

Ancienne console

Pour arrêter et démarrer une instance basée sur Amazon EBS à l'aide de la console

1. Lorsque vous arrêtez une instance, les données contenues sur les volumes de stockage d'instance sont effacées. Avant d'arrêter une instance, vérifiez que vous avez copié toutes les données dont vous avez besoin à partir de vos volumes de stockage d'instance vers un stockage persistant, tel que Amazon EBS ou Amazon S3.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez l'instance.
3. Choisissez Actions, Instance State, Stop. Si cette option est désactivée, l'instance est déjà arrêtée ou son périphérique racine est un volume de stockage d'instance.
4. Lorsque vous êtes invité à confirmer, choisissez Oui, arrêter. L'arrêt de l'instance peut prendre quelques minutes.
5. (Facultatif) Pendant que l'instance est en cours d'arrêt, vous pouvez modifier certains de ses attributs. Pour plus d'informations, consultez Modification d'une instance arrêtée (p. 628).
6. Pour démarrer l'instance arrêtée, sélectionnez l'instance et choisissez Actions, Instance State (État de l'instance) et Start (Démarrer).
7. Dans la boîte de dialogue de confirmation, sélectionnez Oui, démarrer. Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`.

Pour arrêter et démarrer une instance basée sur Amazon EBS à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- stop-instances et start-instances (AWS CLI)
- Stop-EC2Instance et Start-EC2Instance (Outils AWS pour Windows PowerShell)

Modification d'une instance arrêtée

Vous pouvez modifier le type d'instance, les données utilisateur et les attributs d'optimisation EBS d'une instance arrêtée à l'aide d'AWS Management Console ou de l'interface ligne de commande. Vous ne pouvez pas utiliser AWS Management Console pour modifier les attributs `DeleteOnTermination`, de noyau ou de disque RAM.

Pour modifier un attribut d'instance

- Pour changer le type d'instance, consultez Modification du type d'instance (p. 313).
- Pour modifier les données utilisateur de votre instance, consultez Utilisation des données utilisateur d'instance (p. 716).
- Pour activer ou désactiver l'optimisation EBS de votre instance, consultez Modification de l'optimisation EBS (p. 1229).
- Pour modifier l'attribut `DeleteOnTermination` du volume racine de votre instance, consultez Mise à jour du mappage de périphérique de stockage en mode bloc d'une instance en cours d'exécution (p. 1297). Vous n'êtes pas obligé d'arrêter l'instance pour modifier cet attribut.

Pour modifier un attribut d'instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur l'interface ligne de commande, consultez Accès à Amazon EC2 (p. 3).

- modify-instance-attribute (AWS CLI)
- Edit-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

Dépannage

Si vous avez arrêté votre instance basée sur Amazon EBS et que celle-ci semble « bloquée » à l'état `stopping`, vous pouvez forcer son arrêt. Pour de plus amples informations, veuillez consulter Résolution des problèmes d'arrêt de votre instance (p. 1335).

Mettre votre instance Linux en veille prolongée

Lorsque vous mettez une instance en veille prolongée, Amazon EC2 indique au système d'exploitation de procéder à la mise en veille prolongée (suspend-to-disk). La mise en veille prolongée enregistre le contenu de la mémoire d'instance (RAM) sur votre volume racine Amazon Elastic Block Store (Amazon EBS). Amazon EC2 conserve le volume racine EBS de l'instance et les volumes de données EBS attachés. Lorsque vous démarrez votre instance :

- Le volume racine EBS est restauré à l'état précédent.
- Le contenu de la mémoire RAM est chargé à nouveau.
- Les processus qui s'exécutaient précédemment sur l'instance reprennent.
- Les volumes de données précédemment attachés sont attachés à nouveau et l'instance conserve son ID d'instance.

Vous pouvez mettre une instance en veille prolongée uniquement si celle-ci est activée pour la mise en veille prolongée (p. 634) et si elle répond aux prérequis de mise en veille prolongée (p. 630).

Si les actions d'amorçage d'une instance et de création d'une empreinte mémoire afin de devenir complètement productive prennent du temps, vous pouvez utiliser la mise en veille prolongée pour préchauffer l'instance. Pour préchauffer l'instance, vous :

1. La lancez avec la mise en veille prolongée activée.
2. La placez dans l'état souhaité.
3. La mettez en veille prolongée, prête à reprendre dans le même état selon les besoins.

L'utilisation d'une instance mise en veille prolongée lorsqu'elle est à l'état `stopped` n'entraîne pas de facturation. Vous recevez une facture correspondant à l'utilisation de l'instance pendant que cette dernière est à l'état `stopping`, quand le contenu de la mémoire RAM est transféré vers le volume racine EBS. (Ce n'est pas comme lorsque vous arrêtez une instance (p. 625) sans la mettre en veille prolongée.) Le transfert de données n'est pas facturé. Cependant, vous recevez une facture correspondant au stockage de tout volume EBS, y compris le stockage du contenu de la mémoire RAM.

Si vous n'avez plus besoin d'une instance, vous pouvez la résilier à tout moment, y compris quand elle est à un état `stopped` (en veille prolongée). Pour plus d'informations, consultez Résilier une instance (p. 645).

> **Note**
>
> Pour de plus amples informations sur l'utilisation de la mise en veille prolongée sur des instances Windows, veuillez consulter Mettre votre instance Windows en veille prolongée dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Sommaire

- Présentation de la mise en veille prolongée (p. 629)
- Prérequis de la mise en veille prolongée (p. 630)
- Limites (p. 631)
- Configuration d'une AMI existante pour prendre en charge la mise en veille prolongée (p. 632)
- Activation de la mise en veille prolongée pour une instance (p. 634)
- Désactivation de KASLR sur une instance (Ubuntu uniquement) (p. 637)
- Mise en veille prolongée d'une instance (p. 638)
- Démarrage d'une instance mise en veille prolongée (p. 640)
- Résolution des problèmes de mise en veille prolongée (p. 641)

Présentation de la mise en veille prolongée

Le diagramme suivant fournit une présentation de base du processus de mise en veille prolongée.

Lorsque vous mettez en veille prolongée une instance en cours d'exécution, voici ce qui se produit :

- Lorsque vous lancez la mise en veille prolongée, l'instance passe à l'état `stopping`. Amazon EC2 indique au système d'exploitation de procéder à la mise en veille prolongée (suspend-to-disk). La mise en veille prolongée fige tous les processus, enregistre le contenu de la mémoire RAM sur le volume racine EBS, puis exécute un arrêt normal.
- Lorsque l'arrêt est terminé, l'instance passe à l'état `stopped`.

- Les volumes EBS restent attachés à l'instance et leurs données persistent, y compris le contenu enregistré de la mémoire RAM.
- Tous les volumes de stockage d'instance Amazon EC2 restent attachés à l'instance, mais les données des volumes de stockage d'instance sont perdues.
- L'instance est généralement migrée vers un nouvel ordinateur hôte sous-jacent lorsqu'elle démarre. Cela se produit également lorsque vous arrêtez et démarrez une instance.
- Lorsque vous démarrez l'instance, celle-ci démarre et le système d'exploitation lit le contenu de la mémoire RAM depuis le volume racine EBS avant de « défiger » les processus pour qu'ils reprennent leur état.
- L'instance conserve ses adresses IPv4 privées, ainsi que les adresses IPv6. Lorsque vous démarrez l'instance, elle continue de conserver ses adresses IPv4 privées et toutes les adresses IPv6.
- Amazon EC2 publie l'adresse IPv4 publique. Lorsque vous démarrez l'instance, Amazon EC2 attribue une nouvelle adresse IPv4 publique à l'instance.
- L'instance conserve les adresses IP Elastic qui lui sont associées. Les adresses IP Elastic qui sont associées à une instance mise en veille prolongée vous sont facturées. Avec EC2-Classic, une adresse IP Elastic est dissociée de votre instance lorsque vous la mettez en veille prolongée. Pour plus d'informations, consultez EC2-Classic (p. 941).
- Lorsque vous mettez une instance ClassicLink en veille prolongée, celle-ci est détachée du VPC auquel elle était liée. Vous devez lier à nouveau l'instance au VPC après l'avoir démarrée. Pour plus d'informations, consultez ClassicLink (p. 949).

Pour plus d'informations sur les différences entre la mise en veille prolongée, et le redémarrage, l'arrêt et la résiliation, consultez Différences entre redémarrage, arrêt, mise en veille prolongée et résiliation (p. 528).

Prérequis de la mise en veille prolongée

Pour qu'une instance puisse être mise en veille prolongée, les prérequis suivants doivent être réunis :

- Familles d'instances prises en charge - C3, C4, C5, I3, M3, M4, M5, M5a, M5ad, R3, R4, R5, R5a, R5ad et T2.
- Taille de la mémoire RAM de l'instance - doit être inférieure à 150 Go.
- Taille de l'instance - non prise en charge pour les instances nues.
- AMI prises en charge (doit être une AMI HVM prenant en charge la mise en veille prolongée) :
  - AMI Amazon Linux 2 publiée le 29 août 2019 ou version ultérieure
  - AMI Amazon Linux de mars 2018 publiée le 16 novembre 2018 ou version ultérieure
  - AMI Ubuntu 18.04 LTS - Bionic publiée avec le numéro de série 20190722.1 ou version ultérieure.*
  - Ubuntu 16.04 LTS - Xenial AMI.* (Configuration supplémentaire est requis.)

    * Nous recommandons de désactiver KASLR sur les instances avec Ubuntu 18.04 LTS - Bionic et Ubuntu 16.04 LTS - Xenial. Pour plus d'informations, consultez Désactivation de KASLR sur une instance (Ubuntu uniquement) (p. 637).

  Pour configurer votre propre AMI afin de prendre en charge la mise en veille prolongée, consultez Configuration d'une AMI existante pour prendre en charge la mise en veille prolongée (p. 632).

  La prise en charge d'autres versions d'Ubuntu et d'autres systèmes d'exploitation sera bientôt disponible.

  Pour de plus amples informations sur les AMI prises en charge pour Windows, veuillez consulter Conditions préalables à la mise en veille prolongée dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.
- Type de volume racine - doit être un volume EBS, et non un volume de stockage d'instance.
- Types de volume EBS pris en charge : SSD à usage général (`gp2`) ou SSD IOPS provisionnés (`io1` ou `io2`). Si vous choisissez un type de volume SSD IOPS provisionnés (`io1` ou `io2`), pour obtenir des

performances optimales pour la mise en veille prolongée, vous devez provisionner le volume EBS avec les IOPS appropriées. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).

- Taille du volume racine EBS : Doit être suffisamment grand pour stocker le contenu de la mémoire RAM et prendre en compte l'utilisation que vous prévoyez, par exemple, le système d'exploitation ou des applications. Si vous activez la mise en veille prolongée, un espace est alloué sur le volume racine au lancement pour stocker la mémoire RAM.
- Chiffrement du volume racine EBS : Pour que vous puissiez utiliser la mise en veille prolongée, le volume racine doit être chiffré afin d'assurer la protection du contenu sensible qui se trouve en mémoire au moment de la mise en veille prolongée. Lorsque les données de la mémoire RAM sont transférées vers le volume racine EBS, celui-ci est toujours chiffré. Le chiffrement du volume racine est appliqué au lancement de l'instance. L'une des trois options suivantes permet de s'assurer que le volume racine est une volume EBS chiffré :
  - Chiffrement EBS en une étape : vous pouvez lancer des instances EC2 soutenues par EBS depuis une AMI chiffrée et activer l'hibernation en même temps. Pour plus d'informations, consultez Utiliser le chiffrement avec des AMI basées sur EBS (p. 168).
  - Chiffrement EBS par défaut : vous pouvez activer le chiffrement EBS par défaut afin de vous assurer que tous les nouveaux volumes EBS de votre compte AWS sont chiffrés. De cette façon, vous pouvez activer l'hibernation pour vos instances sans spécifier d'intention de chiffrement au moment du lancement de l'instance. Pour de plus amples informations, veuillez consulter Chiffrement par défaut (p. 1196).
  - AMI chiffrée : vous pouvez activer le chiffrement EBS en utilisant une AMI chiffrée pour lancer votre instance. Si votre AMI ne dispose d'aucun volume racine chiffré, vous pouvez le copier sur le nouvel AMI et demander son chiffrement. Pour de plus amples informations, veuillez consulter Chiffrement d'une image non chiffrée pendant la copie (p. 172) et Copie d'une AMI (p. 177).
- Activer la mise en veille prolongée au lancement - Vous ne pouvez pas activer la mise en veille prolongée sur une instance existante (en cours d'exécution ou arrêtée). Pour de plus amples informations, veuillez consulter Activation de la mise en veille prolongée pour une instance (p. 634).
- Options d'achat - Cette fonction est disponible pour les Instances à la demande et les Instances réservées. Elle n'est pas disponible pour les Instances Spot. Pour de plus amples informations, veuillez consulter Mise en veille prolongée des Instances Spot interrompues (p. 454).

Limites

- Les actions suivantes ne sont pas prises en charge pour la mise en veille prolongée :
  - La modification du type ou de la taille d'instance d'une instance mise en veille prolongée
  - La création d'instantanés ou d'AMI à partir d'instances pour lesquelles la mise en veille prolongée est activée
  - La création d'instantanés ou d'AMI à partir d'instances mises en veille prolongée
- Lorsque vous mettez en veille une instance, les données contenues sur les volumes de stockage d'instance sont perdues.
- Vous ne pouvez pas mettre en veille prolongée une instance qui dispose de moins de 150 Go de mémoire RAM.
- Vous ne pouvez pas mettre en veille prolongée une instance faisant partie d'un groupe Auto Scaling ou utilisée par Amazon ECS. Si votre instance est dans un groupe Auto Scaling et que vous essayez de la mettre en veille prolongée, le service Amazon EC2 Auto Scaling marque l'instance arrêtée comme étant non saine, et peut la résilier et lancer une instance de remplacement. Pour de plus amples informations, veuillez consulter Vérification de l'état des instances Auto Scaling dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.
- Nous ne prenons pas en charge la conservation d'une instance mise en veille prolongée au-delà de 60 jours. Pour conserver l'instance mise en veille prolongée au-delà de 60 jours, vous devez la démarrer, l'arrêter, puis la démarrer.

- Nous mettons à jour en permanence notre plateforme avec des mises à niveau et des correctifs de sécurité qui peuvent être en conflit avec des instances mises en veille prolongée existantes. Nous vous avertissons des mises à niveau critiques qui nécessitent un démarrage des instances mises en veille prolongée pour que nous puissions effectuer un arrêt ou un redémarrage afin d'appliquer les mises à niveau et les correctifs de sécurité requis.

Configuration d'une AMI existante pour prendre en charge la mise en veille prolongée

Pour mettre en veille prolongée une instance qui avait été lancée à l'aide de votre propre AMI, vous devez d'abord configurer votre AMI pour prendre en charge la mise en veille prolongée. Pour de plus amples informations, veuillez consulter Mise à jour du logiciel d'instance sur votre instance Amazon Linux (p. 655).

Si vous utilisez l'une des AMI prises en charge (p. 630) (à l'exception d'Ubuntu 16.04 LTS), ou si vous créez une AMI basée sur l'une des AMI prises en charge, vous n'avez pas besoin de configurer celle-ci pour prendre en charge la mise en veille prolongée. Ces AMI sont préconfigurées pour prendre en charge la mise en veille prolongée. Pour configurer Ubuntu 16.04 LTS pour la prise en charge la mise en veille prolongée, vous devez installer le paquet du noyau linux-aws-hwe version 4.15.0-1058-aws ou ultérieure et ec2-hibinit-agent. Pour les étapes de configuration, choisissez l'onglet Ubuntu 16.04 - Xenial ci-dessous.

Amazon Linux 2

Pour configurer une AMI Amazon Linux 2 afin de prendre en charge la mise en veille prolongée

1. Mettez à jour le dernier noyau vers la version 4.14.138-114.102 ou version ultérieure à l'aide de la commande suivante :

```
[ec2-user ~]$ sudo yum update kernel
```

2. Installez le package `ec2-hibinit-agent` à partir des référentiels à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo yum install ec2-hibinit-agent
```

3. Redémarrez l'instance à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo reboot
```

4. Vérifiez que la version du noyau a été mise à jour vers la version 4.14.138-114.102 ou version ultérieure à l'aide de la commande suivante.

```
[ec2-user ~]$ uname -a
```

5. Arrêtez l'instance et créez une AMI. Pour de plus amples informations, veuillez consulter Création d'une AMI Linux à partir d'une instance (p. 133).

Amazon Linux

Pour configurer une AMI Amazon Linux afin de prendre en charge la mise en veille prolongée

1. Mettez à jour le dernier noyau vers la version 4.14.77-70.59 ou version ultérieure à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo yum update kernel
```

2. Installez le package `ec2-hibinit-agent` à partir des référentiels à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo yum install ec2-hibinit-agent
```

3. Redémarrez l'instance à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo reboot
```

4. Vérifiez que la version du noyau est mise à jour vers la version 4.14.77-70.59 ou version ultérieure à l'aide de la commande suivante.

```
[ec2-user ~]$ uname -a
```

5. Arrêtez l'instance et créez une AMI. Pour de plus amples informations, veuillez consulter Création d'une AMI Linux à partir d'une instance (p. 133).

Ubuntu 18.04 - Bionic

Pour configurer une AMI Ubuntu 18.04 LTS afin de prendre en charge la mise en veille prolongée

1. Mettez à jour le dernier noyau vers la version 4.15.0-1044 ou version ultérieure à l'aide des commandes suivantes.

```
[ec2-user ~]$ sudo apt update
[ec2-user ~]$ sudo apt dist-upgrade
```

2. Installez le package `ec2-hibinit-agent` à partir des référentiels à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo apt install ec2-hibinit-agent
```

3. Redémarrez l'instance à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo reboot
```

4. Vérifiez que la version du noyau a été mise à jour vers la version 4.15.0-1044 ou version ultérieure à l'aide de la commande suivante.

```
[ec2-user ~]$ uname -a
```

Ubuntu 16.04 - Xenial

Pour configurer une AMI Ubuntu 16.04 LTS afin de prendre en charge la mise en veille prolongée

1. Mettez à jour le dernier noyau vers la version 4.15.0-1058-aws ou version ultérieure à l'aide des commandes suivantes.

```
[ec2-user ~]$ sudo apt update
[ec2-user ~]$ sudo apt install linux-aws-hwe
```

Note

Le paquet du noyau linux-aws-hwe est entièrement pris en charge par Canonical. Le package continuera à recevoir des mises à jour régulières jusqu'à ce que la prise en charge standard d'Ubuntu 16.04 LTS se termine en avril 2021, et recevra des mises à jour de sécurité supplémentaires jusqu'à ce que la prise en charge de la maintenance de sécurité étendue prenne fin en 2024. Pour plus d'informations, consultez Amazon EC2 Hibernation for Ubuntu 16.04 LTS now available sur le blog Canonical Ubuntu.

2. Installez le package `ec2-hibinit-agent` à partir des référentiels à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo apt install ec2-hibinit-agent
```

3. Redémarrez l'instance à l'aide de la commande suivante.

```
[ec2-user ~]$ sudo reboot
```

4. Vérifiez que la version du noyau a été mise à jour vers la version 4.15.0-1058-aws ou version ultérieure à l'aide de la commande suivante.

```
[ec2-user ~]$ uname -a
```

Activation de la mise en veille prolongée pour une instance

Pour que vous puissiez mettre une instance en veille prolongée, celle-ci doit d'abord être activée pour la mise en veille prolongée. L'activation de la mise en veille prolongée doit être effectuée lors du lancement de l'instance.

Important

Vous ne pouvez pas activer ou désactiver la mise en veille prolongée pour une instance après son lancement.

Console

Pour activer la mise en veille prolongée à l'aide de la console

1. Suivez la procédure Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).
2. Sur la page Sélection d'une Amazon Machine Image (AMI), sélectionnez une AMI qui prend en charge la mise en veille prolongée. Pour plus d'informations sur les AMI supportées, consultez Prérequis de la mise en veille prolongée (p. 630).
3. Sur la page Choisir un type d'instance, sélectionnez un type d'instance pris en charge, puis choisissez Suivant : Configurer les détails de l'instance. Pour plus d'informations sur les types d'instance pris en charge, consultez Prérequis de la mise en veille prolongée (p. 630).
4. Sur la page Configurer les détails de l'instance, pour Stop - Hibernate Behavior (Arrêter - Comportement de mise en veille prolongée), cochez la case Enable hibernation as an additional stop behavior (Activer la mise en veille prolongée comme comportement d'arrêt supplémentaire).
5. Dans la page Ajouter un stockage pour le volume racine, spécifiez les informations suivantes :
   - Pour Taille (Go), saisissez la taille du volume EBS racine. Le volume doit être suffisamment grand pour stocker le contenu de la mémoire RAM et prendre en compte l'utilisation que vous prévoyez.

- Pour Type de volume, sélectionnez un type de volume EBS pris en charge (SSD à usage général (`gp2`) ou SSD IOPS provisionnés (`io1` ou `io2`).
- Pour Chiffrement, sélectionnez la clé de chiffrement du volume. Si vous avez activé le chiffrement par défaut dans cette région AWS, la clé de chiffrement par défaut est sélectionnée.

Pour de plus amples informations sur les prérequis relatifs au volume racine, veuillez consulter Prérequis de la mise en veille prolongée (p. 630).

6. Continuez comme indiqué par l'assistant. Lorsque vous avez terminé de vérifier vos options sur la page Examiner le lancement de l'instance, choisissez Lancer. Pour de plus amples informations, veuillez consulter Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

AWS CLI

Pour activer la mise en veille prolongée à l'aide d'AWS CLI

Utilisez la commande run-instances pour lancer une instance. Spécifiez les paramètres du volume racine EBS à l'aide du paramètre `--block-device-mappings file://mapping.json` et activez la mise en veille prolongée à l'aide du paramètre `--hibernation-options Configured=true`.

```
aws ec2 run-instances \
    --image-id ami-0abcdef1234567890 \
    --instance-type m5.large \
    --block-device-mappings file://mapping.json \
    --hibernation-options Configured=true \
    --count 1 \
    --key-name MyKeyPair
```

Spécifiez les éléments suivants dans `mapping.json` :

```
[
    {
        "DeviceName": "/dev/xvda",
        "Ebs": {
            "VolumeSize": 30,
            "VolumeType": "gp2",
            "Encrypted": true
        }
    }
]
```

Note

La valeur de `DeviceName` doit correspondre au nom du périphérique racine associé à l'AMI. Pour trouver le nom du périphérique racine, utilisez la commande describe-images comme suit :

```
aws ec2 describe-images --image-id ami-0abcdef1234567890
```

Si vous avez activé le chiffrement par défaut dans cette région AWS, vous pouvez omettre `"Encrypted": true`.

PowerShell

Pour activer la mise en veille prolongée à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande New-EC2Instance pour lancer une instance. Spécifiez le volume racine EBS en définissant d'abord le mappage au périphérique de stockage en mode bloc, puis en l'ajoutant à la

commande à l'aide du paramètre `-BlockDeviceMappings`. Activez la mise en veille prolongée à l'aide du paramètre `-HibernationOptions_Configured $true`.

```
PS C:\> $ebs_encrypt = New-Object Amazon.EC2.Model.BlockDeviceMapping
PS C:\> $ebs_encrypt.DeviceName = "/dev/xvda"
PS C:\> $ebs_encrypt.Ebs = New-Object Amazon.EC2.Model.EbsBlockDevice
PS C:\> $ebs_encrypt.Ebs.VolumeSize = 30
PS C:\> $ebs_encrypt.Ebs.VolumeType = "gp2"
PS C:\> $ebs_encrypt.Ebs.Encrypted = $true

PS C:\> New-EC2Instance `
             -ImageId ami-0abcdef1234567890 `
             -InstanceType m5.large `
             -BlockDeviceMappings $ebs_encrypt `
             -HibernationOptions_Configured $true `
             -MinCount 1 `
             -MaxCount 1 `
             -KeyName MyKeyPair
```

Note

La valeur de `DeviceName` doit correspondre au nom du périphérique racine associé à l'AMI. Pour trouver le nom du périphérique racine, utilisez la commande Get-EC2Image comme suit :

```
Get-EC2Image -ImageId ami-0abcdef1234567890
```

Si vous avez activé le chiffrement par défaut dans cette région AWS, vous pouvez omettre `Encrypted = $true` du mappage au périphérique de stockage en mode bloc.

Nouvelle console

Pour voir si une instance est activée pour la mise en veille prolongée à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et, sous l'onglet Détails de la section Détails de l'instance, inspectez le comportement Stop-hibernate. La valeur Enabled (Activé) indique que l'instance est activée pour la mise en veille prolongée.

Ancienne console

Pour voir si une instance est activée pour la mise en veille prolongée à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et dans le volet des détails, examinez Stop - Hibernation behavior (Arrêter - Comportement de mise en veille prolongée). La valeur Enabled (Activé) indique que l'instance est activée pour la mise en veille prolongée.

AWS CLI

Pour voir si une instance est activée pour la mise en veille prolongée à l'aide d'AWS CLI

Utilisez la commande describe-instances et spécifiez le paramètre `--filters "Name=hibernation-options.configured,Values=true"` pour filtrer les instances qui sont activées pour la mise en veille prolongée.

```
aws ec2 describe-instances \
    --filters "Name=hibernation-options.configured,Values=true"
```

Le champ suivant dans le résultat indique que l'instance est activée pour la mise en veille prolongée.

```
"HibernationOptions": {
    "Configured": true
}
```

PowerShell

Pour voir si une instance est activée pour la mise en veille prolongée à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande Get-EC2Instance et spécifiez le paramètre `-Filter @{ Name="hibernation-options.configured"; Value="true"}` pour filtrer les instances qui sont activées pour la mise en veille prolongée.

```
Get-EC2Instance `
    -Filter @{ Name="hibernation-options.configured"; Value="true"}
```

Le résultat répertorie les instances EC2 qui sont activées pour l'hibernation.

Désactivation de KASLR sur une instance (Ubuntu uniquement)

Pour exécuter une mise en veille prolongée sur une instance récemment lancée avec Ubuntu 16.04 LTS - Xenial ou Ubuntu 18.04 LTS - Bionic publiée avec le numéro de série 20190722.1 ou ultérieur, il est recommandé de désactiver KASLR (Kernel Address Space Layout Randomization). Sur Ubuntu 16.04 LTS ou Ubuntu 18.04 LTS, KASLR est activé par défaut. KASLR est une fonction de sécurité du noyau Linux standard qui permet d'atténuer l'exposition aux vulnérabilités d'accès à la mémoire pas encore découvertes, et leurs ramifications, en randomisant la valeur de l'adresse de base du noyau. En activant KASLR, il est possible que l'instance ne reprenne pas son exécution après sa mise en veille prolongée.

Pour en savoir plus sur KASLR, veuillez consulter la page relative aux fonctionnalités d'Ubuntu.

Pour désactiver KASLR sur une instance lancée avec Ubuntu

1. Connectez-vous à votre instance à l'aide de SSH. Pour de plus amples informations, veuillez consulter Connexion à votre instance Linux à l'aide de SSH (p. 601).
2. Ouvrez le fichier `/etc/default/grub.d/50-cloudimg-settings.cfg` dans l'éditeur de votre choix. Éditez la ligne `GRUB_CMDLINE_LINUX_DEFAULT` de sorte à ajouter l'option `nokaslr` à la fin de la ligne, comme illustré dans l'exemple suivant.

```
GRUB_CMDLINE_LINUX_DEFAULT="console=tty1 console=ttyS0 nvme_core.io_timeout=4294967295
 nokaslr"
```

3. Enregistrez le fichier et quittez votre éditeur.
4. Exécutez la commande suivante pour recréer la configuration Grub.

```
[ec2-user ~]$ sudo update-grub
```

5. Redémarrez l'instance.

```
[ec2-user ~]$ sudo reboot
```

6. Confirmez que l'option `nokaslr` a été ajoutée lors de l'exécution de la commande suivante.

```
[ec2-user ~]$ cat /proc/cmdline
```

Le résultat de la commande doit inclure l'option `nokaslr`.

Mise en veille prolongée d'une instance

Vous pouvez mettre une instance en veille prolongée uniquement si celle-ci est activée pour la mise en veille prolongée (p. 634) et si elle répond aux conditions préalables à la mise en veille prolongée (p. 630). Si une instance ne peut pas être mise en veille prolongée, un arrêt normal a lieu.

Nouvelle console

Pour mettre en veille prolongée une instance basée sur Amazon EBS à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une instance et choisissez État de l'instance, Mettre en veille prolongée les instances. Si Mettre l'instance en veille prolongée est désactivé, l'instance est déjà en veille prolongée ou arrêtée, ou elle ne peut pas être mise en veille prolongée. Pour de plus amples informations, veuillez consulter Prérequis de la mise en veille prolongée (p. 630).
4. Lorsque vous êtes invité à confirmer l'opération, choisissez Mettre en veille prolongée. La mise en veille prolongée de l'instance peut prendre quelques minutes. L'état de l'instance passe à Arrêt pendant que l'instance est mise en veille prolongée, puis à Arrêté(e) lorsque l'instance est en veille prolongée.

Ancienne console

Pour mettre en veille prolongée une instance basée sur Amazon EBS à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une instance et choisissez Actions, État de l'instance, Arrêter. Si Stop - Hibernate (Arrêter - Mettre en veille prolongée) est désactivé, l'instance est déjà en veille prolongée ou arrêtée, ou elle ne peut pas être mise en veille prolongée. Pour plus d'informations, consultez Prérequis de la mise en veille prolongée (p. 630).
4. Dans la boîte de dialogue de confirmation, sélectionnez Yes, Stop - Hibernate (Oui, Arrêter - Mettre en veille prolongée). La mise en veille prolongée de l'instance peut prendre quelques minutes. L'État de l'instance passe à Arrêt pendant que l'instance est mise en veille prolongée, puis Arrêtée lorsque l'instance a été mise en veille prolongée.

AWS CLI

Pour mettre en veille prolongée une instance basée sur Amazon EBS à l'aide d'AWS CLI

Utilisez la commande stop-instances et spécifiez le paramètre `--hibernate`.

```
aws ec2 stop-instances \
```

```
    --instance-ids i-1234567890abcdef0 \
    --hibernate
```

PowerShell

Pour mettre en veille prolongée une instance basée sur Amazon EBS à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande Stop-EC2Instance et spécifiez le paramètre `-Hibernate $true`.

```
Stop-EC2Instance `
    -InstanceId i-1234567890abcdef0 `
    -Hibernate $true
```

Nouvelle console

Pour voir si la mise en veille prolongée est initiée sur une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et, sous l'onglet Détails de la section Détails de l'instance, inspectez le message de transition d'état. Le message Client.UserInitiatedHibernate : User initiated hibernate (L'utilisateur a initié la mise en veille prolongée) indique que la mise en veille prolongée a été initiée sur l'instance.

Ancienne console

Pour voir si la mise en veille prolongée est initiée sur une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et dans le volet des détails, examinez Message de motif de transition de l'état. Le message Client.UserInitiatedHibernate : User initiated hibernate (L'utilisateur a initié la mise en veille prolongée) indique que la mise en veille prolongée a été initiée sur l'instance.

AWS CLI

Pour voir si la mise en veille prolongée est initiée sur une instance à l'aide d'AWS CLI

Utilisez la commande describe-instances et spécifiez le filtre `state-reason-code` pour afficher les instances sur lesquelles la mise en veille prolongée est initiée.

```
aws ec2 describe-instances \
    --filters "Name=state-reason-code,Values=Client.UserInitiatedHibernate"
```

Le champ suivant dans le résultat indique que la mise en veille prolongée a été initiée sur l'instance.

```
"StateReason": {
    "Code": "Client.UserInitiatedHibernate"
}
```

PowerShell

Pour voir si la mise en veille prolongée est initiée sur une instance à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande Get-EC2Instance et spécifiez le filtre `state-reason-code` pour afficher les instances sur lesquelles la mise en veille prolongée est initiée.

```
Get-EC2Instance `
    -Filter @{Name="state-reason-code";Value="Client.UserInitiatedHibernate"}
```

Le résultat répertorie les instances EC2 sur lesquelles la mise en veille prolongée a été initiée.

Démarrage d'une instance mise en veille prolongée

Démarrez une instance mise en veille prolongée comme vous le feriez pour une instance arrêtée.

Nouvelle console

Pour démarrer une instance mise en veille prolongée à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une instance mise en veille prolongée et choisissez État de l'instance, Démarrer l'instance. Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`. Pendant ce temps, les contrôles de statut (p. 743) de l'instance montrent l'instance à un état d'échec jusqu'à ce que l'instance ait démarré.

Ancienne console

Pour démarrer une instance mise en veille prolongée à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une instance mise en veille prolongée et choisissez Actions, État de l'instance, Début. Il peut s'écouler quelques minutes avant que l'instance ne passe à l'état `running`. Pendant ce temps, les contrôles de statut (p. 743) de l'instance montrent l'instance à un état d'échec jusqu'à ce que l'instance ait démarré.

AWS CLI

Pour démarrer une instance mise en veille prolongée à l'aide d'AWS CLI

Utilisez la commande start-instances.

```
aws ec2 start-instances \
    --instance-ids i-1234567890abcdef0
```

PowerShell

Pour démarrer une instance mise en veille prolongée à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande Start-EC2Instance.

```
Start-EC2Instance `
    -InstanceId i-1234567890abcdef0
```

Résolution des problèmes de mise en veille prolongée

Utilisez ces informations pour diagnostiquer et résoudre les problèmes courants que vous pourriez rencontrer lors de la mise en veille prolongée d'une instance.

Impossible d'effectuer une mise en veille prolongée immédiatement après le lancement

Si vous essayez de mettre en veille prolongée une instance trop rapidement après l'avoir lancée, vous obtiendrez une erreur.

Vous devez patienter environ deux minutes après le lancement avant de la mettre en veille prolongée.

Le passage de `stopping` à `stopped` prend du temps et l'état de la mémoire n'est pas restauré après le démarrage

Si votre instance mise en veille prolongée prend du temps pour passer de l'état `stopping` à `stopped`, et si l'état de la mémoire n'est pas restauré après que vous avez démarré, cela peut indiquer que la mise en veille prolongée n'a pas été configurée correctement.

Consultez le journal système de l'instance et recherchez les messages liés à la mise en veille prolongée. Pour accéder au journal système, connectez-vous (p. 598) à l'instance ou utilisez la commande get-console-output. Recherchez les lignes de journal de l'agent `hibinit-agent`. Si les lignes de journal indiquent un échec ou si les lignes de journal sont manquantes, il est probable qu'un échec de la configuration de la mise en veille prolongée au lancement ait eu lieu.

Par exemple, le message suivant indique que le volume racine de l'instance n'est pas suffisamment grand :
`hibinit-agent: Insufficient disk space. Cannot create setup for hibernation. Please allocate a larger root device.`

Si la dernière ligne de journal de l'agent `hibinit-agent` est `hibinit-agent: Running: swapoff /swap`, la mise en veille prolongée a été configurée avec succès.

Si vous ne voyez aucun journal issu de ces processus, votre AMI ne prend pas en charge la mise en veille prolongée. Pour plus d'informations sur les AMI supportées, consultez Prérequis de la mise en veille prolongée (p. 630). Si vous avez utilisé votre propre AMI, vérifiez que vous avez suivi les instructions pour Configuration d'une AMI existante pour prendre en charge la mise en veille prolongée (p. 632).

Instance « bloquée » dans l'état d'arrêt

Si vous avez mis votre instance en veille prolongée que celle-ci semble « bloquée » à l'état `stopping`, vous pouvez forcer son arrêt. Pour de plus amples informations, veuillez consulter Résolution des problèmes d'arrêt de votre instance (p. 1335).

Redémarrer votre instance

Le redémarrage d'une instance est similaire à celui d'un système d'exploitation. Dans la plupart des cas, il suffit de quelques minutes pour redémarrer votre instance. Lorsque vous redémarrez une instance, elle conserve son nom DNS public (IPv4), son adresse IPv4 privée, son adresse IPv6 (le cas échéant) et toutes les données se trouvant sur ses volumes de stockage d'instance.

Le redémarrage d'une instance ne déclenche pas de nouvelle de facturation d'instance (avec frais d'une minute minimum), contrairement à l'arrêt, puis au démarrage d'une instance.

Il peut nous arriver de planifier le redémarrage d'une instance pour effectuer des tâches de maintenance, par exemple pour appliquer des mises à jour qui requièrent un redémarrage. Le cas échéant, aucune action n'est requise de votre part. Nous vous recommandons d'attendre simplement le redémarrage dans le créneau horaire prévu. Pour plus d'informations, consultez Événements planifiés pour vos instances. (p. 749).

Nous vous recommandons d'utiliser la console Amazon EC2, un outil de ligne de commande ou l'API Amazon EC2 pour réamorcer votre instance au lieu d'exécuter la commande de réamorçage du système d'exploitation à partir de votre instance. Si vous utilisez la console Amazon EC2, un outil de ligne de commande ou l'API Amazon EC2 pour réamorcer votre instance, nous procédons à un redémarrage matériel si l'instance ne s'arrête pas correctement en quelques minutes. En revanche, si vous avez recours à AWS CloudTrail, l'utilisation d'Amazon EC2 pour redémarrer votre instance crée également un enregistrement d'API du moment où votre instance a été redémarrée.

Nouvelle console

Pour redémarrer une instance à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, État de l'instance, Redémarrer l’instance.
4. Lorsque vous êtes invité à confirmer l'opération, sélectionnez Redémarrer. L'instance reste dans l'état en cours d'exécution.

Ancienne console

Pour redémarrer une instance à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Etat de l'instance, Redémarrer.
4. Lorsque vous êtes invité à confirmer l'opération, sélectionnez Yes, Reboot. L'instance reste dans l'état en cours d'exécution.

Pour redémarrer une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- reboot-instances (AWS CLI)
- Restart-EC2Instance (Outils AWS pour Windows PowerShell)

Mise hors service d'instance

Une instance est planifiée pour être mise hors service quand AWS détecte une défaillance irréparable du matériel sous-jacent hébergeant l'instance. Quand une instance atteint sa date de mise hors service planifiée, elle est arrêtée ou terminée par AWS.

- Si le périphérique racine de votre instance est un volume Amazon EBS, l'instance est arrêtée et vous pouvez la redémarrer à tout moment. Le démarrage de l'instance arrêtée la migre vers un nouveau matériel.
- Si le périphérique racine de votre instance est un volume de stockage d'instance, l'instance est terminée et ne peut pas être utilisée à nouveau.

Pour plus d'informations sur les types d'événements d'instance, consultez Événements planifiés pour vos instances. (p. 749).

Sommaire

- Identification des instances planifiées pour une mise hors service (p. 643)
- Mesures à prendre sur les instances basées sur EBS dont la mise hors service est prévue (p. 644)
- Mesures à prendre pour les instances sauvegardées dans le stockage d'instances dont la mise hors service est prévue (p. 645)

Identification des instances planifiées pour une mise hors service

Si votre instance est planifiée pour une mise hors service, vous recevez un courrier électronique préalable à l'événement avec l'ID d'instance et la date de mise hors service. Vous pouvez également rechercher les instances planifiées pour une mise hors service à l'aide de la console Amazon EC2 ou de la ligne de commande.

> **Important**
>
> Si une instance est programmée pour une mise hors service, nous vous recommandons de prendre des mesures dès que possible car elle peut être inaccessible. (La notification par e-mail que vous recevez indique ce qui suit : « En raison de cette dégradation, votre instance pourrait déjà être inaccessible. ») Pour plus d'informations sur les mesures recommandées, consultez Vérifiez si votre instance est accessible.

Comment identifier des instances prévues pour une mise hors service

- Notification par e-mail (p. 643)
- Identification par la console (p. 643)

Notification par e-mail

Si votre instance est planifiée pour une mise hors service, vous recevez un courrier électronique préalable à l'événement avec l'ID d'instance et la date de mise hors service.

L'e-mail est envoyé à l'adresse associée à votre compte. Il s'agit de la même adresse e-mail que vous utilisez pour vous connecter à AWS Management Console. Pour mettre à jour les informations de contact de votre compte, accédez à la page Account Settings.

Identification par la console

Si vous utilisez un compte de messagerie que vous ne consultez pas régulièrement en rapport avec les notifications de mise hors service d'instances, vous pouvez utiliser la console Amazon EC2 ou la ligne de commande pour déterminer si l'une de vos instances est planifiée pour la mise hors service.

Pour identifier les instances planifiées pour une mise hors service à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, choisissez Tableau de bord du EC2. Sous Événements planifiés, vous pouvez voir les événements associés à vos instances et volumes Amazon EC2, organisés par région.

3. Si vous avez une instance avec un événement planifié affiché, sélectionnez le lien sous le nom de la région pour accéder à la page Événements.
4. La page Events répertorie toutes les ressources qui ont des événements associés. Pour afficher les instances planifiées pour une mise hors service, sélectionnez Instance resources dans la première liste de filtres, puis Instance stop or retirement dans la deuxième liste de filtres.
5. Si les résultats du filtre affichent une instance planifiée pour une mise hors service, sélectionnez-la et notez les date et heure dans le champ Start time du volet des détails. Il s'agit de la date de mise hors service de votre instance.

Pour identifier les instances planifiées pour une mise hors service à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-instance-status (AWS CLI)
- Get-EC2InstanceStatus (Outils AWS pour Windows PowerShell)

Mesures à prendre sur les instances basées sur EBS dont la mise hors service est prévue

Pour conserver les données de votre instance mise hors service, vous pouvez effectuer l'une des actions suivantes. Il est important que vous preniez cette action avant la date de mise hors service de l'instance, afin de prévenir tout arrêt et perte de données imprévus.

Pour savoir si votre instance est basée sur EBS ou sur le stockage d'instances, veuillez consulter Détermination du type de périphérique racine de votre instance (p. 24).

Vérifiez si votre instance est accessible

Lorsque vous êtes averti que votre instance est programmée pour une mise hors service, nous vous recommandons de prendre les mesures suivantes dès que possible :

- Vérifiez si votre instance est accessible en vous connectant (p. 598) ou en envoyant une demande ping à celle-ci.
- Si votre instance est accessible, vous devez prévoir de l'arrêter/la démarrer à un moment approprié avant la date de mise hors service prévue, lorsque l'impact est minime. Pour de plus amples informations sur l'arrêt et le redémarrage de votre instance, et sur ce que vous devez escompter quand votre instance est arrêtée, comme les conséquences sur les adresses publiques, privées et IP Elastic associées à votre instance, veuillez consulter Arrêt et démarrage de votre instance (p. 625). Veuillez noter que les données sur les volumes de stockage d'instance sont perdues lorsque vous arrêtez et démarrez votre instance.
- Si votre instance est inaccessible, vous devez prendre des mesures immédiates et effectuer un arrêt/démarrage (p. 625) pour la récupérer.
- Sinon, si vous souhaitez mettre fin (p. 645) à votre instance, prévoyez de le faire dès que possible, afin de cesser d'engager des frais pour cette dernière.

Créez une sauvegarde de votre instance

Pour disposer d'une sauvegarde, créez une AMI basée sur EBS à partir de votre instance. Pour garantir l'intégrité des données, arrêtez l'instance avant de créer l'AMI. Vous pouvez attendre la date de mise hors service planifiée quand l'instance est arrêtée ou arrêtez l'instance vous-même avant la date de mise hors service. Vous pouvez redémarrer l'instance à tout moment. Pour plus d'informations, consultez Création d'une AMI Linux basée sur Amazon EBS (p. 132).

Lancement d'une instance de remplacement

Après avoir créé une AMI à partir de votre instance, vous pouvez utiliser l'AMI pour lancer une instance de remplacement. Dans la console Amazon EC2, sélectionnez votre nouvelle AMI, puis choisissez Actions, Launch (Lancer). Suivez l'assistant pour lancer l'instance. Pour de plus amples informations sur chaque étape de l'assistant, veuillez consulter Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Mesures à prendre pour les instances sauvegardées dans le stockage d'instances dont la mise hors service est prévue

Pour conserver les données de votre instance mise hors service, vous pouvez effectuer l'une des actions suivantes. Il est important que vous preniez cette action avant la date de mise hors service de l'instance, afin de prévenir tout arrêt et perte de données imprévus.

> Warning
>
> Si votre instance basée sur le stockage d'instance dépasse sa date de mise hors service, elle est terminée et vous ne pouvez pas récupérer l'instance ou les données qui y étaient stockées. Quel que soit le périphérique racine de votre instance, les données des volumes de stockage d'instance sont perdues quand l'instance est mise hors service, même si les volumes sont attachés à une instance basée sur EBS.

Vérifiez si votre instance est accessible

Lorsque vous êtes averti que votre instance est programmée pour une mise hors service, nous vous recommandons de prendre les mesures suivantes dès que possible :

- Vérifiez si votre instance est accessible en vous connectant (p. 598) ou en envoyant une demande ping à celle-ci.
- Si votre instance est inaccessible, les chances de la récupérer sont vraiment très réduites. Pour plus d'informations, consultez Résolution d'un problème d'instance inaccessible (p. 1360). AWS mettra fin à votre instance à la date de mise hors service prévue. Par conséquent, pour une instance inaccessible, vous pouvez immédiatement mettre fin (p. 645) à l'instance vous-même.

Lancement d'une instance de remplacement

Créez une AMI basée sur le stockage d'instance à partir de votre instance à l'aide des outils AMI, comme décrit dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136). Dans la console Amazon EC2, sélectionnez votre nouvelle AMI, puis choisissez Actions, Launch (Lancer). Suivez l'assistant pour lancer l'instance. Pour de plus amples informations sur chaque étape de l'assistant, veuillez consulter Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Convertir votre instance en instance basée sur EBS

Transférez vos données vers un volume EBS, prenez un instantané du volume, puis créez une AMI à partir de l'instantané. Vous pouvez lancer une instance de remplacement à partir de votre nouvel AMI. Pour plus d'informations, consultez Conversion d'une AMI basée sur un stockage d'instance en AMI basée sur des volumes Amazon EBS (p. 148).

Résilier une instance

Vous pouvez supprimer votre instance lorsque vous n'en avez plus besoin. Cette opération est appelée mise hors service (ou résiliation) de votre instance. Dès que l'état d'une instance passe à `shutting-down` ou `terminated`, l'instance ne vous est plus facturée.

Vous ne pouvez pas vous connecter à une instance mise hors service, ni la démarrer. Toutefois, vous pouvez lancer des instances supplémentaires à l'aide de la même AMI. Si vous préférez arrêter et démarrer votre instance, ou la mettre en veille prolongée, consultez Arrêt et démarrage de votre instance (p. 625) ou Mettre votre instance Linux en veille prolongée (p. 628). Pour plus d'informations, consultez Différences entre redémarrage, arrêt, mise en veille prolongée et résiliation (p. 528).

Sommaire

- Terminaison d'instance (p. 646)
- Ce qui se passe lorsque vous résiliez une instance (p. 646)
- Mise hors service d'une instance (p. 647)
- Activation de la protection contre la résiliation (p. 648)
- Modification du comportement d'arrêt lancé de l'instance (p. 649)
- Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649)
- Dépannage (p. 652)

Terminaison d'instance

Une fois que vous avez mis une instance hors service, elle demeure visible sur la console pendant un court instant, puis l'entrée est supprimée automatiquement. Vous ne pouvez pas supprimer vous-même l'entrée de l'instance mise hors service. Une fois qu'une instance est résiliée, les ressources telles que les volumes et les balises en sont progressivement dissociées, et elles ne seront plus visibles dans l'instance terminée après un certain temps.

Lorsqu'une instance est mise hors service, les données des volumes de stockage des instances associées à cette instance sont supprimées.

Par défaut, les volumes du périphérique racine Amazon EBS sont supprimés automatiquement lorsque l'instance est mise hors service. Toutefois, par défaut, tout volume EBS supplémentaire attaché lors du lancement, ou tout volume EBS attaché à une instance existante, persiste même après la mise hors service de l'instance. Ce comportement est contrôlé par l'attribut `DeleteOnTermination` du volume, un attribut que vous pouvez modifier. Pour plus d'informations, consultez Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649).

Vous pouvez éviter que quelqu'un mette hors service une instance accidentellement en utilisant AWS Management Console, la CLI et l'API. Cette fonctionnalité est disponible pour les instances du stockage d'instances Amazon EC2 et pour les instances Amazon EBS. Chaque instance a un attribut `DisableApiTermination` avec la valeur par défaut `false` (l'instance peut être mise hors service via Amazon EC2). Vous pouvez modifier cet attribut d'instance pendant que l'instance est en cours d'exécution ou arrêtée (dans le cas des instances Amazon EBS). Pour plus d'informations, consultez Activation de la protection contre la résiliation (p. 648).

Vous pouvez déterminer si une instance doit être arrêtée ou mise hors service lorsque l'arrêt est lancé depuis l'instance à l'aide d'une commande du système d'exploitation pour l'arrêt du système. Pour plus d'informations, consultez Modification du comportement d'arrêt lancé de l'instance (p. 649).

Si vous exécutez un script de la mise hors service d'une instance, il est possible que cette dernière soit mise hors service de façon anormale dans la mesure où nous ne pouvons pas garantir le bon fonctionnement des scripts d'arrêt. Amazon EC2 tente d'arrêter une instance proprement et d'exécuter les scripts d'arrêt du système. Toutefois, certains événements (par exemple, une panne matérielle) peuvent empêcher l'exécution de ces scripts d'arrêt du système.

Ce qui se passe lorsque vous résiliez une instance

Lorsqu'une instance EC2 est résiliée à l'aide de la commande `terminate-instances`, les éléments suivants sont enregistrés au niveau du système d'exploitation :

- La demande d'API envoie un événement d'appui sur un bouton à l'invité.
- Divers services du système sont arrêtés suite à l'événement d'appui sur un bouton. systemd gère un arrêt normal du système. L'arrêt normal est déclenché par l'événement d'appui sur un bouton d'arrêt ACPI à partir de l'hyperviseur.
- L'arrêt ACPI est initialisé.
- L'instance s'arrête lorsque le processus d'arrêt normal se termine. L'heure d'arrêt du système d'exploitation n'est pas configurable.

Mise hors service d'une instance

Vous pouvez résilier une instance à l'aide d'AWS Management Console ou de la ligne de commande.

Par défaut, lorsque vous initiez une fermeture à partir d'une instance basée sur Amazon EBS (à l'aide de la commande shutdown ou poweroff), l'instance s'arrête. La commande halt ne déclenche pas un arrêt. Si elle est utilisée, l'instance n'est pas résiliée. Au lieu de cela, elle place l'UC à l'état `HLT` et l'instance continue de s'exécuter.

Nouvelle console

Pour résilier une instance à l'aide de la console

1. Avant de résilier une instance, vérifiez que vous ne perdrez aucune donnée en vous assurant que vos volumes Amazon EBS ne seront pas supprimés lors de la résiliation et que vous avez copié les données dont vous avez besoin des volumes du stockage d'instance vers un stockage persistant, par exemple Amazon EBS ou Amazon S3.
2. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
3. Dans le panneau de navigation, choisissez Instances.
4. Sélectionnez l'instance, puis choisissez Actions, État de l'instance, Résilier l'instance.
5. Choisissez Résilier lorsque vous êtes invité à confirmer.

Ancienne console

Pour résilier une instance à l'aide de la console

1. Avant de résilier une instance, vérifiez que vous ne perdrez aucune donnée en vous assurant que vos volumes Amazon EBS ne seront pas supprimés lors de la résiliation et que vous avez copié les données dont vous avez besoin des volumes du stockage d'instance vers un stockage persistant, par exemple Amazon EBS ou Amazon S3.
2. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
3. Dans le volet de navigation, choisissez Instances.
4. Sélectionnez l'instance et choisissez Actions, État de l'instance, Résilier.
5. Sélectionnez Oui, résilier lorsque vous êtes invité à confirmer l'opération.

Pour résilier une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- terminate-instances (AWS CLI)
- Stop-EC2Instance (Outils AWS pour Windows PowerShell)

Activation de la protection contre la résiliation

Par défaut, vous pouvez résilier votre instance à l'aide de la console Amazon EC2, de l'interface ligne de commande ou de l'API. Pour éviter que votre instance ne soit résiliée accidentellement à l'aide d'Amazon EC2, vous pouvez activer la protection contre la résiliation pour l'instance. L'attribut `DisableApiTermination` permet de vérifier si l'instance peut être résiliée à l'aide de la console, de l'interface ligne de commande ou de l'API. Par défaut, la protection contre la résiliation est désactivée pour votre instance. Vous pouvez définir la valeur de cet attribut lorsque vous lancez l'instance, pendant l'exécution de l'instance ou une fois l'instance arrêtée (pour les instances Amazon EBS).

L'attribut `DisableApiTermination` ne vous empêche pas de résilier une instance en lançant l'arrêt à partir de l'instance (à l'aide d'une commande du système d'exploitation pour l'arrêt système) lorsque l'attribut `InstanceInitiatedShutdownBehavior` est défini. Pour plus d'informations, consultez Modification du comportement d'arrêt lancé de l'instance (p. 649).

Limites

Vous ne pouvez pas activer la protection contre la résiliation pour les Instances Spot. En effet, une instance Spot est résiliée lorsque le prix Spot dépasse le montant que vous êtes prêt à payer pour les Instances Spot. Vous pouvez toutefois préparer votre application afin qu'elle soit en mesure de gérer les interruptions d'instance Spot. Pour plus d'informations, consultez Interruptions d'une instance Spot (p. 452).

L'attribut `DisableApiTermination` n'empêche pas Amazon EC2 Auto Scaling de résilier une instance. Pour les instances d'un groupe Auto Scaling, utilisez les fonctions Amazon EC2 Auto Scaling suivantes au lieu de la fonctionnalité de protection contre la résiliation Amazon EC2 :

- Pour empêcher les instances qui font partie d'un groupe Auto Scaling d'être résiliées lors d'une mise à l'échelle, utilisez la fonctionnalité de protection de l'instance. Pour plus d'informations, consultez Protection de l'instance dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.
- Pour empêcher Amazon EC2 Auto Scaling de résilier les instances défectueuses, interrompez le processus `ReplaceUnhealthy`. Pour plus d'informations, consultez Suspension et reprise des processus de mise à l'échelle dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.
- Pour spécifier les instances qu'Amazon EC2 Auto Scaling doit résilier en premier, choisissez une stratégie de résiliation. Pour plus d'informations, consultez Personnalisation de la stratégie de résiliation dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Pour activer la protection contre la résiliation d'une instance lors du lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le tableau de bord, sélectionnez Lancer une instance et suivez les instructions de l'assistant.
3. Sur la page Configurer les détails de l'instance, activez la case à cocher Activer la protection de la résiliation.

Pour activer la protection contre la résiliation d'une instance en cours d'exécution ou arrêtée

1. Sélectionnez l'instance, puis Actions, Instance Settings (Paramètres de l'instance) et Change Termination Protections (Changer la protection de la résiliation).
2. Choisissez Yes, Enable (Oui, Activer).

Pour désactiver la protection contre la résiliation d'une instance en cours d'exécution ou arrêtée

1. Sélectionnez l'instance, puis Actions, Instance Settings (Paramètres de l'instance) et Change Termination Protections (Changer la protection de la résiliation).
2. Choisissez Oui, désactiver.

Pour activer ou désactiver la protection contre la résiliation à l'aide de la ligne de commande.

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur l'interface ligne de commande, consultez Accès à Amazon EC2 (p. 3).

- modify-instance-attribute (AWS CLI)
- Edit-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

Modification du comportement d'arrêt lancé de l'instance

Par défaut, lorsque vous initiez un arrêt à partir d'une instance basée sur Amazon EBS (à l'aide d'une commande telle que shutdown ou poweroff), l'instance s'arrête (notez que halt n'émet pas de commande poweroff et, le cas échéant, l'instance n'est pas résiliée ; au lieu de cela, l'UC bascule en mode HLT et l'instance poursuit son exécution). Vous pouvez modifier ce comportement à l'aide de l'attribut `InstanceInitiatedShutdownBehavior` afin de mettre l'instance hors service. Vous pouvez mettre à jour cet attribut tandis que l'instance est en cours d'exécution ou arrêtée.

Vous pouvez mettre à jour l'attribut `InstanceInitiatedShutdownBehavior` à l'aide de la console Amazon EC2 ou de la ligne de commande. L'attribut `InstanceInitiatedShutdownBehavior` s'applique uniquement lorsque vous arrêtez le système d'exploitation de l'instance elle-même. Il ne s'applique pas lorsque vous arrêtez une instance à l'aide de l'API `StopInstances` ou de la console Amazon EC2.

Pour modifier le comportement d'arrêt d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Choisissez Actions, Paramètres d'instance, Modifier le comportement d'arrêt. Le comportement actuel est sélectionné.
5. Pour modifier le comportement, sélectionnez Arrêter ou Résilier à partir de Arrêter le comportement, puis choisissez Appliquer.

Pour modifier le comportement d'arrêt d'une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur l'interface ligne de commande, consultez Accès à Amazon EC2 (p. 3).

- modify-instance-attribute (AWS CLI)
- Edit-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

Conserver les volumes Amazon EBS lors de la résiliation d'une instance

Lorsqu'une instance est résiliée, Amazon EC2 utilise la valeur de l'attribut `DeleteOnTermination` pour chaque volume Amazon EBS attaché afin de déterminer si le volume doit être conservé ou supprimé.

La valeur par défaut de l' attribut `DeleteOnTermination` diffère selon que le volume est le volume racine de l'instance ou un volume non racine attaché à l'instance.

Volume racine

Par défaut, l'attribut `DeleteOnTermination` du volume racine d'une instance est défini sur `true`. Par conséquent, l'action par défaut consiste à supprimer le volume racine de l'instance lorsque

celle-ci est résiliée. L'attribut `DeleteOnTermination` peut être défini par le créateur d'une AMI ou par la personne qui lance une instance. Lorsque l'attribut est modifié par le créateur d'une AMI ou par la personne qui lance une instance, le nouveau paramètre remplace le paramètre par défaut d'origine de l'AMI. Nous vous recommandons de vérifier le paramètre par défaut de l'attribut `DeleteOnTermination` après avoir lancé une instance avec une AMI.

Volume non racine

Par défaut, lorsque vous attachez un volume EBS à une instance (p. 1109), son attribut `DeleteOnTermination` est défini sur `false`. L'action par défaut consiste donc à conserver ces volumes. Une fois l'instance mise hors service, vous pouvez prendre un instantané du volume conservé ou attacher celui-ci à une autre instance. Vous devez supprimer un volume pour éviter de générer des frais supplémentaires. Pour de plus amples informations, veuillez consulter Suppression d'un volume Amazon EBS (p. 1129).

Pour vérifier la valeur de l'attribut `DeleteOnTermination` d'un volume EBS en cours d'utilisation, examinez le mappage de périphérique de stockage en mode bloc de l'instance. Pour plus d'informations, consultez Affichage des volumes EBS dans le mappage de périphérique de stockage en mode bloc d'une instance (p. 1298).

Vous pouvez modifier la valeur de l'attribut `DeleteOnTermination` d'un volume lorsque vous lancez l'instance ou pendant l'exécution de celle-ci.

Exemples

- Modification du volume racine afin qu'il soit conservé lors du lancement à l'aide de la console (p. 650)
- Modification du volume racine afin qu'il soit conservé lors du lancement à l'aide de la ligne de commande (p. 651)
- Modification du volume racine d'une instance en cours d'exécution afin de le conserver à l'aide de la ligne de commande (p. 651)

Modification du volume racine afin qu'il soit conservé lors du lancement à l'aide de la console

À l'aide de la console, vous pouvez modifier l'attribut `DeleteOnTermination` lorsque vous lancez une instance. Pour modifier cet attribut lorsqu'il est associé à une instance en cours d'exécution, vous devez utiliser la ligne de commande.

Pour modifier le volume racine d'une instance afin de le conserver lors du lancement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le tableau de bord de la console, sélectionnez Lancer une instance.
3. Sur la page Sélection d'une Amazon Machine Image (AMI), choisissez une AMI, puis sélectionnez Sélectionner.
4. Suivez l'Assistant pour compléter les pages Choisir un type d'instance et Configurer les détails de l'instance.
5. Sur la page Ajouter le stockage, décochez la case Supprimer à la résiliation pour le volume racine.
6. Complétez les pages restantes de l'Assistant, puis sélectionnez Lancer.

Vous pouvez vérifier le paramètre en consultant les détails du volume du périphérique racine dans le volet des détails de l'instance. En regard de Périphériques de stockage en mode bloc, choisissez l'entrée du volume du périphérique racine. Par défaut, Supprimer à la résiliation a la valeur `True`. Si vous modifiez le comportement par défaut, Supprimer à la résiliation a la valeur `False`.

Modification du volume racine afin qu'il soit conservé lors du lancement à l'aide de la ligne de commande

Lorsque vous lancez une instance basée sur EBS, vous pouvez utiliser l'une des commandes suivantes afin de modifier le volume du périphérique racine à conserver. Pour obtenir plus d'informations sur l'interface ligne de commande, consultez Accès à Amazon EC2 (p. 3).

- run-instances (AWS CLI)
- New-EC2Instance (Outils AWS pour Windows PowerShell)

Par exemple, ajoutez l'option suivante à votre commande `run-instances` :

```
--block-device-mappings file://mapping.json
```

Spécifiez les éléments suivants dans `mapping.json` :

```
[
  {
    "DeviceName": "/dev/sda1",
    "Ebs": {
      "DeleteOnTermination": false,
      "SnapshotId": "snap-1234567890abcdef0",
      "VolumeType": "gp2"
    }
  }
]
```

Modification du volume racine d'une instance en cours d'exécution afin de le conserver à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes pour modifier le volume du périphérique racine d'une instance basée sur EBS en cours d'exécution afin de le conserver. Pour obtenir plus d'informations sur l'interface ligne de commande, consultez Accès à Amazon EC2 (p. 3).

- modify-instance-attribute (AWS CLI)
- Edit-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

Par exemple, utilisez la commande suivante :

```
aws ec2 modify-instance-attribute --instance-id i-1234567890abcdef0 --block-device-mappings
 file://mapping.json
```

Spécifiez les éléments suivants dans `mapping.json` :

```
[
  {
    "DeviceName": "/dev/sda1",
    "Ebs": {
      "DeleteOnTermination": false
    }
  }
]
```

Dépannage

Si vous arrêtez votre instance et qu'une autre instance démarre, vous avez probablement configuré la mise à l'échelle automatique via une fonctionnalité comme Flotte EC2 ou Amazon EC2 Auto Scaling.

Si votre instance garde l'état `shutting-down` pendant plus longtemps que d'habitude, elle finit par être nettoyée (mise hors service) par les processus automatisés du service Amazon EC2. Pour de plus amples informations, veuillez consulter Mise à fin d'instance retardée (p. 1337).

Récupération de votre instance

Vous pouvez créer une alarme Amazon CloudWatch qui surveille une instance Amazon EC2 et récupère automatiquement l'instance si cette dernière est dégradée suite à une défaillance du matériel sous-jacent ou à un problème nécessitant une intervention d'AWS pour sa résolution. Les instances mises hors service ne peuvent pas être récupérées. Une instance récupérée est identique à l'instance d'origine, y compris pour l'ID d'instance, les adresses IP privées, les adresses IP Elastic et toutes les métadonnées de l'instance. Si l'instance dégradée se trouve dans un groupe de placement, l'instance récupérée s'exécute dans le groupe de placement. Pour de plus amples informations sur l'utilisation d'alarmes Amazon CloudWatch pour récupérer une instance, consultez Ajout d'actions de récupération aux alarmes Amazon CloudWatch (p. 789). Pour résoudre les problèmes de défaillance de la récupération des instances, veuillez consulter Résolution des défaillances de récupération d'instance (p. 652).

Lorsque l'alarme `StatusCheckFailed_System` est déclenchée et que l'action de récupération est initiée, vous en êtes averti par la rubrique Amazon SNS que vous avez sélectionnée quand vous avez créé l'alarme et associé l'action de récupération. Lors de la récupération d'instance, l'instance est migrée pendant un redémarrage d'instance, et toutes les données en mémoire sont perdues. Lorsque le processus est terminé, les informations sont publiées dans la rubrique SNS que vous avez configurée pour l'alarme. Toutes les personnes abonnées à cette rubrique SNS recevront une notification par e-mail qui inclut le statut de la tentative de récupération et les éventuelles instructions supplémentaires. Vous remarquerez un redémarrage d'instance sur l'instance récupérée.

Voici quelques exemples de problèmes entraînant l'échec des contrôles de statut du système :

- Perte de connectivité réseau
- Perte d'alimentation système
- Problèmes logiciels sur un hôte physique
- Problèmes matériels sur un hôte physique ayant un impact sur l'accessibilité du réseau

Si votre instance a une adresse IPv4 publique, elle conserve la même adresse IPv4 publique après la récupération.

Exigences

L'opération de récupération est prise en charge uniquement sur les instances présentant les caractéristiques suivantes :

- Utilise l'un des types d'instance suivants : A1, C3, C4, C5, C5a, C5n, C6g, Inf1, M3, M4, M5, M5a, M5n, M6g, P3, R3, R4, R5, R5a, R5n, R6g, T2, T3, T3a, X1 ou X1e
- S'exécute dans un Virtual Private Cloud (VPC)
- Utilise une location d'instance `default` ou `dedicated`
- Possède uniquement des volumes EBS (ne pas configurer des volumes de stockage d'instance).

Résolution des défaillances de récupération d'instance

Les problèmes suivants peuvent causer l'échec de la récupération automatique de votre instance :

- Capacité temporaire, insuffisante du matériel de remplacement.
- L'instance possède un stockage d'instance attaché ce qui est une configuration non prise en charge pour la récupération automatique d'instance.
- Il existe un évènement continu sur le Tableau de bord de l'état des services qui empêchait le processus de récupération d'être bien exécuté. Reportez-vous à http://status.aws.amazon.com/ pour obtenir les dernières informations sur la disponibilité des services.
- L'instance a atteint l'autorisation quotidienne maximale de trois tentatives de récupération.

Le processus de récupération automatique tente de récupérer votre instance pour trois défaillances distinctes par jour au maximum. Si l'échec du contrôle de statut du système d'instance persiste, nous vous recommandons d'arrêter et de démarrer manuellement l'instance. Pour plus d'informations, consultez Arrêt et démarrage de votre instance (p. 625).

Votre instance peut ensuite supprimée si la récupération automatique échoue et il a été déterminé qu'une dégradation matérielle est la cause première de l'échec du contrôle de statut du système d'origine.

Personnalisation de votre instance Amazon Linux

Après avoir lancé votre instance Amazon Linux et vous être connecté avec succès, vous pouvez effectuer des changements. Il existe de nombreuses façons différentes de configurer une instance afin de répondre aux besoins d'une application spécifique. Les tâches suivantes comptent parmi celles couramment utilisées pour vous permettre de débuter.

Sommaire

- Scénarios de configuration courants (p. 653)
- Gestion des logiciels sur votre instance Amazon Linux (p. 654)
- Gestion des comptes d'utilisateur sur votre instance Amazon Linux (p. 660)
- Contrôle des états du processeur pour votre instance EC2 (p. 662)
- Réglage de l'heure pour votre instance Linux (p. 668)
- Optimisations des options d'UC (p. 673)
- Changement du nom d'hôte de votre instance Amazon Linux (p. 690)
- Réglage de DNS dynamique sur votre instance Amazon Linux (p. 692)
- Exécution de commandes sur votre instance Linux lors du lancement (p. 694)
- Métadonnées d'instance et données utilisateur (p. 701)

Scénarios de configuration courants

La distribution de base de Amazon Linux contient beaucoup de packages logiciels et d'utilitaires qui sont nécessaires aux opérations basiques du serveur. Néanmoins, beaucoup plus de packages logiciels sont disponibles dans différents référentiels de logiciels et encore plus de packages sont disponibles pour permettre la création à partir du code source. Pour plus d'informations sur l'installation et la création de logiciels à partir de ces emplacements, consultez le didacticiel Gestion des logiciels sur votre instance Amazon Linux (p. 654).

Les instances Amazon Linux sont préconfigurées avec un compte `ec2-user`, mais il se peut que vous souhaitiez ajouter d'autres comptes d'utilisateur qui n'ont pas de privilèges du super-utilisateur. Pour plus d'informations sur l'ajout et la suppression des comptes d'utilisateur, consultez le didacticiel Gestion des comptes d'utilisateur sur votre instance Amazon Linux (p. 660).

La configuration horaire par défaut pour les instances Amazon Linux utilise l'Amazon Time Sync Service pour définir l'heure du système sur une instance. Le fuseau horaire par défaut est UTC. Pour plus d'informations sur le réglage du fuseau horaire pour une instance ou l'utilisation de votre serveur horaire, consultez le didacticiel Réglage de l'heure pour votre instance Linux (p. 668).

Si vous possédez votre propre réseau avec un nom de domaine enregistré, vous pouvez changer le nom d'hôte d'une instance pour l'identifier dans le cadre de ce domaine. Vous pouvez aussi changer l'invite du système pour avoir un nom plus descriptif sans changer les réglages du nom d'hôte. Pour plus d'informations, consultez Changement du nom d'hôte de votre instance Amazon Linux (p. 690). Vous pouvez configurer une instance pour utiliser un fournisseur de services DNS dynamiques. Pour plus d'informations, consultez Réglage de DNS dynamique sur votre instance Amazon Linux (p. 692).

Lorsque vous lancez une instance dans Amazon EC2, vous avez la possibilité de transmettre les données des utilisateurs vers l'instance qui peut être utilisée pour effectuer des tâches de configuration communes et même exécuter des scripts après le démarrage de l'instance. Vous pouvez transmettre deux types de données utilisateur vers Amazon EC2 : des directives cloud-init et des scripts shell. Pour plus d'informations, consultez Exécution de commandes sur votre instance Linux lors du lancement (p. 694).

Gestion des logiciels sur votre instance Amazon Linux

La distribution de base de Amazon Linux contient beaucoup de packages logiciels et d'utilitaires qui sont nécessaires aux opérations basiques du serveur. Néanmoins, beaucoup plus de packages logiciels sont disponibles dans différents référentiels de logiciels et encore plus de packages sont disponibles pour permettre la création à partir du code source.

Table des matières

- Mise à jour du logiciel d'instance sur votre instance Amazon Linux (p. 655)
- Ajout de référentiels sur une instance Amazon Linux (p. 656)
- Recherche de packages logiciels sur une instance Amazon Linux (p. 657)
- Installation de packages logiciels sur une instance Amazon Linux (p. 658)
- Préparation à la compilation de logiciels sur une instance Amazon Linux (p. 659)

Il est important de garder les logiciels à jour. Beaucoup de packages dans une distribution Linux sont souvent mis à jour pour résoudre les bogues, ajouter des fonctions et protéger contre le code malveillant. Pour plus d'informations, consultez Mise à jour du logiciel d'instance sur votre instance Amazon Linux (p. 655).

Par défaut, les instances Amazon Linux se lancent avec les référentiels suivants activés :

- Amazon Linux 2 : `amzn2-core` et `amzn2extra-docker`
- Amazon Linux AMI : `amzn-main` et `amzn-updates`

Tandis qu'il existe beaucoup de packages disponibles dans ces référentiels qui sont mis à jour par Amazon Web Services, vous pouvez trouver un package dans un autre référentiel que vous souhaitez installer. Pour plus d'informations, consultez Ajout de référentiels sur une instance Amazon Linux (p. 656). Pour obtenir de l'aide pour trouver des packages dans les référentiels activés, consultez Recherche de packages logiciels sur une instance Amazon Linux (p. 657). Pour plus d'informations sur l'installation des logiciels sur une instance Amazon Linux, consultez le didacticiel Installation de packages logiciels sur une instance Amazon Linux (p. 658).

Les logiciels ne sont pas tous disponibles dans les packages logiciels stockés dans les référentiels. Certains logiciels doivent être compilés sur une instance à partir de son code source. Pour plus d'informations, consultez Préparation à la compilation de logiciels sur une instance Amazon Linux (p. 659).

Les instances Amazon Linux gèrent leur logiciel à l'aide d'un gestionnaire de package yum. Le gestionnaire de packages yum peut installer, supprimer et mettre à jour les logiciels ainsi que gérer l'ensemble des dépendances pour chaque package. Des distributions Linux basées sur Debian, comme Ubuntu, utilisent la commande apt-get et le gestionnaire de package dpkg, donc les exemples yum dans les sections suivantes ne fonctionnent pas pour ces distributions.

Mise à jour du logiciel d'instance sur votre instance Amazon Linux

Il est important de garder les logiciels à jour. Beaucoup de packages dans une distribution Linux sont souvent mis à jour pour résoudre les bogues, ajouter des fonctions et protéger contre le code malveillant. Lorsque vous lancez et vous vous connectez pour la première fois à une instance Amazon Linux, il se peut que vous voyez un message vous demandant de mettre à jour des packages logiciels à des fins de sécurité. Cette section explique comment mettre à jour l'ensemble d'un système ou juste un seul package.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

Pour mettre à jour tous les packages sur une instance Amazon Linux

1. (Facultatif) Lancez une session screen dans votre fenêtre shell. Il se peut que vous connaissiez parfois une interruption du réseau qui peut déconnecter la connexion SSH à votre instance. Si ce problème arrive pendant une longue mise à jour logicielle, cela peut laisser l'instance dans un état récupérable bien que désorienté. Une session screen vous permet de continuer à exécuter la mise à jour même si votre connexion est interrompue et vous pouvez vous reconnecter à la session plus tard sans problème.

   a. Exécutez la commande screen pour démarrer la session.

   ```
   [ec2-user ~]$ screen
   ```

   b. Si votre session est déconnectée, reconnectez-vous à votre instance et énumérez les écrans disponibles.

   ```
   [ec2-user ~]$ screen -ls
   There is a screen on:
    17793.pts-0.ip-12-34-56-78 (Detached)
   1 Socket in /var/run/screen/S-ec2-user.
   ```

   c. Reconnectez-vous à l'écran à l'aide de la commande screen -r et l'ID du processus de la commande précédente.

   ```
   [ec2-user ~]$ screen -r 17793
   ```

   d. Lorsque vous avez terminé d'utiliser screen, servez-vous de la commande exit pour fermer la session.

   ```
   [ec2-user ~]$ exit
   [screen is terminating]
   ```

2. Exécutez la commande yum update. Le cas échéant, vous pouvez ajouter l'indicateur `--security` pour appliquer uniquement les mises à jour de sécurité.

   ```
   [ec2-user ~]$ sudo yum update
   ```

3. Vérifiez les packages énumérés, et tapez **y** puis Entrée pour accepter les mises à jour. La mise à jour de tous les packages d'un système peut prendre plusieurs minutes. Les résultats yum montrent le statut de la mise à jour pendant son exécution.

4. (Facultatif) Redémarrez votre instance pour vous assurer que vous utilisez les derniers packages et bibliothèques de votre mise à jour ; les mises à jour noyau ne sont pas chargées jusqu'au prochain redémarrage. Les mises à jour de n'importe quelle bibliothèque `glibc` devraient aussi être suivies d'un redémarrage. Pour les mises à jour des packages qui contrôlent les services, il peut s'avérer suffisant de redémarrer les services pour récupérer les mises à jour, mais un redémarrage du système assure que toutes les mises à jour précédentes des packages et des bibliothèques sont terminées.

Pour mettre à jour un seul package sur une instance Amazon Linux

Utilisez cette procédure pour mettre à jour un seul package (et ses dépendances) et non l'ensemble du système.

1. Exécutez la commande yum update avec le nom du package que vous souhaiteriez mettre à jour.

```
[ec2-user ~]$ sudo yum update openssl
```

2. Vérifiez les détails relatifs aux packages énumérés, et tapez **y** puis Entrée pour accepter les mises à jour. Il peut arriver qu'il y ait plus d'un package énuméré s'il existe des dépendances de packages qui doivent être résolues. Les résultats yum montrent le statut de la mise à jour pendant son exécution.
3. (Facultatif) Redémarrez votre instance pour vous assurer que vous utilisez les derniers packages et bibliothèques de votre mise à jour ; les mises à jour noyau ne sont pas chargées jusqu'au prochain redémarrage. Les mises à jour de n'importe quelle bibliothèque `glibc` devraient aussi être suivies d'un redémarrage. Pour les mises à jour des packages qui contrôlent les services, il peut s'avérer suffisant de redémarrer les services pour récupérer les mises à jour, mais un redémarrage du système assure que toutes les mises à jour précédentes des packages et des bibliothèques sont terminées.

Ajout de référentiels sur une instance Amazon Linux

Par défaut, les instances Amazon Linux se lancent avec deux référentiels activés : `amzn-main` et `amzn-updates`. Tandis qu'il existe beaucoup de packages disponibles dans ces référentiels qui sont mis à jour par Amazon Web Services, vous pouvez trouver un package dans un autre référentiel que vous souhaitez installer.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

Pour installer un package d'un référentiel différent avec la commande yum, vous devez ajouter les détails relatifs au référentiel au fichier `/etc/yum.conf` ou à son propre fichier *`repository`*`.repo` dans le répertoire `/etc/yum.repos.d`. Vous pouvez le faire manuellement, mais la plupart des référentiels yum ont leur propre fichier *`repository`*`.repo` sur l'URL de leur référentiel.

Pour déterminer quels référentiels yum sont déjà installés

- Utilisez la commande suivante pour répertorier les référentiels yum installés :

```
[ec2-user ~]$ yum repolist all
```

  La sortie obtenue répertorie les référentiels installés et indique l'état de chacun. Les référentiels activés affichent le nombre de packages qu'ils contiennent.

Pour ajouter un référentiel yum à **`/etc/yum.repos.d`**

1. Recherchez l'emplacement du fichier `.repo`. Il variera selon le référentiel que vous ajoutez. Dans cet exemple, le fichier `.repo` se trouve à l'adresse `https://www.`*`example`*`.com/`*`repository`*`.repo`.

2. Ajoutez le référentiel à l'aide de la commande yum-config-manager.

```
[ec2-user ~]$ sudo yum-config-manager --add-repo https://
www.example.com/repository.repo
Loaded plugins: priorities, update-motd, upgrade-helper
adding repo from: https://www.example.com/repository.repo
grabbing file https://www.example.com/repository.repo to /etc/
yum.repos.d/repository.repo
repository.repo                                          | 4.0 kB     00:00
repo saved to /etc/yum.repos.d/repository.repo
```

Après avoir installé un référentiel, vous devez l'activer, comme décrit dans la procédure suivante.

Pour activer un référentiel yum dans `/etc/yum.repos.d`

- Utilisez la commande yum-config-manager avec l'indicateur `--enable repository`. La commande suivante active le référentiel EPEL (Extra Packages for Enterprise Linux) à partir du projet Fedora. Par défaut, ce référentiel est présent dans `/etc/yum.repos.d` sur les instances Amazon Linux AMI, mais il n'est pas activé.

```
[ec2-user ~]$ sudo yum-config-manager --enable epel
```

> **Note**
>
> Pour activer le référentiel EPEL sur Amazon Linux 2, utilisez la commande suivante :
>
> ```
> [ec2-user ~]$ sudo yum install https://dl.fedoraproject.org/pub/epel/epel-
> release-latest-7.noarch.rpm
> ```
>
> Pour obtenir des informations sur l'activation du référentiel EPEL sur d'autres distributions comme Red Hat et CentOS, consultez la documentation au sujet d'EPEL sur https://fedoraproject.org/wiki/EPEL.

Recherche de packages logiciels sur une instance Amazon Linux

Vous pouvez utiliser la commande yum search pour rechercher les descriptions des packages qui sont disponibles dans vos référentiels configurés. Elle est particulièrement utile si vous ne connaissez pas le nom exact du package que vous voulez installer. Il suffit de joindre la recherche de mots clés à la commande ; pour les recherches de plusieurs mots, entourez la requête de recherche avec des guillemets.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

```
[ec2-user ~]$ sudo yum search "find"
```

Voici un exemple de sortie pour Amazon Linux 2.

```
Loaded plugins: extras_suggestions, langpacks, priorities, update-motd
============================= N/S matched: find ==============================
findutils.x86_64 : The GNU versions of find utilities (find and xargs)
gedit-plugin-findinfiles.x86_64 : gedit findinfiles plugin
ocaml-findlib-devel.x86_64 : Development files for ocaml-findlib
```

```
perl-File-Find-Rule.noarch : Perl module implementing an alternative interface to
 File::Find
robotfindskitten.x86_64 : A game/zen simulation. You are robot. Your job is to find kitten.
mlocate.x86_64 : An utility for finding files by name
ocaml-findlib.x86_64 : Objective CAML package manager and build helper
perl-Devel-Cycle.noarch : Find memory cycles in objects
perl-Devel-EnforceEncapsulation.noarch : Find access violations to blessed objects
perl-File-Find-Rule-Perl.noarch : Common rules for searching for Perl things
perl-File-HomeDir.noarch : Find your home and other directories on any platform
perl-IPC-Cmd.noarch : Finding and running system commands made easy
perl-Perl-MinimumVersion.noarch : Find a minimum required version of perl for Perl code
texlive-xesearch.noarch : A string finder for XeTeX
valgrind.x86_64 : Tool for finding memory management bugs in programs
valgrind.i686 : Tool for finding memory management bugs in programs
```

Voici un exemple de sortie pour Amazon Linux.

```
Loaded plugins: priorities, security, update-motd, upgrade-helper
============================= N/S Matched: find ==============================
findutils.x86_64 : The GNU versions of find utilities (find and xargs)
perl-File-Find-Rule.noarch : Perl module implementing an alternative interface to
 File::Find
perl-Module-Find.noarch : Find and use installed modules in a (sub)category
libpuzzle.i686 : Library to quickly find visually similar images (gif, png, jpg)
libpuzzle.x86_64 : Library to quickly find visually similar images (gif, png, jpg)
mlocate.x86_64 : An utility for finding files by name
```

Les demandes de recherche de plusieurs mots entre guillemets donnent uniquement des résultats qui correspondent à la requête exacte. Si vous ne voyez pas le package attendu, simplifiez votre recherche en utilisant un mot clé, puis analyser les résultats. Vous pouvez aussi des synonymes des mots clés pour élargir votre recherche.

Pour de plus amples informations sur les packages pour Amazon Linux 2 et Amazon Linux, veuillez consulter :

- Référentiel de package (p. 189)
- Bibliothèque Extras (Amazon Linux 2) (p. 191)

Installation de packages logiciels sur une instance Amazon Linux

Le gestionnaire de package yum est un outil formidable pour l'installation de logiciels, car il peut chercher l'ensemble de vos référentiels activés pour différents packages logiciels mais aussi gérer toutes les dépendances dans le processus d'installation des logiciels.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

Pour installer un package à partir d'un référentiel, utilisez la commande yum install ***package*** en remplaçant *package* par le nom du logiciel à installer. Par exemple, pour installer le navigateur web à base de texte links, saisissez la commande suivante.

```
[ec2-user ~]$ sudo yum install links
```

Vous pouvez aussi utiliser yum install pour installer les fichiers du package RPM que vous avez téléchargé sur Internet. Pour cela, il vous suffit de joindre le nom du chemin d'un fichier RPM à la commande d'installation au lieu du nom d'un package de référentiel.

```
[ec2-user ~]$ sudo yum install my-package.rpm
```

Préparation à la compilation de logiciels sur une instance Amazon Linux

Il existe une mine de logiciels open source disponibles sur Internet qui n'ont pas été pré-compilés et proposés au téléchargement à partir d'un référentiel de logiciels. Il est possible que vous découvriez un package logiciel que vous devez compiler vous-même, à partir de son code source. Pour que votre système puisse compiler des logiciels, vous devez installer plusieurs outils de développement comme make, gcc et autoconf.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

Comme la compilation de logiciels n'est pas une tâche que chaque instance Amazon EC2 nécessite, ces outils ne sont pas installés par défaut, mais ils sont disponibles dans un groupe de packages appelé « Development tools » (Outils de développement) qui s'ajoute facilement à une instance avec la commande yum groupinstall.

```
[ec2-user ~]$ sudo yum groupinstall "Development Tools"
```

Les packages de code source de logiciels sont souvent disponibles pour le téléchargement (à partir des sites web comme https://github.com/ et http://sourceforge.net/) sous forme de fichier d'archives compressé, appelé un tarball. Ces tarballs portent généralement l'extension de fichier `.tar.gz`. Vous pouvez décompresser ces archives avec la commande tar.

```
[ec2-user ~]$ tar -xzf software.tar.gz
```

Après avoir décompressé et désarchivé le package de code source, vous devriez rechercher un fichier `README` ou `INSTALL` dans le répertoire du code source qui peut vous fournir plus d'instructions pour la compilation et l'installation du code source.

Pour récupérer le code source des packages Amazon Linux

Amazon Web Services fournit le code source pour les packages gérés. Vous pouvez télécharger le code source pour n'importe quel package installé avec la commande yumdownloader --source.

- Exécutez la commande yumdownloader --source ***package*** pour télécharger le code source pour *package*. Par exemple, pour télécharger le code source du package `htop`, saisissiez la commande suivante.

```
[ec2-user ~]$ yumdownloader --source htop

Loaded plugins: priorities, update-motd, upgrade-helper
Enabling amzn-updates-source repository
Enabling amzn-main-source repository
amzn-main-source
                        | 1.9 kB  00:00:00
amzn-updates-source
                        | 1.9 kB  00:00:00
(1/2): amzn-updates-source/latest/primary_db
                        |  52 kB  00:00:00
(2/2): amzn-main-source/latest/primary_db
                        | 734 kB  00:00:00
```

```
htop-1.0.1-2.3.amzn1.src.rpm
```

L'emplacement du fichier RPM est dans le répertoire à partir duquel vous avez exécuté la commande.

Gestion des comptes d'utilisateur sur votre instance Amazon Linux

Chaque type d'instance Linux est lancé avec un compte d'utilisateur du système Linux par défaut. Le nom d'utilisateur par défaut est déterminé par l'AMI qui a été spécifiée au moment du lancement de l'instance.

- Pour Amazon Linux 2 ou l'AMI Amazon Linux, le nom d'utilisateur est `ec2-user`.
- Pour une AMI CentOS, le nom d'utilisateur est `centos`.
- Pour une AMI Debian, le nom d'utilisateur est `admin`.
- Pour une AMI Fedora, le nom d'utilisateur est `ec2-user` ou `fedora`.
- Pour une AMI RHEL, le nom d'utilisateur est `ec2-user` ou `root`.
- Pour une AMI SUSE, le nom d'utilisateur est `ec2-user` ou `root`.
- Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.
- Si `ec2-user` et `root` ne fonctionnent pas, renseignez-vous auprès de votre fournisseur AMI.

> **Note**
>
> Les utilisateurs du système Linux ne doivent pas être confondus avec les utilisateurs AWS Identity and Access Management (IAM). Pour plus d'informations, consultez la section Utilisateurs et groupes IAM dans le IAM Guide de l'utilisateur.

Table des matières

- Considérations (p. 660)
- Création d'un compte d'utilisateur (p. 660)
- Suppression d'un compte utilisateur (p. 662)

Considérations

L'utilisation du compte d'utilisateur par défaut convient à de nombreux applications. Toutefois, vous pouvez décider d'ajouter des comptes d'utilisateur afin que les individus puissent disposer de leurs propres fichiers et espaces de travail. Par ailleurs, la création de comptes d'utilisateur pour de nouveaux utilisateurs est beaucoup plus sécurisée que l'octroi à plusieurs utilisateurs (probablement inexpérimentés) de l'accès au compte utilisateur par défaut, car ce compte peut engendrer beaucoup de dommages à un système lorsqu'il est mal utilisé. Pour plus d'informations, consultez Conseils pour sécuriser votre instance EC2.

Pour activer pour les utilisateurs l'accès SSH à votre instance EC2 à l'aide d'un compte d'utilisateur du système Linux, vous devez partager la clé SSH avec l'utilisateur. Vous pouvez également utiliser EC2 Instance Connect pour fournir l'accès aux utilisateurs sans devoir partager et gérer les clés SSH. Pour plus d'informations, consultez Connexion à votre instance Linux à l'aide de EC2 Instance Connect (p. 604).

Création d'un compte d'utilisateur

Créez d'abord le compte utilisateur, puis ajoutez la clé publique SSH qui permet à l'utilisateur de se connecter à l'instance.

Pour créer un compte utilisateur

1. Créez une nouvelle paire de clés (p. 1048). Vous devez fournir le fichier `.pem` à l'utilisateur pour lequel vous créez le compte d'utilisateur. Ils doivent utiliser ce fichier pour se connecter à l'instance.
2. Récupérez la clé publique de la paire de clés que vous avez créée à l'étape précédente.

```
$  ssh-keygen -y -f /path_to_key_pair/key-pair-name.pem
```

La commande renvoie la clé publique, comme indiqué dans l'exemple suivant.

```
ssh-rsa
 AAAAB3NzaC1yc2EAAAADAQABAAABAQClKsfkNkuSevGj3eYhCe53pcjqP3maAhDFcvBS7O6Vhz2ItxCih
+PnDSUaw+WNQn/mZphTk/a/gU8jEzoOWbkM4yxyb/wB96xbiFveSFJuOp/
d6RJhJOI0iBXrlsLnBItntckiJ7FbtxJMXLvvwJryDUilBMTjYtwB+QhYXUMOzce5Pjz5/i8SeJtjnV3iAoG/
cQk+0FzZqaeJAAHco+CY/5WrUBkrHmFJr6HcXkvJdWPkYQS3xqC0+FmUZofz221CBt5IMucxXPkX4rWi
+z7wB3RbBQoQzd8v7yeb7OzlPnWOyN0qFU0XA246RA8QFYiCNYwI3f05p6KLxEXAMPLE
```

3. Connectez-vous à l'instance.
4. Utilisez la commande adduser pour créer le compte utilisateur et l'ajouter au système (avec une entrée dans le fichier `/etc/passwd`). Cette commande crée également un groupe et un répertoire de base pour le compte. Dans cet exemple, le compte utilisateur s'appelle *`newuser`*.

   - Amazon Linux et Amazon Linux 2

   ```
   [ec2-user ~]$ sudo adduser newuser
   ```

   - Ubuntu

     Incluez le paramètre `--disabled-password` pour créer le compte d'utilisateur sans mot de passe.

   ```
   [ubuntu ~]$ sudo adduser newuser --disabled-password
   ```

5. Passez au nouveau compte pour que l'annuaire et le fichier que vous créez aient la propriété appropriée.

```
[ec2-user ~]$ sudo su - newuser
```

L'invite passe de `ec2-user` à *`newuser`* pour indiquer que vous avez basculé de la session shell au nouveau compte.

6. Ajoutez la clé publique SSH au compte utilisateur. Créez d'abord un répertoire dans le répertoire personnel de l'utilisateur pour le fichier de clé SSH, puis créez le fichier de clé, et enfin collez la clé publique dans le fichier de clé, conformément aux sous-étapes suivantes.

   a. Créez un répertoire `.ssh` dans le répertoire de base *`newuser`* et modifiez ses autorisations de fichier en `700` (seul le propriétaire peut ouvrir le répertoire et y lire ou y écrire).

   ```
   [newuser ~]$ mkdir .ssh
   ```

   ```
   [newuser ~]$ chmod 700 .ssh
   ```

   > **Important**
   >
   > Sans les autorisations de fichier exactes, l'utilisateur ne pourra pas se connecter.

   b. Créez un fichier nommé `authorized_keys` dans le répertoire `.ssh` et modifiez ses autorisations de fichier en `600` (seul le propriétaire peut lire le fichier ou y écrire).

```
[newuser ~]$ touch .ssh/authorized_keys
```

```
[newuser ~]$ chmod 600 .ssh/authorized_keys
```

> **Important**
>
> Sans les autorisations de fichier exactes, l'utilisateur ne pourra pas se connecter.

c. Ouvrez le fichier `authorized_keys` avec votre éditeur de texte préféré (comme vim ou nano).

```
[newuser ~]$ nano .ssh/authorized_keys
```

Collez la clé publique que vous avez récupérée à l' étape 2 dans le fichier et enregistrez les modifications.

> **Important**
>
> Assurez-vous que vous collez la clé publique dans une ligne continue. La clé publique ne doit pas être divisée sur plusieurs lignes.

L'utilisateur doit pouvoir se connecter au compte *`newuser`* de votre instance à l'aide de la clé privée qui correspond à la clé publique que vous avez ajoutée au fichier `authorized_keys`. Pour de plus amples informations sur les différentes méthodes de connexion à une instance Linux, veuillez consulter Connectez-vous à votre instance Linux (p. 598).

Suppression d'un compte utilisateur

Si un compte d'utilisateur n'est plus nécessaire, vous pouvez supprimer ce compte pour qu'il ne puisse plus être utilisé.

Utilisez la commande userdel pour supprimer le compte utilisateur du système. Quand vous spécifiez le paramètre `-r`, le répertoire de base et le fichier temporaire des e-mails de l'utilisateur sont supprimés. Pour conserver le répertoire de base et le fichier temporaire des e-mails de l'utilisateur, omettez le paramètre `-r`.

```
[ec2-user ~]$ sudo userdel -r olduser
```

Contrôle des états du processeur pour votre instance EC2

Les états « C-states » contrôlent les niveaux de veille dans lesquels un cœur peut entrer lorsqu'il est inutilisé. Les états « C-states » sont numérotés de C0 (l'état le plus superficiel lorsque le cœur est totalement éveillé et exécute les instructions) à C6 (l'état de veille le plus profond lorsqu'un cœur est arrêté). Les états « P-states » contrôlent les performances souhaitées (dans la fréquence de l'UC) à partir d'un cœur. La numérotation des états « P-states » commence à P0 (paramètre de performance le plus élevé dans lequel le cœur peut utiliser la technologie Intel Turbo Boost pour améliorer la fréquence si possible) et va de P1 (état « P-state » qui demande la fréquence de base maximale) à P15 (fréquence la plus basse possible).

Les types d'instance EC2 suivants permettent à un système d'exploitation de contrôler les états « C-states » et « P-states » des processeurs.

- Usage général : `m4.10xlarge` | `m4.16xlarge` | `m5.metal` | `m5d.metal`
- Calcul optimisé : `c4.8xlarge` | `c5.metal` | `c5n.metal`

- Mémoire optimisée : `r4.8xlarge` | `r4.16xlarge` | `r5.metal` | `r5d.metal` | `u-6tb1.metal` | `u-9tb1.metal` | `u-12tb1.metal` | `u-18tb1.metal` | `u-24tb1.metal` `x1.16xlarge` | `x1.32xlarge` | `x1e.8xlarge` | `x1e.16xlarge` | `x1e.32xlarge` | `z1d.metal`
- Stockage optimisé : `d2.8xlarge` | `i3.8xlarge` | `i3.16xlarge` | `i3.metal` | `i3en.metal` | `h1.8xlarge` | `h1.16xlarge`
- Calcul accéléré : `f1.16xlarge` | `g3.16xlarge` | `g4dn.metal` | `p2.16xlarge` | `p3.16xlarge`

Les types d'instance suivants permettent à un système d'exploitation de contrôler les états « C-states » des processeurs :

- Usage général : `m5.12xlarge` | `m5.24xlarge` | `m5d.12xlarge` | `m5d.24xlarge` | `m5n.12xlarge` | `m5n.24xlarge` | `m5dn.12xlarge` | `m5dn.24xlarge`
- Calcul optimisé : `c5.9xlarge` | `c5.12xlarge` | `c5.18xlarge` | `c5.24xlarge` | `c5a.24xlarge` | `c5ad.24xlarge` | `c5d.9xlarge` | `c5d.12xlarge` | `c5d.18xlarge` | `c5d.24xlarge` | `c5n.9xlarge` | `c5n.18xlarge`
- Mémoire optimisée : `r5.12xlarge` | `r5.24xlarge` | `r5d.12xlarge` | `r5d.24xlarge` | `r5n.12xlarge` | `r5n.24xlarge` | `r5dn.12xlarge` | `r5dn.24xlarge` | `z1d.6xlarge` | `z1d.12xlarge`
- Stockage optimisé : `i3en.12xlarge` | `i3en.24xlarge`
- Calcul accéléré : `inf1.24xlarge` | `p3dn.24xlarge`

Les processeurs AWS Graviton disposent de modes d'économie d'énergie intégrés et fonctionnent à une fréquence fixe. Par conséquent, ils ne permettent pas au système d'exploitation de contrôler les états C et les états P.

Il se peut que vous vouliez changer les paramètres « C-state » ou « P-state » pour améliorer la cohérence des performances du processeur, réduire la latence ou ajuster votre instance pour une charge de travail spécifique. Les paramètres « C-state » ou « P-state » par défaut offre des performances maximales qui sont optimales pour la plupart des charges de travail. Cependant, si votre application tirerait avantage de la latence réduite pour un coût de fréquences simple ou double cœur plus hautes ou des performances cohérentes à des fréquences plus basses au lieu des fréquences Turbo Boost transmises en paquets, pensez à essayer les paramètres « C-state » ou « P-state » qui sont disponibles pour ces instances.

Les sections suivantes décrivent les différentes configurations d'états du processeur et les façons de surveiller les effets de votre configuration. Ces procédures ont été écrites pour et s'appliquent à Amazon Linux. Néanmoins, elles peuvent aussi être adaptées aux autres distributions Linux avec une version noyau Linux de 3.9 ou plus récente. Pour obtenir plus d'informations sur les autres distributions Linux et le contrôle des états du processeur, consultez la documentation spécifique à votre système.

> **Note**
>
> Les exemples sur cette page utilisent l'utilitaire turbostat (qui est disponible sur Amazon Linux par défaut) pour afficher la fréquence du processus et les informations relatives à l'état « C-state » ainsi que la commande stress (qui peut être installée en exécutant sudo yum install -y stress) pour simuler une charge de travail.
> Si la sortie n'affiche pas les informations relatives à l'état « C-state », incluez l'option --debug dans la commande (sudo turbostat --debug stress ***<options>***).

Sommaire

- La meilleure performance avec la fréquence Turbo Boost maximale (p. 664)
- Haute performance et faible latence en limitant les états « C-states » plus profonds (p. 665)
- Performances de base avec les variations les plus faibles (p. 666)

La meilleure performance avec la fréquence Turbo Boost maximale

Il s'agit de la configuration de contrôle d'état du processeur par défaut pour Amazon Linux AMI et il est recommandé pour la plupart des charges de travail. Cette configuration fournit les meilleures performances avec des variations plus faibles. Le fait de permettre aux cœurs inactifs d'entrer dans des états de veille plus profonds offre le dégagement thermique nécessaire aux processeurs simple ou double cœur d'atteindre leur potentiel Turbo Boost maximal.

L'exemple suivant montre une instance `c4.8xlarge` avec deux cœurs qui fonctionnent activement et atteignent la fréquence Turbo Boost maximale de leur processeur.

```
[ec2-user ~]$ sudo turbostat stress -c 2 -t 10
stress: info: [30680] dispatching hogs: 2 cpu, 0 io, 0 vm, 0 hdd
stress: info: [30680] successful run completed in 10s
pk cor CPU    %c0  GHz  TSC SMI    %c1    %c3    %c6    %c7   %pc2   %pc3   %pc6   %pc7
 Pkg_W RAM_W PKG_% RAM_%
             5.54 3.44 2.90   0   9.18   0.00  85.28   0.00   0.00   0.00   0.00   0.00
 94.04 32.70 54.18  0.00
 0   0   0   0.12 3.26 2.90   0   3.61   0.00  96.27   0.00   0.00   0.00   0.00   0.00
 48.12 18.88 26.02  0.00
 0   0  18   0.12 3.26 2.90   0   3.61
 0   1   1   0.12 3.26 2.90   0   4.11   0.00  95.77   0.00
 0   1  19   0.13 3.27 2.90   0   4.11
 0   2   2   0.13 3.28 2.90   0   4.45   0.00  95.42   0.00
 0   2  20   0.11 3.27 2.90   0   4.47
 0   3   3   0.05 3.42 2.90   0  99.91   0.00   0.05   0.00
 0   3  21  97.84 3.45 2.90   0   2.11
...
 1   1  10   0.06 3.33 2.90   0  99.88   0.01   0.06   0.00
 1   1  28  97.61 3.44 2.90   0   2.32
...
10.002556 sec
```

Dans cet exemple, vCPU 21 et 28 fonctionnent à leur fréquence Turbo Boost maximale, car les autres cœurs sont entrés dans l'état de veille `C6` pour économiser de l'énergie et offrir une marge de puissance et un dégagement thermique pour les cœurs en fonctionnement. vCPU 3 et 10 (chacun partageant un cœur de processeur avec vCPU 21 et 28) possèdent l'état `C1` et attendent des instructions.

Dans l'exemple suivant, les 18 cœurs fonctionnent activement. Il n'existe donc aucune marge pour la fréquence Turbo Boost maximale, mais ils sont tous exécutés à la vitesse « Turbo Boost » de 3.2 GHz lorsque tous les cœurs sont utilisés.

```
[ec2-user ~]$ sudo turbostat stress -c 36 -t 10
stress: info: [30685] dispatching hogs: 36 cpu, 0 io, 0 vm, 0 hdd
stress: info: [30685] successful run completed in 10s
pk cor CPU    %c0  GHz  TSC SMI    %c1    %c3    %c6    %c7   %pc2   %pc3   %pc6   %pc7
 Pkg_W RAM_W PKG_% RAM_%
            99.27 3.20 2.90   0   0.26   0.00   0.47   0.00   0.00   0.00   0.00   0.00
 228.59 31.33 199.26  0.00
 0   0   0  99.08 3.20 2.90   0   0.27   0.01   0.64   0.00   0.00   0.00   0.00   0.00
 114.69 18.55 99.32  0.00
 0   0  18  98.74 3.20 2.90   0   0.62
 0   1   1  99.14 3.20 2.90   0   0.09   0.00   0.76   0.00
 0   1  19  98.75 3.20 2.90   0   0.49
 0   2   2  99.07 3.20 2.90   0   0.10   0.02   0.81   0.00
 0   2  20  98.73 3.20 2.90   0   0.44
 0   3   3  99.02 3.20 2.90   0   0.24   0.00   0.74   0.00
 0   3  21  99.13 3.20 2.90   0   0.13
 0   4   4  99.26 3.20 2.90   0   0.09   0.00   0.65   0.00
 0   4  22  98.68 3.20 2.90   0   0.67
```

```
 0    5    5  99.19 3.20 2.90   0   0.08   0.00   0.73   0.00
 0    5   23  98.58 3.20 2.90   0   0.69
 0    6    6  99.01 3.20 2.90   0   0.11   0.00   0.89   0.00
 0    6   24  98.72 3.20 2.90   0   0.39
...
```

Haute performance et faible latence en limitant les états « C-states » plus profonds

Les états « C-states » contrôlent les niveaux de veille dans lesquels un cœur peut entrer lorsqu'il est inutilisé. Il se peut que vous vouliez contrôler les états « C-states » pour ajuster la latence de votre système par rapport aux performances. La mise en veille de cœurs prend du temps. Même si un cœur en veille donne plus de marge pour qu'un autre cœur fonctionne à une fréquence plus élevée, ce cœur en veille prend du temps pour se remettre en route et fonctionner. Par exemple, si un cœur qui est assigné à la gestion d'interruptions de paquets est en veille, il se peut que la prise en charge de cette interruption soit retardée. Vous pouvez configurer le système pour qu'il n'utilise pas les états « C-states » plus profonds ce qui réduit la latence de réaction du processeur, mais également la marge disponible pour la fréquence Turbo Boost des autres cœurs.

Un scénario commun pour la désactivation d'états de veille plus profonds est une application de la base de données Redis qui stocke la base de données dans la mémoire système pour un temps de réponse aux requêtes le plus rapide possible.

Pour limiter les états de veille plus profonds sur Amazon Linux 2

1. Ouvrez le fichier `/etc/default/grub` avec l'éditeur de votre choix.

   ```
   [ec2-user ~]$ sudo vim /etc/default/grub
   ```

2. Modifiez la ligne `GRUB_CMDLINE_LINUX_DEFAULT` et ajoutez l'option `intel_idle.max_cstate=1` pour définir `C1` comme l'état « C-state » le plus profond pour les cœurs inutilisés.

   ```
   GRUB_CMDLINE_LINUX_DEFAULT="console=tty0 console=ttyS0,115200n8 net.ifnames=0
    biosdevname=0 nvme_core.io_timeout=4294967295 intel_idle.max_cstate=1"
   GRUB_TIMEOUT=0
   ```

3. Enregistrez le fichier et quittez votre éditeur.
4. Exécutez la commande suivante pour recréer la configuration du démarrage.

   ```
   [ec2-user ~]$ sudo grub2-mkconfig -o /boot/grub2/grub.cfg
   ```

5. Redémarrez votre instance pour activer la nouvelle option noyau.

   ```
   [ec2-user ~]$ sudo reboot
   ```

Pour limiter les états de veille plus profonds sur Amazon Linux AMI

1. Ouvrez le fichier `/boot/grub/grub.conf` avec l'éditeur de votre choix.

   ```
   [ec2-user ~]$ sudo vim /boot/grub/grub.conf
   ```

2. Modifiez la ligne `kernel` de la première entrée et ajoutez l'option `intel_idle.max_cstate=1` pour définir `C1` comme l'état « C-state » le plus profond pour les cœurs inutilisés.

   ```
   # created by imagebuilder
   ```

```
default=0
timeout=1
hiddenmenu

title Amazon Linux 2014.09 (3.14.26-24.46.amzn1.x86_64)
root (hd0,0)
kernel /boot/vmlinuz-3.14.26-24.46.amzn1.x86_64 root=LABEL=/ console=ttyS0
 intel_idle.max_cstate=1
initrd /boot/initramfs-3.14.26-24.46.amzn1.x86_64.img
```

3. Enregistrez le fichier et quittez votre éditeur.
4. Redémarrez votre instance pour activer la nouvelle option noyau.

```
[ec2-user ~]$ sudo reboot
```

L'exemple suivant montre une instance `c4.8xlarge` avec deux cœurs qui fonctionnent activement à la fréquence « Turbo Boost » lorsque tous les cœurs sont utilisés.

```
[ec2-user ~]$ sudo turbostat stress -c 2 -t 10
stress: info: [5322] dispatching hogs: 2 cpu, 0 io, 0 vm, 0 hdd
stress: info: [5322] successful run completed in 10s
pk cor CPU    %c0  GHz  TSC SMI    %c1    %c3    %c6    %c7   %pc2   %pc3   %pc6   %pc7
 Pkg_W RAM_W PKG_% RAM_%
             5.56 3.20 2.90   0  94.44   0.00   0.00   0.00   0.00   0.00   0.00   0.00
 131.90 31.11 199.47  0.00
 0   0   0   0.03 2.08 2.90   0  99.97   0.00   0.00   0.00   0.00   0.00   0.00   0.00
 67.23 17.11 99.76  0.00
 0   0  18   0.01 1.93 2.90   0  99.99
 0   1   1   0.02 1.96 2.90   0  99.98   0.00   0.00   0.00
 0   1  19  99.70 3.20 2.90   0   0.30
...
 1   1  10   0.02 1.97 2.90   0  99.98   0.00   0.00   0.00
 1   1  28  99.67 3.20 2.90   0   0.33
 1   2  11   0.04 2.63 2.90   0  99.96   0.00   0.00   0.00
 1   2  29   0.02 2.11 2.90   0  99.98
...
```

Dans cet exemple, les cœurs pour vCPU 19 et 28 fonctionnent à 3,2 GHz tandis que les autres cœurs possèdent l'état « C-state » `C1` et attendent des instructions. Même si les cœurs en fonctionnement n'atteignent pas leur fréquence Turbo Boost maximale, les cœurs inactifs seront beaucoup plus rapides à répondre aux nouvelles requêtes que s'ils possédaient l'état « C-state » `C6` plus profond.

Performances de base avec les variations les plus faibles

Vous pouvez réduire les variations de la fréquence du processeur avec des états « P-states ». Les états « P-states » contrôlent les performances souhaitées (dans la fréquence de l'UC) à partir d'un cœur. La plupart des charges de travail fonctionnent mieux avec l'état P0 ce qui demande une fréquence Turbo Boost. Cependant, il se peut que vous souhaitiez adapter votre système pour obtenir une performance cohérente plus que transmise en paquets ce qui peut se produire lorsque les fréquences Turbo Boost sont activées.

Les charges de travail Intel Advanced Vector Extensions (AVX ou AVX2) peuvent fonctionner convenablement à des fréquences plus basses et les instructions pour AVX peuvent utiliser plus de puissance. L'exécution du processeur à une fréquence plus basse en désactivant la fréquence Turbo Boost peut réduire la quantité d'énergie utilisée et conserver la cohérence de la vitesse. Pour obtenir plus d'informations sur l'optimisation de la configuration et la charge de travail de votre instance pour AVX, consultez http://www.intel.com/content/dam/www/public/us/en/documents/white-papers/performance-xeon-e5-v3-advanced-vector-extensions-paper.pdf.

Cette section décrit comment limiter les états de veille plus longs et désactiver la fréquence Turbo Boost (en demandant l'état « P-state » `P1`) pour offrir une latence faible et la variation de vitesse du processeur la plus faible pour ces types de charges de travail.

Pour limiter les états de veille plus profonds et désactiver la fréquence Turbo Boost sur Amazon Linux 2

1. Ouvrez le fichier `/etc/default/grub` avec l'éditeur de votre choix.

```
[ec2-user ~]$ sudo vim /etc/default/grub
```

2. Modifiez la ligne `GRUB_CMDLINE_LINUX_DEFAULT` et ajoutez l'option `intel_idle.max_cstate=1` pour définir `C1` comme l'état « C-state » le plus profond pour les cœurs inutilisés.

```
GRUB_CMDLINE_LINUX_DEFAULT="console=tty0 console=ttyS0,115200n8 net.ifnames=0
 biosdevname=0 nvme_core.io_timeout=4294967295 intel_idle.max_cstate=1"
GRUB_TIMEOUT=0
```

3. Enregistrez le fichier et quittez votre éditeur.
4. Exécutez la commande suivante pour recréer la configuration du démarrage.

```
[ec2-user ~]$ grub2-mkconfig -o /boot/grub2/grub.cfg
```

5. Redémarrez votre instance pour activer la nouvelle option noyau.

```
[ec2-user ~]$ sudo reboot
```

6. Lorsque vous avez besoin des variations faibles de la vitesse du processeur que l'état « P-state » `P1` offre, exécutez la commande suivante pour désactiver la fréquence Turbo Boost.

```
[ec2-user ~]$ sudo sh -c "echo 1 > /sys/devices/system/cpu/intel_pstate/no_turbo"
```

7. Lorsque votre charge de travail est terminée, vous pouvez réactiver la fréquence Turbo Boost avec la commande suivante.

```
[ec2-user ~]$ sudo sh -c "echo 0 > /sys/devices/system/cpu/intel_pstate/no_turbo"
```

Pour limiter les états de veille plus profonds et désactiver la fréquence Turbo Boost sur Amazon Linux AMI

1. Ouvrez le fichier `/boot/grub/grub.conf` avec l'éditeur de votre choix.

```
[ec2-user ~]$ sudo vim /boot/grub/grub.conf
```

2. Modifiez la ligne `kernel` de la première entrée et ajoutez l'option `intel_idle.max_cstate=1` pour définir `C1` comme l'état « C-state » le plus profond pour les cœurs inutilisés.

```
# created by imagebuilder
default=0
timeout=1
hiddenmenu

title Amazon Linux 2014.09 (3.14.26-24.46.amzn1.x86_64)
root (hd0,0)
kernel /boot/vmlinuz-3.14.26-24.46.amzn1.x86_64 root=LABEL=/ console=ttyS0
 intel_idle.max_cstate=1
```

```
initrd /boot/initramfs-3.14.26-24.46.amzn1.x86_64.img
```

3. Enregistrez le fichier et quittez votre éditeur.
4. Redémarrez votre instance pour activer la nouvelle option noyau.

```
[ec2-user ~]$ sudo reboot
```

5. Lorsque vous avez besoin des variations faibles de la vitesse du processeur que l'état « P-state » `P1` offre, exécutez la commande suivante pour désactiver la fréquence Turbo Boost.

```
[ec2-user ~]$ sudo sh -c "echo 1 > /sys/devices/system/cpu/intel_pstate/no_turbo"
```

6. Lorsque votre charge de travail est terminée, vous pouvez réactiver la fréquence Turbo Boost avec la commande suivante.

```
[ec2-user ~]$ sudo sh -c "echo 0 > /sys/devices/system/cpu/intel_pstate/no_turbo"
```

L'exemple suivant montre une instance `c4.8xlarge` avec deux vCPU qui fonctionnent activement à la fréquence de base avec aucune fréquence Turbo Boost cœur.

```
[ec2-user ~]$ sudo turbostat stress -c 2 -t 10
stress: info: [5389] dispatching hogs: 2 cpu, 0 io, 0 vm, 0 hdd
stress: info: [5389] successful run completed in 10s
pk cor CPU    %c0  GHz  TSC SMI    %c1    %c3    %c6    %c7   %pc2   %pc3   %pc6   %pc7
 Pkg_W RAM_W PKG_% RAM_%
             5.59 2.90 2.90   0  94.41   0.00   0.00   0.00   0.00   0.00   0.00   0.00
 128.48 33.54 200.00  0.00
 0   0   0   0.04 2.90 2.90   0  99.96   0.00   0.00   0.00   0.00   0.00   0.00   0.00
 65.33 19.02 100.00  0.00
 0   0  18   0.04 2.90 2.90   0  99.96
 0   1   1   0.05 2.90 2.90   0  99.95   0.00   0.00   0.00
 0   1  19   0.04 2.90 2.90   0  99.96
 0   2   2   0.04 2.90 2.90   0  99.96   0.00   0.00   0.00
 0   2  20   0.04 2.90 2.90   0  99.96
 0   3   3   0.05 2.90 2.90   0  99.95   0.00   0.00   0.00
 0   3  21  99.95 2.90 2.90   0   0.05
...
 1   1  28  99.92 2.90 2.90   0   0.08
 1   2  11   0.06 2.90 2.90   0  99.94   0.00   0.00   0.00
 1   2  29   0.05 2.90 2.90   0  99.95
```

Les cœurs pour vCPU 21 et 28 fonctionnent activement à la vitesse du processeur de base de 2,9 GHz, et tous les cœurs inactifs fonctionnent aussi à la vitesse de base dans l'état « C-state » `C1` et sont prêts à accepter les instructions.

Réglage de l'heure pour votre instance Linux

Une référence de temps cohérente et précise est essentielle pour beaucoup de tâches et de processus de serveurs. La plupart des journaux de système comportent un horodatage que vous pouvez utiliser pour déterminer lorsque les problèmes surviennent et l'ordre de ces évènements. Si vous utilisez l'AWS CLI ou un kit SDK AWS pour effectuer des demandes à partir de votre instance, ces outils signent des demandes à votre place. Si la date et l'heure de votre instance ne sont pas correctement définies, il se peut que la date de la signature ne corresponde pas à la date de la demande et que AWS rejette la demande.

Amazon fournit l'Amazon Time Sync Service, qui est accessible à partir de toutes les instances EC2, et est également utilisé par d'autres services AWS. Ce service utilise un ensemble d'horloges de référence atomiques et connectées à des satellites, dans chaque région pour fournir des mesures temporelles précises selon la norme internationale UTC (temps universel coordonné) via le NTP (Network Time

Protocol). Amazon Time Sync Service lisse automatiquement les secondes intercalaires qui sont ajoutées au temps UTC.

Amazon Time Sync Service est disponible via NTP à l'adresse IP `169.254.169.123` pour toute instance s'exécutant dans un VPC. Votre instance ne requiert pas d'accès à Internet et vous n'avez pas à configurer les règles de votre groupe de sécurité ou de votre ACL réseau pour autoriser l'accès. Les dernières versions d'Amazon Linux 2 et des AMI Amazon Linux se synchronisent par défaut avec le service Amazon Time Sync Service.

Utilisez les procédures suivantes pour configurer Amazon Time Sync Service sur votre instance à l'aide du client `chrony`. Vous pouvez également utiliser des sources NTP externes. Pour plus d'informations sur NTP et les sources temporelles publiques consultez http://www.ntp.org/. Une instance a besoin d'accéder à Internet pour faire fonctionner les sources temporelles NTP externes.

Rubriques

- Configuration d'Amazon Time Sync Service sur Amazon Linux AMI (p. 669)
- Configuration d'Amazon Time Sync Service sur Ubuntu (p. 670)
- Configuration d'Amazon Time Sync Service sur SUSE Linux (p. 672)
- Changement du fuseau horaire sur Amazon Linux (p. 672)

Configuration d'Amazon Time Sync Service sur Amazon Linux AMI

> **Note**
>
> Sur Amazon Linux 2, la configuration `chrony` par défaut est déjà configurée pour l'utilisation de l'adresse IP d'Amazon Time Sync Service.

Avec l'Amazon Linux AMI, vous devez modifier le fichier de configuration `chrony` pour ajouter une entrée de serveur pour Amazon Time Sync Service.

Pour configurer votre instance pour qu'elle utilise Amazon Time Sync Service

1. Connectez-vous à votre instance et désinstallez le service NTP.

```
[ec2-user ~]$ sudo yum erase 'ntp*'
```

2. Installez le package `chrony`.

```
[ec2-user ~]$ sudo yum install chrony
```

3. Ouvrez le fichier `/etc/chrony.conf` avec un éditeur de texte (tel que vim ou nano). Vérifiez que le fichier contienne la ligne suivante :

```
server 169.254.169.123 prefer iburst minpoll 4 maxpoll 4
```

   Si la ligne est présente, alors Amazon Time Sync Service est déjà configuré et vous pouvez passer à l'étape suivante. Si ce n'est pas le cas, ajoutez la ligne après toute autre instruction `server` ou `pool` déjà présente dans le fichier, puis enregistrer les changements.

4. Relancez le démon `chrony` (`chronyd`).

```
[ec2-user ~]$ sudo service chronyd restart
```

```
Starting chronyd:                                          [  OK  ]
```

Note

Sur RHEL ou CentOS (jusqu'à la version 6), le nom du service est `chrony` au lieu de `chronyd`.

5. Utilisez la commande `chkconfig` pour configurer `chronyd` afin de lancer ce service à chaque démarrage système.

```
[ec2-user ~]$ sudo chkconfig chronyd on
```

6. Vérifiez que `chrony` utilise l'adresse IP `169.254.169.123` pour synchroniser le temps.

```
[ec2-user ~]$ chronyc sources -v
```

```
210 Number of sources = 7

  .-- Source mode  '^' = server, '=' = peer, '#' = local clock.
 / .- Source state '*' = current synced, '+' = combined , '-' = not combined,
| /   '?' = unreachable, 'x' = time may be in error, '~' = time too variable.
||                                                 .- xxxx [ yyyy ] +/- zzzz
||      Reachability register (octal) -.           |  xxxx = adjusted offset,
||      Log2(Polling interval) --.      |          |  yyyy = measured offset,
||                                \     |          |  zzzz = estimated error.
||                                 |    |           \
MS Name/IP address         Stratum Poll Reach LastRx Last sample
===============================================================================
^* 169.254.169.123               3   6    17    43    -30us[ -226us] +/-  287us
^- ec2-12-34-231-12.eu-west>     2   6    17    43   -388us[ -388us] +/-   11ms
^- tshirt.heanet.ie              1   6    17    44   +178us[  +25us] +/- 1959us
^? tbag.heanet.ie                0   6     0     -     +0ns[   +0ns] +/-    0ns
^? bray.walcz.net                0   6     0     -     +0ns[   +0ns] +/-    0ns
^? 2a05:d018:c43:e312:ce77:>     0   6     0     -     +0ns[   +0ns] +/-    0ns
^? 2a05:d018:dab:2701:b70:b>     0   6     0     -     +0ns[   +0ns] +/-    0ns
```

Dans le résultat retourné, `^*` indique la source de temps préférée.

7. Vérifiez les métriques de synchronisation du temps présentées par `chrony`.

```
[ec2-user ~]$ chronyc tracking
```

```
Reference ID    : A9FEA97B (169.254.169.123)
Stratum         : 4
Ref time (UTC)  : Wed Nov 22 13:18:34 2017
System time     : 0.000000626 seconds slow of NTP time
Last offset     : +0.002852759 seconds
RMS offset      : 0.002852759 seconds
Frequency       : 1.187 ppm fast
Residual freq   : +0.020 ppm
Skew            : 24.388 ppm
Root delay      : 0.000504752 seconds
Root dispersion : 0.001112565 seconds
Update interval : 64.4 seconds
Leap status     : Normal
```

Configuration d'Amazon Time Sync Service sur Ubuntu

Vous devez modifier le fichier de configuration `chrony` pour ajouter une entrée de serveur pour Amazon Time Sync Service.

Pour configurer votre instance pour qu'elle utilise Amazon Time Sync Service

1. Connectez-vous à votre instance et utilisez `apt` pour installer le package `chrony`.

```
ubuntu:~$ sudo apt install chrony
```

   Note

   Si nécessaire, mettez d'abord à jour votre instance en exécutant `sudo apt update`.

2. Ouvrez le fichier `/etc/chrony/chrony.conf` avec un éditeur de texte (tel que vim ou nano). Ajoutez la ligne suivante avant toute autre instruction `server` ou `pool` déjà présente dans le fichier, puis enregistrer les changements :

```
server 169.254.169.123 prefer iburst minpoll 4 maxpoll 4
```

3. Redémarrez le service `chrony`.

```
ubuntu:~$ sudo /etc/init.d/chrony restart
```

```
[ ok ] Restarting chrony (via systemctl): chrony.service.
```

4. Vérifiez que `chrony` utilise l'adresse IP `169.254.169.123` pour synchroniser le temps.

```
ubuntu:~$ chronyc sources -v
```

```
210 Number of sources = 7

  .-- Source mode  '^' = server, '=' = peer, '#' = local clock.
 / .- Source state '*' = current synced, '+' = combined , '-' = not combined,
| /   '?' = unreachable, 'x' = time may be in error, '~' = time too variable.
||                                                 .- xxxx [ yyyy ] +/- zzzz
||      Reachability register (octal) -.           |  xxxx = adjusted offset,
||      Log2(Polling interval) --.      |          |  yyyy = measured offset,
||                                   \     |          |  zzzz = estimated error.
||                                    |    |           \
MS Name/IP address         Stratum Poll Reach LastRx Last sample
===============================================================================
^* 169.254.169.123               3   6    17    12    +15us[  +57us] +/-  320us
^- tbag.heanet.ie                1   6    17    13  -3488us[-3446us] +/- 1779us
^- ec2-12-34-231-12.eu-west-     2   6    17    13   +893us[ +935us] +/- 7710us
^? 2a05:d018:c43:e312:ce77:6     0   6     0   10y     +0ns[   +0ns] +/-    0ns
^? 2a05:d018:d34:9000:d8c6:5     0   6     0   10y     +0ns[   +0ns] +/-    0ns
^? tshirt.heanet.ie              0   6     0   10y     +0ns[   +0ns] +/-    0ns
^? bray.walcz.net                0   6     0   10y     +0ns[   +0ns] +/-    0ns
```

   Dans le résultat retourné, `^*` indique la source de temps préférée.

5. Vérifiez les métriques de synchronisation du temps présentées par `chrony`.

```
ubuntu:~$ chronyc tracking
```

```
Reference ID    : 169.254.169.123 (169.254.169.123)
Stratum         : 4
Ref time (UTC)  : Wed Nov 29 07:41:57 2017
System time     : 0.000000011 seconds slow of NTP time
Last offset     : +0.000041659 seconds
RMS offset      : 0.000041659 seconds
```

```
Frequency       : 10.141 ppm slow
Residual freq   : +7.557 ppm
Skew            : 2.329 ppm
Root delay      : 0.000544 seconds
Root dispersion : 0.000631 seconds
Update interval : 2.0 seconds
Leap status     : Normal
```

Configuration d'Amazon Time Sync Service sur SUSE Linux

Installez chrony depuis https://software.opensuse.org/package/chrony.

Ouvrez le fichier `/etc/chrony.conf` avec un éditeur de texte (tel que vim ou nano). Vérifiez que le fichier contient la ligne suivante :

```
server 169.254.169.123 prefer iburst minpoll 4 maxpoll 4
```

Si la ligne n'est pas présente, vous pouvez l'ajouter manuellement. Placez en commentaire les autres lignes sur le serveur ou le groupe (pool). Ouvrez yast et activez le service chrony.

Changement du fuseau horaire sur Amazon Linux

Les instances Amazon Linux sont définies sur le fuseau horaire UTC (temps universel coordonné) par défaut. Vous pouvez modifier l'heure d'une instance à l'heure locale ou à un autre fuseau horaire de votre réseau.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

Pour changer le fuseau horaire sur une instance

1. Identifiez le fuseau horaire à utiliser sur l'instance. Le répertoire `/usr/share/zoneinfo` contient une hiérarchie de fichiers de données sur le fuseau horaire. Parcourez la structure du répertoire à l'endroit où vous recherchez un fichier pour votre fuseau horaire.

   ```
   [ec2-user ~]$ ls /usr/share/zoneinfo
   Africa      Chile    GB         Indian       Mideast   posixrules  US
   America     CST6CDT  GB-Eire    Iran         MST       PRC         UTC
   Antarctica  Cuba     GMT        iso3166.tab  MST7MDT   PST8PDT     WET
   Arctic      EET      GMT0       Israel       Navajo    right       W-SU
   ...
   ```

   Certaines entrées à cet endroit sont des répertoires (comme `America`), et ces répertoires contiennent des fichiers sur le fuseau horaire pour des villes spécifiques. Recherchez votre ville (ou une ville de votre fuseau horaire) à utiliser pour l'instance.
2. Mettez à jour le fichier `/etc/sysconfig/clock` avec le nouveau fuseau horaire. Dans cet exemple, nous utilisons le fichier de données de fuseau horaire pour Los Angeles, `/usr/share/zoneinfo/America/Los_Angeles`.

   a. Ouvrez le fichier `/etc/sysconfig/clock` avec votre éditeur de texte préféré (comme vim ou nano). Vous devez utiliser sudo avec la commande de votre éditeur, car `/etc/sysconfig/clock` est détenu par `root`.

      ```
      [ec2-user ~]$ sudo nano /etc/sysconfig/clock
      ```

b. Recherchez l'entrée `ZONE` et remplacez-la par le fichier sur le fuseau horaire (en omettant la section `/usr/share/zoneinfo` du chemin). Par exemple, pour passer au fuseau horaire de Los Angeles, remplacez l'entrée `ZONE` par ce qui suit:

```
ZONE="America/Los_Angeles"
```

> **Note**
>
> Ne remplacez pas l'entrée `UTC=true` par une autre valeur. Cette entrée concerne l'horloge matérielle et n'a pas besoin d'être corrigée lorsque vous définissez un autre fuseau horaire sur votre instance.

c. Enregistrez le fichier et quittez l'éditeur de texte.

3. Créez un lien symbolique entre `/etc/localtime` et le fichier de fuseau horaire pour que l'instance trouve le fichier de fuseau horaire lorsqu'il fait référence à des informations sur l'heure locale.

```
[ec2-user ~]$ sudo ln -sf /usr/share/zoneinfo/America/Los_Angeles /etc/localtime
```

4. Redémarrez le système pour récupérer les informations sur la nouvelle heure locale dans l'ensemble des services et des applications.

```
[ec2-user ~]$ sudo reboot
```

5. (Facultatif) Vérifiez que le fuseau horaire actuel est mis à jour vers le nouveau fuseau horaire à l'aide de la commande date. Le fuseau horaire actuel apparaît dans la sortie. Dans l'exemple suivant, le fuseau horaire actuel est PDT, c'est-à-dire le fuseau horaire de Los Angeles.

```
[ec2-user ~]$ date
Sun Aug 16 05:45:16 PDT 2020
```

Optimisations des options d'UC

Les instances Amazon EC2 prennent en charge le multithreading, qui permet l'exécution simultanée de plusieurs threads sur un seul cœur d'UC. Chaque thread est représenté comme UC virtuelle (vCPU) sur l'instance. Une instance possède un certain nombre par défaut de cœurs d'UC, qui varie en fonction du type d'instance. Par exemple, un type d'instance `m5.xlarge` a deux cœurs d'UC et deux threads par cœur par défaut, quatre vCPU— au total.

> **Note**
>
> Chaque vCPU est un thread d'un cœur de processeur, à l'exception des instances T2 et des instances alimentées par des processeurs AWS Graviton2.

Dans la plupart des cas, il y a un type d'instance Amazon EC2 qui possède une combinaison de mémoire et d'un certain nombre de vCPU pour convenir à vos charges de travail. Cependant, vous pouvez spécifier les options d'UC suivantes pour optimiser votre instance pour des besoins métier ou des charges de travail spécifiques :

- Nombre de cœurs d'UC : vous pouvez personnaliser le nombre de cœurs d'UC pour l'instance. Vous pourriez agir ainsi pour optimiser potentiellement les coûts de licence de vos logiciels avec une instance ayant une quantité suffisante de RAM pour les charges de travail exigeantes en mémoire, mais moins de cœurs d'UC.
- Threads per core (Threads par cœur) : vous pouvez désactiver le multithreading en spécifiant un seul thread par cœur d'UC. Vous pourriez agir ainsi pour certaines charges de travail, telles que les charges de travail de calcul haute performance (HPC).

Vous pouvez spécifier ces options d'UC lors du lancement de l'instance. Il n'y a pas de frais supplémentaires ou réduits pour spécifier des options d'UC. Vous êtes facturé de la même façon que pour les instances lancées avec les options d'UC par défaut.

Table des matières

- Règles pour spécifier les options d'UC (p. 674)
- Cœurs d'UC et threads par cœur d'UC par type d'instance (p. 674)
- Spécification des options d'UC pour votre instance (p. 687)
- Affichage des options d'UC pour votre instance (p. 688)

Règles pour spécifier les options d'UC

Pour spécifier les options d'UC pour votre instance, soyez conscient des règles suivantes :

- Les options d'UC ne peuvent être spécifiées que pendant un lancement d'instance et ne peuvent pas être modifiées après le lancement.
- Lorsque vous lancez une instance, vous devez spécifier le nombre de cœurs d'UC et de threads par cœur dans la demande. Pour obtenir des exemples de requête, consultez Spécification des options d'UC pour votre instance (p. 687).
- Le nombre de vCPU pour l'instance est égal au nombre de cœurs d'UC multiplié par le nombre de threads par cœur. Pour spécifier un nombre personnalisé de vCPU, vous devez spécifier un nombre valide de cœurs d'UC et de threads par cœur pour le type d'instance. Vous ne pouvez pas dépasser le nombre de vCPU par défaut pour l'instance. Pour plus d'informations, consultez Cœurs d'UC et threads par cœur d'UC par type d'instance (p. 674).
- Pour désactiver le multithreading, spécifiez un seul thread par cœur.
- Si vous modifiez le type d'une instance (p. 313) existante, les options d'UC se changent automatiquement en options d'UC par défaut pour le nouveau type d'instance.
- Les options d'UC spécifiées sont conservées après que vous arrêtez, démarrez ou redémarrez une instance.

Cœurs d'UC et threads par cœur d'UC par type d'instance

Les tableaux suivants répertorient les types d'instance qui prennent en charge la spécification des options d'UC. Pour chaque type, le tableau indique les valeurs par défaut, ainsi que le nombre pris en charge de cœurs d'UC et de threads par cœur.

Instances à calcul accéléré

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`f1.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`f1.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`f1.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`g3.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`g3.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2
`g3.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`g3s.xlarge`	4	2	2	1, 2	1, 2
`g4dn.xlarge`	4	2	2	1, 2	1, 2
`g4dn.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`g4dn.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`g4dn.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2
`g4dn.12xlarge`	48	24	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`g4dn.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`inf1.xlarge`	4	2	2	2	1, 2
`inf1.2xlarge`	8	4	2	2, 4	1, 2
`inf1.6xlarge`	24	12	2	2, 4, 6, 8, 10, 12	1, 2
`inf1.24xlarge`	96	48	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`p2.xlarge`	4	2	2	1, 2	1, 2
`p2.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`p2.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`p3.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`p3.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2
`p3.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`p3dn.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2

Instances de calcul optimisé

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`c4.large`	2	1	2	1	1, 2
`c4.xlarge`	4	2	2	1, 2	1, 2
`c4.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`c4.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`c4.8xlarge`	36	18	2	2, 4, 6, 8, 10, 12, 14, 16, 18	1, 2
`c5.large`	2	1	2	1	1, 2
`c5.xlarge`	4	2	2	2	1, 2
`c5.2xlarge`	8	4	2	2, 4	1, 2
`c5.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`c5.9xlarge`	36	18	2	2, 4, 6, 8, 10, 12, 14, 16, 18	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`c5.12xlarge`	48	24	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`c5.18xlarge`	72	36	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36	1, 2
`c5.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`c5a.large`	2	1	2	1	1, 2
`c5a.xlarge`	4	2	2	1, 2	1, 2
`c5a.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`c5a.4xlarge`	16	8	2	1, 2, 3, 4, 8	1, 2
`c5a.8xlarge`	32	16	2	1, 2, 3, 4, 8, 12, 16	1, 2
`c5a.12xlarge`	48	24	2	1, 2, 3, 4, 8, 12, 16, 20, 24	1, 2
`c5a.16xlarge`	64	32	2	1, 2, 3, 4, 8, 12, 16, 20, 24, 28, 32	1, 2
`c5a.24xlarge`	96	48	2	1, 2, 3, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48	1, 2
`c5ad.large`	2	1	2	1	1, 2
`c5ad.xlarge`	4	2	2	1, 2	1, 2
`c5ad.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`c5ad.4xlarge`	16	8	2	1, 2, 3, 4, 8	1, 2
`c5ad.8xlarge`	32	16	2	1, 2, 3, 4, 8, 12, 16	1, 2
`c5ad.12xlarge`	48	24	2	1, 2, 3, 4, 8, 12, 16, 20, 24	1, 2
`c5ad.16xlarge`	64	32	2	1, 2, 3, 4, 8, 12, 16, 20, 24, 28, 32	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`c5ad.24xlarge`	96	48	2	1, 2, 3, 4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48	1, 2
`c5d.large`	2	1	2	1	1, 2
`c5d.xlarge`	4	2	2	2	1, 2
`c5d.2xlarge`	8	4	2	2, 4	1, 2
`c5d.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`c5d.9xlarge`	36	18	2	2, 4, 6, 8, 10, 12, 14, 16, 18	1, 2
`c5d.12xlarge`	48	24	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`c5d.18xlarge`	72	36	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36	1, 2
`c5d.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`c5n.large`	2	1	2	1	1, 2
`c5n.xlarge`	4	2	2	2	1, 2
`c5n.2xlarge`	8	4	2	2, 4	1, 2
`c5n.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`c5n.9xlarge`	36	18	2	2, 4, 6, 8, 10, 12, 14, 16, 18	1, 2
`c5n.18xlarge`	72	36	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36	1, 2

Instances à usage général

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`m5.large`	2	1	2	1	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`m5.xlarge`	4	2	2	2	1, 2
`m5.2xlarge`	8	4	2	2, 4	1, 2
`m5.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`m5.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`m5.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`m5.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`m5.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`m5a.large`	2	1	2	1	1, 2
`m5a.xlarge`	4	2	2	2	1, 2
`m5a.2xlarge`	8	4	2	2, 4	1, 2
`m5a.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`m5a.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`m5a.12xlarge`	48	24	2	6, 12, 18, 24	1, 2
`m5a.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`m5a.24xlarge`	96	48	2	12, 18, 24, 36, 48	1, 2
`m5ad.large`	2	1	2	1	1, 2
`m5ad.xlarge`	4	2	2	2	1, 2
`m5ad.2xlarge`	8	4	2	2, 4	1, 2
`m5ad.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`m5ad.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`m5ad.12xlarge`	48	24	2	6, 12, 18, 24	1, 2
`m5ad.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`m5ad.24xlarge`	96	48	2	12, 18, 24, 36, 48	1, 2
`m5d.large`	2	1	2	1	1, 2
`m5d.xlarge`	4	2	2	2	1, 2
`m5d.2xlarge`	8	4	2	2, 4	1, 2
`m5d.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`m5d.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`m5d.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`m5d.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`m5d.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`m5dn.large`	2	1	2	1	1, 2
`m5dn.xlarge`	4	2	2	2	1, 2
`m5dn.2xlarge`	8	4	2	2, 4	1, 2
`m5dn.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`m5dn.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`m5dn.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`m5dn.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`m5dn.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`m5n.large`	2	1	2	1	1, 2
`m5n.xlarge`	4	2	2	2	1, 2
`m5n.2xlarge`	8	4	2	2, 4	1, 2
`m5n.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`m5n.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`m5n.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`m5n.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`m5n.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`t3.nano`	2	1	2	1	1, 2
`t3.micro`	2	1	2	1	1, 2
`t3.small`	2	1	2	1	1, 2
`t3.medium`	2	1	2	1	1, 2
`t3.large`	2	1	2	1	1, 2
`t3.xlarge`	4	2	2	2	1, 2
`t3.2xlarge`	8	4	2	2, 4	1, 2
`t3a.nano`	2	1	2	1	1, 2
`t3a.micro`	2	1	2	1	1, 2
`t3a.small`	2	1	2	1	1, 2
`t3a.medium`	2	1	2	1	1, 2
`t3a.large`	2	1	2	1	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`t3a.xlarge`	4	2	2	2	1, 2
`t3a.2xlarge`	8	4	2	2, 4	1, 2

Instances de mémoire optimisée

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`r4.large`	2	1	2	1	1, 2
`r4.xlarge`	4	2	2	1, 2	1, 2
`r4.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`r4.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`r4.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2
`r4.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`r5.large`	2	1	2	1	1, 2
`r5.xlarge`	4	2	2	2	1, 2
`r5.2xlarge`	8	4	2	2, 4	1, 2
`r5.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`r5.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`r5.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`r5.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`r5.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`r5a.large`	2	1	2	1	1, 2
`r5a.xlarge`	4	2	2	2	1, 2
`r5a.2xlarge`	8	4	2	2, 4	1, 2
`r5a.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`r5a.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`r5a.12xlarge`	48	24	2	6, 12, 18, 24	1, 2
`r5a.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`r5a.24xlarge`	96	48	2	12, 18, 24, 36, 48	1, 2
`r5ad.large`	2	1	2	1	1, 2
`r5ad.xlarge`	4	2	2	2	1, 2
`r5ad.2xlarge`	8	4	2	2, 4	1, 2
`r5ad.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`r5ad.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`r5ad.12xlarge`	48	24	2	6, 12, 18, 24	1, 2
`r5ad.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`r5ad.24xlarge`	96	48	2	12, 18, 24, 36, 48	1, 2
`r5d.large`	2	1	2	1	1, 2
`r5d.xlarge`	4	2	2	2	1, 2
`r5d.2xlarge`	8	4	2	2, 4	1, 2
`r5d.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`r5d.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`r5d.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`r5d.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`r5d.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`r5dn.large`	2	1	2	1	1, 2
`r5dn.xlarge`	4	2	2	2	1, 2
`r5dn.2xlarge`	8	4	2	2, 4	1, 2
`r5dn.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`r5dn.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`r5dn.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`r5dn.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`r5dn.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`r5n.large`	2	1	2	1	1, 2
`r5n.xlarge`	4	2	2	2	1, 2
`r5n.2xlarge`	8	4	2	2, 4	1, 2
`r5n.4xlarge`	16	8	2	2, 4, 6, 8	1, 2
`r5n.8xlarge`	32	16	2	2, 4, 6, 8, 10, 12, 14, 16	1, 2
`r5n.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`r5n.16xlarge`	64	32	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`r5n.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2
`x1.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`x1.32xlarge`	128	64	2	4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60, 64	1, 2
`x1e.xlarge`	4	2	2	1, 2	1, 2
`x1e.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`x1e.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`x1e.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2
`x1e.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`x1e.32xlarge`	128	64	2	4, 8, 12, 16, 20, 24, 28, 32, 36, 40, 44, 48, 52, 56, 60, 64	1, 2
`z1d.large`	2	1	2	1	1, 2
`z1d.xlarge`	4	2	2	2	1, 2
`z1d.2xlarge`	8	4	2	2, 4	1, 2
`z1d.3xlarge`	12	6	2	2, 4, 6	1, 2
`z1d.6xlarge`	24	12	2	2, 4, 6, 8, 10, 12	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`z1d.12xlarge`	48	24	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2

Instances de stockage optimisé

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`d2.xlarge`	4	2	2	1, 2	1, 2
`d2.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`d2.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`d2.8xlarge`	36	18	2	2, 4, 6, 8, 10, 12, 14, 16, 18	1, 2
`h1.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`h1.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`h1.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2
`h1.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`i3.large`	2	1	2	1	1, 2
`i3.xlarge`	4	2	2	1, 2	1, 2
`i3.2xlarge`	8	4	2	1, 2, 3, 4	1, 2
`i3.4xlarge`	16	8	2	1, 2, 3, 4, 5, 6, 7, 8	1, 2
`i3.8xlarge`	32	16	2	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16	1, 2
`i3.16xlarge`	64	32	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32	1, 2
`i3en.large`	2	1	2	1	1, 2

Type d'instance	vCPU par défaut	Cœurs d'UC par défaut	Threads par défaut par cœur	Nombre valide de cœurs d'UC	Nombre valide de threads par cœur
`i3en.xlarge`	4	2	2	2	1, 2
`i3en.2xlarge`	8	4	2	2, 4	1, 2
`i3en.3xlarge`	12	6	2	2, 4, 6	1, 2
`i3en.6xlarge`	24	12	2	2, 4, 6, 8, 10, 12	1, 2
`i3en.12xlarge`	48	24	2	2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24	1, 2
`i3en.24xlarge`	96	48	2	4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48	1, 2

Spécification des options d'UC pour votre instance

Vous pouvez spécifier les options d'UC lors du lancement de l'instance. Les exemples suivants concernent un type d'instance `r4.4xlarge`, qui possède les valeurs par défaut suivantes (p. 682) :

- Cœurs d'UC par défaut : 8
- Threads par défaut par cœur : 2
- vCPU par défaut : 16 (8*2)
- Nombre valide de cœurs d'UC : 1, 2, 3, 4, 5, 6, 7, 8
- Nombre valide de threads par cœur : 1, 2

Désactivation du multithreading

Pour désactiver le multithreading, spécifiez un seul thread par cœur.

Désactiver le multithreading pendant le lancement d'une instance (console)

1. Suivez la procédure Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).
2. Sur la page Configurer les détails de l'instance, sous CPU options (Options d'UC), choisissez Specify CPU options (Spécifier les options d'UC).
3. Sous Core count (Nombre de cœurs), choisissez le nombre de cœurs d'UC requis. Dans cet exemple, pour spécifier le nombre de cœurs d'UC par défaut pour une instance `r4.4xlarge`, choisissez `8`.
4. Pour désactiver le multithreading, sous Threads per core (Threads par cœur), sélectionnez 1.
5. Continuez comme indiqué par l'assistant. Lorsque vous avez terminé de vérifier vos options sur la page Examiner le lancement de l'instance, choisissez Lancer. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Désactiver le multithreading pendant le lancement d'une instance (AWS CLI)

Utilisez la commande run-instances de l'AWS CLI et spécifiez la valeur `1` pour `ThreadsPerCore` pour le paramètre `--cpu-options`. Pour `CoreCount`, spécifiez le nombre de cœurs d'UC. Dans cet exemple, pour spécifier le nombre de cœurs d'UC par défaut pour une instance `r4.4xlarge`, spécifiez la valeur `8`.

```
aws ec2 run-instances --image-id ami-1a2b3c4d --instance-type r4.4xlarge --cpu-options
 "CoreCount=8,ThreadsPerCore=1" --key-name MyKeyPair
```

Définition d'un nombre de vCPU personnalisé

Vous pouvez personnaliser le nombre de cœurs d'UC et de threads par cœur pour l'instance.

Pour spécifier un nombre personnalisé de vCPU lors du lancement de l'instance (console)

L'exemple suivant lance une instance `r4.4xlarge` avec six vCPU.

1. Suivez la procédure Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).
2. Sur la page Configurer les détails de l'instance, sous CPU options (Options d'UC), choisissez Specify CPU options (Spécifier les options d'UC).
3. Pour obtenir six vCPU, spécifiez trois cœurs d'UC et deux threads par cœur, comme suit :
   - Sous Core count (Nombre de cœurs), choisissez 3.
   - Sous Threads per core (Threads par cœur), choisissez 2.
4. Continuez comme indiqué par l'assistant. Lorsque vous avez terminé de vérifier vos options sur la page Examiner le lancement de l'instance, choisissez Lancer. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Pour spécifier un nombre personnalisé de vCPU lors du lancement de l'instance (AWS CLI)

L'exemple suivant lance une instance `r4.4xlarge` avec six vCPU.

Utilisez la commande run-instances de l'AWS CLI et spécifiez le nombre de cœurs d'UC et le nombre de threads dans le paramètre `--cpu-options`. Vous pouvez spécifier trois cœurs d'UC et deux thread par cœur pour obtenir six vCPU.

```
aws ec2 run-instances --image-id ami-1a2b3c4d --instance-type r4.4xlarge --cpu-options
 "CoreCount=3,ThreadsPerCore=2" --key-name MyKeyPair
```

Vous pouvez aussi spécifier six cœurs d'UC et un thread par cœur (désactivez le multithreading) pour obtenir six vCPU :

```
aws ec2 run-instances --image-id ami-1a2b3c4d --instance-type r4.4xlarge --cpu-options
 "CoreCount=6,ThreadsPerCore=1" --key-name MyKeyPair
```

Affichage des options d'UC pour votre instance

Vous pouvez afficher les options d'UC pour une instance existante dans la console Amazon EC2 ou en décrivant l'instance à l'aide de l'AWS CLI.

Nouvelle console

Pour afficher les options d'UC d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation sur la gauche, choisissez Instances, puis sélectionnez l'instance.
3. Sous l'onglet Détails, sous Hôte et groupe de placement, recherchez Nombre de vCPU.
4. Pour afficher le nombre de cœurs et les threads par cœur, choisissez la valeur de Nombre de vCPU.

Ancienne console

Pour afficher les options d'UC d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation sur la gauche, choisissez Instances, puis sélectionnez l'instance.
3. Choisissez Description et recherchez Nombre de vCPU.
4. Pour afficher le nombre de cœurs et les threads par cœur, choisissez la valeur de Nombre de vCPU.

Pour afficher les options d'UC pour une instance (AWS CLI)

Utilisez la commande describe-instances.

```
aws ec2 describe-instances --instance-ids i-123456789abcde123
```

```
...
    "Instances": [
        {
            "Monitoring": {
                "State": "disabled"
            },
            "PublicDnsName": "ec2-198-51-100-5.eu-central-1.compute.amazonaws.com",
            "State": {
                "Code": 16,
                "Name": "running"
            },
            "EbsOptimized": false,
            "LaunchTime": "2018-05-08T13:40:33.000Z",
            "PublicIpAddress": "198.51.100.5",
            "PrivateIpAddress": "172.31.2.206",
            "ProductCodes": [],
            "VpcId": "vpc-1a2b3c4d",
            "CpuOptions": {
                "CoreCount": 34,
                "ThreadsPerCore": 1
            },
            "StateTransitionReason": "",
            ...
        }
    ]
...
```

Dans le résultat retourné, le champ `CoreCount` indique le nombre de cœurs pour l'instance. Le champ `ThreadsPerCore` indique le nombre de threads par cœur.

Vous pouvez aussi vous connecter à votre instance et utiliser le , un outil tel que lscpu pour afficher les informations d'UC pour votre instance.

Vous pouvez utiliser AWS Config pour enregistrer, évaluer et auditer les changements de configuration des instances, y compris celles qui sont terminées. Pour plus d'informations, consultez Démarrez avec AWS Config dans le AWS Config Developer Guide.

Changement du nom d'hôte de votre instance Amazon Linux

Lorsque vous lancez une instance, elle reçoit un nom d'hôte qui est une forme d'adresse IPv4 privée interne. Un nom DNS privé Amazon EC2 classique ressemble à `ip-12-34-56-78.us-west-2.compute.internal`, dans lequel le nom se compose du domaine interne, du service (dans ce cas, `compute`), de la région et d'une forme d'adresse IPv4 privée. Une partie de ce nom d'hôte est affiché sur l'invite shell lorsque vous vous connectez dans votre instance (par exemple, `ip-12-34-56-78`). Chaque fois que vous arrêtez et relancez votre instance Amazon EC2 (à moins que vous n'utilisiez une adresse IP Elastic), l'adresse IPv4 publique change, au même titre que le nom DNS public, le nom d'hôte du système et l'invite shell.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

Changement du nom d'hôte du système

Si vous avez un nom DNS public enregistré pour l'adresse IP de votre instance (comme `webserver.mydomain.com`), vous pouvez régler le nom d'hôte du système pour que votre instance s'identifie comme une partie de ce domaine. Ceci modifie également l'invite prompt pour qu'il affiche la première portion de ce nom au lieu du nom d'hôte fourni par AWS (par exemple, `ip-12-34-56-78`). Si vous n'avez pas de nom DNS public enregistré, vous pouvez toujours changer le nom d'hôte, mais le processus est un peu différent.

Pour remplacer le nom d'hôte du système par un nom DNS public

Suivez cette procédure si vous avez déjà un nom DNS public enregistré.

1. - Pour Amazon Linux 2 : Utilisez la commande hostnamectl afin de définir votre nom d'hôte pour refléter le nom de domaine complet (comme **`webserver.mydomain.com`**).

     ```
     [ec2-user ~]$ sudo hostnamectl set-hostname webserver.mydomain.com
     ```

   - Pour Amazon Linux AMI : Sur votre instance, ouvrez le fichier de configuration `/etc/sysconfig/network` dans votre éditeur de texte et modifiez l'entrée `HOSTNAME` pour refléter le nom de domaine complet (comme **`webserver.mydomain.com`**).

     ```
     HOSTNAME=webserver.mydomain.com
     ```

2. Redémarrez l'instance pour récupérer le nouveau nom d'hôte.

   ```
   [ec2-user ~]$ sudo reboot
   ```

   Vous pouvez également redémarrer à l'aide de la console Amazon EC2 (sur la page Instances, sélectionnez l'instance et choisissez État de l'instance, Redémarrer l'instance.

3. Connectez-vous à votre instance et vérifiez que le nom d'hôte a été mis à jour. Votre invite devrait indiquer le nouveau no d'hôte (jusqu'au premier « . ») et la commande hostname doit afficher le nom de domaine complet.

   ```
   [ec2-user@webserver ~]$ hostname
   webserver.mydomain.com
   ```

Pour remplacer le nom d'hôte du système sans nom DNS public

1. • Pour Amazon Linux 2 : Utilisez la commande hostnamectl afin de définir votre nom d'hôte pour refléter le nom d'hôte du système souhaité (comme **webserver**).

```
[ec2-user ~]$ sudo hostnamectl set-hostname webserver.localdomain
```

   • Pour Amazon Linux AMI : Sur votre instance, ouvrez le fichier de configuration `/etc/sysconfig/network` dans votre éditeur de texte préféré et modifiez l'entrée `HOSTNAME` pour refléter le nom d'hôte du système souhaité (comme **webserver**).

```
HOSTNAME=webserver.localdomain
```

2. Ouvrez le fichier `/etc/hosts` dans votre éditeur de texte préféré et modifiez l'entrée commençant par **127.0.0.1** pour correspondre à l'exemple ci-dessous, en remplaçant votre propre nom d'hôte.

```
127.0.0.1 webserver.localdomain webserver localhost4 localhost4.localdomain4
```

3. Redémarrez l'instance pour récupérer le nouveau nom d'hôte.

```
[ec2-user ~]$ sudo reboot
```

   Vous pouvez également redémarrer à l'aide de la console Amazon EC2 (sur la page Instances, sélectionnez l'instance et choisissez État de l'instance, Redémarrer l'instance.

4. Connectez-vous à votre instance et vérifiez que le nom d'hôte a été mis à jour. Votre invite devrait indiquer le nouveau no d'hôte (jusqu'au premier « . ») et la commande hostname doit afficher le nom de domaine complet.

```
[ec2-user@webserver ~]$ hostname
webserver.localdomain
```

Changement de l'invite shell sans affecter le nom d'hôte

Si vous ne voulez pas modifier le nom d'hôte pour votre instance, mais que vous souhaitez un nom de système affiché plus utile (comme **webserver**) que le nom privé fourni par AWS (par exemple, `ip-12-34-56-78`), vous pouvez modifier les fichiers de configuration de l'invite shell pour afficher le pseudonyme de votre système au lieu du nom d'hôte.

Pour remplacer l'invite shell par un pseudonyme d'hôte

1. Créez un fichier dans `/etc/profile.d` qui définit la variable d'environnement appelée `NICKNAME` avec la valeur que vous souhaitez dans l'invite shell. Par exemple, pour définir le pseudonyme du système sur **webserver**, exécutez la commande suivante.

```
[ec2-user ~]$ sudo sh -c 'echo "export NICKNAME=webserver" > /etc/profile.d/prompt.sh'
```

2. Ouvrez le fichier `/etc/bashrc` (Red Hat) ou `/etc/bash.bashrc` (Debian/Ubuntu) dans l'éditeur de texte de votre choix (par exemple, vim ou nano). Vous devez utiliser sudo avec la commande de l'éditeur, car `/etc/bashrc` et `/etc/bash.bashrc` appartiennent à `root`.
3. Modifiez le fichier et changez la variable d'invite shell (`PS1`) pour afficher votre pseudonyme au lieu du nom d'hôte. Recherchez la ligne suivante qui définit l'invite shell dans `/etc/bashrc` ou `/etc/bash.bashrc` (plusieurs lignes sont affichées ci-dessous pour illustrer le contexte ; recherchez la ligne qui commence par `[ "$PS1"`) :

```
# Turn on checkwinsize
shopt -s checkwinsize
[ "$PS1" = "\\s-\\v\\$ " ] && PS1="[\u@\h \W]\\$ "
# You might want to have e.g. tty in prompt (e.g. more virtual machines)
# and console windows
```

Modifiez `\h` (symbole de `hostname`) sur cette ligne en la valeur de la variable `NICKNAME`.

```
# Turn on checkwinsize
shopt -s checkwinsize
[ "$PS1" = "\\s-\\v\\$ " ] && PS1="[\u@$NICKNAME \W]\\$ "
# You might want to have e.g. tty in prompt (e.g. more virtual machines)
# and console windows
```

4. (Facultatif) Pour définir le titre sur les fenêtres shell avec le nouveau pseudonyme, suivez les étapes suivantes.

   a. Créez un fichier nommé `/etc/sysconfig/bash-prompt-xterm`.

   ```
   [ec2-user ~]$ sudo touch /etc/sysconfig/bash-prompt-xterm
   ```

   b. Rendez le fichier exécutable avec la commande suivante.

   ```
   [ec2-user ~]$ sudo chmod +x /etc/sysconfig/bash-prompt-xterm
   ```

   c. Ouvrez le fichier `/etc/sysconfig/bash-prompt-xterm` avec votre éditeur de texte préféré (comme vim ou nano). Vous devez utiliser sudo avec la commande de votre éditeur, car `/etc/sysconfig/bash-prompt-xterm` est détenu par `root`.

   d. Ajoutez la ligne suivante au fichier.

   ```
   echo -ne "\033]0;${USER}@${NICKNAME}:${PWD/#$HOME/~}\007"
   ```

5. Déconnectez-vous puis reconnectez-vous pour récupérer la nouvelle valeur du pseudonyme.

Modifier le nom d'hôte sur d'autres distributions Linux

Les procédures de cette page sont destinées à une utilisation avec Amazon Linux uniquement. Pour plus d'informations sur les autres distributions Linux, consultez leur documentation spécifique et les articles suivants :

- Comment attribuer un nom d'hôte statique à une instance privée Amazon EC2 exécutant RHEL 7 ou Centos 7 ?

Réglage de DNS dynamique sur votre instance Amazon Linux

Lorsque vous lancez une instance EC2, on lui attribue une adresse IP publique et un nom DNS (Domain Name System) public que vous pouvez utiliser pour l'atteindre depuis Internet. Comme il y a tellement d'hôtes dans le domaine Amazon Web Services, ces noms publics doivent être assez longs pour que chaque nom reste unique. Un nom DNS public Amazon EC2 classique ressemble à `ec2-12-34-56-78.us-west-2.compute.amazonaws.com`, dans lequel le nom se compose du domaine Amazon Web Services, du service (dans ce cas, `compute`), de la région et d'une forme d'adresse IP publique.

Les services DNS dynamiques fournissent des noms d'hôte DNS personnalisés dans leur domaine qui peut être facile à mémoriser et aussi plus pertinent vis-à-vis du cas d'utilisation de votre hôte. Certains de ces services sont également gratuits. Vous pouvez utiliser un fournisseur DNS dynamique avec Amazon EC2 et configurez l'instance pour mettre à jour l'adresse IP associée au nom DNS public à chaque fois que l'instance commence. Il existe un choix de plusieurs fournisseurs différents et les détails spécifiques à la sélection d'un fournisseur et à l'enregistrement d'un nom sans eux ne sont pas pris en compte dans le cadre de ce guide.

> **Important**
>
> Ces informations s'appliquent à Amazon Linux. Pour obtenir des informations sur d'autres distributions, consultez leur documentation spécifique.

Pour utiliser DNS dynamique avec Amazon EC2

1. Inscrivez-vous avec un fournisseur de services DNS dynamiques et enregistrez un nom DNS public avec leur service. Cette procédure utilise le service gratuit de noip.com/free comme exemple.
2. Configurez le client de mise à jour de DNS dynamique. Après avoir enregistré un fournisseur de services DNS dynamiques et un nom DNS public avec leur service, reliez le nom DNS à l'adresse IP de votre instance. De nombreux fournisseurs (notamment noip.com) vous permettent de faire cela manuellement depuis la page de votre compte sur leur site web, mais beaucoup prennent également en charge les clients de mise à jour logicielle. Si un client de mise à jour est exécuté sur votre instance EC2, votre enregistrement DNS dynamique est mis à jour à chaque fois que l'adresse IP change, comme après une fermeture et un redémarrage. Dans cet exemple, vous installez le client noip2 qui fonctionne avec le service fourni par noip.com.

   a. Activez le référentiel EPEL (Extra Packages for Enterprise Linux) afin d'accéder au client noip2.

      > **Note**
      >
      > Les instances Amazon Linux possèdent les clés GPG et les détails relatifs au référentiel pour le dépôt EPEL installé par défaut. Cependant, les instances Red Hat et CentOS doivent tout d'abord installer le package `epel-release` avant que vous ne puissiez activer le référentiel EPEL. Pour obtenir plus d'information et télécharger la dernière version de ce package, consultez https://fedoraproject.org/wiki/EPEL.

      - Dans Amazon Linux 2:

      ```
      [ec2-user ~]$ sudo yum install https://dl.fedoraproject.org/pub/epel/epel-release-latest-7.noarch.rpm
      ```

      - Dans Amazon Linux AMI:

      ```
      [ec2-user ~]$ sudo yum-config-manager --enable epel
      ```

   b. Installez le package `noip`.

      ```
      [ec2-user ~]$ sudo yum install -y noip
      ```

   c. Créez le fichier de configuration . Saisissez l'identifiant et le mot de passe lorsque vous y êtes invité et répondez aux questions suivantes pour configurer le client.

      ```
      [ec2-user ~]$ sudo noip2 -C
      ```

3. Activez le service noip.

   - Dans Amazon Linux 2:

```
[ec2-user ~]$ sudo systemctl enable noip.service
```

- Dans Amazon Linux AMI:

```
[ec2-user ~]$ sudo chkconfig noip on
```

4. Lancez le service noip.

   - Dans Amazon Linux 2:

   ```
   [ec2-user ~]$ sudo systemctl start noip.service
   ```

   - Dans Amazon Linux AMI:

   ```
   [ec2-user ~]$ sudo service noip start
   ```

   Cette commande lance le client, qui lit le fichier de configuration (`/etc/no-ip2.conf`) que vous avez créé précédemment et met à jour l'adresse IP du nom du DNS public que vous avez choisi.

5. Vérifiez que le client de mise à jour a défini la bonne adresse IP de votre nom DNS dynamique. Laissez s'écouler quelques minutes pour que les enregistrements DNS se mettent à jour, puis essayez de connecter votre instance en utilisant SSH avec le nom DNS public que vous avez configuré dans cette procédure.

Exécution de commandes sur votre instance Linux lors du lancement

Lorsque vous lancez une instance dans Amazon EC2, vous avez la possibilité de transmettre les données des utilisateurs vers l'instance qui peut être utilisée pour effectuer des tâches de configuration automatisées communes et même exécuter des scripts après le démarrage de l'instance. Vous pouvez transmettre deux types de données utilisateur vers Amazon EC2 : des scripts shell et des directives cloud-init. Vous pouvez aussi transférer ces données dans l'assistant de lancement comme du texte brut, en tant que fichier (cela est utile pour le lancement d'instance à l'aide des outils de ligne de commande) ou en tant que texte encodé base64 (pour les appels API).

Si vous êtes intéressé par les scénarios complexes d'automatisation, pensez à utiliser AWS CloudFormation et AWS OpsWorks. Pour plus d'informations, consultez le AWS CloudFormation Guide de l'utilisateur et le AWS OpsWorks User Guide.

Pour obtenir des informations sur l'exécution de commandes sur votre instance Windows au lancement, consultez Exécution de commandes sur votre instance Windows lors du lancement et Gestion de la configuration des instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Dans les exemples suivants, les commandes du Didacticiel : Installation d'un serveur web LAMP sur Amazon Linux 2 (p. 38) sont converties en script shell et en ensemble de directives cloud-init qui est exécuté lors du lancement de l'instance. Dans chaque exemple, les tâches suivantes sont exécutées par les données de l'utilisateur :

- Les packages logiciels de distribution sont mis à jour.
- Le serveur web, les packages `php` et `mariadb` nécessaires sont installés.
- Le service `httpd` est lancé et activé via la commande systemctl.
- L'utilisateur `ec2-user` est ajouté au groupe Apache.

- La propriété et les autorisations sur les fichiers appropriées sont définies pour le répertoire web et les fichiers qu'il contient.
- Une page Web simple est créée pour tester le serveur Web et le moteur PHP.

Table des matières

- Prérequis (p. 695)
- Données utilisateur et scripts shell (p. 695)
- Données utilisateur et console (p. 696)
- Directives sur les données utilisateur et Cloud-Init (p. 698)
- Données utilisateur et AWS CLI (p. 699)

Prérequis

Les exemples suivants supposent que votre instance possède un nom DNS public que l'on peut atteindre à partir d'Internet. Pour plus d'informations, consultez Étape 1 : Lancement d'une instance (p. 33). Vous devez aussi configurer votre groupe de sécurité pour autoriser les connexions SSH (port 22), HTTP (port 80) et HTTPS (port 443). Pour obtenir plus d'informations sur ces conditions préalables, consultez le didacticiel Configuration avec Amazon EC2 (p. 28).

Par ailleurs, ces instructions sont destinées à Amazon Linux 2, et il se peut que les commandes et les directives ne fonctionnent pas pour d'autres distributions Linux. Pour obtenir plus d'informations sur d'autres distributions, comme leur support pour cloud-init, consultez leur documentation spécifique.

Données utilisateur et scripts shell

Si vous connaissez l'écriture de scripts shell, il s'agit de la méthode la plus simple et la plus complète pour envoyer des instructions à une instance lors du lancement. L'ajout de ces tâches au moment du démarrage augmente le temps que cela prend pour démarrer l'instance. Vous devriez laisser s'écouler quelques minutes supplémentaires pour que les tâches s'effectuent avant de vérifier que le script utilisateur a fini avec succès.

> **Important**
>
> Par défaut, les scripts de données utilisateur et les directives cloud-init s'exécutent uniquement pendant le cycle de démarrage lors du premier lancement d'une instance. Vous pouvez mettre à jour votre configuration afin que vos scripts de données utilisateur et vos directives cloud-init s'exécutent à chaque redémarrage de votre instance. Pour plus d'informations, consultez Comment puis-je exécuter des données utilisateur à chaque redémarrage de mon instance EC2 ? dans le centre de connaissances AWS.

Les scripts shell de données utilisateur doivent commencer par les caractères `#!` et le chemin vers interpréteur que vous avez choisi pour la lecture du script (généralement /bin/bash)). Pour consulter une excellente présentation sur l'écriture de scripts shell, consultez BASH Programming HOW-TO dans le projet de documentation Linux (tldp.org).

Les scripts entrés en tant que données utilisateur sont exécutés comme l'utilisateur `root`, donc n'utilisez pas la commande sudo dans le script. Souvenez que tous les fichiers que vous créez seront détenus par `root`. Si des utilisateurs non-root doivent accéder au fichier, vous devriez modifier les autorisations en fonction du script. Par ailleurs, étant donné que le script n'est pas exécuté de façon interactive, vous ne pouvez pas inclure des commandes qui nécessitent les réactions de l'utilisateur (comme yum update sans l'indicateur `-y`).

Si vous utilisez une API AWS, y compris l'interface de ligne de commande AWS, dans un script de données utilisateur, vous devez utiliser un profil d'instance lors du lancement de l'instance. Un profil d'instance

fournit les informations d'identification AWS appropriées requises par le script de données utilisateur pour exécuter l'appel d'API. Pour de plus amples informations, veuillez consulter Utilisation de profils d'instance dans le IAM Guide de l'utilisateur. Les autorisations que vous attribuez au rôle IAM dépendent des services que vous appelez avec l'API. Pour de plus amples informations, veuillez consulter Rôles IAM pour Amazon EC2.

Le fichier journal de sortie cloud-init (`/var/log/cloud-init-output.log`) capture la sortie de la console, si bien que vous pouvez facilement déboguer vos scripts suite à un lancement si l'instance ne se comporte pas comme vous le vouliez.

Lorsqu'un script de données utilisateur est traité, il est copié dans `/var/lib/cloud/instances/`*`instance-id`*`/` et exécuté à partir de là. Le script n'est pas supprimé après son exécution. Veillez à supprimer les scripts de données utilisateur dans `/var/lib/cloud/instances/`*`instance-id`*`/` avant de créer une AMI à partir de l'instance. Dans le cas contraire, le script figurera dans ce répertoire sur toute instance lancée à partir de l'AMI.

Données utilisateur et console

Vous pouvez spécifier des données utilisateur d'instance lorsque vous lancez l'instance. Si le volume racine de l'instance est un volume EBS, vous pouvez également arrêter l'instance et mettre à jour ses données utilisateur.

Spécification des données utilisateur d'instance au moment du lancement

Suivez la procédure de lancement d'une instance décrite dans Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530), mais, quand vous arrivez à the section called “Étape 3 : Configurer les détails de l'instance” (p. 532) dans cette procédure, copiez votre script shell dans le champ Données utilisateur, puis terminez la procédure de lancement.

Dans l'exemple de script ci-dessous, le script crée et configure notre serveur web.

```
#!/bin/bash
yum update -y
amazon-linux-extras install -y lamp-mariadb10.2-php7.2 php7.2
yum install -y httpd mariadb-server
systemctl start httpd
systemctl enable httpd
usermod -a -G apache ec2-user
chown -R ec2-user:apache /var/www
chmod 2775 /var/www
find /var/www -type d -exec chmod 2775 {} \;
find /var/www -type f -exec chmod 0664 {} \;
echo "<?php phpinfo(); ?>" > /var/www/html/phpinfo.php
```

Laissez assez de temps à l'instance pour lancer et exécuter les commandes dans votre script, puis vérifiez que votre script a terminé les tâches que vous souhaitiez.

Pour notre exemple, dans un navigateur web, saisissez l'URL du fichier test PHP que le script a créé. Cette URL est l'adresse DNS publique de votre instance suivie par une barre oblique et le nom du fichier.

```
http://my.public.dns.amazonaws.com/phpinfo.php
```

Vous devriez voir la page d'informations PHP. Si vous ne pouvez pas voir la page d'informations PHP, vérifiez que le groupe de sécurité que vous utilisez contient une règle pour permettre le trafic HTTP (port 80). Pour plus d'informations, consultez Ajout de règles à un groupe de sécurité (p. 1069).

(Facultatif) Si votre script n'a pas accompli les tâches que vous attendiez ou si vous voulez uniquement vérifier que votre script s'est terminé sans erreur, examinez le fichier journal de sortie cloud-init à

l'emplacement `/var/log/cloud-init-output.log` et recherchez les messages d'erreur dans les résultats.

Pour obtenir des informations supplémentaires sur le débogage, vous pouvez créer une archive Mime en plusieurs parties qui comporte la section de données cloud-init avec la directive suivante :

```
output : { all : '| tee -a /var/log/cloud-init-output.log' }
```

Cette directive envoie la sortie de la commande de votre script à `/var/log/cloud-init-output.log`. Pour plus d'informations sur les formats de données cloud-init et la création d'archives Mime en plusieurs parties, consultez Formats de cloud-init..

Affichage et mise à jour des données utilisateur d'instance

Pour mettre à jour les données de l'utilisateur de l'instance, vous devez d'abord arrêter l'instance. Si l'instance est en cours d'exécution, vous pouvez afficher les données utilisateur, mais vous ne pouvez pas les modifier.

> **Warning**
>
> Lorsque vous arrêtez une instance, les données contenues sur les volumes de stockage d'instance sont effacées. Pour conserver les données des volumes de stockage d'instance, veillez à les sauvegarder sur un stockage permanent.

Nouvelle console

Pour modifier les données utilisateur d'instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez État de l'instance, Arrêter l'instance. Si cette option est désactivée, l'instance est déjà arrêtée ou son périphérique racine est un volume de stockage d'instance.
4. Lorsque vous êtes invité à confirmer l'opération, choisissez Arrêter. L'arrêt de l'instance peut prendre quelques minutes.
5. Alors que l'instance est toujours sélectionnée, choisissez Actions, Instance settings (Paramètres de l'instance), Edit user data (Modifier les données utilisateur).
6. Modifiez les données utilisateur selon vos besoins, puis choisissez Save (Enregistrer).
7. Redémarrez l'instance. Les nouvelles données utilisateur sont visibles sur votre instance, après son redémarrage. Par contre, les scripts de données utilisateur ne sont pas exécutés.

Ancienne console

Pour modifier les données utilisateur d'instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, État de l'instance, Arrêter. Si cette option est désactivée, l'instance est déjà arrêtée ou son périphérique racine est un volume de stockage d'instance.
4. Lorsque vous êtes invité à confirmer, choisissez Oui, arrêter. L'arrêt de l'instance peut prendre quelques minutes.
5. Avec l'instance est toujours sélectionnée, choisissez Actions, Paramètres de l'instance, Afficher/Changer les données utilisateur.

6. Dans la boîte de dialogue Afficher/Changer les données utilisateur, mettez à jour les données utilisateur, puis cliquez sur Enregistrer.
7. Redémarrez l'instance. Les nouvelles données utilisateur sont visibles sur votre instance, après son redémarrage. Par contre, les scripts de données utilisateur ne sont pas exécutés.

Directives sur les données utilisateur et Cloud-Init

Le package cloud-init configure les aspects spécifiques d'une nouvelle instance Amazon Linux lorsqu'elle est lancée. Il configure plus particulièrement le fichier `.ssh/authorized_keys` pour l'utilisateur ec2 afin que vous puissiez vous connecter avec votre clé privée. Pour plus d'informations, consultez cloud-init (p. 192).

Les directives d'utilisateur cloud-init peuvent être transférées vers une instance au moment du lancement tout comme un script, même si la syntaxe est différente. Pour plus d'informations sur cloud-init, accédez à http://cloudinit.readthedocs.org/en/latest/index.html.

> **Important**
>
> Par défaut, les scripts de données utilisateur et les directives cloud-init s'exécutent uniquement pendant le cycle de démarrage lors du premier lancement d'une instance. Vous pouvez mettre à jour votre configuration afin que vos scripts de données utilisateur et vos directives cloud-init s'exécutent à chaque redémarrage de votre instance. Pour plus d'informations, consultez Comment puis-je exécuter des données utilisateur à chaque redémarrage de mon instance EC2 ? dans le centre de connaissances AWS.

L'ajout de ces tâches au moment du démarrage augmente le temps que cela prend pour démarrer une instance. Vous devriez laisser s'écouler quelques minutes supplémentaires pour que les tâches s'effectuent avant de vérifier que vos directives sur les données utilisateur sont terminées.

Pour transférer les directives cloud-init vers une instance avec les données utilisateur

1. Suivez la procédure de lancement d'une instance décrite dans Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530), mais, quand vous arrivez à the section called "Étape 3 : Configurer les détails de l'instance" (p. 532) dans cette procédure, entrez le texte de votre directive could-init dans le champ Données utilisateur, puis terminez la procédure de lancement.

   Pour l'exemple ci-dessous, les directives créent et configurent un serveur Web sur Amazon Linux 2. La ligne `#cloud-config` en haut est requise pour identifier les commandes en tant que directives cloud-init.

```
#cloud-config
repo_update: true
repo_upgrade: all

packages:
 - httpd
 - mariadb-server

runcmd:
 - [ sh, -c, "amazon-linux-extras install -y lamp-mariadb10.2-php7.2 php7.2" ]
 - systemctl start httpd
 - sudo systemctl enable httpd
 - [ sh, -c, "usermod -a -G apache ec2-user" ]
 - [ sh, -c, "chown -R ec2-user:apache /var/www" ]
 - chmod 2775 /var/www
 - [ find, /var/www, -type, d, -exec, chmod, 2775, {}, \; ]
 - [ find, /var/www, -type, f, -exec, chmod, 0664, {}, \; ]
 - [ sh, -c, 'echo "<?php phpinfo(); ?>" > /var/www/html/phpinfo.php' ]
```

2. Laissez assez de temps à l'instance pour lancer et exécuter les directives dans vos données utilisateur, puis vérifiez que vos directives ont terminé les tâches que vous souhaitiez.

   Pour notre exemple, dans un navigateur web, saisissez l'URL du fichier test PHP que les directives ont créé. Cette URL est l'adresse DNS publique de votre instance suivie par une barre oblique et le nom du fichier.

   ```
   http://my.public.dns.amazonaws.com/phpinfo.php
   ```

   Vous devriez voir la page d'informations PHP. Si vous ne pouvez pas voir la page d'informations PHP, vérifiez que le groupe de sécurité que vous utilisez contient une règle pour permettre le trafic HTTP (port 80). Pour plus d'informations, consultez Ajout de règles à un groupe de sécurité (p. 1069).
3. (Facultatif) Si vos directives n'ont pas accompli les tâches que vous attendiez ou si vous voulez uniquement vérifier que vos directives se sont terminées sans erreur, examinez le fichier journal de sortie à l'emplacement `/var/log/cloud-init-output.log` et recherchez les messages d'erreur dans les résultats. Pour plus d'informations sur le débogage, vous pouvez ajouter la ligne suivante à vos directives :

   ```
   output : { all : '| tee -a /var/log/cloud-init-output.log' }
   ```

   Cette directive envoie le résultat runcmd à `/var/log/cloud-init-output.log`.

Données utilisateur et AWS CLI

Vous pouvez utiliser l'AWS CLI pour spécifier, modifier et afficher les données utilisateur de votre instance. Pour plus d'informations sur l'affichage des données utilisateur de votre instance à l'aide des métadonnées d'instance, consultez Récupération des données utilisateur d'instance (p. 716).

Sur Windows, vous pouvez utiliser le Outils AWS pour Windows PowerShell plutôt que l'AWS CLI. Pour plus d'informations, consultez Données utilisateur et Outils pour Windows PowerShell dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Exemple : spécification des données utilisateur au moment du lancement

Pour spécifier les données utilisateur lorsque vous lancez l'instance, utilisez la commande run-instances avec le paramètre `--user-data`. Avec run-instances, l'AWS CLI effectue l'encodage base64 des données utilisateur pour vous.

L'exemple suivant montre comment définir un script en tant que chaîne sur la ligne de commande :

```
aws ec2 run-instances --image-id ami-abcd1234 --count 1 --instance-type m3.medium \
--key-name my-key-pair --subnet-id subnet-abcd1234 --security-group-ids sg-abcd1234 \
--user-data echo user data
```

L'exemple suivant montre comment définir un script en utilisant un fichier texte. Assurez-vous d'utiliser le préfixe `file://` pour spécifier le fichier.

```
aws ec2 run-instances --image-id ami-abcd1234 --count 1 --instance-type m3.medium \
--key-name my-key-pair --subnet-id subnet-abcd1234 --security-group-ids sg-abcd1234 \
--user-data file://my_script.txt
```

L'exemple suivant est celui d'un fichier texte avec un script shell.

```
#!/bin/bash
```

```
yum update -y
service httpd start
chkconfig httpd on
```

Exemple : Modification des données utilisateur d'une instance arrêtée

Vous pouvez modifier les données utilisateur d'une instance arrêtée à l'aide de la commande modify-instance-attribute. Avec modify-instance-attribute, l'AWS CLI n'effectue pas l'encodage base64 des données utilisateur pour vous.

- Sur un ordinateur Linux utilisez la commande base64 pour encoder les données utilisateur.

```
base64 my_script.txt >my_script_base64.txt
```

- Sur un ordinateur Windows, utilisez la commande certutil pour encoder les données utilisateur. Pour pouvoir utiliser ce fichier avec l'AWS CLI, vous devez supprimer la première ligne (BEGIN CERTIFICATE) et la dernière ligne (END CERTIFICATE).

```
certutil -encode my_script.txt my_script_base64.txt
notepad my_script_base64.txt
```

Utilisez les paramètres `--attribute` et `--value` afin d'utiliser le fichier texte encodé pour spécifier les données utilisateur. Assurez-vous d'utiliser le préfixe `file://` pour spécifier le fichier.

```
aws ec2 modify-instance-attribute --instance-id i-1234567890abcdef0 --attribute userData --
value file://my_script_base64.txt
```

Exemple : Effacer les données utilisateur d'une instance arrêtée

Pour supprimer les données utilisateur existantes, utilisez la commande modify-instance-attribute comme suit :

```
aws ec2 modify-instance-attribute --instance-id i-1234567890abcdef0 --user-data Value=
```

Exemple : Affichage des données utilisateur

Pour extraire les données utilisateur pour une instance, utilisez la commande describe-instance-attribute. Avec describe-instance-attribute, l'AWS CLI n'effectue pas le décodage en base64 des données utilisateur pour vous.

```
aws ec2 describe-instance-attribute --instance-id i-1234567890abcdef0 --attribute userData
```

Voici un exemple de sortie avec les données utilisateur base64 encodées.


```
{
    "UserData": {
        "Value":
 "IyEvYmluL2Jhc2gKeXVtIHVwZGF0ZSAteQpzZXJ2aWNlIGh0dHBkIHN0YXJ0CmNoa2NvbmZpZyBodHRwZCBvbg=="
    },
    "InstanceId": "i-1234567890abcdef0"
}
```


- Sur un ordinateur Linux, utilisez l'option `--query` pour obtenir les données utilisateur encodées et la commande base64 pour les décoder.

```
aws ec2 describe-instance-attribute --instance-id i-1234567890abcdef0 --attribute
 userData --output text --query "UserData.Value" | base64 --decode
```

- Sur un ordinateur Windows, utilisez l'option `--query` pour obtenir les données utilisateur codées et la commande certutil pour les décoder. Notez que la sortie encodée est stockée dans un fichier et que la sortie décodée est stockée dans un autre fichier.

```
aws ec2 describe-instance-attribute --instance-id i-1234567890abcdef0 --attribute
 userData --output text --query "UserData.Value" >my_output.txt
certutil -decode my_output.txt my_output_decoded.txt
type my_output_decoded.txt
```

Voici un exemple de sortie.

```
#!/bin/bash
yum update -y
service httpd start
chkconfig httpd on
```

Métadonnées d'instance et données utilisateur

Les métadonnées d'instance sont des données portant sur votre instance que vous pouvez utiliser pour configurer ou gérer l'instance en cours d'exécution. Les métadonnées d'instance sont divisées en catégories, par exemple, nom d'hôte, événements et groupes de sécurité.

Vous pouvez également utiliser les métadonnées d'instance pour accéder aux données utilisateur que vous avez spécifiées au moment du lancement de votre instance. Par exemple, vous pouvez spécifier des paramètres pour la configuration de votre instance ou inclure un script simple. Vous pouvez créer des AMI génériques et utiliser des données utilisateur pour modifier les fichiers de configuration fournis au moment du lancement. Par exemple, si vous exécutez des serveurs web pour plusieurs petites entreprises, ces serveurs peuvent tous utiliser la même AMI générique et récupérer leur contenu à partir du compartiment Amazon S3 que vous spécifiez dans les données utilisateur lors du lancement. Pour ajouter un nouveau client à n'importe quel moment, créez un compartiment pour le client, ajoutez son contenu, puis lancez votre AMI avec l'unique nom de compartiment fourni à votre code dans les données utilisateur. Si vous lancez plus d'une instance en même temps, les données utilisateurs sont disponibles pour toutes les instances de cette réservation. Chaque instance faisant partie de la même réservation possède un numéro `ami-launch-index` unique vous permettant d'écrire du code contrôlant les opérations. Par exemple, le premier hôte peut s'élire comme nœud d'origine dans un cluster. Pour obtenir un exemple détaillé de lancement d'AMI, veuillez consulter Exemple : Valeur d'index de lancement AMI (p. 718).

Les instances EC2 peuvent également comprendre des données dynamiques, par exemple un document d'identité d'instance qui est généré au lancement de l'instance. Pour plus d'informations, consultez Catégories de données dynamiques (p. 728).

> **Important**
>
> Bien que les métadonnées d'instance et les données utilisateur ne soient accessibles qu'au sein de l'instance elle-même, elles ne sont pas protégées par des méthodes d'authentification ou de chiffrement. Toute personne ayant un accès direct à l'instance, et potentiellement tout logiciel s'exécutant sur l'instance, peut afficher ses métadonnées. Vous ne devez donc pas stocker de données sensibles, telles que des mots de passe ou des clés de chiffrement à longue durée, ou des données utilisateur.

Sommaire

- Configuration du service des métadonnées d'instance (p. 702)

- Récupération des métadonnées d'instance (p. 708)
- Utilisation des données utilisateur d'instance (p. 716)
- Récupération de données dynamiques (p. 717)
- Exemple : Valeur d'index de lancement AMI (p. 718)
- Catégories de métadonnées d'instance (p. 721)
- Documents d'identité d'instance (p. 728)

Configuration du service des métadonnées d'instance

Vous pouvez accéder aux métadonnées d'instance à partir d'une instance en cours d'exécution en utilisant l'une des méthodes suivantes :

- Service des métadonnées d'instance Version 1 (IMDSv1) – méthode de demande/réponse
- Service des métadonnées d'instance Version 2 (IMDSv2) – méthode orientée session

Par défaut, vous pouvez utiliser IMDSv1 ou IMDSv2, ou les deux. Le service des métadonnées d'instance fait la distinction entre les demandes IMDSv1 et IMDSv2 pour une demande donnée en déterminant si les en-têtes `PUT` ou `GET`, qui sont uniques pour IMDSv2, sont présents dans cette demande.

Vous pouvez configurer le service des métadonnées d'instance sur chaque instance afin que le code local ou les utilisateurs doivent utiliser IMDSv2. Lorsque vous spécifiez que IMDSv2 doit être utilisé, IMDSv1 ne fonctionne plus. Pour de plus amples informations, veuillez consulter Configuration des options des métadonnées d'instance (p. 705).

Pour récupérer des métadonnées d'instance, veuillez consulter Récupération des métadonnées d'instance (p. 708).

Fonctionnement de Service des métadonnées d'instance Version 2

IMDSv2 utilise des demandes orientées session. Lorsque vous utilisez des demandes orientées session, vous créez un jeton de session qui définit la durée de la session, qui doit être d'une seconde au minimum et de six heures au maximum. Durant la période spécifiée, vous pouvez utiliser le même jeton de session pour les demandes suivantes. Une fois la période spécifiée arrivée à expiration, vous devez créer un nouveau jeton de session à utiliser pour les futures demandes.

L'exemple suivant utilise un script shell Linux et IMDSv2 pour extraire les éléments de métadonnées d'instance de haut niveau. L'exemple de commande :

- Crée un jeton de session d'une durée de six heures (21 600 secondes) en utilisant la demande `PUT`
- Stocke l'en-tête de jeton de session dans une variable nommée `TOKEN`
- Demande les éléments de métadonnées de haut niveau à l'aide du jeton

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-data/
```

Une fois que vous avez créé un jeton, vous pouvez le réutiliser jusqu'à son expiration. Dans l'exemple de commande suivant, qui extrait l'ID de l'AMI utilisée pour lancer l'instance, le jeton stocké dans `$TOKEN` dans l'exemple précédent est réutilisé.

```
[ec2-user ~]$ curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
meta-data/ami-id
```

Lorsque vous utilisez IMDSv2 pour demander les métadonnées d'une instance, la demande doit inclure les éléments suivants :

1. Utilisez une demande `PUT` pour lancer une session sur le service des métadonnées d'instance. La demande `PUT` renvoie un jeton qui doit être inclus dans les demandes `GET` suivantes envoyées au service des métadonnées d'instance. Le jeton est obligatoire pour accéder aux métadonnées à l'aide de IMDSv2.
2. Incluez le jeton dans toutes les demandes `GET` envoyées au service des métadonnées d'instance. Lorsque l'utilisation de jeton est définie sur `required`, les demandes sans jeton valide ou contenant un jeton arrivé à expiration reçoivent un code d'erreur HTTP `401 - Unauthorized`. Pour de plus amples informations sur la modification des conditions d'utilisation des jetons, veuillez consulter modify-instance-metadata-options dans le AWS CLI Command Reference.
   - Le jeton est une clé propre à l'instance. Le jeton n'est pas valide sur les autres instances EC2 et sera rejeté si vous tentez de l'utiliser ailleurs que sur l'instance sur laquelle il a été généré.
   - La demande `PUT` doit inclure un en-tête spécifiant la durée de vie (TTL) du jeton, en secondes, jusqu'à six heures au maximum (21 600 secondes). Le jeton représente une session logique. La durée de vie (TTL) définit la durée de validité du jeton et, par conséquent, la durée de la session.
   - Une fois qu'un jeton est arrivé à expiration, pour pouvoir continuer à accéder aux métadonnées de l'instance, vous devez créer une nouvelle session en utilisant un autre `PUT`.
   - Vous pouvez choisir de réutiliser un jeton ou d'en créer un nouveau pour chaque demande. Pour un faible nombre de demandes, il peut être plus facile de générer et d'utiliser immédiatement un jeton chaque fois que vous avez besoin d'accéder au service des métadonnées d'instance. Cependant, pour une plus grande productivité, vous pouvez spécifier une durée plus longue pour le jeton et le réutiliser plutôt que de devoir écrire une demande `PUT` chaque fois que vous avez besoin de demander des métadonnées d'instance. Il n'existe pas de limite pratique au nombre de jetons simultanés, chacun représentant sa propre session. IMDSv2 est toutefois soumis aux limites normales de connexion du service des métadonnées d'instance. Pour de plus amples informations, veuillez consulter Limitation (p. 714).

Les méthodes HTTP `GET` et `HEAD` sont autorisées dans les demandes de métadonnées d'instance IMDSv2. Les demandes `PUT` sont rejetées si elles contiennent un en-tête X-Forwarded-For.

Par défaut, la réponse aux demandes `PUT` possède une durée de vie (hop limit) de réponse de `1` au niveau du protocole IP. Vous pouvez ajuster cette durée en utilisant la commande `modify-instance-metadata-options` si nécessaire. Par exemple, vous pouvez avoir besoin d'une durée de vie (hop limit) plus élevée pour des raisons de compatibilité en amont avec les services de conteneur s'exécutant sur l'instance. Pour de plus amples informations, veuillez consulter modify-instance-metadata-options dans le AWS CLI Command Reference.

Migration vers Service des métadonnées d'instance Version 2

L'utilisation d'Service des métadonnées d'instance Version 2 (IMDSv2) est facultative. Service des métadonnées d'instance Version 1 (IMDSv1) continuera d'être pris en charge sans limite dans le temps. Si vous choisissez d'effectuer la migration vers IMDSv2, nous vous recommandons d'utiliser les outils et le chemin de transition suivants.

Outils facilitant la migration vers IMDSv2

Si votre logiciel utilise IMDSv1, utilisez les outils suivants pour faciliter sa reconfiguration vers IMDSv2.

- Logiciels AWS : Les dernières versions des kits SDK AWS et des interfaces de ligne de commande AWS prennent en charge IMDSv2. Pour utiliser IMDSv2, veillez à ce que vos instances EC2 possèdent les dernières versions des kits SDK et des interfaces de ligne de commande AWS. Pour de plus amples informations sur la mise à jour de l'interface de ligne de commande, veuillez consulter Mise à niveau vers la dernière version de l'interface de ligne de commande AWS dans le AWS Command Line Interface Guide de l'utilisateur.

- CloudWatch : IMDSv2 utilise des sessions basées sur une jeton, mais pas IMDSv1. La métrique CloudWatch `MetadataNoToken` suit le nombre d'appels au service de métadonnées d'instance qui utilisent IMDSv1. En suivant cette métrique jusqu'à zéro, vous pouvez déterminer si la totalité de votre logiciel a été mis à niveau vers IMDSv2 et le moment auquel cela se produit. Pour de plus amples informations, veuillez consulter Métriques des instances (p. 764).
- Mises à jour des API et des interfaces de ligne de commande EC2 : Pour les instances existantes, vous pouvez utiliser la commande d'interface de ligne de commande modify-instance-metadata-options (ou l'API ModifyInstanceMetadataOptions) pour demander l'utilisation de IMDSv2. Pour les nouvelles instances, vous pouvez utiliser la commande de l'interface de ligne de commande run-instances (ou l'API RunInstances) et le paramètre `metadata-options` pour lancer les nouvelles instances qui nécessitent l'utilisation de IMDSv2.

  Pour exiger l'utilisation de IMDSv2 sur toutes les nouvelles instances lancées par des groupes Auto Scaling, ces derniers peuvent utiliser un modèle de lancement ou une configuration de lancement. Lorsque vous créez un modèle de lancement ou une configuration de lancement, vous devez configurer les paramètres `MetadataOptions` pour exiger l'utilisation de IMDSv2. Après avoir configuré le modèle de lancement ou la configuration de lancement, le groupe Auto Scaling lance de nouvelles instances à l'aide du nouveau modèle de lancement ou de la nouvelle configuration de lancement, mais les instances existantes ne sont pas affectées.

  Utilisez la commande d'interface de ligne de commande modify-instance-metadata-options (ou l'API ModifyInstanceMetadataOptions) pour exiger l'utilisation de IMDSv2 sur les instances existantes, ou terminez les instances et le groupe Auto Scaling lancera de nouvelles instances de remplacement avec les paramètres des options de métadonnées d'instance définis dans le modèle ou la configuration de lancement.

  Pour les groupes Auto Scaling qui utilisent des configurations de lancement, vous pouvez remplacer les configurations de lancement par des modèles de lancement.
- Stratégies IAM et stratégies de contrôle de service (SCP) : Vous pouvez utiliser une condition IAM pour empêcher les utilisateurs IAM de lancer une instance qui n'utilise pas IMDSv2. Vous pouvez également utiliser des conditions IAM pour empêcher les utilisateurs IAM de modifier les instances en cours d'exécution en vue de rétablir IMDSv1, et pour faire en sorte que le services des métadonnées d'instance soit disponible sur l'instance.

  Les clés de condition IAM `ec2:MetadataHttpTokens`, `ec2:MetadataHttpPutResponseHopLimit` et `ec2:MetadataHttpEndpoint` peuvent être utilisées pour contrôler l'utilisation des API RunInstances et ModifyInstanceMetadataOptions et de l'interface de ligne commande correspondante. Si une stratégie est créée et qu'un paramètre de l'appel d'API ne correspond pas à l'état spécifié dans la stratégie à l'aide de la clé de condition, l'appel de l'API ou de l'interface de ligne commande échoue avec la réponse `UnauthorizedOperation`. Ces clés de condition peuvent être utilisées dans les stratégies IAM ou dans les stratégies de contrôle des services (SCP) AWS Organizations.

  Vous pouvez en outre choisir une couche de protection supplémentaire afin d'imposer le passage de IMDSv1 à IMDSv2. Au niveau de la couche de gestion des accès concernant les API appelées via des informations d'identification de rôle EC2, vous pouvez utiliser une nouvelle clé de condition dans les stratégies IAM ou les politiques de contrôle des services (SCP) AWS Organizations. Si vous utilisez la clé de condition de stratégie `ec2:RoleDelivery` avec la valeur `2.0` dans vos stratégies IAM, les appels d'API effectués avec des informations d'identification de rôle EC2 obtenues à partir de IMDSv1 recevront une réponse `UnauthorizedOperation`. Vous pouvez aboutir au même résultat plus généralement avec cette condition requise par une SCP. Cela permet de s'assurer que les informations d'identification fournies via IMDSv1 ne peuvent pas être utilisées pour appeler des API, car tout appel d'API ne respectant pas la condition spécifiée recevra une erreur `UnauthorizedOperation`. Par exemple les stratégies IAM, consultez Utilisation des métadonnées d'instance (p. 1022). Pour de plus amples informations, veuillez consulter Politiques de contrôle des services dans le Guide de l'utilisateur AWS Organizations.

Chemin recommandé pour demander l'accès à IMDSv2

Nous vous recommandons, tout en utilisant les outils mentionnés précédemment, de suivre ce chemin pour la migration vers IMDSv2 :

Etape 1 : Au départ

Mettez à jour les kit SDK, les interfaces de ligne de commande et vos logiciels utilisant des informations d'identification de rôle sur leurs instances EC2 vers des versions de IMDSv2 compatibles. Pour de plus amples informations sur la mise à jour de l'interface de ligne de commande, veuillez consulter Mise à niveau vers la dernière version de l'interface de ligne de commande AWS dans le AWS Command Line Interface Guide de l'utilisateur.

Modifiez ensuite les logiciels accédant directement aux métadonnées de l'instance (en d'autres termes, n'utilisant pas un kit SDK) à l'aide des demandes IMDSv2.

Etape 2 : Pendant la transition

Suivez la progression de votre transition à l'aide de la métrique CloudWatch `MetadataNoToken`. Cette métrique indique le nombre d'appels au service de métadonnées d'instance qui utilisent IMDSv1 sur vos instances. Pour de plus amples informations, veuillez consulter Métriques des instances (p. 764).

Etape 3 : Une fois que tout est prêt sur toutes les instances

Tout est prêt sur toutes les instances lorsque la métrique CloudWatch `MetadataNoToken` enregistre une utilisation nulle d'IMDSv1. A ce stade, voici ce que vous pouvez faire :

- Pour les instances existantes : Vous pouvez imposer l'utilisation d'IMDSv2 via la commande modify-instance-metadata-options. Vous pouvez effectuer ces modifications sur les instances en cours d'exécution. Il n'est pas nécessaire de redémarrer vos instances.
- Pour les nouvelles instances : lors du lancement d'une nouvelle instance, vous pouvez effectuer l'une des opérations suivantes :
  - Dans l'assistant de lancement de la console Amazon EC2, définissez Métadonnées accessibles sur Activé et Version des métadonnées sur V2. Pour de plus amples informations, veuillez consulter Étape 3 : Configurer les détails de l'instance (p. 532).
  - Utilisez la commande run-instances pour spécifier que seul IMDSv2 doit être utilisé.

La mise à jour des options de métadonnées d'instance pour les instances existantes est disponible uniquement via l'API ou AWS CLI. À ce stade, elle n'est pas disponible dans la console Amazon EC2. Pour de plus amples informations, veuillez consulter Configuration des options des métadonnées d'instance (p. 705).

Etape 4 : Une fois opérée la transition de toutes vos instances vers IMDSv2

Les clés de condition IAM `ec2:MetadataHttpTokens`, `ec2:MetadataHttpPutResponseHopLimit` et `ec2:MetadataHttpEndpoint` peuvent être utilisées pour contrôler l'utilisation des API RunInstances et ModifyInstanceMetadataOptions et de l'interface de ligne commande correspondante. Si une stratégie est créée et qu'un paramètre de l'appel d'API ne correspond pas à l'état spécifié dans la stratégie à l'aide de la clé de condition, l'appel de l'API ou de l'interface de ligne commande échoue avec la réponse `UnauthorizedOperation`. Par exemple les stratégies IAM, consultez Utilisation des métadonnées d'instance (p. 1022).

Configuration des options des métadonnées d'instance

Les options de métadonnées d'instance vous permettent de configurer des instances nouvelles ou existantes aux fins suivantes :

- Imposer l'utilisation d'IMDSv2 lorsqu'il s'agit de demander des métadonnées d'instance

- Spécifier la durée de vie (hop limit) de la réponse `PUT`
- Désactiver l'accès aux métadonnées d'instance

Vous pouvez également utiliser des clés de condition IAM dans une stratégie IAM ou SCP pour les besoins suivants :

- Autoriser le lancement d'une instance uniquement si elle est configurée pour exiger l'utilisation d'IMDSv2
- Restreindre le nombre de sauts autorisés
- Désactiver l'accès aux métadonnées d'instance

Vous pouvez configurer les options de métadonnées d'instance lors du lancement de nouvelles instances à partir de la console Amazon EC2. Pour de plus amples informations, veuillez consulter Étape 3 : Configurer les détails de l'instance (p. 532).

Pour configurer les options des métadonnées d'instance sur de nouvelles instances ou des instances existantes, vous pouvez utiliser le kit SDK AWS ou l'AWS CLI. Pour de plus amples informations, veuillez consulter run-instances et modify-instance-metadata-options dans la AWS CLI Command Reference.

> **Note**
>
> Vous devez procéder avec précautions et effectuer des tests méticuleux avant toute modification. Notez les informations suivantes :
>
> - Si vous imposez l'utilisation de IMDSv2, les applications ou agents qui utilisent IMDSv1 pour l'accès aux métadonnées d'instance cesseront de fonctionner.
> - Si vous désactivez tous les accès aux métadonnées d'instance, les applications ou agents dont le fonctionnement repose sur l'accès aux métadonnées d'instance cesseront de fonctionner.

Rubriques

- Configuration des options de métadonnées d'instance pour les nouvelles instances (p. 706)
- Configuration des options de métadonnées d'instance pour les instances existantes (p. 707)

Configuration des options de métadonnées d'instance pour les nouvelles instances

Vous pouvez imposer l'utilisation d'IMDSv2 sur une instance au moment de son lancement. Vous pouvez également créer une stratégie IAM qui empêche les utilisateurs de lancer de nouvelles instances qui n'utilisent pas IMDSv2.

Console

Pour imposer l'utilisation d'IMDSv2 sur une nouvelle instance

- Lors du lancement d'une nouvelle instance dans la console Amazon EC2, sélectionnez les options suivantes sur la page Configurer les détails de l'instance :
  - Sous Détails avancés, pour Métadonnées accessibles, sélectionnez Activé.
  - Pour Version des métadonnées, sélectionnez V2 (jeton obligatoire).

Pour plus d'informations, consultez Étape 3 : Configurer les détails de l'instance (p. 532).

AWS CLI

Pour imposer l'utilisation d'IMDSv2 sur une nouvelle instance

L'exemple run-instances ci-dessous lance une instance `c3.large` avec `--metadata-options` défini sur `HttpTokens=required`. Lorsque vous spécifiez une valeur pour `HttpTokens`, vous devez également définir `HttpEndpoint` sur `enabled`. Comme l'en-tête de jeton sécurisé est défini sur `required` pour les demandes de récupération de métadonnées, cette option permet à l'instance d'imposer l'utilisation d'IMDSv2 lors de la demande de métadonnées d'instance.

```
aws ec2 run-instances
    --image-id ami-0abcdef1234567890
    --instance-type c3.large
 ...
    --metadata-options "HttpEndpoint=enabled,HttpTokens=required"
```

Pour imposer l'utilisation de IMDSv2 sur toutes les nouvelles instances

Pour vous assurer que les utilisateurs IAM peuvent uniquement lancer les instances nécessitant l'utilisation de IMDSv2 lors de la demande de métadonnées d'instance, vous pouvez spécifier que la condition imposant l'utilisation de IMDSv2 devra être remplie pour qu'une instance puisse être lancée. Pour examiner l'exemple de stratégie IAM, veuillez consulter Utilisation des métadonnées d'instance (p. 1022).

Console

Désactivation de l'accès aux métadonnées d'instance

- Pour vous assurer que l'accès à vos métadonnées d'instance est désactivé, quelle que soit la version du service de métadonnées d'instance que vous utilisez, lancez l'instance dans la console Amazon EC2 avec l'option suivante sélectionnée sur la page Configurer les détails de l'instance :
  - Sous Détails avancés, pour Métadonnées accessibles, sélectionnez Désactivé.

Pour plus d'informations, consultez Étape 3 : Configurer les détails de l'instance (p. 532).

AWS CLI

Désactivation de l'accès aux métadonnées d'instance

Pour vous assurer que l'accès aux métadonnées de votre instance est désactivé, quelle que soit la version du service de métadonnées d'instance que vous utilisez, lancez l'instance avec `--metadata-options` défini sur `HttpEndpoint=disabled`. Vous pouvez activer l'accès ultérieurement à l'aide de la commande modify-instance-metadata-options.

```
aws ec2 run-instances
    --image-id ami-0abcdef1234567890
    --instance-type c3.large
    ...
    --metadata-options "HttpEndpoint=disabled"
```

Configuration des options de métadonnées d'instance pour les instances existantes

Vous pouvez imposer l'utilisation d'IMDSv2 sur une instance existante. Vous pouvez également modifier la durée de vie (hop limit) de la réponse PUT et désactiver l'accès aux métadonnées d'instance sur une instance existante. De même, vous pouvez créer une stratégie IAM qui empêche les utilisateurs de modifier les options de métadonnées d'instance sur une instance existante.

Pour exiger l'utilisation de IMDSv2

Vous pouvez faire en sorte que l'utilisation de IMDSv2 soit obligatoire pour pouvoir demander des métadonnées d'instance. Utilisez la commande d'interface de ligne de commande modify-instance-

metadata-options et définissez le paramètre `http-tokens` sur `required`. Lorsque vous spécifiez une valeur pour `http-tokens`, vous devez également définir `http-endpoint` sur `enabled`.

```
aws ec2 modify-instance-metadata-options \
    --instance-id i-1234567898abcdef0 \
    --http-tokens required \
    --http-endpoint enabled
```

Pour modifier la durée de vie (hop limit) de la réponse PUT

Pour les instances existantes, vous pouvez modifier les paramètres de la durée de vie (hop limit) de la réponse `PUT`. Utilisez la commande d'interface de ligne de commande modify-instance-metadata-options et définissez le paramètre `http-put-response-hop-limit` sur la durée de vie (hop limit) requise. Dans l'exemple suivant, la durée de vie (hop limit) est définie `3`. Notez que lorsque vous spécifiez une valeur pour `http-put-response-hop-limit`, vous devez également définir `http-endpoint` sur `enabled`.

```
aws ec2 modify-instance-metadata-options \
    --instance-id i-1234567898abcdef0 \
    --http-put-response-hop-limit 3 \
    --http-endpoint enabled
```

Pour restaurer l'utilisation de IMDSv1 sur une instance utilisant IMDSv2

Vous pouvez utiliser la commande de l'interface de ligne de commande modify-instance-metadata-options avec `http-tokens` défini sur `optional` pour restaurer l'utilisation de IMDSv1 lors de la demande de métadonnées d'instance.

```
aws ec2 modify-instance-metadata-options \
    --instance-id i-1234567898abcdef0 \
    --http-tokens optional \
    --http-endpoint enabled
```

Désactivation de l'accès aux métadonnées d'instance

Vous pouvez désactiver l'accès à vos métadonnées d'instance en désactivant le point de terminaison HTTP du service des métadonnées d'instance, quelle que soit la version de ce dernier que vous utilisiez. Vous pouvez annuler cette modification à tout moment en activant à nouveau le point de terminaison HTTP. Utilisez la commande d'interface de ligne de commande modify-instance-metadata-options et définissez le paramètre `http-endpoint` sur `disabled`.

```
aws ec2 modify-instance-metadata-options \
    --instance-id i-1234567898abcdef0 \
    --http-endpoint disabled
```

Pour contrôler l'utilisation de la commande modify-instance-metadata-options

Pour contrôler les utilisateurs IAM habilités à modifier les options de métadonnées d'instance, spécifiez une stratégie qui empêche tous les utilisateurs d'utiliser l'API ModifyInstanceMetadataOptions hormis ceux dotés d'un rôle déterminé. Pour examiner l'exemple de stratégie IAM, veuillez consulter Utilisation des métadonnées d'instance (p. 1022).

Récupération des métadonnées d'instance

Puisque vos métadonnées d'instance sont disponibles à partir de votre instance en cours d'exécution, vous n'avez pas besoin d'utiliser la console Amazon EC2 ou l'AWS CLI. Cela peut être utile lorsque vous écrivez

des scripts à exécuter depuis votre instance. Par exemple, vous pouvez accéder à l'adresse IP locale de votre instance à partir des métadonnées d'instance afin de gérer une connexion à une application externe.

Les métadonnées d'instance sont divisées en plusieurs catégories. Pour obtenir une description de chaque catégorie de métadonnées d'instance, veuillez consulter Catégories de métadonnées d'instance (p. 721).

Pour voir toutes les catégories de métadonnées d'instance depuis une instance en cours d'exécution, utilisez l'URI ci-après.

```
http://169.254.169.254/latest/meta-data/
```

L'adresse IP `169.254.169.254` est une adresse de lien local et est uniquement valable à partir de l'instance. Pour plus d'informations, consultez Link-local address sur Wikipedia.

Veuillez noter que vous n'êtes pas facturé pour les requêtes HTTP utilisées pour récupérer les métadonnées d'instance et les données utilisateur.

Le format de la commande est différent selon que vous utilisez IMDSv1 ou IMDSv2. Par défaut, vous pouvez utiliser les deux services de métadonnées d'instance. Pour imposer l'utilisation de IMDSv2, veuillez consulter Configuration du service des métadonnées d'instance (p. 702).

Vous pouvez utiliser un outil tel que cURL, comme illustré dans l'exemple suivant.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/
```

Vous pouvez également télécharger l'outil de requête des métadonnées d'instance qui vous permet d'obtenir les métadonnées d'instance via Service des métadonnées d'instance Version 1 sans avoir à saisir la totalité de l'URI ou des noms de catégories.

Réponses et messages d'erreur

Toutes les métadonnées d'instance sont retournées sous forme de texte (type de contenu HTTP `text/plain`).

Une requête pour une ressource de métadonnées spécifique retourne la valeur appropriée ou un code d'erreur HTTP `404 - Not Found` si la ressource n'est pas disponible.

Une requête pour une ressource de métadonnées générale (l'URI se termine par un /) retourne une liste de ressources disponibles ou un code d'erreur HTTP `404 - Not Found` si une telle ressource n'existe pas. Les éléments de la liste se trouvent sur des lignes séparées se terminant par des sauts de ligne (ASCII 10).

Pour les demandes effectuées à l'aide d'Service des métadonnées d'instance Version 2, les codes d'erreur HTTP suivants peuvent être renvoyés :

- `400 - Missing or Invalid Parameters` – La demande `PUT` n'est pas valide.
- `401 - Unauthorized` – La demande `GET` utilise un jeton non valide. Il est recommandé dans ce cas de générer un nouveau jeton.

- `403 - Forbidden` – La demande n'est pas autorisée ou le service des métadonnées d'instance est désactivé.

Exemples de récupération des métadonnées d'instance

Exemples

- Obtenir les versions disponibles des métadonnées d'instance (p. 710)
- Obtenir les éléments de métadonnées de niveau supérieur (p. 711)
- Obtenir la liste des clés publiques disponibles (p. 713)
- Montrer les formats pour lesquels une clé publique 0 est disponible (p. 713)
- Obtenir la clé publique 0 (au format clé OpenSSH) (p. 713)
- Obtenir l'ID de sous-réseau d'une instance (p. 714)

Obtenir les versions disponibles des métadonnées d'instance

Cet exemple permet d'obtenir les versions disponibles des métadonnées d'instance. Ces versions ne coïncident pas nécessairement avec une version API Amazon EC2. Les versions antérieures sont disponibles au cas où vous ayez des scripts reposant sur la structure et les informations présentes dans une version précédente.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/
1.0
2007-01-19
2007-03-01
2007-08-29
2007-10-10
2007-12-15
2008-02-01
2008-09-01
2009-04-04
2011-01-01
2011-05-01
2012-01-12
2014-02-25
2014-11-05
2015-10-20
2016-04-19
2016-06-30
2016-09-02
latest
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/
1.0
2007-01-19
2007-03-01
2007-08-29
2007-10-10
2007-12-15
2008-02-01
2008-09-01
2009-04-04
2011-01-01
```

```
2011-05-01
2012-01-12
2014-02-25
2014-11-05
2015-10-20
2016-04-19
2016-06-30
2016-09-02
latest
```

Obtenir les éléments de métadonnées de niveau supérieur

Cet exemple permet d'obtenir les éléments de métadonnées de niveau supérieur. Pour plus d'informations, consultez Catégories de métadonnées d'instance (p. 721).

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/
ami-id
ami-launch-index
ami-manifest-path
block-device-mapping/
events/
hostname
iam/
instance-action
instance-id
instance-life-cycle
instance-type
local-hostname
local-ipv4
mac
metrics/
network/
placement/
profile
public-hostname
public-ipv4
public-keys/
reservation-id
security-groups
services/
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/
ami-id
ami-launch-index
ami-manifest-path
block-device-mapping/
events/
hostname
iam/
instance-action
instance-id
instance-type
local-hostname
local-ipv4
mac
```

```
metrics/
network/
placement/
profile
public-hostname
public-ipv4
public-keys/
reservation-id
security-groups
services/
```

Les exemples suivants permettent d'extraire les valeurs de certains de éléments de métadonnées de niveau supérieur qui ont été obtenus dans l'exemple précédent. Les demandes IMDSv2 utilisent le jeton stocké qui a été créé dans l'exemple de commande précédent, sous réserve qu'il ne soit pas arrivé à expiration.

IMDSv2

```
[ec2-user ~]$ curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/
latest/meta-data/ami-id
ami-0abcdef1234567890
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/ami-id
ami-0abcdef1234567890
```

IMDSv2

```
[ec2-user ~]$ curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/
latest/meta-data/reservation-id
r-0efghijk987654321
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/reservation-id
r-0efghijk987654321
```

IMDSv2

```
[ec2-user ~]$ curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/
latest/meta-data/local-hostname
ip-10-251-50-12.ec2.internal
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/local-hostname
ip-10-251-50-12.ec2.internal
```

IMDSv2

```
[ec2-user ~]$ curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/
latest/meta-data/public-hostname
ec2-203-0-113-25.compute-1.amazonaws.com
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/public-hostname
ec2-203-0-113-25.compute-1.amazonaws.com
```

Obtenir la liste des clés publiques disponibles

Cet exemple permet d'obtenir la liste des clés publiques disponibles.

IMDSv2

```
[ec2-user ~]$ `curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/public-keys/
0=my-public-key
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/public-keys/
0=my-public-key
```

Montrer les formats pour lesquels une clé publique 0 est disponible

Cet exemple montre les formats pour lesquels une clé publique 0 est disponible.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/public-keys/0/
openssh-key
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/public-keys/0/
openssh-key
```

Obtenir la clé publique 0 (au format clé OpenSSH)

Cet exemple permet d'obtenir la clé publique 0 (au format clé OpenSSH).

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
```

```
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/public-keys/0/openssh-key
ssh-rsa MIICiTCCAfICCQD6m7oRw0uXOjANBgkqhkiG9w0BAQUFADCBiDELMAkGA1UEBhMC
VVMxCzAJBgNVBAgTAldBMRAwDgYDVQQHEwdTZWF0dGxlMQ8wDQYDVQQKEwZBbWF6
b24xFDASBgNVBAsTC0lBTSBDb25zb2xlMRIwEAYDVQQDEwlUZXN0Q2lsYWMxHzAd
BgkqhkiG9w0BCQEWEG5vb25lQGFtYXpvbi5jb20wHhcNMTEwNDI1MjA0NTIxWhcN
MTIwNDI0MjA0NTIxWjCBiDELMAkGA1UEBhMCVVMxCzAJBgNVBAgTAldBMRAwDgYD
VQQHEwdTZWF0dGxlMQ8wDQYDVQQKEwZBbWF6b24xFDASBgNVBAsTC0lBTSBDb25z
b2xlMRIwEAYDVQQDEwlUZXN0Q2lsYWMxHzAdBgkqhkiG9w0BCQEWEG5vb25lQGFt
YXpvbi5jb20wgZ8wDQYJKoZIhvcNAQEBBQADgY0AMIGJAoGBAMaK0dn+a4GmWIWJ
21uUSfwfEvySWtC2XADZ4nB+BLYgVIk60CpiwsZ3G93vUEIO3IyNoH/f0wYK8m9T
rDHudUZg3qX4waLG5M43q7Wgc/MbQITxOUSQv7c7ugFFDzQGBzZswY6786m86gpE
Ibb3OhjZnzcvQAaRHhdlQWIMm2nrAgMBAAEwDQYJKoZIhvcNAQEFBQADgYEAtCu4
nUhVVxYUntneD9+h8Mg9q6q+auNKyExzyLwaxlAoo7TJHidbtS4J5iNmZgXL0Fkb
FFBjvSfpJIlJ00zbhNYS5f6GuoEDmFJl0ZxBHjJnyp378OD8uTs7fLvjx79LjSTb
NYiytVbZPQUQ5Yaxu2jXnimvw3rrszlaEXAMPLE my-public-key
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/public-keys/0/openssh-key
ssh-rsa MIICiTCCAfICCQD6m7oRw0uXOjANBgkqhkiG9w0BAQUFADCBiDELMAkGA1UEBhMC
VVMxCzAJBgNVBAgTAldBMRAwDgYDVQQHEwdTZWF0dGxlMQ8wDQYDVQQKEwZBbWF6
b24xFDASBgNVBAsTC0lBTSBDb25zb2xlMRIwEAYDVQQDEwlUZXN0Q2lsYWMxHzAd
BgkqhkiG9w0BCQEWEG5vb25lQGFtYXpvbi5jb20wHhcNMTEwNDI1MjA0NTIxWhcN
MTIwNDI0MjA0NTIxWjCBiDELMAkGA1UEBhMCVVMxCzAJBgNVBAgTAldBMRAwDgYD
VQQHEwdTZWF0dGxlMQ8wDQYDVQQKEwZBbWF6b24xFDASBgNVBAsTC0lBTSBDb25z
b2xlMRIwEAYDVQQDEwlUZXN0Q2lsYWMxHzAdBgkqhkiG9w0BCQEWEG5vb25lQGFt
YXpvbi5jb20wgZ8wDQYJKoZIhvcNAQEBBQADgY0AMIGJAoGBAMaK0dn+a4GmWIWJ
21uUSfwfEvySWtC2XADZ4nB+BLYgVIk60CpiwsZ3G93vUEIO3IyNoH/f0wYK8m9T
rDHudUZg3qX4waLG5M43q7Wgc/MbQITxOUSQv7c7ugFFDzQGBzZswY6786m86gpE
Ibb3OhjZnzcvQAaRHhdlQWIMm2nrAgMBAAEwDQYJKoZIhvcNAQEFBQADgYEAtCu4
nUhVVxYUntneD9+h8Mg9q6q+auNKyExzyLwaxlAoo7TJHidbtS4J5iNmZgXL0Fkb
FFBjvSfpJIlJ00zbhNYS5f6GuoEDmFJl0ZxBHjJnyp378OD8uTs7fLvjx79LjSTb
NYiytVbZPQUQ5Yaxu2jXnimvw3rrszlaEXAMPLE my-public-key
```

Obtenir l'ID de sous-réseau d'une instance

Cet exemple permet d'obtenir l'ID de sous-réseau pour une instance.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/network/interfaces/macs/02:29:96:8f:6a:2d/subnet-id
subnet-be9b61d7
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/network/interfaces/
macs/02:29:96:8f:6a:2d/subnet-id
subnet-be9b61d7
```

Limitation

Nous limitons les requêtes envoyées par chaque instance au service des métadonnées d'instance et appliquons des limites au nombre de connexions simultanées possible depuis une instance vers le service des métadonnées d'instance.

Si vous utilisez le service des métadonnées d'instance pour récupérer les informations d'identification de sécurité AWS, évitez de demander des informations d'identification pendant chaque transaction ou simultanément depuis un grand nombre de threads ou de processus, car cela risque d'entraîner des restrictions. Nous vous conseillons plutôt de placer les informations d'identification en cache jusqu'à ce que leur date d'expiration approche.

Si vous vous retrouvez limité alors que vous tentez d'accéder au service des métadonnées d'instance, renvoyez une requête avec une stratégie d'interruption exponentielle.

Limitation de l'accès au service des métadonnées d'instance

Vous pouvez envisager d'utiliser des règles de pare-feu locales pour désactiver l'accès au service des métadonnées d'instance à partir de certains ou de tous les processus.

Utilisation d'éléments iptables pour limiter l'accès

L'exemple suivant utilise des éléments Linux iptables et le module `owner` associé pour empêcher le serveur web Apache (en fonction de son ID utilisateur d'installation par défaut `apache`) d'accéder à l'adresse 169.254.169.254. Il utilise une règle deny pour rejeter toutes les demandes de métadonnées d'instance (IMDSv1 ou IMDSv2) de tout processus s'exécutant au nom de cet utilisateur.

```
$ sudo iptables --append OUTPUT --proto tcp --destination 169.254.169.254 --match owner --
uid-owner apache --jump REJECT
```

Vous pouvez aussi envisager d'autoriser uniquement l'accès à des utilisateurs ou des groupes particuliers à l'aide de règles d'autorisation (allow). Les règles allow peuvent être plus faciles à gérer du point de vue de la sécurité, car elles nécessitent que vous déterminiez quels sont les logiciels ayant besoin d'accéder aux métadonnées d'instance. Si vous utilisez des règles allow, vous risquez moins d'autoriser accidentellement un logiciel à accéder au service des métadonnées en cas de modification ultérieure des logiciels ou de la configuration sur une instance. Vous pouvez également combiner une utilisation de groupes avec des règles allow, afin de pouvoir ajouter et supprimer des utilisateurs dans un groupe autorisé sans avoir à modifier la règle du pare-feu.

L'exemple suivant empêche tous les processus d'accéder au service des métadonnées d'instance, à l'exception des processus qui s'exécutent dans le compte utilisateur `trustworthy-user`.

```
$ sudo iptables --append OUTPUT --proto tcp --destination 169.254.169.254 --match owner !
 --uid-owner trustworthy-user --jump REJECT
```

Note

- Pour utiliser des règles de pare-feu locales, vous devez adapter les commandes de l'exemple précédent à vos besoins.
- Par défaut, les règles iptables ne sont pas persistantes après un redémarrage du système. Elles peuvent être rendues persistantes en utilisant des fonctionnalités du système d'exploitation qui ne sont pas décrites ici.
- Le module iptables `owner` correspond uniquement à l'appartenance au groupe si le groupe est le groupe principal d'un utilisateur local donné. Les autres groupes n'ont pas de correspondance.

Utilisation de PF ou de IPFW pour limiter l'accès

Si vous utilisez FreeBSD ou OpenBSD, vous pouvez également envisager d'utiliser PF ou IPFW. Les exemples suivants permettent de limiter l'accès au service des métadonnées d'instance à l'utilisateur racine.

PF

```
$ block out inet proto tcp from any to 169.254.169.254
```

```
$ pass out inet proto tcp from any to 169.254.169.254 user root
```

IPFW

```
$ allow tcp from any to 169.254.169.254 uid root
```

```
$ deny tcp from any to 169.254.169.254
```

> **Note**
>
> L'ordre des commandes PF et IPFW a de l'importance. PF prend par défaut la valeur de la dernière règle correspondante et IPFW prend par défaut la valeur de la première règle correspondante.

Utilisation des données utilisateur d'instance

Lorsque vous utilisez des données utilisateur d'instance, ayez les points suivants à l'esprit :

- Les données utilisateur doivent être codées en base64. La console Amazon EC2 peut effectuer l'encodage base64 pour vous ou accepter les entrées codées en base64.
- Les données d'utilisateur sont limitées à 16 Ko en format brut, avant qu'elles soient encodées en base64. La taille d'une chaîne de longueur n après l'encodage base64 est ceil(n/3)*4.
- Les données utilisateur doivent être décodées en base64 lorsque vous les récupérez. Si vous les récupérez à l'aide des métadonnées d'instance ou de la console, les données sont décodées automatiquement.
- Les données utilisateur sont traitées comme des données opaques : ce que vous donnez est ce que vous obtenez en retour. Il appartient à l'instance d'être capable de l'interpréter.
- Si vous arrêtez une instance, modifiez ses données utilisateur et démarrez l'instance, les données utilisateurs mises à jour ne sont pas exécutées lorsque vous démarrez l'instance.

Spécification des données utilisateur d'instance au moment du lancement

Vous pouvez spécifier des données utilisateur lorsque vous lancez une instance. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530) et Exécution de commandes sur votre instance Linux lors du lancement (p. 694).

Modification des données utilisateur d'instance

Vous pouvez modifier les données utilisateur pour une instance à l'état arrêté si le volume racine est un volume EBS. Pour plus d'informations, consultez Affichage et mise à jour des données utilisateur d'instance (p. 697).

Récupération des données utilisateur d'instance

Pour récupérer des données utilisateur depuis une instance en cours d'exécution, utilisez l'URI ci-après.

```
http://169.254.169.254/latest/user-data
```

Une demande de données utilisateur renvoie les données telles qu'elles sont (type de contenu `application/octet-stream`).

Cet exemple renvoie les données utilisateur qui ont été fournies sous la forme de texte séparé par des virgules.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/user-
data
1234,john,reboot,true | 4512,richard, | 173,,,
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/user-data
1234,john,reboot,true | 4512,richard, | 173,,,
```

Cet exemple renvoie des données utilisateur qui ont été fournies sous la forme d'un script.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/user-
data
#!/bin/bash
yum update -y
service httpd start
chkconfig httpd on
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/user-data
#!/bin/bash
yum update -y
service httpd start
chkconfig httpd on
```

Pour récupérer des données utilisateur pour une instance à partir de votre ordinateur, consultez Données utilisateur et AWS CLI (p. 699)

Récupération de données dynamiques

Pour récupérer des données dynamiques depuis une instance en cours d'exécution, utilisez l'URI ci-après.

```
http://169.254.169.254/latest/dynamic/
```

Cet exemple montre comment récupérer les catégories d'identité d'instance de haut niveau.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/dynamic/
instance-identity/
rsa2048
```

```
pkcs7
document
signature
dsa2048
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/dynamic/instance-identity/
rsa2048
pkcs7
document
signature
dsa2048
```

Pour plus d'informations sur les données dynamiques et pour des exemples sur la façon de les récupérer, consultez Documents d'identité d'instance (p. 728).

Exemple : Valeur d'index de lancement AMI

Cet exemple illustre comment vous pouvez utiliser à la fois les données utilisateur et les métadonnées d'instance pour configurer vos instances.

Alice souhaite lancer quatre instances de son AMI de base de données préférée ; la première instance servant d'instance initiale et les trois autres de réplicas. Lorsqu'elle les lance, elle souhaite ajouter des données utilisateur portant sur la stratégie de réplication pour chaque réplica. Elle sait que ces données seront disponibles pour les quatre instances. Elle a donc besoin de structurer les données utilisateur de sorte que chaque instance reconnaisse quelles parties la concernent. Pour ce faire, elle peut utiliser la valeur de métadonnées d'instance `ami-launch-index` qui sera unique pour chaque instance.

Voici les données utilisateur construites par Alice.

```
replicate-every=1min | replicate-every=5min | replicate-every=10min
```

La donnée `replicate-every=1min` définit la configuration du premier réplica, `replicate-every=5min` définit la configuration du deuxième réplica, et ainsi de suite. Alice décide de fournir ces données sous la forme d'une chaîne ASCII avec un symbole barre verticale (`|`) délimitant les données pour les différentes instances.

Alice lance quatre instances à l'aide de la commande run-instances, en spécifiant les données utilisateur.

```
aws ec2 run-instances \
    --image-id ami-0abcdef1234567890 \
    --count 4 \
    --instance-type t2.micro \
    --user-data "replicate-every=1min | replicate-every=5min | replicate-every=10min"
```

Après leur lancement, toutes les instances ont une copie des données utilisateur et des métadonnées communes présentées ici :

- ID D'AMI : ami-0abcdef1234567890
- ID de réservation : r-1234567890abcabc0
- Clés publiques : aucune
- Nom du groupe de sécurité : par défaut
- Type d'instance : t2.micro

Toutefois, chaque instance possède certaines métadonnées uniques.

Instance 1

Metadonnées	Valeur
instance-id	i-1234567890abcdef0
ami-launch-index	0
public-hostname	ec2-203-0-113-25.compute-1.amazonaws.com
public-ipv4	67.202.51.223
local-hostname	ip-10-251-50-12.ec2.internal
local-ipv4	10.251.50.35

Instance 2

Metadonnées	Valeur
instance-id	i-0598c7d356eba48d7
ami-launch-index	1
public-hostname	ec2-67-202-51-224.compute-1.amazonaws.com
public-ipv4	67.202.51.224
local-hostname	ip-10-251-50-36.ec2.internal
local-ipv4	10.251.50.36

Instance 3

Metadonnées	Valeur
instance-id	i-0ee992212549ce0e7
ami-launch-index	2
public-hostname	ec2-67-202-51-225.compute-1.amazonaws.com
public-ipv4	67.202.51.225
local-hostname	ip-10-251-50-37.ec2.internal
local-ipv4	10.251.50.37

Instance 4

Metadonnées	Valeur
instance-id	i-1234567890abcdef0
ami-launch-index	3
public-hostname	ec2-67-202-51-226.compute-1.amazonaws.com

Metadonnées	Valeur
public-ipv4	67.202.51.226
local-hostname	ip-10-251-50-38.ec2.internal
local-ipv4	10.251.50.38

Alice peut utiliser la valeur `ami-launch-index` pour déterminer quelle portion des données utilisateur est applicable à une instance particulière.

1. Elle se connecte à l'une des instances et récupère `ami-launch-index` pour cette instance afin de s'assurer qu'il s'agit de l'un des réplicas :

   IMDSv2

   ```
   [ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/meta-data/api/token"
    -H "X-aws-ec2-metadata-token-ttl-seconds: 21600"` \
   && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
   data/ami-launch-index
   2
   ```

   Pour les étapes suivantes, les demandes IMDSv2 utilisent le jeton stocké provenant de la commande IMDSv2 précédente, à condition que le jeton ne soit pas arrivé à expiration.

   IMDSv1

   ```
   [ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/ami-launch-index
   2
   ```

2. Elle enregistre les données `ami-launch-index` sous forme de variable.

   IMDSv2

   ```
   [ec2-user ~]$ ami_launch_index=`curl -H "X-aws-ec2-metadata-token: $TOKEN" -v
    http://169.254.169.254/latest/meta-data/ami-launch-index`
   ```

   IMDSv1

   ```
   [ec2-user ~]$ ami_launch_index=`curl http://169.254.169.254/latest/meta-data/ami-
   launch-index`
   ```

3. Elle enregistre les données utilisateur sous forme de variable.

   IMDSv2

   ```
   [ec2-user ~]$ user_data=`curl -H "X-aws-ec2-metadata-token: $TOKEN" -v
    http://169.254.169.254/latest/user-data`
   ```

   IMDSv1

   ```
   [ec2-user ~]$ user_data=`curl http://169.254.169.254/latest/user-data`
   ```

4. Enfin, Alice utilise la commande cut pour extraire la portion de données utilisateur applicable à cette instance.

   IMDSv2

   ```
   [ec2-user ~]$ echo $user_data | cut -d"|" -f"$ami_launch_index"replicate-every=5min
   ```

IMDSv1

```
[ec2-user ~]$ echo $user_data | cut -d"|" -f"$ami_launch_index"
replicate-every=5min
```

Catégories de métadonnées d'instance

Le tableau ci-après répertorie les catégories de métadonnées d'instance.

> **Note**
>
> Lorsque Amazon EC2 libère une nouvelle catégorie de métadonnées d'instance, les métadonnées d'instance de la nouvelle catégorie peuvent ne pas être disponibles pour les instances existantes. Pour vous assurer que les métadonnées d'instance sont disponibles pour une instance existante, vous devez arrêter, puis démarrer (p. 625) l'instance.

> **Important**
>
> Certains des noms de catégorie dans le tableau suivant sont des espaces réservés pour les données qui sont uniques à votre instance. Par exemple, *mac* représente l'adresse MAC de l'interface réseau. Les espaces réservés doivent être remplacés par les valeurs réelles.

non structurées	Description	Date de parution
`ami-id`	L'ID d'AMI utilisé pour lancer l'instance.	Version 1.0
`ami-launch-index`	Si vous avez démarré plus d'une instance en même temps, cette valeur indique l'ordre dans lequel l'instance a été lancée. La valeur 0 indique la première instance lancée.	Version 1.0
`ami-manifest-path`	Chemin d'accès du fichier manifeste d'AMI dans Amazon S3. Si vous avez utilisé une AMI basée sur Amazon EBS pour lancer l'instance, le résultat retourné est `unknown`.	Version 1.0
`ancestor-ami-ids`	Les ID d'AMI de toutes les instances qui ont été regroupées pour créer cette AMI. Cette valeur n'existera que si le fichier manifeste d'AMI contenait une clé `ancestor-amis`.	2007-10-10
`block-device-mapping/ami`	Le périphérique virtuel qui contient le système de fichiers racine/démarrage.	2007-12-15
`block-device-mapping/ebs` N	Les périphériques virtuels associés aux volumes Amazon EBS. Les volumes Amazon EBS ne sont disponibles dans les métadonnées que s'ils étaient présents au moment du lancement ou lorsque l'instance a été démarrée pour la dernière fois. Le N indique l'index du volume Amazon EBS (tel que `ebs1` ou `ebs2`).	2007-12-15

non structurées	Description	Date de parution
`block-device-mapping/ephemeral` N	Les appareils virtuels pour les volumes de stockage d'instances non NVMe. Le N indique l'index de chaque volume. Le nombre de volumes de stockage d'instances dans le mappage d'appareils en bloc peut ne pas correspondre au nombre réel de volumes de stockage d'instances pour l'instance. Le type d'instance détermine le nombre de volumes de stockage d'instances disponibles pour une instance. Si le nombre de volumes de stockage d'instances dans un mappage d'appareils en bloc dépasse le nombre disponible pour une instance, les volumes de stockage d'instances supplémentaires sont ignorés.	2007-12-15
`block-device-mapping/root`	Les périphériques ou partitions virtuels associés aux périphériques ou partitions racines sur le périphérique virtuel où le système de fichiers racine (/ ou C:) est associé avec l'instance donnée.	2007-12-15
`block-device-mapping/swap`	Les périphériques virtuels associés avec `swap`. Pas toujours présents.	2007-12-15
`elastic-gpus/associations/`*`elastic-gpu-id`*	Si un Elastic GPU est attaché à l'instance, contient une chaîne JSON avec des informations sur l'Elastic GPU, notamment son ID et ses informations de connexion.	2016-11-30
`elastic-inference/associations/`*`eia-id`*	Si un accélérateur Elastic Inference est attaché à l'instance, contient une chaîne JSON avec des informations sur l'accélérateur Elastic Inference, notamment son ID et son type.	2018-11-29
`events/maintenance/history`	S'il y a des événements de maintenance terminés ou annulés pour l'instance, contient une chaîne JSON avec des informations sur ces événements. Pour plus d'informations, consultez Pour afficher l'historique des événements terminés ou annulés (p. 753).	2018-08-17

non structurées	Description	Date de parution
`events/maintenance/scheduled`	S'il y a des événements de maintenance activés pour l'instance, contient une chaîne JSON avec des informations sur ces événements. Pour plus d'informations, consultez Affichage des événements planifiés (p. 750).	2018-08-17
`hostname`	Le nom d'hôte DNS IPv4 privé de l'instance. Dans le cas où plusieurs interfaces réseau sont présentes, cela fait référence au périphérique eth0 (le périphérique dont le numéro de périphérique est 0).	Version 1.0
`iam/info`	Si un rôle IAM est associé à l'instance, il contient des informations concernant la dernière mise à jour du profil d'instance, parmi lesquelles la date de dernière mise à jour (LastUpdated), l'InstanceProfileArn et l'InstanceProfileId de l'instance. Sinon, absent.	2012-01-12
`iam/security-credentials/role-name`	Si un rôle IAM est associé à l'instance, *`nom-rôle`* est le nom du rôle et *`nom-rôle`* contient les informations d'identification de sécurité temporaires associées au rôle (pour plus d'informations, consultez Extraction des informations d'identification de sécurité à partir des métadonnées d'instance (p. 1036)). Sinon, absent.	2012-01-12
`identity-credentials/ec2/info`	[Usage interne uniquement] Informations sur les informations d'identification dans `identity-credentials/ec2/security-credentials/ec2-instance`. Ces informations d'identification sont utilisées par des fonctions AWS telles que EC2 Instance Connect et ne disposent pas d'autorisations ou de privilèges d'API AWS supplémentaires au-delà de l'identification de l'instance.	2018-05-23

non structurées	Description	Date de parution
`identity-credentials/ec2/security-credentials/ec2-instance`	[Usage interne uniquement] Informations d'identification permettant au logiciel sur les instances de s'identifier auprès d'AWS pour prendre en charge des fonctions telles que EC2 Instance Connect. Ces informations d'identification ne disposent pas d'autorisations ou de privilèges d'API AWS supplémentaires.	2018-05-23
`instance-action`	Informe l'instance qu'elle devrait redémarrer en vue de la création d'un bundle. Valeurs valides : `none` | `shutdown` | `bundle-pending`.	2008-09-01
`instance-id`	L'ID de cette instance.	Version 1.0
`instance-life-cycle`	Option d'achat de cette instance. Pour plus d'informations, consultez Options d'achat d'instance (p. 323).	01-10-2019
`instance-type`	Le type d'instance. Pour plus d'informations, consultez Types d'instance (p. 213).	2007-08-29
`kernel-id`	L'ID du noyau lancé avec l'instance, le cas échéant.	2008-02-01
`local-hostname`	Le nom d'hôte DNS IPv4 privé de l'instance. Dans le cas où plusieurs interfaces réseau sont présentes, cela fait référence au périphérique eth0 (le périphérique dont le numéro de périphérique est 0).	2007-01-19
`local-ipv4`	L'adresse IPv4 privée de l'instance. Dans le cas où plusieurs interfaces réseau sont présentes, cela fait référence au périphérique eth0 (le périphérique dont le numéro de périphérique est 0).	Version 1.0
`mac`	L'adresse de contrôle d'accès média (MAC) de l'instance. Dans le cas où plusieurs interfaces réseau sont présentes, cela fait référence au périphérique eth0 (le périphérique dont le numéro de périphérique est 0).	2011-01-01
`metrics/vhostmd`	Plus disponible.	2011-05-01

non structurées	Description	Date de parution
`network/interfaces/macs/mac/device-number`	Le numéro de périphérique unique associé à cette interface. Le numéro de périphérique correspond au nom du périphérique, par exemple un `device-number` de 2 est pour le périphérique eth2. Cette catégorie correspond aux champs `DeviceIndex` et `device-index` utilisés par l'API Amazon EC2 et les commandes EC2 pour l'AWS CLI.	2011-01-01
`network/interfaces/macs/mac/interface-id`	L'ID de l'interface réseau.	2011-01-01
`network/interfaces/macs/mac/ipv4-associations/public-ip`	Les adresses IPv4 privées qui sont associées à chaque adresse IP publique et assignées à cette interface.	2011-01-01
`network/interfaces/macs/mac/ipv6s`	Les adresses IPv6 privées associées à l'interface. Retournés uniquement pour les instances lancées dans un VPC.	2016-06-30
`network/interfaces/macs/mac/local-hostname`	Le nom d'hôte local de l'interface.	2011-01-01
`network/interfaces/macs/mac/local-ipv4s`	Les adresses IPv4 privées associées à l'interface.	2011-01-01
`network/interfaces/macs/mac/mac`	L'adresse MAC de l'instance.	2011-01-01
`network/interfaces/macs/mac/owner-id`	L'ID du propriétaire de l'interface réseau. Dans les environnements à interfaces multiples, une interface peut être attachée à un tiers, par exemple Elastic Load Balancing. Le trafic sur l'interface est toujours facturé au propriétaire de l'interface.	2011-01-01
`network/interfaces/macs/mac/public-hostname`	Le DNS public de l'interface (IPv4). Cette catégorie n'est retournée que si l'attribut `enableDnsHostnames` est défini comme `true`. Pour plus d'informations, consultez Utilisation de DNS avec votre VPC.	2011-01-01
`network/interfaces/macs/mac/public-ipv4s`	L'adresse IP publique ou les adresses IP Elastic associées à l'interface. Il peut y avoir plusieurs adresses IPv4 sur une instance.	2011-01-01
`network/interfaces/macs/mac/security-groups`	Les groupes de sécurité auxquels l'interface réseau appartient.	2011-01-01

non structurées	Description	Date de parution
`network/interfaces/macs/mac/security-group-ids`	Les ID des groupes de sécurité auxquels l'interface réseau appartient.	2011-01-01
`network/interfaces/macs/mac/subnet-id`	L'ID du sous-réseau (subnet) dans lequel l'interface réside.	2011-01-01
`network/interfaces/macs/mac/subnet-ipv4-cidr-block`	Le bloc d'adresse CIDR IPv4 du sous-réseau dans lequel l'interface réside.	2011-01-01
`network/interfaces/macs/mac/subnet-ipv6-cidr-blocks`	Le bloc d'adresse CIDR IPv6 du sous-réseau dans lequel l'interface réside.	2016-06-30
`network/interfaces/macs/mac/vpc-id`	L'ID du VPC dans lequel l'interface réside.	2011-01-01
`network/interfaces/macs/mac/vpc-ipv4-cidr-block`	Le bloc d'adresse CIDR IPv4 principal pour le VPC.	2011-01-01
`network/interfaces/macs/mac/vpc-ipv4-cidr-blocks`	Les blocs d'adresse CIDR IPv4 pour le VPC.	2016-06-30
`network/interfaces/macs/mac/vpc-ipv6-cidr-blocks`	Le bloc d'adresse CIDR IPv6 du VPC dans lequel l'interface réside.	2016-06-30
`placement/availability-zone`	La zone de disponibilité dans laquelle l'instance a été lancée.	2008-02-01
`placement/availability-zone-id`	ID de zone de disponibilité statique dans laquelle l'instance est lancée. L'ID de zone de disponibilité est cohérent entre les comptes. Toutefois, il peut être différent de la zone de disponibilité, qui peut varier selon le compte.	2020-08-24
`placement/group-name`	Nom du groupe de placement dans lequel l'instance est lancée.	2020-08-24
`placement/host-id`	ID de l'hôte sur lequel l'instance est lancée. Applicable uniquement aux Hôtes dédiés.	2020-08-24
`placement/partition-number`	Numéro de la partition dans laquelle l'instance est lancée.	2020-08-24
`placement/region`	Région AWS dans laquelle l'instance est lancée.	2020-08-24
`product-codes`	Les éventuels codes produit AWS Marketplace associés à l'instance.	2007-03-01

non structurées	Description	Date de parution
`public-hostname`	Le DNS public de l'instance. Cette catégorie n'est retournée que si l'attribut `enableDnsHostnames` est défini comme `true`. Pour plus d'informations, consultez Utilisation de DNS avec votre VPC dans le Amazon VPC Guide de l'utilisateur.	2007-01-19
`public-ipv4`	L'adresse IPv4 publique. Si une adresse IP Elastic est associée à l'instance, la valeur retournée est l'adresse IP Elastic.	2007-01-19
`public-keys/0/openssh-key`	Clé publique. Disponible uniquement si fournie au moment du lancement de l'instance.	Version 1.0
`ramdisk-id`	L'ID du disque RAM spécifié au moment du lancement, le cas échéant.	2007-10-10
`reservation-id`	L'ID de la réservation.	Version 1.0
`security-groups`	Les noms des groupes de sécurité appliqués à l'instance. Après le lancement, vous pouvez modifier les groupes de sécurité des instances. De tels changements apparaissent ici et dans réseau/interfaces/macs/*mac*/groupes-sécurité.	Version 1.0
`services/domain`	Le domaine des ressources AWS de la région.	2014-02-25
`services/partition`	Partition dans laquelle se trouve la ressource. Pour les régions AWS standard, la partition est `aws`. Si vous avez des ressources dans d'autres partitions, la partition est `aws-`*`partitionname`*. Par exemple, la partition des ressources de la région Chine (Beijing) est `aws-cn`.	2015-10-20
`spot/instance-action`	L'action (hibernation, arrêt ou résiliation) et l'heure approximative (UTC) à laquelle l'action aura lieu. Cet élément est présent uniquement si l'instance Spot a été marquée pour hiberner, ou être arrêtée ou résiliée. Pour plus d'informations, consultez instance-action (p. 458).	2016-11-15

non structurées	Description	Date de parution
`spot/termination-time`	L'heure approximative (indiquée au format UTC) à laquelle le système d'exploitation de votre instance Spot recevra le signal d'arrêt. Cet élément n'est présent et ne contient une valeur temporelle (par exemple, 2015-01-05T18:02:00Z) que si l'instance Spot a été marquée en vue de son arrêt par Amazon EC2. L'élément heure-arrêt n'est pas défini à une heure précise si vous avez mis fin vous-même à l'instance Spot. Pour plus d'informations, consultez termination-time (p. 458).	2014-11-05

Catégories de données dynamiques

Le tableau ci-après répertorie les catégories de données dynamiques.

non structurées	Description	Date de parution
`fws/instance-monitoring`	Valeur indiquant si le client a activé la surveillance détaillée toutes les minutes dans CloudWatch. Valeurs valides : `enabled | disabled`	2009-04-04
`instance-identity/document`	JSON contenant les attributs d'instance, tels que l'ID d'instance, l'adresse IP privée, etc. Voir Documents d'identité d'instance (p. 728).	2009-04-04
`instance-identity/pkcs7`	Utilisé pour vérifier l'authenticité et le contenu du document par rapport à la signature. Voir Documents d'identité d'instance (p. 728).	2009-04-04
`instance-identity/signature`	Les données pouvant être utilisées par d'autres pour vérifier leur origine et leur authenticité. Voir Documents d'identité d'instance (p. 728).	2009-04-04

Documents d'identité d'instance

Chaque instance que vous lancez a un Documents d'identité d'instance qui fournit des informations sur l'instance elle-même. Vous pouvez utiliser le Documents d'identité d'instance pour valider les attributs de l'instance.

Le Documents d'identité d'instance est généré lorsque l'instance est lancée et qu'elle est exposée (au format JSON en texte brut) via le service de métadonnées d'instance. L'adresse IP `169.254.169.254` est une adresse de lien local et est uniquement valable à partir de l'instance. Pour plus d'informations, consultez Link-local address sur Wikipedia.

Vous pouvez récupérer le Documents d'identité d'instance à partir d'une instance en cours d'exécution à tout moment. Le Documents d'identité d'instance contient les informations suivantes :

non structurées	Description
`devpayProductCodes`	Obsolète.
`marketplaceProductCodes`	Code produit AWS Marketplace de l'AMI utilisée pour lancer l'instance.
`availabilityZone`	Zone de disponibilité dans laquelle l'instance est en cours d'exécution.
`privateIp`	L'adresse IPv4 privée de l'instance.
`version`	La version du format du Documents d'identité d'instance
`instanceId`	ID de l'instance.
`billingProducts`	Code produit de facturation de l'AMI utilisée pour lancer l'instance.
`instanceType`	Type de l'instance.
`accountId`	ID du compte AWS qui a lancé l'instance.
`imageId`	ID de l'AMI utilisée pour lancer l'instance.
`pendingTime`	Date et heure auxquelles l'instance a été lancée.
`architecture`	Architecture de l'AMI utilisée pour lancer l'instance (i386 | x86_64 | arm64).
`kernelId`	ID du noyau associé à l'instance, le cas échéant.
`ramdiskId`	ID du disque RAM associé à cette instance, le cas échéant.
`region`	Région dans laquelle l'instance est en cours d'exécution.

Récupérer le Documents d'identité d'instance en texte brut

Pour récupérer le Documents d'identité d'instance en texte brut

Connectez-vous à l'instance et exécutez l'une des commandes suivantes en fonction de la version IMDS (Instance Metadata Service) utilisée par l'instance.

IMDSv2

```
$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-metadata-
token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/dynamic/
instance-identity/document
```

IMDSv1

```
$ curl http://169.254.169.254/latest/dynamic/instance-identity/document
```

Voici un exemple de sortie.

```
{
    "devpayProductCodes" : null,
    "marketplaceProductCodes" : [ "1abc2defghijklm3nopqrs4tu" ],
    "availabilityZone" : "us-west-2b",
    "privateIp" : "10.158.112.84",
    "version" : "2017-09-30",
```

```
    "instanceId" : "i-1234567890abcdef0",
    "billingProducts" : null,
    "instanceType" : "t2.micro",
    "accountId" : "123456789012",
    "imageId" : "ami-5fb8c835",
    "pendingTime" : "2016-11-19T16:32:11Z",
    "architecture" : "x86_64",
    "kernelId" : null,
    "ramdiskId" : null,
    "region" : "us-west-2"
}
```

Vérification du Documents d'identité d'instance

Si vous avez l'intention d'utiliser le contenu du Documents d'identité d'instance à des fins importantes, vous devez vérifier son contenu et son authenticité avant de l'utiliser.

Le Documents d'identité d'instance en texte brut est accompagné de trois signatures hachées et chiffrées. Vous pouvez utiliser ces signatures pour vérifier l'origine et l'authenticité du Documents d'identité d'instance et les informations qu'il contient. Les signatures suivantes sont fournies :

- Signature codée en base64 — Il s'agit d'un hachage SHA256 codé en base64 du Documents d'identité d'instance qui est chiffré à l'aide d'une paire de clés RSA.
- Signature PKCS7 — Il s'agit d'un hachage SHA1 du Documents d'identité d'instance qui est chiffré à l'aide d'une paire de clés DSA.
- Signature RSA-2048 — Il s'agit d'un hachage SHA256 du Documents d'identité d'instance qui est chiffré à l'aide d'une paire de clés RSA-2048.

Chaque signature est disponible à un point de terminaison différent dans les métadonnées de l'instance. Vous pouvez utiliser l'une de ces signatures en fonction de vos exigences de hachage et de chiffrement. Pour vérifier les signatures, vous devez utiliser le certificat AWS public correspondant.

> **Important**
>
> Pour valider le Documents d'identité d'instance avec la signature encodée en base64 ou la signature RSA2048, vous devez demander le certificat AWS public correspondant à AWS Support.

Les rubriques suivantes fournissent des étapes détaillées pour valider le Documents d'identité d'instance avec chaque signature.

- Utilisation de la signature PKCS7 pour vérifier le Documents d'identité d'instance (p. 730)
- Utilisation de la signature codée en base64 pour vérifier le Documents d'identité d'instance (p. 734)
- Utilisation de la signature RSA-2048 pour vérifier le Documents d'identité d'instance (p. 735)

Utilisation de la signature PKCS7 pour vérifier le Documents d'identité d'instance

Cette rubrique explique comment vérifier le Documents d'identité d'instance avec la signature PKCS7 et le certificat public AWS DSA.

Pour vérifier le Documents d'identité d'instance avec la signature PKCS7 et le certificat public AWS DSA

1. Connectez-vous à l'instance.
2. Récupérez la signature PKCS7 à partir des métadonnées de l'instance et ajoutez-la à un fichier nommé `pkcs7`.

   a. Ajoutez l'en-tête `-----BEGIN PKCS7-----` au fichier `pkcs7`.

```
$ echo "-----BEGIN PKCS7-----" > pkcs7
```

b. Récupérez la signature PKCS7 à partir des métadonnées de l'instance et ajoutez-la au fichier `pkcs7`. Utilisez l'une des commandes suivantes en fonction de la version IMDS utilisée par l'instance.

IMDSv2

```
$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
dynamic/instance-identity/pkcs7 >> pkcs7
```

IMDSv1

```
$ curl -s http://169.254.169.254/latest/dynamic/instance-identity/pkcs7
 >> pkcs7
```

c. Ajoutez le pied de page `-----END PKCS7-----` à une nouvelle ligne du fichier `pkcs7`.

```
$ echo "" >> pkcs7
$ echo "-----END PKCS7-----" >> pkcs7
```

3. Ajoutez le contenu de l'Documents d'identité d'instance des métadonnées de l'instance à un fichier nommé `document`. Utilisez l'une des commandes suivantes en fonction de la version IMDS utilisée par l'instance.

IMDSv2

```
$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
metadata-token-ttl-seconds: 21600"` \ && curl -H "X-aws-ec2-metadata-token: $TOKEN"
 -v http://169.254.169.254/latest/dynamic/instance-identity/document > document
```

IMDSv1

```
$ curl -s http://169.254.169.254/latest/dynamic/instance-identity/document
 > document
```

4. Ajoutez le certificat public AWS DSA à un fichier nommé `certificate`.

a. Créez le fichier `certificate`.

```
$ touch certificate
```

b. Ouvrez le fichier `certificate` à l'aide de votre éditeur de texte préféré et ajoutez le contenu du certificat public AWS DSA. Choisissez le certificat approprié pour la région AWS dans laquelle se trouve votre instance.

> **Important**
>
> Si le certificat public AWS DSA pour votre région n'est pas répertorié ci-dessous, contactez AWS Support.

Other AWS Regions

Le certificat AWS public suivant s'applique à toutes les régions AWS, à l'exception de Hong Kong, Bahreïn, Chine et GovCloud.

```
-----BEGIN CERTIFICATE-----
MIIC7TCCAq0CCQCWukjZ5V4aZzAJBgcqhkjOOAQDMFwxCzAJBgNVBAYTAlVTMRkw
FwYDVQQIExBXYXNoaW5ndG9uIFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYD
VQQKExdBbWF6b24gV2ViIFNlcnZpY2VzIExMQzAeFw0xMjAxMDUxMjU2MTJaFw0z
ODAxMDUxMjU2MTJaMFwxCzAJBgNVBAYTAlVTMRkwFwYDVQQIExBXYXNoaW5ndG9u
IFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYDVQQKExdBbWF6b24gV2ViIFNl
cnZpY2VzIExMQzCCAbcwggEsBgcqhkjOOAQBMIIBHwKBgQCjkvcS2bb1VQ4yt/5e
ih5OO6kK/n1Lzllr7D8ZwtQP8fOEpp5E2ng+D6Ud1Z1gYipr58Kj3nssSNpI6bX3
VyIQzK7wLclnd/YozqNNmgIyZecN7EglK9ITHJLP+x8FtUpt3QbyYXJdmVMegN6P
hviYt5JH/nYl4hh3Pa1HJdskgQIVALVJ3ER11+Ko4tP6nwvHwh6+ERYRAoGBAI1j
k+tkqMVHuAFcvAGKocTgsjJem6/5qomzJuKDmbJNu9Qxw3rAotXau8Qe+MBcJl/U
hhy1KHVpCGl9fueQ2s6IL0CaO/buycU1CiYQk40KNHCcHfNiZbdlx1E9rpUp7bnF
lRa2v1ntMX3caRVDdbtPEWmdxSCYsYFDk4mZrOLBA4GEAAKBgEbmeve5f8LIE/Gf
MNmP9CM5eovQOGx5ho8WqD+aTebs+k2tn92BBPqeZqpWRa5P/+jrdKml1qx4llHW
MXrs3IgIb6+hUIB+S8dz8/mmO0bpr76RoZVCXYab2CZedFut7qc3WUH9+EUAH5mw
vSeDCOUMYQR7R9LINYwouHIziqQYMAkGByqGSM44BAMDLwAwLAIUWXBlk40xTwSw
7HX32MxXYruse9ACFBNGmdX2ZBrVNGrN9N2f6ROk0k9K
-----END CERTIFICATE-----
```


Hong Kong Region

Le certificat public AWS de la région Hong Kong est le suivant :


```
-----BEGIN CERTIFICATE-----
MIIC7zCCAq4CCQCO7MJe5Y3VLjAJBgcqhkjOOAQDMFwxCzAJBgNVBAYTAlVTMRkw
FwYDVQQIExBXYXNoaW5ndG9uIFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYD
VQQKExdBbWF6b24gV2ViIFNlcnZpY2VzIExMQzAeFw0xOTAyMDMwMjIxMjFaFw00
NTAyMDMwMjIxMjFaMFwxCzAJBgNVBAYTAlVTMRkwFwYDVQQIExBXYXNoaW5ndG9u
IFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYDVQQKExdBbWF6b24gV2ViIFNl
cnZpY2VzIExMQzCCAbgwggEsBgcqhkjOOAQBMIIBHwKBgQDvQ9RzVvf4MAwGbqfX
blCvCoVb9957OkLGn/04CowHXJ+vTBR7eyIa6AoXltsQXBOmrJswToFKKxT4gbuw
jK7s9QQX4CmTRWcEgO2RXtZSVjOhsUQMh+yf7Ht4OVL97LWnNfGsX2cwjcRWHYgI
7lvnuBNBzLQHdSEwMNq0Bk76PwIVAMan6XIEEPnwr4e6u/RNnWBGKd9FAoGBAOCG
eSNmxpW4QFu4pIlAykm6EnTZKKHT87gdXkAkfoC5fAfOxxhnE2HezZHp9Ap2tMV5
8bWNvoPHvoKCQqwfm+OUBlAxC/3vqoVkKL2mG1KgUH9+hrtpMTkwO3RREnKe7I5O
x9qDimJpOihrL4I0dYvy9xUOoz+DzFAW8+ylWVYpA4GFAAKBgQDbnBAKSxWr9QHY
6Dt+EFdGz6lAZLedeBKpaP53Z1DTO34J0C55YbJTwBTFGqPtOLxnUVDlGiD6GbmC
80f3jvogPR1mSmGsydbNbZnbUEVWrRhe+y5zJ3g9qs/DWmDW0deEFvkhWVnLJkFJ
9pdOu/ibRPH1lE2nz6pK7GbOQtLyHTAJBgcqhkjOOAQDAzAAMC0CFQCoJlwGtJQC
cLoM4p/jtVFOj26xbgIUUS4pDKyHaG/eaygLTtFpFJqzWHc=
-----END CERTIFICATE-----
```


Bahrain Region

Le certificat public AWS de la région Bahreïn est le suivant :


```
-----BEGIN CERTIFICATE-----
MIIC7jCCAq4CCQCVWIgSmP8RhTAJBgcqhkjOOAQDMFwxCzAJBgNVBAYTAlVTMRkw
FwYDVQQIExBXYXNoaW5ndG9uIFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYD
VQQKExdBbWF6b24gV2ViIFNlcnZpY2VzIExMQzAeFw0xOTAyMDUxMzA2MjFaFw00
NTAyMDUxMzA2MjFaMFwxCzAJBgNVBAYTAlVTMRkwFwYDVQQIExBXYXNoaW5ndG9u
IFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYDVQQKExdBbWF6b24gV2ViIFNl
cnZpY2VzIExMQzCCAbgwggEsBgcqhkjOOAQBMIIBHwKBgQDcwojQfgWdV1QliO0B
8n6cLZ38VE7ZmrjZ9OQV//Gst6S1h7euhC23YppKXi1zovefSDwFU54zi3/oJ++q
PHlP1WGL8IZ34BUgRTtG4TVolvp0smjkMvyRu5hIdKtzjV93Ccx15gVgyk+o1IEG
fZ2Kbw/Dd8JfoPS7KaSCmJKxXQIVAIZbIaDFRGa2qcMkW2HWASyNDl7bAoGBANtz
IdhfMq+l2I5iofY2oj3HI21Kj3LtZrWEg3W+/4rVhL3lTm0Nne1rl9yGujrjQwy5
Zp9V4A/w9w2O10Lx4K6hj34Eefy/aQnZwNdNhv/FQP7Az0fju+Yl6L13OOHQrL0z
Q+9cF7zEosekEnBQx3v6psNknKgD3Shgx+GO/LpCA4GFAAKBgQCVS7m77nuNAlZ8
wvUqcooxXMPkxJFl54NxAsAul9KP9KN4svm0O3Zrb7t2FOtXRM8zU3TqMpryq1o5
mpMPsZDg6RXo9BF7Hn0DoZ6PJTamkFA6md+NyTJWJKvXC7iJ8fGDBJqTciUHuCKr
```

```
12AztQ8bFWsrTgTzPE3p6U5ckcgV1TAJBgcqhkjOOAQDAy8AMCwCFB2NZGWm5EDl
86ayV3c1PEDukgQIAhQow38rQkN/VwHVeSW9DqEshXHjuQ==
-----END CERTIFICATE-----
```


GovCloud Regions

Le certificat public AWS pour les régions AWS GovCloud se présente comme suit.


```
-----BEGIN CERTIFICATE-----
MIIC7TCCAq0CCQCWukjZ5V4aZzAJBgcqhkjOOAQDMFwxCzAJBgNVBAYTAlVTMRkw
FwYDVQQIExBXYXNoaW5ndG9uIFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYD
VQQKExdBbWF6b24gV2ViIFNlcnZpY2VzIExMQzAeFw0xMjAxMDUxMjU2MTJaFw0z
ODAxMDUxMjU2MTJaMFwxCzAJBgNVBAYTAlVTMRkwFwYDVQQIExBXYXNoaW5ndG9u
IFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYDVQQKExdBbWF6b24gV2ViIFNl
cnZpY2VzIExMQzCCAbcwggEsBgcqhkjOOAQBMIIBHwKBgQCjkvcS2bb1VQ4yt/5e
ih5OO6kK/n1Lzllr7D8ZwtQP8fOEpp5E2ng+D6Ud1Z1gYipr58Kj3nssSNpI6bX3
VyIQzK7wLclnd/YozqNNmgIyZecN7EglK9ITHJLP+x8FtUpt3QbyYXJdmVMegN6P
hviYt5JH/nYl4hh3Pa1HJdskgQIVALVJ3ER11+Ko4tP6nwvHwh6+ERYRAoGBAI1j
k+tkqMVHuAFcvAGKocTgsjJem6/5qomzJuKDmbJNu9Qxw3rAotXau8Qe+MBcJl/U
hhy1KHVpCGl9fueQ2s6IL0CaO/buycU1CiYQk40KNHCcHfNiZbdlx1E9rpUp7bnF
lRa2v1ntMX3caRVDdbtPEWmdxSCYsYFDk4mZrOLBA4GEAAKBgEbmeve5f8LIE/Gf
MNmP9CM5eovQOGx5ho8WqD+aTebs+k2tn92BBPqeZqpWRa5P/+jrdKml1qx4llHW
MXrs3IgIb6+hUIB+S8dz8/mmO0bpr76RoZVCXYab2CZedFut7qc3WUH9+EUAH5mw
vSeDCOUMYQR7R9LINYwouHIziqQYMAkGByqGSM44BAMDLwAwLAIUWXBlk40xTwSw
7HX32MxXYruse9ACFBNGmdX2ZBrVNGrN9N2f6ROk0k9K
-----END CERTIFICATE-----
```


China Regions

Le certificat public AWS pour les régions Chine (Pékin) et Chine (Ningxia) se présente comme suit.


```
-----BEGIN CERTIFICATE-----
MIIDNjCCAh4CCQD3yZ1w1AVkTzANBgkqhkiG9w0BAQsFADBcMQswCQYDVQQGEwJV
UzEZMBcGA1UECBMQV2FzaGluZ3RvbiBTdGF0ZTEQMA4GA1UEBxMHU2VhdHRsZTEg
MB4GA1UEChMXQW1hem9uIFdlYiBTZXJ2aWNlcyBMTEMwIBcNMTUwNTEzMDk1OTE1
WhgPMjE5NDEwMTYwOTU5MTVaMFwxCzAJBgNVBAYTAlVTMRkwFwYDVQQIExBXYXNo
aW5ndG9uIFN0YXRlMRAwDgYDVQQHEwdTZWF0dGxlMSAwHgYDVQQKExdBbWF6b24g
V2ViIFNlcnZpY2VzIExMQzCCASIwDQYJKoZIhvcNAQEBBQADggEPADCCAQoCggEB
AMWk9vyppSmDU3AxZ2Cy2bvKeK3F1UqNpMuyeriizi+NTsZ8tQqtNloaQcqhto/l
gsw9+QSnEJeYWnmivJWOBdn9CyDpN7cpHVmeGgNJL2fvImWyWe2f2Kq/BL9l7N7C
P2ZT52/sH9orlck1n2zO8xPi7MItgPHQwu3OxsGQsAdWucdxjHGtdchulpo1uJ31
jsTAPKZ3p1/sxPXBBAgBMatPHhRBqhwHO/Twm4J3GmTLWN7oVDds4W3bPKQfnw3r
vtBj/SM4/IgQ3xJslFcl90TZbQbgxIi88R/gWTbs7GsyT2PzstU30yLdJhKfdZKz
/aIzraHvoDTWFaOdy0+OOaECAwEAATANBgkqhkiG9w0BAQsFAAOCAQEAdSzN2+0E
V1BfR3DPWJHWRf1b7zl+1X/ZseW2hYE5r6YxrLv+1VPf/L5I6kB7GEtqhZUqteY7
zAceoLrVu/7OynRyfQetJVGichaaxLNM3lcr6kcxOowb+WQQ84cwrB3keykH4gRX
KHB2rlWSxta+2panSEO1JX2q5jhcFP90rDOtZjlpYv57N/Z9iQ+dvQPJnChdq3BK
5pZlnIDnVVxqRike7BFy8tKyPj7HzoPEF5mh9Kfnn1YoSVu+61lMVv/qRjnyKfS9
c96nE98sYFj0ZVBzXw8Sq4Gh8FiVmFHbQp1peGC19idOUqxPxWsasWxQXO0azYsP
9RyWLHKxH1dMuA==
-----END CERTIFICATE-----
```


c. Enregistrez et fermez le fichier .

5. Utilisez la commande OpenSSL smime pour vérifier la signature. Incluez l'option `-verify` indiquant que la signature doit être vérifiée et l'option `-noverify` indiquant que le certificat n'a pas besoin d'être vérifié.

```
$ openssl smime -verify -in pkcs7 -inform PEM -content document -certfile certificate -
noverify
```

Si la signature est valide, le message `Verification successful` s'affiche. Si la signature ne peut pas être vérifiée, contactez AWS Support.

Utilisation de la signature codée en base64 pour vérifier le Documents d'identité d'instance

Cette rubrique explique comment vérifier le Documents d'identité d'instance avec la signature codée en base64 et le certificat public AWS RSA.

> **Important**
>
> Pour valider le Documents d'identité d'instance avec la signature codée en base64, vous devez demander le certificat public AWS RSA auprès de AWS Support.

Pour valider le Documents d'identité d'instance avec la signature codée en base64 et le certificat public AWS RSA

1. Connectez-vous à l'instance.
2. Récupérez la signature codée en base64 à partir des métadonnées d'instance, convertissez-la en binaire et ajoutez-la à un fichier nommé `signature`. Utilisez l'une des commandes suivantes en fonction de la version IMDS utilisée par l'instance.

   IMDSv2

   ```
   $ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
   metadata-token-ttl-seconds: 21600"` \ && curl -H "X-aws-ec2-metadata-token: $TOKEN"
    -v http://169.254.169.254/latest/dynamic/instance-identity/signature | base64 -d
    > signature
   ```

   IMDSv1

   ```
   $ curl -s http://169.254.169.254/latest/dynamic/instance-identity/signature |
    base64 -d > signature
   ```

3. Récupérez le Documents d'identité d'instance en texte brut à partir des métadonnées de l'instance et ajoutez-le à un fichier nommé `document`. Utilisez l'une des commandes suivantes en fonction de la version IMDS utilisée par l'instance.

   IMDSv2

   ```
   $ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
   metadata-token-ttl-seconds: 21600"` \ && curl -H "X-aws-ec2-metadata-token: $TOKEN"
    -v http://169.254.169.254/latest/dynamic/instance-identity/document > document
   ```

   IMDSv1

   ```
   $ curl -s http://169.254.169.254/latest/dynamic/instance-identity/document
    > document
   ```

4. Ajoutez le certificat public AWS RSA que vous avez reçu d'AWS Support à un fichier nommé `certificate`.
5. Extrayez la clé publique du certificat que vous avez reçu depuis AWS Support et enregistrez-la dans un fichier nommé `key`.

   ```
   $ openssl x509 -pubkey -noout -in certificate > key
   ```

6. Utilisez la commande OpenSSL dgst pour vérifier le Documents d'identité d'instance.

```
$ openssl dgst -sha256 -verify key -signature signature document
```

Si la signature est valide, le message `Verified OK` s'affiche. Si la signature ne peut pas être vérifiée, contactez AWS Support.

Utilisation de la signature RSA-2048 pour vérifier le Documents d'identité d'instance

Cette rubrique explique comment vérifier le Documents d'identité d'instance avec la signature RSA-2048 et le certificat public AWS RSA-2048.

> **Important**
>
> Pour valider le Documents d'identité d'instance avec la signature RSA-2048, vous devez demander le certificat public AWS RSA-2048 auprès de AWS Support.

Pour vérifier l'Documents d'identité d'instance avec la signature RSA-2048 et le certificat public AWS RSA-2048

1. Connectez-vous à l'instance.
2. Récupérez la signature RSA-2048 à partir des métadonnées de l'instance et ajoutez-la à un fichier nommé `rsa2048`.

   a. Ajoutez l'en-tête `-----BEGIN PKCS7-----` au fichier `rsa2048`.

   ```
   $ echo "-----BEGIN PKCS7-----" > rsa2048
   ```

   b. Récupérez la signature RSA-2048 à partir des métadonnées de l'instance et ajoutez-la au fichier `rsa2048`. Utilisez l'une des commandes suivantes en fonction de la version IMDS utilisée par l'instance.

   IMDSv2

   ```
   $ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
   metadata-token-ttl-seconds: 21600"` \
   && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
   dynamic/instance-identity/rsa2048 >> rsa2048
   ```

   IMDSv1

   ```
   $ curl -s http://169.254.169.254/latest/dynamic/instance-identity/rsa2048
    >> rsa2048
   ```

   c. Ajoutez le pied de page `-----END PKCS7-----` à une nouvelle ligne du fichier `rsa2048`.

   ```
   $ echo "" >> rsa2048
   $ echo "-----END PKCS7-----" >> rsa2048
   ```

3. Ajoutez le contenu de l'Documents d'identité d'instance des métadonnées de l'instance à un fichier nommé `document`. Utilisez l'une des commandes suivantes en fonction de la version IMDS utilisée par l'instance.

IMDSv2

```
$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
dynamic/instance-identity/document > document
```

IMDSv1

```
$ curl -s http://169.254.169.254/latest/dynamic/instance-identity/document
 > document
```

4. Ajoutez le certificat public AWS RSA-2048 à un fichier nommé `certificate`.

   a. Créez le fichier `certificate`.

   ```
   $ touch certificate
   ```

   b. Ouvrez le fichier `certificate` à l'aide de votre éditeur de texte préféré et ajoutez le contenu du certificat public AWS RSA-2048 que vous avez reçu depuis AWS Support.

   c. Enregistrez et fermez le fichier .

5. Utilisez la commande OpenSSL smime pour vérifier la signature. Incluez l'option `-verify` indiquant que la signature doit être vérifiée et l'option `-noverify` indiquant que le certificat n'a pas besoin d'être vérifié.

```
$ openssl smime -verify -in rsa2048 -inform PEM -content document -certfile certificate
 -noverify
```

   Si la signature est valide, le message `Verification successful` s'affiche. Si la signature ne peut pas être vérifiée, contactez AWS Support.

Amazon Elastic Inference

Amazon Elastic Inference (EI) est une ressource que vous pouvez attacher à vos instances d'UC Amazon EC2 afin d'accélérer vos charges de travail d'inférence de deep learning (DL). Les accélérateurs Amazon EI sont disponibles en plusieurs tailles ; ils constituent une méthode économique pour intégrer des fonctionnalités intelligentes aux applications s'exécutant sur les instances Amazon EC2.

Amazon EI distribue les opérations de modèle définies par TensorFlow, Apache MXNet et le format ONNX (Open Neural Network Exchange) via MXNet entre les accélérateurs d'inférence DL économiques et l'UC de l'instance.

Pour de plus amples informations sur Amazon Elastic Inference, veuillez consulter le Manuel du développeur Amazon EI.

Identification des instances EC2 Linux

L'application peut avoir à déterminer si elle s'exécute sur une instance EC2.

Pour plus d'informations sur l'identification des instances Windows, consultez Identifier les instances EC2 Windows dans le manuel Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Inspection du Documents d'identité d'instance

Pour identifier une instance EC2 à l'aide d'une méthode définitive incluant un chiffrement, vérifiez le document d'identité d'instance, y compris sa signature. Ces documents sont disponibles sur chaque instance EC2 à l'adresse locale non routable `http://169.254.169.254/latest/dynamic/instance-identity/`. Pour plus d'informations, consultez Documents d'identité d'instance (p. 728).

Inspection de l'UUID du système

Vous pouvez obtenir l'UUID système et rechercher la présence des caractères « ec2 » ou « EC2 » dans l'octet de début de l'UUID. Cette méthode pour déterminer si un système est une instance EC2 est rapide, mais potentiellement inexacte, car il est peu probable qu'un système qui n'est pas une instance EC2 puisse avoir un UUID qui commence par ces caractères. De plus, pour les instances EC2 qui n'utilisent pas Amazon Linux, l'implémentation de la distribution de SMBIOS peut représenter l'UUID au format Little Endian. Par conséquent, les caractères « EC2 » n'apparaissent pas au début de l'UUID.

Example Obtenir l'UUID auprès de l'hyperviseur

Si `/sys/hypervisor/uuid` existe, vous pouvez utiliser la commande suivante :

```
[ec2-user ~]$ cat /sys/hypervisor/uuid
```

Dans l'exemple de sortie suivant, l'UUID commence par « ec2 », ce qui indique que le système est probablement une instance EC2.

```
ec2e1916-9099-7caf-fd21-012345abcdef
```

Example : Obtenir l'UUID à partir de DMI (instances HVM uniquement)

Sur les instances HVM uniquement, vous pouvez utiliser l'interface Desktop Management Interface (DMI).

Vous pouvez utiliser l'outil `dmidecode` pour retourner l'UUID. Sur Amazon Linux, utilisez la commande suivante pour installer l'outil `dmidecode` s'il n'est pas déjà installé sur votre instance :

```
[ec2-user ~]$ sudo yum install dmidecode -y
```

Ensuite, exécutez la commande suivante :

```
[ec2-user ~]$ sudo dmidecode --string system-uuid
```

Vous pouvez également utiliser la commande suivante :

```
[ec2-user ~]$ sudo cat /sys/devices/virtual/dmi/id/product_uuid
```

Dans l'exemple de sortie suivant, l'UUID commence par « EC2 », ce qui indique que le système est probablement une instance EC2.

```
EC2E1916-9099-7CAF-FD21-01234ABCDEF
```

Dans l'exemple de sortie qui suit, l'UUID est représenté au format Little Endian :

```
45E12AEC-DCD1-B213-94ED-01234ABCDEF
```

Sur les instances Nitro, la commande suivante peut être utilisée :

```
[ec2-user ~]$ cat /sys/devices/virtual/dmi/id/board_asset_tag
```

Ceci retourne l'ID de l'instance, qui est unique aux instances EC2 :

```
i-0af01c0123456789a
```

Surveillance d'Amazon EC2

La surveillance constitue une part importante de la gestion de la fiabilité, de la disponibilité et des performances de vos instances Amazon Elastic Compute Cloud (Amazon EC2) et de vos solutions AWS. Vous devez recueillir les données de surveillance de toutes les parties de vos solutions AWS de telle sorte que vous puissiez déboguer plus facilement une éventuelle défaillance à plusieurs points. Cependant, avant de commencer à superviser Amazon EC2, créez un plan de surveillance qui inclut les questions suivantes :

- Quels sont les objectifs de la surveillance ?
- Quelles sont les ressources à surveiller ?
- À quelle fréquence les ressources doivent-elles être surveillées ?
- Quels outils de supervision utiliser ?
- Qui exécute les tâches de supervision ?
- Qui doit être informé en cas de problème ?

Une fois que vous avez défini vos objectifs de surveillance et créé votre plan de surveillance, l'étape suivante consiste à définir une référence pour les performances normales d'Amazon EC2 dans votre environnement. Vous devez mesurer les performances d'Amazon EC2 à différents moments et sous différentes conditions de charge. Lorsque vous supervisez Amazon EC2, stockez l'historique des données de surveillance que vous avez collectées. Vous pouvez comparer les performances actuelles d'Amazon EC2 à leurs données historiques pour vous aider à identifier les modèles de performances normales et les anomalies de performances, et à concevoir les méthodes destinées à les prendre en compte. Par exemple, vous pouvez superviser l'utilisation de l'UC, les E/S de disque et l'utilisation réseau de vos instances EC2. Lorsque les performances se trouvent en dehors de votre référence établie, il se peut que vous ayez besoin de reconfigurer l'instance ou de l'optimiser pour réduire l'utilisation de l'UC, améliorer les E/S disque ou réduire le trafic réseau.

Pour établir une référence, vous devez, au moins, superviser les éléments suivants :

Élément à superviser	Métrique Amazon EC2	Agent de surveillance/ CloudWatch Logs
Utilisation de l'UC	CPUUtilization (p. 764)	
Utilisation réseau	NetworkIn (p. 764) NetworkOut (p. 764)	
Performances disque	DiskReadOps (p. 764) DiskWriteOps (p. 764)	
Lectures/écritures sur disque	DiskReadBytes (p. 764) DiskWriteBytes (p. 764)	
Utilisation de la mémoire, des échanges, de l'espace sur le disque et du fichier d'échange, collecte de journaux		[Instances Linux et Windows Server] Collecte de métriques et de journaux à partir d'instances

Élément à superviser	Métrique Amazon EC2	Agent de surveillance/ CloudWatch Logs
		Amazon EC2 et de serveurs sur site avec l'agent CloudWatch [Migration à partir de l'ancien agent CloudWatch Logs sur des instances Windows Server] Migrer la collecte de journaux d'instances Windows Server vers l'agent CloudWatch

Surveillance automatique et surveillance manuelle

AWS fournit différents outils que vous pouvez utiliser pour superviser Amazon EC2. Vous pouvez configurer certains outils pour qu'ils effectuent la surveillance automatiquement, tandis que d'autres nécessitent une intervention manuelle.

Outils de surveillance

- Outils de surveillance automatique (p. 740)
- Outils de surveillance manuelle (p. 741)

Outils de surveillance automatique

Vous pouvez utiliser les outils de surveillance automatique pour surveiller Amazon EC2 et être informé en cas de problème :

- Contrôles du statut du système – Surveillez les systèmes AWS requis pour utiliser votre instance et assurez-vous qu'ils fonctionnent correctement. Ces contrôles détectent les problèmes liés à votre instance nécessitant une intervention d'AWS pour les résoudre. Lorsqu'un contrôle de statut échoue, vous pouvez choisir d'attendre qu'AWS résolve le problème ou le résoudre vous-même (par exemple, en arrêtant et en redémarrant une instance, ou en la terminant et en la remplaçant). Voici quelques exemples de problèmes entraînant l'échec des contrôles de statut du système :
  - Perte de connectivité réseau
  - Perte d'alimentation système
  - Problèmes logiciels sur un hôte physique
  - Problèmes matériels sur un hôte physique ayant un impact sur l'accessibilité du réseau

  Pour plus d'informations, consultez Contrôles de statut pour vos instances (p. 742).
- Contrôles du statut de l'instance – Surveillez la configuration logicielle et réseau de votre instance. Ces contrôles détectent les problèmes nécessitant votre intervention pour les résoudre. Lorsqu'un contrôle du statut de l'instance échoue, vous devez généralement résoudre le problème vous-même (en redémarrant par exemple l'instance ou en apportant des modifications à votre système d'exploitation). Voici quelques exemples de problèmes susceptibles d'entraîner l'échec des contrôles du statut de l'instance :
  - Échec de contrôles de statut de système
  - Configuration de mise en réseau ou de démarrage incorrecte
  - Mémoire épuisée
  - Système de fichiers corrompu
  - Noyau incompatible

  Pour plus d'informations, consultez Contrôles de statut pour vos instances (p. 742).

- Alarmes Amazon CloudWatch – Surveillez une seule métrique sur une durée définie et exécutez une ou plusieurs actions en fonction de la valeur de la métrique par rapport à un seuil donné sur un certain nombre de durées. L'action est une notification envoyée à une rubrique Amazon Simple Notification Service (Amazon SNS) ou une politique Amazon EC2 Auto Scaling. Les alarmes invoquent des actions uniquement pour des changements d'états maintenus. Les alarmes CloudWatch n'invoqueront pas d'actions simplement parce qu'elles sont dans un état particulier : l'état doit avoir changé et avoir été maintenu pendant un certain nombre de périodes. Pour de plus amples informations, veuillez consulter Surveillance de vos instances avec CloudWatch (p. 761).
- Amazon CloudWatch Events – Automatisez vos services AWS et répondez automatiquement à des événements système. Les événements provenant de services AWS sont fournis à CloudWatch Events presque en temps réel et vous pouvez spécifier des actions automatisées quand un événement correspond à une règle que vous écrivez. Pour plus d'informations, consultez Qu'est-ce que Amazon CloudWatch Events?
- Amazon CloudWatch Logs – Supervisez vos fichiers journaux, stockez-les et accédez-y à partir des instances Amazon EC2, de AWS CloudTrail ou d'autres sources. Pour plus d'informations, consultez la Amazon CloudWatch Logs User Guide.
- Agent CloudWatch – Collectez les journaux et les métriques au niveau du système à partir des hôtes et des invités sur vos instances EC2 et sur les serveurs locaux. Pour de plus amples informations, veuillez consulter Collecte de métriques et de journaux à partir d'instances Amazon EC2 et de serveurs sur site avec l'agent CloudWatch dans le Guide de l'utilisateur Amazon CloudWatch.
- AWS Management Pack pour System Center Operations Manager – lie les instances Amazon EC2 et le système d'exploitation (Microsoft Windows ou Linux) qui s'y exécute. AWS Management Pack est une extension de Microsoft System Center Operations Manager. Il utilise un ordinateur défini de votre centre de données (appelé nœud observateur) et les API Amazon Web Services pour découvrir et collecter à distance les informations sur vos ressources AWS. Pour plus d'informations, veuillez consulter AWS Management Pack for Microsoft System Center.

Outils de surveillance manuelle

Une autre part importante de la surveillance d'Amazon EC2 implique la surveillance manuelle de ces éléments que les scripts, les contrôles de statut et les alarmes CloudWatch ne couvrent pas. Les tableaux de bord des consoles Amazon EC2 et CloudWatch fournissent un aperçu de l'état de votre environnement Amazon EC2.

- Le tableau de bord Amazon EC2 affiche :
  - État du service et événements planifiés par région
  - État de l'instance
  - Contrôles des statuts
  - Statut des alarmes
  - Détails des métriques d'instance (Dans le panneau de navigation, choisissez Instances, sélectionnez une instance, et choisissez l'onglet Surveillance)
  - Détails des métriques de volume (Dans le panneau de navigation, choisissez Volumes, sélectionnez un volume, et choisissez l'onglet Surveillance)
- Le tableau de bord Amazon CloudWatch affiche :
  - Alarmes et statuts en cours
  - Graphiques des alarmes et des ressources
  - Statut d'intégrité du service

  De plus, vous pouvez utiliser CloudWatch pour effectuer les tâches suivantes :
  - Représenter graphiquement les données de surveillance Amazon EC2 pour résoudre les problèmes et découvrir les tendances
  - Rechercher et parcourir toutes les métriques des ressources AWS

- Créer et modifier des alarmes pour être informé des problèmes
- Afficher une présentation de vos alarmes et ressources AWS

Bonnes pratiques de surveillance

Utilisez les bonnes pratiques suivantes pour vous aider dans les tâches de surveillance d'Amazon EC2.

- Faites de la surveillance une priorité pour résoudre les petits problèmes avant qu'ils n'empirent.
- Créez et implémentez un plan de surveillance qui collecte les données de surveillance de toutes les parties de vos solutions AWS de telle sorte que vous puissiez déboguer plus facilement une éventuelle défaillance à plusieurs points. Votre plan de surveillance doit, au moins, traiter les questions suivantes :
  - Quels sont les objectifs de la surveillance ?
  - Quelles sont les ressources à superviser ?
  - A quelle fréquence les ressources doivent-elles être supervisées ?
  - Quels outils de surveillance utiliser ?
  - Qui exécute les tâches de supervision ?
  - Qui doit être informé en cas de problème ?
- Automatisez les tâches de surveillance autant que possible.
- Vérifiez les fichiers journaux de vos instances EC2.

Surveillance du statut de vos instances

Vous pouvez surveiller le statut de vos instances en affichant les contrôles de statut et les événements planifiés pour vos instances.

Un contrôle de statut vous fournit les informations provenant de contrôles automatisés exécutés par Amazon EC2. Ces contrôles automatisés détectent si des problèmes spécifiques concernent vos instances. Les informations de contrôle de statut, avec les données fournies par Amazon CloudWatch, vous donnent une visibilité opérationnelle détaillées sur chacune de vos instances.

Vous pouvez également consulter le statut d' événements spécifiques planifiés pour vos instances. Les statuts des événements fournissent des informations sur les activités à venir planifiées pour vos instances, comme le redémarrage ou la mise hors service. Ils fournissent aussi les heures prévues de début et de fin de chaque événement.

Sommaire

- Contrôles de statut pour vos instances (p. 742)
- Événements planifiés pour vos instances. (p. 749)

Contrôles de statut pour vos instances

Avec la surveillance du statut des instances, vous pouvez rapidement déterminer si Amazon EC2 a détecté des problèmes susceptibles d'empêcher vos instances d'exécuter des applications. Amazon EC2 effectue des contrôles automatisés sur chaque instance EC2 en cours d'exécution pour identifier les problèmes matériels et logiciels. Vous pouvez afficher les résultats de ces contrôles de statut pour identifier des problèmes spécifiques et détectables. Ces données viennent s'ajouter aux informations déjà fournies par Amazon EC2 à propos de l'état de chaque instance (par exemple, `pending`, `running`, `stopping`) ainsi qu'aux métriques d'utilisation surveillées par Amazon CloudWatch (utilisation de l'UC, trafic réseau et activité de disque).

Les contrôles de statut sont exécutés toutes les minutes et chacun d'entre eux renvoie un statut de réussite ou d'échec. Si tous les contrôles réussissent, le statut global de l'instance est OK. Si un ou plusieurs contrôles échouent, le statut global de l'instance est dégradé. Les contrôles de statut sont intégrés à Amazon EC2. Ils ne peuvent donc pas être désactivés ou supprimés.

Lorsqu'un contrôle de statut échoue, la métrique CloudWatch correspondante pour les contrôles de statut est incrémentée. Pour plus d'informations, consultez Métriques de contrôle de statut (p. 770). Vous pouvez utiliser ces métriques pour créer des alarmes CloudWatch qui sont déclenchées en fonction du résultat des contrôles de statut. Par exemple, vous pouvez créer une alarme pour vous avertir si des contrôles de statut échouent sur une instance spécifique. Pour plus d'informations, consultez Création et modification d'alarmes de contrôle de statut (p. 747).

Vous pouvez aussi créer une alarme Amazon CloudWatch qui surveille une instance Amazon EC2 et récupère automatiquement l'instance si cette dernière est dégradée suite à un problème sous-jacent. Pour plus d'informations, consultez Récupération de votre instance (p. 652).

Sommaire

- Types de contrôles de statut (p. 743)
- Affichage des contrôles de statut (p. 744)
- Signalement d'un statut d'instance (p. 746)
- Création et modification d'alarmes de contrôle de statut (p. 747)

Types de contrôles de statut

Il existe deux types de contrôles de statut : les contrôles de statut de système et les contrôles de statut d'instance.

Contrôles de statut de système

Les contrôles de statut du système surveillent les systèmes AWS sur lesquels votre instance s'exécute. Ces contrôles détectent les problèmes sous-jacents liés à votre instance nécessitant une intervention d'AWS pour les résoudre. Lorsqu'un contrôle de statut échoue, vous pouvez choisir d'attendre qu'AWS résolve le problème ou le résoudre vous-même. Pour les instances basées sur Amazon EBS, vous pouvez arrêter et démarrer l'instance vous-même, ce qui, dans la plupart des cas, entraîne la migration de l'instance vers un nouvel hôte. Pour les instances basées sur le stockage d'instance, vous pouvez mettre l'instance hors service et la remplacer.

Voici des exemples de problèmes pouvant entraîner l'échec des contrôles de statut :

- Perte de connectivité réseau
- Perte d'alimentation système
- Problèmes logiciels sur un hôte physique
- Problèmes matériels sur un hôte physique ayant un impact sur l'accessibilité du réseau

Contrôles de statut des instances

Les contrôles de statut des instances surveillent la configuration logicielle et réseau de votre instance individuelle. Amazon EC2 vérifie l'état de l'instance en envoyant une demande de protocole de résolution d'adresse (ARP) à l'interface réseau. Ces contrôles détectent les problèmes nécessitant votre intervention pour les résoudre. Lorsqu'un contrôle de statut d'instance échoue, vous devez généralement résoudre le problème vous-même (par exemple, en redémarrant l'instance ou en effectuant des changements de configuration sur l'instance).

Voici des exemples de problèmes pouvant entraîner l'échec des contrôles d'instance :

- Échec de contrôles de statut de système
- Configuration de mise en réseau ou de démarrage incorrecte
- Mémoire épuisée
- Système de fichiers corrompu
- Noyau incompatible

Affichage des contrôles de statut

Amazon EC2 vous permet de consulter et gérer les contrôles de statut de plusieurs façons.

Affichage du statut à l'aide de la console

Vous pouvez consulter les contrôles de statut à l'aide de la AWS Management Console.

Nouvelle console

Pour afficher les contrôles de statut (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sur la page Instances, la colonne Status check (Vérification de statut) répertorie le statut opérationnel de chaque instance.
4. Pour afficher le statut d'une instance spécifique, sélectionnez l'instance, puis l'onglet Contrôles des statuts.

   Si vous avez une instance dont la vérification de statut est en échec et que l'instance est inaccessible depuis plus de 20 minutes, choisissez Open support case (Ouvrir le cas de support) pour envoyer une demande d'assistance. Pour résoudre vous-même des échecs de contrôle de statut de système ou d'instance, consultez Résolution des instances avec des contrôles de statut échoués (p. 1338).
5. Pour vérifier les métriques CloudWatch associées aux contrôles de statut, sélectionnez l'instance, puis choisissez l'onglet Monitoring (surveillance). Faites défiler jusqu'à ce que les graphiques correspondant aux métriques suivantes s'affichent :

- Status Check Failed (Any) (Échec du contrôle de statut (Quelconque))
- Status check failed (instance) (Échec du contrôle de statut (instance))
- Status check failed (system) (Échec du contrôle de statut (système))

Ancienne console

Pour afficher les contrôles de statut (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sur la page Instances, la colonne Contrôles des statuts répertorie le statut opérationnel de chaque instance.
4. Pour afficher le statut d'une instance spécifique, sélectionnez l'instance, puis l'onglet Contrôles des statuts.

   Si vous avez une instance avec un échec de contrôle de statut et que l'instance est inaccessible depuis plus de 20 minutes, sélectionnez AWS Support pour envoyer une demande d'assistance. Pour résoudre vous-même des échecs de contrôle de statut de système ou d'instance, consultez Résolution des instances avec des contrôles de statut échoués (p. 1338).
5. Pour vérifier les métriques CloudWatch associées aux contrôles de statut, sélectionnez l'instance, puis choisissez l'onglet Monitoring (surveillance). Faites défiler jusqu'à ce que les graphiques correspondant aux métriques suivantes s'affichent :

   - Status Check Failed (Any) [Échec du contrôle de statut (tous)]
   - Status Check Failed (Instance) [Échec du contrôle de statut (instance)]
   - Status Check Failed (System) [Échec du contrôle de statut (système)]

Affichage du statut à l'aide de la ligne de commande

Vous pouvez afficher des contrôles du statut des instances en cours d'exécution à l'aide de la commande describe-instance-status (AWS CLI).

Pour afficher le statut de toutes les instances, utilisez la commande suivante :

```
aws ec2 describe-instance-status
```

Pour obtenir le statut de toutes les instances avec un statut d'instance `impaired`, utilisez la commande suivante.

```
aws ec2 describe-instance-status \
    --filters Name=instance-status.status,Values=impaired
```

Pour obtenir le statut d'une seule instance, utilisez la commande suivante.

```
aws ec2 describe-instance-status \
    --instance-ids i-1234567890abcdef0
```

Vous pouvez également utiliser les commandes suivantes :

- Get-EC2InstanceStatus (Outils AWS pour Windows PowerShell)
- DescribeInstanceStatus (API de requête Amazon EC2)

Si vous avez une instance avec un statut d'échec (failed), consultez Résolution des instances avec des contrôles de statut échoués (p. 1338).

Signalement d'un statut d'instance

Vous pouvez fournir des commentaires si vous êtes confronté à des problèmes avec une instance dont le statut n'apparaît pas comme impaired (dégradé) ou si vous souhaitez envoyer des détails supplémentaires à AWS concernant les problèmes que vous rencontrez avec une instance dégradée.

Nous utilisons ces commentaires pour identifier les problèmes concernant plusieurs clients, mais nous ne répondons pas à des problèmes de compte individuels. Le fait de fournir des commentaires ne modifie pas les résultats de contrôle de statut que vous voyez actuellement pour l'instance.

Envoi de commentaires de statut à l'aide de la console

Nouvelle console

Pour signaler le statut d'une instance (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, choisissez l'onglet Status Checks (Contrôles des statuts), choisissez Actions (deuxième menu Actions dans la moitié inférieure de la page), puis choisissez Report instance status (Signaler le statut de l'instance).
4. Complétez le formulaire Report instance status (Signaler le statut de l'instance), puis sélectionnez Submit (Envoyer).

Ancienne console

Pour signaler le statut d'une instance (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, choisissez l'onglet Contrôles des statuts, puis choisissez Envoyer les commentaires.
4. Complétez le formulaire de signalement du statut d'instance, puis sélectionnez Soumettre.

Création de rapports de commentaires de statut avec la ligne de commande

Utilisez la commande report-instance-status (AWS CLI) pour envoyer des commentaires à propos du statut d'une instance dégradée.

```
aws ec2 report-instance-status \
    --instances i-1234567890abcdef0 \
    --status impaired \
    --reason-codes code
```

Vous pouvez également utiliser les commandes suivantes :

- Send-EC2InstanceStatus (Outils AWS pour Windows PowerShell)
- ReportInstanceStatus (API de requête Amazon EC2)

Création et modification d'alarmes de contrôle de statut

Vous pouvez utiliser les métriques de vérification de statut (p. 770) pour créer des alarmes CloudWatch afin de vous alerter lorsqu'une instance connaît un échec de contrôle de statut.

Création d'une alarme de contrôle de statut à l'aide de la console

Utilisez la procédure suivante pour configurer une alarme qui vous envoie une notification par e-mail, ou arrête, met fin ou récupère une instance en cas d'échec du contrôle de statut de cette dernière.

Nouvelle console

Pour créer une alarme de contrôle de statut (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, choisissez l'onglet Status Checks (Contrôles des statuts), puis choisissez Actions, Create status check alarm (Créer une alarme de contrôle de statut).
4. Sur la page Manage CloudWatch alarms (Gérer les alarmes CloudWatch), sous Add or edit alarm (Ajouter ou modifier une alarme), sélectionnez Create a new alarm (Créer une nouvelle alarme).
5. Pour Alarm notification (Notification d'alarme), activez ou désactivez les notifications Amazon Simple Notification Service (Amazon SNS). Sélectionnez une rubrique Amazon SNS existante ou entrez un nom pour créer une nouvelle rubrique.
6. Activez Alarm action (Action d'alarme) pour spécifier une action à effectuer lorsque l'alarme est déclenchée. Sélectionnez l'action que vous souhaitez effectuer dans la liste déroulante.
7. Pour Seuils d'alarme, sélectionnez la métrique et les critères de l'alarme. Dans Consecutive Period (Périodes consécutives), définissez le nombre de périodes que vous souhaitez évaluer et, dans Period (Période), sélectionnez la période d'évaluation avant de déclencher l'alarme et d'envoyer un e-mail.

   Par exemple, vous pouvez laisser les paramètres par défaut pour Regrouper les échantillons par (Moyenne) et Type de données à échantillonner (utilisation de l'UC). Vous pouvez définir Alarm When (Alarme quand) sur **>=** et entrer **`0.80`** pour Percent (Pourcentage). Pour Consecutive Period (Périodes consécutives), vous pouvez saisir **`1`**. Pour Period (Période), vous pouvez sélectionner 5 Minutes.
8. (Facultatif) Pour Exemple de données de métrique, choisissez Ajouter au tableau de bord.
9. Sélectionnez Créer.

   > **Important**
   >
   > Si vous avez ajouté une adresse e-mail à la liste de destinataires, ou créé une nouvelle rubrique, Amazon SNS envoie un e-mail de confirmation d'abonnement à chaque nouvelle adresse. Chaque destinataire doit confirmer l'abonnement en choisissant le lien contenu dans ce message. Les notifications d'alerte sont envoyées uniquement aux adresses confirmées.

Ancienne console

Pour créer une alarme de contrôle de statut (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, choisissez l'onglet Contrôles des statuts, puis choisissez Créer une alarme de contrôle de statut.
4. Sélectionnez Envoyer une notification à. Choisissez une rubrique SNS existante, ou choisissez créer une rubrique. Si vous créez une rubrique dans Avec ces destinataires, entrez votre adresse e-mail et les adresses des éventuels destinataires supplémentaires, séparées par des virgules.
5. (Facultatif) Sélectionnez Prenez la mesure suivante, puis sélectionnez l'action que vous souhaitez exécuter.
6. Dans le champ Lorsque, sélectionnez le contrôle de statut à propos duquel vous souhaitez être alerté.

   Si vous avez sélectionné Récupérez cette instance à l'étape précédente, sélectionnez Échec du contrôle de statut (système).
7. Dans Pendant au moins, définissez le nombre de périodes que vous souhaitez évaluer et dans période(s) consécutive(s) de, sélectionnez la période d'évaluation avant de déclencher l'alarme et envoyer un e-mail.
8. (Facultatif) Dans Nom de l'alarme, remplacez le nom par défaut par un autre nom pour l'alarme.
9. Sélectionnez Create Alarm.

   > **Important**
   >
   > Si vous avez ajouté une adresse e-mail à la liste de destinataires, ou créé une nouvelle rubrique, Amazon SNS envoie un e-mail de confirmation d'abonnement à chaque nouvelle adresse. Chaque destinataire doit confirmer l'abonnement en choisissant le lien contenu dans ce message. Les notifications d'alerte sont envoyées uniquement aux adresses confirmées.

Si vous devez apporter des modifications à une alarme de statut d'instance, vous pouvez modifier celle-ci.

Nouvelle console

Pour modifier une alarme de contrôle de statut à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Surveillance, Gérer les alarmes CloudWatch.
4. Sur la page Manage CloudWatch alarms (Gérer les alarmes CloudWatch), sous Add or edit alarm (Ajouter ou modifier une alarme), choisissez Edit an existing alarm (Modifier une alarme existante).
5. Dans Search for alarm (Rechercher une alarme), choisissez l'alarme à modifier.
6. Apportez les modifications souhaitées, puis choisissez Update (Mettre à jour).

Ancienne console

Pour modifier une alarme de contrôle de statut à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances.

3. Sélectionnez l'instance et choisissez Actions, Supervision de CloudWatch, Ajouter/Modifier les alarmes.
4. Dans la boîte de dialogue Détails de l'alarme pour, sélectionnez le nom de l'alarme.
5. Dans la boîte de dialogue Modifier l'alarme, apportez les modifications souhaitées, puis sélectionnez Enregistrer.

Création d'une alarme de contrôle de statut à l'aide de la AWS CLI

Dans l'exemple suivant, l'alarme publie une notification dans une rubrique SNS, `arn:aws:sns:us-west-2:111122223333:my-sns-topic`, lorsque l'instance échoue lors d'un contrôle de statut d'instance ou un contrôle de statut de système pour au moins deux périodes consécutives. La métrique CloudWatch utilisée est `StatusCheckFailed`.

Pour créer une alarme de contrôle de statut à l'aide de l'AWS CLI

1. Sélectionnez une rubrique SNS existante ou créez-en une nouvelle. Pour plus d'informations, consultez Utilisation de l'AWS CLI avec Amazon SNS dans le AWS Command Line Interface Guide de l'utilisateur.
2. Utilisez la commande list-metrics suivante afin d'afficher les métriques Amazon CloudWatch disponibles pour Amazon EC2.

```
aws cloudwatch list-metrics --namespace AWS/EC2
```

3. Utilisez la commande put-metric-alarm suivante pour créer l'alarme.

```
aws cloudwatch put-metric-alarm --alarm-name StatusCheckFailed-Alarm-for-
i-1234567890abcdef0 --metric-name StatusCheckFailed --namespace AWS/EC2 --
statistic Maximum --dimensions Name=InstanceId,Value=i-1234567890abcdef0 --unit
 Count --period 300 --evaluation-periods 2 --threshold 1 --comparison-operator
 GreaterThanOrEqualToThreshold --alarm-actions arn:aws:sns:us-west-2:111122223333:my-
sns-topic
```

La période est le temps (en secondes) pendant lequel les métriques Amazon CloudWatch sont recueillies. Dans cet exemple, 300, qui correspond à 60 secondes multipliées par 5 minutes, est utilisé. La période d'évaluation est le nombre de périodes consécutives pour lesquelles la valeur de la métrique doit être comparée au seuil. Dans cet exemple, 2 est utilisé. Les actions d'alarme correspondent aux actions à exécuter lors du déclenchement de cette alarme. Dans cet exemple, l'alarme est configurée pour envoyer un e-mail à l'aide de Amazon SNS.

Événements planifiés pour vos instances.

AWS peut planifier des événements pour vos instances, comme un redémarrage, un arrêt/démarrage ou une mise hors service. Ces événements ne se produisent pas fréquemment. Si l'une de vos instances va être concernée par un événement planifié, AWS envoie un e-mail à l'adresse e-mail associée à votre compte AWS avant cet événement planifié. Cet e-mail fournit des détails concernant l'événement, y compris les dates de début et de fin. En fonction de l'événement, vous pouvez être en mesure d'intervenir pour contrôler le moment de l'événement.

Les événements planifiés sont gérés par AWS. Vous ne pouvez pas planifier d'événements pour vos instances. Vous pouvez afficher les événements planifiés par AWS, personnaliser les notifications d'événements planifiés pour inclure ou supprimer des balises de la notification par e-mail, effectuer des actions lorsqu'une instance est planifiée pour être redémarrée, retirée ou arrêtée.

Pour mettre à jour les informations de contact de votre compte afin d'être sûr d'être averti à propos d'événements planifiés, accédez à la page Account Settings (Paramètres du compte).

Sommaire

- Types d'événements planifiés (p. 750)
- Affichage des événements planifiés (p. 750)
- Personnalisation des notifications d'événements planifiés (p. 754)
- Gestion des instances planifiées pour être arrêtées ou mises hors service (p. 756)
- Gestion des instances planifiées pour un redémarrage (p. 757)
- Gestion des instances planifiées pour une maintenance (p. 759)
- Replanification d'un événement planifié (p. 759)

Types d'événements planifiés

Amazon EC2 peut créer les types d'événements suivants pour vos instances, où l'événement se produit à une heure planifiée :

- Instance stop (Arrêt de l'instance) : à l'heure planifiée, l'instance est arrêté. Lorsque vous la redémarrez, elle est migrée vers un nouvel hôte. S'applique uniquement aux instances basées sur Amazon EBS.
- Instance retirement (Mise hors service d'instance) : à l'heure planifiée, l'instance est arrêtée si elle est soutenue par Amazon EBS ou mise hors service si elle est soutenue par un stockage d'instance.
- Instance reboot (Redémarrage de l'instance) : à l'heure planifiée, l'instance est redémarrée.
- System reboot (Redémarrage du système) : à l'heure planifiée, l'hôte de l'instance est redémarré.
- System maintenance (Maintenance du système) : à l'heure planifiée, l'instance peut être temporairement affectée par une maintenance du réseau ou une maintenance de l'alimentation.

Affichage des événements planifiés

En plus de recevoir une notification des événements planifiés par e-mail, vous pouvez consulter les événements planifiés en utilisant une des méthodes suivantes.

Nouvelle console

Pour afficher les événements planifiés pour vos instances à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Vous pouvez afficher les événements planifiés dans les écrans suivants :
   - Dans le panneau de navigation, sélectionnez Events (Événements). Toutes les ressources avec un événement associé sont affichées. Vous pouvez filtrer par ID de ressource, Type de ressource, Zone de disponibilité, Statut de l'événement ou Type d'événement.

- Sinon, dans le panneau de navigation, vous pouvez sélectionner Tableau de bord EC2. Toutes les ressources avec un événement associé sont affichées sous Événements planifiés.

- Certains événements sont également affichés pour une ressource affectée. Par exemple, dans le panneau de navigation, choisissez Instances et sélectionnez une instance. Si un événement d'arrêt d'instance ou de mise hors service d'instance est associé à l'instance, il est affiché dans le volet inférieur.

Ancienne console

Pour afficher les événements planifiés pour vos instances à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Vous pouvez afficher les événements planifiés dans les écrans suivants :
   - Dans le panneau de navigation, sélectionnez Events (Événements). Toutes les ressources avec un événement associé sont affichées. Vous pouvez filtrer par type de ressource ou par type d'événement spécifique. Vous pouvez sélectionner la ressource pour afficher les détails.

   - Sinon, dans le panneau de navigation, vous pouvez sélectionner Tableau de bord EC2. Toutes les ressources avec un événement associé sont affichées sous Événements planifiés.

   - Certains événements sont également affichés pour une ressource affectée. Par exemple, dans le panneau de navigation, choisissez Instances et sélectionnez une instance. Si un

événement d'arrêt d'instance ou de mise hors service d'instance est associé à l'instance, il est affiché dans le volet inférieur.

AWS CLI

Pour afficher les événements planifiés pour vos instances à l'aide de la AWS CLI

Utilisez la commande describe-instance-status.

```
aws ec2 describe-instance-status \
    --instance-id i-1234567890abcdef0 \
    --query "InstanceStatuses[].Events"
```

L'exemple de sortie suivant montre un événement de redémarrage :

```
[
    "Events": [
        {
            "InstanceEventId": "instance-event-0d59937288b749b32",
            "Code": "system-reboot",
            "Description": "The instance is scheduled for a reboot",
            "NotAfter": "2019-03-15T22:00:00.000Z",
            "NotBefore": "2019-03-14T20:00:00.000Z",
            "NotBeforeDeadline": "2019-04-05T11:00:00.000Z"
         }

    ]
]
```

Voici un exemple de sortie montrant un événement de mise hors service d'instance.

```
[
    "Events": [
        {
            "InstanceEventId": "instance-event-0e439355b779n26",

            "Code": "instance-stop",
            "Description": "The instance is running on degraded hardware",
            "NotBefore": "2015-05-23T00:00:00.000Z"
        }
    ]
]
```

PowerShell

Pour afficher les événements planifiés pour vos instances à l'aide de la Outils AWS pour Windows PowerShell

Utilisez la commande Get-EC2InstanceStatus suivante.

```
PS C:\> (Get-EC2InstanceStatus -InstanceId i-1234567890abcdef0).Events
```

Voici un exemple de sortie montrant un événement de mise hors service d'instance.

```
Code         : instance-stop
```

```
Description  : The instance is running on degraded hardware
NotBefore    : 5/23/2015 12:00:00 AM
```

Instance metadata

Pour afficher les événements planifiés pour vos instances à l'aide des métadonnées de l'instance

Vous pouvez récupérer des informations sur les événements de maintenance actifs pour vos instances à partir des métadonnées de l'instance (p. 701) à l'aide de Service des métadonnées d'instance Version 2 ou Service des métadonnées d'instance Version 1.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/events/maintenance/scheduled
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/events/maintenance/scheduled
```

Voici un exemple de sortie avec des informations sur un événement de redémarrage système planifié, au format JSON.

```
[
  {
    "NotBefore" : "21 Jan 2019 09:00:43 GMT",
    "Code" : "system-reboot",
    "Description" : "scheduled reboot",
    "EventId" : "instance-event-0d59937288b749b32",
    "NotAfter" : "21 Jan 2019 09:17:23 GMT",
    "State" : "active"
  }
]
```

Pour afficher l'historique des événements terminés ou annulés pour vos instances à l'aide des métadonnées de l'instance

Vous pouvez récupérer des informations sur les événements terminés ou annulés pour vos instances à partir des métadonnées de l'instance (p. 701) à l'aide de Service des métadonnées d'instance Version 2 ou Service des métadonnées d'instance Version 1.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/events/maintenance/history
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/events/maintenance/history
```

Voici un exemple de sortie avec des informations sur un événement de redémarrage du système qui a été annulé et un événement de redémarrage du système qui a été terminé, au format JSON.

```
[
```

```
  {
    "NotBefore" : "21 Jan 2019 09:00:43 GMT",
    "Code" : "system-reboot",
    "Description" : "[Canceled] scheduled reboot",
    "EventId" : "instance-event-0d59937288b749b32",
    "NotAfter" : "21 Jan 2019 09:17:23 GMT",
    "State" : "canceled"
  },
  {
    "NotBefore" : "29 Jan 2019 09:00:43 GMT",
    "Code" : "system-reboot",
    "Description" : "[Completed] scheduled reboot",
    "EventId" : "instance-event-0d59937288b749b32",
    "NotAfter" : "29 Jan 2019 09:17:23 GMT",
    "State" : "completed"
  }
]
```

Personnalisation des notifications d'événements planifiés

Vous pouvez personnaliser les notifications d'événements planifiés pour inclure des balises dans la notification par e-mail. Cela facilite l'identification de la ressource affectée (instances ou Hôtes dédiés) et la hiérarchisation des actions pour l'événement à venir.

Lorsque vous personnalisez les notifications d'événements pour inclure des balises, vous pouvez choisir d'inclure :

- Toutes les balises associées à la ressource affectée
- Seules les balises spécifiques associées à la ressource affectée

Par exemple, supposons que vous assignez les balises `application`, `costcenter`, `project` et `owner` à toutes vos instances. Vous pouvez choisir d'inclure toutes les balises dans les notifications d'événements. Sinon, si vous souhaitez afficher uniquement les balises `project` et `owner` dans les notifications d'événements, vous pouvez choisir d'inclure uniquement ces balises.

Après avoir sélectionné les balises à inclure, les notifications d'événement incluront l'ID de ressource (ID d'instance ou Hôte dédié) et les paires clé de balise et valeur associées à la ressource affectée.

Rubriques

- Inclusion de balises dans les notifications d'événements (p. 754)
- Suppression de balises des notifications d'événements (p. 755)
- Affichage des balises à inclure dans les notifications d'événements (p. 756)

Inclusion de balises dans les notifications d'événements

Les balises que vous choisissez d'inclure s'appliquent à toutes les ressources (instances et Hôtes dédiés) de la région sélectionnée. Pour personnaliser les notifications d'événements dans d'autres régions, sélectionnez d'abord la région requise, puis effectuez les étapes suivantes.

Vous pouvez inclure des balises dans les notifications d'événements à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour inclure des balises dans les notifications d'événements

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le volet de navigation, sélectionnez Événements.
3. Choisissez Actions, Manage event notifications (Gérer les notifications d'événements).
4. Sélectionnez Include resource tags in event notifications (Inclure les balises de ressources dans les notifications d'événements).
5. Faites l'une des opérations suivantes, en fonction des balises que vous souhaitez inclure dans les notifications d'événement :
   - Pour inclure toutes les balises associées à l'instance affectée ou Hôte dédié, sélectionnez Include all resource tags (Inclure toutes les balises de ressource).
   - Pour sélectionner manuellement les balises à inclure, sélectionnez Choose the tags to include (Choisir les balises à inclure), puis pour Choose the tags to include (Choisir les balises à inclure), entrez la touche de balise et appuyez sur Entrée.
6. Choisissez Enregistrer.

AWS CLI

Pour inclure toutes les balises dans les notifications d'événements

Utilisez la commande register-instance-event-notification-attributes de l'AWS CLI et définissez le paramètre `IncludeAllTagsOfInstance` sur `true`.

```
aws ec2 register-instance-event-notification-attributes --instance-tag-attribute
 "IncludeAllTagsOfInstance=true"
```

Pour inclure des balises spécifiques dans les notifications d'événements

Utilisez la commande register-instance-event-notification-attributes de l'AWS CLI et spécifiez les balises à inclure avec le paramètre `InstanceTagKeys`.

```
aws ec2 register-instance-event-notification-attributes --instance-tag-attribute
 'InstanceTagKeys=["tag_key_1", "tag_key_2", "tag_key_3"]'
```

Suppression de balises des notifications d'événements

Vous pouvez supprimer les balises des notifications d'événements à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour supprimer les balises des notifications d'événements

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Événements.
3. Choisissez Actions, Manage event notifications (Gérer les notifications d'événements).
4. Effectuer l'une des actions suivantes, en fonction de la balise que vous souhaitez supprimer des notifications d'événement.
   - Pour supprimer toutes les balises des notifications d'événement, désactivez Include resource tags in event notifications (Inclure les balises de ressource dans les notifications d'événement).
   - Pour supprimer des balises spécifiques des notifications d'événements, choisissez Remove (Supprimer) (X) pour les balises répertoriées sous le champ Choose the tags to include (Choisir les balises à inclure).
5. Choisissez Enregistrer.

AWS CLI

Pour supprimer toutes les balises des notifications d'événements

Utilisez la commande deregister-instance-event-notification-attributes de l'AWS CLI et définissez le paramètre `IncludeAllTagsOfInstance` sur `false`.

```
aws ec2 deregister-instance-event-notification-attributes --instance-tag-attribute
 "IncludeAllTagsOfInstance=false"
```

Pour supprimer des balises spécifiques des notifications d'événements

Utilisez la commande deregister-instance-event-notification-attributes de l'AWS CLI et spécifiez les balises à supprimer à l'aide du paramètre `InstanceTagKeys`.

```
aws ec2 deregister-instance-event-notification-attributes --instance-tag-attribute
 'InstanceTagKeys=["tag_key_1", "tag_key_2", "tag_key_3"]'
```

Affichage des balises à inclure dans les notifications d'événements

Vous pouvez afficher les balises qui doivent être incluses dans les notifications d'événement à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour afficher les balises à inclure dans les notifications d'événements

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Événements.
3. Choisissez Actions, Manage event notifications (Gérer les notifications d'événements).

AWS CLI

Pour afficher les balises à inclure dans les notifications d'événements

Utilisez la commande describe-instance-event-notification-attributes de l'AWS CLI.

```
aws ec2 describe-instance-event-notification-attributes
```

Gestion des instances planifiées pour être arrêtées ou mises hors service

Quand AWS détecte une défaillance irrémédiable de l'hôte sous-jacent pour votre instance, il planifie l'arrêt ou la fin de l'instance en fonction du type de périphérique racine pour l'instance. Si le périphérique racine est un volume EBS, l'arrêt de l'instance est planifié. Si le périphérique racine est un volume de stockage d'instance, la fin de l'instance est planifiée. Pour plus d'informations, consultez Mise hors service d'instance (p. 642).

Important

Les données stockées sur des volumes de stockage d'instance sont perdues lorsque l'instance est arrêtée, mise en veille prolongée ou résiliée. Ceci inclut les volumes de stockage d'instance

attachés à une instance ayant un volume EBS comme périphérique racine. Veillez à enregistrer les données de vos volumes de stockage d'instance dont vous aurez besoin ultérieurement avant que l'instance ne soit arrêtée, mise en veille prolongée ou résiliée.

Actions pour les instances basées sur Amazon EBS

Vous pouvez attendre que l'instance s'arrête comme planifié. Sinon, vous pouvez arrêter et démarrer l'instance vous-même, ce qui la migre vers un nouvel ordinateur hôte. Pour plus d'informations sur l'arrêt de votre instance, ainsi que des informations sur les changements apportés à la configuration de votre instance lorsque celle-ci est arrêtée, consultez Arrêt et démarrage de votre instance (p. 625).

Vous pouvez automatiser un arrêt immédiat et un démarrage en réponse à un événement planifié d'arrêt d'instance. Pour plus d'informations, consultez Automatisation d'actions pour des instances EC2 dans le Guide de l'utilisateur AWS Health.

Actions pour les instances basées sur le stockage d'instance

Nous vous recommandons de lancer une instance de remplacement à partir de votre AMI la plus récente et de migrer toutes les données nécessaires vers l'instance de remplacement avant que l'instance ne soit planifiée pour prendre fin. Ensuite, vous pouvez mettre fin à l'instance d'origine ou attendre que l'instance prenne fin comme planifié.

Gestion des instances planifiées pour un redémarrage

Quand AWS doit effectuer des tâches comme installer des mises à niveau ou assurer la maintenance de l'ordinateur hôte sous-jacent, il peut planifier le redémarrage d'une instance ou de l'ordinateur hôte sous-jacent pour l'instance. Vous pouvez reprogrammer la plupart des événements de redémarrage (p. 759) afin que votre instance soit redémarrée à une date et une heure spécifiques qui vous conviennent.

Si vous arrêtez votre instance EC2 classique (p. 952) liée, celle-ci est automatiquement détachée du VPC et les groupes de sécurité VPC ne sont plus associés à l'instance. Vous pourrez lier à nouveau l'instance au VPC après l'avoir redémarrée.

Affichage du type d'événement de redémarrage

Vous pouvez déterminer si l'événement de redémarrage est un redémarrage d'instance ou de système à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour afficher le type d'événement de redémarrage planifié à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Events.
3. Choisissez Type de ressource : instance dans la liste des filtres.
4. Pour chaque instance, affichez la valeur dans la colonne Type d'événement. La valeur est soit system-reboot (redémarrage du système), soit instance-reboot (redémarrage de l'instance).

Ancienne console

Pour afficher le type d'événement de redémarrage planifié à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Events (Événements).
3. Choisissez Instance resources (Ressources d'instance) dans la liste de filtres.

4. Pour chaque instance, affichez la valeur dans la colonne Event Type (Type d'événement). La valeur est soit system-reboot (redémarrage du système), soit instance-reboot (redémarrage de l'instance).

AWS CLI

Pour afficher le type d'événement de redémarrage planifié à l'aide de la AWS CLI

Utilisez la commande describe-instance-status.

```
aws ec2 describe-instance-status --instance-id i-1234567890abcdef0
```

Pour les événements de redémarrage programmés, la valeur de `Code` est soit `system-reboot` ou `instance-reboot`. L'exemple de sortie suivant affiche un événement `system-reboot`.

```
[
    "Events": [
                {
                    "InstanceEventId": "instance-event-0d59937288b749b32",
                    "Code": "system-reboot",
                    "Description": "The instance is scheduled for a reboot",
                    "NotAfter": "2019-03-14T22:00:00.000Z",
                    "NotBefore": "2019-03-14T20:00:00.000Z",
                    "NotBeforeDeadline": "2019-04-05T11:00:00.000Z"
                }
               ]
]
```

Actions pour le redémarrage d'une instance

Vous pouvez attendre que le redémarrage d'instance se produise dans sa fenêtre de maintenance planifiée, reprogrammer (p. 759) le redémarrage d'instance à une date et heure qui vous conviennent, ou redémarrer (p. 641) vous-même l'instance au moment qui vous convient.

Après le redémarrage de votre instance, l'événement planifié pour le redémarrage d'instance est annulé et la description de l'événement est mise à jour. La maintenance en attente pour l'hôte sous-jacent est effectuée et vous pouvez recommencer à utiliser votre instance après son démarrage complet.

Actions pour le redémarrage du système

Vous ne pouvez pas redémarrer le système vous-même. Vous pouvez attendre que le redémarrage de système se produise dans sa fenêtre de maintenance planifiée, ou vous pouvez reprogrammer (p. 759) le redémarrage de système à une date et heure qui vous conviennent. Un redémarrage de système se termine généralement en quelques minutes. Une fois le redémarrage du système effectué, l'instance conserve son adresse IP et son nom DNS, et les données sur les volumes de stockage d'instance locaux sont conservées. Une fois le redémarrage du système achevé, l'événement planifié pour l'instance est effacé et vous pouvez vérifier que les logiciels sur votre instance fonctionnent comme prévu.

Sinon, s'il est nécessaire d'intervenir sur l'instance à un autre moment et que vous ne pouvez pas reprogrammer le redémarrage du système, vous pouvez arrêter et démarrer une instance basée sur les volumes Amazon EBS, ce qui la migre vers un nouvel hôte. Par contre, les données sur les volumes de stockage d'instance locaux ne sont pas conservées. Vous pouvez également automatiser un arrêt d'instance immédiat et un démarrage en réponse à un événement planifié de réinitialisation du système. Pour plus d'informations, consultez Automatisation d'actions pour des instances EC2 dans le Guide de l'utilisateur AWS Health. Dans le cas d'une instance basée sur le stockage d'instance, si vous ne pouvez pas reprogrammer le redémarrage de système, vous pouvez alors lancer une instance de remplacement à

partir de votre AMI la plus récente, migrer toutes les données nécessaires vers l'instance de remplacement avant la fenêtre de maintenance planifiée, puis mettre fin à l'instance d'origine.

Gestion des instances planifiées pour une maintenance

Quand AWS doit effectuer la maintenance sur l'hôte sous-jacent pour une instance, il planifie une maintenance pour l'instance. Il existe deux types d'événements de maintenance : maintenance du réseau et maintenance de l'alimentation.

Lors d'une maintenance du réseau, les instances planifiées perdent leur connectivité réseau pendant une courte période. La connectivité réseau normale vers votre instance est restaurée une fois la maintenance terminée.

Lors d'une maintenance de l'alimentation, les instances planifiées sont mises hors ligne pendant une courte période, puis redémarrées. Lorsqu'un redémarrage est effectué, les paramètres de configuration de votre instance sont conservés.

Une fois que votre instance a redémarré (cela prend normalement quelques minutes), vérifiez que votre application fonctionne comme prévu. À ce stade, votre instance ne devrait plus avoir d'événement planifié associé. Dans le cas contraire, la description de l'événement planifié commence par [Terminé]. Cela peut parfois prendre jusqu'à 1 heure pour que la description de statut de cette instance soit actualisée. Les événements de maintenance terminés restent affichés sur le tableau de bord de la console Amazon EC2 pendant une semaine maximum.

Actions pour les instances basées sur Amazon EBS

Vous pouvez attendre que la maintenance ait lieu comme planifié. Sinon, vous pouvez arrêter et démarrer l'instance, ce qui la migre vers un nouvel hôte. Pour plus d'informations sur l'arrêt de votre instance, ainsi que des informations sur les changements apportés à la configuration de votre instance lorsque celle-ci est arrêtée, consultez Arrêt et démarrage de votre instance (p. 625).

Vous pouvez automatiser un arrêt immédiat et un démarrage en réponse à un événement planifié de maintenance. Pour plus d'informations, consultez Automatisation d'actions pour des instances EC2 dans le Guide de l'utilisateur AWS Health.

Actions pour les instances basées sur le stockage d'instance

Vous pouvez attendre que la maintenance ait lieu comme planifié. Sinon, si vous souhaitez conserver un fonctionnement normal pendant une fenêtre de maintenance planifiée, vous pouvez lancer une instance de remplacement à partir de votre AMI la plus récente, migrer toutes les données nécessaires vers l'instance de remplacement avant la fenêtre de maintenance planifiée, puis mettre fin à l'instance d'origine.

Replanification d'un événement planifié

Vous pouvez replanifier un événement de sorte qu'il se produise à une date et une heure spécifiques qui vous conviennent. Seuls les événements ayant une date d'échéance peuvent être reprogrammés. Il existe d'autres restrictions pour la reprogrammation d'un événement (p. 761).

Vous pouvez replanifier un événement à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour replanifier un événement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Événements.
3. Choisissez Type de ressource : instance dans la liste des filtres.

4. Sélectionnez une ou plusieurs instances, puis sélectionnez Actions, Schedule Event (Programmer un événement).

   Seuls les événements ayant une date d'échéance, indiquée par la valeur Event Deadline (Échéance de l'événement), peuvent être reprogrammés. Si l'un des événements sélectionnés n'a pas de date d'échéance, Actions, Schedule Event (Programmer un événement) est désactivé.
5. Dans New start time (Nouvelle heure de début), saisissez une nouvelle date et une nouvelle heure pour l'événement. La nouvelle date et et la nouvelle heure doivent être antérieures à la valeur de Event Deadline (Échéance de l'événement).
6. Choisissez Enregistrer.

   L'heure de démarrage mise à jour peut prendre 1 à 2 minutes pour s'afficher dans la console.

Ancienne console

Pour replanifier un événement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Events (Événements).
3. Choisissez Instance resources (Ressources d'instance) dans la liste de filtres.
4. Sélectionnez une ou plusieurs instances, puis sélectionnez Actions, Schedule Event (Programmer un événement).

   Seuls les événements dotés d'une date d'échéance d'événement, indiquée par la valeur Event Deadline (Échéance d'événement), peuvent être reprogrammés.
5. Dans Event start time (Heure de début de l'événement), saisissez une nouvelle date et heure pour l'événement. La nouvelle date et et la nouvelle heure doivent être antérieures à la valeur de Event Deadline (Échéance de l'événement).
6. Sélectionnez Programmer un événement.

   L'heure de démarrage mise à jour peut prendre 1 à 2 minutes pour s'afficher dans la console.

AWS CLI

Pour replanifier un événement à l'aide de AWS CLI

1. Seuls les événements dotés d'une date d'échéance d'événement, indiquée par la valeur pour `NotBeforeDeadline`, peuvent être reprogrammés. Utilisez la commande describe-instance-status pour afficher la valeur de paramètre `NotBeforeDeadline`.

   ```
   aws ec2 describe-instance-status --instance-id i-1234567890abcdef0
   ```

   L'exemple de sortie suivant illustre un événement `system-reboot` qui peut être reprogrammé, car `NotBeforeDeadline` contient une valeur.

   ```
   [
       "Events": [
                   {
                       "InstanceEventId": "instance-event-0d59937288b749b32",
                       "Code": "system-reboot",
                       "Description": "The instance is scheduled for a reboot",
                       "NotAfter": "2019-03-14T22:00:00.000Z",
                       "NotBefore": "2019-03-14T20:00:00.000Z",
                       "NotBeforeDeadline": "2019-04-05T11:00:00.000Z"
                   }
   ```

```
                ]
]
```

2. Pour reprogrammer l'événement, utilisez la commande modify-instance-event-start-time. Spécifiez la nouvelle heure de début de l'événement à l'aide du paramètre `not-before`. La nouvelle heure de début doit se situer avant la `NotBeforeDeadline`.

```
aws ec2 modify-instance-event-start-time --instance-id i-1234567890abcdef0
 --instance-event-id instance-event-0d59937288b749b32 --not-
before 2019-03-25T10:00:00.000
```

Cela peut prendre 1 à 2 minutes pour que la commande describe-instance-status renvoie la valeur de paramètre `not-before` mise à jour.

Limites

- Seuls les événements dotés d'une date d'échéance d'événement peuvent être reprogrammés. L'événement peut être reprogrammé jusqu'à la date d'échéance de celui-ci. La colonne Échéance dans la console et le champ `NotBeforeDeadline` dans l'AWS CLI indiquent si l'événement comporte une date d'échéance.
- Seuls les événements n'ayant pas encore démarré peuvent être reprogrammés. La colonne Heure de début dans la console et le champ `NotBefore` dans l'AWS CLI indiquent l'heure de début de l'événement. Les événements programmés pour un lancement dans les 5 prochaines minutes ne peuvent pas être reprogrammés.
- La nouvelle heure de début doit être au moins 60 minutes après l'heure actuelle.
- Si vous reprogrammez plusieurs événements à l'aide de la console, la date d'échéance de l'événement est déterminée par l'événement avec la date d'échéance d'événement la plus proche.

Surveillance de vos instances avec CloudWatch

Vous pouvez surveiller vos instances avec Amazon CloudWatch, qui recueille et traite les données brutes d'Amazon EC2 en métriques lisibles et disponibles presque en temps réel. Ces statistiques sont enregistrées pour une durée de 15 mois et, par conséquent, vous pouvez accéder aux informations historiques et acquérir un meilleur point de vue de la façon dont votre service ou application web s'exécute.

Par défaut, Amazon EC2 envoie les données des métriques à CloudWatch toutes les 5 minutes. Pour envoyer les données des métriques de votre instance à CloudWatch toutes les minutes, vous pouvez activer la surveillance détaillée sur l'instance. Pour plus d'informations, consultez Activer ou désactiver la surveillance détaillée pour vos instances (p. 762).

La console Amazon EC2 affiche un ensemble de graphiques basés sur les données brutes envoyées par Amazon CloudWatch. En fonction de vos besoins, vous pouvez choisir d'utiliser Amazon CloudWatch ou les graphiques de la console pour obtenir les données relatives à vos instances.

Pour plus d'informations sur Amazon CloudWatch, consultez le Guide de l'utilisateur Amazon CloudWatch.

Sommaire

- Activer ou désactiver la surveillance détaillée pour vos instances (p. 762)
- Répertorier les métriques CloudWatch disponibles pour vos instances (p. 764)
- Obtenir les statistiques des métriques de vos instances (p. 775)
- Représenter graphiquement les métriques de vos instances (p. 783)
- Créer une alarme CloudWatch pour une instance (p. 783)
- Créer des alarmes qui arrêtent, finissent, redémarrent ou récupèrent une instance (p. 785)

Activer ou désactiver la surveillance détaillée pour vos instances

Par défaut, la surveillance basique est activée pour votre instance, mais vous pouvez activer la surveillance détaillée si vous le souhaitez. Après que vous avez activé la surveillance détaillée, la console Amazon EC2 affiche les graphiques de surveillance toutes les minutes pour l'instance.

Ce qui suit décrit l'intervalle de données et les frais pour la surveillance de base et détaillée des instances.

Surveillance de base

Les données sont disponibles automatiquement toutes les 5 minutes sans coût aucun.

Surveillance détaillée

Les données sont disponibles toutes les minutes contre des frais supplémentaires.

Pour obtenir le niveau de données, vous devez l'activer spécifiquement pour l'instance. Pour les instances où vous avez activé la surveillance détaillée, vous pouvez également obtenir les donnés agrégées à partir de groupes d'instances similaires.

Frais de surveillance détaillée

Si vous activez la surveillance détaillée, vous êtes facturé par métrique envoyée à CloudWatch. Vous n'êtes pas facturé pour le stockage des données. Pour de plus amples informations sur la tarification pour une surveillance détaillée, veuillez consulter Niveau payant sur la page de tarification Amazon CloudWatch. Pour obtenir un exemple de tarification, veuillez consulter Exemple 1 - Surveillance détaillée EC2 sur la page de tarification Amazon CloudWatch.

Activation de la surveillance détaillée

Vous pouvez activer la surveillance détaillée sur une instance lors de son lancement, ou une fois qu'elle est en cours d'exécution ou arrêtée. L'activation de la surveillance détaillée d'une instance n'affecte pas la surveillance des volumes EBS attachés à l'instance. Pour plus d'informations, consultez Métriques Amazon CloudWatch pour Amazon EBS (p. 1247).

New console

Pour activer la surveillance détaillée d'une instance existante (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Surveillance, Gérer la surveillance détaillée.
4. Sur la page Surveillance détaillée, pour Surveillance détaillée, sélectionnez la case à cocher Activer.
5. Choisissez Enregistrer.

Old console

Pour activer la surveillance détaillée d'une instance existante (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Supervision de CloudWatch, Activer la surveillance détaillée.
4. Dans la boîte de dialogue Enable Detailed Monitoring, choisissez Yes, Enable.

5. Choisissez Fermer.

Pour activer la surveillance détaillée lors du lancement d'une instance (console)

Lors du lancement d'une instance via AWS Management Console, cochez la case Surveillance sur la page Configure Instance Details.

Pour activer la surveillance détaillée d'une instance existante (AWS CLI)

Utilisez la commande monitor-instances suivante pour activer la surveillance détaillée des instances spécifiées.

```
aws ec2 monitor-instances --instance-ids i-1234567890abcdef0
```

Pour activer la surveillance détaillée lors du lancement d'une instance (AWS CLI)

Utilisez la commande run-instances avec l'indicateur `--monitoring` pour activer la surveillance détaillée.

```
aws ec2 run-instances --image-id ami-09092360 --monitoring Enabled=true...
```

Désactivation de la surveillance détaillée

Vous pouvez désactiver la surveillance détaillée sur une instance lors de son lancement, ou une fois qu'elle est en cours d'exécution ou arrêtée.

New console

Pour désactiver la surveillance détaillée (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Surveillance, Gérer la surveillance détaillée.
4. Dans la page Surveillance détaillée, pour Surveillance détaillée, désactivez la case à cocher Activer.
5. Choisissez Enregistrer.

Old console

Pour désactiver la surveillance détaillée (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Supervision de CloudWatch, Désactiver la surveillance détaillée.
4. Dans la boîte de dialogue Disable Detailed Monitoring, choisissez Yes, Disable.
5. Choisissez Fermer.

Pour désactiver la surveillance détaillée (AWS CLI)

Utilisez la commande unmonitor-instances suivante pour désactiver la surveillance détaillée des instances spécifiées.

```
aws ec2 unmonitor-instances --instance-ids i-1234567890abcdef0
```

Répertorier les métriques CloudWatch disponibles pour vos instances

Amazon EC2 envoie les métriques à Amazon CloudWatch. Vous pouvez utiliser AWS Management Console, l'AWS CLI ou une API pour répertorier les métriques qu'Amazon EC2 envoie à CloudWatch. Par défaut, chaque point de données couvre les 5 minutes suivant l'heure de début d'activité de l'instance. Si vous avez activé la surveillance détaillée, chaque point de données couvre la minute suivant l'activité à compte de l'heure de début.

Pour plus d'informations sur la façon d'obtenir les statistiques pour ces métriques, consultez Obtenir les statistiques des métriques de vos instances (p. 775).

Sommaire

- Métriques des instances (p. 764)
- Métriques des crédits UC (p. 767)
- Métriques Amazon EBS pour des instances basées sur Nitro (p. 768)
- Métriques de contrôle de statut (p. 770)
- Métriques de mise en miroir du trafic (p. 771)
- Dimensions de métriques Amazon EC2 (p. 771)
- Métriques d'utilisation Amazon EC2 (p. 771)
- Répertorier les métriques à l'aide de la console (p. 772)
- Répertorier les métriques à l'aide de l'interface ligne de commande AWS CLI (p. 774)

Métriques des instances

L'espace de nom `AWS/EC2` inclut les métriques d'instance suivantes.

Métrique	Description
`CPUUtilization`	Pourcentage d'unités de calcul EC2 allouées actuellement utilisées dans l'instance. Cette métrique identifie la puissance de traitement requise pour exécuter une application sur une instance sélectionnée. Selon le type d'instance, les outils de votre système d'exploitation peuvent afficher un pourcentage plus bas que CloudWatch quand l'instance n'a pas un cœur complet de processeur alloué. Unités : pourcentage
`DiskReadOps`	Opérations de lecture terminées de tous les volumes de stockage d'instance disponibles pour l'instance, au cours de la période spécifiée. Pour calculer la moyenne d'opérations d'E/S par seconde (IOPS) pour la période, divisez le nombre total d'opérations de la période par le nombre de secondes de la période. S'il n'y a pas de volumes de stockage d'instance, la valeur est 0 ou la métrique n'est pas prise en charge. Unités : nombre

Métrique	Description
`DiskWriteOps`	Opérations d'écriture terminées dans tous les volumes de stockage d'instance disponibles pour l'instance, au cours de la période spécifiée. Pour calculer la moyenne d'opérations d'E/S par seconde (IOPS) pour la période, divisez le nombre total d'opérations de la période par le nombre de secondes de la période. S'il n'y a pas de volumes de stockage d'instance, la valeur est 0 ou la métrique n'est pas prise en charge. Unités : nombre
`DiskReadBytes`	Octets lus à partir de tous les volumes de stockage d'instance disponibles pour l'instance. Cette métrique permet de déterminer le volume de données que l'application lit à partir du disque dur de l'instance. Il est ainsi possible de déterminer la vitesse de l'application. Le nombre mentionné correspond au nombre d'octets reçus pendant la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour trouver le nombre d'octets/seconde. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. S'il n'y a pas de volumes de stockage d'instance, la valeur est 0 ou la métrique n'est pas prise en charge. Unités : octets
`DiskWriteBytes`	Octets écrits dans tous les volumes de stockage d'instance disponibles pour l'instance. Cette métrique permet de déterminer le volume de données que l'application écrit sur le disque dur de l'instance. Il est ainsi possible de déterminer la vitesse de l'application. Le nombre mentionné correspond au nombre d'octets reçus pendant la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour trouver le nombre d'octets/seconde. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. S'il n'y a pas de volumes de stockage d'instance, la valeur est 0 ou la métrique n'est pas prise en charge. Unités : octets

Métrique	Description
`NetworkIn`	Nombre d'octets reçus par l'instance sur toutes les interfaces réseau. Cette métrique identifie le volume du trafic réseau entrant d'une seule instance. Le nombre mentionné correspond au nombre d'octets reçus pendant la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour trouver le nombre d'octets/seconde. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. Unités : octets
`NetworkOut`	Nombre d'octets envoyés par l'instance sur toutes les interfaces réseau. Cette métrique identifie le volume du trafic réseau sortant d'une seule instance. Le nombre mentionné correspond au nombre d'octets envoyés pendant la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour trouver le nombre d'octets/seconde. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. Unités : octets
`NetworkPacketsIn`	Nombre de paquets reçus par l'instance sur toutes les interfaces réseau. Cette métrique identifie le volume du trafic entrant en ce qui concerne le nombre de paquets sur une seule instance. Cette métrique est disponible uniquement pour la surveillance basique. Unités : nombre Statistiques : Minimum, Maximum, moyenne
`NetworkPacketsOut`	Nombre de paquets envoyés par l'instance sur toutes les interfaces réseau. Cette métrique identifie le volume du trafic sortant en ce qui concerne le nombre de paquets sur une seule instance. Cette métrique est disponible uniquement pour la surveillance basique. Unités : nombre Statistiques : Minimum, Maximum, moyenne
`MetadataNoToken`	Nombre d'accès réussis au service de métadonnées d'instance en employant une méthode qui n'utilise pas de jeton. Cette métrique est utilisée pour déterminer s'il existe des processus accédant aux métadonnées d'instance qui utilisent Service des métadonnées d'instance Version 1, et qui n'utilisent pas de jeton. Si toutes les demandes utilisent des sessions basées sur un jeton, par ex., Service des métadonnées d'instance Version 2 la valeur est 0. Pour de plus amples informations, consultez Migration vers Service des métadonnées d'instance Version 2 (p. 703). Unités : nombre

Métriques des crédits UC

L'espace de noms `AWS/EC2` inclut les métriques de crédit UC suivantes pour vos instances à capacité extensible (p. 233).

Métrique	Description
`CPUCreditUsage`	Nombre de crédits UC dépensés par l'instance pour l'utilisation de l'UC. Par exemple, un crédit UC est équivalant à un processeur virtuel fonctionnant à 100 % d'utilisation pendant une minute ou une combinaison équivalente de processeurs virtuels, d'utilisation et de temps (par exemple, un processeur virtuel fonctionnant à 50 % d'utilisation pendant deux minutes, ou deux processeurs virtuels fonctionnant à 25 % d'utilisation pendant deux minutes). Les métriques de crédits UC sont disponibles uniquement toutes les 5 minutes. Si vous spécifiez une période supérieure à cinq minutes, utilisez la statistique `Sum` au lieu de la statistique `Average`. Unités : crédits (minutes vCPU)
`CPUCreditBalance`	Nombre de crédits UC gagnés qu'une instance a accumulés depuis son lancement ou son démarrage. Pour les instances T2 Standard, le `CPUCreditBalance` inclut également le nombre de crédits de lancement qui ont été accumulés. Les crédits sont accumulés dans le solde de crédits quand ils sont gagnés et supprimés du solde de crédits lorsqu'ils sont dépensés. Le solde de crédits présente une limite maximum qui est déterminée par la taille de l'instance. Une fois que la limite est atteinte, tous les nouveaux crédits gagnés sont rejetés. Pour les instances T2 Standard, les crédits de lancement ne sont pas comptés dans la limite. L'instance peut dépenser les crédits figurant dans le `CPUCreditBalance` pour dépasser le niveau de base de l'utilisation de l'UC. Les crédits figurant dans le `CPUCreditBalance` d'une instance en cours d'exécution n'expirent pas. Lorsqu'une instance T3 ou T3a s'arrête, la valeur `CPUCreditBalance` est conservée pendant sept jours. Au-delà, tous les crédits accumulés sont perdus. Lorsqu'une instance T2 s'arrête, la valeur de `CPUCreditBalance` n'est pas conservée, et tous les crédits accumulés sont perdus. Les métriques de crédits UC sont disponibles uniquement toutes les 5 minutes. Unités : crédits (minutes vCPU)
`CPUSurplusCreditBalance`	Nombre de crédits excédentaires ayant été dépensés par une instance `unlimited` lorsque la valeur `CPUCreditBalance` est nulle. La valeur de `CPUSurplusCreditBalance` est remboursée progressivement par les crédits UC gagnés. Si le nombre de crédits excédentaires dépasse le nombre maximum de crédits que l'instance

Métrique	Description
	peut gagner en 24 heures, les crédits excédentaires dépensés au-dessus du maximum génèrent des frais supplémentaires. Les métriques de crédits UC sont disponibles uniquement toutes les 5 minutes. Unités : crédits (minutes vCPU)
`CPUSurplusCreditsCharged`	Nombre de crédits excédentaires dépensés qui ne sont pas remboursés progressivement par les crédits UC gagnés et qui génèrent donc des frais supplémentaires. Les crédits excédentaires dépensés sont facturés lorsque l'une des situations suivantes se produit : • Les crédits excédentaires dépensés dépassent le nombre maximum de crédits que l'instance peut gagner sur une période de 24 heures. Les crédits excédentaires dépensés au-dessus de ce maximum sont facturés à la fin de l'heure. • L'instance est arrêtée ou résiliée. • L'instance bascule du mode `unlimited` au mode `standard`. Les métriques de crédits UC sont disponibles uniquement toutes les 5 minutes. Unités : crédits (minutes vCPU)

Métriques Amazon EBS pour des instances basées sur Nitro

L'espace de noms `AWS/EC2` inclut les métriques Amazon EBS suivantes pour les instances basées sur Nitro qui ne sont pas des instances de type matériel nu. Pour obtenir la liste des types d'instances basés sur Nitro, consultez Instances reposant sur le système Nitro (p. 218).

Les valeurs de métrique des instances basées sur Nitro seront toujours des entiers (nombres entiers), alors que les valeurs des instances basées sur Xen prennent en charge les nombres décimaux. Ainsi, une faible utilisation d'UC pour les instances basées sur Nitro peut être arrondie à 0.

Métrique	Description
`EBSReadOps`	Opérations de lecture terminées de tous les volumes Amazon EBS attachés à l'instance au cours de la période spécifiée. Pour calculer la moyenne d'opérations de lecture d'E/S par seconde (IOPS en lecture) pour la période, divisez le nombre total d'opérations de la période par le nombre de secondes de la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour calculer les IOPS en lecture. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. Unité : nombre

Métrique	Description
`EBSWriteOps`	Opérations d'écriture terminées de tous les volumes EBS attachés à l'instance au cours de la période spécifiée. Pour calculer la moyenne d'opérations d'écriture d'E/S par seconde (IOPS en écriture) pour la période, divisez le nombre total d'opérations de la période par le nombre de secondes de la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour calculer les IOPS en écriture. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. Unité : nombre
`EBSReadBytes`	Octets lus de tous les volumes EBS attachés à l'instance au cours de la période spécifiée. Le nombre mentionné correspond au nombre d'octets lus pendant la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour trouver le nombre d'octets lus/seconde. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. Unité : octets
`EBSWriteBytes`	Octets écrits dans tous les volumes EBS attachés à l'instance au cours de la période spécifiée. Le nombre mentionné correspond au nombre d'octets écrits pendant la période. Si vous utilisez une surveillance de base (cinq minutes), vous pouvez diviser ce nombre par 300 pour trouver le nombre d'octets écrits/seconde. Si vous avez recours à une surveillance détaillée (une minute), divisez-le par 60. Unité : octets
`EBSIOBalance%`	Disponible uniquement pour les plus petites tailles d'instance. Fournit des informations sur le pourcentage de crédits d'E/S restant dans le compartiment en rafales. Cette métrique est disponible uniquement pour la surveillance basique. La statistique `Sum` n'est pas applicable pour cette métrique. Unité : pourcentage

Métrique	Description
`EBSByteBalance%`	Disponible uniquement pour les plus petites tailles d'instance. Fournit des informations sur le pourcentage de crédits de débit restant dans le compartiment en rafales. Cette métrique est disponible uniquement pour la surveillance basique. La statistique `Sum` n'est pas applicable pour cette métrique. Unité : pourcentage

Pour plus d'informations sur les métriques fournies pour vos volumes EBS, consultez Métriques Amazon EBS (p. 1247). Pour plus d'informations sur les métriques fournies pour vos parcs d'instances Spot, consultez Métriques CloudWatch pour un Parc d'instances Spot (p. 437).

Métriques de contrôle de statut

L'espace de nom `AWS/EC2` inclut les métriques de contrôle de statut suivantes. Par défaut, les métriques de contrôle de statut sont disponibles à la fréquence d'1 minute sans frais supplémentaires. Pour une instance nouvellement lancée, les données de métriques de contrôle de statut sont disponibles uniquement une fois que l'état d'initialisation de l'instance a pris fin (quelques minutes après que l'instance passe à l'état en cours d'exécution). Pour plus d'informations sur les vérifications de statut EC2, veuillez consulter Contrôles de statut pour vos instances (p. 742).

Métrique	Description
`StatusCheckFailed`	Indique si l'instance a passé avec succès le contrôle de statut d'instance et le contrôle de statut de système au cours de la dernière minute. Cette métrique peut avoir la valeur 0 (succès) ou 1 (échec). Par défaut, cette métrique est disponible à la fréquence d'1 minute sans frais supplémentaires. Unités : nombre
`StatusCheckFailed_Instance`	Indique si l'instance a passé avec succès le contrôle de statut de l'instance de la dernière minute. Cette métrique peut avoir la valeur 0 (succès) ou 1 (échec). Par défaut, cette métrique est disponible à la fréquence d'1 minute sans frais supplémentaires. Unités : nombre
`StatusCheckFailed_System`	Indique si l'instance a passé avec succès le contrôle de statut du système de la dernière minute. Cette métrique peut avoir la valeur 0 (succès) ou 1 (échec). Par défaut, cette métrique est disponible à la fréquence d'1 minute sans frais supplémentaires.

Métrique	Description
	Unités : nombre

Métriques de mise en miroir du trafic

L'espace de noms `AWS/EC2` inclut des métriques pour le trafic mis en miroir. Pour de plus amples informations, veuillez consulter Surveillance du trafic mis en miroir à l'aide de Amazon CloudWatchdans le guide de la mise en miroir Amazon VPC.

Dimensions de métriques Amazon EC2

Vous pouvez utiliser les dimensions suivantes pour affiner les métriques répertoriées dans les tableaux précédents.

Dimension	Description
`AutoScalingGroupName`	Cette dimension filtre les données que vous demandez pour toutes les instances dans un groupe de capacité donné. Un groupe Auto Scaling est un ensemble d'instances que vous définissez si vous utilisez Auto Scaling. Cette dimension est disponible uniquement pour les métriques Amazon EC2 lorsque les instances sont dans un groupe Auto Scaling. Disponible pour les instances avec la surveillance détaillée ou basique activée.
`ImageId`	Cette dimension filtre les données que vous demandez pour toutes les instances exécutant cette AMI (Amazon Machine Image) Amazon EC2. Disponible pour les instances avec la surveillance détaillée activée.
`InstanceId`	Cette dimension filtre les données que vous demandez de l'instance identifiée uniquement. Cela vous aide à identifier une instance exacte à partir de laquelle surveiller les données.
`InstanceType`	Cette dimension filtre les données que vous demandez pour toutes les instances s'exécutant avec ce type d'instance spécifiée. Cela vous permet de classer vos données selon le type d'instance en cours d'exécution. Par exemple, vous pouvez comparer les données issues d'une instance m1.small et d'une instance m1.large pour déterminer qui a la meilleure valeur commerciale pour votre application. Disponible pour les instances avec la surveillance détaillée activée.

Métriques d'utilisation Amazon EC2

Vous pouvez utiliser les métriques d'utilisation CloudWatch pour fournir une visibilité sur l'utilisation des ressources de votre compte. Utilisez ces métriques pour visualiser votre utilisation actuelle du service sur des graphiques et des tableaux de bord CloudWatch.

Les métriques d'utilisation Amazon EC2 correspondent aux quotas de service AWS. Vous pouvez configurer des alarmes qui vous alertent lorsque votre utilisation approche d'un quota de service. Pour plus d'informations sur l'intégration CloudWatch avec les quotas de service, veuillez consulter Métriques d'intégration et d'utilisation des quotas de service.

Amazon EC2 publie les métriques suivantes dans l'espace de noms `AWS/Usage`.

Métrique	Description
`ResourceCount`	Nombre des ressources spécifiées exécutées dans votre compte. Les ressources sont définies par les dimensions associées à la métrique. La statistique la plus utile pour cette métrique est `MAXIMUM`, qui représente le nombre maximal de ressources utilisées pendant la période d'une minute.

Les dimensions suivantes permettent d'affiner les métriques d'utilisation publiées par Amazon EC2.

Dimension	Description
`Service`	Nom du service AWS contenant la ressource. Pour les métriques d'utilisation d'Amazon EC2, la valeur de cette dimension est `EC2`.
`Type`	Type d'entité faisant l'objet d'un rapport. Actuellement, la seule valeur valide pour les métriques d'utilisation d'Amazon EC2 est `Resource`.
`Resource`	Type de ressource en cours d'exécution. Actuellement, la seule valeur valide pour les métriques d'utilisation d'Amazon EC2 est `vCPU`, qui renvoie des informations sur les instances en cours d'exécution.
`Class`	Classe de ressource suivie. Pour les métriques d'utilisation d'Amazon EC2 avec `vCPU` comme valeur de la dimension `Resource`, les valeurs valides sont `Standard/OnDemand`, `F/OnDemand`, `G/OnDemand`, `Inf/OnDemand`, `P/OnDemand` et `X/OnDemand`. Les valeurs de cette dimension définissent la première lettre des types d'instance signalés par la métrique. Par exemple, `Standard/OnDemand` renvoie des informations sur toutes les instances en cours d'exécution dont les types commencent par A, C, D, H, I, M, R, T et Z, et `G/OnDemand` renvoie des informations sur toutes les instances en cours d'exécution dont les types commencent par G.

Répertorier les métriques à l'aide de la console

Les métriques sont d'abord regroupées par espace de noms, puis par les différentes combinaisons de dimension au sein de chaque espace de noms. Par exemple, vous pouvez afficher toutes les métriques fournies par Amazon EC2 ou les métriques regroupées par ID d'instance, type d'instance, ID d'image (AMI) ou groupe Auto Scaling.

Pour afficher les métriques disponibles par catégorie (console)

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, choisissez Métriques.
3. Choisissez l'espace de nom de métrique EC2.

4. Sélectionnez une dimension de métrique (Per-Instance Metrics (Métriques par instance) par exemple).

5. Pour trier les métriques, utilisez l'en-tête de colonne. Pour représenter graphiquement une métrique, cochez la case en regard de la métrique. Pour filtrer par ressource, sélectionnez l'ID de ressource, puis Add to search. Pour filtrer par métrique, choisissez le nom de la métrique, puis Add to search.

Répertorier les métriques à l'aide de l'interface ligne de commande AWS CLI

Utilisez la commande list-metrics afin de répertorier les métriques CloudWatch pour vos instances.

Pour répertorier toutes les métriques disponibles pour Amazon EC2 (AWS CLI)

L'exemple suivant spécifie l'espace de noms `AWS/EC2` pour afficher toutes les métriques pour Amazon EC2.

```
aws cloudwatch list-metrics --namespace AWS/EC2
```

Voici un exemple de sortie :

```
{
  "Metrics": [
    {
        "Namespace": "AWS/EC2",
        "Dimensions": [
            {
                "Name": "InstanceId",
                "Value": "i-1234567890abcdef0"
            }
        ],
        "MetricName": "NetworkOut"
    },
    {
        "Namespace": "AWS/EC2",
```

```
        "Dimensions": [
            {
                "Name": "InstanceId",
                "Value": "i-1234567890abcdef0"
            }
        ],
        "MetricName": "CPUUtilization"
    },
    {
        "Namespace": "AWS/EC2",
        "Dimensions": [
            {
                "Name": "InstanceId",
                "Value": "i-1234567890abcdef0"
            }
        ],
        "MetricName": "NetworkIn"
    },
    ...
  ]
}
```

Pour répertorier toutes les métriques disponibles pour une instance (AWS CLI)

L'exemple suivant spécifie l'espace de nom `AWS/EC2` et la dimension `InstanceId` pour afficher les résultats uniquement pour l'instance spécifiée.

```
aws cloudwatch list-metrics --namespace AWS/EC2 --dimensions
 Name=InstanceId,Value=i-1234567890abcdef0
```

Pour répertorier une métrique dans toutes les instances (AWS CLI)

L'exemple suivant spécifie l'espace de nom `AWS/EC2` et un nom de métrique pour afficher les résultats uniquement pour la métrique spécifiée.

```
aws cloudwatch list-metrics --namespace AWS/EC2 --metric-name CPUUtilization
```

Obtenir les statistiques des métriques de vos instances

Vous pouvez obtenir les statistiques des métriques CloudWatch de vos instances.

Sommaire

- Vue d'ensemble des statistiques (p. 775)
- Obtenir les statistiques d'une instance spécifique (p. 776)
- Regrouper les statistiques à travers les instances (p. 779)
- Regroupement de statistiques par groupe Auto Scaling (p. 781)
- Regroupement de statistiques par AMI (p. 782)

Vue d'ensemble des statistiques

Les statistiques sont des regroupements de données de métriques sur des périodes spécifiées. CloudWatchfournit des statistiques basées sur les points de données des métriques fournies par vos

données personnalisées ou fournies à CloudWatch par d'autres services AWS. Les regroupements sont effectués en utilisant l'espace de noms, le nom métrique, les dimensions et l'unité de mesure des points de données, pendant la période spécifiée. Le tableau suivant décrit les statistiques disponibles.

Statistique	Description
`Minimum`	La valeur la plus basse observée pendant la période spécifiée. Vous pouvez utiliser cette valeur pour déterminer les faibles volumes d'activité pour votre application.
`Maximum`	La valeur la plus haute observée pendant la période spécifiée. Vous pouvez utiliser cette valeur pour déterminer les volumes d'activité élevés pour votre application.
`Sum`	Toutes les valeurs soumises pour la métrique correspondante ajoutées ensemble. Cette statistique peut être utile pour déterminer le volume total d'une métrique.
`Average`	La valeur de `Sum` / `SampleCount` pendant la période spécifiée. En comparant cette statistique à `Minimum` et à `Maximum`, vous pouvez déterminer l'ampleur d'une métrique et si l'utilisation moyenne est proche de `Minimum` ou de `Maximum`. Cette comparaison vous permet de savoir quand augmenter ou diminuer vos ressources en fonction des besoins.
`SampleCount`	Le compte (nombre) des points de données utilisé pour le calcul statistique.
`pNN.NN`	Valeur du centile spécifié. Vous pouvez spécifier un centile en utilisant jusqu'à deux décimales (par exemple, p95.45).

Obtenir les statistiques d'une instance spécifique

Les exemples ci-dessous montrent comment déterminer l'utilisation maximale de l'UC d'une instance EC2 avec AWS Management Console ou à l'aide de l'AWS CLI.

Prérequis

- Vous devez avoir l'ID de l'instance. Vous pouvez obtenir l'ID d'instance en utilisant AWS Management Console ou la commande describe-instances.
- Par défaut, la surveillance basique est activée, mais vous pouvez activer la surveillance détaillée. Pour plus d'informations, consultez Activer ou désactiver la surveillance détaillée pour vos instances (p. 762).

Pour afficher l'utilisation d'UC d'une instance spécifique (console)

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, choisissez Métriques.
3. Choisissez l'espace de nom de métrique EC2.

4. Choisissez la dimension Per-Instance Metrics (Métriques par instance).

5. Dans le champ de recherche, entrez **`CPUUtilization`**, puis appuyez sur Entrée. Choisissez la ligne de l'instance spécifique, qui contient un graphique pour la métrique CPUUtilization de l'instance. Pour nommer le graphique, choisissez l'icône en forme de crayon. Pour modifier la plage de temps, sélectionnez l'une des valeurs prédéfinies ou choisissez custom.

6. Pour modifier la statistique ou la période pour la métrique, choisissez l'onglet Graphed metrics. Sélectionnez l'en-tête de colonne ou une valeur individuelle et choisissez une autre valeur.

Pour obtenir l'utilisation d'UC pour une instance spécifique (AWS CLI)

Utilisez la commande suivante get-metric-statistics afin d'obtenir la métrique CPUUtilization pour l'instance spécifiée à l'aide de la période et de l'intervalle de temps spécifiés :

```
aws cloudwatch get-metric-statistics --namespace AWS/EC2 --metric-name CPUUtilization  --
period 3600 \
--statistics Maximum --dimensions Name=InstanceId,Value=i-1234567890abcdef0 \
--start-time 2016-10-18T23:18:00 --end-time 2016-10-19T23:18:00
```

Voici un exemple de sortie. Chaque valeur représente le pourcentage d'utilisation maximale de l'UC pour une seule instance EC2.

```
{
    "Datapoints": [
        {
            "Timestamp": "2016-10-19T00:18:00Z",
            "Maximum": 0.33000000000000002,
            "Unit": "Percent"
        },
        {
            "Timestamp": "2016-10-19T03:18:00Z",
            "Maximum": 99.670000000000002,
            "Unit": "Percent"
        },
        {
            "Timestamp": "2016-10-19T07:18:00Z",
            "Maximum": 0.34000000000000002,
            "Unit": "Percent"
        },
        {
            "Timestamp": "2016-10-19T12:18:00Z",
            "Maximum": 0.34000000000000002,
            "Unit": "Percent"
        },
        ...
    ],
    "Label": "CPUUtilization"
}
```

Regrouper les statistiques à travers les instances

Les statistiques agrégées sont disponibles pour les instances pour lesquelles la surveillance détaillée a été activée. Les instances qui utilisent la surveillance basique ne sont pas incluses dans les regroupements. De plus, Amazon CloudWatch ne regroupe pas les données de différentes régions. Par conséquent, les métriques sont totalement séparées entre les régions. Avant de pouvoir obtenir des statistiques regroupées entre les instances, vous devez activer la surveillance détaillée (avec coût additionnel) qui fournit des données toutes les minutes.

Cet exemple vous montre comment obtenir l'utilisation moyenne de l'UC pour vos instances EC2 à l'aide de la surveillance détaillée. Comme aucune dimension n'est spécifiée, CloudWatch retourne les statistiques pour toutes les dimensions de l'espace de noms `AWS/EC2`.

> **Important**
>
> Cette technique d'extraction de toutes les dimensions à travers un espace de noms AWS ne fonctionne pas pour les espaces de noms que vous publiez sur Amazon CloudWatch. Avec les espaces de noms personnalisés, vous devez spécifier l'ensemble complet des dimensions associées à un point de données particulier pour pouvoir extraire les statistiques qui incluent le point de données.

Pour afficher l'utilisation moyenne de l'UC dans vos instances (console)

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, choisissez Métriques.

3. Choisissez l'espace de noms EC2, puis choisissez Across All Instances (Sur toutes les instances).
4. Choisissez la ligne contenant CPUUtilization qui affiche un graphique pour la métrique pour toutes vos instances EC2. Pour nommer le graphique, choisissez l'icône en forme de crayon. Pour modifier la plage de temps, sélectionnez l'une des valeurs prédéfinies ou choisissez custom.

5. Pour modifier la statistique ou la période pour la métrique, choisissez l'onglet Graphed metrics. Sélectionnez l'en-tête de colonne ou une valeur individuelle et choisissez une autre valeur.

Pour obtenir l'utilisation moyenne de l'UC sur vos instances (AWS CLI)

Utilisez la commande get-metric-statistics comme suit pour obtenir l'utilisation moyenne de la métrique CPUUtilization dans vos instances.

```
aws cloudwatch get-metric-statistics \
    --namespace AWS/EC2 \
    --metric-name CPUUtilization \
    --period 3600  --statistics "Average" "SampleCount" \
    --start-time 2016-10-11T23:18:00 \
    --end-time 2016-10-12T23:18:00
```

Voici un exemple de sortie :

```
{
    "Datapoints": [
        {
            "SampleCount": 238.0,
            "Timestamp": "2016-10-12T07:18:00Z",
            "Average": 0.038235294117647062,
            "Unit": "Percent"
        },
```

```
        {
            "SampleCount": 240.0,
            "Timestamp": "2016-10-12T09:18:00Z",
            "Average": 0.16670833333333332,
            "Unit": "Percent"
        },
        {
            "SampleCount": 238.0,
            "Timestamp": "2016-10-11T23:18:00Z",
            "Average": 0.041596638655462197,
            "Unit": "Percent"
        },
        ...
    ],
    "Label": "CPUUtilization"
}
```

Regroupement de statistiques par groupe Auto Scaling

Vous pouvez regrouper des statistiques pour les instances EC2 dans un groupe Auto Scaling. Notez qu'Amazon CloudWatch ne peut pas regrouper de données entre les régions. Les métriques sont totalement séparées d'une région à une autre.

Cet exemple vous montre comment récupérer le nombre total d'octets écrits sur disque pour un groupe Auto Scaling. Le total est calculé par durée d'une minute sur une période de 24 heures pour toutes les instances EC2 dans le groupe Auto Scaling spécifié.

Pour afficher la métrique DiskWriteBytes pour les instances d'un groupe Auto Scaling (console)

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, choisissez Métriques.
3. Choisissez l'espace de noms EC2, puis choisissez By Auto Scaling Group (Par groupe Auto Scaling).
4. Choisissez la ligne pour la métrique DiskWriteBytes et le groupe Auto Scaling spécifique, qui affiche un graphique pour la métrique pour les instances faisant partie du groupe Auto Scaling. Pour nommer le graphique, choisissez l'icône en forme de crayon. Pour modifier la plage de temps, sélectionnez l'une des valeurs prédéfinies ou choisissez custom.
5. Pour modifier la statistique ou la période pour la métrique, choisissez l'onglet Graphed metrics. Sélectionnez l'en-tête de colonne ou une valeur individuelle et choisissez une autre valeur.

Pour afficher la métrique DiskWriteBytes pour les instances d'un groupe Auto Scaling (AWS CLI)

Utilisez la commande get-metric-statistics comme suit.

```
aws cloudwatch get-metric-statistics --namespace AWS/EC2 --metric-name DiskWriteBytes --
period 360 \
--statistics "Sum" "SampleCount" --dimensions Name=AutoScalingGroupName,Value=my-asg --
start-time 2016-10-16T23:18:00 --end-time 2016-10-18T23:18:00
```

Voici un exemple de sortie :

```
{
    "Datapoints": [
        {
            "SampleCount": 18.0,
            "Timestamp": "2016-10-19T21:36:00Z",
            "Sum": 0.0,
            "Unit": "Bytes"
```

```
        },
        {
            "SampleCount": 5.0,
            "Timestamp": "2016-10-19T21:42:00Z",
            "Sum": 0.0,
            "Unit": "Bytes"
        }
    ],
    "Label": "DiskWriteBytes"
}
```

Regroupement de statistiques par AMI

Vous pouvez regrouper des statistiques pour les instances dont la surveillance détaillée est activée. Les instances qui utilisent une surveillance basique ne sont pas incluses. Notez qu'Amazon CloudWatch ne peut pas regrouper de données entre les régions. Les métriques sont totalement séparées d'une région à une autre.

Avant de pouvoir obtenir des statistiques regroupées entre les instances, vous devez activer la surveillance détaillée (avec coût additionnel) qui fournit des données toutes les minutes. Pour plus d'informations, consultez Activer ou désactiver la surveillance détaillée pour vos instances (p. 762).

Cet exemple vous montre comment déterminer l'utilisation moyenne de l'UC pour toutes les instances qui utilisent une Amazon Machine Image (AMI) spécifique. La moyenne est calculée par intervalles de 60 secondes pour une période d'un jour.

Pour afficher l'utilisation moyenne de l'UC par AMI (console)

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, choisissez Métriques.
3. Choisissez l'espace de noms EC2, puis choisissez By Image (AMI) Id (Par ID d'image (AMI)).
4. Choisissez la ligne de la métrique CPUUtilization et l'AMI spécifique, qui affiche un graphique pour la métrique pour l'AMI spécifiée. Pour nommer le graphique, choisissez l'icône en forme de crayon. Pour modifier la plage de temps, sélectionnez l'une des valeurs prédéfinies ou choisissez custom.
5. Pour modifier la statistique ou la période pour la métrique, choisissez l'onglet Graphed metrics. Sélectionnez l'en-tête de colonne ou une valeur individuelle et choisissez une autre valeur.

Pour obtenir l'utilisation moyenne de l'UC pour un ID d'image (AWS CLI)

Utilisez la commande get-metric-statistics comme suit.

```
aws cloudwatch get-metric-statistics --namespace AWS/EC2 --metric-name CPUUtilization  --
period 3600 \
--statistics Average --dimensions Name=ImageId,Value=ami-3c47a355 --start-
time 2016-10-10T00:00:00 --end-time 2016-10-11T00:00:00
```

Voici un exemple de sortie. Chaque valeur représente le pourcentage d'utilisation moyenne de l'UC pour les instances EC2 exécutant l'AMI spécifiée.

```
{
    "Datapoints": [
        {
            "Timestamp": "2016-10-10T07:00:00Z",
            "Average": 0.041000000000000009,
            "Unit": "Percent"
        },
        {
```

```
            "Timestamp": "2016-10-10T14:00:00Z",
            "Average": 0.079579831932773085,
            "Unit": "Percent"
        },
        {
            "Timestamp": "2016-10-10T06:00:00Z",
            "Average": 0.036000000000000011,
            "Unit": "Percent"
        },
        ...
    ],
    "Label": "CPUUtilization"
}
```

Représenter graphiquement les métriques de vos instances

Après avoir lancé une instance, vous pouvez ouvrir la console Amazon EC2 et afficher les graphiques de surveillance d'une instance dans l'onglet Surveillance. Chaque graphique s'appuie sur l'une des métriques Amazon EC2 disponibles.

Les graphiques suivants sont disponibles :

- Utilisation moyenne de l'UC (pourcentage)
- Lectures moyennes sur disque (octets)
- Ecritures moyennes sur disque (octets)
- Nombre maximal entrées réseau (octets)
- Nombre maximal sorties réseau (octets)
- Récapitulatif des opérations de lecture sur disque (nombre)
- Récapitulatif des opérations d'écriture sur disque (nombre)
- Récapitulatif des statuts (quels qu'il soient)
- Récapitulatif des statuts d'instance (nombre)
- Récapitulatif des statuts système (nombre)

Pour plus d'informations sur les métriques et les données qu'elles leur fournissent, consultez Répertorier les métriques CloudWatch disponibles pour vos instances (p. 764).

Graphique de métriques à l'aide de la console CloudWatch

Vous pouvez également utiliser la console CloudWatch pour représenter graphiquement les données métriques générées par Amazon EC2 et d'autres services AWS. Pour plus d'informations, consultez la section Graphique de métriques du Guide de l'utilisateur Amazon CloudWatch.

Créer une alarme CloudWatch pour une instance

Vous pouvez créer une alarme CloudWatch qui surveille les métriques CloudWatch pour l'une de vos instances. CloudWatch vous envoie automatiquement une notification quand la métrique atteint un seuil que vous spécifiez. Vous pouvez créer une alarme CloudWatch à l'aide de la console Amazon EC2 ou en utilisant les options plus avancées de la console CloudWatch.

Pour créer une alarme à l'aide de la console CloudWatch

Pour obtenir des exemples, consultez la section Création d'alarmes Amazon CloudWatch du Guide de l'utilisateur Amazon CloudWatch.

New console

Pour créer une alarme à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance et choisissez Actions, Surveillance, Gérer les alarmes CloudWatch.
4. Sur la page détaillée Gérer les alarmes CloudWatch, sous Ajouter ou modifier une alarme, sélectionnez Créer une nouvelle alarme.
5. Pour Notification d'alarme, choisissez si vous souhaitez activer ou désactiver l'option pour configurer les notifications Amazon Simple Notification Service (Amazon SNS). Entrez une rubrique Amazon SNS existante ou entrez un nom pour créer une nouvelle rubrique.
6. Pour Action d'alarme, choisissez d'activer ou de désactiver la bascule pour spécifier une action à effectuer lorsque l'alarme est déclenchée. Sélectionnez une action dans la liste déroulante.
7. Pour Seuils d'alarme, sélectionnez la métrique et les critères de l'alarme. Par exemple, vous pouvez laisser les paramètres par défaut pour Regrouper les échantillons par (Moyenne) et Type de données à échantillonner (utilisation de l'UC). Pour Alarme quand, choisissez **>=** et entrez **`0.80`**. Pour Période consécutive, entrez **`1`**. Pour Période, sélectionnez 5 minutes.
8. (Facultatif) Pour Exemple de données de métrique, choisissez Ajouter au tableau de bord.
9. Sélectionnez Créer.

Old console

Pour créer une alarme à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Dans l'onglet Surveillance situé en bas de la page, choisissez Créer une alarme. Autre possibilité : dans la liste déroulante Actions, choisissez Surveillance CloudWatch, Ajouter/Modifier une alarme.
5. Dans la boîte de dialogue Créer une alarme, exécutez l'une des actions suivantes :
   a. Choisissez create topic. Attribuez un nom à la rubrique SNS dans Envoyer une notification à :. Saisissez une ou plusieurs adresses e-mail auxquelles envoyer les notifications dans Avec ces destinataires.
   b. Spécifiez la métrique et les critères de la stratégie. Par exemple, vous pouvez conserver les paramètres par défaut pour Whenever (utilisation moyenne du processeur). Pour est, choisissez `>=` et saisissez `80` %. Pour Pendant au moins, saisissez `1` période consécutive de `5 Minutes`.
   c. Sélectionnez Create Alarm.

Créer des alarmes qui arrêtent, finissent, redémarrent ou récupèrent une instance

A l'aide des actions d'alarme Amazon CloudWatch, vous pouvez créer des alarmes qui arrêtent, mettent hors service, redémarrent ou récupèrent vos instances automatiquement. Vous pouvez utiliser les actions d'arrêt ou de terminaison pour vous permettre d'économiser de l'argent quand vous n'avez plus besoin qu'une instance s'exécute. De même, les actions de redémarrage et de récupération vous permettent de redémarrer automatiquement ces instances ou de les récupérer sur un nouveau matériel en cas de déficience du nouveau matériel.

Le rôle lié à un service `AWSServiceRoleForCloudWatchEvents` permet à AWS d'exécuter des actions d'alarme en votre nom. La première fois que vous créez une alarme dans AWS Management Console, l'interface de ligne de commande IAM ou l'API IAM, CloudWatch crée automatiquement le rôle lié à un service.

Il existe un certain nombre de scénarios dans lesquels vous pourriez vouloir arrêter ou terminer automatiquement votre instance. Par exemple, vous pourriez avoir des instances dédiées aux tâches de traitement différé de la paie ou de calcul scientifique qui s'exécutent pendant une durée, puis achèvent leur travail. Plutôt que de laisser ces instances demeurer inactives (et d'accumuler les frais), vous pouvez les arrêter ou les terminer, ce qui peut vous aider à économiser de l'argent. La principale différence entre l'utilisation des actions d'alarme « stop » et « terminate » est que vous pouvez facilement démarrer une instance arrêtée si vous devez l'exécuter à nouveau ultérieurement, et que vous pouvez conserver les

mêmes ID d'instance et volume racine. Cependant, vous ne pouvez pas démarrer une instance résiliée. Vous devez à la place lancer une nouvelle instance.

Vous pouvez ajouter les actions d'arrêt, de terminaison, de redémarrage ou de récupération à toute alarme définie sur une métrique Amazon EC2 par instance, y compris les métriques de surveillance élémentaire ou détaillée fournies par Amazon CloudWatch (dans l'espace de noms `AWS/EC2`), ainsi que toute métrique personnalisée incluant la dimension `InstanceId`, aussi longtemps que sa valeur se réfère à une instance Amazon EC2 valide en cours d'exécution.

Prise en charge de la console

Vous pouvez créer des alarmes à l'aide de la console Amazon EC2 ou de la console CloudWatch. Les procédures décrites dans cette documentation utilisent la console Amazon EC2. Pour voir les procédures qui utilisent la console CloudWatch, consultez Création d'alarmes qui arrêtent, mettent hors service, redémarrent ou récupèrent une instance dans le Guide de l'utilisateur Amazon CloudWatch.

Autorisations

Si vous êtes un utilisateur AWS Identity and Access Management (IAM), vous devez disposer des autorisations suivantes pour créer ou modifier une alarme :

- `iam:CreateServiceLinkedRole`, `iam:GetPolicy`, `iam:GetPolicyVersion` et `iam:GetRole` – Pour toutes les alarmes avec des actions Amazon EC2
- `ec2:DescribeInstanceStatus` et `ec2:DescribeInstances` – Pour toutes les alarmes sur des métriques de statut d'instance Amazon EC2
- `ec2:StopInstances` – Pour les alarmes avec des actions d'arrêt
- `ec2:TerminateInstances` – Pour les alarmes avec des actions de fin
- Aucune autorisation spécifique n'est nécessaire pour les alarmes avec des actions de récupération.

Si vous disposez des autorisations en lecture/écriture pour Amazon CloudWatch, mais pas pour Amazon EC2, vous pouvez toujours créer une alarme, mais les actions d'arrêt ou de fin ne seront pas exécutées sur l'instance Amazon EC2. Toutefois, si vous êtes autorisé ultérieurement à utiliser les API Amazon EC2 associées, les actions d'alarme que vous avez créées précédemment sont exécutées. Pour de plus amples informations sur les autorisations IAM, veuillez consulter Autorisations et stratégies dans le IAM Guide de l'utilisateur.

Sommaire

- Ajout d'actions d'arrêt aux alarmes Amazon CloudWatch (p. 786)
- Ajout d'actions de fin aux alarmes Amazon CloudWatch (p. 787)
- Ajout d'actions de redémarrage aux alarmes Amazon CloudWatch (p. 788)
- Ajout d'actions de récupération aux alarmes Amazon CloudWatch (p. 789)
- Utilisation de la console Amazon CloudWatch pour afficher l'historique des alarmes et des actions (p. 791)
- Scénarios d'action d'alarme Amazon CloudWatch (p. 791)

Ajout d'actions d'arrêt aux alarmes Amazon CloudWatch

Vous pouvez créer une alarme qui arrête une instance Amazon EC2 quand un certain seuil a été atteint. Par exemple, vous pouvez exécuter des instances de développement ou de test, et, à l'occasion, oublier de les fermer. Vous pouvez créer une alarme qui est déclenchée quand le pourcentage moyen d'utilisation de l'UC a été inférieur à 10 % pendant 24 heures, indiquant que l'instance est inactive et n'est plus en cours d'utilisation. Vous pouvez ajuster le seuil, la durée et la période en fonction de vos besoins ; de plus, vous

pouvez ajouter une notification Amazon Simple Notification Service (Amazon SNS) de façon à recevoir un courrier électronique quand l'alarme est déclenchée.

Les instances qui utilisent un volume Amazon EBS comme périphérique racine peuvent être arrêtées ou résiliées, tandis que celles qui recourent au stockage d'instance comme périphérique racine peuvent uniquement être résiliées.

Pour créer une alarme afin d'arrêter une instance inactive (console Amazon EC2)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance. Sous l'onglet Surveillance, choisissez Créer une alarme.
4. Dans la boîte de dialogue Créer une alarme, exécutez l'une des actions suivantes :

   a. Pour recevoir un e-mail quand l'alarme est déclenchée, dans la zone Envoyer une notification à, choisissez une rubrique Amazon SNS existante ou cliquez sur créer une rubrique pour en créer une.

      Pour créer une rubrique, attribuez-lui un nom dans Envoyer une notification à, puis entrez les adresses e-mail des destinataires (séparées par une virgule) dans Avec ces destinataires. Après avoir créé l'alarme, vous recevrez un e-mail de confirmation d'abonnement que vous devrez accepter avant de pouvoir recevoir les notifications associées à cette rubrique.
   b. Choisissez Prendre la mesure suivante, choisissez Arrêtez cette instance.
   c. Pour Lorsque, choisissez la statistique que vous voulez utiliser, puis sélectionnez la métrique. Dans cet exemple, choisissez Moyenne et Utilisation de la CPU.
   d. Pour Est, définissez le seuil de la métrique. Pour cet exemple, entrez 10 %.
   e. Pour Pendant au moins, choisissez la période d'évaluation de l'alarme. Pour cet exemple, entrez 24 périodes consécutives d'1 heure.
   f. Pour modifier le nom de l'alarme, entrez un nouveau nom dans Nom de l'alarme. Les noms d'alarme doivent contenir uniquement des caractères ASCII.

      Si vous n'attribuez pas de nom à l'alarme, Amazon CloudWatch en crée un automatiquement.

      Note

      Vous pouvez régler la configuration de l'alarme en fonction de vos propres besoins avant de créer l'alarme, ou pouvez la modifier ultérieurement. Les paramètres de configuration incluent ceux de métrique, de seuil, de durée, d'action et de notification. Cependant, après avoir créé une alarme, vous ne pourrez pas modifier son nom par la suite.
   g. Sélectionnez Create Alarm.

Ajout d'actions de fin aux alarmes Amazon CloudWatch

Vous pouvez créer une alarme qui finit automatiquement une instance EC2 quand un certain seuil a été atteint (aussi longtemps que la protection de fin n'est pas activée pour l'instance). Par exemple, il se peut que vous vouliez finir une instance quand elle a terminé son travail et que vous n'avez pas besoin de l'instance à nouveau. Si vous souhaitez utiliser l'instance par la suite, vous devez arrêter l'instance, et non y mettre fin. Pour de plus amples informations sur l'activation et la désactivation de la protection de résiliation pour une instance, veuillez consulter Activation de la protection contre la résiliation (p. 648).

Pour créer une alarme afin de résilier une instance inactive (console Amazon EC2)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance. Sous l'onglet Surveillance, choisissez Créer une alarme.

4. Dans la boîte de dialogue Créer une alarme, exécutez l'une des actions suivantes :

   a. Pour recevoir un e-mail quand l'alarme est déclenchée, dans la zone Envoyer une notification à, choisissez une rubrique Amazon SNS existante ou cliquez sur créer une rubrique pour en créer une.

      Pour créer une rubrique, attribuez-lui un nom dans Envoyer une notification à, puis entrez les adresses e-mail des destinataires (séparées par une virgule) dans Avec ces destinataires. Après avoir créé l'alarme, vous recevrez un e-mail de confirmation d'abonnement que vous devrez accepter avant de pouvoir recevoir les notifications associées à cette rubrique.

   b. Choisissez Prendre la mesure suivante, puis Résilier cette instance.

   c. Pour Lorsque, sélectionnez une statistique, puis choisissez la métrique. Dans cet exemple, choisissez Moyenne et Utilisation de la CPU.

   d. Pour Est, définissez le seuil de la métrique. Pour cet exemple, entrez 10 %.

   e. Pour Pendant au moins, choisissez la période d'évaluation de l'alarme. Pour cet exemple, entrez 24 périodes consécutives d'1 heure.

   f. Pour modifier le nom de l'alarme, entrez un nouveau nom dans Nom de l'alarme. Les noms d'alarme doivent contenir uniquement des caractères ASCII.

      Si vous n'attribuez pas de nom à l'alarme, Amazon CloudWatch en crée un automatiquement.

      > **Note**
      >
      > Vous pouvez régler la configuration de l'alarme en fonction de vos propres besoins avant de créer l'alarme, ou pouvez la modifier ultérieurement. Les paramètres de configuration incluent ceux de métrique, de seuil, de durée, d'action et de notification. Cependant, après avoir créé une alarme, vous ne pourrez pas modifier son nom par la suite.

   g. Sélectionnez Create Alarm.

Ajout d'actions de redémarrage aux alarmes Amazon CloudWatch

Vous pouvez créer une alarme Amazon CloudWatch qui surveille une instance Amazon EC2 et la redémarre automatiquement. L'action d'alarme de redémarrage est recommandée pour les défaillances de vérification de l'état d'instance (par opposition à l'action d'alarme de récupération, qui convient aux défaillances de la vérification de l'état du système). Le redémarrage d'une instance est similaire à celui d'un système d'exploitation. Dans la plupart des cas, il suffit de quelques minutes pour redémarrer votre instance. Lorsque vous redémarrez une instance, elle reste sur le même hôte physique, ce qui signifie qu'elle conserve son nom DNS public, son adresse IP privée et toutes les données se trouvant sur ses volumes de stockage d'instance.

Le redémarrage d'une instance ne déclenche pas de nouvelle de facturation d'instance (avec frais d'une minute minimum), contrairement à l'arrêt, puis au redémarrage d'une instance. Pour plus d'informations, consultez Redémarrer votre instance dans le Amazon EC2 Guide de l'utilisateur pour les instances Linux.

> **Important**
>
> Pour prévenir toute condition de concurrence entre les actions de redémarrage et de récupération, évitez de définir le même nombre de périodes d'évaluation pour une alarme de redémarrage et une alarme de récupération. Nous vous recommandons de définir des alarmes de redémarrage sur trois périodes d'évaluation d'une minute chacune. Pour plus d'informations, consultez Évaluation d'une alarme dans le Guide de l'utilisateur Amazon CloudWatch.

Pour créer une alarme afin de redémarrer une instance (console Amazon EC2)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance. Sous l'onglet Surveillance, choisissez Créer une alarme.
4. Dans la boîte de dialogue Créer une alarme, exécutez l'une des actions suivantes :

   a. Pour recevoir un e-mail quand l'alarme est déclenchée, dans la zone Envoyer une notification à, choisissez une rubrique Amazon SNS existante ou cliquez sur créer une rubrique pour en créer une.

      Pour créer une rubrique, attribuez-lui un nom dans Envoyer une notification à, puis entrez les adresses e-mail des destinataires (séparées par une virgule) dans Avec ces destinataires. Après avoir créé l'alarme, vous recevrez un e-mail de confirmation d'abonnement que vous devrez accepter avant de pouvoir recevoir les notifications associées à cette rubrique.
   b. Choisissez Prendre la mesure suivante, puis Redémarrez cette instance.
   c. Pour Lorsque, puis Échec du contrôle de statut (instance).
   d. Pour Pendant au moins, choisissez la période d'évaluation de l'alarme. Pour cet exemple, entrez 3 périodes consécutives d'1 minute.
   e. Pour modifier le nom de l'alarme, entrez un nouveau nom dans Nom de l'alarme. Les noms d'alarme doivent contenir uniquement des caractères ASCII.

      Si vous n'attribuez pas de nom à l'alarme, Amazon CloudWatch en crée un automatiquement.
   f. Sélectionnez Create Alarm.

Ajout d'actions de récupération aux alarmes Amazon CloudWatch

Vous pouvez créer une alarme Amazon CloudWatch qui surveille une instance Amazon EC2. Si l'instance est dégradée suite à une défaillance du matériel sous-jacent ou à un problème nécessitant une intervention d'AWS pour sa résolution, vous pouvez créer récupérer automatiquement l'instance. Les instances mises hors service ne peuvent pas être récupérées. Une instance récupérée est identique à l'instance d'origine, y compris pour l'ID d'instance, les adresses IP privées, les adresses IP Elastic et toutes les métadonnées de l'instance.

CloudWatch vous empêche d'ajouter une action de récupération à une alarme figurant sur une instance qui ne prend pas en charge les actions de récupération.

Lorsque l'alarme `StatusCheckFailed_System` est déclenchée et que l'action de récupération est initiée, vous en êtes averti par la rubrique Amazon SNS que vous avez choisie quand vous avez créé l'alarme et associé l'action de récupération. Lors de la récupération d'instance, l'instance est migrée pendant un redémarrage d'instance, et toutes les données en mémoire sont perdues. Lorsque le processus est terminé, les informations sont publiées dans la rubrique SNS que vous avez configurée pour l'alarme. Toutes les personnes abonnées à cette rubrique SNS reçoivent une notification par e-mail qui inclut le statut de la tentative de récupération et les éventuelles instructions supplémentaires. Vous remarquez un redémarrage d'instance sur l'instance récupérée.

L'action de récupération ne peut être utilisée qu'avec `StatusCheckFailed_System`, pas avec `StatusCheckFailed_Instance`.

Les problèmes suivants peuvent entraîner l'échec des contrôles de statut de système :

- Perte de connectivité réseau
- Perte d'alimentation système
- Problèmes logiciels sur un hôte physique
- Problèmes matériels sur un hôte physique ayant un impact sur l'accessibilité du réseau

L'opération de récupération est prise en charge uniquement sur les instances présentant les caractéristiques suivantes :

- Utilisez l'un des types d'instance suivants : A1, C3, C4, C5, C5a, C5n, C6g, Inf1, M3, M4, M5, M5a, M5n, M6g, P3, R3, R4, R5, R5a, R5n, R6g, T2, T3, T3a, X1 ou X1e
- Utilisation de la location d'instance `default` ou `dedicated`
- Utilisation de volumes EBS uniquement (ne pas configurer des volumes de stockage d'instance). Pour plus d'informations, consultez « Récupérez cette instance » est désactivé.

Si votre instance a une adresse IP publique, elle la conserve après la récupération.

> **Important**
>
> Pour prévenir toute condition de concurrence entre les actions de redémarrage et de récupération, évitez de définir le même nombre de périodes d'évaluation pour une alarme de redémarrage et une alarme de récupération. Nous vous recommandons de définir des alarmes de récupération sur deux périodes d'évaluation d'une minute chacune. Pour plus d'informations, consultez Évaluation d'une alarme dans le Guide de l'utilisateur Amazon CloudWatch.

Pour créer une alarme afin de récupérer une instance (console Amazon EC2)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance. Sous l'onglet Surveillance, choisissez Créer une alarme.
4. Dans la boîte de dialogue Créer une alarme, exécutez l'une des actions suivantes :

   a. Pour recevoir un e-mail quand l'alarme est déclenchée, dans la zone Envoyer une notification à, choisissez une rubrique Amazon SNS existante ou cliquez sur créer une rubrique pour en créer une.

      Pour créer une rubrique, attribuez-lui un nom dans Envoyer une notification à, puis entrez les adresses e-mail des destinataires (séparées par une virgule) dans Avec ces destinataires. Après avoir créé l'alarme, vous recevrez un e-mail de confirmation d'abonnement que vous devrez accepter avant de pouvoir recevoir les messages électroniques associés à cette rubrique.

      > **Note**
      >
      > - Les utilisateurs doivent s'abonner à la rubrique SNS spécifiée pour recevoir des notifications par e-mail lorsque l'alarme se déclenche.
      > - L'utilisateur racine du compte AWS reçoit toujours des notifications par e-mail lorsque des actions de récupération d'instance automatique sont exécutées, même si une rubrique SNS n'est pas spécifiée.
      > - L'utilisateur racine du compte AWS reçoit toujours des notifications par e-mail lorsque des actions de récupération d'instance automatique sont exécutées, même s'il n'est pas abonné à la rubrique SNS spécifiée.

   b. Choisissez Prendre la mesure suivante, choisissez Récupérez cette instance.

   c. Pour Lorsque, choisissez Échec du contrôle de statut (système).

   d. Pour Pendant au moins, choisissez la période d'évaluation de l'alarme. Pour cet exemple, entrez 2 périodes consécutives d'1 minute.

   e. Pour modifier le nom de l'alarme, entrez un nouveau nom dans Nom de l'alarme. Les noms d'alarme doivent contenir uniquement des caractères ASCII.

      Si vous n'attribuez pas de nom à l'alarme, Amazon CloudWatch en crée un automatiquement.

   f. Sélectionnez Create Alarm.

Utilisation de la console Amazon CloudWatch pour afficher l'historique des alarmes et des actions

Vous pouvez afficher l'alarme et l'historique des actions historique dans la console Amazon CloudWatch. Amazon CloudWatch conserve l'historique des alarmes et des actions des deux dernières semaines.

Pour afficher l'historique des actions et des alarmes déclenchées (console CloudWatch)

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Dans le panneau de navigation, cliquez sur Alarms.
3. Sélectionnez une alarme.
4. L'onglet Détails affiche la transition d'état la plus récente, ainsi que les valeurs de date et de métrique.
5. Choisissez l'onglet Historique pour afficher les entrées les plus récentes de l'historique.

Scénarios d'action d'alarme Amazon CloudWatch

Vous pouvez utiliser la console Amazon EC2 pour créer des actions d'alarme qui arrêtent ou finissent une instance Amazon EC2 quand certaines conditions sont satisfaites. Dans la capture d'écran suivante de la page de la console où vous avez défini les actions d'alarme, nous avons numéroté les paramètres. Nous avons également numéroté les paramètres des scénarios qui suivent afin de vous aider à créer les actions appropriées.

Scénario 1 : arrêter le développement inactif et tester les instances

Créez une alarme qui arrête une instance utilisée pour le développement ou le test de logiciels quand elle a été inactive pendant au moins une heure.

Paramètre	Value
1	Arrêter
2	Maximum
3	CPUUtilization
4	<=
5	10 %
6	60 minutes
7	1

Scénario 2 : arrêter les instances inactives

Créez une alarme qui arrête une instance et envoie un courrier électronique quand l'instance est inactive depuis 24 heures.

Paramètre	Value
1	Stop and email
2	Moyenne
3	CPUUtilization
4	<=
5	5 %
6	60 minutes
7	24

Scénario 3 : envoyer un e-mail relatif aux serveurs Web ayant un trafic inhabituellement élevé

Créez une alarme qui envoie un courrier électronique quand une instance dépasse 10 Go de trafic réseau sortant par jour.

Paramètre	Value
1	E-mail
2	Somme
3	NetworkOut
4	>
5	10 Go
6	1 jour
7	1

Scénario 4 : arrêter les serveurs Web avec un trafic inhabituellement élevé

Créez un alarme qui arrête une instance et envoie un SMS quand le trafic sortant excède 1 Go par heure.

Paramètre	Value
1	Stop and send SMS
2	Somme
3	NetworkOut
4	>

Paramètre	Value
5	1 Go
6	1 heure
7	1

Scénario 5 : arrêter une instance confrontée à une fuite mémoire

Créez une alarme qui arrête une instance quand l'utilisation de la mémoire atteint ou dépasse 90 %, de telle sorte que les journaux d'application puissent être extraits pour le dépannage.

> **Note**
>
> La métrique MemoryUtilization est une métrique personnalisée. Pour pouvoir utiliser la métrique MemoryUtilization, vous devez installer les scripts Perl pour les instances Linux. Pour plus d'informations, veuillez consulter Supervision des métriques de la mémoire et du disque pour les instances Amazon EC2 Linux.

Paramètre	Value
1	Arrêter
2	Maximum
3	MemoryUtilization
4	>=
5	90 %
6	1 minute
7	1

Scénario 6 : arrêter une instance déficiente

Créez une alarme qui arrête une instance après qu'elle a échoué à trois contrôles de statut consécutifs (exécutés à 5 minutes d'intervalle).

Paramètre	Value
1	Arrêter
2	Moyenne
3	StatusCheckFailed_System
4	>=
5	1
6	15 minutes
7	1

Scénario 7 : terminer les instances quand les tâches de traitement par batch sont terminés

Créez une alarme qui finit une instance exécutant des traitements par batch quand elle n'envoie plus de données de résultat.

Paramètre	Value
1	Terminer
2	Maximum
3	NetworkOut
4	<=
5	100,000 bytes
6	5 minutes
7	1

Automatisation d'Amazon EC2 avec EventBridge

Amazon EventBridge vous permet d'automatiser vos services AWS et de répondre automatiquement à des événements système tels que des problèmes de disponibilité d'application ou des modifications de ressource. Les événements des services AWS sont fournis à EventBridge presque en temps réel. Vous pouvez écrire des règles simples pour indiquer quels événements vous intéressent et les actions automatisées à effectuer quand un événement correspond à une règle. Les actions pouvant être déclenchées automatiquement sont les suivantes :

- Appel d'une fonction AWS Lambda
- Appel de la fonctionnalité Exécuter la commande d'Amazon EC2
- Relais de l'événement à Amazon Kinesis Data Streams
- Activation d'une machine d'état AWS Step Functions
- Notification d'une rubrique Amazon SNS ou d'une file d'attente Amazon SQS

Voici quelques exemples d'utilisation de EventBridge avec Amazon EC2 :

- Activation d'une fonction Lambda à chaque démarrage d'une nouvelle instance Amazon EC2.
- Notification d'une rubrique Amazon SNS quand un volume Amazon EBS est créé ou modifié.
- Envoi d'une commande à une ou plusieurs instances Amazon EC2 à l'aide de la commande Amazon EC2 Exécuter chaque fois qu'un certain événement se produit dans un autre service AWS.

Pour plus d'informations, consultez le manuel Guide de l'utilisateur Amazon EventBridge.

Surveillance des métriques de la mémoire et du disque pour les instances Amazon EC2 Linux

Vous pouvez utiliser Amazon CloudWatch pour collecter des métriques et des journaux à partir des systèmes d’exploitation pour vos instances EC2.

> **Important**
>
> Les scripts de surveillance sont obsolètes. Nous vous recommandons d'utiliser l'agent CloudWatch pour collecter des métriques et des fichiers journaux. Nous fournissons ces informations sur les scripts de surveillance pour les clients qui migrent encore des scripts de surveillance obsolètes vers l'agent CloudWatch.

Collecte de métriques à l'aide de l'agent CloudWatch

Vous pouvez utiliser l'agent CloudWatch pour collecter des métriques système et des fichiers journaux à partir d'instances Amazon EC2 et de serveurs sur site. L'agent prend en charge Windows Server et Linux, et vous permet de sélectionner les métriques à collecter, notamment des métriques de sous-ressource, par exemple, cœur par UC. Nous vous recommandons d'utiliser l'agent pour collecter des métriques et des fichiers journaux plutôt que d'utiliser les scripts de surveillance. Pour plus d'informations, consultez Collecte de métriques à partir d'instances Amazon EC2 et de serveurs sur site avec l'agent CloudWatch dans le Guide de l'utilisateur Amazon CloudWatch.

Obsolète : Collecte de métriques à l'aide des scripts de surveillance CloudWatch

> **Important**
>
> Nous fournissons des informations sur les scripts de surveillance pour les clients qui n'ont pas encore migré des scripts de surveillance obsolètes vers l'agent CloudWatch.

Les scripts de surveillance montrent comment produire et consommer des métriques personnalisées pour Amazon CloudWatch. Ces exemples de scripts Perl comportent un exemple entièrement fonctionnel qui indique les métriques d'utilisation de la mémoire, des échanges et de l'espace sur le disque pour une instance Linux.

Les coûts d'utilisation Amazon CloudWatch standard pour les métriques personnalisées s'appliquent à votre utilisation de ces scripts. Pour plus d'informations, consultez la page de tarification Amazon CloudWatch.

Sommaire

- Systèmes pris en charge (p. 796)
- Autorisations requises (p. 797)
- Installation des packages obligatoires (p. 797)
- Installation des scripts de surveillance (p. 798)
- mon-put-instance-data.pl (p. 799)
- mon-get-instance-stats.pl (p. 802)
- Visualisation de vos métriques personnalisées dans la console (p. 803)
- Dépannage (p. 804)

Systèmes pris en charge

Les scripts de surveillance ont été testés sur des instances utilisant les systèmes suivants. L'utilisation des scripts de surveillance sur un autre système d'exploitation n'est pas prise en charge.

- Amazon Linux 2
- Amazon Linux AMI 2014.09.2 ou version suivante

- Red Hat Enterprise Linux 6.9 et 7.4
- SUSE Linux Enterprise Server 12
- Ubuntu Server 14.04 et 16.04

Autorisations requises

Assurez-vous que les scripts disposent de l'autorisation d'appeler les actions suivantes en associant un rôle IAM à votre instance :

- cloudwatch:PutMetricData
- cloudwatch:GetMetricStatistics
- cloudwatch:ListMetrics
- ec2:DescribeTags

Pour plus d'informations, consultez Utilisation des rôles IAM (p. 1037).

Installation des packages obligatoires

Avec certaines versions de Linux, vous devez installer des modules Perl supplémentaires avant d'utiliser les scripts de surveillance.

Pour installer les packages obligatoires sur l'AMI Amazon Linux 2 et Amazon Linux

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. A l'invite de commande, installez les packages comme suit :

```
sudo yum install -y perl-Switch perl-DateTime perl-Sys-Syslog perl-LWP-Protocol-https
 perl-Digest-SHA.x86_64
```

Pour installer les packages requis sur Ubuntu

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. A l'invite de commande, installez les packages comme suit :

```
sudo apt-get update
sudo apt-get install unzip
sudo apt-get install libwww-perl libdatetime-perl
```

Pour installer les packages obligatoires sur Red Hat Enterprise Linux 7

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. A l'invite de commande, installez les packages comme suit :

```
sudo yum install perl-Switch perl-DateTime perl-Sys-Syslog perl-LWP-Protocol-https
 perl-Digest-SHA --enablerepo="rhui-REGION-rhel-server-optional" -y
sudo yum install zip unzip
```

Pour installer les packages requis sur Red Hat Enterprise Linux 6.9

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. A l'invite de commande, installez les packages comme suit :

```
sudo yum install perl-DateTime perl-CPAN perl-Net-SSLeay perl-IO-Socket-SSL perl-
Digest-SHA gcc -y
sudo yum install zip unzip
```

3. Exécutez CPAN en tant qu'utilisateur élevé :

```
sudo cpan
```

   Appuyez sur ENTER via les invites jusqu'à ce que l'invite suivante s'affiche :

```
cpan[1]>
```

4. Lors de l'invite CPAN, exécutez chacune des invites suivantes : exécutez une commande et son installation et lorsque vous retournez à l'invite CPAN, exécutez la commande suivante. Appuyez sur ENTER lorsque vous êtes invité à continuer au cours du processus :

```
cpan[1]> install YAML
cpan[2]> install LWP::Protocol::https
cpan[3]> install Sys::Syslog
cpan[4]> install Switch
```

Pour installer les packages requis sur SUSE

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Sur les serveurs exécutant SUSE Linux Enterprise Server 12, vous aurez peut-être besoin de télécharger le package `perl-Switch`. Vous pouvez télécharger et installer ce package avec les commandes suivantes :

```
wget http://download.opensuse.org/repositories/devel:/languages:/perl/SLE_12_SP3/
noarch/perl-Switch-2.17-32.1.noarch.rpm
sudo rpm -i perl-Switch-2.17-32.1.noarch.rpm
```

3. Installez les packages obligatoires comme suit :

```
sudo zypper install perl-Switch perl-DateTime
sudo zypper install -y "perl(LWP::Protocol::https)"
```

Installation des scripts de surveillance

Les étapes suivantes vous montrent comment télécharger, décompresser et configurer les scripts de surveillance CloudWatch sur une instance EC2 Linux.

Pour télécharger, installer et configurer les scripts de surveillance

1. A une invite de commande, placez-vous dans le dossier où vous voulez stocker les scripts de surveillance, puis tapez la commande suivante pour les télécharger :

```
curl https://aws-cloudwatch.s3.amazonaws.com/downloads/
CloudWatchMonitoringScripts-1.2.2.zip -O
```

2. Exécutez les commandes suivantes pour installer les scripts de surveillance que vous avez téléchargés :

```
unzip CloudWatchMonitoringScripts-1.2.2.zip && \
rm CloudWatchMonitoringScripts-1.2.2.zip && \
cd aws-scripts-mon
```

Le package pour les scripts de surveillance contient les fichiers suivants :

- CloudWatchClient.pm – Module Perl partagé qui simplifie l'appel de Amazon CloudWatch à partir d'autres scripts.
- mon-put-instance-data.pl – Collecte des métriques du système sur une instance Amazon EC2 (utilisation de la mémoire, des échanges et de l'espace sur le disque) et les envoie vers Amazon CloudWatch.
- mon-get-instance-stats.pl – Envoie des requêtes à Amazon CloudWatch et affiche les statistiques d'utilisation les plus récentes pour l'instance EC2 sur laquelle ce script est exécuté.
- awscreds.template – Modèle de fichier pour les informations d'identification AWS qui stocke votre ID de clé d'accès rapide et votre clé d'accès secrète.
- LICENSE.txt – Fichier texte contenant la licence Apache 2.0.
- NOTICE.txt – Mention du droit d'auteur.

mon-put-instance-data.pl

Ce script collecte les données d'utilisation de la mémoire, des échanges et de l'espace sur le disque en ce qui concerne le système actuel. Il effectue ensuite un appel distant de Amazon CloudWatch pour présenter les données collectées en tant que métriques personnalisées.

Options

Nom	Description
`--mem-util`	Collecte et envoie les métriques MemoryUtilization en pourcentages. Cette métrique comptabilise la mémoire allouée par les applications et le système d'exploitation telle qu'elle est utilisée, et inclut également la mémoire cache et tampon utilisée si vous spécifiez l'option `--mem-used-incl-cache-buff`.
`--mem-used`	Collecte et envoie les métriques MemoryUsed, présentées en mégaoctets. Cette métrique comptabilise la mémoire allouée par les applications et le système d'exploitation telle qu'elle est utilisée, et inclut également la mémoire cache et tampon utilisée si vous spécifiez l'option `--mem-used-incl-cache-buff`.
`--mem-used-incl-cache-buff`	Si vous incluez cette option, la mémoire actuellement utilisée pour le cache et les tampons est comptabilisée comme « utilisée » lorsque les métriques sont enregistrées pour `--mem-util`, `--mem-used` et `--mem-avail`.
`--mem-avail`	Collecte et envoie les métriques MemoryAvailable, présentées en mégaoctets. Cette métrique comptabilise la mémoire allouée par les applications et le système d'exploitation telle qu'elle est utilisée,

Nom	Description
	et inclut également la mémoire cache et tampon utilisée si vous spécifiez l'option `--mem-used-incl-cache-buff`.
`--swap-util`	Collecte et envoie les métriques SwapUtilization, présentées en pourcentages.
`--swap-used`	Collecte et envoie les métriques SwapUsed, présentées en mégaoctets.
`--disk-path=PATH`	Sélectionne le disque sur lequel effectuer le rapport. PATH peut spécifier un point de montage ou n'importe quel fichier situé sur un point de montage pour le système de fichiers qui doit être présenté. Pour sélectionner plusieurs disques, spécifiez un `--disk-path=PATH` pour chacun d'entre eux. Pour sélectionner un disque pour les systèmes de fichiers montés sur `/` et `/home`, utilisez les paramètres suivants : `--disk-path=/ --disk-path=/home`
`--disk-space-util`	Collecte et envoie la métrique DiskSpaceUtilization pour les disques sélectionnés. Cette métrique est présentée en pourcentages. Notez que les métriques d'utilisation des disques calculées par ce script diffèrent des valeurs calculées par la commande df -k -l. Si vous trouvez les valeurs de la commande df -k -l plus utiles, vous pouvez modifier les calculs dans le script.
`--disk-space-used`	Collecte et envoie la métrique DiskSpaceUsed pour les disques sélectionnés. Cette métrique est présentée par défaut en gigaoctets. En raison de l'espace réservé sur le disque dans les systèmes d'exploitation Linux, l'espace utilisé et l'espace disponible sur le disque ne s'ajoutent peut-être pas correctement à la quantité de l'espace total du disque.
`--disk-space-avail`	Collecte et envoie la métrique DiskSpaceAvailable pour les disques sélectionnés. Cette métrique est présentée en gigaoctets. En raison de l'espace réservé sur le disque dans les systèmes d'exploitation Linux, l'espace utilisé et l'espace disponible sur le disque ne s'ajoutent peut-être pas correctement à la quantité de l'espace total du disque.
`--memory-units=UNITS`	Spécifie les unités avec lesquelles il faut indiquer l'utilisation de la mémoire. Si rien n'est spécifié, la mémoire est présentée en mégaoctets. Parmi les UNITES, on compte les suivantes : octets, kilo-octets, mégaoctets, gigaoctets.
`--disk-space-units=UNITS`	Spécifie les unités avec lesquelles il faut indiquer l'utilisation de l'espace sur le disque. Si rien n'est spécifié, l'espace sur le disque est présenté en gigaoctets. Parmi les UNITES, on compte les suivantes : octets, kilo-octets, mégaoctets, gigaoctets.

Nom	Description
`--aws-credential-file=PATH`	Indique l'emplacement du fichier contenant les informations d'identification AWS. Ce paramètre ne peut pas être utilisé avec les paramètres `--aws-access-key-id` et `--aws-secret-key`.
`--aws-access-key-id=VALUE`	Spécifie l'ID de clé d'accès rapide AWS à utiliser pour identifier le mandataire. Doit être utilisé avec l'option `--aws-secret-key`. N'utilisez pas cette option avec le paramètre `--aws-credential-file`.
`--aws-secret-key=VALUE`	Spécifie la clé d'accès secrète AWS à utiliser pour assigner la demande vers CloudWatch. Doit être utilisé avec l'option `--aws-access-key-id`. N'utilisez pas cette option avec le paramètre `--aws-credential-file`.
`--aws-iam-role=VALUE`	Spécifie le rôle IAM utilisé pour fournir des informations d'identification AWS. La valeur `=VALUE` est obligatoire. Si aucune information d'identification n'est spécifié, le rôle IAM par défaut associé à l'instance EC2 est appliqué. Seul un rôle IAM peut être utilisé. Si aucun rôle IAM n'est trouvé ou si plus d'un rôle IAM est trouvé, le script retournera une erreur. N'utilisez pas cette option avec les paramètres `--aws-credential-file`, `--aws-access-key-id` ou `--aws-secret-key`.
`--aggregated[=only]`	Ajoute des métriques regroupées pour le type d'instance, l'ID d'AMI et en règle générale pour la région. La valeur `=only` est optionnelle. Si cela est indiqué, le script indique uniquement des métriques regroupées.
`--auto-scaling[=only]`	Ajoute des métriques regroupées pour le groupe Auto Scaling. La valeur `=only` est optionnelle. Si cela est indiqué, le script indique uniquement des métriques Auto Scaling. La stratégie IAM associée au compte ou au rôle IAM utilisant les scripts doit avoir les autorisations d'appeler l'action EC2 DescribeTags.
`--verify`	Effectue un test du script qui collecte les métriques, prépare une demande HTTP complète, mais n'appelle pas réellement CloudWatch pour présenter les données. Cette option vérifie également que les informations d'identification sont fournies. Lorsqu'elle est exécutée en mode commentaire, cette option produit les métriques qui seront envoyées vers CloudWatch.
`--from-cron`	Utilisez cette option lorsque vous appelez le script à partir de Cron. Lorsque cette option est utilisée, tous les résultats du diagnostic sont supprimés, mais les messages d'erreur sont envoyés au journal du système local du compte d'utilisateur.
`--verbose`	Affiche des informations détaillées sur ce que fait le script.
`--help`	Affiche les informations d'utilisation.
`--version`	Affiche le numéro de version du script.

Exemples

Les exemples suivants partent du principe que vous avez fourni un rôle IAM ou un fichier `awscreds.conf`. Sinon, vous devez fournir les informations d'identification utilisateur à l'aide des paramètres `--aws-access-key-id` et `--aws-secret-key` pour ces commandes.

L'exemple suivant exécute un test simple sans publier les données dans CloudWatch.

```
./mon-put-instance-data.pl --mem-util --verify --verbose
```

L'exemple suivant collecte toutes les métriques de mémoire disponibles et les envoie vers CloudWatch, en comptabilisant la mémoire cache et tampon comme utilisée

```
./mon-put-instance-data.pl --mem-used-incl-cache-buff --mem-util --mem-used --mem-avail
```

L'exemple suivant collecte les métriques regroupées pour un groupe Auto Scaling et les envoyer vers Amazon CloudWatch sans présenter de métriques d'instance individuelles

```
./mon-put-instance-data.pl --mem-util --mem-used --mem-avail --auto-scaling=only
```

L'exemple suivant collecte des métriques regroupées pour un type d'instance, un ID d'AMI et une région, et les envoyer vers Amazon CloudWatch sans présenter des métriques d'instance individuelles

```
./mon-put-instance-data.pl --mem-util --mem-used --mem-avail --aggregated=only
```

Pour établir une planification cron pour des métriques signalées à CloudWatch, commencez par modifier le crontab à l'aide de la commande crontab -e. Ajoutez la commande suivante pour indiquer l'utilisation de la mémoire et de l'espace sur le disque à CloudWatch toutes les cinq minutes :

```
*/5 * * * * ~/aws-scripts-mon/mon-put-instance-data.pl --mem-used-incl-cache-buff --mem-
util --disk-space-util --disk-path=/ --from-cron
```

Si le script rencontre une erreur, il écrit le message d'erreur dans le journal du système.

mon-get-instance-stats.pl

Ces requêtes de script CloudWatch pour les statistiques concernant les métriques de la mémoire, des échanges et de l'espace sur le disque dans un intervalle de temps donné à l'aide du nombre des heures les plus récentes. Ces données sont fournies pour l'instance Amazon EC2 sur laquelle ce script est exécuté.

Options

Nom	Description
`--recent-hours=N`	Spécifie le nombre d'heures récentes sur lesquelles établir un rapport, comme ce qui est représenté par `N` où N est un nombre entier.
`--aws-credential-file=PATH`	Indique l'emplacement du fichier contenant les informations d'identification AWS.
`--aws-access-key-id=VALUE`	Spécifie l'ID de clé d'accès rapide AWS à utiliser pour identifier le mandataire. Doit être utilisé avec l'option `--aws-secret-key`. N'utilisez pas cette option avec `--aws-credential-file`.

Nom	Description
`--aws-secret-key=VALUE`	Spécifie la clé d'accès secrète AWS à utiliser pour assigner la demande vers CloudWatch. Doit être utilisé avec l'option `--aws-access-key-id`. N'utilisez pas cette option avec `--aws-credential-file`.
`--aws-iam-role=VALUE`	Spécifie le rôle IAM utilisé pour fournir des informations d'identification AWS. La valeur `=VALUE` est obligatoire. Si aucune information d'identification n'est spécifié, le rôle IAM par défaut associé à l'instance EC2 est appliqué. Seul un rôle IAM peut être utilisé. Si aucun rôle IAM n'est trouvé ou si plus d'un rôle IAM est trouvé, le script retournera une erreur. N'utilisez pas cette option avec les paramètres `--aws-credential-file`, `--aws-access-key-id` ou `--aws-secret-key`.
`--verify`	Effectue un test du script. Cette option vérifie également que les informations d'identification sont fournies.
`--verbose`	Affiche des informations détaillées sur ce que fait le script.
`--help`	Affiche les informations d'utilisation.
`--version`	Affiche le numéro de version du script.

Exemple

Pour obtenir les statistiques d'utilisation au cours des 12 dernières heures, exécutez la commande suivante :

```
./mon-get-instance-stats.pl --recent-hours=12
```

Voici un exemple de réponse :

```
Instance metric statistics for the last 12 hours.

CPU Utilization
    Average: 1.06%, Minimum: 0.00%, Maximum: 15.22%

Memory Utilization
    Average: 6.84%, Minimum: 6.82%, Maximum: 6.89%

Swap Utilization
    Average: N/A, Minimum: N/A, Maximum: N/A

Disk Space Utilization on /dev/xvda1 mounted as /
    Average: 9.69%, Minimum: 9.69%, Maximum: 9.69%
```

Visualisation de vos métriques personnalisées dans la console

Après avoir exécuté avec succès le script `mon-put-instance-data.pl`, vous pouvez afficher vos métriques personnalisées dans la console Amazon CloudWatch.

Pour consulter les métriques personnalisées

1. Exécutez `mon-put-instance-data.pl` comme décrit précédemment.

2. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
3. Sélectionnez View Metrics (Afficher les métriques).
4. Pour Viewing (Affichage), vos métriques personnalisées publiées par le script sont affichées avec le préfixe `System/Linux`.

Dépannage

Le module CloudWatchClient.pm met en cache localement les métadonnées de l'instance. Si vous créez une AMI à partir d'une instance où vous avez exécuté les scripts de surveillance, n'importe quelle instance lancée à partir de l'AMI dans le cache TTL (par défaut : six heures, 24 heures pour les groupes Auto Scaling) émettra des métriques à l'aide de l'ID d'instance de l'instance d'origine. Après la période de cache TTL, le script récupère des données actualisées et les scripts de surveillance utilisent l'ID d'instance de l'instance actuelle. Pour corriger immédiatement cela, supprimez les données mises en cache à l'aide de la commande suivante :

```
rm /var/tmp/aws-mon/instance-id
```

Journalisation des appels d'API Amazon EC2 et Amazon EBS avec AWS CloudTrail

Amazon EC2 et Amazon EBS sont intégrés à AWS CloudTrail, un service qui fournit un enregistrement des actions réalisées par un utilisateur, un rôle ou un service AWS dans Amazon EC2 et Amazon EBS. CloudTrail capture tous les appels d'API pour Amazon EC2 et Amazon EBS sous la forme d'événements, y compris les appels passés depuis la console et les appels de code vers les API. Si vous créez un journal d'activité, vous pouvez activer la diffusion en continu des événements CloudTrail sur un compartiment Amazon S3, y compris les événements pour Amazon EC2 et Amazon EBS. Si vous ne configurez pas de journal de suivi, vous pouvez toujours afficher les événements les plus récents dans la console CloudTrail dans Event history (Historique des événements). À l'aide des informations collectées par CloudTrail, vous pouvez déterminer la demande qui a été envoyée à Amazon EC2 et Amazon EBS, l'adresse IP à partir de laquelle la demande a été effectuée, l'auteur de la demande, la date de la demande, ainsi que d'autres informations.

Pour en savoir plus sur CloudTrail, consultez AWS CloudTrail User Guide.

Informations sur Amazon EC2 et Amazon EBS dans CloudTrail

CloudTrail est activé sur votre compte AWS lorsque vous créez le compte. Lorsqu'une activité a lieu dans Amazon EC2 et Amazon EBS, elle est enregistrée dans un événement CloudTrail avec d'autres événements de service AWS dans Event history (Historique des événements). Vous pouvez afficher, rechercher et télécharger les événements récents dans votre compte AWS. Pour plus d'informations, consultez Affichage des événements avec l'historique des événements CloudTrail .

Pour obtenir un enregistrement continu des événements dans votre compte AWS, y compris les événements pour Amazon EC2 et Amazon EBS, créez un journal d'activité. Un journal de suivi permet à CloudTrail de livrer les fichiers journaux dans un compartiment Amazon S3. Par défaut, lorsque vous créez un journal d'activité dans la console, il s'applique à toutes les régions. Le journal de suivi consigne les événements de toutes les régions dans la partition AWS et livre les fichiers journaux dans le compartiment Amazon S3 de votre choix. En outre, vous pouvez configurer d'autres services AWS pour analyser plus en profondeur les données d'événement collectées dans les journaux CloudTrail et agir sur celles-ci. Pour plus d'informations, consultez :

- Présentation de la création d'un journal de suivi
- Intégrations et services pris en charge par CloudTrail
- Configuration des Notifications de Amazon SNS pour CloudTrail
- Réception de fichiers journaux CloudTrail de plusieurs régions et Réception de fichiers journaux CloudTrail de plusieurs comptes

Toutes les actions Amazon EC2 et les actions de gestion Amazon EBS sont consignées par CloudTrail et documentées dans l'Amazon EC2 API Reference. Par exemple, les appels aux actions RunInstances, DescribeInstances ou CreateImage génèrent des entrées dans les fichiers journaux CloudTrail.

Chaque événement ou entrée du journal contient des informations sur la personne qui a généré la demande. Les informations relatives à l'identité permettent de déterminer les éléments suivants :

- Si la demande a été effectuée avec les informations d'identification utilisateur racine ou IAM.
- Si la demande a été effectuée avec des informations d'identification de sécurité temporaires pour un rôle ou un utilisateur fédéré.
- Si la requête a été effectuée par un autre service AWS.

Pour plus d'informations, consultez la section Élément userIdentity CloudTrail.

Compréhension des entrées du fichier journal Amazon EC2 et Amazon EBS

Un journal de suivi est une configuration qui active la livraison d'événements en tant que fichiers journaux à un compartiment Amazon S3 que vous spécifiez. Les fichiers journaux CloudTrail contiennent une ou plusieurs entrées de journal. Un événement représente une demande individuelle à partir d'une source quelconque et comprend des informations sur l'action demandée, la date et l'heure, les paramètres de la demande, etc. Les fichiers journaux CloudTrail ne sont pas des séries ordonnées retraçant les appels d'API publics. Ils ne suivent aucun ordre précis.

L'enregistrement de fichier journal suivant montre qu'un utilisateur a résilié une instance.

```
{
   "Records":[
      {
         "eventVersion":"1.03",
         "userIdentity":{
            "type":"Root",
            "principalId":"123456789012",
            "arn":"arn:aws:iam::123456789012:root",
            "accountId":"123456789012",
            "accessKeyId":"AKIAIOSFODNN7EXAMPLE",
            "userName":"user"
         },
         "eventTime":"2016-05-20T08:27:45Z",
         "eventSource":"ec2.amazonaws.com",
         "eventName":"TerminateInstances",
         "awsRegion":"us-west-2",
         "sourceIPAddress":"198.51.100.1",
         "userAgent":"aws-cli/1.10.10 Python/2.7.9 Windows/7botocore/1.4.1",
         "requestParameters":{
            "instancesSet":{
               "items":[{
                  "instanceId":"i-1a2b3c4d"
               }]
            }
```

```
        },
        "responseElements":{
           "instancesSet":{
              "items":[{
                 "instanceId":"i-1a2b3c4d",
                 "currentState":{
                    "code":32,
                    "name":"shutting-down"
                 },
                 "previousState":{
                    "code":16,
                    "name":"running"
                 }
              }]
           }
        },
        "requestID":"be112233-1ba5-4ae0-8e2b-1c302EXAMPLE",
        "eventID":"6e12345-2a4e-417c-aa78-7594fEXAMPLE",
        "eventType":"AwsApiCall",
        "recipientAccountId":"123456789012"
     }
   ]
}
```

Utilisation de AWS CloudTrail pour auditer les utilisateurs qui se connectent via EC2 Instance Connect

Utilisez AWS CloudTrail pour auditer les utilisateurs qui se connectent à vos instances via EC2 Instance Connect.

Pour auditer l'activité SSH via EC2 Instance Connect à l'aide de la console AWS CloudTrail

1. Ouvrez la console AWS CloudTrail à l'adresse https://console.aws.amazon.com/cloudtrail/.
2. Vérifiez que vous êtes dans la région correcte.
3. Dans le volet de navigation, sélectionnez Event history (Historique des événements).
4. Pour Filter (Filtre), choisissez Event source (Source de l'événement), ec2-instance-connect.amazonaws.com.
5. (Facultatif) Pour Time range (Plage de temps), sélectionnez une plage de temps.
6. Choisissez l'icône Refresh events (Actualiser les événements).
7. La page affiche les événements correspondant aux appels d'API `SendSSHPublicKey`. Développez un événement à l'aide de la flèche pour afficher des détails supplémentaires, comme le nom d'utilisateur et la clé d'accès AWS qui ont été utilisés pour établir la connexion SSH, ainsi que l'adresse IP source.
8. Pour afficher toutes les informations sur l'événement au format JSON, choisissez Afficher l'événement. Le champ requestParameters contient l'ID de l'instance de destination, le nom d'utilisateur OS et la clé publique qui ont été utilisés pour établir la connexion SSH.

```
{
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "ABCDEFGONGNOMOOCB6XYTQEXAMPLE",
        "arn": "arn:aws:iam::1234567890120:user/IAM-friendly-name",
        "accountId": "123456789012",
        "accessKeyId": "ABCDEFGUKZHNAW4OSN2AEXAMPLE",
        "userName": "IAM-friendly-name",
```

```
        "sessionContext": {
            "attributes": {
                "mfaAuthenticated": "false",
                "creationDate": "2018-09-21T21:37:58Z"}
        }
    },
    "eventTime": "2018-09-21T21:38:00Z",
    "eventSource": "ec2-instance-connect.amazonaws.com",
    "eventName": "SendSSHPublicKey ",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "123.456.789.012",
    "userAgent": "aws-cli/1.15.61 Python/2.7.10 Darwin/16.7.0 botocore/1.10.60",
    "requestParameters": {
        "instanceId": "i-0123456789EXAMPLE",
        "osUser": "ec2-user",
        "SSHKey": {
            "publicKey": "ssh-rsa ABCDEFGHIJKLMNO01234567890EXAMPLE"
        }
     },
    "responseElements": null,
    "requestID": "1a2s3d4f-bde6-11e8-a892-f7ec64543add",
    "eventID": "1a2w3d4r5-a88f-4e28-b3bf-30161f75be34",
    "eventType": "AwsApiCall",
    "recipientAccountId": "0987654321"
}
```


Si vous avez configuré votre compte AWS de façon à collecter les événements CloudTrail dans un compartiment S3, vous pouvez télécharger et auditer les informations par programmation. Pour de plus amples informations, veuillez consulter Obtention et consultation des fichiers journaux de CloudTrail dans le AWS CloudTrail User Guide.

Mise en réseau dans Amazon EC2

Amazon EC2 offre les fonctionnalités de mise en réseau suivantes :

Fonctionnalités

- Adressage IP des instances Amazon EC2 (p. 808)
- Fourniture de vos propres adresses IP (BYOIP) dans Amazon EC2 (p. 825)
- Adresses IP Elastic (p. 831)
- Interfaces réseau Elastic (p. 840)
- Mise en réseau améliorée sur Linux (p. 864)
- Elastic Fabric Adapter (EFA) (p. 891)
- Groupes de placement (p. 924)
- Unité de transmission maximale (MTU) du réseau pour votre instance EC2 (p. 937)
- Clouds privés virtuels (p. 940)
- EC2-Classic (p. 941)

Adressage IP des instances Amazon EC2

Amazon EC2 et Amazon VPC prennent en charge les protocoles d'adressage IPv4 et IPv6. Par défaut, Amazon EC2 et Amazon VPC utilisent le protocole adressage IPv4 ; vous ne pouvez pas désactiver ce comportement. Lorsque vous créez un VPC, vous devez spécifier un bloc d'adresse CIDR IPv4 (une plage d'adresses IPv4 privées). Vous pouvez également attribuer un bloc d'adresse CIDR IPv6 à votre VPC et vos sous-réseaux, et attribuer des adresses IPv6 de ce bloc aux instances de votre sous-réseau. Les adresses IPv6 sont accessibles sur Internet. Pour plus d'informations sur IPv6, consultez Adressage IP de votre VPC dans le Amazon VPC Guide de l'utilisateur.

Sommaire

- Adresses IPv4 privées et noms d'hôte DNS internes (p. 808)
- Adresses IPv4 publiques et noms d'hôte DNS externes (p. 809)
- Adresses IP Elastic (IPv4) (p. 810)
- Serveur Amazon DNS (p. 810)
- Adresses IPv6 (p. 810)
- Utilisation des adresses IPv4 pour vos instances (p. 811)
- Utilisation des adresses IPv6 pour vos instances (p. 815)
- Plusieurs adresses IP (p. 817)

Adresses IPv4 privées et noms d'hôte DNS internes

Une adresse IPv4 privée est une adresse IP qui ne peut pas être atteinte via Internet. Vous pouvez utiliser des adresses IPv4 privées et un nom d'hôte DNS interne pour toute communication entre des instances du même VPC. Pour plus d'informations sur les normes et spécifications des adresses IPv4 privées, consultez RFC 1918. Nous allouons des adresses IPv4 privées aux instances à l'aide de DHCP.

> **Note**
>
> Vous pouvez créer un VPC avec un bloc d'adresse CIDR publiquement routable ne faisant pas partie des plages d'adresses IPv4 privées spécifiées dans la norme RFC 1918. Toutefois, dans le cadre de cette documentation, nous faisons référence à des adresses IPv4 privées (ou « adresses IP privées ») en tant qu'adresses IP se trouvant dans la plage CIDR IPv4 de votre VPC.

Lorsque vous lancez une instance, nous allouons une adresse IPv4 privée principale pour l'instance. Chaque instance se voit également attribuer un nom d'hôte DNS interne qui est résolu en adresse IPv4 privée principale, par exemple `ip-10-251-50-12.ec2.internal`. Vous pouvez utiliser le nom d'hôte DNS interne pour les communications entre les instances du même VPC, mais vous ne pouvez pas résoudre le nom d'hôte DNS interne en dehors du VPC.

Une instance reçoit une adresse IP privée principale de la plage d'adresses IPv4 du sous-réseau. Pour plus d'informations, consultez Dimensionnement des VPC et des sous-réseaux dans le Amazon VPC Guide de l'utilisateur. Si vous ne spécifiez pas d'adresse IP privée principale lorsque vous lancez l'instance, nous sélectionnons une adresse IP disponible dans la plage IPv4 de sous-réseaux à votre place. Chaque instance d'un VPC comporte une interface réseau par défaut (eth0) à laquelle une adresse IPv4 privée principale est attribuée. Vous pouvez également spécifier des adresses IPv4 privées supplémentaires, connues sous le nom d'adresses IPv4 privées secondaires. Contrairement aux adresses IP privées principales, les adresses IP privées secondaires peuvent être réaffectées d'une instance à une autre. Pour plus d'informations, consultez Plusieurs adresses IP (p. 817).

Qu'il s'agisse d'une adresse principale ou secondaire, une adresse IPv4 privée reste associée à l'interface réseau lorsque l'instance est arrêtée (ou mise en veille) et redémarrée, ainsi que lorsqu'elle est libérée lors de la désactivation de l'instance.

Adresses IPv4 publiques et noms d'hôte DNS externes

Une adresse IP publique est une adresse IPv4, qui est accessible depuis Internet. Vous pouvez utiliser des adresses publiques pour les communications entre vos instances et Internet.

Chaque instance qui reçoit une adresse IP publique se voit également attribuer un nom d'hôte DNS externe, par exemple `ec2-203-0-113-25.compute-1.amazonaws.com`. Nous résolvons un nom d'hôte DNS externe en adresse IP publique de l'instance en dehors de son VPC, et en adresse IPv4 privée de l'instance dans son VPC. L'adresse IP publique est mappée à l'adresse IP privée principale par le biais de NAT (Network Address Translation, traduction d'adresses réseau). Pour plus d'informations, consultez RFC 1631: The IP Network Address Translator (NAT).

Lorsque vous lancez une instance dans un VPC par défaut, nous lui attribuons une adresse IP publique par défaut. Lorsque vous lancez une instance sur un VPC autre qu'un VPC par défaut, le sous-réseau a un attribut qui détermine si les instances lancées sur ce sous-réseau reçoivent une adresse IP publique à partir du groupe d'adresses IPv4 publiques. Par défaut, nous n'attribuons aucune adresse IP publique aux instances lancées dans un sous-réseau autre que celui défini par défaut.

Vous pouvez contrôler si votre instance reçoit une adresse IP publique en procédant comme suit :

- Modifier l'attribut d'adressage IP public de votre sous-réseau. Pour plus d'informations, consultez Modification de l'attribut d'adressage IPv4 public de votre sous-réseau dans le Amazon VPC Guide de l'utilisateur.
- Activer ou désactiver la fonction d'adressage IP public pendant le lancement, ce qui remplace l'attribut d'adressage IP public du sous-réseau. Pour plus d'informations, consultez Attribution d'une adresse IPv4 publique lors du lancement d'une instance (p. 813).

Une adresse IP publique est attribuée à votre instance à partir du pool d'adresses IPv4 publiques d'Amazon ; elle n'est pas associée à votre compte AWS. Quand une adresse IP publique est dissociée de votre instance, elle est réintégrée dans le pool d'adresses IPv4 publiques et vous ne pouvez plus la réutiliser.

Vous ne pouvez pas associer manuellement une adresse IP publique à votre instance ou la dissocier de cette dernière. Par contre, dans certains cas, nous libérons l'adresse IP publique de votre instance ou nous lui en affectons une nouvelle:

- Nous libérons l'adresse IP publique de votre instance lorsqu'elle est arrêtée, mise en veille ou mise hors service. Toute instance arrêtée ou mise en veille de manière prolongée reçoit une nouvelle adresse IP publique au démarrage.
- L'adresse IP publique de votre instance est libérée lorsque vous lui associez une adresse IP Elastic. Lorsque vous dissociez l'adresse IP Elastic de votre instance, cette dernière reçoit une nouvelle adresse IP publique.
- Si l'adresse IP publique de votre instance sur un VPC a été libérée, cette instance ne recevra pas de nouvelle adresse si plusieurs interfaces réseau sont attachées à l'instance.
- Si l'adresse IP publique de votre instance est libérée alors qu'elle a une adresse IP privée secondaire associée à une adresse IP Elastic, l'instance ne reçoit pas de nouvelle adresse IP publique.

Si vous avez besoin d'une adresse IP publique permanente qui peut être associée aux instances et en être dissociée comme vous le souhaitez, utilisez plutôt une adresse IP Elastic.

Si vous utilisez DNS dynamique pour mapper un nom DNS existant à l'adresse IP publique d'une nouvelle instance, cela peut prendre jusqu'à 24 heures pour que l'adresse IP soit propagée via Internet. De ce fait, de nouvelles instances peuvent ne pas recevoir le trafic alors que des instances terminées continuent de recevoir des demandes. Pour résoudre ce problème, utilisez une adresse IP Elastic. Vous pouvez allouer votre propre adresse IP Elastic, puis l'associer à votre instance. Pour plus d'informations, consultez Adresses IP Elastic (p. 831).

Si vous attribuez une adresse IP Elastic à une instance, elle reçoit un nom d'hôte DNS IPv4 si les noms d'hôte DNS sont activés. Pour plus d'informations, consultez Utilisation de DNS avec votre VPC dans le Amazon VPC Guide de l'utilisateur.

> Note
>
> Les instances qui accèdent à d'autres instances via leur adresse IP NAT publique doivent régler le transfert de données régional ou Internet (selon que les instances se trouvent dans la même région ou non).

Adresses IP Elastic (IPv4)

Une adresse IP Elastic est une adresse IPv4 publique que vous pouvez allouer à votre compte. Vous pouvez l'associer à des instances et le dissocier selon vos besoins. Elle est allouée à votre compte jusqu'à ce que vous choisissiez de la libérer. Pour plus d'informations sur les adresses IP Elastic et leur utilisation, consultez Adresses IP Elastic (p. 831).

Nous ne prenons pas en charge les adresses IP Elastic pour IPv6.

Serveur Amazon DNS

Amazon fournit un serveur DNS qui résout les noms d'hôte DNS IPv4 fournis par Amazon aux adresses IPv4. Le serveur Amazon DNS se trouve à la base de votre plage réseau VPC plus deux. Pour plus d'informations, consultez Serveur Amazon DNS dans le Amazon VPC Guide de l'utilisateur.

Adresses IPv6

Le cas échéant, vous pouvez associer un bloc d'adresse CIDR IPv6 à votre VPC, et associer les blocs d'adresse CIDR IPv6 à vos sous-réseaux. Le bloc d'adresse CIDR IPv6 de votre VPC est automatiquement

attribué à partir du pool d'adresses IPv6 d'Amazon ; vous ne pouvez pas choisir la plage vous-même. Pour plus d'informations, consultez les rubriques suivantes dans le Amazon VPC Guide de l'utilisateur :

- Dimensionnement des VPC et des sous-réseaux pour IPv6
- Association d'un bloc d'adresse CIDR IPv6 à votre VPC
- Association d'un bloc d'adresse CIDR IPv6 à votre sous-réseau

Les adresses IPv6 sont globalement uniques, et par conséquent accessibles via Internet. Votre instance reçoit une adresse IPv6 si un bloc d'adresse CIDR IPv6 est associé à votre VPC et votre sous-réseau, et si l'une des conditions suivantes est vraie :

- Votre sous-réseau est configuré pour attribuer automatiquement une adresse IPv6 à une instance lors du lancement. Pour plus d'informations, consultez Modification de l'attribut d'adressage IPv6 public de votre sous-réseau.
- Vous attribuez une adresse IPv6 à votre instance lors du lancement.
- Vous attribuez une adresse IPv6 à l'interface réseau principale de votre instance après son lancement.
- Vous attribuez une adresse IPv6 à une interface réseau dans le même sous-réseau et vous liez l'interface réseau à votre instance après son lancement.

Lorsque votre instance reçoit une adresse IPv6 lors du lancement, l'adresse est associée à l'interface réseau principale (eth0) de l'instance. Vous pouvez dissocier l'adresse IPv6 de l'interface réseau. Nous ne prenons pas en charge les noms d'hôte DNS IPv6 pour votre instance.

Une adresse IPv6 persiste lorsque vous arrêtez (ou mettez en veille) et démarrez votre instance, et est libérée lorsque vous désactivez votre instance. Vous ne pouvez pas réattribuer une adresse IPv6 si elle est déjà attribuée à une autre interface réseau — vous devez d'abord annuler l'attribution.

Vous pouvez attribuer des adresses IPv6 supplémentaires à votre instance en les attribuant à une interface réseau attachée à votre instance. Le nombre d'adresses IPv6 que vous pouvez assigner à une interface réseau et le nombre d'interfaces réseau que vous pouvez lier à une instance varient en fonction du type d'instance. Pour plus d'informations, consultez Adresses IP par interface réseau et par type d'instance (p. 841).

Utilisation des adresses IPv4 pour vos instances

Vous pouvez attribuer une adresse IPv4 publique à votre instance lorsque vous la lancez. Vous pouvez afficher les adresses IPv4 de votre dans la console via la page Instances ou la page Interfaces réseau.

Sommaire

- Affichage des adresses IPv4 (p. 811)
- Attribution d'une adresse IPv4 publique lors du lancement d'une instance (p. 813)

Affichage des adresses IPv4

Vous pouvez utiliser la console Amazon EC2 afin d'afficher les adresses IPv4 privées, les adresses IPv4 publiques et les adresses IP Elastic de vos instances. Vous pouvez également déterminer les adresses IPv4 publiques et privées de votre instance depuis cette dernière en utilisant ses métadonnées. Pour plus d'informations, consultez Métadonnées d'instance et données utilisateur (p. 701).

L'adresse IPv4 publique est affichée comme propriété de l'interface réseau dans la console, mais elle est mappée à l'adresse IPv4 privée principale via la traduction d'adresses réseau (NAT, Network Address Translation). Par conséquent, si vous inspectez les propriétés de votre interface réseau sur votre instance, par exemple via `ifconfig` (Linux) ou `ipconfig` (Windows), l'adresse IPv4 publique ne s'affiche pas.

Pour déterminer l'adresse IPv4 publique de votre instance à partir d'une instance, utilisez les métadonnées d'instance.

Nouvelle console

Pour afficher les adresses IPv4 d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Instances, puis choisissez votre instance.
3. Les informations suivantes sont disponibles dans l'onglet Mise en réseau :
   - Adresse IPv4 publique — Adresse IPv4 publique. Si vous avez associé une adresse IP Elastic à l'instance ou à l'interface réseau principale, il s'agit de l'adresse IP Elastic.
   - DNS IPv4 public — Nom d'hôte DNS externe.
   - Adresses IPv4 privées — Adresse IPv4 privée.
   - DNS IPv4 privé — Nom d'hôte DNS interne.
   - Adresses IPv4 privées secondaires — Toutes adresse IPv4 privée secondaire.
   - Adresses IP élastiques — Toutes les adresses IP Elastic associées.
4. Sinon, sous Interfaces réseau, sous l'onglet Mise en réseau, choisissez l'ID d'interface de l'interface réseau principale (par exemple, eni-123abc456def78901). Les informations suivantes sont disponibles :
   - DNS privé (IPv4) — Nom d'hôte DNS interne.
   - IP IPv4 privée primaire — Adresse IPv4 privée primaire.
   - IP IPv4 privées secondaires — Toute adresse IPv4 privée secondaire.
   - DNS public — Nom d'hôte DNS externe.
   - IPv4 IP publique — Adresse IPv4 publique. Si vous avez associé une adresse IP Elastic à l'instance ou à l'interface réseau principale, il s'agit de l'adresse IP Elastic.
   - Adresses IP Elastic — Adresses IP élastiques associées.

Ancienne console

Pour afficher les adresses IPv4 d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Instances, puis choisissez votre instance.
3. Les informations suivantes sont disponibles dans l'onglet Description :
   - DNS privé — Nom d'hôte DNS interne.
   - IP privées — Adresse IPv4 privée.
   - IP privées secondaires — Toute adresse IPv4 privée secondaire.
   - DNS public — Nom d'hôte DNS externe.
   - IPv4 IP publique — Adresse IPv4 publique. Si vous avez associé une adresse IP Elastic à l'instance ou à l'interface réseau principale, il s'agit de l'adresse IP Elastic.
   - Adresses IP Elastic — Adresses IP élastiques associées.
4. Vous pouvez également afficher les adresses IPv4 de l'instance à l'aide de l'interface réseau principale. Sous Interfaces réseau, sous l'onglet Description choisissez eth0, puis choisissez l'ID d'interface (par exemple, eni-123abc456def78901). Les informations suivantes sont disponibles :
   - DNS privé (IPv4) — Nom d'hôte DNS interne.
   - IP IPv4 privée primaire — Adresse IPv4 privée primaire.
   - IP IPv4 privées secondaires — Toute adresse IPv4 privée secondaire.

- DNS public — Nom d'hôte DNS externe.
- IPv4 IP publique — Adresse IPv4 publique. Si vous avez associé une adresse IP Elastic à l'instance ou à l'interface réseau principale, il s'agit de l'adresse IP Elastic.
- Adresses IP Elastic — Adresses IP élastiques associées.

Pour afficher les adresses IPv4 d'une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-instances (AWS CLI)
- Get-EC2InstanceOutils AWS pour Windows PowerShell

Pour déterminer les adresses IPv4 publiques de votre instance à l'aide de ses métadonnées

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Utilisez la commande suivante pour accéder à l'adresse IP privée :

   IMDSv2

   ```
   [ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-
   aws-ec2-metadata-token-ttl-seconds: 21600"` \
   && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
   meta-data/local-ipv4
   ```

   IMDSv1

   ```
   [ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/local-ipv4
   ```

3. Utilisez la commande suivante pour accéder à l'adresse IP publique :

   IMDSv2

   ```
   [ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-
   aws-ec2-metadata-token-ttl-seconds: 21600"` \
   && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
   meta-data/public-ipv4
   ```

   IMDSv1

   ```
   [ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/public-ipv4
   ```

   Notez que si une adresse IP Elastic est associée à l'instance, la valeur renvoyée est celle de l'adresse IP Elastic.

Attribution d'une adresse IPv4 publique lors du lancement d'une instance

Chaque sous-réseau a un attribut qui détermine si une adresse IP publique est attribuée aux instances lancées dans ce sous-réseau. Par défaut, cet attribut est configuré sur false pour les sous-réseaux personnalisés et sur true pour les sous-réseaux par défaut. Lorsque vous lancez une instance, une fonction

d'adressage IPv4 public vous permet également de vérifier si votre instance dispose d'une adresse IPv4 publique. Vous pouvez remplacer le comportement par défaut de l'attribut d'adressage IP du sous-réseau. L'adresse IPv4 publique est attribuée à partir du pool d'adresses IPv4 publiques d'Amazon, et est attribuée à l'interface réseau avec l'index du périphérique eth0. Cette fonction dépend de certaines conditions au moment du lancement de votre instance.

Considérations

- Vous ne pouvez pas dissocier manuellement l'adresse IP publique de votre instance après le lancement. A la place, elle est libérée automatiquement dans certains cas, après quoi vous ne pouvez pas la réutiliser. Pour plus d'informations, consultez Adresses IPv4 publiques et noms d'hôte DNS externes (p. 809). Si vous avez besoin d'une adresse IP publique permanente que vous pouvez associer ou dissocier comme vous le souhaitez, attribuez plutôt une adresse IP Elastic à l'instance après le lancement. Pour plus d'informations, consultez Adresses IP Elastic (p. 831).
- Vous ne pouvez pas attribuer automatiquement une adresse IP publique si vous spécifiez plusieurs interfaces réseau. En outre, vous ne pouvez pas remplacer le paramètre de sous-réseau à l'aide de la fonction « auto-assign IP public », si vous spécifiez une interface réseau existante pour eth0.
- La fonction d'adressage IP public est uniquement disponible lors du lancement. Cependant, que vous attribuiez ou non une adresse IP publique à votre instance lors du lancement, vous pouvez associer une adresse IP Elastic à votre instance après son lancement. Pour plus d'informations, consultez Adresses IP Elastic (p. 831). Vous pouvez également modifier le comportement de l'adressage IPv4 public de votre sous-réseau. Pour de plus amples informations, veuillez consulter Modification de l'attribut d'adressage IPv4 public de votre sous-réseau.

Pour activer ou désactiver la fonctionnalité d'adressage IP publique à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Sélectionnez une AMI et un type d'instance, puis choisissez Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
4. Sur la page Configurer les détails de l'instance, sélectionnez un VPC dans le champ Réseau. La liste Attribuer automatiquement l'adresse IP publique s'affiche. Sélectionnez Activer ou Désactiver pour remplacer les paramètres par défaut du sous-réseau.
5. Suivez les étapes des pages suivantes de l'assistant pour effectuer la configuration de votre instance. Pour plus d'informations sur les options de configuration de l'assistant, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530). Sur la page finale Examiner le lancement de l'instance, vérifiez les paramètres, puis choisissez Lancer pour sélectionner une paire de clés et lancer votre instance.
6. Sur la page Instances, sélectionnez votre nouvelle instance et accédez à son adresse IP publique dans le champ IP publique IPv4 du volet de détails.

Pour activer ou désactiver la fonctionnalité d'adressage IP publique à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- Utilisez l'option `--associate-public-ip-address` ou `--no-associate-public-ip-address` avec la commande run-instances (AWS CLI)
- Utilisez le paramètre `-AssociatePublicIp` avec la commande New-EC2Instance (Outils AWS pour Windows PowerShell)

Utilisation des adresses IPv6 pour vos instances

Vous pouvez afficher les adresses IPv6 affectées à votre instance, attribuer une adresse IPv6 publique à votre instance ou annuler l'affectation d'une adresse IPv6 à votre instance. Vous pouvez afficher ces adresses dans la console via la page Instances ou la page Interfaces réseau.

Sommaire

- Affichage des adresses IPv6 (p. 815)
- Attribution d'une adresse IPv6 à une instance (p. 816)
- Annulation de l'attribution d'une adresse IPv6 d'une instance (p. 817)

Affichage des adresses IPv6

Vous pouvez utiliser la console Amazon EC2, AWS CLI, et les métadonnées d'instance pour afficher les adresses IPv6 de vos instances.

Nouvelle console

Pour afficher les adresses IPv4 d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Sous l'onglet Mise en réseau, recherchez les adresses IPv6.
5. Sinon, sous Interfaces réseau dans l'onglet Mise en réseau, choisissez l'ID de l'interface réseau (par exemple, eni-123abc456def78901). Recherchez IP IPv6.

Ancienne console

Pour afficher les adresses IPv4 d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Sous l'onglet Mise en réseau, recherchez IP IPv6.
5. Sinon, sous Interfaces réseau, sous l'onglet Description choisissez eth0, puis choisissez l'ID d'interface (par exemple, eni-123abc456def78901). Recherchez IP IPv6.

Pour afficher les adresses IPv6 d'une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-instances (AWS CLI)
- Get-EC2InstanceOutils AWS pour Windows PowerShell

Pour afficher les adresses IPv6 d'une instance à l'aide de métadonnées d'instance

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).

2. Utilisez la commande suivante pour afficher l'adresse IPv6 (vous pouvez obtenir l'adresse MAC à partir de `http://169.254.169.254/latest/meta-data/network/interfaces/macs/`).

   IMDSv2

   ```
   [ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-
   aws-ec2-metadata-token-ttl-seconds: 21600"` \
   && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
   meta-data/network/interfaces/macs/mac-address/ipv6s
   ```

   IMDSv1

   ```
   [ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/network/interfaces/
   macs/mac-address/ipv6s
   ```

Attribution d'une adresse IPv6 à une instance

Si votre VPC et votre sous-réseau disposent de blocs d'adresse CIDR IPv6 associés, vous pouvez attribuer une adresse IPv6 à votre instance pendant ou après le lancement. L'adresse IPv6 publique est attribuée à partir de plage d'adresses IPv6 publiques du sous-réseau à l'interface réseau avec l'index du périphérique eth0.

Pour attribuer une adresse IPv6 à une instance lors du lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sélectionnez une AMI et un type d'instance qui prend en charge IPv6, puis choisissez Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
3. Sur la page Configurer les détails de l'instance, sélectionnez un VPC dans le champ Réseau et un sous-réseau dans le champ Sous-réseau. Dans le champ Attribuer automatiquement l'adresse IP IPv6, choisissez Activer.
4. Complétez les étapes suivantes de l'assistant pour lancer votre instance.

Pour attribuer une adresse IPv6 à une instance après son lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez votre instance et choisissez Actions, Mise en réseau, puis Gérer les adresses IP privées.
4. Sélectionnez l'interface réseau. Sous Adresses IPv6, choisissez Attribuer une nouvelle adresse IP. Entrez une adresse IPv6 de la plage du sous-réseau ou laissez le champ vide pour permettre à Amazon de choisir une adresse IPv6 pour vous.
5. Choisissez Enregistrer.

Pour attribuer une adresse IPv6 à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- Utilisez l'option `--ipv6-addresses` avec la commande run-instances (AWS CLI)
- Utilisez la propriété `Ipv6Addresses` pour `-NetworkInterface` dans la commande New-EC2Instance (Outils AWS pour Windows PowerShell)
- assign-ipv6-addresses (AWS CLI)

- Register-EC2Ipv6AddressList (Outils AWS pour Windows PowerShell)

Annulation de l'attribution d'une adresse IPv6 d'une instance

Vous pouvez à tout moment annuler l'affectation d'une adresse IPv6 à partir d'une instance.

Pour annuler l'affectation d'une adresse IPv6 à une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez votre instance et choisissez Actions, Mise en réseau, puis Gérer les adresses IP privées.
4. Sélectionnez l'interface réseau. Sous Adresses IPv6, choisissez Annuler l'attribution en regard de l'adresse IPv6.
5. Choisissez Enregistrer.

Pour annuler l'affectation d'une adresse IPv6 à une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- unassign-ipv6-addresses (AWS CLI)
- Unregister-EC2Ipv6AddressList (Outils AWS pour Windows PowerShell).

Plusieurs adresses IP

Vous pouvez spécifier plusieurs adresses IPv4 et IPv6 privées pour vos instances. Le nombre d'interfaces réseau et d'adresses IPv4 et IPv6 privées que vous pouvez spécifier pour une instance dépend du type d'instance. Pour plus d'informations, consultez Adresses IP par interface réseau et par type d'instance (p. 841).

Il peut être utile d'attribuer plusieurs adresses IP privées à une instance sur votre VPC pour effectuer les opérations suivantes :

- Héberger plusieurs sites web sur un seul serveur en utilisant plusieurs certificats SSL sur un seul serveur et en associant chaque certificat à une adresse IP spécifique.
- Faire fonctionner les composants des réseaux tels que les pare-feu ou les équilibreurs de charge qui ont plusieurs adresses IP pour chaque interface réseau.
- Rediriger le trafic interne vers une instance de secours en cas d'échec de votre instance, en réattribuant l'adresse IP secondaire à l'instance de secours.

Sommaire

- Utilisation de plusieurs adresses IP (p. 817)
- Utilisation de plusieurs adresses IPv4 (p. 818)
- Utilisation de plusieurs adresses IPv6 (p. 822)

Utilisation de plusieurs adresses IP

La liste suivante explique le fonctionnement de plusieurs adresses IP avec les interfaces réseau :

- Vous pouvez attribuer une adresse IPv4 privée secondaire à n'importe quelle interface réseau. L'interface réseau n'a pas besoin d'être attachée à l'instance.

- Vous pouvez attribuer plusieurs adresses IPv6 à une interface réseau qui se trouve dans un sous-réseau disposant d'un bloc d'adresse CIDR IPv6 associé.
- Vous devez choisir une adresse IPv4 secondaire à partir de la plage du bloc d'adresse CIDR IPv4 du sous-réseau pour l'interface réseau.
- Vous devez choisir les adresses IPv6 à partir de la plage du bloc d'adresse CIDR IPv6 du sous-réseau de l'interface réseau.
- Vous associez des groupes de sécurité aux interfaces réseau, pas d'adresses IP individuelles. Par conséquent, chaque adresse IP que vous spécifiez dans une interface réseau est soumise au groupe de sécurité de son interface réseau.
- Plusieurs adresses IP peuvent être attribuées aux interfaces réseau liées aux instances en cours d'exécution ou arrêtées, ou leur attribution à ces interfaces peut être annulée.
- Les adresses IPv4 privées secondaires attribuées à une interface réseau peuvent être réattribuées à une autre interface si vous l'autorisez explicitement.
- Une adresse IPv6 ne peut pas être réattribuée à une autre interface réseau ; vous devez tout d'abord annuler l'attribution de l'adresse IPv6 à partir de l'interface réseau existante.
- Lorsque vous attribuez plusieurs adresses IP à une interface réseau à l'aide des outils ou de l'API de ligne de commande, l'opération complète échoue si l'une des adresses IP ne peut pas être attribuée.
- Les adresses IPv4 privées principales, les adresses IPv4 privées secondaires, les adresses IP Elastic et les adresses IPv6 restent avec une interface réseau secondaire lorsque celle-ci est dissociée d'une instance ou attachée à une instance.
- Même si vous ne pouvez pas détacher l'interface réseau principale à partir d'une instance, vous pouvez réattribuer l'adresse IPv4 privée secondaire de l'interface réseau principale à une autre interface réseau.

La liste suivante explique le fonctionnement de plusieurs adresses IP avec les adresses IP Elastic (IPv4 uniquement) :

- Chaque adresse IPv4 privée peut être associée à une seule adresse IP Elastic, et inversement.
- Lorsqu'une adresse IPv4 privée secondaire est réattribuée à une autre interface, l'adresse IPv4 privée secondaire conserve son association à une adresse IP Elastic.
- Lorsque l'attribution d'une adresse IPv4 privée secondaire est annulée à partir d'une interface, une adresse IP Elastic associée est automatiquement dissociée de l'adresse IPv4 privée secondaire.

Utilisation de plusieurs adresses IPv4

Vous pouvez attribuer l'adresse IPv4 privée secondaire à une instance, associer une adresse IPv4 Elastic à l'adresse IPv4 privée secondaire et annuler l'attribution d'une adresse IPv4 privée secondaire.

Sommaire

- Attribution d'une adresse IPv4 privée secondaire (p. 818)
- Configuration du système d'exploitation sur votre instance pour reconnaître l'adresse IPv4 privée secondaire (p. 820)
- Association d'une adresse IP Elastic à l'adresse IPv4 privée secondaire (p. 821)
- Consultation de vos adresses IPv4 privées secondaires (p. 821)
- Annulation de l'attribution d'une adresse IPv4 privée secondaire (p. 821)

Attribution d'une adresse IPv4 privée secondaire

Vous pouvez attribuer l'adresse IPv4 privée secondaire à l'interface réseau pour une instance au moment du lancement de l'instance ou après celui-ci. Cette section comprend les procédures suivantes.

- Pour attribuer une adresse IPv4 privée secondaire lors du lancement d'une instance (p. 819)
- Pour attribuer une adresse IPv4 secondaire lors du lancement à l'aide de la ligne de commande (p. 820)
- Pour attribuer une adresse IPv4 privée secondaire à une interface réseau (p. 820)
- Pour attribuer une adresse IPv4 privée secondaire à une instance existante à l'aide de la ligne de commande (p. 820)

Pour attribuer une adresse IPv4 privée secondaire lors du lancement d'une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Sélectionnez une AMI et un type d'instance, puis choisissez Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
4. Sur la page Configurer les détails de l'instance, sélectionnez un VPC dans le champ Réseau et un sous-réseau dans le champ Sous-réseau.
5. Dans la section Interfaces réseau, effectuez les opérations suivantes, puis choisissez Next: Add Storage (Suivant : Ajouter le stockage) :

   - Pour ajouter une interface réseau, choisissez Ajouter périphérique. La console vous permet de spécifier jusqu'à deux interfaces réseau lorsque vous lancez une instance. Une fois l'instance lancée, choisissez Interfaces réseau dans le panneau de navigation pour ajouter des interfaces réseau. Le nombre total d'interfaces réseau que vous pouvez attacher varie en fonction du type d'instance. Pour plus d'informations, consultez Adresses IP par interface réseau et par type d'instance (p. 841).

     > **Important**
     >
     > Lorsque vous ajoutez une deuxième interface réseau, le système ne peut plus attribuer automatiquement d'adresse IPv4 publique. Vous ne pourrez pas connecter l'instance sur IPv4 à moins d'attribuer une adresse IP Elastic à l'interface réseau principale (eth0). Vous pouvez attribuer l'adresse IP Elastic une fois l'assistant de lancement terminé. Pour plus d'informations, consultez Utilisation d'adresses IP Elastic (p. 832).

   - Pour chaque interface réseau, sous Adresses IP secondaires, choisissez Ajouter une IP, puis entrez une adresse IP privée de la plage de sous-réseaux ou acceptez la valeur par défaut, `Auto-assign`, pour permettre à Amazon de sélectionner une adresse.
6. Sur la page suivante Ajouter le stockage, vous pouvez spécifier les volumes que vous souhaitez attacher à l'instance, outre ceux spécifiés par l'AMI (par exemple, le volume du périphérique racine), puis sélectionner Next: Add Tags (Suivant : Ajouter des balises).
7. Sur la page Ajouter des balises, spécifiez des balises pour l'instance, par exemple un nom évocateur, puis sélectionnez Suivant : Configurer le groupe de sécurité.
8. Sur la page Configurer le groupe de sécurité, sélectionnez un groupe de sécurité existant ou créez-en un nouveau. Choisissez Vérifier et lancer.
9. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis choisissez Lancer pour sélectionner une paire de clés et lancer votre instance. Si vous débutez avec Amazon EC2 et que vous n'avez pas encore créé de paire de clés, l'assistant vous invite à en créer une.

> **Important**
>
> Après avoir ajouté une adresse IP privée secondaire à une interface réseau, vous devez connecter l'instance et configurer l'adresse IP privée secondaire sur l'instance elle-même. Pour plus d'informations, consultez Configuration du système d'exploitation sur votre instance pour reconnaître l'adresse IPv4 privée secondaire (p. 820).

Pour attribuer une adresse IPv4 secondaire lors du lancement à l'aide de la ligne de commande

- Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).
  - L'option `--secondary-private-ip-addresses` avec la commande run-instances (AWS CLI)
  - Définissez `-NetworkInterface` et spécifiez le paramètre `PrivateIpAddresses` avec la commande New-EC2Instance (Outils AWS pour Windows PowerShell).

Pour attribuer une adresse IPv4 privée secondaire à une interface réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Interfaces réseau dans le panneau de navigation, puis sélectionnez l'interface réseau attachée à l'instance.
3. Choisissez Actions, Gérer les adresses IP.
4. Sous Adresses IPv4, choisissez Attribuer une nouvelle adresse IP.
5. Entrez une adresse IPv4 spécifique située dans la plage de sous-réseaux de l'instance ou laissez le champ vide pour laisser Amazon sélectionner l'adresse IP à votre place.
6. (Facultatif) Sélectionnez Autoriser la réattribution pour permettre la réattribution de l'adresse IP privée secondaire si elle était déjà attribuée à une autre interface réseau.
7. Choisissez Oui, mettre à jour.

Vous pouvez également attribuer une adresse IPv4 privée secondaire à une instance. Choisissez Instances dans le panneau de navigation, sélectionnez l'instance et choisissez Actions, sélectionnez Mise en réseau, puis Gérer les adresses IP. Vous pouvez configurer les mêmes informations que précédemment. L'adresse IP est attribuée à l'interface réseau principale (eth0) pour l'instance.

Pour attribuer une adresse IPv4 privée secondaire à une instance existante à l'aide de la ligne de commande

- Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).
  - assign-private-ip-addresses (AWS CLI)
  - Register-EC2PrivateIpAddress (Outils AWS pour Windows PowerShell)

Configuration du système d'exploitation sur votre instance pour reconnaître l'adresse IPv4 privée secondaire

Une fois que vous avez attribué une adresse IPv4 privée secondaire à votre instance, vous devez configurer le système d'exploitation de cette dernière afin qu'il puisse reconnaître l'adresse IP privée secondaire.

- Si vous utilisez Amazon Linux, le package ec2-net-utils peut effectuer cette opération. Il configure les interfaces réseau supplémentaires associées pendant l'exécution de l'instance, il rafraîchit les adresses IPv4 secondaires au cours du renouvellement du bail DHCP et il met à jour les règles de routage associées. Vous pouvez rafraîchir immédiatement la liste d'interfaces en utilisant la commande `sudo service network restart`, puis consulter la liste mise à jour en utilisant `ip addr li`. Si vous avez besoin d'un contrôle manuel sur votre configuration réseau, vous pouvez supprimer le package ec2-net-utils. Pour plus d'informations, consultez Configuration de votre interface réseau à l'aide de ec2-net-utils (p. 862).
- Si vous utilisez une autre distribution Linux, consultez la documentation correspondante. Recherchez des informations sur la configuration d'interfaces réseau et d'adresses IPv4 secondaires supplémentaires.

Si l'instance a deux ou plusieurs interfaces sur le même sous-réseau, recherchez des informations sur l'utilisation des règles de routage pour contourner le routage asymétrique.

Pour de plus amples informations sur la configuration d'une instance Windows, veuillez consulter Configuration d'une adresse IP privée secondaire pour votre instance Windows dans un VPC dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Association d'une adresse IP Elastic à l'adresse IPv4 privée secondaire

Pour associer une adresse IP Elastic à une adresse IPv4 privée secondaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Adresses IP Elastic.
3. Sélectionnez Actions, puis Associer l'adresse.
4. Pour Interface réseau, sélectionnez l'interface réseau, puis sélectionnez l'adresse IP secondaire dans la liste IP privée.
5. Choisissez Associate.

Pour associer une adresse IP Elastic à une adresse IPv4 privée secondaire à l'aide de la ligne de commande

- Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).
  - associate-address (AWS CLI)
  - Register-EC2Address (Outils AWS pour Windows PowerShell)

Consultation de vos adresses IPv4 privées secondaires

Pour consulter les adresses IPv4 privées attribuées à une interface réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau avec les adresses IP privées à afficher.
4. Sur l'onglet Détails du volet de détails, vérifiez les champs IP privée principale IPv4 et IP privées secondaires IPv4 de l'adresse IPv4 privée principale et des adresses IPv4 privées secondaires attribuées à l'interface réseau.

Pour consulter les adresses IPv4 privées attribuées à une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance avec les adresses IPv4 privées à afficher.
4. Sur l'onglet Description du volet de détails, vérifiez les champs IP privées et IP privées secondaires de l'adresse IPv4 privée principale et des adresses IPv4 privées secondaires attribuées à l'instance via son interface réseau.

Annulation de l'attribution d'une adresse IPv4 privée secondaire

Si vous n'avez plus besoin d'une adresse IPv4 privée secondaire, vous pouvez annuler son attribution à partir de l'instance ou de l'interface réseau. Lorsque l'attribution d'une adresse IPv4 privée secondaire est annulée à partir d'une interface réseau, l'adresse IP Elastic (si elle existe) est également dissociée.

Pour annuler l'attribution d'une adresse IPv4 privée secondaire depuis une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une instance, puis choisissez Actions, Mise en réseau, Gérer les adresses IP.
4. Sous Adresses IPv4, choisissez Annuler l'attribution pour l'adresse IPv4 dont l'attribution est à annuler.
5. Choisissez Oui, mettre à jour.

Pour annuler l'attribution d'une adresse IPv4 privée secondaire depuis une interface réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau, puis choisissez Actions, Gérer les adresses IP.
4. Sous Adresses IPv4, choisissez Annuler l'attribution pour l'adresse IPv4 dont l'attribution est à annuler.
5. Choisissez Oui, mettre à jour.

Pour annuler l'attribution d'une adresse IPv4 privée secondaire à l'aide de la ligne de commande

- Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).
  - unassign-private-ip-addresses (AWS CLI)
  - Unregister-EC2PrivateIpAddress (Outils AWS pour Windows PowerShell)

Utilisation de plusieurs adresses IPv6

Vous pouvez attribuer plusieurs adresses IPv6 à votre instance, afficher les adresses IPv6 attribuées à votre instance et annuler l'attribution d'adresses IPv6 à partir de votre instance.

Sommaire

- Attribution de plusieurs adresses IPv6 (p. 822)
- Affichage de vos adresses IPv6 (p. 824)
- Annulation de l'attribution d'une adresse IPv6 (p. 824)

Attribution de plusieurs adresses IPv6

Vous pouvez attribuer une ou plusieurs adresses IPv6 à votre instance pendant ou après le lancement. Pour attribuer une adresse IPv6 à une instance, le VPC et le sous-réseau dans lequel vous lancez l'instance doivent disposer d'un bloc d'adresse CIDR IPv6 associé. Pour plus d'informations, consultez VPC et sous-réseaux dans le Amazon VPC Guide de l'utilisateur.

Pour attribuer plusieurs adresses IPv6 lors du lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sur le tableau de bord, choisissez Lancer une instance.
3. Sélectionnez une AMI, choisissez un type d'instance, puis choisissez Next: Configure Instance Details (Suivant : Configurer les détails de l'instance). Vérifiez que vous choisissez un type d'instance qui prend en charge IPv6. Pour plus d'informations, consultez Types d'instance (p. 213).
4. Sur la page Configurer les détails de l'instance, sélectionnez votre VPC dans la liste Réseau et le sous-réseau dans la liste Sous-réseau.

5. Dans la section Interfaces réseau, effectuez les opérations suivantes, puis choisissez Next: Add Storage (Suivant : Ajouter le stockage) :
   - Pour attribuer une seule adresse IPv6 à l'interface réseau principale (eth0), sous Adresses IP IPv6, choisissez Ajouter une IP. Pour ajouter une adresse IPv6 secondaire, choisissez à nouveau Ajouter une IP. Vous pouvez entrer une adresse IPv6 de la plage du sous-réseau, ou conserver la valeur par défaut Attribution automatique afin de laisser Amazon choisir à votre place une adresse IPv6 du sous-réseau.
   - Choisissez Ajouter périphérique pour ajouter une autre interface réseau et répétez les étapes ci-dessus pour ajouter une ou plusieurs adresses IPv6 à l'interface réseau. La console vous permet de spécifier jusqu'à deux interfaces réseau lorsque vous lancez une instance. Une fois l'instance lancée, choisissez Interfaces réseau dans le panneau de navigation pour ajouter des interfaces réseau. Le nombre total d'interfaces réseau que vous pouvez attacher varie en fonction du type d'instance. Pour plus d'informations, consultez Adresses IP par interface réseau et par type d'instance (p. 841).
6. Suivez les étapes suivantes de l'Assistant pour attacher des volumes et attribuer des balises à votre instance.
7. Sur la page Configurer le groupe de sécurité, sélectionnez un groupe de sécurité existant ou créez-en un nouveau. Si vous voulez que votre instance soit accessible via IPv6, assurez-vous que votre groupe de sécurité comporte des règles autorisant l'accès à partir des adresses IPv6. Pour plus d'informations, consultez Référence des règles de groupe de sécurité (p. 1075). Choisissez Vérifier et lancer.
8. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis choisissez Lancer pour sélectionner une paire de clés et lancer votre instance. Si vous débutez avec Amazon EC2 et que vous n'avez pas encore créé de paire de clés, l'assistant vous invite à en créer une.

Vous pouvez utiliser l'écran Instances de la console Amazon EC2 pour attribuer plusieurs adresses IPv6 à une instance existante. Ceci attribue les adresses IPv6 à l'interface réseau principale (eth0) pour l'instance. Pour attribuer une adresse IPv6 spécifique à l'instance, vérifiez que l'adresse IPv6 n'est pas déjà attribuée à une autre instance ou interface réseau.

Pour attribuer plusieurs adresses IPv6 à une instance existante

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez votre instance et choisissez Actions, Mise en réseau, puis Gérer les adresses IP.
4. Sous Adresses IPv6, choisissez Attribuer une nouvelle adresse IP pour chaque adresse IPv6 à ajouter. Vous pouvez spécifier une adresse IPv6 de la plage du sous-réseau, ou conserver la valeur Attribution automatique pour laisser Amazon choisir une adresse IPv6 à votre place.
5. Choisissez Oui, mettre à jour.

Vous pouvez également attribuer plusieurs adresses IPv6 à une interface réseau existante. L'interface réseau doit avoir été créée dans un sous-réseau qui dispose d'un bloc d'adresse CIDR IPv6 associé. Pour attribuer une adresse IPv6 spécifique à l'interface réseau, vérifiez que l'adresse IPv6 n'est pas déjà affectée à une autre interface réseau.

Pour attribuer plusieurs adresses IPv6 à une interface réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez votre interface réseau, choisissez Actions, puis Gérer les adresses IP.
4. Sous Adresses IPv6, choisissez Attribuer une nouvelle adresse IP pour chaque adresse IPv6 à ajouter. Vous pouvez spécifier une adresse IPv6 de la plage du sous-réseau, ou conserver la valeur Attribution automatique pour laisser Amazon choisir une adresse IPv6 à votre place.

5. Choisissez Oui, mettre à jour.

Présentation de l'interface de ligne de commande (CLI)

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- Attribuer une adresse IPv6 lors du lancement :
  - Utilisez les options `--ipv6-addresses` ou `--ipv6-address-count` avec la commande run-instances (AWS CLI)
  - Définissez `-NetworkInterface` et spécifiez les paramètres `Ipv6Addresses` ou `Ipv6AddressCount` avec la commande New-EC2Instance (Outils AWS pour Windows PowerShell).
- Attribuer une adresse IPv6 à une interface réseau :
  - assign-ipv6-addresses (AWS CLI)
  - Register-EC2Ipv6AddressList (Outils AWS pour Windows PowerShell)

Affichage de vos adresses IPv6

Vous pouvez afficher les adresses IPv6 d'une instance ou d'une interface réseau.

Pour consulter les adresses IPv6 attribuées à une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez votre instance. Dans le volet des détails, vérifiez le champ Adresses IP IPv6.

Pour afficher les adresses IPv6 attribuées à une interface réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez votre interface réseau. Dans le volet des détails, vérifiez le champ Adresses IP IPv6.

Présentation de l'interface de ligne de commande (CLI)

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- Afficher les adresses IPv6 d'une instance :
  - describe-instances (AWS CLI)
  - Get-EC2InstanceOutils AWS pour Windows PowerShell
- Afficher les adresses IPv6 d'une interface réseau :
  - describe-network-interfaces (AWS CLI)
  - Get-EC2NetworkInterface (Outils AWS pour Windows PowerShell)

Annulation de l'attribution d'une adresse IPv6

Vous pouvez annuler l'attribution d'une adresse IPv6 à partir de l'interface réseau principale d'une instance, ou vous pouvez annuler l'attribution d'une adresse IPv6 à partir d'une interface réseau.

Pour annuler l'attribution d'une adresse IPv6 à partir d'une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez votre instance et choisissez Actions, Mise en réseau, puis Gérer les adresses IP.
4. Sous Adresses IPv6, choisissez Annuler l'attribution pour l'adresse IPv6 dont l'attribution est à annuler.
5. Choisissez Oui, mettre à jour.

Pour annuler l'attribution d'une adresse IPv6 à partir d'une interface réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez votre interface réseau, choisissez Actions, puis Gérer les adresses IP.
4. Sous Adresses IPv6, choisissez Annuler l'attribution pour l'adresse IPv6 dont l'attribution est à annuler.
5. Choisissez Enregistrer.

Présentation de l'interface de ligne de commande (CLI)

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- unassign-ipv6-addresses (AWS CLI)
- Unregister-EC2Ipv6AddressList (Outils AWS pour Windows PowerShell).

Fourniture de vos propres adresses IP (BYOIP) dans Amazon EC2

Vous pouvez fournir tout ou partie de votre plage d'adresses IPv4 ou IPv6 publiques depuis votre réseau sur site vers votre compte AWS. La plage d'adresses vous appartient toujours, mais AWS la publie sur Internet par défaut. Une fois que vous avez fourni la plage d'adresses à AWS, celle-ci s'affiche dans votre compte en tant que groupe d'adresses.

La fonctionnalité BYOIP n'est pas disponible dans toutes les régions ni pour toutes les ressources. Pour obtenir la liste des régions prises en charge, veuillez consulter les Questions fréquentes (FAQ) sur la fonctionnalité Bring Your Own IP.

> **Note**
>
> Les étapes suivantes décrivent comment fournir votre propre plage d'adresses IP pour l'utiliser dans Amazon EC2 uniquement. Pour connaître les étapes à suivre pour fournir votre propre plage d'adresses IP pour l'utiliser dans AWS Global Accelerator, veuillez consulter Fourniture de vos propres adresses IP (BYOIP) dans le Manuel du développeur AWS Global Accelerator.

Rubriques

- Exigences (p. 826)
- Préparation à l'importation de votre plage d'adresses dans votre compte AWS (p. 826)
- Mise en service de la plage d'adresses pour une utilisation avec AWS (p. 828)
- Publication de la plage d'adresses via AWS (p. 829)
- Utiliser votre plage d'adresses (p. 830)
- Mise hors service de la plage d'adresses (p. 831)

Exigences

- La plage d'adresses doit être enregistrée auprès de votre registre Internet régional (RIR), tel que l'ARIN (American Registry for Internet Numbers), le RIPE NCC (Réseaux IP Européens Network Coordination Centre) ou l'APNIC (Asia-Pacific Network Information Centre). Elle doit être enregistrée pour une entreprise ou une entité institutionnelle et ne peut pas être enregistrée pour une personne individuelle.
- La plage d'adresses IPv4 la plus spécifique que vous pouvez apporter est /24.
- La plage d'adresses IPv6 la plus spécifique que vous pouvez apporter est /48 pour les CIDR annoncés publiquement et /56 pour les CIDR qui ne sont pas annoncés publiquement (p. 829).
- Vous ne pouvez importer chaque plage d'adresses que dans une région à la fois.
- Vous pouvez apporter un total de cinq plages d'adresses IPv4 et IPv6 par région à votre compte AWS.
- L'historique des adresses de la plage d'adresses IP doit être propre. Nous pouvons enquêter sur la réputation de la plage d'adresses IP et nous réserver le droit de rejeter une plage d'adresses IP si elle contient une adresse IP qui a une mauvaise réputation ou qui est associée à un comportement malveillant.
- Vous devez être propriétaire de l'adresse IP que vous utilisez. Cela signifie que seuls les éléments suivants sont pris en charge :
  - ARIN - Types de réseaux « Direct Allocation » et « Direct Assignment »
  - RIPE - Statuts d'allocation « ALLOCATED PA », « LEGACY » et « ASSIGNED PI » et « ALLOCATED-BY-RIR »
  - APNIC - Statuts d'allocation « ALLOCATED PORTABLE » et « ASSIGNED PORTABLE »

Préparation à l'importation de votre plage d'adresses dans votre compte AWS

Pour vérifier que vous seul pouvez importer votre plage d'adresses dans votre compte AWS, vous devez autoriser Amazon à publier la plage d'adresses. Vous devez également fournir la preuve que vous êtes propriétaire de la plage d'adresses via un message d'autorisation signé.

Une ROA (Route Origin Authorization) est une instruction de chiffrement concernant les annonces de routage que vous pouvez créer via votre RIR. Ce document contient la plage d'adresses, les numéros de système autonome (ASN) qui sont autorisés à publier la plage d'adresses et une date d'expiration. Une ROA autorise Amazon à publier une plage d'adresses sous un numéro AS spécifique. Elle n'autorise toutefois pas votre compte AWS à importer la plage d'adresses dans AWS. Pour autoriser votre compte AWS à importer une plage d'adresses dans AWS, vous devez publier un certificat X509 auto-signé dans les remarques RDAP (Registry Data Access Protocol) pour la plage d'adresses. Le certificat contient une clé publique utilisée par AWS pour vérifier la signature de contexte d'autorisation que vous fournissez. Il est recommandé de sécuriser la clé privée et de l'utiliser pour signer le message de contexte d'autorisation.

Les commandes dans ces tâches sont prises en charge sous Linux. Sous Windows, vous pouvez utiliser Windows Subsystem for Linux pour exécuter des commandes Linux.

Tâches

- Créer un objet ROA (p. 827)
- Créer un certificat X509 auto-signé (p. 827)
- Créer un message d'autorisation signé (p. 827)

Créer un objet ROA

Créez un objet ROA pour autoriser les ASN 16509 et 14618 d'Amazon à publier votre plage d'adresses et les ASN qui sont actuellement autorisés à publier la plage d'adresses. Vous devez définir la longueur maximale du plus petit préfixe que vous souhaitez importer (/24 par exemple). La mise à disposition de la ROA sur Amazon peut prendre jusqu'à 24 heures. Pour plus d'informations, consultez les ressources suivantes :

- ARIN — ROA Requests
- RIPE — Managing ROAs
- APNIC — Route Management

Créer un certificat X509 auto-signé

Utilisez la procédure suivante pour créer un certificat X509 auto-signé et l'ajouter à l'enregistrement RDAP de votre RIR. Les commandes openssl exigent OpenSSL version 1.0.2 ou ultérieure.

Copiez les commandes ci-dessous et remplacez uniquement les valeurs d'espace réservé (en italique et en couleur).

Pour créer un certificat X509 auto-signé et l'ajouter à l'enregistrement RDAP

1. Générez une paire de clés RSA 2048 bits comme indiqué ci-après.

```
openssl genrsa -out private.key 2048
```

2. Créez un certificat public X509 à partir de la paire de clés en utilisant la commande suivante. Dans cet exemple, le certificat expire dans 365 jours, après quoi il n'est plus fiable. Veillez donc à définir l'expiration de façon appropriée. Lorsque vous êtes invité à saisir des informations, vous pouvez accepter les valeurs par défaut.

```
openssl req -new -x509 -key private.key -days 365 | tr -d "\n" > publickey.cer
```

3. Mettez à jour l'enregistrement RDAP pour votre RIR avec le certificat X509. Assurez-vous de copier les éléments `-----BEGIN CERTIFICATE-----` et `-----END CERTIFICATE-----` à partir du certificat. Si ce n'est pas déjà fait, assurez-vous de supprimer les caractères de saut de ligne à l'aide des commandes tr -d "\n" aux étapes précédentes. Pour afficher votre certificat, exécutez la commande suivante.

```
cat publickey.cer
```

   Pour l'ARIN, ajoutez le certificat dans la section « Public Comments » pour votre plage d'adresses. Ne l'ajoutez pas à la section des commentaires de votre organisation.

   Pour le RIPE, ajoutez le certificat sous la forme d'un nouveau champ « descr » pour votre plage d'adresses. Ne l'ajoutez pas à la section des commentaires de votre organisation.

   Pour l'APNIC, envoyez la clé publique par e-mail à l'adresse helpdesk@apnic.net afin de l'ajouter manuellement au champs « remarks (remarques) » pour votre plage d'adresses. Envoyez l'e-mail en utilisant le contact autorisé APNIC pour les adresses IP.

Créer un message d'autorisation signé

Le format du message d'autorisation signé est le suivant (la date est la date d'expiration du message).

```
1|aws|account|cidr|YYYYMMDD|SHA256|RSAPSS
```

Pour créer un message d'autorisation signé

1. Créez un message d'autorisation en texte brut et stockez-le dans une variable nommée `text_message` comme illustré dans l'exemple suivant. Copiez l'exemple suivant et remplacez uniquement le numéro de compte, la plage d'adresses et la date d'expiration de l'exemple par vos propres valeurs.

```
text_message="1|aws|123456789012|198.51.100.0/24|20191201|SHA256|RSAPSS"
```

2. Signez le message d'autorisation dans `text_message` en utilisant la paire de clés que vous avez créée, puis stockez-le dans une variable nommée `signed_message`.

```
signed_message=$(echo $text_message | tr -d "\n" | openssl dgst -sha256 -sigopt
 rsa_padding_mode:pss -sigopt rsa_pss_saltlen:-1 -sign private.key -keyform PEM |
 openssl base64 | tr -- '+=/' '-_~' | tr -d "\n")
```

> **Important**
>
> Nous vous recommandons de copier et de coller cette commande. Ne modifiez ni ne remplacez aucune des valeurs.

Mise en service de la plage d'adresses pour une utilisation avec AWS

Lorsque vous mettez en service une plage d'adresses pour une utilisation avec AWS, vous confirmez que vous êtes propriétaire de la plage d'adresses et vous autorisez Amazon à la publier. Nous vérifions également que vous possédez la plage d'adresses via un message d'autorisation signé. Ce message est signé avec la paire de clés X509 auto-signée que vous avez utilisée lors de la mise à jour de l'enregistrement RDAP avec le certificat X509.

Pour mettre en service la plage d'adresses, utilisez la commande provision-byoip-cidr suivante. Remplacez la plage d'adresses de l'exemple par votre propre plage d'adresses. L'option `--cidr-authorization-context` utilise les variables que vous avez créées précédemment, pas le message ROA.

```
aws ec2 provision-byoip-cidr --cidr address-range --cidr-authorization-context
 Message="$text_message",Signature="$signed_message"
```

La mise en service d'une plage d'adresses est une opération asynchrone : l'appel est immédiatement renvoyé, mais la plage d'adresses ne peut pas être utilisée tant que son statut ne bascule pas de `pending-provision` à `provisioned`. Le processus d'allocation peut durer jusqu'à trois semaines. Pour surveiller le statut des plages d'adresses que vous allouez, utilisez la commande describe-byoip-cidrs suivante.

```
aws ec2 describe-byoip-cidrs --max-results 5
```

S'il y a des problèmes pendant la mise en service et que l'état passe à `failed-provision`, vous devez exécuter à nouveau la commande `provision-byoip-cidr` une fois que les problèmes ont été résolus.

Provisionner une plage d'adresses IPv6 qui n'est pas annoncée publiquement

Par défaut, une plage d'adresses est provisionnée pour être annoncée publiquement sur Internet. Vous pouvez provisionner une plage d'adresses IPv6 qui ne sera pas publiée publiquement. Lorsque vous associez un bloc CIDR IPv6 d'une plage d'adresses non publique à un VPC, le CIDR IPv6 ne peut être accessible que via une connexion AWS Direct Connect.

Un ROA n'est pas nécessaire pour fournir une plage d'adresses non publique.

Pour provisionner une plage d'adresses IPv6 qui ne sera pas publiée publiquement, utilisez la commande provision-byoip-cidr suivante.

```
aws ec2 provision-byoip-cidr --cidr address-range --cidr-authorization-context
 Message="$text_message",Signature="$signed_message" --no-publicly-advertisable
```

> **Important**
>
> Vous ne pouvez définir l'indicateur `publicly-advertisable` ou `no-publicly-advertisable` que pendant le provisionnement. Vous ne pouvez pas modifier le statut annoncé d'une plage d'adresses ultérieurement.

Publication de la plage d'adresses via AWS

Une fois que la plage d'adresses est mise en service, elle est prête à être publiée. Vous devez publier la plage d'adresses exacte que vous avez mise en service. Vous ne pouvez pas publier seulement une portion de la plage d'adresses mise en service.

Si vous avez provisionné une plage d'adresses IPv6 qui ne sera pas publiée publiquement, vous n'avez pas besoin de terminer cette étape.

Nous vous recommandons d'arrêter de publier la plage d'adresses à partir d'autres emplacements avant de la publier via AWS. Si vous continuez à publier votre plage d'adresses IP à partir d'autres emplacements, sa prise en charge ne sera pas assurée de manière fiable ou les problèmes associés ne pourront pas être identifiés et résolus. Plus précisément, nous ne pouvons pas garantir que le trafic vers la plage d'adresses entrera dans notre réseau.

Pour réduire les temps d'arrêt, vous pouvez configurer vos ressources AWS de sorte à utiliser une adresse issue de votre groupe d'adresses avant d'être publiée, puis simultanément, arrêter de la publier à partir de l'emplacement actuel et commencer à la publier via AWS. Pour plus d'informations sur l'allocation d'une adresse IP Elastic à partir de votre groupe d'adresses, consultez Allocation d'une adresse IP Elastic (p. 833).

Pour publier la plage d'adresses, utilisez la commande advertise-byoip-cidr suivante.

```
aws ec2 advertise-byoip-cidr --cidr address-range
```

> **Important**
>
> Vous pouvez exécuter la commande advertise-byoip-cidr au moins une fois tous les 10 secondes, même si vous spécifiez des plages d'adresses différentes à chaque fois.

Pour arrêter la publication de la plage d'adresses, utilisez la commande withdraw-byoip-cidr suivante.

```
aws ec2 withdraw-byoip-cidr --cidr address-range
```

Important

Vous pouvez exécuter la commande withdraw-byoip-cidr au moins une fois tous les 10 secondes, même si vous spécifiez des plages d'adresses différentes à chaque fois.

Utiliser votre plage d'adresses

Vous pouvez afficher et utiliser les plages d'adresses IPv4 et IPv6 que vous avez provisionnées dans votre compte.

Plages d'adresses IPv4

Vous pouvez créer une adresse IP Elastic à partir de votre groupe d'adresses IPv4 et l'utiliser avec vos ressources AWS, telles que vos instances EC2, vos passerelles NAT et vos Équilibreurs de charge du réseau.

Pour afficher des informations sur les pools d'adresses IPv4 que vous avez provisionnés dans votre compte, utilisez la commande describe-public-ipv4-pools suivante.

```
aws ec2 describe-public-ipv4-pools
```

Pour créer une adresse IP Elastic à partir de votre pool d'adresses, utilisez la commande allocate-address. Vous pouvez utiliser l'option `--public-ipv4-pool` pour spécifier l'ID du groupe d'adresses renvoyé par `describe-byoip-cidrs`. Vous pouvez aussi utiliser l'option `--address` pour spécifier une adresse de la plage d'adresses que vous avez allouée.

Plages d'adresses IPv6

Pour afficher des informations sur les pools d'adresses IPv6 que vous avez provisionnés dans votre compte, utilisez la commande describe-ipv6-pools suivante.

```
aws ec2 describe-ipv6-pools
```

Pour créer un VPC et spécifier un CIDR IPv6 à partir de votre pool d'adresses IPv6, utilisez la commande create-vpc suivante. Pour laisser Amazon choisir le CIDR IPv6 dans votre pool d'adresses IPv6, omettez l'option `--ipv6-cidr-block`.

```
aws ec2 create-vpc --cidr-block 10.0.0.0/16 --ipv6-cidr-block ipv6-cidr --ipv6-pool pool-id
```

Pour associer un bloc CIDR IPv6 de votre pool d'adresses IPv6 à un VPC, utilisez la commande associate-vpc-cidr-block suivante. Pour laisser Amazon choisir le CIDR IPv6 dans votre pool d'adresses IPv6, omettez l'option `--ipv6-cidr-block`.

```
aws ec2 associate-vpc-cidr-block --vpc-id vpc-123456789abc123ab --ipv6-cidr-block ipv6-cidr
 --ipv6-pool pool-id
```

Pour afficher vos VPC et les informations de pool d'adresses IPv6 associées, utilisez la commande describe-vpcs . Pour afficher des informations sur les blocs CIDR IPv6 associés à partir d'un pool d'adresses IPv6 spécifique, utilisez la commande get-associated-ipv6-pool-cidrs suivante.

```
aws ec2 get-associated-ipv6-pool-cidrs --pool-id pool-id
```

Si vous dissociez le bloc CIDR IPv6 de votre VPC, il est libéré dans votre pool d'adresses IPv6.

Pour de plus amples informations sur l'utilisation de blocs CIDR IPv6 dans la console VPC, veuillez consulter Utilisation de VPC et de sous-réseaux dans le Amazon VPC Guide de l'utilisateur.

Mise hors service de la plage d'adresses

Pour arrêter d'utiliser votre plage d'adresses avec AWS, libérez d'abord toutes les adresses IP Elastic et dissociez les blocs CIDR IPv6 qui sont toujours alloués depuis le pool d'adresses. Ensuite, arrêtez la publicité de la plage d'adresses et enfin, mettez hors service la plage d'adresses.

Vous ne pouvez pas mettre hors service une partie de la plage d'adresses. Si vous souhaitez utiliser une plage d'adresses plus spécifique avec AWS, mettez hors service toute la plage d'adresses et mettez en service une plage d'adresses plus spécifique.

Pour libérer chaque adresse IP Elastic, utilisez la commande release-address suivante.

```
aws ec2 release-address --allocation-id eipalloc-12345678abcabcabc
```

(IPv6) Pour dissocier un bloc CIDR IPv6, utilisez la commande disassociate-vpc-cidr-block suivante.

```
aws ec2 disassociate-vpc-cidr-block --association-id vpc-cidr-assoc-12345abcd1234abc1
```

Pour arrêter la publication de la plage d'adresses, utilisez la commande withdraw-byoip-cidr suivante.

```
aws ec2 withdraw-byoip-cidr --cidr address-range
```

Pour annuler l'allocation de la plage d'adresses, utilisez la commande deprovision-byoip-cidr suivante.

```
aws ec2 deprovision-byoip-cidr --cidr address-range
```

La mise hors service d'une plage d'adresses peut prendre jusqu'à un jour.

Adresses IP Elastic

Une adresse IP Elastic est une adresse IP IPv4 statique conçue pour le cloud computing. En utilisant une adresse IP Elastic, vous pouvez contourner un problème de défaillance d'une instance ou d'un logiciel en remappant rapidement l'adresse à une autre instance de votre compte. Une adresse IP Elastic est allouée à votre compte AWS et vous appartient jusqu'à ce que vous la libériez.

Une adresse IP Elastic est une adresse IP IPv4 publique, qui est accessible depuis Internet. Si votre instance ne dispose pas d'une adresse IPv4 publique, vous pouvez lui associer une adresse IP Elastic pour établir la communication avec Internet. Cela vous permet, par exemple, de vous connecter à l'instance à partir de votre ordinateur local.

Actuellement, nous ne prenons pas en charge les adresses IP Elastic pour IPv6.

Sommaire

- Principes de base d'une adresse IP Elastic (p. 831)
- Utilisation d'adresses IP Elastic (p. 832)
- Utilisation d'enregistrements DNS inverse pour les applications de messagerie (p. 839)
- Limite appliquée aux adresses IP Elastic (p. 839)

Principes de base d'une adresse IP Elastic

Les caractéristiques de base d'une adresse IP Elastic sont les suivantes :

- Une adresse IP Elastic est statique ; elle ne change pas au fil du temps.
- Pour utiliser une adresse IP Elastic, commencez par en attribuer une à votre compte, puis associez-la à votre instance ou à une interface réseau.
- Lorsque vous associez une adresse IP Elastic à une instance, elle est également associée à l'interface réseau principale de l'instance. Lorsque vous associez une adresse IP Elastic à une interface réseau attachée à une instance, elle est également associée à l'instance.
- Lorsque vous associez une adresse IP Elastic à une instance ou à son interface réseau principale, l'adresse IPv4 publique de l'instance (si elle en avait une) est réintégrée dans le pool d'adresses IPv4 publiques d'Amazon. Vous ne pouvez pas réutiliser une adresse IPv4 publique ou la convertir en adresse IP Elastic. Pour plus d'informations, consultez Adresses IPv4 publiques et noms d'hôte DNS externes (p. 809).
- Vous pouvez dissocier une adresse IP Elastic d'une ressource et la réassocier à une autre ressource. Pour éviter un comportement inattendu, assurez-vous que toutes les connexions actives à la ressource nommée dans l'association existante sont fermées avant d'effectuer la modification. Une fois que vous avez associé votre adresse IP Elastic à une ressource différente, vous pouvez rouvrir vos connexions à la ressource nouvellement associée.
- Une adresse IP Elastic dissociée demeure attribuée à votre compte jusqu'à ce que vous la libériez explicitement.
- Pour garantir une utilisation efficace des adresses IP Elastic, nous imposons des frais horaires minimes si une adresse IP Elastic n'est pas associée à une instance en cours d'exécution, ou si elle est associée à une instance arrêtée ou à une interface réseau détachée. Pendant que votre instance s'exécute, une adresse IP Elastic associée à l'instance ne vous sera pas facturée, contrairement à toute adresse IP Elastic supplémentaire associée à cette instance. Pour de plus amples informations, veuillez consulter la section relative aux adresses IP Elastic de la page Tarification Amazon EC2, Tarification à la demande.
- Lorsque vous associez une adresse IP Elastic à une instance qui avait une adresse IPv4 publique, le nom d'hôte DNS public de l'instance est mis à jour pour correspondre à l'adresse IP Elastic.
- Nous résolvons un nom d'hôte DNS public en adresse IPv4 publique ou en l'adresse IP Elastic de l'instance en dehors du réseau de cette dernière et nous la résolvons en adresse IPv4 privée de l'instance depuis le réseau de cette dernière.
- Une adresse IP Elastic provient du groupe d'adresses IPv4 publiques d'Amazon, ou d'un groupe d'adresses IP personnalisé que vous avez importé dans votre compte AWS.
- Lorsque vous allouez une adresse IP Elastic à partir d'un groupe d'adresses IP que vous avez importé dans votre compte AWS, elle n'est pas prise en compte dans votre limite d'adresses IP Elastic. Pour plus d'informations, consultez Limite appliquée aux adresses IP Elastic (p. 839).
- Lorsque vous allouez les adresses IP Elastic, vous pouvez les associer à un groupe de bordure réseau. C'est l'endroit à partir duquel nous publions le bloc d'adresse CIDR. La définition du groupe de bordure réseau limite le bloc d'adresse CIDR à ce groupe. Si vous ne spécifiez pas le groupe de bordure réseau, nous définissons le groupe de bordure contenant toutes les zones de disponibilité de la région (par exemple, `us-west-2`).
- Une adresse IP Elastic ne peut être utilisée que dans un groupe de frontière de réseau spécifique.
- Une adresse IP Elastic est destinée uniquement à une région spécifique et ne peut pas être déplacée vers une autre région.

Utilisation d'adresses IP Elastic

Les sections suivantes expliquent comment utiliser les adresses IP Elastic.

Tâches

- Allocation d'une adresse IP Elastic (p. 833)
- Description de vos adresses IP Elastic (p. 834)
- Balisage d'une adresse IP Elastic (p. 834)

- Association d'une adresse IP Elastic à une instance en cours d'exécution ou à une interface réseau (p. 836)
- Dissociation d'une adresse IP Elastic (p. 837)
- Libération d'une adresse IP Elastic (p. 837)
- Récupération d'une adresse IP Elastic (p. 838)

Allocation d'une adresse IP Elastic

Vous pouvez allouer une adresse IP Elastic à partir du groupe d'adresses IPv4 publiques d'Amazon, ou à partir d'un groupe d'adresses IP personnalisé que vous avez importé dans votre compte AWS. Pour plus d'informations sur l'importation de votre propre plage d'adresses IP dans votre compte AWS, consultez Fourniture de vos propres adresses IP (BYOIP) dans Amazon EC2 (p. 825).

Vous pouvez allouer une adresse IP Elastic à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour allouer une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Elastic IPs.
3. Choisissez Allocate Elastic IP address (Allouer l'adresse IP Elastic).
4. Pour Scope (Champ d'application), choisissez VPC ou EC2-Classic en fonction du champ d'application dans lequel il sera utilisé.
5. (VPC uniquement) Pour le groupe d'adresses IPv4 public, choisissez l'une des options suivantes :
   - Amazon's pool of IP addresses (Groupes d'adresses IP d'Amazon) — À utiliser si vous souhaitez qu'une adresse IPv4 soit allouée à partir du groupe d'adresses IP d'Amazon.
   - My pool of public IPv4 addresses (Mon groupe d'adresses IPv4 publiques) — À utiliser si vous souhaitez allouer une adresse IPv4 à partir d'un groupe d'adresses IP que vous avez ajouté à votre compte AWS. Cette option est désactivée si vous ne disposez pas de groupes d'adresses IP.
   - Pool d'adresses IPv4 appartenant au client—Si vous souhaitez allouer une adresse IPv4 depuis un pool créé à partir de votre réseau local pour une utilisation avec un AWS Outpost. Cette option est désactivée si vous n'avez pas d'AWS Outpost.
6. Choisissez Allocate.

Ancienne console

Pour allouer une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Elastic IPs.
3. Choisissez Allouer une nouvelle adresse.
4. Pour IPv4 address pool (Groupe d'adresses IPv4), choisissez Amazon pool (Groupe Amazon).
5. Choisissez Allouer et fermez l'écran de confirmation.

AWS CLI

Pour allouer une adresse IP Elastic

Utilisez la commande allocate-address de l'AWS CLI.

PowerShell

Pour allouer une adresse IP Elastic

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Address.

Description de vos adresses IP Elastic

Vous pouvez décrire une adresse IP Elastic à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour décrire vos adresses IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic à afficher, puis choisissez Actions, View details (Afficher les détails).

Ancienne console

Pour décrire vos adresses IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez un filtre dans la liste Resource Attribute pour commencer la recherche. Vous pouvez utiliser plusieurs filtres dans une seule recherche.

AWS CLI

Pour décrire vos adresses IP Elastic

Utilisez la commande describe-addresses de l'AWS CLI.

PowerShell

Pour décrire vos adresses IP Elastic

Utilisez la commande Outils AWS pour Windows PowerShell Get-EC2Address.

Balisage d'une adresse IP Elastic

Vous pouvez allouer des balises personnalisées à vos adresses IP Elastic pour classer celles-ci de différentes façons, par exemple, par objectif, par propriétaire ou par environnement. Cela vous aide à retrouver rapidement une adresse IP Elastic spécifique en fonction des balises personnalisées que vous lui avez attribuées.

Vous pouvez uniquement baliser les adresses IP Elastic qui se trouvent dans le champ d'application du VPC.

Note

Le suivi d'allocation des coûts à l'aide des balises d'adresse IP Elastic n'est pas pris en charge.

Vous pouvez baliser une adresse IP Elastic à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour baliser une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic à baliser, puis choisissez Actions, View details (Afficher les détails).
4. Dans la section Tags (Balises) choisissez Manage tags (Gérer les balises).
5. Spécifiez une paire de clé et de valeur de balise.
6. (Facultatif) Choisissez Add tag (Ajouter une balise) pour ajouter des balises supplémentaires.
7. Choisissez Enregistrer.

Ancienne console

Pour baliser une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic à baliser et choisissez Balises.
4. Sélectionnez Ajouter/Modifier des balises.
5. Dans la boîte de dialogue Ajouter/Modifier des balises, choisissez Créer une balise, puis spécifiez la clé et la valeur de la balise.
6. (Facultatif) Choisissez Créer une balise pour ajouter des balises supplémentaires à l'adresse IP Elastic.
7. Choisissez Enregistrer.

AWS CLI

Pour baliser une adresse IP Elastic

Utilisez la commande create-tags de l'AWS CLI.

```
aws ec2 create-tags --resources eipalloc-12345678 --tags Key=Owner,Value=TeamA
```

PowerShell

Pour baliser une adresse IP Elastic

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Tag.

La commande `New-EC2Tag` nécessite un paramètre `Tag`, qui spécifie la paire clé-valeur à utiliser pour la balise d'adresse IP Elastic. Les commandes suivantes créent le paramètre `Tag` :

```
PS C:\> $tag = New-Object Amazon.EC2.Model.Tag
PS C:\> $tag.Key = "Owner"
PS C:\> $tag.Value = "TeamA"
```

```
PS C:\> New-EC2Tag -Resource eipalloc-12345678 -Tag $tag
```

Association d'une adresse IP Elastic à une instance en cours d'exécution ou à une interface réseau

Si vous associez une adresse IP Elastic à votre instance pour permettre la communication avec Internet, vous devez également vous assurer que votre instance se trouve dans un sous-réseau public. Pour plus d'informations, consultez Passerelles Internet dans le Amazon VPC Guide de l'utilisateur.

Vous pouvez associer une adresse IP Elastic à une instance ou à une interface réseau à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour associer une adresse IP Elastic à une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic à associer, puis choisissez Actions, Associate Elastic IP address (Associer l'adresse IP Elastic).
4. Pour Resource type (Type de ressource), choisissez Instance.
5. Par exemple, choisissez l'instance à laquelle vous souhaitez associer l'adresse IP Elastic. Vous pouvez également entrer du texte pour rechercher une instance spécifique.
6. (Facultatif) Pour Private IP address (Adresse IP privée), spécifiez une adresse IP privée à laquelle associer l'adresse IP Elastic.
7. Choisissez Associate.

Pour associer une adresse IP Elastic à une interface réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic à associer, puis choisissez Actions, Associate Elastic IP address (Associer l'adresse IP Elastic).
4. Pour Type de ressource, choisissez Interface réseau.
5. Dans Network interface (Interface réseau), choisissez l'interface réseau à laquelle associer l'adresse IP Elastic. Vous pouvez également entrer du texte pour rechercher une interface réseau spécifique.
6. (Facultatif) Pour Private IP address (Adresse IP privée), spécifiez une adresse IP privée à laquelle associer l'adresse IP Elastic.
7. Choisissez Associate.

Ancienne console

Pour associer une adresse IP Elastic à une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez une adresse IP Elastic et choisissez Actions, puis Associer l'adresse.
4. Sélectionnez l'instance dans Instance, puis choisissez Associer.

AWS CLI

Pour associer une adresse IP Elastic

Utilisez la commande associate-address de l'AWS CLI.

PowerShell

Pour associer une adresse IP Elastic

Utilisez la commande Outils AWS pour Windows PowerShell Register-EC2Address.

Dissociation d'une adresse IP Elastic

Vous pouvez dissocier une adresse IP Elastic d'une instance ou d'une interface réseau à tout moment. Après avoir dissocié l'adresse IP Elastic, vous pouvez la réassocier à une autre ressource.

Vous pouvez dissocier une adresse IP Elastic à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour dissocier et réassocier une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic à dissocier, puis choisissez Actions, Disassociate Elastic IP address (Dissocier l'adresse IP Elastic).
4. Choisissez Dissocier.

Ancienne console

Pour dissocier et réassocier une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic, choisissez Actions, puis sélectionnez Dissocier l'adresse.
4. Choisissez Dissocier l'adresse.

AWS CLI

Dissocier une adresse IP Elastic

Utilisez la commande disassociate-address de l'AWS CLI.

PowerShell

Dissocier une adresse IP Elastic

Utilisez la commande Outils AWS pour Windows PowerShell Unregister-EC2Address.

Libération d'une adresse IP Elastic

Si vous n'avez plus besoin d'une adresse IP Elastic, nous vous recommandons de la libérer via l'une des méthodes suivantes. L'adresse à publier ne doit pas être actuellement associée à une ressource AWS, telle qu'une instance EC2, une passerelle NAT ou Équilibreur de charge du réseau.

Nouvelle console

Libérer une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le volet de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic à libérer, puis choisissez Actions, Release Elastic IP addresses (Libérer des adresses IP Elastic).
4. Choisissez Release (Libérer).

Ancienne console

Libérer une adresse IP Elastic

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic, choisissez Actions, puis sélectionnez Libérer des adresses. Choisissez Libérer lorsque vous y êtes invité.

AWS CLI

Libérer une adresse IP Elastic

Utilisez la commande release-address de l'AWS CLI.

PowerShell

Libérer une adresse IP Elastic

Utilisez la commande Outils AWS pour Windows PowerShell Remove-EC2Address.

Récupération d'une adresse IP Elastic

Si vous avez libéré votre adresse IP Elastic, vous pouvez essayer de la récupérer. Les règles suivantes s'appliquent :

- Vous ne pouvez pas récupérer une adresse IP Elastic si celle-ci a été allouée à un autre compte AWS ou si cela risque d'entraîner un dépassement de votre limite d'adresses IP Elastic.
- Vous ne pouvez pas récupérer les balises associées à une adresse IP Elastic.
- Il n'est possible de récupérer une adresse IP Elastic qu'à l'aide de l'API Amazon EC2 ou d'un outil de ligne de commande.

AWS CLI

Pour récupérer une adresse IP Elastic

Utilisez la commande allocate-address de l'AWS CLI et précisez l'adresse IP à l'aide du paramètre `--address` comme suit.

```
aws ec2 allocate-address --domain vpc --address 203.0.113.3
```

PowerShell

Pour récupérer une adresse IP Elastic

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2Address et précisez l'adresse IP à l'aide du paramètre `-Address` comme suit.

```
PS C:\> New-EC2Address -Address 203.0.113.3 -Domain vpc -Region us-east-1
```

Utilisation d'enregistrements DNS inverse pour les applications de messagerie

Si vous prévoyez d'envoyer des e-mails à des tiers à partir d'une instance, nous vous conseillons de mettre en service une ou plusieurs adresses IP Elastic et de les fournir à AWS. AWS collabore avec des fournisseurs d'accès Internet et des organisations de lutte contre les courriers indésirables (anti-spam) pour réduire le risque que vos e-mails envoyés à partir de ces adresses soient signalés comme du spam.

De plus, l'affectation d'un enregistrement DNS inverse statique à l'adresse IP Elastic utilisée pour envoyer des e-mails contribue à éviter que ces derniers soient considérés comme du spam par les organisations de lutte anti-spam. Notez que l'enregistrement DNS avant correspondant (enregistrement de type A) pointant vers votre adresse IP Elastic doit exister avant que vous créiez votre enregistrement DNS inverse.

Si un enregistrement DNS inverse est associé à une adresse IP Elastic, cette dernière et verrouillée pour votre compte et ne peut pas être libérée tant que l'enregistrement n'est pas supprimé.

Pour supprimer des limites d'envoi d'e-mail ou nous fournir vos adresses IP Elastic et vos enregistrements DNS inverses, accédez à la page Demande de suppression des limites d'envoi d'e-mail.

Limite appliquée aux adresses IP Elastic

Par défaut, tous les comptes AWS sont limités à cinq (5) adresses IP Elastic par région, car les adresses Internet publiques (IPv4) sont des ressources publiques rares. Nous vous encourageons vivement à utiliser une adresse IP Elastic principalement pour pouvoir remapper l'adresse à une autre instance dans le cas de la défaillance d'une instance et d'utiliser des noms d'hôte DNS pour toutes les autres communications internœuds.

Pour vérifier le nombre d'adresses IP Elastic utilisées

Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/ et choisissez Elastic IPs (Adresses IP Elastic) dans le panneau de navigation.

Pour vérifier votre limite de compte actuelle pour les adresses IP Elastic

Vous pouvez vérifier votre limite dans la console Amazon EC2 ou dans la console Quotas de service. Effectuez l'une des actions suivantes :

- Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

  Choisissez Limits (Limites) dans le panneau de navigation, puis entrez **`IP`** dans le champ de recherche. La limite est EC2-VPC Elastic IPs (Adresses IP Elastic EC2-VPC). Si vous avez accès à EC2-Classic, il existe une limite supplémentaire, EC2-Classic Elastic IPs (Adresses IP Elastic EC2-Classic).
- Ouvrez la console Quotas de service à l'adresse https://console.aws.amazon.com/servicequotas/.

  Sur le tableau de bord, choisissez Amazon Elastic Compute Cloud (Amazon EC2). Si Amazon Elastic Compute Cloud (Amazon EC2) n'est pas répertorié dans le tableau de bord, choisissez Services AWS, saisissez **`EC2`** dans le champ de recherche, puis choisissez Amazon Elastic Compute Cloud (Amazon EC2).

  Sur la page des quotas du service Amazon EC2, entrez **`IP`** dans le champ de recherche. La limite est EC2-VPC Elastic IPs (Adresses IP Elastic EC2-VPC). Si vous avez accès à EC2-Classic, il existe une limite supplémentaire, EC2-Classic Elastic IPs (Adresses IP Elastic EC2-Classic). Pour de plus amples informations, choisissez la limite.

Si vous pensez que votre architecture justifie l'utilisation d'adresses IP Elastic supplémentaires, vous pouvez demander une augmentation de quota directement à partir de la console Quotas de service.

Interfaces réseau Elastic

Une interface réseau élastique est un composant réseau logique dans un VPC qui représente une carte réseau virtuelle. Elle peut inclure les attributs suivants :

- Une adresse IPv4 privée principale de la plage d'adresses IPv4 de votre VPC.
- Une ou plusieurs adresses IPv4 privées secondaires de la plage d'adresses IPv4 de votre VPC.
- Une adresse IP Elastic (IPv4) par adresse IPv4 privée
- Une adresse IPv4 publique
- Une ou plusieurs adresses IPv6
- Un ou plusieurs groupes de sécurité
- Une adresse MAC
- Un indicateur de vérification de source/destination
- Une description

Vous pouvez créer et configurer des interfaces réseau dans votre compte et les attacher à des instances de votre VPC. Votre compte peut également compter des interfaces réseau requester-managed gérées par le demandeur, qui sont créées et gérées par les services AWS afin de vous permettre d'utiliser d'autres ressources et services. Vous ne pouvez pas gérer ces interfaces réseau vous-même. Pour plus d'informations, consultez Interfaces réseau gérées par demandeur (p. 863).

Cette ressource AWS est appelée interface réseau dans AWS Management Console et l'API Amazon EC2. Par conséquent, nous utilisons « interface réseau » dans cette documentation au lieu d'indiquer « interface réseau élastique ». L'expression « interface réseau » dans cette documentation signifie toujours « interface réseau élastique ».

Sommaire

- Notions fondamentales concernant l'interface réseau (p. 840)
- Adresses IP par interface réseau et par type d'instance (p. 841)
- Utilisation des interfaces réseau (p. 853)
- Scénarios pour les interfaces réseau (p. 859)
- Meilleures pratiques pour la configuration des interfaces réseau (p. 861)
- Interfaces réseau gérées par demandeur (p. 863)

Notions fondamentales concernant l'interface réseau

Vous pouvez créer une interface réseau, l'attacher à une instance, la détacher d'une instance et l'attacher à une autre instance. Les attributs d'une interface réseau la suivent lorsque celle-ci est attachée à une instance ou détachée d'une instance, puis rattachée à une autre instance. Lorsque vous déplacez une interface réseau d'une instance vers une autre, le trafic réseau est redirigé vers la nouvelle instance.

Interface réseau principale

Chaque instance a une interface réseau par défaut, appelée l'interface réseau principale. Vous ne pouvez pas détacher une interface réseau principale d'une instance. Vous pouvez créer et attacher des Network Interfaces supplémentaires. Le nombre maximal d'interfaces réseau que vous pouvez utiliser varie en fonction du type d'instance. Pour plus d'informations, consultez Adresses IP par interface réseau et par type d'instance (p. 841).

Adresses IPv4 publiques pour les interfaces réseau

Dans un VPC, tous les sous-réseaux ont un attribut modifiable qui détermine si les interfaces réseau créées dans ce sous-réseau (et, par conséquent, les instances lancées dans ce sous-réseau) sont attribuées à une adresse IPv4 publique. Pour de plus amples informations, veuillez consulter Comportement de l'adressage IP public de votre sous-réseau dans le Amazon VPC Guide de l'utilisateur. L'adresse IPv4 publique est attribuée à partir du pool d'adresses IPv4 publiques d'Amazon. Lorsque vous lancez une instance, l'adresse IP est attribuée à l'interface réseau principale qui est créée.

Lorsque vous créez une interface réseau, elle hérite l'attribut d'adressage IPv4 public du sous-réseau. Si vous modifiez par la suite l'attribut d'adressage IPv4 public du sous-réseau, l'interface réseau conserve le paramètre qui était en vigueur lorsqu'elle a été créée. Si vous lancez une instance et spécifiez une interface réseau existante comme interface réseau principale, l'attribut d'adresse IPv4 publique est déterminé par cette interface réseau.

Pour plus d'informations, consultez Adresses IPv4 publiques et noms d'hôte DNS externes (p. 809).

Adresses IP Elastic pour l'interface réseau

Si vous disposez d'une adresse IP Elastic, vous pouvez l'associer à l'une des adresses IPv4 privées de l'interface réseau. Vous pouvez associer une adresse IP Elastic à chaque adresse IPv4 privée.

Si vous dissociez une adresse IP Elastic d'une interface réseau, vous pouvez la libérer dans le pool d'adresses. C'est la seule façon d'associer une adresse IP Elastic à une instance d'un sous-réseau ou d'un VPC différent, car les interfaces réseau sont spécifiques aux sous-réseaux.

Adresses IPv6 pour les interfaces réseau

Si vous associez des blocs d'adresse CIDR IPv6 à votre VPC et à votre sous-réseau, vous pouvez attribuer une ou plusieurs adresses IPv6 de la plage du sous-réseau à une interface réseau. Chaque adresse IPv6 peut être attribuée à une interface réseau.

Tous les sous-réseaux ont un attribut modifiable qui détermine si les interfaces réseau créées dans ce sous-réseau (et, par conséquent, les instances lancées dans ce sous-réseau) reçoivent automatiquement une adresse IPv6 de la plage du sous-réseau. Pour de plus amples informations, veuillez consulter Comportement de l'adressage IP public de votre sous-réseau dans le Amazon VPC Guide de l'utilisateur. Lorsque vous lancez une instance, l'adresse IPv6 est attribuée à l'interface réseau principale qui est créée.

Pour plus d'informations, consultez Adresses IPv6 (p. 810).

Comportement de résiliation

Vous pouvez définir le comportement de résiliation d'une interface réseau attachée à une instance. Vous pouvez spécifier si l'interface réseau doit être supprimée automatiquement lorsque vous résiliez l'instance à laquelle celle-ci est attachée.

Vérification origine/destination

La désactivation de la vérification origine/destination permet à une instance de traiter le trafic réseau qui ne lui est pas destiné spécifiquement. Par exemple, les instances qui exécutent des services comme la traduction d'adresses réseau, le routage ou un pare-feu doivent désactiver l'attribut de contrôle source/destination. Cet attribut est activé par défaut.

Surveillance du trafic IP

Vous pouvez activer un journal de flux VPC sur votre interface réseau pour capturer des informations sur le trafic IP circulant vers et depuis l'interface réseau. Une fois que vous avez créé un journal de flux, vous pouvez afficher et extraire ses données dans Amazon CloudWatch Logs. Pour plus d'informations, consultez Journaux de flux VPC dans le Amazon VPC Guide de l'utilisateur.

Adresses IP par interface réseau et par type d'instance

Le tableau suivant répertorie le nombre maximal d'interfaces réseau par type d'instance et le nombre maximal d'adresses IPv4 privées et d'adresses IPv6 par interface réseau. La limite pour les adresses IPv6

est distincte de la limite pour les adresses IPv4 privées par interface réseau. Certains types d'instance ne prennent pas en charge l'adressage IPv6. Les interfaces réseau, les adresses IPv4 privées multiples et les adresses IPv6 ne sont disponibles que pour les instances qui s'exécutent dans un VPC. Les adresses IPv6 sont publiques et accessibles via Internet. Pour plus d'informations, consultez Plusieurs adresses IP (p. 817). Pour de plus amples informations sur IPv6 dans un VPC, veuillez consulter Adressage IP de votre VPC dans le Amazon VPC Guide de l'utilisateur.

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`a1.medium`	2	4	4
`a1.large`	3	10	10
`a1.xlarge`	4	15	15
`a1.2xlarge`	4	15	15
`a1.4xlarge`	8	30	30
`a1.metal`	8	30	30
`c1.medium`	2	6	IPv6 non pris en charge
`c1.xlarge`	4	15	IPv6 non pris en charge
`c3.large`	3	10	10
`c3.xlarge`	4	15	15
`c3.2xlarge`	4	15	15
`c3.4xlarge`	8	30	30
`c3.8xlarge`	8	30	30
`c4.large`	3	10	10
`c4.xlarge`	4	15	15
`c4.2xlarge`	4	15	15
`c4.4xlarge`	8	30	30
`c4.8xlarge`	8	30	30
`c5.large`	3	10	10
`c5.xlarge`	4	15	15
`c5.2xlarge`	4	15	15
`c5.4xlarge`	8	30	30
`c5.9xlarge`	8	30	30
`c5.12xlarge`	8	30	30
`c5.18xlarge`	15	50	50
`c5.24xlarge`	15	50	50

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`c5.metal`	15	50	50
`c5a.large`	3	10	10
`c5a.xlarge`	4	15	15
`c5a.2xlarge`	4	15	15
`c5a.4xlarge`	8	30	30
`c5a.8xlarge`	8	30	30
`c5a.12xlarge`	8	30	30
`c5a.16xlarge`	15	50	50
`c5a.24xlarge`	15	50	50
`c5ad.large`	3	10	10
`c5ad.xlarge`	4	15	15
`c5ad.2xlarge`	4	15	15
`c5ad.4xlarge`	8	30	30
`c5ad.8xlarge`	8	30	30
`c5ad.12xlarge`	8	30	30
`c5ad.16xlarge`	15	50	50
`c5ad.24xlarge`	15	50	50
`c5d.large`	3	10	10
`c5d.xlarge`	4	15	15
`c5d.2xlarge`	4	15	15
`c5d.4xlarge`	8	30	30
`c5d.9xlarge`	8	30	30
`c5d.12xlarge`	8	30	30
`c5d.18xlarge`	15	50	50
`c5d.24xlarge`	15	50	50
`c5d.metal`	15	50	50
`c5n.large`	3	10	10
`c5n.xlarge`	4	15	15
`c5n.2xlarge`	4	15	15
`c5n.4xlarge`	8	30	30

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`c5n.9xlarge`	8	30	30
`c5n.18xlarge`	15	50	50
`c5n.metal`	15	50	50
`c6g.medium`	2	4	4
`c6g.large`	3	10	10
`c6g.xlarge`	4	15	15
`c6g.2xlarge`	4	15	15
`c6g.4xlarge`	8	30	30
`c6g.8xlarge`	8	30	30
`c6g.12xlarge`	8	30	30
`c6g.16xlarge`	15	50	50
`c6g.metal`	15	50	50
`c6gd.medium`	2	4	4
`c6gd.large`	3	10	10
`c6gd.xlarge`	4	15	15
`c6gd.2xlarge`	4	15	15
`c6gd.4xlarge`	8	30	30
`c6gd.8xlarge`	8	30	30
`c6gd.12xlarge`	8	30	30
`c6gd.16xlarge`	15	50	50
`c6gd.metal`	15	50	50
`cc2.8xlarge`	8	30	IPv6 non pris en charge
`cr1.8xlarge`	8	30	IPv6 non pris en charge
`d2.xlarge`	4	15	15
`d2.2xlarge`	4	15	15
`d2.4xlarge`	8	30	30
`d2.8xlarge`	8	30	30
`f1.2xlarge`	4	15	15
`f1.4xlarge`	8	30	30
`f1.16xlarge`	8	50	50

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`g2.2xlarge`	4	15	IPv6 non pris en charge
`g2.8xlarge`	8	30	IPv6 non pris en charge
`g3s.xlarge`	4	15	15
`g3.4xlarge`	8	30	30
`g3.8xlarge`	8	30	30
`g3.16xlarge`	15	50	50
`g4dn.xlarge`	3	10	10
`g4dn.2xlarge`	3	10	10
`g4dn.4xlarge`	3	10	10
`g4dn.8xlarge`	4	15	15
`g4dn.12xlarge`	8	30	30
`g4dn.16xlarge`	4	15	15
`g4dn.metal`	15	50	50
`h1.2xlarge`	4	15	15
`h1.4xlarge`	8	30	30
`h1.8xlarge`	8	30	30
`h1.16xlarge`	15	50	50
`hs1.8xlarge`	8	30	IPv6 non pris en charge
`i2.xlarge`	4	15	15
`i2.2xlarge`	4	15	15
`i2.4xlarge`	8	30	30
`i2.8xlarge`	8	30	30
`i3.large`	3	10	10
`i3.xlarge`	4	15	15
`i3.2xlarge`	4	15	15
`i3.4xlarge`	8	30	30
`i3.8xlarge`	8	30	30
`i3.16xlarge`	15	50	50
`i3.metal`	15	50	50
`i3en.large`	3	10	10

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`i3en.xlarge`	4	15	15
`i3en.2xlarge`	4	15	15
`i3en.3xlarge`	4	15	15
`i3en.6xlarge`	8	30	30
`i3en.12xlarge`	8	30	30
`i3en.24xlarge`	15	50	50
`i3en.metal`	15	50	50
`inf1.xlarge`	4	10	10
`inf1.2xlarge`	4	10	10
`inf1.6xlarge`	8	30	30
`inf1.24xlarge`	15	30	30
`m1.small`	2	4	IPv6 non pris en charge
`m1.medium`	2	6	IPv6 non pris en charge
`m1.large`	3	10	IPv6 non pris en charge
`m1.xlarge`	4	15	IPv6 non pris en charge
`m2.xlarge`	4	15	IPv6 non pris en charge
`m2.2xlarge`	4	30	IPv6 non pris en charge
`m2.4xlarge`	8	30	IPv6 non pris en charge
`m3.medium`	2	6	IPv6 non pris en charge
`m3.large`	3	10	IPv6 non pris en charge
`m3.xlarge`	4	15	IPv6 non pris en charge
`m3.2xlarge`	4	30	IPv6 non pris en charge
`m4.large`	2	10	10
`m4.xlarge`	4	15	15
`m4.2xlarge`	4	15	15
`m4.4xlarge`	8	30	30
`m4.10xlarge`	8	30	30
`m4.16xlarge`	8	30	30
`m5.large`	3	10	10
`m5.xlarge`	4	15	15

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`m5.2xlarge`	4	15	15
`m5.4xlarge`	8	30	30
`m5.8xlarge`	8	30	30
`m5.12xlarge`	8	30	30
`m5.16xlarge`	15	50	50
`m5.24xlarge`	15	50	50
`m5.metal`	15	50	50
`m5a.large`	3	10	10
`m5a.xlarge`	4	15	15
`m5a.2xlarge`	4	15	15
`m5a.4xlarge`	8	30	30
`m5a.8xlarge`	8	30	30
`m5a.12xlarge`	8	30	30
`m5a.16xlarge`	15	50	50
`m5a.24xlarge`	15	50	50
`m5ad.large`	3	10	10
`m5ad.xlarge`	4	15	15
`m5ad.2xlarge`	4	15	15
`m5ad.4xlarge`	8	30	30
`m5ad.8xlarge`	8	30	30
`m5ad.12xlarge`	8	30	30
`m5ad.16xlarge`	15	50	50
`m5ad.24xlarge`	15	50	50
`m5d.large`	3	10	10
`m5d.xlarge`	4	15	15
`m5d.2xlarge`	4	15	15
`m5d.4xlarge`	8	30	30
`m5d.8xlarge`	8	30	30
`m5d.12xlarge`	8	30	30
`m5d.16xlarge`	15	50	50

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`m5d.24xlarge`	15	50	50
`m5d.metal`	15	50	50
`m5dn.large`	3	10	10
`m5dn.xlarge`	4	15	15
`m5dn.2xlarge`	4	15	15
`m5dn.4xlarge`	8	30	30
`m5dn.8xlarge`	8	30	30
`m5dn.12xlarge`	8	30	30
`m5dn.16xlarge`	15	50	50
`m5dn.24xlarge`	15	50	50
`m5n.large`	3	10	10
`m5n.xlarge`	4	15	15
`m5n.2xlarge`	4	15	15
`m5n.4xlarge`	8	30	30
`m5n.8xlarge`	8	30	30
`m5n.12xlarge`	8	30	30
`m5n.16xlarge`	15	50	50
`m5n.24xlarge`	15	50	50
`m6g.medium`	2	4	4
`m6g.large`	3	10	10
`m6g.xlarge`	4	15	15
`m6g.2xlarge`	4	15	15
`m6g.4xlarge`	8	30	30
`m6g.8xlarge`	8	30	30
`m6g.12xlarge`	8	30	30
`m6g.16xlarge`	15	50	50
`m6g.metal`	15	50	50
`m6gd.medium`	2	4	4
`m6gd.large`	3	10	10
`m6gd.xlarge`	4	15	15

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`m6gd.2xlarge`	4	15	15
`m6gd.4xlarge`	8	30	30
`m6gd.8xlarge`	8	30	30
`m6gd.12xlarge`	8	30	30
`m6gd.16xlarge`	15	50	50
`m6gd.metal`	15	50	50
`p2.xlarge`	4	15	15
`p2.8xlarge`	8	30	30
`p2.16xlarge`	8	30	30
`p3.2xlarge`	4	15	15
`p3.8xlarge`	8	30	30
`p3.16xlarge`	8	30	30
`p3dn.24xlarge`	15	50	50
`r3.large`	3	10	10
`r3.xlarge`	4	15	15
`r3.2xlarge`	4	15	15
`r3.4xlarge`	8	30	30
`r3.8xlarge`	8	30	30
`r4.large`	3	10	10
`r4.xlarge`	4	15	15
`r4.2xlarge`	4	15	15
`r4.4xlarge`	8	30	30
`r4.8xlarge`	8	30	30
`r4.16xlarge`	15	50	50
`r5.large`	3	10	10
`r5.xlarge`	4	15	15
`r5.2xlarge`	4	15	15
`r5.4xlarge`	8	30	30
`r5.8xlarge`	8	30	30
`r5.12xlarge`	8	30	30

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`r5.16xlarge`	15	50	50
`r5.24xlarge`	15	50	50
`r5.metal`	15	50	50
`r5a.large`	3	10	10
`r5a.xlarge`	4	15	15
`r5a.2xlarge`	4	15	15
`r5a.4xlarge`	8	30	30
`r5a.8xlarge`	8	30	30
`r5a.12xlarge`	8	30	30
`r5a.16xlarge`	15	50	50
`r5a.24xlarge`	15	50	50
`r5ad.large`	3	10	10
`r5ad.xlarge`	4	15	15
`r5ad.2xlarge`	4	15	15
`r5ad.4xlarge`	8	30	30
`r5ad.8xlarge`	8	30	30
`r5ad.12xlarge`	8	30	30
`r5ad.16xlarge`	15	50	50
`r5ad.24xlarge`	15	50	50
`r5d.large`	3	10	10
`r5d.xlarge`	4	15	15
`r5d.2xlarge`	4	15	15
`r5d.4xlarge`	8	30	30
`r5d.8xlarge`	8	30	30
`r5d.12xlarge`	8	30	30
`r5d.16xlarge`	15	50	50
`r5d.24xlarge`	15	50	50
`r5d.metal`	15	50	50
`r5dn.large`	3	10	10
`r5dn.xlarge`	4	15	15

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`r5dn.2xlarge`	4	15	15
`r5dn.4xlarge`	8	30	30
`r5dn.8xlarge`	8	30	30
`r5dn.12xlarge`	8	30	30
`r5dn.16xlarge`	15	50	50
`r5dn.24xlarge`	15	50	50
`r5n.large`	3	10	10
`r5n.xlarge`	4	15	15
`r5n.2xlarge`	4	15	15
`r5n.4xlarge`	8	30	30
`r5n.8xlarge`	8	30	30
`r5n.12xlarge`	8	30	30
`r5n.16xlarge`	15	50	50
`r5n.24xlarge`	15	50	50
`r6g.medium`	2	4	4
`r6g.large`	3	10	10
`r6g.xlarge`	4	15	15
`r6g.2xlarge`	4	15	15
`r6g.4xlarge`	8	30	30
`r6g.8xlarge`	8	30	30
`r6g.12xlarge`	8	30	30
`r6g.16xlarge`	15	50	50
`r6g.metal`	15	50	50
`r6gd.medium`	2	4	4
`r6gd.large`	3	10	10
`r6gd.xlarge`	4	15	15
`r6gd.2xlarge`	4	15	15
`r6gd.4xlarge`	8	30	30
`r6gd.8xlarge`	8	30	30
`r6gd.12xlarge`	8	30	30

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`r6gd.16xlarge`	15	50	50
`r6gd.metal`	15	50	50
`t1.micro`	2	2	IPv6 non pris en charge
`t2.nano`	2	2	2
`t2.micro`	2	2	2
`t2.small`	3	4	4
`t2.medium`	3	6	6
`t2.large`	3	12	12
`t2.xlarge`	3	15	15
`t2.2xlarge`	3	15	15
`t3.nano`	2	2	2
`t3.micro`	2	2	2
`t3.small`	3	4	4
`t3.medium`	3	6	6
`t3.large`	3	12	12
`t3.xlarge`	4	15	15
`t3.2xlarge`	4	15	15
`t3a.nano`	2	2	2
`t3a.micro`	2	2	2
`t3a.small`	2	4	4
`t3a.medium`	3	6	6
`t3a.large`	3	12	12
`t3a.xlarge`	4	15	15
`t3a.2xlarge`	4	15	15
`u-6tb1.metal`	5	30	30
`u-9tb1.metal`	5	30	30
`u-12tb1.metal`	5	30	30
`u-18tb1.metal`	15	50	50
`u-24tb1.metal`	15	50	50
`x1.16xlarge`	8	30	30

Type d'instance	Nombre maximal d'interfaces réseau	Adresses IPv4 privées par interface	Adresses IPv6 par interface
`x1.32xlarge`	8	30	30
`x1e.xlarge`	3	10	10
`x1e.2xlarge`	4	15	15
`x1e.4xlarge`	4	15	15
`x1e.8xlarge`	4	15	15
`x1e.16xlarge`	8	30	30
`x1e.32xlarge`	8	30	30
`z1d.large`	3	10	10
`z1d.xlarge`	4	15	15
`z1d.2xlarge`	4	15	15
`z1d.3xlarge`	8	30	30
`z1d.6xlarge`	8	30	30
`z1d.12xlarge`	15	50	50
`z1d.metal`	15	50	50

Utilisation des interfaces réseau

Vous pouvez travailler avec les interfaces réseau à l'aide de la console Amazon EC2 ou de la ligne de commande.

Sommaire

- Création d'une interface réseau (p. 853)
- Affichage des détails relatifs à une interface réseau (p. 854)
- Attachement d'une interface réseau à une instance (p. 855)
- Détachement d'une interface réseau d'une instance (p. 855)
- Gestion des adresses IP (p. 856)
- Modification des attributs d'interface réseau (p. 857)
- Ajout ou modification de balises (p. 858)
- Suppression d'une interface réseau (p. 859)

Création d'une interface réseau

Vous pouvez créer une interface réseau dans un sous-réseau. Vous ne pouvez pas déplacer une interface réseau vers un autre sous-réseau une fois que celle-ci a été créée et vous pouvez uniquement attacher l'interface à réseau à des instances dans la même zone de disponibilité.

Pour créer une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez Créer une interface réseau.
4. Pour Description, saisissez un nom descriptif.
5. Pour Sous-réseau, sélectionnez le sous-réseau.
6. Pour IP privée (ou IP privée IPv4), entrez l'adresse IPv4 privée principale. Si vous ne spécifiez pas d'adresse IPv4, nous sélectionnons une adresse IPv4 privée disponible dans le sous-réseau sélectionné.
7. (IPv6 uniquement) Si vous avez sélectionné un sous-réseau qui a un bloc d'adresse CIDR IPv6 associé, vous pouvez le cas échéant spécifier une adresse IPv6 dans le champ IP IPv6.
8. Pour créer un Elastic Fabric Adapter (EFA), sélectionnez Elastic Fabric Adapter (EFA).
9. Pour Groupes de sécurité, sélectionnez un ou plusieurs groupes de sécurité.
10. (Facultatif) Choisissez Ajouter une balise et entrez une clé et une valeur de balise.
11. Choisissez Yes, Create.

Pour créer une interface réseau à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-network-interface (AWS CLI)
- New-EC2NetworkInterface (Outils AWS pour Windows PowerShell)

Affichage des détails relatifs à une interface réseau

Vous pouvez afficher toutes les interfaces réseau dans votre compte.

Pour décrire une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau.
4. Pour afficher les détails, choisissez Détails.

Pour décrire une interface réseau à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-network-interfaces (AWS CLI)
- Get-EC2NetworkInterface (Outils AWS pour Windows PowerShell)

Pour décrire un attribut d'interface réseau à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-network-interface-attribute (AWS CLI)
- Get-EC2NetworkInterfaceAttribute (Outils AWS pour Windows PowerShell)

Attachement d'une interface réseau à une instance

Vous pouvez attacher une interface réseau à toute instance arrêtée ou en cours d'exécution en utilisant les pages Instances ou Interfaces réseau de la console Amazon EC2. Vous pouvez également spécifier une interface réseau existante ou attacher une interface réseau supplémentaire lorsque vous lancez une instance (p. 530).

Si l'adresse IPv4 publique de votre instance est libérée, elle n'en reçoit pas de nouvelle si plusieurs interfaces réseau sont attachées à l'instance. Pour plus d'informations sur le comportement des adresses IPv4 publiques, consultez Adresses IPv4 publiques et noms d'hôte DNS externes (p. 809).

Pour attacher une interface réseau à une instance à l'aide de la page Instances

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Sélectionnez Actions, Mise en réseau, Attacher l'interface réseau.
5. Sélectionnez l'interface réseau, puis choisissez Attacher.

Pour attacher une interface réseau à une instance à l'aide de la page Interfaces réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau, puis choisissez Attacher.
4. Sélectionnez l'instance, puis choisissez Attacher.

Pour attacher une interface réseau à une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- attach-network-interface (AWS CLI)
- Add-EC2NetworkInterface (Outils AWS pour Windows PowerShell)

Détachement d'une interface réseau d'une instance

Vous pouvez détacher une interface réseau secondaire à tout moment en utilisant la page Instances ou Interfaces réseau de la console Amazon EC2.

Pour détacher une interface réseau d'une instance à l'aide de la page Instances

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Sélectionnez Actions, Mise en réseau, Détacher l'interface réseau.
5. Sélectionnez l'interface réseau, puis choisissez Détacher.

Vous ne pouvez pas utiliser la console Amazon EC2 pour détacher une interface réseau attachée à une ressource d'un autre service, telle qu'un équilibreur de charge Elastic Load Balancing, une fonction

Lambda, un espace de travail ou une passerelle NAT. Les interfaces réseau de ces ressources seront supprimées lorsque la ressource est supprimée.

Pour détacher une interface réseau d'une instance à l'aide de la page Interfaces réseau

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau et contrôlez la description pour vérifier que l'interface réseau est attachée à une instance et non à un autre type de ressource. Si la ressource est une instance EC2, choisissez Détacher.

   Si l'interface réseau est l'interface réseau principale de l'instance, le bouton Détacher est désactivé.
4. Lorsque vous êtes invité à confirmer l'opération, choisissez Oui, détacher.
5. Si vous ne parvenez pas à détacher l'interface réseau de l'instance, choisissez Forcer le détachement, puis réessayez. Nous vous recommandons de choisir cette option uniquement en dernier recours. Forcer un détachement peut vous empêcher d'attacher une interface réseau différente sur le même index jusqu'à ce que vous redémarriez l'instance. Cela peut également empêcher les métadonnées de l'instance de refléter que l'interface réseau a été détachée jusqu'à ce que vous redémarriez l'instance.

Pour détacher une interface réseau à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- detach-network-interface (AWS CLI)
- Dismount-EC2NetworkInterface (Outils AWS pour Windows PowerShell)

Gestion des adresses IP

Vous pouvez gérer les adresses IP suivantes pour vos interfaces réseau :

- Adresses IP Elastic (une par adresse IPv4 privée)
- Adresses IPv4
- Adresses IPv6

Pour agir sur les adresses IP Elastic d'une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau.
4. Pour associer une adresse IP Elastic, procédez comme suit :

   a. Sélectionnez Actions, Associate Address.
   b. Pour Address (Adresse), sélectionnez l'adresse IP Elastic.
   c. Pour Associer à l'adresse IP privée, sélectionnez l'adresse IPv4 privée à associer à l'adresse IP Elastic.
   d. Sélectionnez Autoriser la réassociation pour permettre à l'adresse IP Elastic d'être associée à l'interface réseau spécifiée si elle est actuellement associée à une autre instance ou interface réseau, puis choisissez Associer l'adresse.
5. Pour dissocier une adresse IP Elastic, procédez comme suit :

   a. Choisissez Actions, Disassociate address (Dissocier l'adresse).

b. Dans la boîte de dialogue Dissocier l'adresse IP, sélectionnez Oui, dissocier.

Pour gérer les adresses IPv4 et IPv6 d'une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau.
4. Choisissez Actions, Gérer les adresses IP.
5. Pour IPv4 Addresses (Adresses IPv4), modifiez les adresses IP selon vos besoins. Pour attribuer une adresse IPv4, choisissez Assign new IP (Attribuer une nouvelle adresse IP), puis spécifiez une adresse IPv4 dans la plage de sous-réseaux ou laissez AWS en choisir une pour vous. Pour annuler l'attribution d'une adresse IPv4, choisissez Unassign (Annuler l'attribution) en regard de l'adresse.
6. Pour IPv6 Addresses (Adresses IPv6), modifiez les adresses IP selon vos besoins. Pour attribuer une adresse IPv6, choisissez Assign new IP (Attribuer une nouvelle adresse IP), puis spécifiez une adresse IPv6 dans la plage de sous-réseaux ou laissez AWS en choisir une pour vous. Pour annuler l'attribution d'une adresse IPv6, choisissez Unassign (Annuler l'attribution) en regard de l'adresse.
7. Choisissez Oui, mettre à jour.

Pour gérer les adresses IP d'une interface réseau à l'aide de l'AWS CLI

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- assign-ipv6-addresses
- associate-address
- disassociate-address
- unassign-ipv6-addresses

Pour gérer les adresses IP d'une interface réseau à l'aide de Outils pour Windows PowerShell

Vous pouvez utiliser l'une des commandes suivantes.

- Register-EC2Address
- Register-EC2Ipv6AddressList
- Unregister-EC2Address
- Unregister-EC2Ipv6AddressList

Modification des attributs d'interface réseau

Vous pouvez modifier les attributs d'interface réseau suivants :

- Description (p. 857)
- Groupes de sécurité (p. 858)
- Supprimer à la résiliation (p. 858)
- Contrôle source/destination (p. 858)

Pour changer la description d'une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau, puis choisissez Actions, Modifier la description.
4. Dans la boîte de dialogue Change Description (Modifier la description), entrez une description pour l'interface réseau, puis choisissez Save (Enregistrer).

Pour changer les groupes de sécurité d'une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau, puis choisissez Actions, Changer les groupes de sécurité.
4. Pour Change Security Groups (Modifier les groupes de sécurité), sélectionnez les groupes de sécurité à utiliser, puis choisissez Save (Enregistrer).

   Le groupe de sécurité et l'interface réseau doivent être créés pour le même VPC. Pour modifier le groupe de sécurité pour les interfaces appartenant à d'autres services, par exemple Elastic Load Balancing, faites-le via ce service.

Pour modifier le comportement de résiliation d'une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau, puis choisissez Actions, Modifier le comportement de résiliation.
4. Dans la boîte de dialogue Modifier le comportement de résiliation, cochez la case Supprimer à la résiliation si vous souhaitez que l'interface réseau soit supprimée lorsque vous résiliez une instance.

Pour changer le contrôle de la source/destination d'une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau, puis choisissez Actions, Changer le contrôle de la source/ dest..
4. Dans la boîte de dialogue, sélectionnez Activé (pour une activation) ou Désactivé (pour une désactivation), puis Enregistrer.

Pour modifier les attributs d'interface réseau à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- modify-network-interface-attribute (AWS CLI)
- Edit-EC2NetworkInterfaceAttribute (Outils AWS pour Windows PowerShell)

Ajout ou modification de balises

Les balises sont les métadonnées que vous pouvez ajouter à une interface réseau. Les balises sont privées et sont uniquement visibles pour votre compte. Chaque balise est constituée d'une clé et d'une valeur facultative. Pour en savoir plus sur les balises, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Pour ajouter ou changer des balises pour une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau.
4. Dans le volet des détails, sous l'onglet Balises, sélectionnez Ajouter/Modifier des balises.
5. Dans la boîte de dialogue Ajouter/Modifier des balises, sélectionnez Créer une balise pour chaque balise à créer, puis entrez une clé et une valeur facultative. Lorsque vous avez terminé, sélectionnez Enregistrer.

Pour ajouter ou changer des balises pour une interface réseau à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-tags (AWS CLI)
- New-EC2Tag (Outils AWS pour Windows PowerShell)

Suppression d'une interface réseau

Pour supprimer une instance, vous devez d'abord détacher l'interface réseau. La suppression d'une interface réseau libère tous les attributs qui lui sont associés, ainsi que toute adresse IP privée ou adresse IP Elastic à utiliser par une autre instance.

Pour supprimer une interface réseau à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez une interface réseau, puis Supprimer.
4. Dans la boîte de dialogue Supprimer l'interface réseau, sélectionnez Oui, supprimer.

Pour supprimer une interface réseau à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- delete-network-interface (AWS CLI)
- Remove-EC2NetworkInterface (Outils AWS pour Windows PowerShell)

Scénarios pour les interfaces réseau

Il peut être utile d'attacher plusieurs Network Interfaces à une instance si vous souhaitez :

- créer un réseau de gestion ;
- utiliser des composants de réseau et de sécurité dans votre VPC ;
- créer des instances à deux interfaces réseau avec des charges de travail/rôles sur des sous-réseaux distincts ;
- créer une solution haute disponibilité à faible coût.

Création d'un réseau de gestion

Vous pouvez créer un réseau de gestion à l'aide des interfaces réseau. Dans ce scénario, comme illustré dans l'image suivante :

- L'interface réseau principale (eth0) sur l'instance gère le trafic public.
- L'interface réseau secondaire (eth1) gère le trafic de gestion du backend et est connectée à un sous-réseau distinct de votre VPC qui dispose de contrôles d'accès plus restrictifs.

L'interface publique, qui peut ou non se trouver derrière un équilibreur de charge, a un groupe de sécurité associé qui autorise l'accès au serveur à partir d'Internet (par exemple, autoriser les ports TCP 80 et 443 depuis `0.0.0.0/0` ou depuis l'équilibreur de charge).

L'interface privée est associée à un groupe de sécurité associé permettant l'accès SSH uniquement à partir d'une plage autorisée d'adresses IP, soit au sein du VPC, soit à partir d'Internet, d'un sous-réseau privé au sein du VPC ou d'une passerelle privée virtuelle.

Pour assurer les fonctions de basculement, pensez à utiliser une adresse IPv4 privée secondaire pour le trafic entrant sur une interface réseau. Dans le cas de la défaillance d'une instance, vous pouvez déplacer l'interface et/ou l'adresse IPv4 privée secondaire vers une instance de secours.

Utilisation de composants de réseau et de sécurité dans votre VPC

Certains composants de réseau et de sécurité tiers, tels que des équilibreurs de charge, des serveurs NAT (Network Address Translation, traduction d'adresses réseau) et des serveurs proxy, doivent, de préférence, être configurés avec plusieurs Network Interfaces. Vous pouvez créer et attacher des Network Interfaces secondaires à des instances d'un VPC qui exécutent ces types d'applications, et configurer les interfaces supplémentaires avec leurs propres adresses IP publiques et privées, leurs propres groupes de sécurité et leur propre source/destination checking.

Création d'instances à deux interfaces réseau avec des charges de travail/rôles sur des sous-réseaux distincts

Vous pouvez placer une interface réseau sur chacun de vos serveurs web qui se connecte à un réseau de niveau intermédiaire où réside un serveur d'applications. Le serveur d'applications peut également avoir deux interfaces réseau sur le réseau principal (sous-réseau) où réside le serveur de base de données. Au lieu d'acheminer des paquets réseau via les instances à deux interfaces réseau, chaque instance à deux interfaces réseau reçoit et traite les demandes sur le serveur frontal, établit une connexion au serveur principal, puis envoie les demandes aux serveurs se trouvant sur le réseau principal.

Création d'une solution haute disponibilité à faible coût

Si l'une de vos instances remplissant une fonction particulière subit une défaillance, son Network Interface peut être attachée à une instance de remplacement ou de hot standby préconfigurée pour le même rôle afin de récupérer rapidement le service. Par exemple, vous pouvez utiliser une interface réseau comme interface réseau principale ou secondaire d'un service critique tel qu'une instance de base de données ou une instance NAT. Si une instance subit une défaillance, vous (ou, plus probablement, le code s'exécutant pour votre compte) pouvez attacher l'interface réseau à une instance de secours à chaud. Comme l'interface conserve ses adresses IP privées, ses adresses IP Elastic et son adresse MAC, le trafic réseau commence à passer vers l'instance de secours dès que vous attachez l'interface réseau à l'instance de remplacement. Les utilisateurs rencontreront une brève perte de connectivité entre le moment où l'instance subit la défaillance et celui où l'interface réseau est attachée à l'instance de secours, mais aucune modification de la table de routage VPC ou de votre serveur DNS n'est requise.

Meilleures pratiques pour la configuration des interfaces réseau

- Vous pouvez attacher une interface réseau à une instance lorsqu'elle est en cours d'exécution (attachement de secours), arrêtée (attachement à chaud) ou en cours de lancement (attachement à froid).
- Vous pouvez détacher les interfaces réseau secondaires lorsque l'instance s'exécute ou est arrêtée. Toutefois, vous ne pouvez pas détacher l'interface réseau principale.
- Vous pouvez déplacer une interface réseau d'une instance à une autre si les instances sont dans la même zone de disponibilité et le même VPC mais dans des sous-réseaux différents.
- Lors du lancement d'une instance à l'aide de l'interface de ligne de commande, de l'API ou d'un kit SDK, vous pouvez spécifier l'interface réseau principale et des interfaces réseau supplémentaires.
- Le lancement d'une instance Amazon Linux ou Windows Server avec plusieurs interfaces réseau configure automatiquement les interfaces, les adresses IPv4 privées et les tables de routage du système d'exploitation de l'instance.
- Un attachement à chaud ou de secours d'une interface réseau supplémentaire peut nécessiter que vous mettiez en place manuellement la deuxième interface, configuriez l'adresse IPv4 privée et modifiez

la table de routage en conséquence. Les instances qui exécutent Amazon Linux ou Windows Server reconnaissent automatiquement l'attachement à chaud ou de secours et se configurent elles-mêmes.

- L'attachement d'une autre interface réseau à une instance (par exemple, une configuration d'association de cartes réseau) ne peut pas être utilisée comme méthode pour augmenter ou doubler la bande passante réseau vers ou depuis l'instance à deux interfaces réseau.
- Si vous attachez plusieurs interfaces réseau du même sous-réseau à une instance, vous pouvez être confronté à des problèmes de mise en réseau comme le routage asymétrique. Si possible, utilisez plutôt une adresse IPv4 privée secondaire sur l'interface réseau principale.

Configuration de votre interface réseau à l'aide de ec2-net-utils

Les AMI Amazon Linux contiennent des scripts supplémentaires installés par AWS, appelés ec2-net-utils. Ces scripts automatisent le cas échéant la configuration de vos interfaces réseau. Ces scripts sont disponibles pour Amazon Linux uniquement.

Utilisez la commande suivante pour installer le package sur Amazon Linux s'il n'est pas déjà installé, ou mettez-le à jour s'il est installé et que des mises à jour supplémentaires sont disponibles :

```
$ yum install ec2-net-utils
```

Les composants suivants font partie de ec2-net-utils :

udev rules (`/etc/udev/rules.d`)

Identifie les interfaces réseau qui sont attachées, détachées ou rattachées à une instance en cours d'exécution, et s'assure que le script hotplug s'exécute (`53-ec2-network-interfaces.rules`). Mappe l'adresse MAC sur un nom de périphérique (`75-persistent-net-generator.rules`, qui génère `70-persistent-net.rules`).

script hotplug

Génère un fichier de configuration d'interface adapté à une utilisation avec DHCP (`/etc/sysconfig/network-scripts/ifcfg-eth`N). Génère également un fichier de configuration de route (`/etc/sysconfig/network-scripts/route-eth`N).

script DHCP

Chaque fois qu'une Network Interface reçoit un nouveau bail DHCP, ce script interroge les métadonnées d'instance pour des adresses IP Elastic. Pour chaque adresse IP Elastic, il ajoute une règle à la base de données de stratégies de routage pour s'assurer que le trafic sortant à partir de cette adresse utilise l'interface réseau correcte. Il ajoute également chaque adresse IP privée à l'interface réseau comme adresse secondaire.

ec2ifup ethN

Étend la fonctionnalité de la commande ifup standard. Une fois que ce script a réécrit les fichiers de configuration `ifcfg-eth`N et `route-eth`N, il exécute ifup.

ec2ifdown ethN

Étend la fonctionnalité de la commande ifdown standard. Une fois que le script a supprimé les règles pour l'interface réseau de la base de données des stratégies de routage, il exécute ifdown.

ec2ifscan

Recherche les interfaces réseau qui n'ont pas été configurées et les configure.

Ce script n'est pas disponible dans la version initiale de ec2-net-utils.

Pour répertorier les fichiers de configuration générés par ec2-net-utils, utilisez la commande suivante :

```
$ ls -l /etc/sysconfig/network-scripts/*-eth?
```

Pour désactiver l'automatisation instance par instance, vous pouvez ajouter `EC2SYNC=no` au fichier `ifcfg-eth`N correspondant. Par exemple, utilisez la commande suivante pour désactiver l'automatisation pour l'interface eth1 :

```
$ sed -i -e 's/^EC2SYNC=yes/EC2SYNC=no/' /etc/sysconfig/network-scripts/ifcfg-eth1
```

Pour désactiver complètement l'automatisation, vous pouvez supprimer le package à l'aide de la commande suivante :

```
$ yum remove ec2-net-utils
```

Interfaces réseau gérées par demandeur

Une interface réseau gérée par demandeur est une interface réseau qu'un service AWS crée dans votre VPC. Cette interface réseau peut représenter une instance pour un autre service, par exemple une instance Amazon RDS, ou elle peut vous permettre d'accéder à un autre service ou une autre ressource, comme un service AWS PrivateLink ou une tâche Amazon ECS.

Vous ne pouvez pas modifier ou détacher une interface réseau gérée par demandeur. Si vous supprimez la ressource que l'interface réseau représente, le service AWS détache et supprime l'interface réseau. Pour changer les groupes de sécurité d'une interface réseau gérée par demandeur, vous devrez peut-être utiliser la console ou les outils de ligne de commande correspondant à ce service. Pour plus d'informations, consultez la documentation propre au service concerné.

Vous pouvez baliser une interface réseau gérée par demandeur. Pour plus d'informations, consultez Ajout ou modification de balises (p. 858).

Vous pouvez afficher les interfaces réseau gérées par demandeur présentes dans votre compte.

Pour afficher les interfaces réseau gérées par demandeur à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Sélectionnez l'interface réseau et affichez les informations suivantes dans le volet des détails :
   - Propriétaire de l'attachement : si vous avez créé l'interface réseau, ce champ présente l'ID de votre compte AWS. Dans le cas contraire, il affiche un alias ou un ID du principal ou service qui a créé l'interface réseau.
   - Description : fournit des informations concernant l'objet de l'interface réseau ; par exemple, « Interface de point de terminaison d'un VPC ».

Pour afficher les interfaces réseau gérées par demandeur à l'aide de la ligne de commande

1. Utilisez la commande describe-network-interfaces de l'AWS CLI pour décrire les interfaces réseau dans votre compte.

   ```
   aws ec2 describe-network-interfaces
   ```

2. Dans la sortie, le champ `RequesterManaged` affiche `true` si l'interface réseau est gérée par un autre service AWS.

   ```
   {
   ```

```
    "Status": "in-use",
     ...
    "Description": "VPC Endpoint Interface vpce-089f2123488812123",
    "NetworkInterfaceId": "eni-c8fbc27e",
    "VpcId": "vpc-1a2b3c4d",
    "PrivateIpAddresses": [
        {
            "PrivateDnsName": "ip-10-0-2-227.ec2.internal",
            "Primary": true,
            "PrivateIpAddress": "10.0.2.227"
        }
    ],
    "RequesterManaged": true,
     ...
}
```

Vous pouvez également utiliser la commande Get-EC2NetworkInterface Outils pour Windows PowerShell.

Mise en réseau améliorée sur Linux

La mise en réseau améliorée utilise la virtualisation d'I/O d'une racine unique (SR-IOV) pour fournir des fonctionnalités de mise en réseau hautes performances sur les types d'instance pris en charge (p. 864). La méthode SR-IOV de virtualisation des appareils fournit de meilleures performances des E/S et une utilisation de la CPU réduite par rapport aux interfaces réseau virtualisées traditionnelles. La mise en réseau améliorée offre une bande passante supérieure, des performances de paquet par seconde (PPS) nettement plus élevées, ainsi que des latences réduites entre les instances. L'utilisation de la mise en réseau améliorée n'implique aucun coût supplémentaire.

Pour plus d'informations sur la vitesse réseau prise en charge pour chaque type d'instance, consultez Types d'instances Amazon EC2.

Sommaire

- Prise en charge de la mise en réseau améliorée (p. 864)
- Activation de la mise en réseau améliorée sur une instance (p. 865)
- Activation de la mise en réseau améliorée avec l'adaptateur Elastic Network Adapter (ENA) sur les instances Linux (p. 865)
- Activation de la mise en réseau améliorée avec l'interface Intel 82599 VF sur les instances Linux (p. 878)
- Dépannage d'Elastic Network Adapter (ENA) (p. 884)

Prise en charge de la mise en réseau améliorée

Tous les types d'instance de génération actuelle (p. 214) prennent en charge la mise en réseau améliorée, à l'exception des instances T2.

Vous pouvez activer la mise en réseau améliorée à l'aide de l'un des mécanismes suivants :

Elastic Network Adapter (ENA)

Elastic Network Adapter (ENA) prend en charge des vitesses réseau allant jusqu'à 100 Gbit/s pour les types d'instances pris en charge.

Les instances de la génération actuelle utilisent ENA pour la mise en réseau améliorée, à l'exception des instances C4, D2 et M4 plus petites que `m4.16xlarge`.

Interface Intel 82599 Virtual Function (VF)

L'interface Intel 82599 Virtual Function prend en charge les vitesses réseau allant jusqu'à 10 Gbit/s pour les types d'instance pris en charge.

Les types d'instance suivants utilisent l'interface Intel 82599 VF pour la mise en réseau améliorée : C3, C4, D2, I2, M4 (sauf m4.16xlarge) et R3.

Pour accéder à un résumé des mécanismes de mise en réseau améliorée par type d'instance, veuillez consulter Résumé des fonctions de réseautage et de stockage (p. 220).

Activation de la mise en réseau améliorée sur une instance

Si votre type d'instance prend en charge l'adaptateur Elastic Network Adapter pour la mise en réseau améliorée, suivez les procédures décrites dans la section Activation de la mise en réseau améliorée avec l'adaptateur Elastic Network Adapter (ENA) sur les instances Linux (p. 865).

Si votre type d'instance prend en charge l'interface Intel 82599 VF pour la mise en réseau améliorée, suivez les procédures décrites dans la section Activation de la mise en réseau améliorée avec l'interface Intel 82599 VF sur les instances Linux (p. 878).

Activation de la mise en réseau améliorée avec l'adaptateur Elastic Network Adapter (ENA) sur les instances Linux

Amazon EC2 offre des fonctionnalités de mise en réseau améliorée via l'adaptateur Elastic Network Adapter (ENA). Pour utiliser la mise en réseau améliorée, vous devez installer le module ENA requis et activer la prise en charge ENA.

Table des matières

- Prérequis (p. 865)
- Performances réseau améliorées (p. 866)
- Test de l'activation de la mise en réseau améliorée (p. 866)
- Activation de la mise en réseau améliorée sur Amazon Linux AMI (p. 868)
- Activation de la mise en réseau améliorée sur Ubuntu (p. 869)
- Activation de la mise en réseau améliorée sur Linux (p. 871)
- Activation de la mise en réseau améliorée sur Ubuntu via DKMS (p. 873)
- Dépannage (p. 875)
- Optimisations du système d'exploitation (p. 875)

Prérequis

Pour préparer la mise en réseau améliorée à l'aide de l'adaptateur réseau ENA, configurez votre instance comme suit :

- Lancez l'instance à l'aide d'un type d'instance de génération actuelle (p. 214), autre que des instances C4, D2 et M4 plus petites que `m4.16xlarge`, ou T2.

- Lancez l'instance à l'aide d'une version prise en charge du noyau Linux et d'une distribution prise en charge pour activer automatiquement la mise en réseau améliorée ENA pour votre instance. Pour de plus amples informations, veuillez consulter ENA Linux Kernel Driver Release Notes.
- Vérifiez que l'instance a une connectivité Internet.
- Installez et configurez l'interface de ligne de commande AWS CLI ou les Outils AWS pour Windows PowerShell sur l'ordinateur de votre choix, de préférence votre ordinateur de bureau ou ordinateur portable local. Pour plus d'informations, consultez Accès à Amazon EC2 (p. 3). La gestion de la mise en réseau améliorée n'est pas possible à partir de la console Amazon EC2.
- Si l'instance comporte des données importantes que vous souhaitez conserver, vous devez les sauvegarder maintenant en créant une AMI à partir de votre instance. La mise à jour des noyaux et des modules noyau, ainsi que l'activation de l'attribut `enaSupport`, peuvent rendre les instances incompatibles ou les systèmes d'exploitation inaccessibles. Si cela se produit et que vous disposez d'une sauvegarde récente, vos données continueront d'être conservées.

Performances réseau améliorées

La documentation suivante fournit un résumé des performances réseau pour les types d'instance qui prennent en charge la mise en réseau améliorée ENA :

- Performances réseau pour les instances de calcul accéléré (p. 301)
- Performances réseau pour les instances optimisées pour le calcul (p. 273)
- Performances réseau pour les instances à usage général (p. 227)
- Performances réseau pour les instances optimisées en mémoire (p. 283)
- Performances réseau pour les instances optimisées pour le stockage (p. 292)

Test de l'activation de la mise en réseau améliorée

Les AMI suivantes incluent le module ENA requis et la prise en charge ENA est activée :

- Amazon Linux 2
- Amazon Linux AMI 2018.03
- Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure
- Red Hat Enterprise Linux 7.4 ou version ultérieure
- SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
- CentOS 7.4.1708 ou version ultérieure
- FreeBSD 11.1 ou version ultérieure
- Debian GNU/Linux 9 ou version ultérieure

Pour tester si la mise en réseau améliorée est déjà activée, vérifiez que le `ena` module est installé sur votre instance et que l'attribut `enaSupport` est défini. Si votre instance remplit ces deux conditions, la commande ethtool -i eth*n* doit afficher que le module est en cours d'utilisation sur l'interface réseau.

Module noyau (ena)

Pour vérifier que le module `ena` est installé, utilisez la commande modinfo comme illustré dans l'exemple suivant :

```
[ec2-user ~]$ modinfo ena
filename:       /lib/modules/4.14.33-59.37.amzn2.x86_64/kernel/drivers/amazon/net/ena/
ena.ko
```

```
version:        1.5.0g
license:        GPL
description:    Elastic Network Adapter (ENA)
author:         Amazon.com, Inc. or its affiliates
srcversion:     692C7C68B8A9001CB3F31D0
alias:          pci:v00001D0Fd0000EC21sv*sd*bc*sc*i*
alias:          pci:v00001D0Fd0000EC20sv*sd*bc*sc*i*
alias:          pci:v00001D0Fd00001EC2sv*sd*bc*sc*i*
alias:          pci:v00001D0Fd00000EC2sv*sd*bc*sc*i*
depends:
retpoline:      Y
intree:         Y
name:           ena
...
```

Dans le cas d'Amazon Linux ci-dessus, le module `ena` est installé.

```
ubuntu:~$ modinfo ena
ERROR: modinfo: could not find module ena
```

Dans l'instance Ubuntu ci-dessus, le module n'est pas installé. Vous devez donc commencer par l'installer. Pour plus d'informations, consultez Activation de la mise en réseau améliorée sur Ubuntu (p. 869).

Attribut de l'instance (enaSupport)

Pour vérifier si l'attribut de mise en réseau améliorée `enaSupport` est défini sur une instance, utilisez l'une des commandes suivantes. Si l'attribut est défini, la réponse est true.

- describe-instances (AWS CLI)

```
aws ec2 describe-instances --instance-ids instance_id --query
 "Reservations[].Instances[].EnaSupport"
```

- Get-EC2InstanceOutils pour Windows PowerShell

```
(Get-EC2Instance -InstanceId instance-id).Instances.EnaSupport
```

Attribut image (enaSupport)

Pour vérifier si l'attribut de mise en réseau améliorée `enaSupport` est déjà défini sur une AMI, utilisez l'une des commandes suivantes. Si l'attribut est défini, la réponse est true.

- describe-images (AWS CLI)

```
aws ec2 describe-images --image-id ami_id --query "Images[].EnaSupport"
```

- Get-EC2Image (Outils pour Windows PowerShell)

```
(Get-EC2Image -ImageId ami_id).EnaSupport
```

Pilote d'interface réseau

Utilisez la commande suivante pour vérifier que le module `ena` est utilisé sur une interface particulière, en remplaçant le nom de l'interface par celui que vous voulez contrôler. Si vous utilisez une seule interface (par défaut), ce sera `eth0`. Si le système d'exploitation prend en charge les noms de réseau prévisibles (p. 871), il peut s'agir d'un nom tel que `ens5`.

Dans l'exemple suivant, le module `ena` n'est pas chargé, car le pilote affiché est `vif`.

```
[ec2-user ~]$ ethtool -i eth0
driver: vif
version:
firmware-version:
bus-info: vif-0
supports-statistics: yes
supports-test: no
supports-eeprom-access: no
supports-register-dump: no
supports-priv-flags: no
```

Dans cet exemple, le module `ena` est chargé avec la version minimale recommandée. La mise en réseau améliorée est correctement configurée pour cette instance.

```
[ec2-user ~]$ ethtool -i eth0
driver: ena
version: 1.5.0g
firmware-version:
expansion-rom-version:
bus-info: 0000:00:05.0
supports-statistics: yes
supports-test: no
supports-eeprom-access: no
supports-register-dump: no
supports-priv-flags: no
```

Activation de la mise en réseau améliorée sur Amazon Linux AMI

Amazon Linux 2 et les dernières versions de Amazon Linux AMI incluent le module requis pour la mise en réseau améliorée avec ENA installé et la prise en charge ENA activée. Par conséquent, si vous lancez une instance avec la dernière version HVM d'Amazon Linux sur un type d'instance pris en charge, la mise en réseau améliorée est déjà activée pour votre instance. Pour de plus amples informations, veuillez consulter Test de l'activation de la mise en réseau améliorée (p. 866).

Si vous avez lancé votre instance avec une version d'Amazon Linux AMI plus ancienne et que la mise en réseau améliorée n'est pas activée sur cette dernière, utilisez le procédure suivante pour l'activer.

Pour activer la mise en réseau améliorée sur Amazon Linux AMI

1. Connectez-vous à votre instance.
2. Depuis l'instance, exécutez la commande suivante pour mettre à jour votre instance avec le noyau et les modules noyau les plus récents, y compris `ena` :

   ```
   [ec2-user ~]$ sudo yum update
   ```

3. Depuis votre ordinateur local, redémarrez votre instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : reboot-instances (AWS CLI), ou Restart-EC2Instance (Outils AWS pour Windows PowerShell).
4. Connectez-vous à nouveau à votre instance et vérifiez que le module `ena` est installé et possède la version minimale recommandée à l'aide de la commande modinfo ena depuis Test de l'activation de la mise en réseau améliorée (p. 866).
5. [EBS-backed instance] À partir de votre ordinateur local, arrêtez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : stop-instances (AWS CLI), Stop-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous

devez arrêter l'instance de la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

[Instance basée sur le stockage d'instance] Vous ne pouvez pas arrêter l'instance pour modifier l'attribut. Vous devez utiliser cette procédure : Pour activer la mise en réseau améliorée sur Amazon Linux AMI (instances basées sur le stockage d'instance) (p. 869).

6. Depuis votre ordinateur local, activez l'attribut de mise en réseau améliorée à l'aide de l'une des commandes suivantes:

   - modify-instance-attribute (AWS CLI)

   ```
   aws ec2 modify-instance-attribute --instance-id instance_id --ena-support
   ```

   - Edit-EC2InstanceAttribute (Outils pour Windows PowerShell)

   ```
   Edit-EC2InstanceAttribute -InstanceId instance-id -EnaSupport $true
   ```

7. (Facultatif) Créez une AMI à partir de l'instance, comme décrit dans Création d'une AMI Linux basée sur Amazon EBS (p. 132). L'AMI hérite de l'attribut `enaSupport` de mise en réseau améliorée de l'instance. Par conséquent, vous pouvez utiliser cet AMI pour lancer une autre instance avec la mise en réseau améliorée activée par défaut.
8. Depuis votre ordinateur local, démarrez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : start-instances (AWS CLI) ou Start-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez démarrer l'instance dans la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.
9. Connectez-vous à votre instance et vérifiez que le module `ena` est installé et chargé sur votre interface réseau à l'aide de la commande ethtool -i eth*n* depuis Test de l'activation de la mise en réseau améliorée (p. 866).

   Si vous ne parvenez pas à vous connecter à votre instance après avoir activé la mise en réseau améliorée, consultez Dépannage d'Elastic Network Adapter (ENA) (p. 884).

Pour activer la mise en réseau améliorée sur Amazon Linux AMI (instances basées sur le stockage d'instance)

Suivez la procédure précédente jusqu'à l'étape à laquelle vous avez arrêté l'instance. Créez une nouvelle AMI comme décrit dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136), en veillant à activer l'attribut de mise en réseau améliorée lors de l'enregistrement de l'AMI.

- register-image (AWS CLI)

```
aws ec2 register-image --ena-support ...
```

- Register-EC2Image (Outils AWS pour Windows PowerShell)

```
Register-EC2Image -EnaSupport $true ...
```

Activation de la mise en réseau améliorée sur Ubuntu

Les dernières AMI HVM Ubuntu incluent le module requis pour la mise en réseau améliorée avec ENA installé et la prise en charge ENA activée. Par conséquent, si vous lancez une instance avec la dernière AMI HVM Ubuntu sur un type d'instance pris en charge, la mise en réseau améliorée est déjà activée pour votre instance. Pour plus d'informations, consultez Test de l'activation de la mise en réseau améliorée (p. 866).

Si vous avez lancé votre instance à l'aide d'une AMI plus ancienne et que la mise en réseau améliorée n'est pas déjà activée pour celle-ci, vous pouvez installer le package noyau `linux-aws` pour obtenir les pilotes de mise en réseau améliorée les plus récents et mettre à jour l'attribut requis.

Pour installer le package noyau linux-aws (Ubuntu 16.04 ou version ultérieure)

Ubuntu 16.04 et 18.04 sont fournis avec le noyau personnalisé Ubuntu (package noyau linux-aws). Pour utiliser un autre noyau, contactez AWS Support.

Pour installer le package noyau linux-aws (Ubuntu Trusty 14.04)

1. Connectez-vous à votre instance.
2. Mettez à jour le cache du package et les packages.

```
ubuntu:~$ sudo apt-get update && sudo apt-get upgrade -y linux-aws
```

   > **Important**
   >
   > Si, lors d'une mise à jour, vous êtes invité à installer `grub`, utilisez `/dev/xvda` pour y installer `grub`, puis choisissez de conserver la version courante de `/boot/grub/menu.lst`.

3. [EBS-backed instance] À partir de votre ordinateur local, arrêtez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : stop-instances (AWS CLI), Stop-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez arrêter l'instance de la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

   [Instance basée sur le stockage d'instance] Vous ne pouvez pas arrêter l'instance pour modifier l'attribut. Vous devez utiliser cette procédure : Pour activer la mise en réseau améliorée sur Ubuntu (instances basées sur le stockage d'instance) (p. 870).
4. Depuis votre ordinateur local, activez l'attribut de mise en réseau améliorée à l'aide de l'une des commandes suivantes:

   - modify-instance-attribute (AWS CLI)

   ```
   aws ec2 modify-instance-attribute --instance-id instance_id --ena-support
   ```

   - Edit-EC2InstanceAttribute (Outils pour Windows PowerShell)

   ```
   Edit-EC2InstanceAttribute -InstanceId instance-id -EnaSupport $true
   ```

5. (Facultatif) Créez une AMI à partir de l'instance, comme décrit dans Création d'une AMI Linux basée sur Amazon EBS (p. 132). L'AMI hérite de l'attribut `enaSupport` de mise en réseau améliorée de l'instance. Par conséquent, vous pouvez utiliser cet AMI pour lancer une autre instance avec la mise en réseau améliorée activée par défaut.
6. Depuis votre ordinateur local, démarrez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : start-instances (AWS CLI) ou Start-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez démarrer l'instance dans la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

Pour activer la mise en réseau améliorée sur Ubuntu (instances basées sur le stockage d'instance)

Suivez la procédure précédente jusqu'à l'étape à laquelle vous avez arrêté l'instance. Créez une nouvelle AMI comme décrit dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136), en veillant à activer l'attribut de mise en réseau améliorée lors de l'enregistrement de l'AMI.

- register-image (AWS CLI)

```
aws ec2 register-image --ena-support ...
```

- Register-EC2Image (Outils AWS pour Windows PowerShell)

```
Register-EC2Image -EnaSupport $true ...
```

Activation de la mise en réseau améliorée sur Linux

Les dernières AMI pour Red Hat Enterprise Linux, SUSE Linux Enterprise Server et CentOS incluent le module requis pour une mise en réseau améliorée avec ENA et ont la prise en charge ENA activée. Par conséquent, si vous lancez une instance avec la dernière AMI HVM Ubuntu sur un type d'instance pris en charge, la mise en réseau améliorée est déjà activée pour votre instance. Pour plus d'informations, consultez Test de l'activation de la mise en réseau améliorée (p. 866).

La procédure suivante fournit les étapes générales pour activer la mise en réseau améliorée via ENA sur une distribution Linux autre qu'Amazon Linux AMI ou Ubuntu. Pour de plus amples informations, telles que la syntaxe détaillée des commandes, les emplacements de fichier ou la prise en charge des packages et des outils, veuillez consulter la documentation spécifique de votre distribution Linux.

Pour activer la mise en réseau améliorée sur Linux

1. Connectez-vous à votre instance.
2. Clonez le code source du module `ena` sur votre instance depuis GitHub à l'adresse https://github.com/amzn/amzn-drivers. (SUSE Linux Enterprise Server 12 SP2 et versions ultérieures incluent ENA 2.02 par défaut, de sorte que vous n'êtes pas tenu de télécharger ni de compiler le pilote ENA. Pour SUSE Linux Enterprise Server 12 SP2 et versions ultérieures, vous devez déposer une demande d'ajout de la version du pilote que vous souhaitez dans le noyau de base).

   ```
   git clone https://github.com/amzn/amzn-drivers
   ```

3. Compilez et installez le module `ena` sur votre instance. Ces étapes dépendent de la distribution Linux. Pour plus d'informations sur la compilation du module sur Red Hat Enterprise Linux, consultez l'article du Centre de connaissances AWS.
4. Exécutez la commande sudo depmod pour mettre à jour les dépendances du module.
5. Mettez à jour `initramfs` sur votre instance pour garantir que le nouveau module se charge au démarrage. Par exemple, si votre distribution prend en charge dracut, vous pouvez utiliser la commande suivante :

   ```
   dracut -f -v
   ```

6. Déterminez si par défaut votre système utilise des noms d'interface réseau prévisibles. Les systèmes qui utilisent systemd ou udev version 197 ou supérieure peuvent renommer les périphériques Ethernet et ne garantissent pas qu'une seule interface réseau sera nommée `eth0`. Ce comportement peut entraîner des problèmes de connexion à votre instance. Pour plus d'informations et pour voir les autres options de configuration, consultez la section sur les noms d'interface réseau prévisibles sur le site web de freedesktop.org.

   a. Vous pouvez vérifier les versions systemd ou udev sur les systèmes RPM en utilisant la commande suivante :

   ```
   rpm -qa | grep -e '^systemd-[0-9]\+\|^udev-[0-9]\+'
   systemd-208-11.el7_0.2.x86_64
   ```

Dans l'exemple Red Hat Enterprise Linux 7 ci-dessus, la version systemd est 208, de sorte que les noms d'interface réseau prévisibles doivent être désactivés.

b. Désactivez les noms d'interface réseau prévisibles en ajoutant l'option `net.ifnames=0` à la ligne `GRUB_CMDLINE_LINUX` dans `/etc/default/grub`.

```
sudo sed -i '/^GRUB\_CMDLINE\_LINUX/s/\"$/\ net\.ifnames\=0\"/' /etc/default/grub
```

c. Générez à nouveau le fichier de configuration grub.

```
sudo grub2-mkconfig -o /boot/grub2/grub.cfg
```

7. [EBS-backed instance] À partir de votre ordinateur local, arrêtez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : stop-instances (AWS CLI), Stop-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez arrêter l'instance de la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

   [Instance basée sur le stockage d'instance] Vous ne pouvez pas arrêter l'instance pour modifier l'attribut. Vous devez utiliser cette procédure : Pour activer la mise en réseau améliorée sur Linux (instances basées sur– le stockage d'instance) (p. 873).

8. Depuis votre ordinateur local, activez l'attribut de mise en réseau améliorée `enaSupport` à l'aide de l'une des commandes suivantes:

   - modify-instance-attribute (AWS CLI)

```
aws ec2 modify-instance-attribute --instance-id instance_id --ena-support
```

   - Edit-EC2InstanceAttribute (Outils pour Windows PowerShell)

```
Edit-EC2InstanceAttribute -InstanceId instance-id -EnaSupport $true
```

9. (Facultatif) Créez une AMI à partir de l'instance, comme décrit dans Création d'une AMI Linux basée sur Amazon EBS (p. 132) . L'AMI hérite de l'attribut `enaSupport` de mise en réseau améliorée de l'instance. Par conséquent, vous pouvez utiliser cet AMI pour lancer une autre instance avec la mise en réseau améliorée activée par défaut.

   Important

   Si le système d'exploitation de votre instance contient un fichier `/etc/udev/rules.d/70-persistent-net.rules`, vous devez le supprimer avant de créer l'AMI. Ce fichier contient l'adresse MAC de la carte Ethernet de l'instance d'origine. Si une autre instance démarre avec ce fichier, le système d'exploitation ne pourra pas trouver le périphérique et il se peut qu'`eth0` échoue, entraînant des problèmes de démarrage. Le fichier est à nouveau généré au cycle de démarrage suivant et les instances lancées depuis l'AMI créent leur propre version du fichier.

10. Depuis votre ordinateur local, démarrez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : start-instances (AWS CLI) ou Start-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez démarrer l'instance dans la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

11. (Facultatif) Connectez-vous à votre instance et vérifiez que le module est installé.

    Si vous ne parvenez pas à vous connecter à votre instance après avoir activé la mise en réseau améliorée, consultez Dépannage d'Elastic Network Adapter (ENA) (p. 884).

Pour activer la mise en réseau améliorée sur Linux (instances basées sur– le stockage d'instance)

Suivez la procédure précédente jusqu'à l'étape à laquelle vous avez arrêté l'instance. Créez une nouvelle AMI comme décrit dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136), en veillant à activer l'attribut de mise en réseau améliorée lors de l'enregistrement de l'AMI.

- register-image (AWS CLI)

```
aws ec2 register-image --ena-support ...
```

- Register-EC2Image (Outils AWS pour Windows PowerShell)

```
Register-EC2Image -EnaSupport ...
```

Activation de la mise en réseau améliorée sur Ubuntu via DKMS

Cette méthode est fournie à des fins de test et de rétroaction uniquement. Elle n'est pas destinée à être utilisée avec des déploiements en production. Pour de plus amples informations sur les déploiements en production, veuillez consulter Activation de la mise en réseau améliorée sur Ubuntu (p. 869).

> **Important**
>
> L'utilisation de DKMS annule le contrat de support pour votre abonnement. Il ne doit pas être utilisé pour les déploiements de production.

Pour activer la mise en réseau améliorée via ENA sur Ubuntu (instances basées sur EBS)

1. Suivez les étapes 1 et 2 dans Activation de la mise en réseau améliorée sur Ubuntu (p. 869).
2. Installez les packages `build-essential` pour compiler le module noyau et les packages `dkms` pour que le module `ena` soit recréé chaque fois que votre noyau est mis à jour.

   ```
   ubuntu:~$ sudo apt-get install -y build-essential dkms
   ```

3. Clonez la source pour le module `ena` sur votre instance depuis GitHub, à l'adresse https://github.com/amzn/amzn-drivers.

   ```
   ubuntu:~$ git clone https://github.com/amzn/amzn-drivers
   ```

4. Déplacez le package `amzn-drivers` vers le répertoire `/usr/src/` afin que DKMS puisse le trouver et le générer à chaque mise à jour du noyau. Ajoutez le numéro de version (que vous trouverez dans les notes de version) du code source au nom du répertoire. Par exemple, la version `1.0.0` apparaît dans l'exemple suivant.

   ```
   ubuntu:~$ sudo mv amzn-drivers /usr/src/amzn-drivers-1.0.0
   ```

5. Créez le fichier de configuration DKMS avec les valeurs suivantes, en remplaçant votre version d'`ena`.

   Créez le fichier.

   ```
   ubuntu:~$ sudo touch /usr/src/amzn-drivers-1.0.0/dkms.conf
   ```

   Modifiez le fichier et ajoutez les valeurs suivantes.

   ```
   ubuntu:~$ sudo vim /usr/src/amzn-drivers-1.0.0/dkms.conf
   ```

```
PACKAGE_NAME="ena"
PACKAGE_VERSION="1.0.0"
CLEAN="make -C kernel/linux/ena clean"
MAKE="make -C kernel/linux/ena/ BUILD_KERNEL=${kernelver}"
BUILT_MODULE_NAME[0]="ena"
BUILT_MODULE_LOCATION="kernel/linux/ena"
DEST_MODULE_LOCATION[0]="/updates"
DEST_MODULE_NAME[0]="ena"
AUTOINSTALL="yes"
```

6. Ajoutez, générez et installez le module `ena` sur votre instance à l'aide de DKMS.

   Ajoutez le module à DKMS.

```
ubuntu:~$ sudo dkms add -m amzn-drivers -v 1.0.0
```

   Générez le module avec la commande dkms.

```
ubuntu:~$ sudo dkms build -m amzn-drivers -v 1.0.0
```

   Installez le module avec dkms.

```
ubuntu:~$ sudo dkms install -m amzn-drivers -v 1.0.0
```

7. Générez à nouveau `initramfs` afin que le module approprié soit chargé au démarrage.

```
ubuntu:~$ sudo update-initramfs -u -k all
```

8. Vérifiez que le module `ena` est installé à l'aide de la commande modinfo ena depuis Test de l'activation de la mise en réseau améliorée (p. 866).

```
ubuntu:~$ modinfo ena
filename:       /lib/modules/3.13.0-74-generic/updates/dkms/ena.ko
version:        1.0.0
license:        GPL
description:    Elastic Network Adapter (ENA)
author:         Amazon.com, Inc. or its affiliates
srcversion:     9693C876C54CA64AE48F0CA
alias:          pci:v00001D0Fd0000EC21sv*sd*bc*sc*i*
alias:          pci:v00001D0Fd0000EC20sv*sd*bc*sc*i*
alias:          pci:v00001D0Fd00001EC2sv*sd*bc*sc*i*
alias:          pci:v00001D0Fd00000EC2sv*sd*bc*sc*i*
depends:
vermagic:       3.13.0-74-generic SMP mod_unload modversions
parm:           debug:Debug level (0=none,...,16=all) (int)
parm:           push_mode:Descriptor / header push mode
 (0=automatic,1=disable,3=enable)
     0 - Automatically choose according to device capability (default)
     1 - Don't push anything to device memory
     3 - Push descriptors and header buffer to device memory (int)
parm:           enable_wd:Enable keepalive watchdog (0=disable,1=enable,default=1)
 (int)
parm:           enable_missing_tx_detection:Enable missing Tx completions. (default=1)
 (int)
parm:           numa_node_override_array:Numa node override map
 (array of int)
parm:           numa_node_override:Enable/Disable numa node override (0=disable)
 (int)
```

9. Passez à l'étape 3 dans Activation de la mise en réseau améliorée sur Ubuntu (p. 869).

Dépannage

Pour plus d'informations sur le dépannage de votre adaptateur ENA, consultez la section Dépannage d'Elastic Network Adapter (ENA) (p. 884).

Optimisations du système d'exploitation

Afin d'obtenir des performances réseau optimales sur les instances dont la mise en réseau est améliorée, il se peut que vous ayez besoin de modifier la configuration par défaut du système d'exploitation. Nous vous recommandons d'apporter les modifications suivantes à la configuration des applications qui nécessitent des performances réseau élevées.

Outre ces optimisations de système d'exploitation, vous devez également tenir compte de l'unité de transmission maximale (MTU) de votre trafic réseau et l'ajuster en fonction de votre charge de travail et de l'architecture réseau. Pour plus d'informations, consultez Unité de transmission maximale (MTU) du réseau pour votre instance EC2 (p. 937).

AWS mesure régulièrement les temps de latence aller-retour moyens entre les instances lancées dans un groupe de placement de cluster de 50 u et les temps de latence de queue de 200 u au percentile 99,9. Si vos applications nécessitent des temps de latence constamment faibles, nous vous recommandons d'utiliser la dernière version des pilotes ENA sur des instances Nitro à performances fixes.

Ces procédures ont été écrites pour Amazon Linux 2 et Amazon Linux AMI. Toutefois, elles peuvent également fonctionner pour d'autres distributions Linux avec une version 3.9 ou plus récente du noyau. Pour plus d'informations, consultez la documentation propre à votre système.

Pour optimiser votre instance Amazon Linux pour la mise en réseau améliorée

1. Vérifiez la source d'horloge de votre instance :

```
cat /sys/devices/system/clocksource/clocksource0/current_clocksource
```

2. Si la source d'horloge est `xen`, réalisez les sous-étapes suivantes. Sinon, passez à l'étape Step 3 (p. 875).

   a. Modifiez la configuration GRUB et ajoutez `xen_nopvspin=1` et `clocksource=tsc` aux options de démarrage du noyau.

      - Pour Amazon Linux 2, modifiez le fichier `/etc/default/grub` et ajoutez ces options à la ligne `GRUB_CMDLINE_LINUX_DEFAULT`, comme illustré ci-dessous :

```
GRUB_CMDLINE_LINUX_DEFAULT="console=tty0 console=ttyS0,115200n8 net.ifnames=0
 biosdevname=0 nvme_core.io_timeout=4294967295 xen_nopvspin=1 clocksource=tsc"
GRUB_TIMEOUT=0
```

      - Pour Amazon Linux AMI, modifiez le fichier `/boot/grub/grub.conf` et ajoutez ces options à la ligne `kernel`, comme illustré ci-dessous :

```
kernel /boot/vmlinuz-4.14.62-65.117.amzn1.x86_64 root=LABEL=/ console=tty1
 console=ttyS0 selinux=0 nvme_core.io_timeout=4294967295 xen_nopvspin=1
 clocksource=tsc
```

   b. (Amazon Linux 2 uniquement) Générez à nouveau le fichier de configuration GRUB afin de récupérer ces modifications :

```
sudo grub2-mkconfig -o /boot/grub2/grub.cfg
```

3. Si votre type d'instance est répertorié comme étant pris en charge sur Contrôle des états du processeur pour votre instance EC2 (p. 662), empêchez le système d'utiliser des états « C-states »

plus profonds afin de garantir des performances homogènes. Pour plus d'informations, consultez Haute performance et faible latence en limitant les états « C-states » plus profonds (p. 665).

a. Modifiez la configuration GRUB et ajoutez `intel_idle.max_cstate=1` aux options de démarrage du noyau.

   - Pour Amazon Linux 2, modifiez le fichier `/etc/default/grub` et ajoutez cette option à la ligne `GRUB_CMDLINE_LINUX_DEFAULT`, comme illustré ci-dessous :

```
GRUB_CMDLINE_LINUX_DEFAULT="console=tty0 console=ttyS0,115200n8
 net.ifnames=0 biosdevname=0 nvme_core.io_timeout=4294967295 xen_nopvspin=1
 clocksource=tsc intel_idle.max_cstate=1"
GRUB_TIMEOUT=0
```

   - Pour Amazon Linux AMI, modifiez le fichier `/boot/grub/grub.conf` et ajoutez cette option à la ligne `kernel`, comme illustré ci-dessous :

```
kernel /boot/vmlinuz-4.14.62-65.117.amzn1.x86_64 root=LABEL=/ console=tty1
 console=ttyS0 selinux=0 nvme_core.io_timeout=4294967295 xen_nopvspin=1
 clocksource=tsc intel_idle.max_cstate=1
```

b. (Amazon Linux 2 uniquement) Générez à nouveau le fichier de configuration GRUB afin de récupérer ces modifications :

```
sudo grub2-mkconfig -o /boot/grub2/grub.cfg
```

4. Assurez-vous que la mémoire du noyau réservée est suffisante pour supporter un taux élevé d'allocations de tampon de paquets (la valeur par défaut peut être insuffisante).

   a. Ouvrez (comme `root` ou avec sudo) le fichier `/etc/sysctl.conf` avec l'éditeur de votre choix.

   b. Ajoutez la ligne `vm.min_free_kbytes` au fichier avec la valeur de mémoire du noyau réservée (en kilo-octets) pour votre type d'instance. En règle générale, cette valeur doit être définie entre 1 et 3 % de mémoire système disponible et ajustée à la hausse ou à la baisse en fonction des besoins de votre application.

```
vm.min_free_kbytes = 1048576
```

   c. Appliquez cette configuration avec la commande suivante :

```
sudo sysctl -p
```

   d. Vérifiez que le paramètre a été appliqué à l'aide de la commande suivante :

```
sudo sysctl -a 2>&1 | grep min_free_kbytes
```

5. Redémarrez l'instance pour charger la nouvelle configuration :

```
sudo reboot
```

6. (Facultatif) Distribuez manuellement les interruptions de réception de paquets afin qu'elles soient associées à différents processeurs qui appartiennent tous au même nœud NUMA. Soyez toutefois prudent, car irqbalancer est désactivé globalement.

   Note

   La modification apportée à la configuration à cette étape ne persiste pas après un redémarrage.

a. Créez un fichier appelé `smp_affinity.sh` à l'intérieur duquel vous collerez le bloc de code suivant :

```
#/bin/sh
service irqbalance stop
affinity_values=(00000001 00000002 00000004 00000008 00000010 00000020 00000040
 00000080)
irqs=($(grep eth /proc/interrupts|awk '{print $1}'|cut -d : -f 1))
irqLen=${#irqs[@]}
for (( i=0; i<${irqLen}; i++ ));
do
  echo $(printf "0000,00000000,00000000,00000000,${affinity_values[$i]}") > /proc/
irq/${irqs[$i]}/smp_affinity;
  echo "IRQ ${irqs[$i]} =" $(cat /proc/irq/${irqs[$i]}/smp_affinity);
done
```

b. Exécutez le script avec la commande suivante :

```
sudo bash ./smp_affinity.sh
```

7. (Facultatif) Si les vCPU qui gèrent les IRQ de réception sont surchargés ou si le traitement de votre réseau d'application consomme énormément d'UC, vous pouvez transférer une partie de la charge du traitement du réseau à d'autres cœurs avec la technologie RPS (receive packet steering). Assurez-vous que les cœurs utilisés pour RPS appartiennent au même nœud NUMA afin d'éviter les verrouillages entre nœuds NUMA. Par exemple, pour utiliser les cœurs 8 à 15 pour le traitement de paquets, utilisez la commande suivante.

   Note

   La modification apportée à la configuration à cette étape ne persiste pas après un redémarrage.

```
for i in `seq 0 7`; do echo $(printf "0000,00000000,00000000,00000000,0000ff00") | sudo
 tee /sys/class/net/eth0/queues/rx-$i/rps_cpus; done
```

8. (Facultatif) Si possible, conservez l'ensemble du traitement sur le même nœud NUMA.

   a. Installez numactl :

```
sudo yum install -y numactl
```

   b. Lorsque vous exécutez votre programme de traitement de réseau, liez-le à un seul nœud NUMA. Par exemple, la commande suivante lie le script shell, `run.sh`, au nœud NUMA 0 :

```
numactl --cpunodebind=0 --membind=0 run.sh
```

   c. Si l'hyper-threading est activé, vous pouvez configurer votre application afin qu'elle utilise un seul thread matériel par cœur d'UC.

      - Vous pouvez vérifier quels cœurs d'UC sont mappés à un nœud NUMA à l'aide de la commande lscpu :

```
lscpu | grep NUMA
```

File d'attente:

```
NUMA node(s):      2
NUMA node0 CPU(s): 0-15,32-47
```

```
NUMA node1 CPU(s): 16-31,48-63
```

- Vous pouvez vérifier quels threads matériels appartiennent à une UC physique à l'aide de la commande suivante :

```
cat /sys/devices/system/cpu/cpu0/topology/thread_siblings_list
```

File d'attente:

```
0,32
```

Dans cet exemple, les threads 0 et 32 sont mappés à l'UC 0.

- Afin d'éviter l'exécution sur les threads 32 à 47 (qui sont des threads matériels de la même UC que les threads 0 à 15), utilisez la commande suivante :

```
numactl --physcpubind=+0-15 --membind=0 ./run.sh
```

9. Utilisez plusieurs interfaces réseau Elastic pour différentes classes de trafic. Par exemple, si vous exécutez un serveur Web qui utilise une base de données backend, utilisez une interface réseau Elastic pour le serveur Web frontal et une autre pour la connexion de base de données.

Activation de la mise en réseau améliorée avec l'interface Intel 82599 VF sur les instances Linux

Amazon EC2 fournit des fonctionnalités de mise en réseau améliorée via l'interface Intel 82599 VF, qui utilise le pilote Intel `ixgbevf`.

Sommaire

- Prérequis (p. 878)
- Test de l'activation de la mise en réseau améliorée (p. 879)
- Activation de la mise en réseau améliorée sur Amazon Linux (p. 880)
- Activation de la mise en réseau améliorée sur Ubuntu (p. 881)
- Activation de la mise en réseau améliorée sur les autres distributions Linux (p. 882)
- Dépannage des problèmes de connexion (p. 884)

Prérequis

Pour vous préparer à la mise en réseau améliorée à l'aide de l'interface Intel 82599 VF, configurez l'instance comme suit :

- Effectuez votre sélection parmi les types d'instances pris en charge suivants : C3, C4, D2, I2, M4 (à l'exception de `m4.16xlarge`) et R3.
- Lancez l'instance depuis une AMI HVM avec la version du noyau Linux 2.6.32 (ou version ultérieure). Les AMI HVM Amazon Linux les plus récentes disposent des attributs et des modules requis pour la mise en réseau améliorée. Par conséquent, si vous lancez une instance avec prise en charge de la mise en réseau améliorée et basée sur Amazon EBS–à l'aide d'une AMI HVM –active, la mise en réseau améliorée est déjà Amazon Linux activée pour votre instance.

  Warning

  La mise en réseau améliorée n'est prise en charge que pour les instances HVM. L'activation de la mise en réseau améliorée avec une instance de paravirtualisation peut la rendre inaccessible.

La définition de cet attribut sans le module ou la version de module approprié peut rendre votre instance inaccessible.

- Vérifiez que l'instance a une connectivité Internet.
- Installez et configurez l'interface de ligne de commande AWS CLI ou les Outils AWS pour Windows PowerShell sur l'ordinateur de votre choix, de préférence votre ordinateur de bureau ou ordinateur portable local. Pour plus d'informations, consultez Accès à Amazon EC2 (p. 3). La gestion de la mise en réseau améliorée n'est pas possible à partir de la console Amazon EC2.
- Si l'instance comporte des données importantes que vous souhaitez conserver, vous devez les sauvegarder maintenant en créant une AMI à partir de votre instance. La mise à jour des noyaux et des modules noyau, ainsi que l'activation de l'attribut `sriovNetSupport`, peuvent rendre les instances incompatibles ou les systèmes d'exploitation inaccessibles. Si cela se produit et que vous disposez d'une sauvegarde récente, vos données continueront d'être conservées.

Test de l'activation de la mise en réseau améliorée

La mise en réseau améliorée avec l'interface Intel 82599 VF est activée si le `ixgbevf` module est installé sur votre instance et que l'attribut `sriovNetSupport` est défini.

Attribut d'instance (sriovNetSupport)

Pour vérifier si l'attribut de mise en réseau améliorée `sriovNetSupport` est défini sur une instance, utilisez l'une des commandes suivantes :

- describe-instance-attribute (AWS CLI)

```
aws ec2 describe-instance-attribute --instance-id instance_id --attribute sriovNetSupport
```

- Get-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

```
Get-EC2InstanceAttribute -InstanceId instance-id -Attribute sriovNetSupport
```

Si l'attribut n'est pas défini, `SriovNetSupport` est vide. Si l'attribut est défini, la valeur est simple, comme indiqué dans l'exemple de sortie suivant.

```
"SriovNetSupport": {
    "Value": "simple"
},
```

Attribut d'image (sriovNetSupport)

Pour vérifier si l'attribut de mise en réseau améliorée `sriovNetSupport` est déjà défini sur une AMI, utilisez l'une des commandes suivantes :

- describe-images (AWS CLI)

```
aws ec2 describe-images --image-id ami_id --query "Images[].SriovNetSupport"
```

- Get-EC2Image (Outils AWS pour Windows PowerShell)

```
(Get-EC2Image -ImageId ami-id).SriovNetSupport
```

Si l'attribut n'est pas défini, `SriovNetSupport` est vide. Si l'attribut est défini, la valeur est simple.

Pilote d'interface réseau

Utilisez la commande suivante pour vérifier que le module est utilisé sur une interface particulière, en remplaçant le nom de l'interface par celui que vous voulez contrôler. Si vous utilisez une seule interface (par défaut), ce sera `eth0`. Si le système d'exploitation prend en charge les noms de réseau prévisibles (p. 882), il peut s'agir d'un nom tel que `ens5`.

Dans l'exemple suivant, le module `ixgbevf` n'est pas chargé, car le pilote affiché est `vif`.

```
[ec2-user ~]$ ethtool -i eth0
driver: vif
version:
firmware-version:
bus-info: vif-0
supports-statistics: yes
supports-test: no
supports-eeprom-access: no
supports-register-dump: no
supports-priv-flags: no
```

Dans cet exemple, le module `ixgbevf` est chargé. La mise en réseau améliorée est correctement configurée pour cette instance.

```
[ec2-user ~]$ ethtool -i eth0
driver: ixgbevf
version: 4.0.3
firmware-version: N/A
bus-info: 0000:00:03.0
supports-statistics: yes
supports-test: yes
supports-eeprom-access: no
supports-register-dump: yes
supports-priv-flags: no
```

Activation de la mise en réseau améliorée sur Amazon Linux

Les AMI HVM Amazon Linux les plus récentes disposent du module `ixgbevf` et de l'attribut `sriovNetSupport` requis pour la mise en réseau améliorée. Par conséquent, si vous lancez un type d'instance à l'aide d'une AMI HVM Amazon Linux actuelle, la mise en réseau améliorée est déjà activée pour votre instance. Pour plus d'informations, consultez Test de l'activation de la mise en réseau améliorée (p. 879).

Si vous avez lancé votre instance avec une AMI Amazon Linux plus ancienne où la mise en réseau améliorée n'est pas activée, utilisez le procédure suivante pour l'activer.

> **Warning**
>
> Il n'existe aucun moyen de désactiver l'attribut de mise en réseau améliorée une fois que vous l'avez activé.

Pour activer la mise en réseau améliorée

1. Connectez-vous à votre instance.
2. Depuis l'instance, exécutez la commande suivante pour mettre à jour votre instance avec le noyau et les modules noyau les plus récents, y compris `ixgbevf` :

   ```
   [ec2-user ~]$ sudo yum update
   ```

3. Depuis votre ordinateur local, redémarrez votre instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : reboot-instances (AWS CLI), ou Restart-EC2Instance (Outils AWS pour Windows PowerShell).

4. Connectez-vous à nouveau à votre instance et vérifiez que le module `ixgbevf` est installé et possède la version minimale recommandée à l'aide de la commande modinfo ixgbevf depuis Test de l'activation de la mise en réseau améliorée (p. 879).
5. [EBS-backed instance] À partir de votre ordinateur local, arrêtez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : stop-instances (AWS CLI), Stop-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez arrêter l'instance de la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

   [Instance basée sur le stockage d'instance] Vous ne pouvez pas arrêter l'instance pour modifier l'attribut. Vous devez utiliser cette procédure : Pour activer la mise en réseau améliorée (instances basées sur le stockage d'instance) (p. 881).
6. Depuis votre ordinateur local, activez l'attribut de mise en réseau améliorée à l'aide de l'une des commandes suivantes:

   - modify-instance-attribute (AWS CLI)

   ```
   aws ec2 modify-instance-attribute --instance-id instance_id --sriov-net-support
    simple
   ```

   - Edit-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

   ```
   Edit-EC2InstanceAttribute -InstanceId instance_id -SriovNetSupport "simple"
   ```
7. (Facultatif) Créez une AMI à partir de l'instance, comme décrit dans Création d'une AMI Linux basée sur Amazon EBS (p. 132) . L'AMI hérite l'attribut de mise en réseau améliorée de l'instance. Par conséquent, vous pouvez utiliser cet AMI pour lancer une autre instance avec la mise en réseau améliorée activée par défaut.
8. Depuis votre ordinateur local, démarrez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : start-instances (AWS CLI) ou Start-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez démarrer l'instance dans la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.
9. Connectez-vous à votre instance et vérifiez que le module `ixgbevf` est installé et chargé sur votre interface réseau à l'aide de la commande ethtool -i eth*n* depuis Test de l'activation de la mise en réseau améliorée (p. 879).

Pour activer la mise en réseau améliorée (instances basées sur le stockage d'instance)

Suivez la procédure précédente jusqu'à l'étape à laquelle vous avez arrêté l'instance. Créez une nouvelle AMI comme décrit dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136), en veillant à activer l'attribut de mise en réseau améliorée lors de l'enregistrement de l'AMI.

- register-image (AWS CLI)

```
aws ec2 register-image --sriov-net-support simple ...
```

- Register-EC2Image (Outils AWS pour Windows PowerShell)

```
Register-EC2Image -SriovNetSupport "simple" ...
```

Activation de la mise en réseau améliorée sur Ubuntu

Avant de commencer, vérifiez si la mise en réseau améliorée est déjà activée (p. 879) sur votre instance.

Les AMI HVM Ubuntu Quick Start comprennent les pilotes nécessaires pour la mise en réseau améliorée. Si vous disposez d'une version du fichier `ixgbevf` antérieure à 2.16.4, vous pouvez installer le package noyau `linux-aws` pour obtenir les pilotes de mise en réseau améliorée les plus récents.

La procédure suivante fournit les étapes générales pour la compilation du module `ixgbevf` sur une instance Ubuntu.

Pour installer le package noyau linux-aws

1. Connectez-vous à votre instance.
2. Mettez à jour le cache du package et les packages.

```
ubuntu:~$ sudo apt-get update && sudo apt-get upgrade -y linux-aws
```

> **Important**
>
> Si, lors d'une mise à jour, vous êtes invité à installer `grub`, utilisez `/dev/xvda` pour installer `grub`, puis choisissez de conserver la version actuelle de `/boot/grub/menu.lst`.

Activation de la mise en réseau améliorée sur les autres distributions Linux

Avant de commencer, vérifiez si la mise en réseau améliorée est déjà activée (p. 879) sur votre instance. Les dernières AMI HVM Quick Start comprennent les pilotes nécessaires pour la mise en réseau améliorée. Vous n'avez donc pas besoin d'effectuer des étapes supplémentaires.

La procédure suivante fournit les étapes générales pour si vous devez activer la mise en réseau améliorée avec l'interface Intel 82599 VF sur une distribution Linux autre qu'Amazon Linux ou Ubuntu. Pour plus d'informations, telles que la syntaxe détaillée des commandes, les emplacements de fichier ou la prise en charge des packages et des outils, consultez la documentation spécifique de votre distribution Linux.

Pour activer la mise en réseau améliorée sur Linux

1. Connectez-vous à votre instance.
2. Téléchargez la source pour le module `ixgbevf` sur votre instance depuis Sourceforge, à l'adresse https://sourceforge.net/projects/e1000/files/ixgbevf%20stable/.

   Les versions d'`ixgbevf` antérieures à 2.16.4, notamment la 2.14.2, ne sont pas générées correctement sur certaines distributions Linux, y compris certaines versions d'Ubuntu.
3. Compilez et installez le module `ixgbevf` sur votre instance.

   > **Warning**
   >
   > Si vous compilez le module `ixgbevf` pour votre noyau actuel, puis mettez à niveau le noyau sans générer à nouveau le pilote du nouveau noyau, il se peut que votre système retourne au module `ixgbevf` spécifique à la distribution lors du prochain redémarrage. Cela peut rendre votre système inaccessible si la version propre à la distribution n'est pas compatible avec la mise en réseau améliorée.
4. Exécutez la commande sudo depmod pour mettre à jour les dépendances du module.
5. Mettez à jour `initramfs` sur votre instance pour garantir que le nouveau module se charge au démarrage.
6. Déterminez si par défaut votre système utilise des noms d'interface réseau prévisibles. Les systèmes qui utilisent systemd ou udev version 197 ou supérieure peuvent renommer les périphériques Ethernet et ne garantissent pas qu'une seule interface réseau sera nommée `eth0`. Ce comportement peut entraîner des problèmes de connexion à votre instance. Pour plus d'informations et pour voir les autres

options de configuration, consultez la section sur les noms d'interface réseau prévisibles sur le site web de freedesktop.org.

a. Vous pouvez vérifier les versions systemd ou udev sur les systèmes RPM en utilisant la commande suivante :

```
[ec2-user ~]$ rpm -qa | grep -e '^systemd-[0-9]\+\|^udev-[0-9]\+'
systemd-208-11.el7_0.2.x86_64
```

Dans l'exemple Red Hat Enterprise Linux 7 ci-dessus, la version systemd est 208, de sorte que les noms d'interface réseau prévisibles doivent être désactivés.

b. Désactivez les noms d'interface réseau prévisibles en ajoutant l'option `net.ifnames=0` à la ligne `GRUB_CMDLINE_LINUX` dans `/etc/default/grub`.

```
[ec2-user ~]$ sudo sed -i '/^GRUB\_CMDLINE\_LINUX/s/\"$/\ net\.ifnames\=0\"/' /etc/
default/grub
```

c. Générez à nouveau le fichier de configuration grub.

```
[ec2-user ~]$ sudo grub2-mkconfig -o /boot/grub2/grub.cfg
```

7. [EBS-backed instance] À partir de votre ordinateur local, arrêtez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : stop-instances (AWS CLI), Stop-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez arrêter l'instance de la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

   [Instance basée sur le stockage d'instance] Vous ne pouvez pas arrêter l'instance pour modifier l'attribut. Vous devez utiliser cette procédure : Pour activer la mise en réseau améliorée (instances basées –sur le stockage d'instance) (p. 884).

8. Depuis votre ordinateur local, activez l'attribut de mise en réseau améliorée à l'aide de l'une des commandes suivantes:

   - modify-instance-attribute (AWS CLI)

```
aws ec2 modify-instance-attribute --instance-id instance_id --sriov-net-support
 simple
```

   - Edit-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

```
Edit-EC2InstanceAttribute -InstanceId instance_id -SriovNetSupport "simple"
```

9. (Facultatif) Créez une AMI à partir de l'instance, comme décrit dans Création d'une AMI Linux basée sur Amazon EBS (p. 132) . L'AMI hérite l'attribut de mise en réseau améliorée de l'instance. Par conséquent, vous pouvez utiliser cet AMI pour lancer une autre instance avec la mise en réseau améliorée activée par défaut.

   Important

   Si le système d'exploitation de votre instance contient un fichier `/etc/udev/rules.d/70-persistent-net.rules`, vous devez le supprimer avant de créer l'AMI. Ce fichier contient l'adresse MAC de la carte Ethernet de l'instance d'origine. Si une autre instance démarre avec ce fichier, le système d'exploitation ne pourra pas trouver le périphérique et il se peut qu'`eth0` échoue, entraînant des problèmes de démarrage. Le fichier est à nouveau généré au cycle de démarrage suivant et les instances lancées depuis l'AMI créent leur propre version du fichier.

10. Depuis votre ordinateur local, démarrez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : start-instances (AWS CLI) ou Start-EC2Instance (Outils AWS pour Windows

PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez démarrer l'instance dans la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

11. (Facultatif) Connectez-vous à votre instance et vérifiez que le module est installé.

Pour activer la mise en réseau améliorée (instances basées –sur le stockage d'instance)

Suivez la procédure précédente jusqu'à l'étape à laquelle vous avez arrêté l'instance. Créez une nouvelle AMI comme décrit dans Création d'une AMI Linux basée sur le stockage d'instance (p. 136), en veillant à activer l'attribut de mise en réseau améliorée lors de l'enregistrement de l'AMI.

- register-image (AWS CLI)

```
aws ec2 register-image --sriov-net-support simple ...
```

- Register-EC2Image (Outils AWS pour Windows PowerShell)

```
Register-EC2Image -SriovNetSupport "simple" ...
```

Dépannage des problèmes de connexion

Si vous perdez la connexion tout en activant la mise en réseau améliorée, il se peut que le module `ixgbevf` ne soit pas compatible avec le noyau. Essayez d'installer la version du module `ixgbevf` fournie avec la distribution de Linux pour votre instance.

Si vous activez la mise en réseau améliorée pour une instance de paravirtualisation ou une AMI, votre instance peut devenir inaccessible.

Pour de plus amples informations, veuillez consulter la section relative à la procédure d'activation et de configuration de la mise en réseau améliorée sur les instances EC2.

Dépannage d'Elastic Network Adapter (ENA)

Elastic Network Adapter (ENA) est conçu pour améliorer l'intégrité du système d'exploitation et réduire les risques de perturbations à long terme en raison d'un comportement inattendu de matériel ou de défaillances. L'architecture ENA assure une transparence optimale des défaillances de périphériques ou de pilotes auprès du système. Cette rubrique fournit des informations de dépannage pour ENA.

Si vous ne pouvez pas vous connecter à votre instance, commencez par la section Dépannage des problèmes de connexion (p. 885).

Si vous ne parvenez pas à vous connecter à votre instance, recueillez des informations de diagnostic à l'aide des mécanismes de détection des défaillances et de récupération couverts dans des sections ultérieures de cette rubrique.

Sommaire

- Dépannage des problèmes de connexion (p. 885)
- Mécanisme Keep-alive (p. 886)
- Expiration du délai d'attente des opérations de lecture (p. 887)
- Statistiques (p. 887)
- Journaux d'erreur de pilote dans syslog (p. 890)

Dépannage des problèmes de connexion

Si vous perdez la connexion lors de l'activation de la mise en réseau améliorée, il se peut que le module `ena` ne soit pas compatible avec le noyau de votre instance. Cela peut se produire si vous installez le module pour une version de noyau spécifique (sans dkms ou avec un fichier dkms.conf mal configuré), puis que le noyau de votre instance est mis à jour. Si le module `ena` du noyau de l'instance qui est chargé au moment du démarrage n'est pas correctement installé, votre instance ne reconnaît pas la carte réseau et devient inaccessible.

Si vous activez la mise en réseau améliorée pour une instance de paravirtualisation (PV) ou une AMI, votre instance peut devenir inaccessible.

Si votre instance devient inaccessible après l'activation de la mise en réseau améliorée via ENA, vous pouvez désactiver l'attribut `enaSupport` pour votre instance afin qu'elle utilise une autre carte réseau à la place.

Pour désactiver la mise en réseau améliorée via ENA (instances basées sur EBS)

1. Depuis votre ordinateur local, arrêtez votre instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : stop-instances (AWS CLI) ou Stop-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez arrêter l'instance de la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.

   > **Important**
   >
   > Si vous utilisez une instance basée sur le stockage d'instance, vous ne pouvez pas l'arrêter. A la place, passez à Pour désactiver la mise en réseau améliorée via ENA (instances basées sur le stockage d'instance) (p. 885).

2. Depuis votre ordinateur local, désactivez l'attribut de mise en réseau améliorée à l'aide de la commande suivante.

   - modify-instance-attribute (AWS CLI)

   ```
   $ aws ec2 modify-instance-attribute --instance-id instance_id --no-ena-support
   ```

3. Depuis votre ordinateur local, démarrez l'instance à l'aide de la console Amazon EC2 ou de l'une des commandes suivantes : start-instances (AWS CLI) ou Start-EC2Instance (Outils AWS pour Windows PowerShell). Si votre instance est gérée par AWS OpsWorks, vous devez démarrer l'instance dans la console AWS OpsWorks de telle sorte que l'état de l'instance demeure synchronisé.
4. (Facultatif) Connectez-vous à votre instance et essayez de réinstaller le module `ena` avec votre version de noyau actuelle en suivant les étapes décrites dans la section Activation de la mise en réseau améliorée avec l'adaptateur Elastic Network Adapter (ENA) sur les instances Linux (p. 865).

Pour désactiver la mise en réseau améliorée via ENA (instances basées sur le stockage d'instance)

Si votre instance est basée sur le stockage d'instance, créez une AMI, comme décrit dans la section Création d'une AMI Linux basée sur le stockage d'instance (p. 136). Veillez à désactiver l'attribut `enaSupport` de mise en réseau améliorée lorsque vous inscrivez l'AMI.

- register-image (AWS CLI)

```
$ aws ec2 register-image --no-ena-support ...
```

- Register-EC2Image (Outils AWS pour Windows PowerShell)

```
C:\> Register-EC2Image -EnaSupport $false ...
```

Mécanisme Keep-alive

Le dispositif ENA publie des événements keep-alive selon une fréquence fixe (généralement une fois par seconde). Le pilote ENA implémente un mécanisme de surveillance, qui recherche la présence de ces messages keep-alive. Si un ou plusieurs messages sont présents, la surveillance est réarmée. Dans le cas contraire, le pilote conclut que l'appareil a subi une défaillance et effectue alors les opérations suivantes :

- Il envoie ses statistiques dans le journal système.
- Il réinitialise le dispositif ENA.
- Il réinitialise l'état du pilote ENA.

La procédure de réinitialisation ci-dessus peut entraîner une perte de trafic pour une courte période de temps (la récupération des connexions TCP doit être possible), mais ne devrait pas affecter l'utilisateur.

Le dispositif ENA peut également demander indirectement une procédure de réinitialisation de l'appareil. Dans ce cas, il n'envoie pas de notification keep-alive. Cela est possible, par exemple, si le périphérique ENA atteint un état inconnu après le chargement d'une configuration irrécupérable.

Voici un exemple de la procédure de réinitialisation :

```
[18509.800135] ena 0000:00:07.0 eth1: Keep alive watchdog timeout. // The watchdog process
 initiates a reset
[18509.815244] ena 0000:00:07.0 eth1: Trigger reset is on
[18509.825589] ena 0000:00:07.0 eth1: tx_timeout: 0 // The driver logs the current
 statistics
[18509.834253] ena 0000:00:07.0 eth1: io_suspend: 0
[18509.842674] ena 0000:00:07.0 eth1: io_resume: 0
[18509.850275] ena 0000:00:07.0 eth1: wd_expired: 1
[18509.857855] ena 0000:00:07.0 eth1: interface_up: 1
[18509.865415] ena 0000:00:07.0 eth1: interface_down: 0
[18509.873468] ena 0000:00:07.0 eth1: admin_q_pause: 0
[18509.881075] ena 0000:00:07.0 eth1: queue_0_tx_cnt: 0
[18509.888629] ena 0000:00:07.0 eth1: queue_0_tx_bytes: 0
[18509.895286] ena 0000:00:07.0 eth1: queue_0_tx_queue_stop: 0
.......
........
[18511.280972] ena 0000:00:07.0 eth1: free uncompleted tx skb qid 3 idx 0x7 // At the end
 of the down process, the driver discards incomplete packets.
[18511.420112] [ENA_COM: ena_com_validate_version] ena device version: 0.10 //The driver
 begins its up process
[18511.420119] [ENA_COM: ena_com_validate_version] ena controller version: 0.0.1
 implementation version 1
[18511.420127] [ENA_COM: ena_com_admin_init] ena_defs : Version:[b9692e8] Build date [Wed
 Apr  6 09:54:21 IDT 2016]
[18512.252108] ena 0000:00:07.0: Device watchdog is Enabled
[18512.674877] ena 0000:00:07.0: irq 46 for MSI/MSI-X
[18512.674933] ena 0000:00:07.0: irq 47 for MSI/MSI-X
[18512.674990] ena 0000:00:07.0: irq 48 for MSI/MSI-X
[18512.675037] ena 0000:00:07.0: irq 49 for MSI/MSI-X
[18512.675085] ena 0000:00:07.0: irq 50 for MSI/MSI-X
[18512.675141] ena 0000:00:07.0: irq 51 for MSI/MSI-X
[18512.675188] ena 0000:00:07.0: irq 52 for MSI/MSI-X
[18512.675233] ena 0000:00:07.0: irq 53 for MSI/MSI-X
[18512.675279] ena 0000:00:07.0: irq 54 for MSI/MSI-X
[18512.772641] [ENA_COM: ena_com_set_hash_function] Feature 10 isn't supported
[18512.772647] [ENA_COM: ena_com_set_hash_ctrl] Feature 18 isn't supported
[18512.775945] ena 0000:00:07.0: Device reset completed successfully // The reset process
 is complete
```

Expiration du délai d'attente des opérations de lecture

L'architecture ENA propose une utilisation limitée des opérations de lecture des E/S mappées par la mémoire (MMIO). Le pilote du périphérique ENA n'accède aux registres MMIO que lors de la procédure d'initialisation.

Si les journaux du pilote (disponibles dans la sortie dmesg) indiquent une défaillance des opérations de lecture, un pilote incompatible ou mal compilé, un dispositif saturé ou une défaillance matérielle peuvent en être la cause.

Les entrées de journal intermittentes qui indiquent des défaillances des opérations de lecture ne sont pas problématiques. Dans ce cas, le pilote réessaie de les traiter. Toutefois, une série d'entrées de journal contenant des défaillances de lecture indique un problème de pilote ou de matériel.

Voici un exemple d'entrée de journal pilote indiquant une défaillance des opérations de lecture en raison de l'expiration d'un délai d'attente :

```
[ 47.113698] [ENA_COM: ena_com_reg_bar_read32] reading reg failed for timeout. expected:
 req id[1] offset[88] actual: req id[57006] offset[0]
[ 47.333715] [ENA_COM: ena_com_reg_bar_read32] reading reg failed for timeout. expected:
 req id[2] offset[8] actual: req id[57007] offset[0]
[ 47.346221] [ENA_COM: ena_com_dev_reset] Reg read32 timeout occurred
```

Statistiques

Si vous rencontrez des problèmes de latence ou si les performances réseau sont insuffisantes, vous devez récupérer les statistiques de l'appareil et les examiner. Pour obtenir ces statistiques, utilisez ethtool, comme indiqué ci-dessous :

```
[ec2-user ~]$ ethtool -S ethN
NIC statistics:
     tx_timeout: 0
     io_suspend: 0
     io_resume: 0
     wd_expired: 0
     interface_up: 1
     interface_down: 0
     admin_q_pause: 0
     queue_0_tx_cnt: 4329
     queue_0_tx_bytes: 1075749
     queue_0_tx_queue_stop: 0
...
```

Les paramètres de sortie de commande suivants sont décrits ci-dessous :

`tx_timeout: N`

Nombre de fois que la surveillance Netdev a été activée.

`io_suspend: N`

Non pris en charge. Cette valeur doit toujours correspondre à zéro.

`io_resume: N`

Non pris en charge. Cette valeur doit toujours correspondre à zéro.

`wd_expired: N`

Nombre de fois que le pilote n'a pas reçu l'événement keep-alive au cours des trois secondes précédentes.

`interface_up: N`

Nombre de fois que l'interface ENA a été affichée.

`interface_down: N`

Nombre de fois que l'interface ENA a été fermée.

`admin_q_pause: N`

L'état de la file d'attente d'administration est instable. Cette valeur doit toujours correspondre à zéro.

`queue_N_tx_cnt: N`

Nombre de paquets transmis pour la file d'attente *N*.

`queue_N_tx_bytes: N`

Nombre d'octets transmis pour la file d'attente *N*.

`queue_N_tx_queue_stop: N`

Nombre de fois que la file d'attente *N* était pleine et qu'elle a été arrêtée.

`queue_N_tx_queue_wakeup: N`

Nombre de fois que la file d'attente *N* a repris après avoir été arrêtée.

`queue_N_tx_dma_mapping_err: N`

Nombre d'erreurs d'accès direct à la mémoire. Si cette valeur ne correspond pas à 0, les ressources système sont faibles.

`queue_N_tx_napi_comp: N`

Nombre de fois que le gestionnaire `napi` a appelé `napi_complete` pour la file d'attente *N*.

`queue_N_tx_poll: N`

Nombre de fois que le gestionnaire `napi` a été planifié pour la file d'attente *N*.

`queue_N_tx_doorbells: N`

Nombre de portes de transmission pour la file d'attente *N*.

`queue_N_tx_linearize: N`

Nombre de fois que la linéarisation SKB a été tentée pour la file d'attente *N*.

`queue_N_tx_linearize_failed: N`

Nombre de fois que la linéarisation SKB a échoué pour la file d'attente *N*.

`queue_N_tx_prepare_ctx_err: N`

Nombre de fois que `ena_com_prepare_tx` a échoué pour la file d'attente *N*. Cette valeur doit toujours être égale à zéro. Si ce n'est pas le cas, consultez les journaux du pilote.

`queue_N_tx_missing_tx_comp: codeN`

Nombre de paquets qui n'ont pas été traités entièrement pour la file d'attente *N*. Cette valeur doit toujours correspondre à zéro.

`queue_N_tx_bad_req_id: N`

Valeur `req_id` non valide pour la file d'attente *N*. La valeur `req_id` valide est égale à zéro, moins la valeur `queue_size`, moins 1.

`queue_N_rx_cnt: N`

Nombre de paquets reçus pour la file d'attente *N*.

`queue_N_rx_bytes: N`

Nombre d'octets reçus pour la file d'attente *N*.

`queue_N_rx_refil_partial: N`

Nombre de fois que le pilote n'a pas réussi à remplir la portion vide de la file d'attente `rx` avec les tampons pour la file d'attente *N*. Si cette valeur n'est pas égale à zéro, les ressources mémoire sont faibles.

`queue_N_rx_bad_csum: N`

Nombre de fois que la file d'attente `rx` a reçu un mauvais total de contrôle pour la file d'attente *N* (uniquement si le déchargement du total de contrôle `rx` est pris en charge).

`queue_N_rx_page_alloc_fail: N`

Nombre de fois que l'allocation des pages a échoué pour la file d'attente *N*. Si cette valeur n'est pas égale à zéro, les ressources mémoire sont faibles.

`queue_N_rx_skb_alloc_fail: N`

Nombre de fois que l'allocation SKB a échoué pour la file d'attente *N*. Si cette valeur n'est pas égale à zéro, les ressources système sont faibles.

`queue_N_rx_dma_mapping_err: N`

Nombre d'erreurs d'accès direct à la mémoire. Si cette valeur ne correspond pas à 0, les ressources système sont faibles.

`queue_N_rx_bad_desc_num: N`

Trop de tampons par paquet. Si cette valeur n'est pas égale à 0, cela indique l'utilisation de très petits tampons.

`queue_N_rx_small_copy_len_pkt: N`

Optimisation : pour les paquets qui se trouvent sous ce seuil, ce qui est défini par `sysfs`, le paquet est copié directement dans la stack pour éviter l'allocation d'une nouvelle page.

`ena_admin_q_aborted_cmd: N`

Nombre de commandes d'administration qui ont été abandonnées. Généralement, cela se produit lors de la procédure de récupération automatique.

`ena_admin_q_submitted_cmd: N`

Nombre de portes d'administration de la file d'attente.

`ena_admin_q_completed_cmd: N`

Nombre de finalisations de la file d'attente d'administration.

`ena_admin_q_out_of_space: N`

Nombre de fois que le pilote a essayé de présenter la nouvelle commande d'administration, mais que la file d'attente était pleine.

`ena_admin_q_no_completion: N`

Nombre de fois que l'administration du pilote n'a pas été terminée pour une commande.

Journaux d'erreur de pilote dans syslog

Le pilote ENA écrit les messages journaux dans syslog pendant le démarrage du système. Vous pouvez examiner ces journaux pour rechercher les erreurs si vous rencontrez des problèmes. Voici un exemple d'informations enregistrées par le pilote ENA dans syslog pendant le démarrage du système, ainsi que des annotations pour certains messages.

```
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  478.416939] [ENA_COM: ena_com_validate_version]
 ena device version: 0.10
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  478.420915] [ENA_COM: ena_com_validate_version]
 ena controller version: 0.0.1 implementation version 1
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.256831] ena 0000:00:03.0: Device watchdog is
 Enabled
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.672947] ena 0000:00:03.0: creating 8 io
 queues. queue size: 1024
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.680885] [ENA_COM:
 ena_com_init_interrupt_moderation] Feature 20 isn't supported  // Interrupt moderation is
 not supported by the device
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.691609] [ENA_COM: ena_com_get_feature_ex]
 Feature 10 isn't supported // RSS HASH function configuration is not supported by the
 device
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.694583] [ENA_COM: ena_com_get_feature_ex]
 Feature 18 isn't supported //RSS HASH input source configuration is not supported by the
 device
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.697433] [ENA_COM:
 ena_com_set_host_attributes] Set host attribute isn't supported
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.701064] ena 0000:00:03.0 (unnamed
 net_device) (uninitialized): Cannot set host attributes
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  479.704917] ena 0000:00:03.0: Elastic Network
 Adapter (ENA) found at mem f3000000, mac addr 02:8a:3c:1e:13:b5 Queues 8
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  480.805037] EXT4-fs (xvda1): re-mounted. Opts:
 (null)
Jun  3 22:37:46 ip-172-31-2-186 kernel: [  481.025842] NET: Registered protocol family 10
```

Quelles sont les erreurs que je peux ignorer ?

Les avertissements suivants qui peuvent apparaître dans les journaux d'erreur de votre système peuvent être ignorées pour Elastic Network Adapter :

Set host attribute isn't supported

Les attributs de l'hôte ne sont pas pris en charge pour cet appareil.

failed to alloc buffer for rx queue

Il s'agit d'une erreur récupérable. Elle indique qu'il y a peut-être eu n problème de pression de mémoire lorsque l'erreur a été lancée.

Feature *X* isn't supported

La fonctionnalité référencée n'est pas prise en charge par Elastic Network Adapter. Les valeurs possibles pour *X* incluent :

- **10** : la configuration de la fonction de hachage RSS n'est pas prise en charge pour cet appareil.
- **12** : la configuration de la table d'indirection RSS n'est pas prise en charge pour cet appareil.
- **18** : la configuration des entrées de hachage RSS n'est pas prise en charge pour cet appareil.
- **20** : la modération d'interruption n'est pas prise en charge pour cet appareil.
- **27** : le pilote ENA (Elastic Network Adapter) ne prend pas en charge l'interrogation des fonctions Ethernet à partir de snmpd.

Failed to config AENQ

Elastic Network Adapter ne prend pas en charge la configuration AENQ.

Trying to set unsupported AENQ events

Cette erreur indique une tentative de définition d'un groupe d'événements AENQ qui n'est pas pris en charge par Elastic Network Adapter.

Elastic Fabric Adapter (EFA)

Elastic Fabric Adapter (EFA) (EFA) est un périphérique réseau que vous pouvez attacher à votre instance Amazon EC2 pour accélérer les applications HPC (Calcul Haute Performance). EFA vous permet d'atteindre les performances d'application d'un cluster HPC sur site, avec l'évolutivité, la flexibilité et l'élasticité fournies par le cloud AWS.

EFA offre une latence plus faible et plus cohérente avec un débit plus élevé que le transport TCP utilisé traditionnellement dans des systèmes HPC basés sur le cloud. Il améliore les performances des communications entre instances, ce qui est essentiel pour la mise à l'échelle des applications HPC et de Machine Learning. Il est optimisé pour fonctionner sur l'infrastructure réseau AWS existante et peut être mis à l'échelle en fonction des exigences des applications.

EFA s'intègre à Libfabric 1.11.1 et prend en charge Open MPI 4.0.5 et Intel MPI Mise à jour 7 de 2019 pour les applications HPC, et Nvidia Collective Communications Library (NCCL) pour les applications de Machine Learning.

Note

Les capacités de contournement du système d'exploitation d'EFAs ne sont pas prises en charge sur des instances Windows. Si vous attachez un EFA à une instance Windows, l'instance fonctionne en tant qu'adaptateur Elastic Network Adapter sans les capacités EFA ajoutées.

Sommaire

- Principes de base EFA (p. 891)
- Interfaces et bibliothèques prises en charge (p. 893)
- Types d'instance pris en charge (p. 893)
- AMIs prises en charge (p. 893)
- Restrictions liées à EFA (p. 894)
- Démarrer avec EFA et MPI (p. 894)
- Démarrer avec EFA et NCCL (p. 902)
- Utilisation d’EFA (p. 920)
- Surveillance d'un EFA (p. 922)
- Vérification du programme d'installation EFA à l'aide d'un total de contrôle (p. 923)

Principes de base EFA

Un EFA est un adaptateur Elastic Network Adapter (ENA) avec des capacités ajoutées. Il offre toutes les fonctionnalités d'un ENA, avec des capacités de contournement du système d'exploitation supplémentaires. Le contournement du système d'exploitation est un modèle d'accès qui permet aux applications HPC et de Machine Learning de communiquer directement avec le matériel de l'interface réseau pour offrir des fonctionnalités de transport fiable à faible latence.

Traditional HPC software stack in EC2

Traditionnellement, les applications HPC utilisent Message Passing Interface (MPI) pour servir d'interface avec le transport réseau du système. Dans le cloud AWS, cela signifiait que les applications communiquaient avec MPI, qui utilisait alors la pile TCP/IP du système d'exploitation et le pilote de périphérique ENA pour permettre la communication réseau entre les instances.

Avec EFA, les applications HPC ou NCCL utilisent MPI pour servir d'interface avec l'API Libfabric. L'API Libfabric contourne le noyau du système d'exploitation et communique directement avec l'appareil EFA pour placer les paquets sur le réseau. Cela réduit la surcharge et permet à l'application HPC de s'exécuter plus efficacement.

> **Note**
>
> Libfabric est un composant de base de l'infrastructure OpenFabrics Interfaces (OFI), qui définit et exporte l'API d'espace utilisateur d'OFI. Pour plus d'informations, consultez le site web de Libfabric OpenFabrics.

Différences entre les EFAs et les adaptateurs ENA

Les adaptateurs Elastic Network Adapter (ENAs) fournissent les fonctionnalités de mise en réseau IP classiques qui sont requises pour prendre en charge la mise en réseau VPC. Les adaptateurs EFAs fournissent les mêmes fonctionnalités de mise en réseau IP que les adaptateurs ENA, mais ils prennent

également en charge les capacités de contournement du système d'exploitation. Le contournement du système d'exploitation permet aux applications HPC et de Machine Learning de contourner le noyau du système d'exploitation et de communiquer directement avec l'appareil EFA.

Interfaces et bibliothèques prises en charge

EFA prend en charge les interfaces et bibliothèques suivantes :

- Open MPI 4.0.5
- Intel MPI Mise à jour 7 de 2019
- NVIDIA Collective Communications Library (NCCL) 2.4.2 et versions ultérieures

Types d'instance pris en charge

Les types d'instance suivants prennent en charge EFAs :

- Usage général : `m5dn.24xlarge` | `m5n.24xlarge`
- Calcul optimisé : `c5n.18xlarge` | `c5n.metal`
- Mémoire optimisée: `r5dn.24xlarge` | `r5n.24xlarge`
- Stockage optimisé : `i3en.24xlarge` | `i3en.metal`
- Calcul accéléré : `g4dn.metal` | `inf1.24xlarge` | `p3dn.24xlarge`

Les types d'instance disponibles varient selon la région. Pour voir les types d'instance disponibles qui prennent EFA en charge dans une région, utilisez la commande describe-instance-types avec l'option `--region` et le code régional approprié.

```
aws ec2 describe-instance-types --region us-east-2 --filters Name=network-info.efa-
supported,Values=true --query InstanceTypes[*].[InstanceType] --output text
```

Voici un exemple de sortie.

```
g4dn.metal
i3en.24xlarge
r5n.24xlarge
c5n.18xlarge
m5n.24xlarge
inf1.24xlarge
m5dn.24xlarge
c5n.metal
p3dn.24xlarge
i3en.metal
r5dn.24xlarge
```

AMIs prises en charge

LesAMI suivantes prennent en charge les EFAs :

- Amazon Linux et Amazon Linux 2
- CentOS 7
- RHEL 7.6, 7.7 et 7.8
- Ubuntu 16.04 et 18.04

Restrictions liées à EFA

Les restrictions suivantes s'appliquent à EFA :

- Vous ne pouvez attacher qu'un seul EFA par instance.
- Le trafic de contournement du système d'exploitation EFA est limité à un seul sous-réseau. En d'autres termes, le trafic EFA ne peut pas être envoyé d'un sous-réseau à un autre. Le trafic IP normal de l'EFA peut être envoyé d'un sous-réseau à un autre.
- Le trafic de contournement du système d'exploitation EFA n'est pas routable. Le trafic IP normal de l'EFA reste routable.
- L'EFA doit appartenir à un groupe de sécurité qui autorise tout le trafic entrant et sortant vers et depuis le groupe de sécurité proprement dit.

Démarrer avec EFA et MPI

Ce didacticiel vous permet de lancer un cluster d'instances EFA et compatible MPI pour les charges de travail HPC. Dans ce didacticiel, vous exécuterez les étapes suivantes :

Sommaire

- Étape 1 : Préparer un groupe de sécurité activé pour les EFA (p. 894)
- Étape 2 : Lancer une instance temporaire (p. 895)
- Étape 3 : Installer le logiciel EFA (p. 895)
- Étape 4 : Désactiver la protection ptrace (p. 898)
- Étape 5 : (Facultatif) Installer Intel MPI (p. 899)
- Étape 6 : Installer votre application HPC (p. 900)
- Étape 7 : Créer uneAMI activée pour EFA (p. 900)
- Étape 8 : Lancer des instances activées pour EFA dans un groupe de placement de cluster (p. 900)
- Étape 9 : Résilier l'instance temporaire (p. 901)
- Étape 10 : Activer SSH sans mot de passe (p. 902)

Étape 1 : Préparer un groupe de sécurité activé pour les EFA

L'EFA a besoin d'un groupe de sécurité qui autorise tout le trafic entrant et sortant vers et depuis le groupe de sécurité proprement dit. La procédure suivante autorise tout le trafic entrant et sortant à des fins de test uniquement. Pour d'autres scénarios, veuillez consulter Référence des règles de groupe de sécurité (p. 1075).

Pour créer un groupe de sécurité activé pour EFA

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité, puis Créer un groupe de sécurité.
3. Dans la fenêtre Créer un groupe de sécurité, procédez comme suit :
   a. Pour Nom du groupe de sécurité, entrez un nom descriptif pour le groupe de sécurité, tel que `EFA-enabled security group`.
   b. (Facultatif) Dans le champ Description, saisissez une brève description du groupe de sécurité.
   c. Pour VPC, sélectionnez le VPC dans lequel vous prévoyez de lancer vos instances activées pour EFA.
   d. Sélectionnez Créer.

4. Sélectionnez le groupe de sécurité que vous avez créé et dans l'onglet Description, copiez l'ID du groupe.
5. Dans l'onglet Entrant procédez comme suit :
   a. Choisissez Modifier.
   b. Pour Type, sélectionnez Tout le trafic.
   c. Dans Source, choisissez Personnaliser et collez l'ID du groupe de sécurité que vous avez copié dans le champ.
   d. Choisissez Enregistrer.
6. Dans l'onglet Sortant procédez comme suit :
   a. Choisissez Modifier.
   b. Pour Type, sélectionnez Tout le trafic.
   c. Pour Destination, choisissez Personnaliser et collez l'ID du groupe de sécurité que vous avez copié dans le champ.
   d. Choisissez Save.

Étape 2 : Lancer une instance temporaire

Lancez une instance temporaire que vous pouvez utiliser pour installer et configurer les composants logiciels EFA. Vous utilisez cette instance pour créer une AMI activée pour EFA depuis laquelle vous pouvez lancer vos instances activées pour EFA.

Pour lancer une instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Sur la page Choisir une AMI, choisissez Sélectionner pour l'une des AMIs prises en charge (p. 893).
4. Sur la page Choisir un type d'instance, sélectionnez l'un des types d'instance pris en charge (p. 893), puis choisissez Suivant : Configurer les détails de l'instance.
5. Sur la page Configurer les détails de l'instance, procédez de la façon suivante :
   a. Pour Sous-réseau, choisissez le sous-réseau dans lequel lancer l'instance.
   b. Pour Elastic Fabric Adapter (EFA), choisissez Enable (Activer).
   c. Dans la section Interfaces réseau, pour l'appareil eth0, choisissez Nouvelle interface réseau.
   d. Choisissez Suivant : Ajouter un stockage.
6. Sur la page Add Storage (Ajouter un stockage), spécifiez les volumes à attacher aux instances, outre ceux spécifiés par l'AMI (par exemple, le volume du périphérique racine). Choisissez ensuite Suivant : Ajouter des balises.
7. Sur la page Ajouter des balises, spécifiez une balise que vous pouvez utiliser pour identifier l'instance temporaire, puis choisissez }Suivant : Configurer le groupe de sécurité.
8. Sur la page Configurer le groupe de sécurité, cliquez sur Attribuer un groupe de sécurité, choisissez Sélectionner un groupe de sécurité existant, puis le groupe de sécurité que vous avez créé à l'étape 1.
9. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis sélectionnez Lancer pour choisir une paire de clés et lancer votre instance.

Étape 3 : Installer le logiciel EFA

Installez le noyau activé pour EFA, les pilotes EFA, Libfabric et la pile Open MPI qui est requise pour prendre en charge EFA sur votre instance temporaire.

Les étapes varient selon que vous avez l'intention d'utiliser EFA avec Open MPI ou avec Intel MPI, ou avec Open MPI et Intel MPI.

Pour installer le logiciel EFA

1. Connectez-vous à l'instance que vous avez lancée. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Pour vous assurer que tous vos packages logiciels sont mis à jour, effectuez une mise à jour logicielle rapide sur votre instance. Ce processus peut prendre quelques minutes.
   - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

     ```
     $ sudo yum update -y
     ```

   - Ubuntu 16.04 et Ubuntu 18.04

     ```
     $ sudo apt-get update -y
     ```

     ```
     $ sudo apt-get upgrade -y
     ```

3. Téléchargez les fichiers d'installation du logiciel EFA. Les fichiers d'installation du logiciel sont packagés dans un fichier tarball compressé (`.tar.gz`). Pour télécharger la version stable la plus récente, utilisez la commande suivante.

   ```
   $ curl -O https://efa-installer.amazonaws.com/aws-efa-installer-1.10.0.tar.gz
   ```

   Vous pouvez aussi obtenir la dernière version en remplaçant le numéro de version par `latest` dans la commande ci-dessus.
4. (Facultatif) Vérifiez l'authenticité et l'intégrité du fichier tarball EFA (`.tar.gz`). Nous vous recommandons de le faire pour vérifier l'identité de l'éditeur du logiciel et pour vérifier que le fichier n'a pas été modifié ou endommagé depuis sa publication. Si vous ne souhaitez pas vérifier le fichier tarball, ignorez cette étape.

   Note

   Par ailleurs, si vous préférez vérifier le fichier tarball à l'aide d'un total de contrôle MD5 ou SHA256 à la place, veuillez consulter Vérification du programme d'installation EFA à l'aide d'un total de contrôle (p. 923).

   a. Téléchargez la clé GPG publique et importez-la dans votre porte-clés.

      ```
      $ wget https://efa-installer.amazonaws.com/aws-efa-installer.key
      ```

      ```
      $ gpg --import aws-efa-installer.key
      ```

      La commande doit renvoyer une valeur de clé. Notez la valeur de la clé, car vous en aurez besoin lors de l'étape suivante.

   b. Vérifiez l'empreinte digitale de la clé GPG. Exécutez la commande suivante et spécifiez la valeur de clé de l'étape précédente.

      ```
      $ gpg --fingerprint key_value
      ```

La commande doit renvoyer une empreinte digitale identique à `4E90 91BC BB97 A96B 26B1 5E59 A054 80B1 DD2D 3CCC`. Si l'empreinte digitale ne correspond pas, n'exécutez pas le script d'installation EFA et contactez AWS Support.

c. Téléchargez le fichier de signature et vérifiez la signature du fichier tarball EFA.

```
$ wget https://efa-installer.amazonaws.com/aws-efa-installer-1.10.0.tar.gz.sig
```

```
$ gpg --verify ./aws-efa-installer-1.10.0.tar.gz.sig
```

Voici un exemple de sortie.

```
gpg: Signature made Wed 29 Jul 2020 12:50:13 AM UTC using RSA key ID DD2D3CCC
gpg: Good signature from "Amazon EC2 EFA <ec2-efa-maintainers@amazon.com>"
gpg: WARNING: This key is not certified with a trusted signature!
gpg:          There is no indication that the signature belongs to the owner.
Primary key fingerprint: 4E90 91BC BB97 A96B 26B1  5E59 A054 80B1 DD2D 3CCC
```

Si le résultat comprend `Good signature` et l'empreinte correspond à celle renvoyée à l'étape précédente, passez à l'étape suivante. Si ce n'est pas le cas, n'exécutez pas le script d'installation EFA et contactez AWS Support.

5. Procédez à l'extraction des fichiers à partir du fichier compressé `.tar.gz` et accédez au répertoire extrait.

```
$ tar -xf aws-efa-installer-1.10.0.tar.gz
```

```
$ cd aws-efa-installer
```

6. Installez le logiciel EFA. Effectuez l'une des actions suivantes en fonction de votre cas d'utilisation.

- Open MPI et Intel MPI

  Si vous avez l'intention d'utiliser EFA avec Open MPI et Intel MPI, vous devez installer le logiciel EFA avec Libfabric et Open MPI, et vous devez réaliser l'Étape 5 : (Facultatif) Installer Intel MPI. Pour installer le logiciel EFA avec Libfabric et Open MPI, exécutez la commande suivante.

```
$ sudo ./efa_installer.sh -y
```

  Libfabric est installé dans le répertoire `/opt/amazon/efa`, tandis qu'Open MPI est installé dans le répertoire `/opt/amazon/openmpi`.

- Open MPI uniquement

  Si vous avez l'intention d'utiliser EFA avec Open MPI uniquement, vous devez installer le logiciel EFA avec Libfabric et Open MPI, et vous pouvez ignorer l'Étape 5 : (Facultatif) Installer Intel MPI. Pour installer le logiciel EFA avec Libfabric et Open MPI, exécutez la commande suivante.

```
$ sudo ./efa_installer.sh -y
```

  Libfabric est installé dans le répertoire `/opt/amazon/efa`, tandis qu'Open MPI est installé dans le répertoire `/opt/amazon/openmpi`.

- Intel MPI uniquement

Si vous avez l'intention d'utiliser EFA uniquement avec Intel MPI, vous pouvez installer le logiciel EFA sans Libfabric ni Open MPI. Dans ce cas, Intel MPI utilise son Libfabric intégré. Si vous optez pour cette solution, vous devez effectuer l'Étape 5 : (Facultatif) Installer Intel MPI.

Pour installer le logiciel EFA sans Libfabric ni Open MPI, exécutez la commande suivante.

```
$ sudo ./efa_installer.sh -y --minimal
```

7. Déconnectez-vous de l'instance, puis reconnectez-vous.
8. Vérifiez que les composants logiciels EFA ont été installés avec succès.

```
$ fi_info -p efa
```

La commande doit renvoyer des informations sur les interfaces EFA Libfabric. L'exemple suivant illustre la sortie de la commande.

```
provider: efa
    fabric: EFA-fe80::94:3dff:fe89:1b70
    domain: efa_0-rdm
    version: 2.0
    type: FI_EP_RDM
    protocol: FI_PROTO_EFA
provider: efa
    fabric: EFA-fe80::94:3dff:fe89:1b70
    domain: efa_0-dgrm
    version: 2.0
    type: FI_EP_DGRAM
    protocol: FI_PROTO_EFA
provider: efa;ofi_rxd
    fabric: EFA-fe80::94:3dff:fe89:1b70
    domain: efa_0-dgrm
    version: 1.0
    type: FI_EP_RDM
    protocol: FI_PROTO_RXD
```

Étape 4 : Désactiver la protection ptrace

Pour améliorer les performances de votre application HPC, Libfabric utilise la mémoire locale de l'instance pour les communications interprocessus lorsque les processus s'exécutent sur la même instance.

La fonction de mémoire partagée utilise Cross Memory Attach (CMA), non pris en charge avec la protection ptrace. Si vous utilisez une distribution Linux dans laquelle la protection ptrace est activée par défaut, telle que Ubuntu, vous devez la désactiver. Si la protection ptrace n'est pas activée par défaut dans votre distribution Linux, ignorez cette étape.

Pour désactiver la protection ptrace

Effectuez l'une des actions suivantes :

- Pour désactiver temporairement la protection ptrace à des fins de test, exécutez la commande suivante.

```
$ sudo sysctl -w kernel.yama.ptrace_scope=0
```

- Pour désactiver définitivement la protection ptrace, ajoutez `kernel.yama.ptrace_scope = 0`à `/etc/sysctl.d/10-ptrace.conf` et redémarrez l'instance.

Étape 5 : (Facultatif) Installer Intel MPI

> **Important**
>
> Si vous avez l'intention d'utiliser uniquement Open MPI, passez cette étape. Vous ne devez effectuer cette étape que si vous avez l'intention d'utiliser Intel MPI.

Intel MPI nécessite une installation et une configuration de variable d'environnement supplémentaires.

Prérequis

Vérifiez que l'utilisateur qui exécute les étapes suivantes dispose des autorisations sudo.

Pour installer Intel MPI

1. Pour télécharger les fichiers d'installation d'Intel MPI, veuillez consulter le site web Intel Developer Zone.

   Vous devez vous enregistrer pour pouvoir télécharger les fichiers d'installation. Une fois que vous êtes enregistré, procédez comme suit :

   a. Pour Product (Produit), choisissez Intel MPI Library for Linux.
   b. Pour Version, choisissez Mise à jour 7 de 2019, puis Full Product.
2. Les fichiers d'installation sont packagés dans un fichier `.tar.gz` compressé. Procédez à l'extraction des fichiers à partir du fichier compressé `.tar.gz` et accédez au répertoire extrait.

   ```
   $ tar -xf file_name.tgz
   ```

   ```
   $ cd directory_name
   ```
3. Ouvrez `silent.cfg` à l'aide de votre éditeur de texte préféré. Sur la ligne 10, remplacez `ACCEPT_EULA=decline` par `ACCEPT_EULA=accept`. Enregistrez les modifications, puis fermez le fichier.
4. Exécutez le script d'installation.

   ```
   $ sudo ./install.sh -s silent.cfg
   ```

   Intel MPI est installé dans le répertoire `/opt/intel/impi/` par défaut.
5. Ajoutez les variables d'environnement Intel MPI aux scripts de démarrage de shell correspondants afin de vous assurer qu'ils sont définis à chaque démarrage de l'instance. Effectuez l'une des actions suivantes en fonction de votre shell.

   - Pour bash, ajoutez la variable d'environnement suivante à `/home/username/.bashrc` et `/home/username/.bash_profile`.

     ```
     source /opt/intel/compilers_and_libraries/linux/mpi/intel64/bin/mpivars.sh
     ```
   - Pour csh et tcsh, ajoutez la variable d'environnement suivante à `/home/username/.cshrc`.

     ```
     source /opt/intel/compilers_and_libraries/linux/mpi/intel64/bin/mpivars.csh
     ```
6. Déconnectez-vous de l'instance, puis reconnectez-vous.
7. Exécutez la commande suivante pour vérifier qu'Intel MPI a été installé avec succès.

   ```
   $ which mpicc
   ```

Assurez-vous que le chemin d'accès renvoyé inclut le sous-répertoire `/opt/intel/`.

Note

Si vous ne souhaitez plus utiliser Intel MPI, supprimez les variables d'environnement des scripts de démarrage de shell.

Étape 6 : Installer votre application HPC

Installez l'application HPC sur l'instance temporaire. La procédure d'installation varie selon l'application HPC. Pour plus d'informations sur l'installation du logiciel sur votre instance Linux, consultez Gestion de logiciels sur votre instance Linux.

Note

Vous pouvez avoir besoin de vous reporter à la documentation de votre application HPC pour obtenir des instructions d'installation.

Étape 7 : Créer uneAMI activée pour EFA

Une fois que vous avez installé les composants logiciels requis, vous devez créer une AMI que vous pouvez réutiliser pour lancer vos instances activées pour EFA.

Pour créer une AMI à partir de votre instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance temporaire que vous avez créée et choisissez Actions, Image, Create image (Créer une image).
4. Pour Create image (Créer une image), procédez comme suit :
   a. Pour Image name (Nom de l'image), entrez un nom descriptif pour l'AMI.
   b. (Facultatif) Dans le champ Image description (Description de l'image), saisissez une brève description de l'objet de l'AMI.
   c. Choisissez Create image (Créer une image).
5. Dans le volet de navigation, choisissez AMIs.
6. Recherchez l'AMI que vous avez créée dans la liste. Attendez que le statut passe de `pending` à `available` avant de poursuivre avec l'étape suivante.

Étape 8 : Lancer des instances activées pour EFA dans un groupe de placement de cluster

Lancez vos instances activées pour EFA dans un groupe de placement de cluster à l'aide de l'AMI activée pour EFA que vous avez créée à l'Étape 7 et le groupe de sécurité activé pour EFA que vous avez créé à l'Étape 1.

Note

Vous ne devez pas impérativement lancer vos instances EFA dans un groupe de placement de cluster. Toutefois, nous vous recommandons d'exécuter vos instances activées pour EFA dans un

groupe de placement de cluster, car cela lance celles-ci dans un groupe à faible latence au sein d'une zone de disponibilité unique.

Pour lancer des instances activées pour EFA dans un groupe de placement de cluster

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Sur la page Choisir une AMI, choisissez Mes AMI, recherchez l'AMI que vous avez créée à l'Étape 7, puis choisissez Sélectionner.
4. Sur la page Choisir un type d'instance, sélectionnez l'un des types d'instance pris en charge (p. 893), puis choisissez Suivant : Configurer les détails de l'instance.
5. Sur la page Configurer les détails de l'instance, procédez de la façon suivante :
   a. Pour Nombre d'instances, entrez le nombre d'instances activées pour EFA que vous voulez lancer.
   b. Pour Réseau et Sous-réseau, sélectionnez le VPC et le sous-réseau dans lesquels lancer les instances.
   c. Pour le Groupe de placement, sélectionnez Ajoutez une instance au groupe de placement.
   d. Pour Nom du groupe de placement, sélectionnez Ajoutez à un nouveau groupe de placement, entrez un nom descriptif pour le groupe de placement, puis pour Stratégie de groupe de placement, choisissez cluster.
   e. Pour EFA, choisissez Enable (Activer).
   f. Dans la section Interfaces réseau, pour l'appareil eth0, choisissez Nouvelle interface réseau. Vous pouvez éventuellement entrer une adresse IPv4 principale et une ou plusieurs adresses IPv4 secondaires. Si vous lancez l'instance dans un sous-réseau auquel un bloc CIDR IPv6 est associé, vous pouvez éventuellement spécifier une adresse IPv6 principale et une ou plusieurs adresses IPv6 secondaires.
   g. Choisissez Suivant : Ajouter un stockage.
6. Sur la page Ajouter un stockage, spécifiez les volumes à attacher aux instances, outre ceux spécifiés par l'AMI (par exemple, le volume du périphérique racine), puis sélectionner Suivant : Ajouter des balises.
7. Sur la page Ajouter des balises, spécifiez des balises pour l'instance, par exemple un nom évocateur, puis sélectionnez Suivant : Configurer le groupe de sécurité.
8. Sur la page Configurer le groupe de sécurité, cliquez sur Attribuer un groupe de sécurité, choisissez Sélectionner un groupe de sécurité existant, puis le groupe de sécurité que vous avez créé à l'étape 1.
9. Choisissez Vérifier et lancer.
10. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis choisissez Lancer pour sélectionner une paire de clés et lancer votre instance.

Étape 9 : Résilier l'instance temporaire

À ce stade, vous n'avez plus besoin de l'instance temporaire que vous avez lancée. Vous pouvez résilier l'instance pour ne plus payer celle-ci.

Pour résilier l'instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance temporaire, puis choisissez Actions, Instance state (État de l'instance), Terminate instance (Mettre fin à l'instance).
4. Lorsque vous êtes invité à confirmer, choisissez Terminate (Mettre fin).

Étape 10 : Activer SSH sans mot de passe

Pour permettre à vos applications de s'exécuter sur toutes les instances de votre cluster, vous devez activer l'accès SSH sans mot de passe entre le nœud principal et les nœuds membres. Le nœud principal est l'instance à partir de laquelle vous exécutez vos applications. Les instances restantes du cluster sont les nœuds membres.

Pour activer l'accès SSH sans mot de passe entre les instances du cluster

1. Sélectionnez une instance du cluster en tant que nœud principal et connectez-vous y.
2. Désactivez `strictHostKeyChecking` et activez `ForwardAgent` sur le nœud principal. Ouvrez `~/.ssh/config` à l'aide de l'éditeur de texte de votre choix et ajoutez ce qui suit.

```
Host *
    ForwardAgent yes
Host *
    StrictHostKeyChecking no
```

3. Générez une paire de clés RSA.

```
$ ssh-keygen -t rsa -N "" -f ~/.ssh/id_rsa
```

   La paire de clés est créée dans le répertoire `$HOME/.ssh/`.
4. Modifiez les autorisations de la clé privée sur le nœud principal.

```
$ chmod 600 ~/.ssh/id_rsa
chmod 600 ~/.ssh/config
```

5. Ouvrez `~/.ssh/id_rsa.pub` à l'aide de l'éditeur de texte de votre choix et copiez la clé.
6. Pour chaque nœud membre du cluster, procédez comme suit :
   a. Connectez-vous à l'instance.
   b. Ouvrez `~/.ssh/authorized_keys` à l'aide de l'éditeur de texte de votre choix et ajoutez la clé publique que vous avez copiée précédemment.
7. Pour tester si l'accès SSH sans mot de passe fonctionne comme prévu, connectez-vous à votre nœud principal et exécutez la commande suivante.

```
$ ssh member_node_private_ip
```

   Vous devez vous connecter au nœud membre sans être invité à entrer une clé ou un mot de passe.

Démarrer avec EFA et NCCL

La Nvidia Collective Communications Library (NCCL) est une bibliothèque de routines de communication collectives standard pour plusieurs GPU sur un nœud ou plusieurs nœuds. La NCCL peut être utilisée conjointement avec EFA, Libfabric et MPI pour prendre en charge différentes charges de travail de Machine Learning. Pour de plus amples informations, veuillez consulter le site web NCCL.

Note

- NCCL avec EFA est prise en charge avec les instances p3dn.24xlarge uniquement.
- Seule NCCL 2.4.2 et les versions ultérieures sont prises en charge avec EFA.

Les didacticiels suivants vous permettent de lancer un cluster d'instances EFA et NCCL pour les charges de travail de Machine Learning.

- Utilisation d'une AMI de base (p. 903)
- Utilisation d'une AMI AWS Deep Learning (p. 914)

Utilisation d'une AMI de base

Les étapes suivantes vous permettent de commencer à utiliser l'une des AMIs de base prises en charge (p. 893).

Sommaire

- Étape 1 : Préparer un groupe de sécurité activé pour les EFA (p. 903)
- Étape 2 : Lancer une instance temporaire (p. 904)
- Étape 3 : Installer le logiciel EFA (p. 904)
- Étape 4 : Installer les pilotes GPU Nvidia et le Nvidia CUDA Toolkit (p. 907)
- Étape 5 : Installer CuDNN (p. 908)
- Étape 6 : Installer NCCL (p. 909)
- Étape 7 : Installer le plug-in aws-ofi-nccl (p. 909)
- Étape 8 : Installer les tests NCCL (p. 910)
- Étape 9 : Tester votre configuration EFA et NCCL (p. 911)
- Étape 10 : Installer vos applications de Machine Learning (p. 912)
- Étape 11 : Créer une AMI activée pour EFA et NCCL (p. 912)
- Étape 12 : Résilier l'instance temporaire (p. 913)
- Étape 13 : Lancer les instances activées pour EFA et NCCL dans un groupe de placement de cluster (p. 913)
- Étape 14 : Activer SSH sans mot de passe (p. 914)

Étape 1 : Préparer un groupe de sécurité activé pour les EFA

L'EFA a besoin d'un groupe de sécurité qui autorise tout le trafic entrant et sortant vers et depuis le groupe de sécurité proprement dit. La procédure suivante autorise tout le trafic entrant et sortant à des fins de test uniquement. Pour d'autres scénarios, veuillez consulter Référence des règles de groupe de sécurité (p. 1075).

Pour créer un groupe de sécurité activé pour EFA

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité, puis Créer un groupe de sécurité.
3. Dans la fenêtre Créer un groupe de sécurité, procédez comme suit :
   a. Pour Nom du groupe de sécurité, entrez un nom descriptif pour le groupe de sécurité, tel que `EFA-enabled security group`.
   b. (Facultatif) Dans le champ Description, saisissez une brève description du groupe de sécurité.
   c. Pour VPC, sélectionnez le VPC dans lequel vous prévoyez de lancer vos instances activées pour EFA.
   d. Sélectionnez Créer.
4. Sélectionnez le groupe de sécurité que vous avez créé et dans l'onglet Description, copiez l'ID du groupe.
5. Dans l'onglet Entrant procédez comme suit :

   a. Choisissez Modifier.
   b. Pour Type, sélectionnez Tout le trafic.
   c. Dans Source, choisissez Personnaliser et collez l'ID du groupe de sécurité que vous avez copié dans le champ.
   d. Choisissez Enregistrer.
6. Dans l'onglet Sortant procédez comme suit :
   a. Choisissez Modifier.
   b. Pour Type, sélectionnez Tout le trafic.
   c. Pour Destination, choisissez Personnaliser et collez l'ID du groupe de sécurité que vous avez copié dans le champ.
   d. Choisissez Save.

Étape 2 : Lancer une instance temporaire

Lancez une instance temporaire que vous pouvez utiliser pour installer et configurer les composants logiciels EFA. Vous utilisez cette instance pour créer une AMI activée pour EFA depuis laquelle vous pouvez lancer vos instances activées pour EFA.

Pour lancer une instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Sur la page Choisir une AMI, choisissez l'une des AMIs prises en charge.
4. Sur la page Choisir un type d'instance, sélectionnez `p3dn.24xlarge`, puis choisissez Suivant : Configurer les détails d'instance.
5. Sur la page Configurer les détails de l'instance, procédez de la façon suivante :
   a. Pour Sous-réseau, choisissez le sous-réseau dans lequel lancer l'instance.
   b. Pour Elastic Fabric Adapter (EFA), choisissez Enable (Activer).
   c. Dans la section Interfaces réseau, pour l'appareil eth0, choisissez Nouvelle interface réseau.
   d. Choisissez Suivant : Ajouter un stockage.
6. Sur la page Ajouter un stockage, spécifiez les volumes à attacher aux instances, outre ceux spécifiés par l'AMI (par exemple, le volume du périphérique racine). Choisissez ensuite Suivant : Ajouter des balises.
7. Sur la page Ajouter des balises, spécifiez une balise que vous pouvez utiliser pour identifier l'instance temporaire, puis choisissez }Suivant : Configurer le groupe de sécurité.
8. Sur la page Configurer un groupe de sécurité, pour Attribuer un groupe de sécurité, choisissez Sélectionner un groupe de sécurité existant. Sélectionnez ensuite le groupe de sécurité que vous avez créé lors de l'étape 1.
9. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis sélectionnez Lancer pour choisir une paire de clés et lancer votre instance.

Étape 3 : Installer le logiciel EFA

Installez le noyau activé pour EFA, les pilotes EFA, Libfabric et la pile Open MPI qui est requise pour prendre en charge EFA sur votre instance temporaire.

Pour installer le logiciel EFA

1. Connectez-vous à l'instance que vous avez lancée. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Pour vous assurer que tous vos packages logiciels sont mis à jour, effectuez une mise à jour logicielle rapide sur votre instance. Ce processus peut prendre quelques minutes.
   - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

     ```
     $ sudo yum update -y
     ```

   - Ubuntu 16.04 et Ubuntu 18.04

     ```
     $ sudo apt-get update -y
     ```

     ```
     $ sudo apt-get upgrade -y
     ```

3. Téléchargez les fichiers d'installation du logiciel EFA. Les fichiers d'installation du logiciel sont packagés dans un fichier tarball compressé (`.tar.gz`). Pour télécharger la version stable la plus récente, utilisez la commande suivante.

   ```
   $ curl -O https://efa-installer.amazonaws.com/aws-efa-installer-1.10.0.tar.gz
   ```

   Vous pouvez aussi obtenir la dernière version en remplaçant le numéro de version par `latest` dans la commande ci-dessus.
4. (Facultatif) Vérifiez l'authenticité et l'intégrité du fichier tarball EFA (`.tar.gz`). Nous vous recommandons de le faire pour vérifier l'identité de l'éditeur du logiciel et pour vérifier que le fichier n'a pas été modifié ou endommagé depuis sa publication. Si vous ne souhaitez pas vérifier le fichier tarball, ignorez cette étape.

   Note

   Par ailleurs, si vous préférez vérifier le fichier tarball à l'aide d'un total de contrôle MD5 ou SHA256 à la place, veuillez consulter Vérification du programme d'installation EFA à l'aide d'un total de contrôle (p. 923).

   a. Téléchargez la clé GPG publique et importez-la dans votre porte-clés.

      ```
      $ wget https://efa-installer.amazonaws.com/aws-efa-installer.key
      ```

      ```
      $ gpg --import aws-efa-installer.key
      ```

      La commande doit renvoyer une valeur de clé. Notez la valeur de la clé, car vous en aurez besoin lors de l'étape suivante.

   b. Vérifiez l'empreinte digitale de la clé GPG. Exécutez la commande suivante et spécifiez la valeur de clé de l'étape précédente.

      ```
      $ gpg --fingerprint key_value
      ```

      La commande doit renvoyer une empreinte digitale identique à `4E90 91BC BB97 A96B 26B1 5E59 A054 80B1 DD2D 3CCC`. Si l'empreinte digitale ne correspond pas, n'exécutez pas le script d'installation EFA et contactez AWS Support.

   c. Téléchargez le fichier de signature et vérifiez la signature du fichier tarball EFA.

```
$ wget https://efa-installer.amazonaws.com/aws-efa-installer-1.10.0.tar.gz.sig
```

```
$ gpg --verify ./aws-efa-installer-1.10.0.tar.gz.sig
```

Voici un exemple de sortie.

```
gpg: Signature made Wed 29 Jul 2020 12:50:13 AM UTC using RSA key ID DD2D3CCC
gpg: Good signature from "Amazon EC2 EFA <ec2-efa-maintainers@amazon.com>"
gpg: WARNING: This key is not certified with a trusted signature!
gpg:          There is no indication that the signature belongs to the owner.
Primary key fingerprint: 4E90 91BC BB97 A96B 26B1  5E59 A054 80B1 DD2D 3CCC
```

Si le résultat comprend `Good signature` et l'empreinte correspond à celle renvoyée à l'étape précédente, passez à l'étape suivante. Si ce n'est pas le cas, n'exécutez pas le script d'installation EFA et contactez AWS Support.

5. Procédez à l'extraction des fichiers à partir du fichier compressé `.tar.gz` et accédez au répertoire extrait.

```
$ tar -xf aws-efa-installer-1.10.0.tar.gz
```

```
$ cd aws-efa-installer
```

6. Exécutez le script d'installation du logicielEFA.

```
$ sudo ./efa_installer.sh -y
```

Libfabric est installé dans le répertoire `/opt/amazon/efa`, tandis qu'Open MPI est installé dans le répertoire `/opt/amazon/openmpi`.

7. Déconnectez-vous de l'instance, puis reconnectez-vous.
8. Vérifiez que les composants logiciels EFA ont été installés avec succès.

```
$ fi_info -p efa
```

La commande doit renvoyer des informations sur les interfaces EFA Libfabric. L'exemple suivant illustre la sortie de la commande.

```
provider: efa
    fabric: EFA-fe80::94:3dff:fe89:1b70
    domain: efa_0-rdm
    version: 2.0
    type: FI_EP_RDM
    protocol: FI_PROTO_EFA
provider: efa
    fabric: EFA-fe80::94:3dff:fe89:1b70
    domain: efa_0-dgrm
    version: 2.0
    type: FI_EP_DGRAM
    protocol: FI_PROTO_EFA
provider: efa;ofi_rxd
    fabric: EFA-fe80::94:3dff:fe89:1b70
    domain: efa_0-dgrm
    version: 1.0
    type: FI_EP_RDM
```

```
    protocol: FI_PROTO_RXD
```

Étape 4 : Installer les pilotes GPU Nvidia et le Nvidia CUDA Toolkit

Pour installer les pilotes GPU Nvidia et le Nvidia CUDA Toolkit

1. Installez les utilitaires nécessaires pour l'installation des pilotes GPU Nvidia et du Nvidia CUDA toolkit.
   - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

     ```
     $ sudo yum groupinstall 'Development Tools' -y
     ```

   - Ubuntu 16.04 et Ubuntu 18.04

     ```
     $ sudo apt-get install build-essential -y
     ```

2. Pour utiliser le pilote GPU Nvidia, vous devez d'abord désactiver les pilotes open source `nouveau`.
   a. Installez le compilateur gcc et le package d'en-têtes de noyau correspondant à la version du noyau que vous utilisez actuellement.
      - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

        ```
        $ sudo yum install -y gcc kernel-devel-$(uname -r)
        ```

      - Ubuntu 16.04 et Ubuntu 18.04

        ```
        $ sudo apt-get install -y gcc make linux-headers-$(uname -r)
        ```

   b. Ajouter `nouveau` au `/etc/modprobe.d/blacklist.conf` fichier de liste de refus.

      ```
      $ cat << EOF | sudo tee --append /etc/modprobe.d/blacklist.conf
      blacklist vga16fb
      blacklist nouveau
      blacklist rivafb
      blacklist nvidiafb
      blacklist rivatv
      EOF
      ```

   c. Ouvrez le fichier `/etc/default/grub` à l'aide de l'éditeur de texte de votre choix et ajoutez ce qui suit.

      ```
      GRUB_CMDLINE_LINUX="rdblacklist=nouveau"
      ```

   d. Générez à nouveau la configuration Grub.
      - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

        ```
        $ sudo grub2-mkconfig -o /boot/grub2/grub.cfg
        ```

      - Ubuntu 16.04 et Ubuntu 18.04

        ```
        $ sudo update-grub
        ```

3. Redémarrez l'instance et reconnectez-vous à celle-ci.
4. Téléchargez le programme d'installation du Nvidia CUDA Toolkit.

```
$ wget https://developer.download.nvidia.com/compute/cuda/11.0.3/local_installers/
cuda_11.0.3_450.51.06_linux.run
```

5. Exécutez le programme d'installation du Nvidia CUDA Toolkit.

```
$ sudo sh cuda_11.0.3_450.51.06_linux.run
```

   Lorsque vous êtes invité à accepter l'accord de licence, entrez `accept` et appuyez sur Entrée.
6. Dans le menu du programme d'installation CUDA, assurez-vous que tous les éléments sont sélectionnés, choisissez Installer, puis appuyez sur Entrée.
7. Ajoutez les instructions suivantes aux scripts de démarrage de shell afin de vous assurer que les chemins CUDA sont définis à chaque démarrage de l'instance.

```
export PATH=/usr/local/cuda/bin:/usr/local/cuda/NsightCompute-2019.1${PATH:+:${PATH}}
export LD_LIBRARY_PATH=/usr/local/cuda/lib64\${LD_LIBRARY_PATH:+:${LD_LIBRARY_PATH}}
```

   - Pour les shells bash, ajoutez les instructions à `/home/`*`username`*`/.bashrc` et `/home/`*`username`*`/.bash_profile`.
   - Pour les shells tcsh, ajoutez les instructions à `/home/`*`username`*`/.cshrc`.
8. Pour vérifier que les pilotes GPU Nvidia sont fonctionnels, exécutez la commande suivante.

```
$ nvidia-smi -q | head
```

   La commande doit renvoyer des informations sur les GPU Nvidia, les pilotes GPU Nvidia et le Nvidia CUDA Toolkit.

Étape 5 : Installer CuDNN

Installez la bibliothèque CuDNN. La bibliothèque CuDNN est disponible sous la forme de deux packages, une bibliothèque de développeur et une bibliothèque d'exécution. Selon votre cas d'utilisation, vous devrez peut-être installer une seule bibliothèque, ou les deux.

Pour installer CuDNN

1. Accédez à votre répertoire de base.

```
$ cd $HOME
```

2. Installez les référentiels de Machine Learning Nvidia.

   - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

```
$ sudo yum install -y https://developer.download.nvidia.com/compute/machine-learning/
repos/rhel7/x86_64/nvidia-machine-learning-repo-rhel7-1.0.0-1.x86_64.rpm
```

   - Ubuntu 16.04 et Ubuntu 18.04

```
$ wget https://developer.download.nvidia.com/compute/machine-learning/repos/
ubuntu1804/x86_64/nvidia-machine-learning-repo-ubuntu1804_1.0.0-1_amd64.deb
```

```
$ sudo apt-key adv --fetch-keys http://developer.download.nvidia.com/compute/machine-
learning/repos/ubuntu1804/x86_64/7fa2af80.pub
```

```
$ sudo apt install ./nvidia-machine-learning-repo-ubuntu1804_1.0.0-1_amd64.deb
```

3. Installez les bibliothèques.

   - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

   ```
   $ sudo yum install -y libcudnn8 libcudnn8-devel
   ```

   - Ubuntu 16.04 et Ubuntu 18.04

   ```
   $ sudo apt update && sudo apt install libcudnn8 libcudnn8-dev -y
   ```

Étape 6 : Installer NCCL

Installez la NCCL. Pour de plus amples informations sur la NCCL, veuillez consultez le référentiel NCCL.

Pour installer la NCCL.

1. Accédez au répertoire `/opt`.

   ```
   $ cd /opt
   ```

2. Clonez le référentiel officiel de la NCCL dans l'instance et accédez au référentiel cloné local.

   ```
   $ sudo git clone https://github.com/NVIDIA/nccl.git
   ```

   ```
   $ cd nccl
   ```

3. Créez et installez la NCCL et spécifiez le répertoire d'installation CUDA.

   ```
   $ sudo make -j src.build CUDA_HOME=/usr/local/cuda
   ```

Étape 7 : Installer le plug-in aws-ofi-nccl

Le plug-in aws-ofi-nccl mappe les API de transport orientées connexion de la NCCL à l'interface fiable sans connexion de Libfabric. Cela vous permet d'utiliser Libfabric comme fournisseur de réseau tout en exécutant des applications basées sur la NCCL. Pour de plus amples informations sur le plug-in aws-ofi-nccl, veuillez consulter le référentiel aws-ofi-nccl.

Pour installer le plug-in aws-ofi-nccl

1. Accédez à votre répertoire de base.

   ```
   $ cd $HOME
   ```

2. (Ubuntu 16.04 and Ubuntu 18.04) Installez les utilitaires nécessaires pour installer le plug-in aws-ofi-nccl. Pour installer les utilitaires requis, exécutez la commande suivante.

   ```
   $ sudo apt-get install libtool autoconf -y
   ```

3. Clonez la branche `aws` du référentiel aws-ofi-nccl AWS officiel à l'instance et accédez au référentiel cloné local.

```
$ git clone https://github.com/aws/aws-ofi-nccl.git -b aws
```

```
$ cd aws-ofi-nccl
```

4. Pour générer le script `configure`, exécutez le script `autogen.sh`.

```
$ ./autogen.sh
```

5. Pour générer les fichiers make, exécutez le script `configure` et spécifiez les répertoires d'installation MPI, Libfabric, NCCL et CUDA.

```
$ ./configure --prefix=/opt/aws-ofi-nccl --with-mpi=/opt/amazon/openmpi \
--with-libfabric=/opt/amazon/efa --with-nccl=/opt/nccl/build \
--with-cuda=/usr/local/cuda
```

6. Ajoutez le répertoire Open MPI à la variable `PATH`.

```
$ export PATH=/opt/amazon/openmpi/bin/:$PATH
```

7. Installez le plug-in aws-ofi-nccl.

```
$ make
```

```
$ sudo make install
```

Étape 8 : Installer les tests NCCL

Installez les tests NCCL. Les tests NCCL vous permettent de vous assurer que NCCL a été installée correctement et qu'elle fonctionne normalement. Pour de plus amples informations sur les tests NCCL, veuillez consulter le référentiel nccl-tests.

Pour installer les tests NCCL

1. Accédez à votre répertoire de base.

```
$ cd $HOME
```

2. Clonez le référentiel officiel nccl-tests dans l'instance et accédez au référentiel cloné local.

```
$ git clone https://github.com/NVIDIA/nccl-tests.git
```

```
$ cd nccl-tests
```

3. Ajoutez le répertoire Libfabric à la variable `LD_LIBRARY_PATH`.

   - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

     ```
     $ export LD_LIBRARY_PATH=/opt/amazon/efa/lib64:$LD_LIBRARY_PATH
     ```

   - Ubuntu 16.04 et Ubuntu 18.04

     ```
     $ export LD_LIBRARY_PATH=/opt/amazon/efa/lib:$LD_LIBRARY_PATH
     ```

4. Installez les tests NCCL et spécifiez les répertoires d'installation MPI, NCCL et CUDA.

```
$ make MPI=1 MPI_HOME=/opt/amazon/openmpi NCCL_HOME=/opt/nccl/build CUDA_HOME=/usr/
local/cuda
```

Étape 9 : Tester votre configuration EFA et NCCL

Exécutez un test afin de vous assurer que votre instance temporaire est configurée correctement pour EFA et NCCL.

Pour tester votre configuration EFA et NCCL

1. Créez un fichier hôte qui spécifie les hôtes sur lesquels les tests doivent être exécutés. La commande suivante crée un fichier hôte nommé `my-hosts` qui inclut une référence à l'instance elle-même.

   IMDSv2

   ```
   [ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-
   aws-ec2-metadata-token-ttl-seconds: 21600"` \
   && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
   meta-data/local-ipv4 >> my-hosts
   ```

   IMDSv1

   ```
   [ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/local-ipv4 >> my-hosts
   ```

2. Exécutez le test et spécifiez le fichier hôte (`--hostfile`) et le nombre de GPU à utiliser (`-n`). La commande suivante exécute le test `all_reduce_perf` sur 8 GPU sur l'instance elle-même et spécifie les variables d'environnement suivantes.

   - `FI_PROVIDER="efa"`—spécifie le fournisseur d'interface fabric. Cette valeur doit être définie sur `"efa"`.
   - `FI_EFA_TX_MIN_CREDITS=64`—spécifie le nombre de crédits d'envois demandés par l'expéditeur au destinataire. `64` est la valeur recommandée pour les tâches NCCL qui utilisent EFA. La valeur doit uniquement être augmentée pour les transferts de messages dont le volume est supérieur à 256 Mo.
   - `FI_EFA_ENABLE_SHM_TRANSFER=0` - Permet au fournisseur SHM d'établir la communication entre tous les processus intra-nœud.
   - `NCCL_DEBUG=INFO`—permet des sorties de débogage détaillées. Vous pouvez également spécifier `VERSION` pour imprimer uniquement la version NCCL au début du test ou `WARN` pour recevoir uniquement des messages d'erreur.
   - `NCCL_ALGO=ring` - Utilise l'algorithme d'anneau pour les opérations collectives.

   Pour de plus amples informations sur les arguments de test NCCL, veuillez consulter le LISEZ-MOI sur les tests NCCL dans le référentiel nccl-tests officiel.

   - Amazon Linux, Amazon Linux 2, RHEL 7.6/7.7/7.8, CentOS 7

   ```
   $ /opt/amazon/openmpi/bin/mpirun \
       -x FI_PROVIDER="efa" \
       -x FI_EFA_TX_MIN_CREDITS=64 \
       -x FI_EFA_ENABLE_SHM_TRANSFER=0 \
       -x LD_LIBRARY_PATH=/opt/nccl/build/lib:/usr/local/cuda/lib64:/opt/amazon/efa/
   lib64:/opt/amazon/openmpi/lib64:/opt/aws-ofi-nccl/lib:$LD_LIBRARY_PATH \
       -x NCCL_DEBUG=INFO \
   ```

```
    -x NCCL_ALGO=ring \
    -x RDMAV_FORK_SAFE=1 \
    --hostfile my-hosts -n 8 -N 8 \
    --mca btl tcp,self --mca btl_tcp_if_exclude lo,docker0 --bind-to none \
    $HOME/nccl-tests/build/all_reduce_perf -b 8 -e 1G -f 2 -g 1 -c 1 -n 100
```

- Ubuntu 16.04 et Ubuntu 18.04

```
$ /opt/amazon/openmpi/bin/mpirun \
    -x FI_PROVIDER="efa" \
    -x FI_EFA_TX_MIN_CREDITS=64 \
    -x FI_EFA_ENABLE_SHM_TRANSFER=0 \
    -x LD_LIBRARY_PATH=/opt/nccl/build/lib:/usr/local/cuda/lib64:/opt/amazon/efa/
lib64:/opt/amazon/openmpi/lib64:/opt/aws-ofi-nccl/lib:$LD_LIBRARY_PATH \
    -x NCCL_DEBUG=INFO \
    -x NCCL_ALGO=ring \
    -x RDMAV_FORK_SAFE=1 \
    --hostfile my-hosts -n 8 -N 8 \
    --mca btl tcp,self --mca btl_tcp_if_exclude lo,docker0 --bind-to none \
    $HOME/nccl-tests/build/all_reduce_perf -b 8 -e 1G -f 2 -g 1 -c 1 -n 100
```

3. Vous pouvez confirmer que EFA est actif en tant que fournisseur sous-jacent pour NCCL lorsque le journal `NCCL_DEBUG` est imprimé.

```
$ ip-192-168-2-54:14:14 [0] NCCL INFO NET/OFI Forcing AWS OFI ndev 4
*ip-192-168-2-54:14:14 [0] NCCL INFO NET/OFI Selected Provider is efa*
```

Étape 10 : Installer vos applications de Machine Learning

Installez les applications de machine learning sur l'instance temporaire. La procédure d'installation varie selon l'application de machine learning spécifique. Pour plus d'informations sur l'installation du logiciel sur votre instance Linux, consultez Gestion de logiciels sur votre instance Linux.

> **Note**
>
> Vous pouvez avoir besoin de vous reporter à la documentation de votre application de machine learning pour obtenir des instructions d'installation.

Étape 11 : Créer une AMI activée pour EFA et NCCL

Une fois que vous avez installé les composants logiciels requis, vous devez créer une AMI que vous pouvez réutiliser pour lancer vos instances activées pour EFA.

Pour créer une AMI à partir de votre instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance temporaire que vous avez créée et choisissez Actions, Image, Create image (Créer une image).
4. Pour Create image (Créer une image), procédez comme suit :
   a. Pour Image name (Nom de l'image), entrez un nom descriptif pour l'AMI.
   b. (Facultatif) Dans le champ Image description (Description de l'image), saisissez une brève description de l'objet de l'AMI.
   c. Choisissez Create image (Créer une image).
5. Dans le volet de navigation, choisissez AMIs.

6. Recherchez l'AMI que vous avez créée dans la liste. Attendez que le statut passe de `pending` à `available` avant de poursuivre avec l'étape suivante.

Étape 12 : Résilier l'instance temporaire

À ce stade, vous n'avez plus besoin de l'instance temporaire que vous avez lancée. Vous pouvez résilier l'instance pour ne plus payer celle-ci.

Pour résilier l'instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance temporaire, puis choisissez Actions, Instance state (État de l'instance), Terminate instance (Mettre fin à l'instance).
4. Lorsque vous êtes invité à confirmer, choisissez Terminate (Mettre fin).

Étape 13 : Lancer les instances activées pour EFA et NCCL dans un groupe de placement de cluster

Lancez vos instances activées pour EFA et NCCL dans un groupe de placement de cluster à l'aide de l'AMI activée pour EFA et le groupe de sécurité activé pour EFA que vous avez créé précédemment.

Pour lancer vos instances activées pour EFA et NCCL dans un groupe de placement de cluster

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Dans la page Choose an AMI (Choisir une AMI), choisissez My AMIss (Mes AMI), recherchez l'AMI que vous avez créée précédemment, puis choisissez Select (Sélectionner).
4. Sur la page Choose an Instance Type (Choisir un type d'instance), sélectionnez p3dn.24xlarge et choisissez Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
5. Sur la page Configurer les détails de l'instance, procédez de la façon suivante :
   a. Pour Number of instances (Nombre d'instances), entrez le nombre d'instances activées pour EFA et NCCL que vous voulez lancer.
   b. Pour Réseau et Sous-réseau, sélectionnez le VPC et le sous-réseau dans lesquels lancer les instances.
   c. Pour le Groupe de placement, sélectionnez Ajoutez une instance au groupe de placement.
   d. Pour Placement group name (Nom du groupe de placement), sélectionnez Add to a new placement group (Ajoutez à un nouveau groupe de placement) et entrez un nom descriptif pour le groupe de placement. Ensuite, pour Placement group strategy (Stratégie de groupe de placement), sélectionnez cluster.
   e. Pour EFA, choisissez Enable (Activer).
   f. Dans la section Interfaces réseau, pour l'appareil eth0, choisissez Nouvelle interface réseau. Vous pouvez éventuellement entrer une adresse IPv4 principale et une ou plusieurs adresses IPv4 secondaires. Si vous lancez l'instance dans un sous-réseau auquel un bloc CIDR IPv6 est associé, vous pouvez éventuellement spécifier une adresse IPv6 principale et une ou plusieurs adresses IPv6 secondaires.
   g. Choisissez Suivant : Ajouter un stockage.
6. Sur la page Add Storage (Ajouter un stockage), spécifiez les volumes à attacher aux instances, outre ceux spécifiés par l'AMI (par exemple, le volume du périphérique racine). Choisissez ensuite Next: Add Tags (Suivant : Ajouter des balises).

7. Sur la page Ajouter des balises, spécifiez des balises pour l'instance, par exemple un nom évocateur, puis sélectionnez Suivant : Configurer le groupe de sécurité.
8. Sur la page Configure Security Group (Configurer un groupe de sécurité), cliquez sur Assign a security group (Attribuer un groupe de sécurité), choisissez Select an existing security group (Sélectionner un groupe de sécurité existant), puis le groupe de sécurité que vous avez créé précédemment.
9. Choisissez Vérifier et lancer.
10. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis choisissez Lancer pour sélectionner une paire de clés et lancer votre instance.

Étape 14 : Activer SSH sans mot de passe

Pour permettre à vos applications de s'exécuter sur toutes les instances de votre cluster, vous devez activer l'accès SSH sans mot de passe entre le nœud principal et les nœuds membres. Le nœud principal est l'instance à partir de laquelle vous exécutez vos applications. Les instances restantes du cluster sont les nœuds membres.

Pour activer l'accès SSH sans mot de passe entre les instances du cluster

1. Sélectionnez une instance du cluster en tant que nœud principal et connectez-vous y.
2. Désactivez `strictHostKeyChecking` et activez `ForwardAgent` sur le nœud principal. Ouvrez `~/.ssh/config` à l'aide de l'éditeur de texte de votre choix et ajoutez ce qui suit.

```
Host *
    ForwardAgent yes
Host *
    StrictHostKeyChecking no
```

3. Générez une paire de clés RSA.

```
$ ssh-keygen -t rsa -N "" -f ~/.ssh/id_rsa
```

   La paire de clés est créée dans le répertoire `$HOME/.ssh/`.
4. Modifiez les autorisations de la clé privée sur le nœud principal.

```
$ chmod 600 ~/.ssh/id_rsa
chmod 600 ~/.ssh/config
```

5. Ouvrez `~/.ssh/id_rsa.pub` à l'aide de l'éditeur de texte de votre choix et copiez la clé.
6. Pour chaque nœud membre du cluster, procédez comme suit :
   a. Connectez-vous à l'instance.
   b. Ouvrez `~/.ssh/authorized_keys` à l'aide de l'éditeur de texte de votre choix et ajoutez la clé publique que vous avez copiée précédemment.
7. Pour tester si l'accès SSH sans mot de passe fonctionne comme prévu, connectez-vous à votre nœud principal et exécutez la commande suivante.

```
$ ssh member_node_private_ip
```

   Vous devez vous connecter au nœud membre sans être invité à entrer une clé ou un mot de passe.

Utilisation d'une AMI AWS Deep Learning

Les étapes suivantes vous permettent de démarrer avec l'une des AMI AWS Deep Learning suivantes :

- Deep Learning AMI (Amazon Linux 2) Version 25.0 et versions ultérieures
- Deep Learning AMI (Amazon Linux) Version 25.0 et versions ultérieures
- Deep Learning AMI (Ubuntu 18.04) Version 25.0 et versions ultérieures
- Deep Learning AMI (Ubuntu 16.04) Version 25.0 et versions ultérieures

Pour de plus amples informations, veuillez consulter le Guide de l'utilisateur AWS Deep Learning AMI.

Sommaire

- Étape 1 : Préparer un groupe de sécurité activé pour les EFA (p. 915)
- Étape 2 : Lancer une instance temporaire (p. 916)
- Étape 3 : Tester votre configuration EFA et NCCL (p. 916)
- Étape 4 : Installer vos applications de Machine Learning (p. 917)
- Étape 5 : Créer une AMI activée pour EFA et NCCL (p. 917)
- Étape 6 : Résilier l'instance temporaire (p. 918)
- Étape 7 : Lancer les instances activées pour EFA et NCCL dans un groupe de placement de cluster (p. 918)
- Étape 8 : Activer SSH sans mot de passe (p. 919)

Étape 1 : Préparer un groupe de sécurité activé pour les EFA

L'EFA a besoin d'un groupe de sécurité qui autorise tout le trafic entrant et sortant vers et depuis le groupe de sécurité proprement dit. La procédure suivante autorise tout le trafic entrant et sortant à des fins de test uniquement. Pour d'autres scénarios, veuillez consulter Référence des règles de groupe de sécurité (p. 1075).

Pour créer un groupe de sécurité activé pour EFA

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité, puis Créer un groupe de sécurité.
3. Dans la fenêtre Créer un groupe de sécurité, procédez comme suit :
   a. Pour Nom du groupe de sécurité, entrez un nom descriptif pour le groupe de sécurité, tel que `EFA-enabled security group`.
   b. (Facultatif) Dans le champ Description, saisissez une brève description du groupe de sécurité.
   c. Pour VPC, sélectionnez le VPC dans lequel vous prévoyez de lancer vos instances activées pour EFA.
   d. Sélectionnez Créer.
4. Sélectionnez le groupe de sécurité que vous avez créé et dans l'onglet Description, copiez l'ID du groupe.
5. Dans l'onglet Entrant procédez comme suit :
   a. Choisissez Modifier.
   b. Pour Type, sélectionnez Tout le trafic.
   c. Dans Source, choisissez Personnaliser et collez l'ID du groupe de sécurité que vous avez copié dans le champ.
   d. Choisissez Enregistrer.
6. Dans l'onglet Sortant procédez comme suit :
   a. Choisissez Modifier.

b. Pour Type, sélectionnez Tout le trafic.
c. Pour Destination, choisissez Personnaliser et collez l'ID du groupe de sécurité que vous avez copié dans le champ.
d. Choisissez Save.

Étape 2 : Lancer une instance temporaire

Lancez une instance temporaire que vous pouvez utiliser pour installer et configurer les composants logiciels EFA. Vous utilisez cette instance pour créer une AMI activée pour EFA depuis laquelle vous pouvez lancer vos instances activées pour EFA.

Pour lancer une instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Sur la page Choisir une AMI, choisissez une AWS Deep Learning AMI Version 25.0 ou ultérieure prise en charge.
4. Sur la page Choisir un type d'instance, sélectionnez `p3dn.24xlarge`, puis choisissez Suivant : Configurer les détails d'instance.
5. Sur la page Configurer les détails de l'instance, procédez de la façon suivante :
   a. Pour Sous-réseau, choisissez le sous-réseau dans lequel lancer l'instance.
   b. Pour Elastic Fabric Adapter (EFA), choisissez Enable (Activer).
   c. Dans la section Interfaces réseau, pour l'appareil eth0, choisissez Nouvelle interface réseau.
   d. Choisissez Suivant : Ajouter un stockage.
6. Sur la page Ajouter un stockage, spécifiez les volumes à attacher aux instances, outre ceux spécifiés par l'AMI (par exemple, le volume du périphérique racine). Choisissez ensuite Suivant : Ajouter des balises.
7. Sur la page Ajouter des balises, spécifiez une balise que vous pouvez utiliser pour identifier l'instance temporaire, puis choisissez }Suivant : Configurer le groupe de sécurité.
8. Sur la page Configurer un groupe de sécurité, pour Attribuer un groupe de sécurité, choisissez Sélectionner un groupe de sécurité existant. Sélectionnez ensuite le groupe de sécurité que vous avez créé lors de l'étape 1.
9. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis sélectionnez Lancer pour choisir une paire de clés et lancer votre instance.

Étape 3 : Tester votre configuration EFA et NCCL

Exécutez un test afin de vous assurer que votre instance temporaire est configurée correctement pour EFA et NCCL.

Pour tester votre configuration EFA et NCCL

1. Créez un fichier hôte qui spécifie les hôtes sur lesquels les tests doivent être exécutés. La commande suivante crée un fichier hôte nommé `my-hosts` qui inclut une référence à l'instance elle-même.

   IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-
aws-ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
meta-data/local-ipv4 >> my-hosts
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/local-ipv4 >> my-hosts
```

2. Exécutez le test et spécifiez le fichier hôte (`--hostfile`) et le nombre de GPU à utiliser (`-n`). La commande suivante exécute le test `all_reduce_perf` sur 8 GPU sur l'instance elle-même et spécifie les variables d'environnement suivantes.

   - `FI_PROVIDER="efa"`—spécifie le fournisseur d'interface fabric. Cette valeur doit être définie sur `"efa"`.
   - `FI_EFA_TX_MIN_CREDITS=64`—spécifie le nombre de crédits d'envois demandés par l'expéditeur au destinataire. `64` est la valeur recommandée pour les tâches NCCL qui utilisent EFA. La valeur doit uniquement être augmentée pour les transferts de messages dont le volume est supérieur à 256 Mo.
   - `NCCL_DEBUG=INFO`—permet des sorties de débogage détaillées. Vous pouvez également spécifier `VERSION` pour imprimer uniquement la version NCCL au début du test ou `WARN` pour recevoir uniquement des messages d'erreur.
   - `NCCL_TREE_THRESHOLD=0`—désactive les algorithmes d'arborescence pour le test.

   Pour de plus amples informations sur les arguments de test NCCL, veuillez consulter le LISEZ-MOI sur les tests NCCL dans le référentiel nccl-tests officiel.

```
/opt/amazon/openmpi/bin/mpirun \
 -x FI_PROVIDER="efa" \
 -x FI_EFA_TX_MIN_CREDITS=64 \
 -x NCCL_DEBUG=INFO \
 -x NCCL_TREE_THRESHOLD=0 \
 --hostfile my-hosts -n 8 -N 8 \
 --mca btl tcp,self --mca btl_tcp_if_exclude lo,docker0 --bind-to none \
 $HOME/src/bin/efa-tests/efa-cuda-10.0/all_reduce_perf -b 8 -e 1G -f 2 -g 1 -c 1 -n 100
```

Étape 4 : Installer vos applications de Machine Learning

Installez les applications de machine learning sur l'instance temporaire. La procédure d'installation varie selon l'application de machine learning spécifique. Pour plus d'informations sur l'installation du logiciel sur votre instance Linux, consultez Gestion de logiciels sur votre instance Linux.

Note

Vous pouvez avoir besoin de vous reporter à la documentation de votre application de machine learning pour obtenir des instructions d'installation.

Étape 5 : Créer une AMI activée pour EFA et NCCL

Une fois que vous avez installé les composants logiciels requis, vous devez créer une AMI que vous pouvez réutiliser pour lancer vos instances activées pour EFA.

Pour créer une AMI à partir de votre instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance temporaire que vous avez créée et choisissez Actions, Image, Create image (Créer une image).
4. Pour Create image (Créer une image), procédez comme suit :

   a. Pour Image name (Nom de l'image), entrez un nom descriptif pour l'AMI.
   b. (Facultatif) Dans le champ Image description (Description de l'image), saisissez une brève description de l'objet de l'AMI.
   c. Choisissez Create image (Créer une image).
5. Dans le volet de navigation, choisissez AMIs.
6. Recherchez l'AMI que vous avez créée dans la liste. Attendez que le statut passe de `pending` à `available` avant de poursuivre avec l'étape suivante.

Étape 6 : Résilier l'instance temporaire

À ce stade, vous n'avez plus besoin de l'instance temporaire que vous avez lancée. Vous pouvez résilier l'instance pour ne plus payer celle-ci.

Pour résilier l'instance temporaire

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance temporaire, puis choisissez Actions, Instance state (État de l'instance), Terminate instance (Mettre fin à l'instance).
4. Lorsque vous êtes invité à confirmer, choisissez Terminate (Mettre fin).

Étape 7 : Lancer les instances activées pour EFA et NCCL dans un groupe de placement de cluster

Lancez vos instances activées pour EFA et NCCL dans un groupe de placement de cluster à l'aide de l'AMI activée pour EFA et le groupe de sécurité activé pour EFA que vous avez créé précédemment.

Pour lancer vos instances activées pour EFA et NCCL dans un groupe de placement de cluster

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Dans la page Choose an AMI (Choisir une AMI), choisissez My AMIss (Mes AMI), recherchez l'AMI que vous avez créée précédemment, puis choisissez Select (Sélectionner).
4. Sur la page Choose an Instance Type (Choisir un type d'instance), sélectionnez p3dn.24xlarge et choisissez Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
5. Sur la page Configurer les détails de l'instance, procédez de la façon suivante :
   a. Pour Number of instances (Nombre d'instances), entrez le nombre d'instances activées pour EFA et NCCL que vous voulez lancer.
   b. Pour Réseau et Sous-réseau, sélectionnez le VPC et le sous-réseau dans lesquels lancer les instances.
   c. Pour le Groupe de placement, sélectionnez Ajoutez une instance au groupe de placement.
   d. Pour Placement group name (Nom du groupe de placement), sélectionnez Add to a new placement group (Ajoutez à un nouveau groupe de placement) et entrez un nom descriptif pour le groupe de placement. Ensuite, pour Placement group strategy (Stratégie de groupe de placement), sélectionnez cluster.
   e. Pour EFA, choisissez Enable (Activer).
   f. Dans la section Interfaces réseau, pour l'appareil eth0, choisissez Nouvelle interface réseau. Vous pouvez éventuellement entrer une adresse IPv4 principale et une ou plusieurs adresses IPv4 secondaires. Si vous lancez l'instance dans un sous-réseau auquel un bloc CIDR IPv6 est

associé, vous pouvez éventuellement spécifier une adresse IPv6 principale et une ou plusieurs adresses IPv6 secondaires.

g. Choisissez Suivant : Ajouter un stockage.

6. Sur la page Add Storage (Ajouter un stockage), spécifiez les volumes à attacher aux instances, outre ceux spécifiés par l'AMI (par exemple, le volume du périphérique racine). Choisissez ensuite Next: Add Tags (Suivant : Ajouter des balises).
7. Sur la page Ajouter des balises, spécifiez des balises pour l'instance, par exemple un nom évocateur, puis sélectionnez Suivant : Configurer le groupe de sécurité.
8. Sur la page Configure Security Group (Configurer un groupe de sécurité), cliquez sur Assign a security group (Attribuer un groupe de sécurité), choisissez Select an existing security group (Sélectionner un groupe de sécurité existant), puis le groupe de sécurité que vous avez créé précédemment.
9. Choisissez Vérifier et lancer.
10. Sur la page Examiner le lancement de l'instance, vérifiez les paramètres, puis choisissez Lancer pour sélectionner une paire de clés et lancer votre instance.

Étape 8 : Activer SSH sans mot de passe

Pour permettre à vos applications de s'exécuter sur toutes les instances de votre cluster, vous devez activer l'accès SSH sans mot de passe entre le nœud principal et les nœuds membres. Le nœud principal est l'instance à partir de laquelle vous exécutez vos applications. Les instances restantes du cluster sont les nœuds membres.

Pour activer l'accès SSH sans mot de passe entre les instances du cluster

1. Sélectionnez une instance du cluster en tant que nœud principal et connectez-vous y.
2. Désactivez `strictHostKeyChecking` et activez `ForwardAgent` sur le nœud principal. Ouvrez `~/.ssh/config` à l'aide de l'éditeur de texte de votre choix et ajoutez ce qui suit.

```
Host *
    ForwardAgent yes
Host *
    StrictHostKeyChecking no
```

3. Générez une paire de clés RSA.

```
$ ssh-keygen -t rsa -N "" -f ~/.ssh/id_rsa
```

La paire de clés est créée dans le répertoire `$HOME/.ssh/`.

4. Modifiez les autorisations de la clé privée sur le nœud principal.

```
$ chmod 600 ~/.ssh/id_rsa
chmod 600 ~/.ssh/config
```

5. Ouvrez `~/.ssh/id_rsa.pub` à l'aide de l'éditeur de texte de votre choix et copiez la clé.
6. Pour chaque nœud membre du cluster, procédez comme suit :

   a. Connectez-vous à l'instance.

   b. Ouvrez `~/.ssh/authorized_keys` à l'aide de l'éditeur de texte de votre choix et ajoutez la clé publique que vous avez copiée précédemment.

7. Pour tester si l'accès SSH sans mot de passe fonctionne comme prévu, connectez-vous à votre nœud principal et exécutez la commande suivante.

```
$ ssh member_node_private_ip
```

Vous devez vous connecter au nœud membre sans être invité à entrer une clé ou un mot de passe.

Utilisation d'EFA

Vous pouvez créer, utiliser et gérer un EFA tout comme n'importe quelle interface réseau Elastic dans Amazon EC2. En revanche, contrairement aux interfaces réseau Elastic, les EFAs ne peuvent pas être attachés à une instance ou détachés de celle-ci à l'état d'exécution.

Exigences relatives à EFA

Pour utiliser EFA, vous devez procéder comme suit :

- Choisissez l'un des types d'instance pris en charge (p. 893).
- Utilisez l'une des AMIs prises en charge (p. 893).
- Installez les composants logiciels EFA. Pour plus d'informations, consultez Étape 3 : Installer le logiciel EFA (p. 895) et Étape 5 : (Facultatif) Installer Intel MPI (p. 899).
- Utilisez un groupe de sécurité qui autorise tout le trafic entrant et sortant vers et depuis le groupe de sécurité proprement dit. Pour plus d'informations, consultez Étape 1 : Préparer un groupe de sécurité activé pour les EFA (p. 894).

Table des matières

- Création d'un EFA (p. 920)
- Attachement d'une instance EFA à une instance arrêtée (p. 921)
- Attachement d'une instance EFA lors du lancement d'une instance (p. 921)
- Ajout d'une instance EFA à un modèle de lancement (p. 921)
- Gestion des adresses IP pour un EFA (p. 922)
- Modification du groupe de sécurité pour un EFA (p. 922)
- Détachement d'un EFA (p. 922)
- Affichage d'un EFAs (p. 922)
- Suppression d'un EFA (p. 922)

Création d'un EFA

Vous pouvez créer un EFA dans un sous-réseau au sein d'un VPC. Vous ne pouvez pas déplacer l'EFA vers un autre sous-réseau une fois qu'il a été créé et vous pouvez uniquement l'attacher à des instances arrêtées dans la même zone de disponibilité.

Pour créer un(e) EFA à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Network Interfaces.
3. Sélectionnez Créer une interface réseau.
4. Pour Description, saisissez un nom descriptif pour l'EFA.
5. Pour Sous-réseau, sélectionnez le sous-réseau dans lequel créer l'EFA.
6. Pour IP privée, saisissez l'adresse IPv4 privée principale. Si vous ne spécifiez pas d'adresse IPv4, nous sélectionnons une adresse IPv4 privée disponible dans le sous-réseau sélectionné.
7. (IPv6 uniquement) Si vous avez sélectionné un sous-réseau qui a un bloc d'adresse CIDR IPv6 associé, vous pouvez le cas échéant spécifier une adresse IPv6 dans le champ IP IPv6.

8. Pour Groupes de sécurité, sélectionnez un ou plusieurs groupes de sécurité.
9. Pour EFA, choisissez Activé.
10. Choisissez Yes, Create.

Pour créer un(e) EFA à l'aide de la AWS CLI

Utilisez la commande create-network-interface et pour `interface-type`, spécifiez `efa`, comme dans l'exemple suivant.

```
aws ec2 create-network-interface --subnet-id subnet-01234567890 --description example_efa
 --interface-type efa
```

Attachement d'une instance EFA à une instance arrêtée

Vous pouvez attacher un EFA à toute instance prise en charge à l'état `stopped`. Vous ne pouvez pas attacher un EFA à une instance à l'état `running`. Pour plus d'informations sur les types d'instance pris en charge, consultez Types d'instance pris en charge (p. 893).

Vous attachez un EFA à une instance de la même manière que vous attachez une interface réseau à une instance. Pour de plus amples informations, veuillez consulter Attachement d'une interface réseau à une instance (p. 855).

Attachement d'une instance EFA lors du lancement d'une instance

Pour attacher un EFA existant lors du lancement d'une instance (AWS CLI)

Utilisez la commande run-instances et pour NetworkInterfaceId, spécifiez l'ID de l'EFA, comme dans l'exemple suivant.

```
aws ec2 run-instances --image-id ami_id --count 1 --instance-
type c5n.18xlarge --key-name my_key_pair --network-interfaces
 DeviceIndex=0,NetworkInterfaceId=efa_id,Groups=sg_id,SubnetId=subnet_id
```

Pour attacher un nouvel EFA lors du lancement d'une instance (AWS CLI)

Utilisez la commande run-instances et pour InterfaceType, spécifiez `efa`, comme dans l'exemple suivant.

```
aws ec2 run-instances --image-id ami_id --count 1 --instance-
type c5n.18xlarge --key-name my_key_pair --network-interfaces
 DeviceIndex=0,InterfaceType=efa,Groups=sg_id,SubnetId=subnet_id
```

Ajout d'une instance EFA à un modèle de lancement

Vous pouvez créer un modèle de lancement contenant les informations de configuration nécessaires pour lancer des instances activées pour EFA. Pour créer un modèle de lancement activé pour EFA, créez un nouveau modèle de lancement, et spécifiez un type d'instance pris en charge, votre AMI activée pour EFA et un groupe de sécurité activé pour EFA. Pour plus d'informations, consultez Démarrer avec EFA et MPI (p. 894).

Vous pouvez tirer parti des modèles de lancement pour lancer des instances activées pour EFA avec d'autres services AWS comme AWS Batch.

Pour plus d'informations sur la création de modèles de lancement, consultez Création d'un modèle de lancement (p. 539).

Gestion des adresses IP pour un EFA

Vous pouvez modifier les adresses IP associées à un EFA. Si vous avez une adresse IP Elastic, vous pouvez l'associer à un EFA. Si votre EFA est mis en service dans un sous-réseau auquel un bloc d'adresse CIDR IPv6 est associé, vous pouvez attribuer une ou plusieurs adresses IPv6 à l'EFA.

Vous attribuez des adresses IP Elastic (IPv4) et IPv6 à un EFA tout comme vous attribuez une adresse IP à une interface réseau Elastic. Pour de plus amples informations, veuillez consulter Gestion des adresses IP (p. 856).

Modification du groupe de sécurité pour un EFA

Vous pouvez modifier le groupe de sécurité associé à un EFA. Pour que vous puissiez activer la fonctionnalité de contournement du système d'exploitation, l'EFA doit appartenir à un groupe de sécurité qui autorise tout le trafic entrant et sortant vers et depuis le groupe de sécurité proprement dit.

Vous pouvez modifier le groupe de sécurité associé à un EFA comme vous le feriez pour un groupe de sécurité associé à une interface réseau Elastic. Pour de plus amples informations, veuillez consulter Modification du groupe de sécurité (p. 858).

Détachement d'un EFA

Pour détacher un EFA d'une instance, vous devez d'abord arrêter l'instance. Vous ne pouvez pas détacher un EFA d'une instance à l'état d'exécution.

Vous détachez un EFA d'une instance tout comme vous détachez une interface réseau Elastic d'une instance. Pour plus d'informations, consultez Détachement d'une interface réseau d'une instance (p. 855).

Affichage d'un EFAs

Vous pouvez afficher tous les EFAs de votre compte.

Vous affichez les EFAs comme vous le feriez pour les interfaces réseau Elastic. Pour plus d'informations, consultez Affichage des détails relatifs à une interface réseau (p. 854).

Suppression d'un EFA

Pour supprimer un EFA, vous devez d'abord le détacher de l'instance. Vous ne pouvez pas supprimer un EFA pendant qu'il est attaché à une instance.

Vous supprimez les EFAs comme vous le feriez pour les interfaces réseau Elastic. Pour plus d'informations, consultez Suppression d'une interface réseau (p. 859).

Surveillance d'un EFA

Vous pouvez utiliser les fonctions suivantes pour surveiller les performances de vos Elastic Fabric Adapters.

Journaux de flux Amazon VPC

Vous pouvez créer un journal de flux Amazon VPC pour capturer des informations sur le trafic entrant ou sortant de votre EFA. Les données des journaux de flux peuvent être publiées dans Amazon CloudWatch Logs et Amazon S3. Une fois que vous avez créé un journal de flux, vous pouvez extraire et afficher ses

données dans la destination choisie. Pour plus d'informations, consultez Journaux de flux VPC dans le Amazon VPC Guide de l'utilisateur.

Vous créez un journal de flux pour un EFA comme vous le feriez pour une interface réseau Elastic. Pour plus d'informations, consultez Création d'un journal de flux dans le Amazon VPC Guide de l'utilisateur.

Dans les entrées de journal de flux, le trafic EFA est identifié par des adresses `srcAddress` et `destAddress`, qui sont formatées comme des adresses MAC, comme dans l'exemple suivant.

```
version accountId  eniId        srcAddress        destAddress       sourcePort destPort
 protocol packets bytes start      end        action log-status
2       3794735123 eni-10000001 01:23:45:67:89:ab 05:23:45:67:89:ab -          -        -
      9       5689  1521232534 1524512343 ACCEPT OK
```

Amazon CloudWatch

Amazon CloudWatch propose des métriques qui vous permettent de surveiller vos EFAs 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. Pour plus d'informations, consultez Surveillance de vos instances avec CloudWatch (p. 761).

Vérification du programme d'installation EFA à l'aide d'un total de contrôle

Vous pouvez éventuellement vérifier l'archive EFA (fichier .tar.gz) à l'aide d'un total de contrôle MD5 ou SHA256. Nous vous recommandons de le faire pour vérifier l'identité de l'éditeur du logiciel et pour vérifier que l'application n'a pas été modifiée ou endommagée depuis sa publication.

Pour vérifier l'archive

Utilisez l'utilitaire md5sum pour le total de contrôle MD5 ou l'utilitaire sha256sum pour le total de contrôle SHA256 et spécifiez le nom du fichier tarball. Vous devez exécuter la commande à partir du répertoire dans lequel vous avez enregistré le fichier tarball.

- MD5

```
$  md5sum tarball_filename.tar.gz
```

- SHA256

```
$  sha256sum tarball_filename.tar.gz
```

Les commandes doivent renvoyer une valeur du total de contrôle au format suivant.

```
checksum_value tarball_filename.tar.gz
```

Comparez la valeur du total de contrôle renvoyée par la commande avec la valeur du total de contrôle fournie dans le tableau ci-dessous. Si les totaux de contrôle correspondent, on peut alors exécuter le script d'installation en toute sécurité. Si les totaux de contrôle ne correspondent pas, n'exécutez pas le script d'installation et contactez AWS Support.

Par exemple, la commande suivante vérifie l'archive EFA 1.9.4 à l'aide du total de contrôle SHA256.

```
$  sha256sum aws-efa-installer-1.9.4.tar.gz
```

```
1009b5182693490d908ef0ed2c1dd4f813cc310a5d2062ce9619c4c12b5a7f14 aws-efa-
installer-1.9.4.tar.gz
```

Le tableau suivant répertorie les totaux de contrôle des versions récentes de EFA.

Version	Télécharger le kit URL	Totaux de contrôle
EFA 1.10.0	https://efa-installer.amazonaws.com/aws-efa-installer-1.10.0.tar.gz	MD5 : `46f73f5a7afe41b4bb918c81888fefa9` SHA256 : `136612f96f2a085a7d98296da0afb6fa807b38142`
EFA 1.9.5	https://efa-installer.amazonaws.com/aws-efa-installer-1.9.5.tar.gz	MD5 : `95edb8a209c18ba8d250409846eb6ef4` SHA256 : `a4343308d7ea4dc943ccc21bcebed913e8868e59b`
EFA 1.9.4	https://efa-installer.amazonaws.com/aws-efa-installer-1.9.4.tar.gz	MD5 : `f26dd5c350422c1a985e35947fa5aa28` SHA256 : `1009b5182693490d908ef0ed2c1dd4f813cc310a5`
EFA 1.9.3	https://efa-installer.amazonaws.com/aws-efa-installer-1.9.3.tar.gz	MD5 : `95755765a097802d3e6d5018d1a5d3d6` SHA256 : `46ce732d6f3fcc9edf6a6e9f9df0ad136054328e2`
EFA 1.8.4	https://efa-installer.amazonaws.com/aws-efa-installer-1.8.4.tar.gz	MD5 : `85d594c41e831afc6c9305263140457e` SHA256 : `0d974655a09b213d7859e658965e56dc4f23a0eee`

Groupes de placement

Lorsque vous lancez une nouvelle instance EC2, le service EC2 tente de placer l'instance de façon à ce que toutes vos instances soient réparties sur la matériel sous-jacent pour minimiser les échecs corrélés. Vous pouvez utiliser des groupes de placement pour influencer le placement d'un groupe d'instances interdépendantes afin de répondre aux besoins de votre charge de travail. Selon le type de charge de travail, vous pouvez créer un groupe de placement à l'aide de l'une des stratégies de placement suivantes :

- Cluster – regroupe des instances rapprochées à l'intérieur d'une Zone de disponibilité. Cette stratégies permet aux charges de travail d'atteindre les performances réseau à faible latence nécessaires à une communication de nœud à nœud étroitement couplée, typique des applications HPC.
- Partition – répartit les instances entre les partitions logiques de façon à ce que des groupes d'instances d'une partition ne partagent pas le matériel sous-jacent avec des groupes d'instances d'autres partitions. Cette stratégie est généralement utilisée par les grandes charges de travail distribuées et répliquées telles que Hadoop, Cassandra, et Kafka.

- Répartition – place strictement un petit groupe d'instances sur un matériel sous-jacent distinct pour réduire les défaillances corrélées.

Il n'y a aucuns frais pour la création d'un groupe de placement.

Sommaire

- Groupes de placement du cluster (p. 925)
- Groupes de placement par partition (p. 926)
- Groupes de placement par répartition (p. 927)
- Règles et restrictions des groupes de placement (p. 927)
- Création d'un groupe de placement (p. 929)
- Balisage d'un groupe de placement (p. 930)
- Lancement d'instances dans un groupe de placement (p. 932)
- Description des instances dans un groupe de placement (p. 933)
- Modification du groupe de placement d'une instance (p. 935)
- Suppression d'un groupe de placement (p. 936)

Groupes de placement du cluster

Un groupe de placement du cluster est un regroupement logique d'instances dans une même zone de disponibilité. Un groupe de placement de cluster peut s'étendre sur plusieurs VPC appairés dans la même région. Les instances du même groupe de placement de cluster bénéficient d'une limite de débit par flux supérieure pour le trafic TCP/IP et sont placées dans le même segment de bande passante haute bissection du réseau.

L'image ci-après illustre les instances placées dans un groupe de placement du cluster.

Les groupes de placement de cluster sont recommandés pour les applications qui bénéficient d'une latence réseau faible, d'un débit réseau élevé, ou des deux. Ils sont également recommandés lorsque la majorité du trafic réseau est échangé entre les instances du groupe. Pour assurer la plus faible latence et les meilleures performances réseau de paquets par seconde pour votre groupe de placement, choisissez un type d'instance qui prend en charge la mise en réseau améliorée. Pour plus d'informations, consultez Gestion de réseau améliorée (p. 864).

Nous vous recommandons de lancer vos instances de la façon suivante :

- Utilisez une seule demande de lancement pour lancer le nombre d'instances dont vous avez besoin dans le groupe de placement.
- Utilisez le même type d'instance pour toutes les instances du groupe de placement.

Si vous essayez d'ajouter ultérieurement des instances supplémentaires au groupe de placement, ou si vous essayez de lancer plusieurs types d'instance dans le groupe de placement, vous augmentez les risques d'obtenir une erreur de capacité insuffisante.

Si vous arrêtez une instance dans un groupe de placement, puis que vous la relancez, elle s'exécute encore au sein de celui-ci. Par contre, le démarrage échoue si la capacité est insuffisante pour l'instance.

Si vous recevez une erreur de capacité lorsque vous lancez une instance dans un groupe de placement dont des instances sont déjà en cours d'exécution, arrêtez et démarrez toutes les instances dans le groupe de placement, puis réessayez le lancement. Le redémarrage des instances peut entraîner leur migration vers un matériel qui dispose d'une capacité suffisante pour toutes les instances demandées.

Groupes de placement par partition

Les groupes de placement de partitions permettent de réduire la probabilité de défaillances de matériel corrélé pour votre application. Lorsque vous utilisez des groupes de placement de partitions, Amazon EC2 divise chaque groupe en segments logiques, appelés partitions. Amazon EC2 assure que chaque partition dans un groupe de placement dispose de son propre ensemble de racks, Chaque rack est doté de son propre réseau et de sa propre alimentation. Aucune partition dans un même groupe de placement ne dispose du même portant, ce qui vous permet ainsi d'isoler l'impact d'échecs matériels dans votre application.

L'image suivante est une représentation visuelle simplifiée d'un groupe de placement de partitions dans une seule Zone de disponibilité. Elle représente des instances placées dans un groupe de placement par partition composé de trois partitions—Partition 1, Partition 2 et Partition 3. Chaque partition comprend plusieurs instances. Les instances d'une partition ne partagent pas de portants avec les instances des autres partitions, ce qui vous permet de limiter l'impact des pannes matérielles à une seule partition.

Il est possible d'utiliser les groupes de placement par partition afin de déployer des charges de travail distribuées et répliquées volumineuses (telles que HDFS, HBase et Cassandra) sur différents portants. Lorsque vous lancez des instances dans un groupe de placement par partition, Amazon EC2 tente de distribuer uniformément les instances sur toutes les partitions que vous spécifiez. Vous avez également la possibilité de lancer des instances d'une partition donnée afin de mieux contrôler l'emplacement des instances.

Un groupe de placement par partition peut disposer de partitions dans plusieurs Zones de disponibilité de la même région. Un groupe de placement par partition peut contenir jusqu'à sept partitions par zone de disponibilité. Seules les restrictions de votre compte limitent le nombre d'instances pouvant être lancées dans un groupe de placement par partition.

De plus, les groupes de placement par partition vous permettent de voir le détail des partitions — types d'instance présents dans telle ou telle partition. Vous pouvez partager ces informations avec des applications prenant en compte la topologie, telles que HDFS, HBase et Cassandra. Ces applications utilisent ces informations pour prendre des décisions informées sur la réplication des données dans le but d'accroître la disponibilité et la durabilité de ces dernières.

Si vous exécutez ou lancez une instance dans un groupe de placement par partition et que le matériel nécessaire au traitement de la demande est insuffisant, la demande échoue. Amazon EC2 met à disposition davantage de matériel distinct au fil du temps, de sorte que vous pouvez renouveler votre demande plus tard.

Groupes de placement par répartition

Un groupe de placement par répartition est un groupe d'instances placées chacune sur des portants différents, chacun de ces derniers doté de son propre réseau et de sa propre alimentation.

L'image ci-après représente sept instances au sein d'une seule zone de disponibilité qui sont placées dans un groupe de placement par répartition. Les sept instances sont placées sur sept portants différents.

Les groupes de placement par répartition sont recommandés pour les applications ayant un petit nombre d'instances critiques, qui doivent être séparées les unes des autres. Le lancement d'instances dans un groupe de placement par répartition réduit le risque de défaillances simultanées, qui peuvent se produire lorsque les instances partagent les mêmes portants. Les groupes de placement par répartition fournissent un accès à des portants différents et sont par conséquent adaptés à l'association de différents types d'instance et au lancement d'instances au fil du temps.

Un groupe de placement par répartition peut également s'étendre sur plusieurs Zones de disponibilité dans la même région. Vous pouvez disposer de jusqu'à sept instances en cours d'exécution par Zone de disponibilité et par groupe.

Si vous exécutez ou lancez une instance dans un groupe de placement par répartition et que le matériel nécessaire à la satisfaction de la requête est insuffisant, la requête échoue. Amazon EC2 met à disposition davantage de matériel distinct au fil du temps, vous pouvez donc renouveler votre demande plus tard.

Règles et restrictions des groupes de placement

Règles et restrictions générales

Avant d'utiliser des groupes de placement, vous devez être conscient des règles suivantes :

- Le nom que vous spécifiez pour un groupe de placement doit être unique au sein de votre compte AWS pour la région.
- Vous ne pouvez pas fusionner des groupes de placement.
- Une instance peut être lancée dans un seul groupe de placement à la fois ; elle ne peut pas s'étendre sur plusieurs groupes de placement.
- Réservation de capacité à la demande (p. 506) et les Instances réservées zonales (p. 332) fournissent une réservation de capacité pour les instances EC2 dans une zone de disponibilité spécifique. La réservation de capacité peut être utilisée par les instances d'un groupe de placement. Toutefois, il n'est pas possible de réserver explicitement de la capacité pour un groupe de placement.

- Vous ne pouvez pas lancer Hôtes dédiés dans les groupes de placement.

Règles et restrictions des groupes de placement du cluster

Les règles suivantes s'appliquent aux groupes de placement du cluster :

- Pour les instances d'un groupe de placement de cluster, vous devez utiliser les types d'instance pris en charge suivants :
  - Les instances de la génération actuelle (p. 214), à l'exception des instances à capacité extensible (p. 233) (par exemple, T2).
  - Les instances de la génération précédente (p. 217) suivantes : C3, `cc2.8xlarge`, G2 `cr1.8xlarge`, G2, `hs1.8xlarge`, I2 et R3.
- Un groupe de placement du cluster ne peut pas s'étendre sur plusieurs zones de disponibilité.
- La vitesse de débit réseau maximale du trafic entre deux instances au sein d'un groupe de placement du cluster est limitée par la plus lente des deux instances. Pour les applications très exigeantes en débit, choisissez un type d'instance avec une connectivité réseau qui correspond à vos besoins.
- Pour les instances pour lesquelles la mise en réseau améliorée est active, les règles suivantes s'appliquent :
  - Les instances se trouvant dans un groupe de placement du cluster peuvent utiliser jusqu'à 10 Gbit/s pour le trafic à flux unique. Les instances qui ne se trouvent pas dans un groupe de placement du cluster peuvent utiliser jusqu'à 5 Gbit/s pour le trafic à flux unique.
  - Le trafic vers et depuis des compartiments Amazon S3 de la même région via l'espace d'adressage IP public ou un point de terminaison d'un VPC peut utiliser la totalité de la bande passante cumulée disponible pour l'instance.
- Vous pouvez lancer plusieurs types d'instance dans un groupe de placement du cluster. Toutefois, cela réduit la probabilité de disponibilité de la capacité requise pour que votre lancement réussisse. Nous vous recommandons d'utiliser le même type d'instance pour toutes les instances d'un groupe de placement du cluster.
- Le trafic réseau vers Internet et via une connexion AWS Direct Connect à destination de ressources sur site est limité à 5 Gbits/s.

Règles et restrictions des groupes de placement par partition

Les règles suivantes s'appliquent aux groupes de placement par partition :

- Un groupe de placement par partition prend en charge jusqu'à sept partitions par zone de disponibilité. Seules les restrictions de votre compte limitent le nombre d'instances pouvant être lancées dans un groupe de placement par partition.
- Lorsque vous lancez des instances dans un groupe de placement par partition, Amazon EC2 tente de distribuer uniformément les instances sur toutes les partitions. Toutefois, Amazon EC2 ne garantit pas une distribution uniforme des instances entre toutes les partitions.
- Un groupe de placement par partition avec des Instances dédiées peut comprendre deux partitions au maximum.

Règles et restrictions des groupes de placement par répartition

Les règles suivantes s'appliquent aux groupes de placement par répartition :

- Un groupe de placement par répartition, prend en charge un maximum de 7 instances en cours d'exécution pour une même zone de disponibilité. Par exemple, dans une région possédant trois zones de disponibilité, vous pouvez exécuter un total de 21 instances dans le groupe (sept par zone). Si vous

essayez de lancer une huitième instance dans la même zone de disponibilité et dans le même groupe de placement par répartition, le lancement échoue. Si vous avez besoin de plus de sept instances dans une zone de disponibilité, nous vous recommandons d'utiliser plusieurs groupes de placement par répartition. L'utilisation de plusieurs groupes de placement par répartition ne garantit pas la répartition des instances entre les groupes, mais cela permet de limiter l'impact de certains types d'incidents pour la répartition dans chaque groupe.

- Les groupes de placement par répartition ne sont pas pris en charge pour les Instances dédiées.

Création d'un groupe de placement

Vous pouvez créer un groupe de placement en employant l'une des méthodes suivantes.

> **Note**
>
> Vous pouvez marquer un groupe de placement lors de la création à l'aide des outils de ligne de commande uniquement.

Nouvelle console

Pour créer un groupe de placement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Groupes de placement, puis Créer un groupe de placement.
3. Spécifiez le nom du groupe.
4. Choisissez la stratégie de placement du groupe. Si vous choisissez Partition, choisissez le nombre de partitions au sein du groupe.
5. Choisissez Create group.

Ancienne console

Pour créer un groupe de placement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Groupes de placement, puis Créer un groupe de placement.
3. Spécifiez le nom du groupe.
4. Choisissez la stratégie de placement du groupe. Si vous choisissez Partition, spécifiez le nombre de partitions au sein du groupe.
5. Sélectionnez Create.

AWS CLI

Pour créer un groupe de placement à l'aide de l'AWS CLI

Utilisez la commande create-placement-group. L'exemple suivant crée un groupe de placement nommé `my-cluster` qui utilise la stratégie de placement `cluster` et applique une balise avec une clé `purpose` et une valeur `production`.

```
aws ec2 create-placement-group --group-name my-cluster --strategy cluster --tag-
specifications 'ResourceType=placement-group,Tags={Key=purpose,Value=production}'
```

Pour créer un groupe de placement par partition à l'aide de l'AWS CLI

Utilisez la commande create-placement-group. Spécifiez le paramètre `--strategy` avec la valeur `partition` et le paramètre `--partition-count` avec le nombre de partitions souhaité. Dans cet exemple, le groupe de placement par partition est nommé `HDFS-Group-A` et créé avec cinq partitions.

```
aws ec2 create-placement-group --group-name HDFS-Group-A --strategy partition --
partition-count 5
```

PowerShell

Pour créer un groupe de placement à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande New-EC2PlacementGroup.

Balisage d'un groupe de placement

Pour vous aider à classer et à gérer vos groupes de placement existants, vous pouvez les baliser avec des métadonnées personnalisées. Pour plus d'informations sur le fonctionnement des balises, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Lorsque vous balisez un groupe de placement, les instances lancées dans le groupe de placement ne sont pas automatiquement balisées. Vous devez baliser explicitement les instances lancées dans le groupe de placement. Pour plus d'informations, consultez Ajout d'une balise lorsque vous lancez une instance (p. 1317).

Vous pouvez afficher, ajouter et supprimer des balises à l'aide de la nouvelle console et des outils de ligne de commande.

Nouvelle console

Pour afficher, ajouter ou supprimer une balise pour un groupe de placement existant

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Groupes de placement.
3. Sélectionnez un groupe de placement, puis choisissez Actions, Gérer les balises.
4. La section Gérer les balises affiche toutes les balises affectées au groupe de placement. Pour ajouter ou supprimer des balises, procédez comme suit :
   - Pour ajouter une balise, choisissez Ajouter la balise, puis entrez la clé et la valeur de la balise. Vous pouvez ajouter jusqu'à 50 balises par groupe de placement. Pour de plus amples informations, veuillez consulter Restrictions liées aux balises (p. 1313).
   - Pour supprimer une balise, choisissez Supprimer en regard de la balise à supprimer.
5. Sélectionnez Save Changes.

AWS CLI

Pour afficher les balises des groupe de placement

Utilisez la commande describe-tags pour afficher les balises de la ressource spécifiée. Dans l'exemple suivant, vous décrivez les balises de tous vos groupes de placement.

```
aws ec2 describe-tags \
    --filters Name=resource-type,Values=placement-group
```

```
{
```

```
    "Tags": [
        {
            "Key": "Environment",
            "ResourceId": "pg-0123456789EXAMPLE",
            "ResourceType": "placement-group",
            "Value": "Production"
        },
        {
            "Key": "Environment",
            "ResourceId": "pg-9876543210EXAMPLE",
            "ResourceType": "placement-group",
            "Value": "Production"
        }
    ]
}
```

Vous pouvez également utiliser la commande describe-tags pour afficher les balises d'un groupe de placement en spécifiant son ID. Dans l'exemple suivant, vous décrivez les balises pour `pg-0123456789EXAMPLE`.

```
aws ec2 describe-tags \
    --filters Name=resource-id,Values=pg-0123456789EXAMPLE
```

```
{
    "Tags": [
        {
            "Key": "Environment",
            "ResourceId": "pg-0123456789EXAMPLE",
            "ResourceType": "placement-group",
            "Value": "Production"
        }
    ]
}
```

Vous pouvez également afficher les balises d'un groupe de placement en décrivant le groupe de placement.

Utilisez la commande describe-placement-groups pour afficher la configuration du groupe de placement spécifié, qui inclut toutes les balises définies pour le groupe de placement.

```
aws ec2 describe-placement-groups \
    --group-name my-cluster
```

```
{
    "PlacementGroups": [
        {
            "GroupName": "my-cluster",
            "State": "available",
            "Strategy": "cluster",
            "GroupId": "pg-0123456789EXAMPLE",
            "Tags": [
                {
                    "Key": "Environment",
                    "Value": "Production"
                }
            ]
        }
    ]
}
```

Pour baliser un groupe de placement existant à l'aide de l'AWS CLI

Utilisez la commande create-tags pour baliser les ressources existantes. Dans l'exemple suivant, le groupe de placement existant est balisé avec Key=Cost-Center et Value=CC-123.

```
aws ec2 create-tags \
    --resources pg-0123456789EXAMPLE \
    --tags Key=Cost-Center,Value=CC-123
```

Pour supprimer une balise d'un groupe de placement à l'aide de l'AWS CLI

Vous pouvez utiliser la commande delete-tags pour supprimer des balises de ressources existantes. Pour obtenir des exemples, reportez-vous à la section Exemples dans le document AWS CLI Command Reference.

PowerShell

Pour afficher les balises des groupe de placement

Utilisez la commande Get-EC2Tag.

Pour décrire les balises d'un groupe de placement spécifique

Utilisez la commande Get-EC2PlacementGroup.

Pour baliser un groupe de placement existant

Utilisez la commande New-EC2Tag.

Pour supprimer une balise d'un groupe de placement

Utilisez la commande Remove-EC2Tag.

Lancement d'instances dans un groupe de placement

Vous pouvez lancer une instance dans un groupe de placement si les règles et les limitations de groupe de placement sont respectées (p. 927) en employant l'une des méthodes suivantes.

Console

Pour lancer des instances dans un groupe de placement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Choisissez Launch Instances. Exécutez l'assistant selon les instructions en veillant à procéder comme suit :
   - Sur la page Choisir un type d'instance, sélectionnez un type d'instance pouvant être lancé dans un groupe de placement.
   - Sur la page Configurer les détails de l'instance, les champs suivants s'appliquent aux groupes de placement :
     - Pour le Number of instances (Nombre d'instances), saisissez le nombre total d'instances dont vous aurez besoin dans le groupe de placement, car vous ne pourrez peut-être pas ajouter des instances ultérieurement dans celui-ci.
     - Pour le Groupe de placement, cochez la case Add instance to placement group (Ajouter l'instance au groupe de placement) Si Placement group (Groupe de placement) ne figure pas

sur cette page, vérifiez que vous avez sélectionné un type d'instance qui peut être lancé dans un groupe de placement. Sinon, cette option n'est pas disponible.

- Pour Placement group name (Nom de groupe de placement), vous pouvez choisir d'ajouter les instances à un groupe de placement existant ou à un nouveau groupe de placement que vous créez.
- Choisissez une stratégie adaptée pour Placement group strategy (Stratégie de groupe de placement). Si vous choisissez partition, pour Target partition (Partition cible), choisissez Auto distribution (Distribution automatique) pour permettre à Amazon EC2 de répartir les instances aussi équitablement que possible entre toutes les partitions du groupe. Vous pouvez également spécifier la partition dans laquelle les instances seront lancées.

AWS CLI

Pour lancer les instances dans un groupe de placement à l'aide de l'AWS CLI

Utilisez la commande run-instances et spécifiez le nom du groupe de placement à l'aide du paramètre `--placement "GroupName = my-cluster"`. Dans cet exemple, le groupe de placement est nommé `my-cluster`.

```
aws ec2 run-instances --placement "GroupName = my-cluster"
```

Pour lancer les instances dans une partition spécifique d'un groupe de placement par partition à l'aide de l'AWS CLI

Utilisez la commande run-instances et spécifiez le nom du groupe de placement et la partition à l'aide du paramètre `--placement "GroupName = HDFS-Group-A, PartitionNumber = 3"`. Dans cet exemple, le groupe de placement est nommé `HDFS-Group-A` et il contient `3` partitions.

```
aws ec2 run-instances --placement "GroupName = HDFS-Group-A, PartitionNumber = 3"
```

PowerShell

Pour lancer les instances dans un groupe de placement à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande New-EC2Instance et spécifiez le nom du groupe de placement à l'aide du paramètre `-Placement_GroupName`.

Description des instances dans un groupe de placement

Vous pouvez afficher les informations de placement de vos instances en employant l'une des méthodes suivantes. Vous pouvez également filtrer les groupes de placement par partition par nombre de partitions à l'aide de l'AWS CLI.

Nouvelle console

Pour afficher le groupe de placement et le nombre de partitions d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.

4. Dans l'onglet Description sous Hôte et groupe de placement, recherchez Groupe de placement. Le champ est vide si l'instance ne figure pas dans un groupe de placement. Sinon, il contient le nom du groupe de placement. Si le groupe de placement est un groupe de placement, Numéro de partition contient le numéro de partition de l'instance.

Ancienne console

Pour afficher le groupe de placement et le nombre de partitions d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Dans l'onglet Description recherchez Groupe de placement. Le champ est vide si l'instance ne figure pas dans un groupe de placement. Sinon, il contient le nom du groupe de placement. Si le groupe de placement est un groupe de placement, Numéro de partition contient le numéro de partition de l'instance.

AWS CLI

Pour afficher le nombre de partitions d'une instance dans un groupe de placement par partition à l'aide de l'AWS CLI

Utilisez la commande describe-instances et spécifiez le paramètre `--instance-id`.

```
aws ec2 describe-instances --instance-id i-0123a456700123456
```

La réponse inclut les informations sur le placement, notamment le nom du groupe de placement et le nombre de partitions correspondant à l'instance.

```
"Placement": {
     "AvailabilityZone": "us-east-1c",
     "GroupName": "HDFS-Group-A",
     "PartitionNumber": 3,
     "Tenancy": "default"
}
```

Pour filtrer les instances pour un groupe de placement par partition et un nombre de partitions spécifiques à l'aide de l'AWS CLI

Utilisez la commande describe-instances et spécifiez le paramètre `--filters` avec les filtres `placement-group-name` et `placement-partition-number`. Dans cet exemple, le groupe de placement est nommé `HDFS-Group-A` et il contient `7` partitions.

```
aws ec2 describe-instances --filters "Name = placement-group-name, Values = HDFS-Group-
A" "Name = placement-partition-number, Values = 7"
```

La réponse répertorie toutes les instances qui figurent dans la partition désignée au sein du groupe de placement spécifié. Voici un exemple de sortie présentant uniquement l'ID et le type d'instance ainsi que les informations sur le placement pour les instances retournées.

```
"Instances": [
                {
                    "InstanceId": "i-0a1bc23d4567e8f90",
```

```
                "InstanceType": "r4.large",
                },

                "Placement": {
                    "AvailabilityZone": "us-east-1c",
                    "GroupName": "HDFS-Group-A",
                    "PartitionNumber": 7,
                    "Tenancy": "default"
                }

            {
                "InstanceId": "i-0a9b876cd5d4ef321",
                "InstanceType": "r4.large",
                },

                "Placement": {
                    "AvailabilityZone": "us-east-1c",
                    "GroupName": "HDFS-Group-A",
                    "PartitionNumber": 7,
                    "Tenancy": "default"
                }
        ],
```

Modification du groupe de placement d'une instance

Vous pouvez changer le groupe de placement d'une instance de l'une des façons suivantes :

- Déplacement d'une instance existante vers un groupe de placement
- Déplacement d'une instance d'un groupe de placement vers un autre
- Suppression d'une instance d'un groupe de placement

Avant de déplacer ou de supprimer l'instance, celle-ci doit être à l'état `stopped`. Vous pouvez déplacer ou supprimer une instance à l'aide de l'AWS CLI ou d'un kit SDK AWS.

AWS CLI

Pour déplacer une instance vers un groupe de placement à l'aide de l'AWS CLI

1. Arrêtez l'instance à l'aide de la commande stop-instances.
2. Utilisez la commande modify-instance-placement et indiquez le nom du groupe de placement vers lequel déplacer l'instance.

```
aws ec2 modify-instance-placement --instance-id i-0123a456700123456 --group-
name MySpreadGroup
```

3. Démarrez l'instance à l'aide de la commande start-instances.

PowerShell

Pour déplacer une instance vers un groupe de placement à l'aide d'Outils AWS pour Windows PowerShell

1. Arrêtez l'instance à l'aide de la commande Stop-EC2Instance.
2. Utilisez la commande Edit-EC2InstancePlacement et indiquez le nom du groupe de placement vers lequel déplacer l'instance.
3. Démarrez l'instance à l'aide de la commande Start-EC2Instance.

AWS CLI

Pour supprimer une instance d'un groupe de placement à l'aide de l'AWS CLI

1. Arrêtez l'instance à l'aide de la commande stop-instances.
2. Utilisez la commande modify-instance-placement et spécifiez une chaîne vide pour le nom du groupe.

```
aws ec2 modify-instance-placement --instance-id i-0123a456700123456 --group-name ""
```

3. Démarrez l'instance à l'aide de la commande start-instances.

PowerShell

Pour supprimer une instance d'un groupe de placement à l'aide d'Outils AWS pour Windows PowerShell

1. Arrêtez l'instance à l'aide de la commande Stop-EC2Instance.
2. Utilisez la commande Edit-EC2InstancePlacement et spécifiez une chaîne vide pour le nom du groupe de placement.
3. Démarrez l'instance à l'aide de la commande Start-EC2Instance.

Suppression d'un groupe de placement

Si vous avez besoin de supprimer un groupe de placement ou si vous n'en avez plus besoin, vous pouvez le supprimer. Vous pouvez supprimer un groupe de placement en employant l'une des méthodes suivantes.

Exigence

Pour pouvoir être supprimé, un groupe de placement ne doit pas contenir d'instances. Vous pouvez résilier (p. 647) toutes les instances que vous avez lancées dans le groupe de placement, les déplacer (p. 935) dans un autre groupe de placement ou les supprimer (p. 936) du groupe de placement.

Nouvelle console

Pour supprimer un groupe de placement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Groupes de placement.
3. Sélectionnez le groupe de placement et choisissez Actions, Supprimer.
4. Lorsque vous êtes invité à confirmer, entrez **`Delete`**, puis choisissez Delete (Supprimer).

Ancienne console

Pour supprimer un groupe de placement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Groupes de placement.
3. Sélectionnez le groupe de placement et choisissez Actions, Supprimer le groupe de placement.
4. Lorsque vous êtes invité à confirmer l'opération, choisissez Supprimer.

AWS CLI

Pour supprimer un groupe de placement à l'aide de l'AWS CLI

Utilisez la commande delete-placement-group et indiquez le nom du groupe de placement à supprimer. Dans cet exemple, le nom du groupe de placement est `my-cluster`.

```
aws ec2 delete-placement-group --group-name my-cluster
```

PowerShell

Pour supprimer un groupe de placement à l'aide d'Outils AWS pour Windows PowerShell

Utilisez la commande Remove-EC2PlacementGroup pour supprimer le groupe de placement.

Unité de transmission maximale (MTU) du réseau pour votre instance EC2

L'unité de transmission maximale (MTU) d'une connexion réseau correspond à la taille, en octets, du paquet le plus volumineux susceptible d'être transmis via la connexion. Plus la MTU d'une connexion est élevée, plus la quantité de données pouvant être transmises dans un seul paquet est importante. Les paquets Ethernet sont composés de la trame, ou des données réelles que vous envoyez, et des informations réseau générales associées.

Les trames Ethernet peuvent avoir différents formats, le plus courant étant le format de trame standard Ethernet v2. Il prend en charge une MTU de 1500, c'est-à-dire la taille de paquet Ethernet la plus importante prise en charge presque partout sur Internet. La MTU maximum prise en charge pour une instance dépend du type d'instance. Tous les types d'instance Amazon EC2 prennent en charge une MTU de 1500 et de nombreuses tailles d'instance actuelles prennent en charge une MTU de 9001 ou les trames jumbo.

Si votre instance s'exécute dans une zone de longueur d'onde, la valeur MTU maximale est de 1300.

Pour afficher les informations de la MTU du réseau pour les instances Windows, consultez cette page dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows : Unité de transmission maximale (MTU) du réseau pour votre instance EC2.

Sommaire

- Trames jumbo (MTU de 9001) (p. 937)
- Détection de la MTU du chemin (p. 938)
- Vérifier la MTU du chemin entre deux hôtes (p. 938)
- Vérification et définition de la MTU sur votre instance Linux (p. 939)
- Dépannage (p. 940)

Trames jumbo (MTU de 9001)

Les trames jumbo permettent d'utiliser plus de 1 500 octets de données en augmentant la charge utile par paquet, et donc en augmentant le pourcentage de paquet qui ne constitue pas des frais supplémentaires. Moins de paquets sont nécessaires pour envoyer le même volume de données utilisables. Toutefois, en dehors d'une région AWS donnée (EC2-Classic), d'un seul VPC ou d'une connexion d'appairage de VPC, vous rencontrerez un chemin maximum d'une MTU de 1500. Les connexions VPN et le trafic transmis via une passerelle Internet sont limités à une MTU de 1500. Si la taille des paquets dépasse 1 500 octets, ceux-ci sont fragmentés ou abandonnés si l'indicateur `Don't Fragment` est défini dans l'en-tête IP.

Les trames jumbo doivent être utilisées avec prudence pour le trafic Internet ou pour tout trafic quittant un VPC. Les paquets sont fragmentés par des systèmes intermédiaires, ce qui ralentit le trafic. Pour utiliser les trames jumbo dans un VPC et éviter de ralentir le trafic destiné à sortir du VPC, vous pouvez configurer la taille de MTU par routage ou utiliser plusieurs interfaces réseau Elastic avec différentes tailles de MTU et différents routages.

Pour les instances situées dans un même groupe de placement du cluster, les trames jumbo permettent d'atteindre le débit réseau maximum possible et elles sont recommandées dans ce cas. Pour plus d'informations, consultez Groupes de placement (p. 924).

Vous pouvez utiliser des trames jumbo pour le trafic entre vos VPC et vos réseaux sur site via AWS Direct Connect. Pour plus d'informations et pour savoir comment vérifier la capacité de trame Jumbo, consultez Définition de la MTU du réseau dans le AWS Direct Connect Guide de l'utilisateur.

Toutes les instances de la génération actuelle (p. 220) prennent en charge des trames jumbo. Les instances suivantes de l'ancienne génération prennent en charge des trames jumbo : C3, G2, I2, M3 et R3.

Pour plus d'informations sur les tailles MTU prises en charge pour les passerelles de transit, veuillez consulter MTU dans Passerelles de transit Amazon VPC.

Détection de la MTU du chemin

Path MTU Discovery (PMTUD) est utilisé pour déterminer l'unité de transmission maximale (MTU) d'un chemin réseau. Le MTU de chemin est la taille maximale du paquet entre l'hôte d'origine et l'hôte de réception. Si un hôte envoie un paquet dont la taille est plus importante que la MTU définie pour l'hôte destinataire ou que celle d'un appareil se trouvant sur le chemin, l'hôte ou l'appareil destinataire retourne le message ICMP suivant : `Destination Unreachable: Fragmentation Needed and Don't Fragment was Set` (Type 3, Code 4). Ce message demande à l'hôte initial de modifier la MTU jusqu'à ce que le paquet puisse être transmis.

Par défaut, les groupes de sécurité n'autorisent pas le trafic ICMP entrant. Toutefois, les groupes de sécurité sont avec état. Par conséquent, les réponses ICMP aux demandes sortantes sont autorisées à circuler, quelles que soient les règles du groupe de sécurité. Par conséquent, vous n'avez pas besoin d'ajouter explicitement une règle ICMP entrante pour vous assurer que votre instance peut recevoir la réponse du message ICMP. Pour de plus amples informations sur la configuration des règles ICMP dans une liste ACL réseau, veuillez consulter Découverte MTU de chemin dans le Amazon VPC Guide de l'utilisateur.

> **Important**
>
> Path MTU Discovery ne garantit pas que les trames jumbo ne seront pas abandonnées par certains routeurs. Une passerelle Internet sur votre VPC transmettra uniquement les paquets de 1 500 octets au maximum. Les paquets dont la MTU est de 1500 sont recommandés pour le trafic Internet.

Vérifier la MTU du chemin entre deux hôtes

Vous pouvez vérifier la MTU du chemin entre deux hôtes à l'aide de la commande tracepath, qui fait partie du package `iputils` disponible par défaut sur de nombreuses distributions Linux, dont Amazon Linux.

Pour vérifier la MTU du chemin à l'aide de tracepath

Utilisez la commande suivante pour vérifier la MTU du chemin entre votre instance EC2 et un autre hôte. Vous pouvez utiliser un nom DNS ou une adresse IP comme destination. Si la destination est une autre instance EC2, vérifiez que le groupe de sécurité autorise le trafic UDP entrant. Cet exemple vérifie la MTU du chemin entre une instance EC2 et `amazon.com`.

```
[ec2-user ~]$ tracepath amazon.com
 1?: [LOCALHOST]      pmtu 9001
```

```
 1:  ip-172-31-16-1.us-west-1.compute.internal (172.31.16.1)   0.187ms pmtu 1500
 1:  no reply
 2:  no reply
 3:  no reply
 4:  100.64.16.241 (100.64.16.241)                              0.574ms
 5:  72.21.222.221 (72.21.222.221)                             84.447ms asymm 21
 6:  205.251.229.97 (205.251.229.97)                           79.970ms asymm 19
 7:  72.21.222.194 (72.21.222.194)                             96.546ms asymm 16
 8:  72.21.222.239 (72.21.222.239)                             79.244ms asymm 15
 9:  205.251.225.73 (205.251.225.73)                           91.867ms asymm 16
...
31:  no reply
     Too many hops: pmtu 1500
     Resume: pmtu 1500
```

Dans cet exemple, la MTU du chemin est 1500.

Vérification et définition de la MTU sur votre instance Linux

Certaines instances sont configurées de façon à utiliser les trames jumbo, tandis que d'autres sont configurées de façon à utiliser les tailles de trame standard. Vous pouvez utiliser les trames jumbo pour le trafic réseau au sein de votre VPC ou utiliser des trames standard pour le trafic Internet. Quel que soit le cas de figure, nous vous recommandons de vérifier que votre instance se comportera comme vous le souhaitez. Vous pouvez utiliser les procédures de cette section afin de vérifier le paramètre MTU de votre interface réseau et de le modifier si nécessaire.

Pour vérifier le paramètre MTU sur une instance Linux

Vous pouvez vérifier valeur actuelle de la MTU à l'aide de la commande ip suivante. Notez que dans l'exemple de sortie, *`mtu 9001`* indique que cette instance utilise des trames jumbo.

```
[ec2-user ~]$ ip link show eth0
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 9001 qdisc pfifo_fast state UP mode DEFAULT
 group default qlen 1000
    link/ether 02:90:c0:b7:9e:d1 brd ff:ff:ff:ff:ff:ff
```

Pour définir la valeur de la MTU sur une instance Linux

1. Vous pouvez définir la valeur de la MTU à l'aide de la commande ip. La commande suivante définit la valeur souhaitée pour la MTU jusqu'à 1500, mais vous pouvez utiliser 9001 à la place.

   ```
   [ec2-user ~]$ sudo ip link set dev eth0 mtu 1500
   ```

2. (Facultatif) Pour conserver le paramètre de la MTU du réseau après le redémarrage, modifiez les fichiers de configuration suivants en fonction de votre type de système d'exploitation.

   - Pour Amazon Linux 2, ajoutez la ligne suivante au fichier `/etc/sysconfig/network-scripts/ifcfg-eth0` :

     ```
     MTU=1500
     ```

     Ajoutez la ligne suivante dans le fichier `/etc/dhcp/dhclient.conf` :

     ```
     request subnet-mask, broadcast-address, time-offset, routers, domain-name, domain-
     search, domain-name-servers, host-name, nis-domain, nis-servers, ntp-servers;
     ```

- Pour Amazon Linux, ajoutez les lignes suivantes à votre fichier `/etc/dhcp/dhclient-eth0.conf`.

```
interface "eth0" {
supersede interface-mtu 1500;
}
```

- Pour les autres distributions Linux, consultez leur documentation spécifique.

3. (Facultatif) Redémarrez l'instance et vérifiez que le paramètre MTU est correct.

Dépannage

Si vous rencontrez des problèmes de connectivité entre votre instance EC2 et un cluster Amazon Redshift lorsque vous utilisez les trames Jumbo, consultez Des requêtes semblent se bloquer dans le Amazon Redshift Cluster Management Guide

Clouds privés virtuels

Amazon Virtual Private Cloud (Amazon VPC) vous permet de définir un réseau virtuel dans votre propre domaine isolé de manière logique dans le cloud AWS, appelé virtual private cloud (VPC). Vous pouvez lancer vos ressources Amazon EC2, comme des instances dans les sous-réseaux de votre VPC. Votre VPC ressemble beaucoup à un réseau traditionnel que vous pourriez exécuter dans votre propre centre de données, et présente l'avantage d'utiliser l'infrastructure scalable d'AWS. Vous pouvez configurer votre VPC en sélectionnant sa plage d'adresses IP, en créant des sous-réseaux et en configurant des tables de routage, des passerelles réseau et des paramètres de sécurité. Vous pouvez connecter des instances dans votre VPC à Internet ou à votre propre centre de données.

Lorsque vous créez votre compte AWS, nous créons un VPC par défaut pour vous dans chaque région. Un VPC par défaut est un VPC déjà configuré et prêt à être utilisé par vous. Vous pouvez lancer des instances dans votre VPC par défaut immédiatement. Vous pouvez également créer votre propre VPC personnalisé et le configurer selon vos besoins.

Si vous avez créé votre compte AWS avant le 4 décembre 2013, vous pouvez bénéficier de la prise en charge de la plateforme EC2-Classic dans certaines régions. Si vous avez créé votre compte AWS après le 4 décembre 2013, il ne prend pas en charge EC2-Classic. Vous devez donc lancer vos ressources dans un VPC. Pour plus d'informations, consultez EC2-Classic (p. 941).

Documentation Amazon VPC

Pour plus d'information sur Amazon VPC, consultez la documentation suivante.

Guide	Description
Amazon VPC Guide de l'utilisateur	Décrit les concepts clés et fournit des instructions d'utilisation pour les fonctions d'Amazon VPC.
Amazon VPC Peering Guide	Décrit les connexions d'appairage de VPC et fournit des instructions d'utilisation.
Passerelles de transit Amazon VPC	Décrit les passerelles de transit et fournit des instructions pour leur configuration et leur utilisation.

Guide	Description
Guide de l'utilisateur AWS Site-to-Site VPN	Décrit les connexions Site-to-Site VPN et fournit des instructions pour leur configuration et leur utilisation.

EC2-Classic

Avec EC2-Classic, vos instances s'exécutent dans un réseau plat unique partagé avec d'autres clients. Avec Amazon VPC, vos instances s'exécutent dans un VPC (virtual private cloud) sur votre compte AWS.

La plateforme EC2-Classic a été lancée dans la version d'origine d'Amazon EC2. Si vous avez créé votre compte AWS après le 4 décembre 2013, il ne prend pas en charge EC2-Classic. Vous devez donc lancer vos instances Amazon EC2 dans un VPC.

Si votre compte ne prend pas en charge EC2-Classic, nous créons un VPC par défaut pour vous. Par défaut, lorsque vous lancez une instance, nous la lançons dans votre VPC par défaut. Vous pouvez également créer un VPC autre qu'un VPC par défaut et le spécifier lorsque vous lancez une instance.

Suppression des plateformes prises en charge

La console Amazon EC2 indique dans quelles plateformes vous pouvez lancer des instances pour la région sélectionnée et si vous disposez d'un VPC par défaut dans cette région.

Assurez-vous que la région à utiliser est bien sélectionnée dans la barre de navigation. Sur le tableau de bord de la console Amazon EC2, recherchez la section Plateformes prises en charge sous Attributs du compte.

Comptes prenant en charge EC2-Classic

Le tableau de bord affiche les informations suivantes sous Attributs du compte pour indiquer que le compte prend uniquement en charge la plateforme EC2-Classic et les VPC dans cette région, mais que cette dernière n'a pas de VPC par défaut.

La sortie de la commande describe-account-attributes inclut les valeurs `EC2` et `VPC` de l'attribut `supported-platforms`.

```
aws ec2 describe-account-attributes --attribute-names supported-platforms
{
    "AccountAttributes": [
        {
            "AttributeName": "supported-platforms",
            "AttributeValues": [
                {
                    "AttributeValue": "EC2"
                },
                {
                    "AttributeValue": "VPC"
                }
```

```
            ]
        }
    ]
}
```

Comptes qui requièrent un VPC

Le tableau de bord affiche les informations suivantes sous Attributs du compte pour indiquer que le compte requiert un VPC pour lancer des instances dans cette région, ne prend pas en charge la plateforme EC2-Classic dans cette région, et que cette dernière a un VPC par défaut associé à l'identifiant `vpc-1a2b3c4d`.

La sortie de la commande describe-account-attributes pour la région spécifiée inclut seulement la valeur `VPC` pour l'attribut `supported-platforms`.

```
aws ec2 describe-account-attributes --attribute-names supported-platforms --region us-
east-2
{
    "AccountAttributes": [
        {
            "AttributeValues": [
                {
                    "AttributeValue": "VPC"
                }
            ]
            "AttributeName": "supported-platforms",
        }
    ]
}
```

Types d'instances disponibles dans EC2-Classic

La plupart des instances les plus récentes requièrent VPC. Les seuls types d'instance pris en charge dans EC2-Classic sont les suivants :

- Usage général : M1, M3 et T1
- Calcul optimisé : C1, C3 et CC2
- Mémoire optimisée : CR1, M2 et R3
- Stockage optimisé : D2, HS1 et I2
- Calcul accéléré : G2

Si votre compte prend en charge EC2-Classic, mais que vous n'avez pas créé de VPC personnalisé, vous pouvez effectuer l'une des actions suivantes pour lancer des instances qui requièrent un VPC :

- Créez un VPC personnalisé et lancez-y votre instance pour VPC uniquement en spécifiant un ID de sous-réseau (subnet) ou un ID d'interface réseau dans la requête. Notez que vous devez créer un VPC personnalisé si vous n'avez aucun VPC par défaut et que vous utilisez l'AWS CLI, l'API Amazon EC2 ou un kit SDK AWS pour lancer une instance pour VPC uniquement.

- Lancez votre instance pour VPC uniquement à l'aide de la console Amazon EC2. La console Amazon EC2 crée un VPC personnalisé dans votre compte et lance l'instance dans le sous-réseau, dans la première zone de disponibilité. La console crée le VPC avec les attributs suivants :
  - Un sous-réseau de chaque zone de disponibilité avec l'attribut de l'adressage IPv4 public défini sur `true` afin que les instances reçoivent une adresse IPv4 publique. Pour plus d'informations, consultez Adressage IP dans votre VPC dans le Amazon VPC Guide de l'utilisateur.
  - Une passerelle Internet et une table de routage principale qui achemine le trafic dans le VPC vers la passerelle Internet. Cela permet aux instances que vous lancez dans le VPC de communiquer via Internet. Pour plus d'informations, consultez Passerelles Internet dans le Amazon VPC Guide de l'utilisateur.
  - Un groupe de sécurité par défaut pour le VPC et une liste ACL réseau par défaut qui est associée à chaque sous-réseau (subnet). Pour plus d'informations, consultez la section Groupes de sécurité pour votre VPC dans le manuel Amazon VPC Guide de l'utilisateur.

Si vous avez d'autres ressources dans EC2-Classic, vous pouvez les migrer vers un VPC. Pour plus d'informations, consultez Migration d'EC2-Classic vers un VPC (p. 961).

Différences entre les instances d'EC2-Classic et d'un VPC

Le tableau ci-après récapitule les différences entre les instances lancées dans EC2-Classic, les instances lancées dans un VPC par défaut et les instances lancées dans un VPC autre qu'un VPC par défaut.

Caractéristiques	EC2-Classic	VPC par défaut	VPC personnalisé
Adresse IPv4 publique (à partir du pool d'adresses IP publiques d'Amazon)	Votre instance reçoit une adresse IPv4 publique du groupe d'adresses IPv4 publiques d'EC2-Classic.	Votre instance lancée dans un sous-réseau par défaut reçoit une adresse IPv4 publique par défaut, à moins que vous ne spécifiez une autre option lors du lancement ou que vous ne modifiiez l'attribut d'adresse IPv4 publique du sous-réseau.	Votre instance ne reçoit pas d'adresse IPv4 publique par défaut, à moins que vous ne spécifiez une autre option lors du lancement ou que vous ne modifiiez l'attribut d'adresse IPv4 publique du sous-réseau.
Adresse IPv4 privée	Votre instance reçoit une adresse IPv4 privée de la plage d'adresses EC2-Classic à chacun de ses démarrages.	Votre instance reçoit une adresse IPv4 privée statique de la plage d'adresses de votre VPC par défaut.	Votre instance reçoit une adresse IPv4 privée statique de la plage d'adresses de votre VPC.
Adresses IPv4 privées multiples	Nous sélectionnons une adresse IP privée unique pour votre instance. Les adresses IP multiples ne sont pas prises en charge.	Vous pouvez affecter plusieurs adresses IPv4 privées à votre instance.	Vous pouvez affecter plusieurs adresses IPv4 privées à votre instance.
Adresse IP Elastic (IPv4)	Une adresse IP Elastic est dissociée de l'instance lorsque vous arrêtez cette dernière.	Une adresse IP Elastic reste associée à l'instance lorsque vous arrêtez cette dernière.	Une adresse IP Elastic reste associée à l'instance lorsque vous arrêtez cette dernière.
Association d'une	Vous associez une adresse IP Elastic à une instance.	Une adresse IP Elastic est une propriété d'une	Une adresse IP Elastic est une propriété d'une

Caractéristiques	EC2-Classic	VPC par défaut	VPC personnalisé
adresse IP Elastic		interface réseau. Vous pouvez associer une adresse IP Elastic à une instance en mettant à jour l'interface réseau attachée à l'instance.	interface réseau. Vous pouvez associer une adresse IP Elastic à une instance en mettant à jour l'interface réseau attachée à l'instance.
Réassociation d'une adresse IP Elastic	Si l'adresse IP Elastic est déjà associée à une autre instance, l'adresse est automatiquement associée à la nouvelle instance.	Si l'adresse IP Elastic est déjà associée à une autre instance, l'adresse est automatiquement associée à la nouvelle instance.	Si l'adresse IP Elastic est déjà associée à une autre instance, cette opération aboutit uniquement si vous avez autorisé la réassociation.
Balisage d'adresses IP Elastic	Vous ne pouvez pas appliquer des balises à une adresse IP Elastic.	Vous pouvez appliquer des balises à une adresse IP Elastic.	Vous pouvez appliquer des balises à une adresse IP Elastic.
Noms d'hôte DNS	Les noms d'hôte DNS sont activés par défaut.	Les noms d'hôte DNS sont activés par défaut.	Les noms d'hôte DNS sont désactivés par défaut.
Groupe de sécurité	Un groupe de sécurité peut référencer des groupes de sécurité appartenant à d'autres comptes AWS.	Vous pouvez référencer des groupes de sécurité de votre VPC ou d'un VPC pair dans une connexion d'appairage VPC.	Un groupe de sécurité peut référencer des groupes de sécurité pour votre VPC uniquement.
Association d'un groupe de sécurité	Vous ne pouvez pas changer les groupes de sécurité de votre instance en cours d'exécution. Vous pouvez modifier les règles des groupes de sécurité assignés, ou remplacer l'instance par une nouvelle instance (créez une AMI depuis l'instance, lancez une nouvelle instance depuis cette AMI avec les groupes de sécurité dont vous avez besoin, dissociez toutes les adresses IP Elastic de l'instance d'origine et associez-les à la nouvelle instance, puis mettez fin à l'instance d'origine).	Vous pouvez assigner jusqu'à 5 groupes de sécurité à une instance. Vous pouvez assigner des groupes de sécurité à votre instance lorsque vous la lancez et pendant son exécution.	Vous pouvez assigner jusqu'à 5 groupes de sécurité à une instance. Vous pouvez assigner des groupes de sécurité à votre instance lorsque vous la lancez et pendant son exécution.
Règles des groupes de sécurité	Vous pouvez ajouter des règles pour le trafic entrant uniquement.	Vous pouvez ajouter des règles pour le trafic entrant et sortant.	Vous pouvez ajouter des règles pour le trafic entrant et sortant.
Location	Votre instance s'exécute sur un matériel partagé.	Vous pouvez exécuter l'instance sur un matériel partagé ou sur un matériel à client unique.	Vous pouvez exécuter l'instance sur un matériel partagé ou sur un matériel à client unique.

Caractéristiques	EC2-Classic	VPC par défaut	VPC personnalisé
Accès à Internet	Votre instance peut accéder à Internet. Votre instance reçoit automatiquement une adresse IP publique et peut accéder à Internet directement depuis le bout du réseau AWS.	Par défaut, votre instance peut accéder à Internet. Votre instance reçoit une adresse IP publique par défaut. Une passerelle Internet est attachée à votre VPC par défaut, et votre sous-réseau (subnet) par défaut possède une route vers la passerelle Internet.	Par défaut, votre instance ne peut pas accéder à Internet. Votre instance ne reçoit pas d'adresse IP publique par défaut. Votre VPC peut avoir une passerelle Internet, en fonction de la façon dont il a été créé.
Adressage IPv6	L'adressage IPv6 n'est pas pris en charge. Vous ne pouvez pas attribuer d'adresses IPv6 à vos instances.	Vous pouvez éventuellement associer un bloc d'adresses CIDR IPv6 à votre VPC et attribuer des adresses IPv6 aux instances de votre VPC.	Vous pouvez éventuellement associer un bloc d'adresses CIDR IPv6 à votre VPC et attribuer des adresses IPv6 aux instances de votre VPC.

Groupes de sécurité pour EC2-Classic

Si vous utilisez EC2-Classic, vous devez employer les groupes de sécurité créés spécifiquement pour EC2-Classic. Lorsque vous démarrez une instance dans EC2-Classic, vous devez spécifier un groupe de sécurité de la même région que l'instance. Vous ne pouvez pas spécifier un groupe de sécurité que vous avez créé pour un VPC quand vous lancez une instance dans EC2-Classic.

Après avoir lancé une instance dans EC2-Classic, vous ne pouvez pas modifier ses groupes de sécurité. Cependant, vous pouvez ajouter des règles à un groupe de sécurité ou en supprimer. Les modifications sont appliquées automatiquement à toutes les instances associées au groupe de sécurité après une brève période.

Votre compte AWS possède automatiquement un groupe de sécurité par défaut par région pour EC2-Classic. Si vous essayez de supprimer le groupe de sécurité par défaut, vous obtenez l'erreur suivante : Client.InvalidGroup.Reserved: The security group 'default' is reserved.

Vous pouvez créer des groupes de sécurité personnalisés. Le nom du groupe de sécurité doit être unique dans votre compte pour la région. Pour créer un groupe de sécurité à utiliser dans EC2-Classic, choisissez Pas de VPC pour le VPC.

Vous pouvez ajouter des règles entrantes à vos groupes de sécurité par défaut et personnalisés. Vous ne pouvez pas modifier les règles sortantes pour un groupe de sécurité EC2-Classic. Quand vous créez une règle de groupe de sécurité, vous pouvez utiliser un groupe de sécurité différent pour EC2-Classic dans la même région que la source ou la destination. Pour spécifier un groupe de sécurité pour un autre compte AWS (ajoutez l'ID de compte AWS comme préfixe ; par exemple, `111122223333/sg-edcd9784`).

Dans EC2-Classic, vous pouvez avoir jusqu'à 500 groupes de sécurité dans chaque région pour chaque compte. Vous pouvez ajouter jusqu'à 100 règles à un groupe de sécurité. Vous pouvez avoir jusqu'à 800 règles de groupe de sécurité par instance. Cette limite est calculée avec la formule règles par groupe de sécurité multiplié par groupes de sécurité par instance. Si vous référencez d'autres groupes de sécurité dans vos règles de groupe de sécurité, nous vous recommandons d'utiliser des noms de groupes de sécurité de 22 caractères maximum.

Adressage IP et DNS

Amazon fournit un serveur DNS qui résout les noms d'hôte DNS IPv4 fournis par Amazon aux adresses IPv4. Dans EC2-Classic, le serveur Amazon DNS se trouve à l'adresse `172.16.0.23`.

Si vous créez une configuration de pare-feu personnalisée dans EC2-Classic, vous devez créer dans votre pare-feu une règle permettant le trafic entrant depuis le port 53 (DNS), avec un port de destination issu de la plage éphémère, à partir de l'adresse du serveur Amazon DNS. Si vous ne procédez pas ainsi, la résolution DNS interne de vos instances échouera. Si votre pare-feu n'autorise pas automatiquement les réponses aux requêtes DNS, vous devez autoriser le trafic provenant de l'adresse IP du serveur Amazon DNS. Pour obtenir l'adresse IP du serveur Amazon DNS, utilisez la commande suivante depuis votre instance :

```
grep nameserver /etc/resolv.conf
```

Adresses IP Elastic

Si votre compte prend en charge EC2-Classic, il existe un groupe d'adresses IP Elastic à utiliser avec la plateforme EC2-Classic et un autre groupe correspondant à vos VPC. Vous ne pouvez pas associer une adresse IP Elastic que vous avez allouée pour être utilisée avec un VPC à une instance dans EC2-Classic, et inversement. Par contre, vous pouvez migrer une adresse IP Elastic que vous avez allouée pour être utilisée sur la plateforme EC2-Classic à utiliser avec un VPC. Vous ne pouvez pas migrer une adresse IP Elastic vers une autre région.

Pour allouer une adresse IP Elastic à utiliser dans EC2-Classic à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Choisissez Allouer une nouvelle adresse.
4. Sélectionnez Classique, puis choisissez Allouer. Fermez l'écran de confirmation.

Migration d'une Adresse IP Elastic à partir de EC2-Classic

Si votre compte prend en charge EC2-Classic, vous pouvez migrer des adresses IP Elastic que vous avez allouées pour une utilisation avec la plateforme EC2-Classic à utiliser avec un VPC, au sein de la même région. Cela peut vous aider à migrer vos ressources de EC2-Classic vers un VPC ; par exemple, vous pouvez lancer de nouveaux serveurs web dans votre VPC, puis utiliser les mêmes adresses IP Elastic que celles de vos serveurs web dans EC2-Classic pour vos nouveaux serveurs web VPC.

Une fois que vous avez migré une adresse IP Elastic vers un VPC, vous ne pouvez pas l'utiliser avec EC2-Classic. En revanche, vous pouvez la restaurer dans EC2-Classic, si nécessaire. Vous ne pouvez pas migrer une adresse IP Elastic allouée initialement pour être utilisée avec un VPC vers EC2-Classic.

Pour que vous puissiez migrer une adresse IP Elastic, celle-ci ne doit pas être associée à une instance. Pour plus d'informations sur la dissociation d'une adresse IP Elastic d'une instance, consultez Dissociation d'une adresse IP Elastic (p. 837).

Vous pouvez migrer autant d'adresses IP Elastic EC2-Classic que peut comporter votre compte. Toutefois, lorsque vous migrez une adresse IP Elastic, celle-ci est comptabilisée dans votre limite d'adresses IP Elastic pour VPC. Vous ne pouvez pas migrer une adresse IP Elastic si cela entraîne un dépassement de votre limite. De même, lorsque vous restaurez une adresse IP Elastic vers EC2-Classic, celle-ci est comptabilisée dans votre limite d'adresses IP Elastic pour EC2-Classic. Pour plus d'informations, consultez Limite appliquée aux adresses IP Elastic (p. 839).

Vous ne pouvez pas migrer une adresse IP Elastic qui a été allouée à votre compte il y a moins de 24 heures.

Vous pouvez migrer une adresse IP Elastic d'EC2-Classic à l'aide de la console Amazon EC2 ou Amazon VPC. Cette option est disponible uniquement si votre compte prend en charge EC2-Classic.

Pour déplacer une adresse IP Elastic à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic et choisissez Actions, puis Transférer dans la portée de VPC.
4. Dans la boîte de dialogue de confirmation, choisissez Transférer l'adresse IP élastique.

Vous pouvez restaurer une adresse IP Elastic sur EC2-Classic à l'aide de la console Amazon EC2 ou Amazon VPC.

Pour restaurer une adresse IP Elastic sur EC2-Classic à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Elastic IPs.
3. Sélectionnez l'adresse IP Elastic, choisissez Actions, puis Restaurer dans la portée EC2.
4. Dans la boîte de dialogue de confirmation, sélectionnez Restaurer.

Une fois que vous avez exécuté la commande pour déplacer ou restaurer votre adresse IP Elastic, le processus de migration de l'adresse IP Elastic peut prendre quelques minutes. Utilisez la commande describe-moving-addresses pour vérifier si votre adresse IP Elastic est encore en cours de déplacement ou si le déplacement est terminé.

Une fois que vous avez déplacé votre adresse IP Elastic, vous pouvez consulter son ID d'allocation sur la page Adresses IP Elastic dans le champ ID d'allocation.

Si l'adresse IP Elastic reste à l'état de déplacement plus de 5 minutes, contactez Premium Support.

Pour déplacer une adresse IP Elastic à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- move-address-to-vpc (AWS CLI)
- Move-EC2AddressToVpc (Outils AWS pour Windows PowerShell)

Pour restaurer une adresse IP Elastic dans EC2-Classic à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- restore-address-to-classic (AWS CLI)
- Restore-EC2AddressToClassic (Outils AWS pour Windows PowerShell)

Pour décrire l'état de vos adresses en déplacement à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-moving-addresses (AWS CLI)
- Get-EC2Address (Outils AWS pour Windows PowerShell)

Partager et accéder à des ressources entre EC2-Classic et un VPC

Certaines ressources et fonctions de votre compte AWS peuvent être accessibles ou partagées entre les plateformes EC2-Classic et un VPC, par exemple via ClassicLink. Pour plus d'informations, consultez ClassicLink (p. 949).

Si votre compte prend en charge EC2-Classic, vous avez peut-être paramétré des ressources à utiliser dans EC2-Classic. Si vous souhaitez effectuer une migration depuis EC2-Classic vers un VPC, vous devez recréer ces ressources dans votre VPC. Pour plus d'informations sur la migration depuis EC2-Classic vers un VPC, consultez Migration d'EC2-Classic vers un VPC (p. 961).

Les ressources suivantes peuvent être accessibles ou partagées entre EC2-Classic et un VPC.

Ressource	Remarques
AMI	
Tâche de bundle	
Volume EBS	
Adresse IP Elastic (IPv4)	Vous pouvez migrer une adresse IP Elastic à partir de EC2-Classic vers un VPC. Vous ne pouvez pas migrer une adresse IP Elastic allouée initialement pour être utilisée dans un VPC vers EC2-Classic. Pour plus d'informations, consultez Migration d'une Adresse IP Elastic à partir de EC2-Classic (p. 946).
Instance	Une instance EC2-Classic peut communiquer avec des instances dans un VPC en utilisant des adresses IPv4 publiques. Vous pouvez également utiliser ClassicLink pour permettre la communication via des adresses IPv4 privées. Vous ne pouvez pas migrer une instance à partir de EC2-Classic vers un VPC. Toutefois, vous pouvez migrer votre application depuis une instance dans EC2-Classic vers une instance dans un VPC. Pour plus d'informations, consultez Migration d'EC2-Classic vers un VPC (p. 961).
Paire de clés	
Équilibreur de charge	Si vous utilisez ClassicLink, vous pouvez enregistrer une instance EC2-Classic liée à un équilibreur de charge dans un VPC, sous réserve que le VPC ait un sous-réseau dans la même zone de disponibilité que l'instance. Vous ne pouvez pas migrer un équilibreur de charge depuis EC2-Classic vers un VPC. Vous ne pouvez pas enregistrer une instance dans un VPC avec un équilibreur de charge dans EC2-Classic.
Groupe de placement	

Ressource	Remarques
Instance réservée	Vous pouvez modifier la plateforme réseau de vos Instances réservées d'EC2-Classic vers un VPC. Pour plus d'informations, consultez Modification des Instances réservées (p. 357).
Groupe de sécurité	Une instance EC2-Classic liée peut utiliser les groupes de sécurité VPC via ClassicLink pour contrôler le trafic en provenance ou à destination du VPC. Les instances VPC ne peuvent pas utiliser les groupes de sécurité EC2-Classic. Vous ne pouvez pas migrer un groupe de sécurité depuis EC2-Classic vers un VPC. Vous pouvez copier les règles d'un groupe de sécurité d'EC2-Classic vers un groupe de sécurité d'un VPC. Pour plus d'informations, consultez Création d'un groupe de sécurité (p. 1067).
Instantané	

Les ressources suivantes ne peuvent être partagées ou déplacées entre EC2-Classic et un VPC :

- Instances Spot

ClassicLink

ClassicLink vous permet de lier des instances EC2-Classic à un VPC de votre compte, au sein de la même région. Si vous associez des groupes de sécurité VPC à l'instance EC2-Classic, cela permet la communication entre votre instance EC2-Classic et les instances de votre VPC à l'aide d'adresses IPv4 privées. ClassicLink élimine la nécessité d'utiliser des adresses IPv4 publiques ou des adresses IP Elastic pour activer la communication entre des instances dans ces plateformes.

ClassicLink est disponible pour tous les utilisateurs avec des comptes prenant en charge la plateforme EC2-Classic et peut être utilisé avec n'importe quelle instance EC2-Classic. Pour plus d'informations sur la migration de vos ressources vers un VPC, consultez Migration d'EC2-Classic vers un VPC (p. 961).

Il n'y a pas frais supplémentaires pour ClassicLink. Des frais standard pour le transfert de données et l'utilisation d'une instance sont applicables.

Sommaire

- Principes de base d'un ClassicLink (p. 949)
- Limites ClassicLink (p. 952)
- Utilisation de ClassicLink (p. 953)
- Exemples de stratégies IAM pour ClassicLink (p. 956)
- Exemple : Configuration d'un groupe de sécurité ClassicLink pour une application web à trois niveaux (p. 958)

Principes de base d'un ClassicLink

La liaison d'une instance EC2-Classic à un VPC à l'aide de ClassicLink s'effectue en deux étapes. D'abord, vous devez activer le VPC pour ClassicLink. Par défaut, tous les VPC de votre compte ne sont pas activés pour ClassicLink, afin de conserver leur isolement. Une fois que vous avez activé le VPC pour ClassicLink,

vous pouvez alors lier toute instance EC2-Classic en cours d'exécution dans la même région de votre compte à ce VPC. La liaison de votre instance inclut la sélection de groupes de sécurité du VPC à associer à votre instance EC2-Classic. Une fois que vous avez lié l'instance, celle-ci peut communiquer avec des instances de votre VPC à l'aide de leurs adresses IP privées, à condition que les groupes de sécurité VPC l'autorisent. Votre instance EC2-Classic ne perd pas son adresse IP privée lorsqu'elle est liée au VPC.

Le fait de lier votre instance à un VPC est parfois appelé attacher votre instance.

Une instance EC2-Classic liée peut communiquer avec des instances d'un VPC, mais elle ne fait pas partie du VPC. Si vous répertoriez vos instances et que vous filtrez par VPC, par exemple, par le biais de la requêtes d'API `DescribeInstances` ou à l'aide de l'écran Instances de la console Amazon EC2, les résultats ne renvoient pas les instances EC2-Classic qui sont liées au VPC. Pour plus d'informations sur l'affichage de vos instances EC2-Classic liées, consultez Affichage de vos VPC activés pour ClassicLink et de vos instances liées (p. 955).

Par défaut, si vous utilisez un nom d'hôte DNS public pour définir l'adresse d'une instance dans un VPC à partir d'une instance EC2-Classic liée, le nom d'hôte est résolu en l'adresse IP publique de l'instance. Il en va de même si vous utilisez un nom d'hôte DNS public pour définir l'adresse d'une instance EC2-Classic liée à partir d'une instance du VPC. Si vous souhaitez que le nom d'hôte DNS public soit résolu en une adresse IP privée, vous pouvez activer la prise en charge de ClassicLink DNS pour le VPC. Pour plus d'informations, consultez Activation de la prise en charge de DNS ClassicLink (p. 955).

Si vous n'avez plus besoin d'une connexion ClassicLink entre votre instance et le VPC, vous pouvez détacher l'instance EC2-Classic du VPC. Vous dissociez ainsi les groupes de sécurité du VPC de l'instance EC2-Classic. Une instance EC2-Classic liée est automatiquement détachée d'un VPC lorsqu'elle est arrêtée. Une fois que vous avez détaché toutes les instances EC2-Classic liées du VPC, vous pouvez désactiver ClassicLink pour le VPC.

Utilisation d'autres services AWS dans votre VPC avec ClassicLink

Les instances EC2-Classic liées peuvent accéder aux services AWS suivants dans le VPC : Amazon Redshift, Amazon ElastiCache, Elastic Load Balancing et Amazon RDS. Par contre, les instances du VPC ne peuvent pas accéder aux services AWS alloués par la plateforme EC2-Classic à l'aide de ClassicLink.

Si vous utilisez Elastic Load Balancing, vous pouvez enregistrer vos instances EC2-Classic liées avec l'équilibreur de charge. Vous devez créer votre équilibreur de charge dans le VPC activé pour ClassicLink VPC et activer la Zone de disponibilité dans laquelle l'instance s'exécute. Si vous mettez fin à l'instance EC2-Classic liée, l'équilibreur de charge annule l'enregistrement de l'instance.

Si vous utilisez Amazon EC2 Auto Scaling, vous pouvez créer un groupe Amazon EC2 Auto Scaling avec des instances qui sont liées automatiquement à un VPC activé pour ClassicLink spécifié au lancement. Pour plus d'informations, consultez Liaison d'instances EC2-Classic à un VPC dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

Si vous utilisez des instances Amazon RDS ou des clusters Amazon Redshift dans votre VPC, et que ceux-ci sont accessibles publiquement (accessibles à partir d'Internet), le point de terminaison dont vous vous servez pour définir l'adresse de ces ressources à partir d'une instance EC2-Classic liée est résolu par défaut en une adresse IP publique. Si ces ressources ne sont pas accessibles publiquement, le point de terminaison est résolu par défaut en une adresse IP privée. Pour définir l'adresse d'une instance RDS ou d'un cluster Redshift accessible publiquement via une adresse IP privée à l'aide de ClassicLink, vous devez utiliser son adresse IP privée ou son nom d'hôte DNS privé, ou vous devez activer la prise en charge de ClassicLink DNS pour le VPC.

Si vous utilisez un nom d'hôte DNS privé ou une adresse IP privée pour définir l'adresse d'une instance RDS, l'instance EC2-Classic liée ne peut pas utiliser la prise en charge du failover (basculement) disponible pour les déploiements multi-AZ.

Vous pouvez utiliser la console Amazon EC2 pour rechercher les adresses IP privées de vos ressources Amazon Redshift, Amazon ElastiCache ou Amazon RDS.

Pour rechercher les adresses IP privées de vos ressources AWS dans votre VPC

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Network Interfaces.
3. Consultez les descriptions des interfaces réseau dans la colonne Description. Le nom du service figurera dans la description pour une interface réseau utilisée par Amazon Redshift, Amazon ElastiCache ou Amazon RDS. Par exemple, une interface réseau attachée à une instance Amazon RDS comportera la description suivante : `RDSNetworkInterface`.
4. Sélectionnez l'interface réseau requise.
5. Dans le volet de détails, obtenez l'adresse IP privée dans le champ IP privée principale IPv4.

Contrôle de l'utilisation des ClassicLink

Par défaut, les utilisateurs IAM ne sont pas autorisés à utiliser ClassicLink. Vous pouvez utiliser une stratégie IAM octroyant aux utilisateurs des autorisations d'activer ou de désactiver un VPC pour ClassicLink, de lier une instance à un VPC activé pour ClassicLink, ou de la détacher et d'afficher les VPC ClassicLink et les instances EC2-Classic liées. Pour plus d'informations sur les stratégies IAM pour Amazon EC2, consultez Stratégies IAM pour Amazon EC2 (p. 979).

Pour plus d'informations sur les stratégies d'utilisation de ClassicLink, consultez l'exemple suivant : Exemples de stratégies IAM pour ClassicLink (p. 956).

Groupes de sécurité dans ClassicLink

La liaison de votre instance EC2-Classic à un VPC n'a pas d'incidence sur vos groupes de sécurité EC2-Classic. Ces derniers continuent de contrôler l'ensemble du trafic vers et depuis l'instance. Cela exclut le trafic vers et depuis les instances dans le VPC, qui est contrôlé par les groupes de sécurité VPC que vous avez associés à l'instance EC2-Classic. Les instances EC2-Classic qui sont liées au même VPC ne peuvent pas communiquer entre elles par le biais du VPC, qu'elles soient ou non associées au même groupe de sécurité VPC. La communication entre les instances EC2-Classic est contrôlée par les groupes de sécurité EC2-Classic que vous avez associés à ces instances. Pour accéder à un exemple de configuration d'un groupe de sécurité, consultez Exemple : Configuration d'un groupe de sécurité ClassicLink pour une application web à trois niveaux (p. 958).

Une fois que vous avez lié votre instance à un VPC, vous ne pouvez pas changer les groupes de sécurité VPC qui sont associés à l'instance. Pour associer d'autres groupes de sécurité à l'instance, vous devez d'abord détacher l'instance, puis lier à nouveau celle-ci au VPC, en sélectionnant les groupes de sécurité requis.

Routage pour ClassicLink

Lorsque vous activez un VPC pour ClassicLink, une route statique est ajoutée à toutes les tables de routage VPC avec la destination `10.0.0.0/8` et la cible `local`. Cela permet la communication entre les instances du VPC et les instances EC2-Classic liées au VPC. Si vous ajoutez une table de routage personnalisée à un VPC activé pour ClassicLink, une route statique est ajoutée automatiquement avec la destination `10.0.0.0/8` et la cible `local`. Lorsque vous désactivez ClassicLink pour un VPC, cette route est supprimée automatiquement dans toutes les tables de routage VPC.

Les VPC compris dans les plages d'adresses IP `10.0.0.0/16` et `10.1.0.0/16` peuvent être activés pour ClassicLink uniquement s'ils ne disposent pas de routes statiques existantes dans les tables de routage de la plage d'adresses IP `10.0.0.0/8`, à l'exclusion des routes locales ajoutées automatiquement lorsque le VPC a été créé. De même, si vous avez activé un VPC pour ClassicLink, vous ne pourrez peut-être pas ajouter des routes plus spécifiques à vos tables de routage dans la plage d'adresses IP `10.0.0.0/8`.

> **Important**
>
> Si le bloc d'adresse CIDR de votre VPC est une plage d'adresses IP publiquement routable, réfléchissez aux implications de sécurité avant de lier une instance EC2-Classic à votre VPC.

Par exemple, si votre instance EC2-Classic liée subit une attaque par saturation (DoS, Denial of Service) à partir d'une adresse IP source faisant partie de la plage d'adresses IP du VPC, le trafic de réponses est envoyé dans votre VPC. Nous vous recommandons vivement de créer votre VPC à l'aide d'une plage d'adresses IP privées, comme spécifié dans la RFC 1918.

Pour plus d'informations sur la création de tables de routage et le routage dans votre VPC, consultez Tables de routage dans le Amazon VPC Guide de l'utilisateur.

Activation d'une connexion d'appairage de VPC pour ClassicLink

Si vous avez une connexion d'appairage de VPC entre deux VPC et qu'une ou plusieurs instances EC2-Classic sont liées à un des VPC ou aux deux via ClassicLink, vous pouvez étendre la connexion d'appairage de VPC pour permettre la communication entre les instances EC2-Classic et les instances VPC de l'autre côté de la connexion d'appairage de VPC. Les instances EC2-Classic et les instances dans le VPC peuvent ainsi communiquer grâce aux adresses IP privées. Pour ce faire, vous pouvez autoriser un VPC local à communiquer avec une instance EC2-Classic liée dans un VPC pair, ou vous pouvez autoriser une instance EC2-Classic locale liée à communiquer avec des instances dans un VPC pair.

Si vous activez un VPC local afin de communiquer avec une instance EC2-Classic liée dans un VPC pair, une route statique est ajoutée automatiquement à vos tables de routage avec la destination `10.0.0.0/8` et la cible `local`.

Pour plus d'informations et d'exemples, consultez la section Configurations avec ClassicLink dans le Amazon VPC Peering Guide.

Limites ClassicLink

Pour utiliser la fonction ClassicLink, vous devez être conscient des restrictions suivantes :

- Vous pouvez lier une instance EC2-Classic à un seul VPC à la fois.
- Si vous arrêtez votre instance EC2-Classic liée, celle-ci est automatiquement détachée du VPC et les groupes de sécurité VPC ne sont plus associés à l'instance. Vous pourrez lier à nouveau l'instance au VPC après l'avoir redémarrée.
- Vous ne pouvez pas lier une instance EC2-Classic à un VPC situé dans une autre région ou un autre compte AWS.
- Vous ne pouvez pas utiliser ClassicLink pour lier une instance VPC à un autre VPC ou à une ressource EC2-Classic. Pour établir une connexion privée entre des VPC, vous pouvez utiliser une connexion d'appairage de VPC. Pour plus d'informations, consultez le Amazon VPC Peering Guide.
- Vous ne pouvez pas associer une adresse IP Elastic de VPC à une instance EC2-Classic liée.
- Vous ne pouvez pas activer les instances EC2-Classic pour la communication IPv6. Vous pouvez associer un bloc d'adresse CIDR IPv6 à votre VPC et assigner l'adresse IPv6 à des ressources de votre VPC. Toutefois, la communication entre une instance ClassicLinked et des ressources du VPC a lieu via IPv4 uniquement.
- Les VPC avec des routes qui entrent en conflit avec la plage d'adresses IP privées EC2-Classic `10/8` ne peuvent pas être activés pour ClassicLink. Ceci n'inclut pas les VPC avec les plages d'adresses IP privées `10.0.0.0/16` et `10.1.0.0/16` dont les tables de routage comportent déjà des routes locales. Pour plus d'informations, consultez Routage pour ClassicLink (p. 951).
- Les VPC configurés pour une location matérielle dédiée ne peuvent pas être activés pour ClassicLink. Contactez AWS Support pour demander que votre VPC de location dédiée soit autorisé à être activé pour ClassicLink.

  Important

  Les instances EC2-Classic s'exécutent sur un matériel partagé. Si vous définissez la location de votre VPC sur `dedicated` en raison d'exigences réglementaires ou de sécurité, lier une instance EC2-Classic à votre VPC peut ne pas respecter ces exigences, car cela permet à

une ressource de location partagée d'accéder à vos ressources isolées directement à l'aide d'adresses IP privées. Si vous devez activer votre VPC dédié pour ClassicLink, fournissez la raison détaillée dans votre demande à AWS Support.

- Si vous liez votre instance EC2-Classic à un VPC de la plage `172.16.0.0/16` et qu'un serveur DNS s'exécute sur l'adresse IP `172.16.0.23/32` dans le VPC, votre instance EC2-Classic liée ne peut pas accéder au serveur DNS du VPC. Pour contourner ce problème, exécutez un serveur DNS sur une autre adresse IP au sein du VPC.
- ClassicLink ne prend pas en charge les relations transitives en dehors du VPC. Votre instance EC2-Classic liée n'a accès à aucune connexion VPN, aucun point de terminaison de passerelle VPC, aucune passerelle NAT, ni aucune passerelle Internet associée au VPC. De même, les ressources situées de l'autre côté d'une connexion VPN ou d'une passerelle Internet n'ont accès à aucune instance EC2-Classic liée.

Utilisation de ClassicLink

Vous pouvez utiliser les consoles Amazon EC2 et Amazon VPC pour vous servir des fonctions de ClassicLink. Vous pouvez activer ou désactiver un VPC pour ClassicLink, et lier des instances EC2-Classic à un VPC et les détacher.

> **Note**
>
> Les fonctions de ClassicLink sont visibles uniquement sur les consoles pour les comptes et les régions prenant en charge EC2-Classic.

Tâches

- Activation d'un VPC pour ClassicLink (p. 953)
- Création d'un VPC avec ClassicLink activé (p. 954)
- Liaison d'une instance à un VPC (p. 954)
- Liaison d'une instance à un VPC au lancement (p. 954)
- Affichage de vos VPC activés pour ClassicLink et de vos instances liées (p. 955)
- Activation de la prise en charge de DNS ClassicLink (p. 955)
- Désactivation de la prise en charge de DNS ClassicLink (p. 955)
- Suppression du lien d'une instance à un VPC (p. 956)
- Désactivation de ClassicLink pour un VPC (p. 956)

Activation d'un VPC pour ClassicLink

Pour lier une instance EC2-Classic à un VPC, vous devez activer le VPC pour ClassicLink. Vous ne pouvez pas activer un VPC pour ClassicLink si le VPC comporte un routage entrant en conflit avec la plage d'adresses IP privées EC2-Classic. Pour plus d'informations, consultez Routage pour ClassicLink (p. 951).

Pour activer un VPC pour ClassicLink

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, sélectionnez Vos VPC.
3. Sélectionnez le VPC.
4. Choisissez Actions, Activer ClassicLink.
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Activer ClassicLink.
6. (Facultatif) Si vous souhaitez que le nom d'hôte DNS public soit résolu en adresse IP privée, activez la prise en charge de DNS ClassicLink pour le VPC avant de lier des instances. Pour plus d'informations, consultez Activation de la prise en charge de DNS ClassicLink (p. 955).

Création d'un VPC avec ClassicLink activé

Vous pouvez créer un VPC et l'activer immédiatement pour ClassicLink à l'aide de l'assistant VPC sur la console Amazon VPC.

Pour créer un VPC avec ClassicLink activé

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Sur le tableau de bord Amazon VPC, choisissez Lancer l'assistant VPC.
3. Sélectionnez l'une des options de configuration de VPC, puis choisissez Sélectionner.
4. Sur la page suivante de l'assistant, sélectionnez Oui pour Activer ClassicLink. Complétez les autres étapes de l'assistant pour créer votre VPC. Pour plus d'informations sur l'assistant VPC, consultez la section relative aux scénarios pour Amazon VPC dans le Amazon VPC Guide de l'utilisateur.
5. (Facultatif) Si vous souhaitez que le nom d'hôte DNS public soit résolu en adresse IP privée, activez la prise en charge de DNS ClassicLink pour le VPC avant de lier des instances. Pour plus d'informations, consultez Activation de la prise en charge de DNS ClassicLink (p. 955).

Liaison d'une instance à un VPC

Une fois que vous avez activé un VPC pour ClassicLink, vous pouvez lier une instance EC2-Classic à celui-ci. L'instance doit être dans l'état `running`.

Si vous souhaitez que le nom d'hôte DNS public soit résolu en adresse IP privée, activez la prise en charge de DNS ClassicLink pour le VPC avant de lier l'instance. Pour plus d'informations, consultez Activation de la prise en charge de DNS ClassicLink (p. 955).

Pour lier une instance à un VPC

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une ou plusieurs instances EC2-Classic en cours d'exécution.
4. Choisissez Actions, ClassicLink, Lier au VPC.
5. Choisissez le VPC. La console affiche uniquement les VPC qui sont activés pour ClassicLink.
6. Sélectionnez un ou plusieurs groupes de sécurité VPC à associer à vos instances. La console affiche uniquement les groupes de sécurité pour les VPC activés pour ClassicLink.
7. Choisissez Lier.

Liaison d'une instance à un VPC au lancement

Vous pouvez utiliser l'assistant de lancement sur la console Amazon EC2 pour lancer une instance EC2-Classic et la lier immédiatement à un VPC activé pour ClassicLink.

Pour lier une instance à un VPC au lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sur le tableau de bord Amazon EC2, sélectionnez Launch Instance (Lancer une instance).
3. Sélectionnez une AMI, puis choisissez un type d'instance pris en charge sur EC2-Classic. Pour plus d'informations, consultez Types d'instances disponibles dans EC2-Classic (p. 942).
4. Sur la page Configurer les détails de l'instance, procédez de la façon suivante :

   a. Pour Réseau, choisissez Lancer dans EC2-Classic. Si cette option est désactivée, le type d'instance n'est pas pris en charge sur EC2-Classic.

   b. Développez Lien vers VPC (ClassicLink) et choisissez un VPC dans Lien vers un VPC. La console affiche uniquement les VPC avec ClassicLink activé.

5. Terminez les étapes restantes de l'assistant, puis lancez votre instance. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Affichage de vos VPC activés pour ClassicLink et de vos instances liées

Vous pouvez afficher tous vos VPC activés pour ClassicLink sur la console Amazon VPC, ainsi que la totalité de vos instances EC2-Classic liées sur la console Amazon EC2.

Pour afficher vos VPC activés pour ClassicLink

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, sélectionnez Vos VPC.
3. Sélectionnez le VPC.
4. Si la valeur de ClassicLink est Activé, le VPC est activé pour ClassicLink.

Activation de la prise en charge de DNS ClassicLink

Vous pouvez activer la prise en charge de DNS ClassicLink pour votre VPC afin que les noms d'hôte DNS qui sont adressés entre les instances EC2-Classic et les instances du VPC soient résolus en adresses IP privées et non en adresses IP publiques. Pour que cette fonction fonctionne, votre VPC doit être activé pour les noms d'hôte DNS et la résolution DNS.

> **Note**
>
> Si vous activez la prise en charge de DNS ClassicLink pour votre VPC, votre instance EC2-Classic liée peut accéder à n'importe quelle zone hébergée privée associée au VPC. Pour plus d'informations, consultez Utilisation des zones hébergées privées dans le Amazon Route 53 Manuel du développeur.

Pour activer la prise en charge de DNS ClassicLink

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, sélectionnez Vos VPC.
3. Sélectionnez le VPC.
4. Choisissez Actions, Modifier la prise en charge ClassicLink DNS.
5. Pour Prise en chargeClassicLink DNS, sélectionnez Activer.
6. Sélectionnez Save Changes.

Désactivation de la prise en charge de DNS ClassicLink

Vous pouvez désactiver la prise en charge de ClassicLink DNS pour votre VPC afin que les noms d'hôte DNS qui sont adressés entre les instances EC2-Classic et les instances du VPC soient résolus en adresses IP publiques et non en adresses IP privées.

Pour désactiver la prise en charge de DNS ClassicLink

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, sélectionnez Vos VPC.
3. Sélectionnez le VPC.
4. Choisissez Actions, Modifier la prise en charge ClassicLink DNS.
5. Pour Prise en chargeClassicLink DNS, désélectionnez Activer.
6. Sélectionnez Save Changes.

Suppression du lien d'une instance à un VPC

Si vous n'avez plus besoin d'une connexion ClassicLink entre votre instance EC2-Classic et votre VPC, vous pouvez supprimer les liens de l'instance du VPC. La suppression du lien de l'instance dissocie les groupes de sécurité du VPC de l'instance.

Une instance arrêtée est automatiquement détachée d'un VPC.

Pour détacher une instance d'un VPC

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une ou plusieurs de vos instances.
4. Choisissez Actions, ClassicLink, Annuler le lien du VPC.
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Annuler le lien.

Désactivation de ClassicLink pour un VPC

Si vous n'avez plus besoin d'une connexion entre des instances EC2-Classic et votre VPC, vous pouvez désactiver ClassicLink sur le VPC. Vous devez d'abord détacher toutes les instances EC2-Classic qui sont liées au VPC.

Pour désactiver ClassicLink pour un VPC

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, sélectionnez Vos VPC.
3. Sélectionnez votre VPC.
4. Choisissez Actions, Désactiver ClassicLink.
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Désactiver ClassicLink.

Exemples de stratégies IAM pour ClassicLink

Vous pouvez activer un VPC pour ClassicLink, puis lier une instance EC2-Classic au VPC. Vous pouvez aussi afficher vos VPC ClassicLink, ainsi que la totalité de vos instances EC2-Classic liées à un VPC. Vous pouvez créer des stratégies avec des autorisations au niveau des ressources pour les actions `ec2:EnableVpcClassicLink`, `ec2:DisableVpcClassicLink`, `ec2:AttachClassicLinkVpc` et `ec2:DetachClassicLinkVpc` afin de contrôler la façon dont ces utilisateurs peuvent utiliser ces actions. Les autorisations au niveau des ressources ne sont pas prises en charge pour les actions `ec2:Describe*`.

Exemples

- Autorisations complètes à utiliser avec ClassicLink (p. 956)
- Activer et désactiver un VPC pour ClassicLink (p. 957)
- Lier les instances (p. 957)
- Supprimer le lien des instances (p. 958)

Autorisations complètes à utiliser avec ClassicLink

La stratégie suivante accorde aux utilisateurs l'autorisation d'afficher les VPC ClassicLink et les instances EC2-Classic liées, d'activer et de désactiver un VPC pour ClassicLink, et de lier des instances à partir d'un VPC ClassicLink ou d'annuler le lien.

```
{
```

```
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeClassicLinkInstances", "ec2:DescribeVpcClassicLink",
        "ec2:EnableVpcClassicLink", "ec2:DisableVpcClassicLink",
        "ec2:AttachClassicLinkVpc", "ec2:DetachClassicLinkVpc"
      ],
      "Resource": "*"
    }
   ]
}
```

Activer et désactiver un VPC pour ClassicLink

La stratégie suivante permet à l'utilisateur d'activer et de désactiver les VPC pour ClassicLink ayant la balise spécifique "`purpose=classiclink`". Les utilisateurs ne peuvent pas activer ou désactiver d'autres VPC pour ClassicLink.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "ec2:*VpcClassicLink",
      "Resource": "arn:aws:ec2:region:account:vpc/*",
      "Condition": {
        "StringEquals": {
          "ec2:ResourceTag/purpose":"classiclink"
        }
      }
    }
  ]
}
```

Lier les instances

La stratégie suivante autorise les utilisateurs à lier les instances à un VPC uniquement si elles sont de type `m3.large`. La deuxième déclaration permet aux utilisateurs d'utiliser les ressources du VPC et du groupe de ressources, qui sont requises pour lier une instance à un VPC.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "ec2:AttachClassicLinkVpc",
      "Resource": "arn:aws:ec2:region:account:instance/*",
      "Condition": {
        "StringEquals": {
          "ec2:InstanceType":"m3.large"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": "ec2:AttachClassicLinkVpc",
      "Resource": [
        "arn:aws:ec2:region:account:vpc/*",
        "arn:aws:ec2:region:account:security-group/*"
      ]
    }
```

```
  ]
}
```

La stratégie suivante autorise les utilisateurs à lier les instances à un VPC spécifique (`vpc-1a2b3c4d`) uniquement et à n'associer que des groupes spécifiques de sécurité du VPC à l'instance (`sg-1122aabb` et `sg-aabb2233`). Les utilisateurs ne peuvent pas lier une instance à un autre VPC ni spécifier un autre des groupes de sécurité du VPC pour l'associer à l'instance dans la demande.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "ec2:AttachClassicLinkVpc",
      "Resource": [
         "arn:aws:ec2:region:account:vpc/vpc-1a2b3c4d",
         "arn:aws:ec2:region:account:instance/*",
         "arn:aws:ec2:region:account:security-group/sg-1122aabb",
         "arn:aws:ec2:region:account:security-group/sg-aabb2233"
       ]
    }
  ]
}
```

Supprimer le lien des instances

La stratégie suivante accorde aux utilisateurs l'autorisation de supprimer les liens de toute instance EC2-Classic liée à partir d'un VPC, mais uniquement si l'instance a la balise "`unlink=true`". La seconde déclaration autorise les utilisateurs à avoir recours à la ressource du VPC, qui est requise pour détacher une instance du VPC.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": "ec2:DetachClassicLinkVpc",
      "Resource": [
         "arn:aws:ec2:region:account:instance/*"
      ],
      "Condition": {
         "StringEquals": {
            "ec2:ResourceTag/unlink":"true"
         }
      }
   },
   {
      "Effect": "Allow",
      "Action": "ec2:DetachClassicLinkVpc",
      "Resource": [
         "arn:aws:ec2:region:account:vpc/*"
         ]
      }
   ]
}
```

Exemple : Configuration d'un groupe de sécurité ClassicLink pour une application web à trois niveaux

Dans cet exemple, vous disposez d'une application avec trois instances : un serveur web public, un serveur d'applications et un serveur de base de données. Votre serveur web accepte le trafic HTTPS d'Internet

et communique avec votre serveur d'applications via le port TCP 6001. Votre serveur d'applications communique alors avec votre serveur de base de données via le port TCP 6004. Vous êtes en train de migrer la totalité de votre application vers un VPC de votre compte. Vous avez déjà migré votre serveur d'applications et votre serveur de base de données vers votre VPC. Votre serveur web est encore dans EC2-Classic et lié à votre VPC via ClassicLink.

Vous souhaitez une configuration de groupe de sécurité qui autorise la circulation du trafic uniquement entre ces instances. Vous disposez de quatre groupes de sécurité : deux pour votre serveur web (`sg-1a1a1a1a` et `sg-2b2b2b2b`), un pour votre serveur d'applications (`sg-3c3c3c3c`) et un pour votre serveur de base de données (`sg-4d4d4d4d`).

Le schéma suivant montre l'architecture de vos instances et leur configuration de groupe de sécurité.

Groupes de sécurité pour votre serveur web (**`sg-1a1a1a1a`** et **`sg-2b2b2b2b`**)

Vous disposez d'un groupe de sécurité dans EC2-Classic et d'un autre dans votre VPC. Vous avez associé le groupe de sécurité du VPC avec votre instance de serveur web lorsque vous avez lié l'instance à votre VPC via ClassicLink. Le groupe de sécurité du VPC vous permet de contrôler le trafic sortant de votre serveur web vers votre serveur d'applications.

Voici les règles de groupe de sécurité pour le groupe de sécurité EC2-Classic (`sg-1a1a1a1a`).

Inbound			
Source	Type	Port Range	Comments
0.0.0.0/0	HTTPS	443	Autorise le trafic Internet à accéder à votre serveur web.

Voici les règles de groupe de sécurité pour le groupe de sécurité du VPC (`sg-2b2b2b2b`).

Outbound			
Destination	Type	Port Range	Comments
sg-3c3c3c3c	TCP	6001	Autorise le trafic sortant de votre serveur web vers votre serveur d'applications dans votre VPC (ou vers toute autre instance associée à `sg-3c3c3c3c`).

Groupe de sécurité pour votre serveur d'applications (**`sg-3c3c3c3c`**)

Voici les règles de groupe de sécurité pour le groupe de sécurité VPC associé à votre serveur d'applications.

Inbound			
Source	Type	Port Range	Comments
sg-2b2b2b2b	TCP	6001	Autorise le type de trafic spécifié de votre serveur web (ou toute autre instance associée à `sg-2b2b2b2b`) à accéder à votre serveur d'applications.
Outbound			
Destination	Type	Port Range	Comments
sg-4d4d4d4d	TCP	6004	Autorise le trafic sortant du serveur d'applications vers le serveur de base de données (ou vers toute autre instance associée à `sg-4d4d4d4d`).

Groupe de sécurité pour votre serveur de base de données (**`sg-4d4d4d4d`**)

Voici les règles de groupe de sécurité pour le groupe de sécurité VPC associé à votre serveur de base de données.

Inbound			
Source	Type	Port Range	Comments

sg-3c3c3c3c	TCP	6004	Autorise le type de trafic spécifié de votre serveur d'applications (ou toute autre instance associée à `sg-3c3c3c3c`) à accéder à votre serveur de base de données.

Migration d'EC2-Classic vers un VPC

Si vous avez créé votre compte AWS avant le 4 décembre 2013, vous pouvez bénéficier de la prise en charge d'EC2-Classic dans certaines régions AWS. Certaines ressources et fonctions Amazon EC2, comme la mise en réseau améliorée et des types d'instance plus récents, requièrent un cloud privé virtuel (VPC). Certaines ressources peuvent être partagées entre EC2-Classic et un VPC, contrairement à d'autres ressources. Pour plus d'informations, consultez Partager et accéder à des ressources entre EC2-Classic et un VPC (p. 948). Nous vous recommandons de migrer vers un VPC pour tirer parti des fonctionnalités uniquement VPC.

Pour migrer d'EC2-Classic vers un VPC, vous devez migrer ou recréer vos ressources EC2-Classic dans un VPC. Vous pouvez migrer et recréer vos ressources dans leur intégralité, ou effectuer une migration incrémentielle au fil du temps à l'aide de ClassicLink.

Sommaire

- Options pour obtenir un VPC par défaut (p. 961)
- Migrer vos ressources vers un VPC (p. 962)
- Utiliser ClassicLink pour une migration incrémentielle (p. 966)
- Exemple : Migrer une application web simple (p. 968)

Options pour obtenir un VPC par défaut

Un VPC par défaut est un VPC configuré et prêt à l'emploi, et uniquement disponible dans les régions qui sont VPC uniquement. Pour les régions qui prennent en charge EC2-Classic, vous pouvez créer un VPC autre que par défaut pour configurer vos ressources. Toutefois, vous pouvez vouloir utiliser un VPC par défaut si vous préférez ne pas en configurer vous-même ou si vous n'avez pas d'exigences spécifiques pour votre configuration VPC. Pour plus d'informations sur les VPC par défaut, consultez VPC par défaut et sous-réseaux par défaut dans le Amazon VPC Guide de l'utilisateur.

Les options suivantes permettent d'utiliser un VPC par défaut lorsque vous disposez d'un compte AWS qui prend en charge EC2-Classic.

Options

- Basculer vers une région VPC uniquement (p. 961)
- Créer un nouveau compte AWS (p. 961)
- Convertir votre compte AWS existant en VPC uniquement (p. 962)

Basculer vers une région VPC uniquement

Utilisez cette option si vous souhaitez utiliser votre compte existant pour configurer vos ressources dans un VPC par défaut et que vous n'avez pas besoin d'utiliser une région spécifique. Pour rechercher une région dotée d'un VPC par défaut, consultez Suppression des plateformes prises en charge (p. 941).

Créer un nouveau compte AWS

Les nouveaux comptes AWS prennent en charge le VPC uniquement. Utilisez cette option si vous souhaitez un compte doté d'un VPC par défaut dans chaque région.

Convertir votre compte AWS existant en VPC uniquement

Utilisez cette option si vous souhaitez un VPC par défaut dans chaque région pour votre compte existant. Avant de convertir votre compte, vous devez supprimer toutes vos ressources EC2-Classic. Vous pouvez également migrer certaines ressources vers un VPC. Pour de plus amples informations, veuillez consulter Migrer vos ressources vers un VPC (p. 962).

Pour convertir votre compte EC2-Classic

1. Supprimez ou migrez (le cas échéant) les ressources que vous avez créées pour être utilisées dans EC2-Classic. Tel est le cas des éléments suivants :
   - Instances Amazon EC2
   - Groupes de sécurité EC2-Classic (à l'exclusion du groupe de sécurité par défaut, que vous ne pouvez pas supprimer vous-même)
   - Adresses IP Elastic EC2-Classic
   - Equilibreurs de charge classiques
   - Ressources Amazon RDS
   - Ressources Amazon ElastiCache
   - Ressources Amazon Redshift
   - Ressources AWS Elastic Beanstalk
   - Ressources AWS Data Pipeline
   - Ressources Amazon EMR
   - Ressources AWS OpsWorks
2. Accédez au Centre de support AWS à l'adresse console.aws.amazon.com/support.
3. Choisissez Create case (Créer une demande).
4. Choisissez Support de compte et facture.
5. Pour Type, choisissez Compte. Pour Catégorie, choisissez Convertir EC2 Classic en VPC.
6. Remplissez les autres détails selon vos besoins, puis choisissez Soumettre. Nous examinerons votre demande et vous contacterons pour vous guider dans les prochaines étapes.

Migrer vos ressources vers un VPC

Vous pouvez migrer ou déplacer certaines de vos ressources vers un VPC. Certaines ressources peuvent uniquement être migrées d'EC2-Classic à partir d'un VPC qui se trouve dans la même région et dans le même compte AWS. Si la ressource ne peut pas être migrée, vous devez créer une ressource à utiliser dans votre VPC.

Prérequis

Avant de commencer, vous devez disposer d'un VPC. Si vous n'avez pas de VPC par défaut, vous pouvez créer un VPC autre que par défaut en utilisant l'une des méthodes suivantes :

- Dans la console Amazon VPC, utilisez l'assistant VPC pour créer un nouveau VPC. Pour plus d'informations, consultez Configurations de l'Assistant de la console Amazon VPC. Utilisez cette option si vous souhaitez configurer rapidement un VPC à l'aide de l'une des options de configuration disponibles.
- Dans la console Amazon VPC, configurez les composants d'un VPC en fonction de vos besoins. Pour plus d'informations, consultez VPC et sous-réseaux. Utilisez cette option si vous avez des exigences spécifiques pour votre VPC, par exemple, un nombre particulier de sous-réseaux.

Rubriques

- Groupes de sécurité (p. 963)

- Adresses IP Elastic (p. 963)
- AMI et instances (p. 964)
- Instances de base de données Amazon RDS (p. 966)

Groupes de sécurité

Si vous souhaitez que les instances de votre VPC aient les mêmes règles de groupe de sécurité que vos instances EC2-Classic, vous pouvez utiliser la console Amazon EC2 pour copier vos règles de groupe de sécurité EC2-Classic existantes dans un nouveau groupe de sécurité VPC.

Vous pouvez uniquement copier des règles de groupe de sécurité vers un nouveau groupe de sécurité VPC du même compte AWS dans la même région. Si vous utilisez une autre région ou un compte AWS différent, vous devez créer un nouveau groupe de sécurité et ajouter manuellement les règles vous-même. Pour de plus amples informations, veuillez consulter Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062).

Pour copier vos règles de groupe de sécurité dans un nouveau groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Groupes de sécurité.
3. Sélectionnez le groupe de sécurité associé à votre instance EC2-Classic et choisissez Actions, puis Copier vers le nouveau.

   Note

   Pour identifier un groupe de sécurité EC2-Classic, vérifiez la colonne ID du VPC. Pour chaque groupe de sécurité EC2-Classic, la valeur de la colonne est vide ou contient un symbole `-`.
4. Dans la boîte de dialogue Créer un groupe de sécurité, indiquez un nom et une description pour votre nouveau groupe de sécurité. Sélectionnez votre VPC dans la liste VPC.
5. L'onglet Entrant est alimenté avec les règles de votre groupe de sécurité EC2-Classic. Vous pouvez modifier les règles si nécessaire. Dans l'onglet Sortant, une règle qui autorise tout le trafic sortant a été créée automatiquement pour vous. Pour plus d'informations sur la modification de règles de groupe de sécurité, consultez Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062).

   Note

   Si vous avez défini une règle dans votre groupe de sécurité EC2-Classic faisant référence à un autre groupe de sécurité, vous ne pouvez pas utiliser la même règle dans votre groupe de sécurité VPC. Modifiez la règle pour qu'elle fasse référence à un autre groupe de sécurité du même VPC.
6. Sélectionnez Create.

Adresses IP Elastic

Vous pouvez migrer une adresse IP Elastic qui est allouée pour une utilisation dans EC2-Classic pour l'utiliser avec un VPC. Vous ne pouvez pas migrer une adresse IP Elastic vers une autre région ou un compte AWS différent. Pour plus d'informations, consultez Migration d'une Adresse IP Elastic à partir de EC2-Classic (p. 946).

Pour identifier une adresse IP Elastic qui est allouée pour une utilisation dans EC2-Classic

Dans la console Amazon EC2, choisissez Elastic IPs dans le volet de navigation. Dans la colonne Portée la valeur est standard.

Vous pouvez également utiliser la commande `describe-addresses` suivante.

```
aws ec2 describe-addresses --filters Name=domain,Values=standard
```

AMI et instances

Une AMI est un template pour le lancement de votre instance Amazon EC2. Vous pouvez créer votre propre AMI basée sur une instance EC2-Classic existante, puis utiliser cette AMI pour lancer des instances dans votre VPC.

Sommaire

- Identification d'instances EC2-Classic (p. 964)
- Créer une AMI (p. 964)
- (Facultatif) Partagez ou copiez votre AMI (p. 965)
- (Facultatif) Stockage de vos données sur des volumes Amazon EBS (p. 965)
- Lancer une instance dans votre VPC (p. 965)

Identification d'instances EC2-Classic

Si vous avez des instances en cours d'exécution dans EC2-Classic et dans un VPC, vous pouvez identifier vos instances EC2-Classic.

Console Amazon EC2

Choisissez Instances dans le volet de navigation. Dans la colonne ID du VPC la valeur de chaque instance EC2-Classic est vide ou contient un symbole `-`. Si la colonne ID du VPC n'est pas présente, choisissez l'icône d'engrenage et rendez la colonne visible.

AWS CLI

Utilisez la commande AWS CLI describe-instances suivante. Le paramètre `--query` affiche uniquement les instances où la valeur pour `VpcId` est `null`.

```
aws ec2 describe-instances --query 'Reservations[*].Instances[?VpcId==`null`]'
```

Créer une AMI

Après avoir identifié votre instance EC2-Classic, vous pouvez créer une AMI à partir de celle-ci.

Pour créer une AMI Windows

Pour plus d'informations, consultez Création d'une AMI Windows personnalisée.

Pour créer une AMI Linux

La méthode que vous utilisez pour créer une AMI Linux dépend du type de périphérique racine de votre instance et de la plateforme de système d'exploitation sur laquelle votre instance s'exécute. Pour identifier le type de périphérique racine de votre instance, accédez à la page Instances, sélectionnez votre instance et examinez les informations du champ Type de périphérique racine dans l'onglet Description. Si la valeur est `ebs`, votre instance est basée sur EBS. Si la valeur est `instance-store`, votre instance est basée sur le stockage d'instance. Vous pouvez également utiliser la commande describe-instances AWS CLI pour déterminer le type de périphérique racine.

Le tableau suivant fournit des options que vous pouvez utiliser pour créer une AMI Linux en fonction du type de périphérique racine de votre instance et de la plateforme logicielle.

> **Important**
>
> Certains types d'instance prennent en charge la virtualisation PV et la virtualisation HVM, alors que d'autres prennent en charge seulement l'une ou l'autre de ces techniques. Si vous prévoyez d'utiliser votre AMI pour lancer un autre type d'instance que celui en cours, vérifiez que le type

d'instance prend en charge le type de virtualisation fourni par votre AMI. Si votre AMI prend en charge la virtualisation PV et que vous souhaitez utiliser un type d'instance prenant en charge la virtualisation HVM, vous devrez peut-être réinstaller votre logiciel sur une AMI HVM de base. Pour plus d'informations sur la virtualisation PV et HVM, consultez Types de virtualisation d'AMI Linux.

Type de périphérique racine de l'instance	Action
EBS	Créez une AMI basée sur EBS à partir de votre instance. Pour en savoir plus, consultez Création d'une AMI Linux basée sur Amazon EBS.
Stockage d'instance	Créez une AMI basée sur le stockage d'instance à partir de votre instance à l'aide des outils AMI. Pour plus d'informations, consultez Création d'une AMI Linux basée sur le stockage d'instance.
Stockage d'instance	Convertissez votre instance basée sur le stockage d'instance en instance basée sur EBS. Pour plus d'informations, consultez Conversion de votre AMI basée sur le stockage d'instance en AMI basée sur Amazon EBS.

(Facultatif) Partagez ou copiez votre AMI

Pour utiliser votre AMI pour lancer une instance dans un nouveau compte AWS, vous devez d'abord partager l'AMI avec votre nouveau compte. Pour de plus amples informations, veuillez consulter Partager une AMI avec des comptes AWS spécifiques (p. 121).

Pour utiliser votre AMI pour lancer une instance dans un VPC dans une autre région, vous devez d'abord copier l'AMI dans cette région. Pour plus d'informations, consultez Copie une AMI (p. 173).

(Facultatif) Stockage de vos données sur des volumes Amazon EBS

Vous pouvez créer un volume Amazon EBS et l'utiliser pour sauvegarder et stocker les données sur une instance, tout comme vous utiliseriez un disque dur physique. Les volumes Amazon EBS peuvent être attachés à et détachés de toute instance de la même zone de disponibilité. Vous pouvez détacher un volume de votre instance dans EC2-Classic et l'attacher à une nouvelle instance que vous lancez dans votre VPC au sein de la même zone de disponibilité.

Pour plus d'informations sur les volumes Amazon EBS, consultez les rubriques suivantes :

- Volumes Amazon EBS (p. 1086)
- Création d'un volume Amazon EBS (p. 1107)
- Attacher un volume Amazon EBS à une instance (p. 1109)

Pour sauvegarder les données sur votre volume Amazon EBS, vous pouvez créer des instantanés périodiques de votre volume. Pour de plus amples informations, veuillez consulter Création d'instantanés Amazon EBS (p. 1132). Si nécessaire, vous pouvez restaurer un volume Amazon EBS à partir de votre instantané. Pour de plus amples informations, veuillez consulter Création d'un volume à partir d'un instantané (p. 1108).

Lancer une instance dans votre VPC

Après avoir créé une AMI, vous pouvez utiliser l'assistant de lancement Amazon EC2 pour lancer une instance dans votre VPC. L'instance aura les mêmes données et configurations que votre instance EC2-Classic existante.

> **Note**
>
> Vous pouvez utiliser cette possibilité pour effectuer une mise à niveau vers un type d'instance de génération actuel. Toutefois, vérifiez que le type d'instance prend en charge le type de

virtualisation que votre AMI propose (PV ou HVM). Pour plus d'informations sur la virtualisation PV et HVM, consultez Types de virtualisation AMI Linux (p. 108).

Pour lancer une instance dans votre VPC

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sur le tableau de bord, choisissez Lancer une instance.
3. Sur la page Choose an Amazon Machine Image (Sélection d'une Amazon Machine Image), sélectionnez la catégorie Mes AMI, puis l'AMI que vous avez créée. Para ailleurs, si vous avez partagé une AMI à partir d'un autre compte, dans la liste de filtre Propriétaire choisissez Partagée avec moi. Sélectionnez l'AMI que vous avez partagée à partir de votre compte EC2-Classic.
4. Sur la page Choisir un type d'instance, sélectionnez le type d'instance, puis Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
5. Sur la page Configurer les détails de l'instance, sélectionnez votre VPC dans la liste Réseau. Sélectionnez le sous-réseau requis dans la liste Sous-réseau. Configurez tous les autres détails nécessaires, puis parcourez les pages suivantes de l'assistant jusqu'à la page Configurer le groupe de sécurité.
6. Choisissez Sélectionner un groupe existant, puis le groupe de sécurité que vous avez créé pour votre VPC. Choisissez Vérifier et lancer.
7. Vérifiez les détails de votre instance, puis sélectionnez Lancer pour spécifier une paire de clés et lancer votre instance.

Pour plus d'informations sur les paramètres que vous pouvez configurer à chaque étape de l'assistant, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).

Instances de base de données Amazon RDS

Vous pouvez déplacer votre instance de base de données EC2-Classic vers un VPC dans la même région, dans le même compte. Pour plus d'informations, consultez Mise à jour du VPC pour une instance de base de données dans le Amazon RDS Guide de l'utilisateur.

Utiliser ClassicLink pour une migration incrémentielle

La fonction ClassicLink facilite la gestion d'une migration incrémentielle vers un VPC. ClassicLink vous permet de lier une instance EC2-Classic à un VPC dans votre compte de la même région. Vos nouvelles ressources VPC peuvent ainsi communiquer avec l'instance EC2-Classic à l'aide d'adresses IPv4 privées. Cela vous permet de transférer les fonctionnalités composant par composant jusqu'à ce que votre application s'exécute intégralement dans votre VPC.

Utilisez cette option si vous ne pouvez pas vous permettre des temps d'arrêt pendant la migration, par exemple si vous avez une application multi-niveaux avec des processus qui ne peuvent pas être interrompus.

Pour plus d'informations sur ClassicLink, consultez ClassicLink (p. 949).

Tâches

- Étape 1 : Préparation de votre séquence de migration (p. 966)
- Étape 2 : Activation de votre VPC pour ClassicLink (p. 967)
- Étape 3 : Liaison de vos instances EC2-Classic à votre VPC (p. 967)
- Étape 4 : Exécution de la migration vers le VPC (p. 967)

Étape 1 : Préparation de votre séquence de migration

Pour utiliser efficacement ClassicLink, vous devez d'abord identifier les composants de votre application qui doivent être migrés vers le VPC et confirmer l'ordre de migration des fonctionnalités correspondantes.

Par exemple, vous disposez d'une application reposant sur un serveur web de présentation, d'un serveur de base de données principal et d'une logique d'authentification pour les transactions. Vous pouvez décider de démarrer le processus de migration avec la logique d'authentification, puis le serveur de base de données et enfin, le serveur web.

Ensuite, vous pouvez commencer à migrer ou à recréer vos ressources. Pour plus d'informations, consultez Migrer vos ressources vers un VPC (p. 962).

Étape 2 : Activation de votre VPC pour ClassicLink

Une fois que vous avez configuré les instances de votre nouveau VPC et rendu les fonctionnalités de votre application disponibles dans le VPC, vous pouvez utiliser ClassicLink pour activer la communication IP privée entre vos nouvelles instances VPC et vos instances EC2-Classic. Tout d'abord, vous devez activer votre VPC pour ClassicLink.

Pour activer un VPC pour ClassicLink

1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
2. Dans le panneau de navigation, sélectionnez Vos VPC.
3. Sélectionnez un VPC.
4. Choisissez Actions, Activer ClassicLink.
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Activer ClassicLink.

Étape 3 : Liaison de vos instances EC2-Classic à votre VPC

Une fois que vous avez activé ClassicLink dans votre VPC, vous pouvez lier vos instances EC2-Classic au VPC. L'instance doit être dans l'état `running`.

Pour lier une instance à un VPC

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une ou plusieurs instances EC2-Classic en cours d'exécution.
4. Choisissez Actions, ClassicLink, Lier au VPC.
5. Choisissez un VPC. La console affiche uniquement les VPC qui sont activés pour ClassicLink.
6. Sélectionnez un ou plusieurs groupes de sécurité VPC à associer à vos instances. La console affiche uniquement les groupes de sécurité pour les VPC activés pour ClassicLink.
7. Choisissez Lier.

Étape 4 : Exécution de la migration vers le VPC

En fonction de la taille de votre application et des fonctionnalités devant être migrées, répétez les étapes précédentes jusqu'à avoir déplacé tous les composants de votre application depuis EC2-Classic vers votre VPC.

Une fois que vous avez activé la communication interne entre vos instances EC2-Classic et VPC, vous devez mettre à jour votre application pour qu'elle pointe vers votre service migré dans le VPC au lieu du service sur la plateforme EC2-Classic. Les étapes précises de cette opération dépendent de la conception de votre application. En général, cela inclut la mise à jour de vos adresses IP de destination pour pointer vers les adresses IP de vos instances VPC au lieu de vos instances EC2-Classic.

Une fois que vous avez terminé cette étape et vérifié que l'application fonctionne à partir de votre VPC, vous pouvez résilier vos instances EC2-Classic et désactiver ClassicLink pour votre VPC. Vous pouvez également nettoyer toutes les ressources EC2-Classic dont vous n'avez plus besoin, pour éviter d'encourir

des frais pour ces dernières. Par exemple, vous pouvez libérer les adresses IP Elastic et supprimer les volumes associés à vos instances EC2-Classic.

Exemple : Migrer une application web simple

Dans cet exemple, vous utilisez AWS pour héberger votre site web de jardinage. Pour gérer votre site web, vous disposez de trois instances en cours d'exécution dans EC2-Classic. Les instances A et B hébergent votre application web destinée au public et vous utilisez Elastic Load Balancing pour équilibrer la charge du trafic entre ces instances. Vous avez attribué des adresses IP Elastic aux instances A et B afin de disposer d'adresses IP statiques pour les tâches de configuration et d'administration sur ces instances. L'instance C contient votre base de données MySQL pour votre site web. Vous avez enregistré le nom de domaine `www.garden.example.com` et vous avez utilisé Route 53 pour créer une zone hébergée avec un ensemble d'enregistrements d'alias associé au nom DNS de votre équilibreur de charge.

La première partie de la migration vers un VPC consiste à choisir un type d'architecture VPC adapté à vos besoins. Dans le cas présent, vous avez choisi l'architecture suivante : un sous-réseau public pour vos serveurs web et un sous-réseau privé pour votre serveur de base de données. Au fur et à mesure que votre site web se développe, vous pouvez ajouter des serveurs web et des serveurs de base de données supplémentaires à vos sous-réseaux. Par défaut, les instances du sous-réseau privé ne peuvent pas accéder à Internet ; cependant, vous pouvez activer un accès Internet par le biais d'un périphérique NAT (Network Address Translation, traduction d'adresses réseau) dans le sous-réseau public. Vous pouvez souhaiter configurer un périphérique NAT pour prendre en charge les mises à jour et les correctifs (patches) périodiques issus d'Internet pour votre serveur de base de données. Vous migrerez vos adresses IP Elastic vers un VPC et créerez un équilibreur de charge dans votre sous-réseau public pour équilibrer la charge du trafic entre vos serveurs web.

Pour migrer votre application web vers un VPC, vous pouvez suivre ces étapes :

- Créer un VPC : dans le cas présent, vous pouvez créer l'assistant VPC sur la console Amazon VPC pour créer votre VPC et vos sous-réseaux. La configuration du deuxième assistant crée un VPC avec un sous-réseau privé et un sous-réseau public, et lance et configure pour vous un périphérique NAT dans le sous-réseau public. Pour plus d'informations, consultez VPC avec des sous-réseaux publics et privés (NAT) dans le Amazon VPC Guide de l'utilisateur.
- Configurer vos groupes de sécurité : dans votre environnement EC2-Classic, vous disposez d'un groupe de sécurité pour vos serveurs web et d'un autre groupe de sécurité pour votre serveur de base de données. Vous pouvez utiliser la console Amazon EC2 pour copier les règles de chaque groupe de sécurité dans de nouveaux groupes de sécurité pour votre VPC. Pour plus d'informations, consultez Groupes de sécurité (p. 963).

  Tip

  Créez d'abord les groupes de sécurité qui sont référencés par d'autres groupes de sécurité.
- Créer des AMI et lancer de nouvelles instances : créez une AMI à partir de l'un de vos serveurs web et une deuxième AMI à partir de votre serveur de base de données. Lancez ensuite des serveurs web de remplacement dans votre sous-réseau public et lancez votre serveur de base de données de remplacement dans votre sous-réseau privé. Pour plus d'informations, consultez Créer une AMI (p. 964).
- Configurer votre périphérique NAT : si vous utilisez une instance NAT, vous devez créer pour celle-ci un groupe de sécurité autorisant le trafic HTTP et HTTPS à partir de votre sous-réseau privé. Pour plus

d'informations, veuillez consulter Instances NAT. Si vous utilisez une passerelle NAT, le trafic issu de votre sous-réseau privé est autorisé automatiquement.

- Configurer votre base de données : lorsque vous avez créé une AMI à partir de votre serveur de base de données dans EC2-Classic, toutes les informations de configuration qui étaient stockées dans cette instance ont été copiées dans l'AMI. Vous devrez peut-être vous connecter à votre nouveau serveur de base de données et mettre à jour les détails de configuration. Par exemple, si vous avez configuré votre base de données pour accorder des autorisations complètes de lecture, d'écriture et de modification à vos serveurs web dans EC2-Classic, vous devez mettre à jour les fichiers de configuration pour octroyer les mêmes autorisations à vos nouveaux serveurs web VPC.
- Configurer vos serveurs web : vos serveurs web auront les mêmes paramètres de configuration que vos instances dans EC2-Classic. Par exemple, si vous avez configuré vos serveurs web pour utiliser la base de données dans EC2-Classic, mettez à jour les paramètres de configuration de vos serveurs web pour pointer vers votre nouvelle instance de base de données.

  Note

  Par défaut, les instances lancées dans un sous-réseau autre que celui par défaut (personnalisé) ne se voient pas affecter une adresse IP publique, sauf si vous spécifiez une autre option au moment du lancement. Votre nouveau serveur de base de données peut ne pas avoir une adresse IP publique. Dans ce cas, vous pouvez mettre à jour le fichier de configuration de vos serveurs web pour utiliser le nom DNS privé de votre nouveau serveur de base de données. Les instances du même VPC peuvent communiquer ensemble via l'adresse IP privée.
- Migrer vos adresses IP Elastic : dissociez vos adresses IP Elastic de vos serveurs web dans EC2-Classic, puis migrez-les vers un VPC. Après les avoir migrées, vous pouvez les associer à vos nouveaux serveurs web dans votre VPC. Pour plus d'informations, consultez Migration d'une Adresse IP Elastic à partir de EC2-Classic (p. 946).
- Créer un nouvel équilibreur de charge : pour continuer à utiliser Elastic Load Balancing afin d'équilibrer la charge du trafic sur vos instances, vous devez bien comprendre les différentes façons dont vous pouvez configurer votre équilibreur de charge dans VPC. Pour de plus amples informations, veuillez consulter Elastic Load Balancing Guide de l'utilisateur.
- Mettre à jour vos enregistrements DNS : après avoir configuré votre équilibreur de charge dans votre sous-réseau public, vérifiez que votre domaine `www.garden.example.com` pointe vers votre nouveau équilibreur de charge. Pour ce faire, mettez à jour vos enregistrements DNS et votre ensemble d'enregistrements alias dans Route 53. Pour plus d'informations sur l'utilisation de Route 53, consultez Démarrez avec Route 53.
- Fermer vos ressources EC2-Classic : après avoir vérifié que votre application web fonctionne à partir de l'architecture VPC, vous pouvez fermer vos ressources EC2-Classic pour que celles-ci ne vous soient plus facturées.

Sécurité dans Amazon EC2

Chez AWS, la sécurité dans le cloud est notre priorité numéro 1. En tant que client AWS, vous bénéficiez d'un centre de données et d'une architecture réseau conçus pour répondre aux exigences des organisations les plus pointilleuses en termes de sécurité.

La sécurité est une responsabilité partagée entre AWS et vous-même. Le modèle de responsabilité partagéede décrit ceci comme la sécurité du cloud et la sécurité dans le cloud :

- La sécurité du cloud – AWS est responsable de la protection de l'infrastructure qui exécute des services AWS dans le cloud AWS. AWS vous fournit également les 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 conformité AWS. Pour en savoir plus sur les programmes de conformité qui s'appliquent à Amazon EC2, veuillez consulter Services AWS concernés par le programme de conformité.
- Sécurité dans le cloud – Votre responsabilité est déterminée par le service AWS que vous utilisez. Vous êtes également responsable d'autres facteurs, y compris la sensibilité de vos données, les exigences de votre entreprise,et la législation et la réglementation applicables.

Cette documentation vous aide à comprendre comment appliquer le modèle de responsabilité partagée lors de l'utilisation de Amazon EC2. Elle vous montre comment configurer Amazon EC2 pour atteindre vos objectifs en matière de sécurité et de conformité. Vous pouvez également apprendre à utiliser d'autres services AWS qui vous aident à surveiller et sécuriser vos ressources Amazon EC2.

Sommaire

- Sécurité de l'infrastructure dans Amazon EC2 (p. 971)
- Amazon EC2 et points de terminaison d'un VPC d'interface (p. 973)
- Résilience dans Amazon EC2 (p. 974)
- Protection des données dans Amazon EC2 (p. 975)
- Identity and Access Management pour Amazon EC2 (p. 976)
- Paires de clés Amazon EC2 et instances Linux (p. 1047)
- Groupes de sécurité Amazon EC2 pour les instances Linux (p. 1062)
- Gestion des mises à jour dans Amazon EC2 (p. 1081)
- Validation de la conformité pour Amazon EC2 (p. 1082)

Sécurité de l'infrastructure dans Amazon EC2

En tant que service géré, Amazon EC2 est protégé par les procédures de sécurité du réseau mondial AWS qui sont décrites dans le livre blanc Amazon Web Services : Présentation des procédures de sécurité.

Vous utilisez les appels d'API publiés par AWS pour accéder à Amazon EC2 via le réseau. Les clients doivent prendre en charge le protocole TLS (Transport Layer Security) 1.0 ou version ultérieure. Nous recommandons TLS 1.2 ou version ultérieure. Les clients doivent également prendre en charge les suites de chiffrement PFS (Perfect Forward Secrecy) comme Ephemeral Diffie-Hellman (DHE) ou Elliptic Curve

Ephemeral Diffie-Hellman (ECDHE) La plupart des systèmes modernes telles que Java 7 et versions ultérieures prennent en charge ces modes.

En outre, les demandes doivent être signées à l'aide d'un ID de clé d'accès et d'une clé d'accès secrète associée à un mandataire IAM. Vous pouvez également utiliser AWS Security Token Service (AWS STS) pour générer des informations d'identification de sécurité temporaires et signer les demandes.

Isolement de réseau

Un Virtual Private Cloud (VPC) est un réseau virtuel situé dans votre propre zone isolée logiquement dans le Cloud AWS. Utilisez des VPC distincts pour isoler l'infrastructure par charge de travail ou entité organisationnelle.

Un sous-réseau est une plage d'adresses IP dans un VPC. Lorsque vous lancez une instance, vous la lancez dans un sous-réseau de votre VPC. Utilisez des sous-réseaux pour isoler les niveaux de votre application (par exemple, web, application et base de données) dans un VPC unique. Utilisez des sous-réseaux privés pour vos instances si elles ne doivent pas être accessibles directement à partir d'Internet.

Pour appeler l'API Amazon EC2 depuis votre VPC sans envoyer de trafic sur le réseau Internet public, utilisez AWS PrivateLink.

Isolation sur les hôtes physiques

Différentes instances EC2 sur un même hôte physique sont isolées les unes des autres comme si elles se trouvaient sur des hôtes physiques distincts. L'hyperviseur isole l'UC et la mémoire, et les instances sont équipées de disques virtuels au lieu d'accéder aux disques bruts.

Lorsque vous arrêtez ou résiliez une instance, la mémoire qui lui est allouée est remise à zéro par l'hyperviseur avant d'être allouée à une nouvelle instance, et chaque bloc de stockage est réinitialisé. Cela permet d'être sûr que vos données ne seront pas accidentellement exposées sur une autre instance.

Les adresses MAC réseau sont affectées de façon dynamique aux instances par l'infrastructure réseau AWS. Les adresses IP sont affectées de façon dynamique aux instances par l'infrastructure réseau AWS ou affectées par un administrateur EC2 via des demandes d'API authentifiées. Le réseau AWS n'autorise les instances à envoyer du trafic qu'à partir des adresses MAC et IP qui leur sont affectées. Dans le cas contraire, le trafic est abandonné.

Par défaut, une instance ne peut pas recevoir un trafic qui ne lui est pas spécifiquement adressé. Si vous avez besoin d'exécuter des services de translation d'adresse réseau (NAT), de routage ou de pare-feu sur votre instance, vous pouvez désactiver la vérification origine/destination pour l'interface réseau.

Contrôle du trafic réseau

Vous devez prendre en compte les éléments suivants pour le contrôle du trafic réseau vers vos instances EC2 :

- Limitez l'accès à vos instances à l'aide de groupes de sécurité (p. 1062). Par exemple, vous pouvez autoriser uniquement le trafic provenant des plages d'adresses de votre réseau d'entreprise.
- Utilisez des sous-réseaux privés pour vos instances si elles ne doivent pas être accessibles directement à partir d'Internet. Utilisez un hôte bastion ou une passerelle NAT pour l'accès Internet à partir d'une instance d'un sous-réseau privé.
- Utilisez AWS Virtual Private Network ou AWS Direct Connect pour établir des connexions privées de vos réseaux distants vers vos VPC. Pour de plus amples informations, veuillez consulter Network-to-Amazon VPC Connectivity Options.
- Utilisez des journaux de flux VPC pour surveiller la trafic atteignant vos instances.

- Utilisez AWS Security Hub pour rechercher les accès réseau non intentionnels à partir de vos instances.
- Utilisez EC2 Instance Connect (p. 604) pour vous connecter à vos instances à l'aide de Secure Shell (SSH) sans devoir partager et gérer des clés SSH.
- Utilisez AWS Systems Manager Session Manager pour accéder à vos instances de façon distante au lieu d'ouvrir des ports SSH entrants et gérer des clés SSH.
- Utilisez AWS Systems Manager Run Command pour automatiser les tâches administratives courantes au lieu d'ouvrir des ports SSH entrants et de gérer des clés SSH.

En plus de restreindre l'accès réseau à chaque instance Amazon EC2, Amazon VPC prend en charge la mise en œuvre de contrôles de sécurité réseau supplémentaires, notamment les passerelles en ligne, les serveurs proxy et diverses options de surveillance réseau.

Pour de plus amples informations, veuillez consulter le livre blanc intitulé AWS Security Best Practices.

Amazon EC2 et points de terminaison d'un VPC d'interface

Vous pouvez améliorer le niveau de sécurité de votre VPC en configurant Amazon EC2 pour utiliser un point de terminaison de VPC d'interface. Les points de terminaison d'interface sont optimisés par AWS PrivateLink, une technologie qui vous permet d'accéder de manière privée aux API Amazon EC2 en limitant tout le trafic réseau entre votre VPC et Amazon EC2 au réseau Amazon. Avec les points de terminaison d'interface, vous n'avez pas non plus besoin d'une passerelle Internet, d'un appareil NAT ou d'une passerelle privée virtuelle.

Il n'est pas obligatoire de configurer AWS PrivateLink, mais c'est recommandé. Pour de plus amples informations sur AWS PrivateLink et les points de terminaison du VPC, veuillez consulter Points de terminaison de l'interface du VPC (AWS PrivateLink).

Rubriques

- Création d'un point de terminaison de VPC d'interface (p. 973)
- Création d'une stratégie de point de terminaison de VPC d'interface (p. 973)

Création d'un point de terminaison de VPC d'interface

Création d'un point de terminaison pour Amazon EC2 à l'aide du nom de service suivant :

- **`com.amazonaws.`*`region`*`.ec2`** — Crée un point de terminaison pour les actions d'API Amazon EC2.

Pour plus d'informations, veuillez consulter Création d'un point de terminaison d'interface dans le Amazon VPC Guide de l'utilisateur.

Création d'une stratégie de point de terminaison de VPC d'interface

Vous pouvez attacher une stratégie à votre point de terminaison de VPC pour contrôler l'accès à l'API Amazon EC2. La stratégie spécifie :

- Le mandataire qui peut exécuter des actions.

- Les actions qui peuvent être effectuées.
- La ressource sur laquelle les actions peuvent être effectuées.

> **Important**
>
> Lorsqu'une stratégie autre que celle définie par défaut est appliquée à un point de terminaison VPC d'interface pour Amazon EC2, certaines demandes d'API ayant échoué, telles que celles qui échouent depuis `RequestLimitExceeded`, ne sont pas nécessairement consignées dans AWS CloudTrail ou Amazon CloudWatch.

Pour de plus amples informations, veuillez consulter Contrôle de l'accès aux services avec des points de terminaison d'un VPC dans le Amazon VPC Guide de l'utilisateur.

L'exemple suivant illustre une stratégie de point de terminaison VPC qui refuse l'autorisation de créer des volumes non chiffrés ou de lancer des instances avec des volumes non chiffrés. L'exemple de stratégie accorde également à tout le monde l'autorisation d'effectuer toutes les autres actions Amazon EC2.

```
{
    "Version": "2012-10-17",
    "Statement": [
    {
        "Action": "ec2:*",
        "Effect": "Allow",
        "Resource": "*",
        "Principal": "*"
    },
    {
        "Action": [
            "ec2:CreateVolume"
        ],
        "Effect": "Deny",
        "Resource": "*",
        "Principal": "*",
        "Condition": {
            "Bool": {
                "ec2:Encrypted": "false"
            }
        }
    },
    {
        "Action": [
            "ec2:RunInstances"
        ],
        "Effect": "Deny",
        "Resource": "*",
        "Principal": "*",
        "Condition": {
            "Bool": {
                "ec2:Encrypted": "false"
            }
        }
    }]
}
```

Résilience dans Amazon EC2

L'infrastructure mondiale d’AWS repose sur des régions et des zones de disponibilité AWS. Les régions fournissent plusieurs zones de disponibilité physiquement séparées et isolées, reliées par un réseau à

latence faible, à débit élevé et à forte redondance. 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 plus hautement disponibles, tolérantes aux pannes et évolutives que les infrastructures traditionnelles à un ou plusieurs centres de données.

Si vous avez besoin de répliquer vos données ou applications sur des distances géographiques plus importantes, utilisez les zones locales AWS. Une zone locale AWS est une extension d'une région AWS située à proximité géographique de vos utilisateurs. Les zones locales ont leurs propres connexions à Internet et prennent en charge AWS Direct Connect. Comme toutes les zones AWS, les régions locales AWS sont totalement isolées des autres zones AWS.

Si vous devez répliquer vos données ou applications dans une zone locale AWS, AWS vous recommande d'utiliser l'une des zones suivantes comme zone de basculement :

- Une autre zone locale
- Une zone de disponibilité dans la région qui n'est pas la zone parent Vous pouvez utiliser la commande describe-availability-zones pour afficher la zone parent.

Pour de plus amples informations sur les régions et les zones de disponibilité AWS, veuillez consulter Infrastructure mondiale AWS.

Outre l'infrastructure mondiale AWS, Amazon EC2, propose les fonctionnalités suivantes pour la prise en charge de la résilience des données :

- Copie des AMI entre régions
- Copie des instantanés EBS entre régions
- Automatisation des instantanés EBS à l'aide d'Gestionnaire de cycle de vie des données Amazon
- Gestion de la santé et de la disponibilité de votre flotte Amazon EC2 Auto Scaling
- Distribution du trafic entrant sur plusieurs instances dans une ou plusieurs zones de disponibilité à l'aide d'Elastic Load Balancing

Protection des données dans Amazon EC2

The AWS shared responsibility model applies to data protection in Amazon Elastic Compute Cloud. As described in this model, AWS is responsible for protecting the global infrastructure that runs all of the AWS Cloud. You are responsible for maintaining control over your content that is hosted on this infrastructure. This content includes the security configuration and management tasks for the AWS services that you use. For more information about data privacy, see the Data Privacy FAQ. For information about data protection in Europe, see the AWS Shared Responsibility Model and GDPR blog post on the AWS Security Blog.

For data protection purposes, we recommend that you protect AWS account credentials and set up individual user accounts with AWS Identity and Access Management (IAM). That way each user is given only the permissions necessary to fulfill their job duties. We also recommend that you secure your data in the following ways:

- Use multi-factor authentication (MFA) with each account.
- Use SSL/TLS to communicate with AWS resources. We recommend TLS 1.2 or later.
- Set up API and user activity logging with AWS CloudTrail.
- Use AWS encryption solutions, along with all default security controls within AWS services.
- Use advanced managed security services such as Amazon Macie, which assists in discovering and securing personal data that is stored in Amazon S3.

- If you require FIPS 140-2 validated cryptographic modules when accessing AWS through a command line interface or an API, use a FIPS endpoint. For more information about the available FIPS endpoints, see Federal Information Processing Standard (FIPS) 140-2.

We strongly recommend that you never put sensitive identifying information, such as your customers' account numbers, into free-form fields such as a Name field. This includes when you work with Amazon EC2 or other AWS services using the console, API, AWS CLI, or AWS SDKs. Any data that you enter into Amazon EC2 or other services might get picked up for inclusion in diagnostic logs. When you provide a URL to an external server, don't include credentials information in the URL to validate your request to that server.

Chiffrement au repos

Le chiffrement Amazon EBS est une solution de chiffrement destinées à vos volumes et instantanés EBS. Il utilise les clés principales client (CMK) AWS Key Management Service (AWS KMS). Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Les données sur les volumes de stockage d'instance NVMe sont chiffrées à l'aide d'un chiffrement XTS-AES-256 implémenté dans un module matériel sur l'instance. Les clés de chiffrement sont générées à l'aide du module matériel et sont uniques pour chaque périphérique de stockage d'instance NVMe. Toutes les clés de chiffrement sont détruites lorsque l'instance est arrêtée ou résiliée et ne peuvent pas être récupérées. Vous ne pouvez pas désactiver le chiffrement et vous ne pouvez pas fournir votre propre clé de chiffrement.

Chiffrement en transit

SSH fournit un canal de communications sécurisé pour l'accès distant à vos instances Linux. L'accès distant à vos instances via AWS Systems Manager Session Manager et Run Command est chiffré à l'aide de TLS 1.2 et les demandes de création d'une connexion sont chiffrées à l'aide de SigV4.

Utilisez un protocole de chiffrement tel que Transport Layer Security (TLS) pour chiffrer les données sensibles en transit entre les clients et vos instances.

AWS assure une connectivité sécurisée et privée entre les instances EC2 de tous les types. En outre, certains types d'instances utilisent les capacités de transfert de charges du matériel sous-jacent pour chiffrer automatiquement le trafic en transit entre les instances, à l'aide d'algorithmes AEAD avec chiffrement 256 bits. Il n'y a aucun impact sur les performances du réseau. Pour assurer le chiffrement supplémentaire du trafic en transit, les exigences suivantes doivent être respectées :

- Les instances utilisent les types d'instance suivants : C5a, C5ad, C5n, G4, I3en, M5dn, M5n, P3dn, R5dn et R5n.
- Les instances se trouvent dans la même région.
- Les instances se trouvent dans le même VPC ou des VPC appairés, et le trafic ne passe pas par un périphérique réseau virtuel, tel qu'un équilibreur de charge ou une passerelle de transit.

Identity and Access Management pour Amazon EC2

Vos informations d'identification de sécurité vous identifient aux services dans AWS et vous accordent une utilisation illimitée de vos ressources AWS, telles que vos ressources Amazon EC2. Vous pouvez utiliser les fonctions d'Amazon EC2 et d'AWS Identity and Access Management (IAM) pour permettre aux autres

utilisateurs, services et applications d'utiliser vos ressources Amazon EC2 sans partager vos informations d'identification de sécurité. Vous pouvez utiliser IAM pour contrôler la façon dont les autres utilisateurs emploient les ressources de votre compte AWS et vous pouvez employer les groupes de sécurité pour contrôler l'accès à vos instances Amazon EC2. Vous pouvez choisir entre une utilisation complète et une utilisation limitée de vos ressources Amazon EC2.

Sommaire

- Accès réseau à votre instance (p. 977)
- Attributs d'autorisation Amazon EC2 (p. 977)
- IAM et Amazon EC2 (p. 977)
- Stratégies IAM pour Amazon EC2 (p. 979)
- Rôles IAM pour Amazon EC2 (p. 1034)
- Autorisation du trafic entrant pour vos instances Linux (p. 1044)

Accès réseau à votre instance

Un groupe de sécurité fonctionne comme un pare-feu qui contrôle le trafic autorisé à atteindre une ou plusieurs instances. Lorsque vous lancez une instance, vous lui attribuez un ou plusieurs groupes de sécurité. Vous ajoutez des règles à chaque groupe de sécurité qui contrôlent le trafic de l'instance. Vous pouvez modifier les règles d'un groupe de sécurité à tout moment. Les nouvelles règles sont appliquées automatiquement à toutes les instances auxquelles le groupe de sécurité est affecté.

Pour plus d'informations, consultez Autorisation du trafic entrant pour vos instances Linux (p. 1044).

Attributs d'autorisation Amazon EC2

Votre organisation peut avoir plusieurs comptes AWS. Amazon EC2 vous permet de spécifier des comptes AWS supplémentaires qui peuvent utiliser vos Amazon Machine Images (AMIs) et vos instantanés Amazon EBS. Ces autorisations fonctionnent seulement au niveau du compte AWS ; vous ne pouvez pas limiter les autorisations pour des utilisateurs spécifiques au sein du compte AWS spécifié. Tous les utilisateurs du compte AWS que vous avez spécifiés peuvent utiliser l'AMI ou l'instantané.

Chaque AMI possède un attribut `LaunchPermission` qui contrôle les comptes AWS pouvant accéder à l'AMI. Pour plus d'informations, consultez Rendre une AMI publique (p. 119).

Chaque instantané Amazon EBS possède un attribut `createVolumePermission` qui contrôle les comptes AWS pouvant utiliser l'instantané. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

IAM et Amazon EC2

IAM vous permet d'effectuer les tâches suivantes :

- Créer des utilisateurs et des groupes sous votre compte AWS
- Attribuer des informations d'identification de sécurité uniques à chaque utilisateur de votre compte AWS
- Contrôler les autorisations de chaque utilisateur pour exécuter les tâches à l'aide des ressources AWS
- Permettre aux utilisateurs d'un autre compte AWS de partager vos ressources AWS
- Créer des rôles pour votre compte AWS et définir les utilisateurs ou services qui peuvent les assumer
- Utiliser les identités existantes de votre entreprise pour attribuer les autorisations d'exécuter des tâches à l'aide des ressources AWS

Grâce à l'utilisation d'IAM avec Amazon EC2, vous pouvez contrôler si les utilisateurs de votre organisation peuvent exécuter une tâche à l'aide d'actions d'API Amazon EC2 particulières et s'ils peuvent utiliser les ressources AWS spécifiques.

Cette rubrique vous aide à répondre aux questions suivantes :

- Comment créer des groupes et des utilisateurs dans IAM ?
- Comment créer une stratégie ?
- De quelles stratégies IAM ai-je besoin pour exécuter des tâches dans Amazon EC2 ?
- Comment accorder des permissions pour exécuter des actions dans Amazon EC2 ?
- Comment attribuer des permissions pour exécuter des actions sur des ressources spécifiques dans Amazon EC2 ?

Création d'un groupe et d'utilisateurs IAM

Pour créer un groupe IAM

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le panneau de navigation, choisissez Groupes, puis Créer un nouveau groupe.
3. Dans la zone Nom du groupe, entrez un nom pour votre groupe, puis choisissez Étape suivante.
4. Sur la page Attacher la stratégie, sélectionnez une stratégie gérée AWS, puis choisissez Étape suivante. Par exemple, pour Amazon EC2, l'une des politiques gérées AWS suivantes peut répondre à vos besoins :
   - PowerUserAccess
   - ReadOnlyAccess
   - AmazonEC2FullAccess
   - AmazonEC2ReadOnlyAccess
5. Choisissez Create Group.

Votre nouveau groupe apparaît sous Nom du groupe.

Pour créer un utilisateur IAM, ajouter l'utilisateur à votre groupe et créer un mot de passe pour l'utilisateur

1. Dans le panneau de navigation, sélectionnez Utilisateurs, Add user (Ajouter un utilisateur).
2. Entrez un nom d'utilisateur dans Nom utilisateur.
3. Pour Access type (Type d'accès), sélectionnez Programmatic access (Accès par programmation) et AWS Management Console access (Accès à l'AWS Console)).
4. Pour Console password (Mot de passe de la console), choisissez l'une des options suivantes :
   - Autogenerated password (Mot de passe généré automatiquement). Chaque utilisateur obtient un mot de passe généré de façon aléatoire qui correspond à la stratégie de mot de passe actuelle en vigueur (le cas échéant). Vous pouvez afficher ou télécharger les mots de passe lorsque vous accédez à la page Final.
   - Custom password (Mot de passe personnalisé). Chaque utilisateur se voit attribuer le mot de passe que vous tapez dans la zone.
5. Choisissez Suivant : Autorisations.
6. Sur la page Réglez les permissions, choisissez Ajouter un utilisateur au groupe. Sélectionnez la case à cocher en regard du groupe que vous avez créé précédemment, puis choisissez Next: Review (Suivant : Vérification).
7. Choisissez Create user.

8. Pour afficher les clés d'accès des utilisateurs (ID de clé d'accès et clés d'accès secrètes), choisissez Afficher à côté de chaque mot de passe et clé d'accès secrète à voir. Pour enregistrer les clés d'accès, choisissez Télécharger le rapport CSV, puis enregistrez le fichier dans un emplacement sûr sur votre ordinateur.

   Important

   Vous ne pouvez pas récupérer la clé d'accès secrète après avoir exécuté cette étape ; si vous l'égarez, vous devrez la recréer.
9. Choisissez Fermer.
10. Donnez à chaque utilisateur ses informations d'identification (clés d'accès et mot de passe) ; ils peuvent ainsi utiliser les services en fonction des autorisations que vous avez spécifiées pour le groupe IAM.

Voir aussi

Pour plus d'informations sur IAM, consultez les ressources suivantes :

- Stratégies IAM pour Amazon EC2 (p. 979)
- Rôles IAM pour Amazon EC2 (p. 1034)
- AWS Identity and Access Management (IAM)
- IAM Guide de l'utilisateur

Stratégies IAM pour Amazon EC2

Par défaut, les utilisateurs IAM n'ont pas l'autorisation de créer ou de modifier des ressources Amazon EC2 ou d'exécuter des tâches à l'aide de l'API Amazon EC2. (Cela signifie qu'ils ne peuvent pas non plus le faire à l'aide de la console ou de l'interface ligne de commande Amazon EC2.) Pour autoriser les utilisateurs IAM à créer ou à modifier des ressources et à exécuter des tâches, vous devez créer les stratégies IAM qui accordent aux utilisateurs IAM l'autorisation d'utiliser les actions d'API et ressources spécifiques dont ils ont besoin, puis d'attacher ces stratégies aux utilisateurs ou groupes IAM qui requièrent ces autorisations.

Quand vous attachez une stratégie à un utilisateur ou à un groupe d'utilisateurs, elle accorde ou refuse aux utilisateurs l'autorisation d'exécuter les tâches spécifiées sur les ressources spécifiées. Pour plus d'informations générales sur les stratégies IAM, consultez Autorisations et stratégies dans le IAM Guide de l'utilisateur. Pour plus d'informations sur la gestion et la création de stratégies IAM personnalisées, consultez Gestion des stratégies IAM.

Mise en route

Une stratégie IAM doit accorder ou refuser les autorisations permettant d'utiliser une ou plusieurs actions Amazon EC2. Elle doit aussi spécifier les ressources qui peuvent être utilisées avec l'action : il peut s'agir de toutes les ressources ou, dans certains cas, de ressources spécifiques. La stratégie peut aussi inclure les conditions que vous appliquez à la ressource.

Amazon EC2 prend partiellement en charge les permissions au niveau des ressources. Cela signifie que pour certaines actions d'API EC2, vous ne pouvez pas spécifier quelle ressource un utilisateur est autorisé à utiliser pour cette action. Au lieu de cela, vous devez autoriser les utilisateurs à utiliser toutes les ressources pour cette action.

Tâche	Sujet
Comprendre la structure de base d'une stratégie	Syntaxe d'une stratégie (p. 980)

Tâche	Sujet
Définir les actions de votre stratégie	Actions pour Amazon EC2 (p. 981)
Définir les ressources spécifiques de votre stratégie	Amazon Resource Names (ARN) pour Amazon EC2 (p. 982)
Appliquer les conditions à l'utilisation des ressources	Clés de condition pour Amazon EC2 (p. 983)
Utiliser les permissions disponibles au niveau des ressources pour Amazon EC2	Actions, ressources et clés de condition pour Amazon EC2 (IAM Guide de l'utilisateur)
Tester votre stratégie	Vérification que les utilisateurs ont les autorisations requises (p. 984)
Exemple de stratégies pour une interface ligne de commande ou un SDK	Exemples de stratégies à utiliser avec l'AWS CLI ou un kit SDK AWS (p. 987)
Exemple de politiques pour la console Amazon EC2	Exemples de stratégies à utiliser sur la console Amazon EC2 (p. 1025)

Structure d'une stratégie

Les rubriques suivantes expliquent la structure d'une stratégie IAM.

Sommaire

- Syntaxe d'une stratégie (p. 980)
- Actions pour Amazon EC2 (p. 981)
- Autorisations au niveau des ressources prises en charge pour les opérations d'API Amazon Amazon EC2 (p. 981)
- Amazon Resource Names (ARN) pour Amazon EC2 (p. 982)
- Clés de condition pour Amazon EC2 (p. 983)
- Vérification que les utilisateurs ont les autorisations requises (p. 984)

Syntaxe d'une stratégie

Une stratégie IAM est un document JSON qui se compose d'une ou de plusieurs déclarations. Chaque déclaration est structurée comme suit :

```
{
  "Statement":[{
    "Effect":"effect",
    "Action":"action",
    "Resource":"arn",
    "Condition":{
      "condition":{
        "key":"value"
        }
      }
    }
  ]
}
```

Une déclaration se compose de différents éléments :

- Effect : effect peut avoir la valeur `Allow` ou `Deny`. Comme, par défaut, les utilisateurs IAM n'ont pas la permission d'utiliser les ressources et les actions d'API, toutes les demandes sont refusées. Une autorisation explicite remplace l'autorisation par défaut. Un refus explicite remplace toute autorisation.
- Action : action désigne l'action d'API spécifique pour laquelle vous accordez ou refusez l'autorisation. Pour en savoir plus sur la spécification d'action, consultez Actions pour Amazon EC2 (p. 981).
- Resource : la ressource affectée par l'action. Certaines actions d'API Amazon EC2 vous permettent d'inclure des ressources spécifiques dans votre politique qui peuvent être créées ou modifiées par l'action. Vous spécifiez une ressource à l'aide d'un Amazon Resource Name (ARN) ou du caractère générique (*) pour indiquer que l'instruction s'applique à toutes les ressources. Pour plus d'informations, consultez Autorisations au niveau des ressources prises en charge pour les opérations d'API Amazon Amazon EC2 (p. 981).
- Condition : les conditions sont facultatives. Elles permettent de contrôler à quel moment votre stratégie est effective. Pour plus d'informations sur la spécification des conditions pour Amazon EC2, consultez Clés de condition pour Amazon EC2 (p. 983).

Pour plus d'informations sur les exemples de déclarations de stratégie IAM pour Amazon EC2, consultez Exemples de stratégies à utiliser avec l'AWS CLI ou un kit SDK AWS (p. 987).

Actions pour Amazon EC2

Dans une déclaration de stratégie IAM, vous pouvez spécifier une action d'API à partir de n'importe quel service prenant en charge IAM. Pour Amazon EC2, utilisez le préfixe suivant avec le nom de l'action d'API : `ec2:`. Par exemple : `ec2:RunInstances` et `ec2:CreateImage`.

Pour spécifier plusieurs actions dans une seule déclaration, séparez-les par des virgules comme suit :

```
"Action": ["ec2:action1", "ec2:action2"]
```

Vous pouvez aussi spécifier plusieurs actions à l'aide de caractères génériques. Par exemple, vous pouvez spécifier toutes les actions dont le nom commence par le mot « Describe » comme suit :

```
"Action": "ec2:Describe*"
```

Pour spécifier toutes les actions d'API Amazon EC2, utilisez le caractère générique * comme suit :

```
"Action": "ec2:*"
```

Pour obtenir la liste des actions Amazon EC2, consultez Actions dans le Amazon EC2 API Reference.

Autorisations au niveau des ressources prises en charge pour les opérations d'API Amazon Amazon EC2

Les autorisations au niveau des ressources désignent la capacité de spécifier les ressources sur lesquelles les utilisateurs sont autorisés à exécuter des actions. Amazon EC2 prend partiellement en charge les autorisations niveau ressources. Cela signifie que pour certaines actions Amazon EC2, vous pouvez contrôler à quel moment les utilisateurs sont autorisés à utiliser ces actions en fonction des conditions qui doivent être satisfaites, ou les ressources spécifiques que les utilisateurs sont autorisés à utiliser. Par exemple, vous pouvez accorder aux utilisateurs les autorisations de lancer des instances, mais uniquement d'un type spécifique et seulement à l'aide d'une AMI spécifique.

Pour spécifier une ressource dans la déclaration de stratégie IAM, vous utilisez son Amazon Resource Name (ARN). Pour plus d'informations sur la spécification de la valeur de l'ARN, consultez Amazon Resource Names (ARN) pour Amazon EC2 (p. 982). Si une action d'API ne prend pas en charge les

ARN individuels, utilisez un caractère générique (*) pour spécifier que toutes les ressources peuvent être concernées par l'action.

Pour afficher les tableaux qui identifient les actions d'API Amazon EC2 qui prennent en charge les autorisations au niveau des ressources, ainsi que les ARN et les clés de condition que vous pouvez utiliser dans une stratégie, consultez Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur. Les clés de condition pour Amazon EC2 sont également expliquées plus en détail dans une section ultérieure.

Notez que vous pouvez appliquer des autorisations au niveau des ressources et basées sur des balises dans les stratégies IAM que vous utilisez pour les actions d'API Amazon EC2. Vous bénéficiez ainsi d'un meilleur contrôle sur les ressources qu'un utilisateur peut créer, modifier ou utiliser. Pour plus d'informations, consultez Octroi de l'autorisation de baliser les ressources lors de la création (p. 985).

Amazon Resource Names (ARN) pour Amazon EC2

Chaque déclaration de stratégie IAM s'applique aux ressources que vous spécifiez à l'aide de leur ARN.

Un ARN obéit à la syntaxe générale suivante :

```
arn:aws:[service]:[region]:[account]:resourceType/resourcePath
```

web

Le service (par exemple, `ec2`).

région

La région de la ressource (par exemple, `us-east-1`).

compte

L'ID du compte AWS, sans trait d'union (par exemple, `123456789012`).

type de ressource

Le type de ressource (par exemple, `instance`).

chemin de la ressource

Un chemin qui identifie la ressource. Vous pouvez utiliser le caractère générique * dans vos chemins.

Par exemple, vous pouvez indiquer une instance spécifique (`i-1234567890abcdef0`) dans votre déclaration à l'aide de son ARN comme suit :

```
"Resource": "arn:aws:ec2:us-east-1:123456789012:instance/i-1234567890abcdef0"
```

Vous pouvez spécifier toutes les instances qui appartiennent à un compte spécifique à l'aide du caractère générique * comme suit :

```
"Resource": "arn:aws:ec2:us-east-1:123456789012:instance/*"
```

Vous pouvez aussi spécifier toutes les ressources Amazon EC2 qui appartiennent à un compte spécifique à l'aide du caractère générique * comme suit :

```
"Resource": "arn:aws:ec2:us-east-1:123456789012:*"
```

Pour spécifier toutes les ressources, ou si une action d'API spécifique ne prend pas en charge les ARN, utilisez le caractère générique * dans l'élément `Resource` comme suit :

```
"Resource": "*"
```

De nombreuses actions d'API Amazon EC2 nécessitent plusieurs ressources. Par exemple, comme `AttachVolume` attache un volume Amazon EBS à une instance, un utilisateur IAM doit avoir les autorisations nécessaires pour utiliser le volume et l'instance. Pour spécifier plusieurs ressources dans une seule déclaration, séparez leurs ARN par des virgules, comme suit :

```
"Resource": ["arn1", "arn2"]
```

Pour obtenir la liste des ARN pour les ressources Amazon EC2, consultez Types de ressources définis par Amazon EC2 dans le IAM Guide de l'utilisateur.

Clés de condition pour Amazon EC2

Dans une déclaration de stratégie, vous pouvez, le cas échéant, spécifier des conditions qui contrôlent à quel moment la déclaration est effective. Chaque condition contient une ou plusieurs paires clé-valeur. Les clés de condition ne sont pas sensibles à la casse. Nous avons défini des conditions de clé à l'échelle d'AWS, ainsi que des clés de condition supplémentaires spécifiques aux services.

Pour obtenir la liste des clés de condition spécifiques au service pour Amazon EC2, veuillez consulter Clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur. Amazon EC2 implémente également les clés de condition à l'échelle d'AWS. Pour plus d'informations, consultez Informations disponibles dans toutes les demandes dans le IAM Guide de l'utilisateur.

Pour utiliser une clé de condition dans votre stratégie IAM, utilisez l'instruction `Condition`. Par exemple, la stratégie suivante accorde aux utilisateurs l'autorisation d'ajouter et de supprimer des règles entrantes et sortantes pour n'importe quel groupe de sécurité. Elle utilise la clé de condition `ec2:Vpc` pour spécifier que ces actions ne peuvent être effectuées que sur des groupes de sécurité dans un VPC spécifique.

```
{
"Version": "2012-10-17",
  "Statement":[{
    "Effect":"Allow",
    "Action": [
       "ec2:AuthorizeSecurityGroupIngress",
       "ec2:AuthorizeSecurityGroupEgress",
       "ec2:RevokeSecurityGroupIngress",
       "ec2:RevokeSecurityGroupEgress"],
     "Resource": "arn:aws:ec2:region:account:security-group/*",
      "Condition": {
        "StringEquals": {
          "ec2:Vpc": "arn:aws:ec2:region:account:vpc/vpc-11223344556677889"
        }
      }
    }
  ]
}
```

Si vous spécifiez plusieurs conditions ou plusieurs clés dans une même condition, elles sont analysées à l'aide d'une opération logique AND. Si vous spécifiez une seule condition avec plusieurs valeurs pour une clé, la condition est analysée à l'aide d'une opération logique OR. Pour que les autorisations soient accordées, toutes les conditions doivent être satisfaites.

Vous pouvez aussi utiliser des espaces réservés quand vous spécifiez des conditions. Par exemple, vous pouvez accorder à un utilisateur IAM l'autorisation d'utiliser des ressources avec une balise qui spécifie

son nom d'utilisateur IAM. Pour plus d'informations, consultez Variables de stratégie dans le IAM Guide de l'utilisateur.

> **Important**
>
> Plusieurs clés de condition sont propres à une ressource et certaines actions d'API utilisent plusieurs ressources. Si vous écrivez une stratégie avec une clé de condition, utilisez l'élément `Resource` de la déclaration pour spécifier la ressource à laquelle la clé de condition s'applique. Dans le cas contraire, la stratégie peut empêcher totalement les utilisateurs d'exécuter l'action, car le contrôle de la condition échoue pour les ressources auxquelles la clé de condition ne s'applique pas. Si vous ne voulez pas spécifier de ressource ou si vous avez écrit l'élément `Action` de votre stratégie pour inclure plusieurs actions d'API, vous devez utiliser le type de condition `...IfExists` pour garantir que la clé de condition est ignorée pour les ressources qui ne l'utilisent pas. Pour plus d'informations, consultez Conditions ...IfExists dans le IAM Guide de l'utilisateur.

Toutes les actions Amazon EC2 prennent en charge les clés de condition `aws:RequestedRegion` et `ec2:Region`. Pour plus d'informations, consultez Exemple : Restriction de l'accès à une région spécifique (p. 988).

La clé `ec2:SourceInstanceARN` peut être utilisée pour les conditions qui spécifient le nom ARN de l'instance à partir de laquelle une demande a été effectuée. Cette clé de condition est disponible sur l'ensemble des services AWS et n'est pas spécifique à un service. Pour examiner des exemples de stratégie, consultez Autorise une instance EC2 à attacher et détacher des volumes et Exemple : Octroi à une instance spécifique de l'autorisation de visualiser des ressources dans d'autres services AWS (p. 1021). La clé `ec2:SourceInstanceARN` ne peut pas être utilisée comme variable pour renseigner le nom ARN de l'élément `Resource` dans une instruction.

Pour obtenir des déclarations de politique pour Amazon EC2, consultez Exemples de stratégies à utiliser avec l'AWS CLI ou un kit SDK AWS (p. 987).

Vérification que les utilisateurs ont les autorisations requises

Après que vous avez créé une stratégie IAM, il vous est recommandé de vérifier si elle accorde aux utilisateurs les autorisations d'utiliser les actions d'API et ressources particulières dont ils ont besoin avant que vous ne placiez la stratégie en production.

D'abord, créez un utilisateur IAM à des fins de test, puis attachez la stratégie IAM que vous avez créée à l'utilisateur test. Ensuite, créez une demande en tant qu'utilisateur test.

Si l'action Amazon EC2 que vous testez crée ou modifie une ressource, vous devez créer la demande à l'aide du paramètre `DryRun` (ou exécuter la commande de l'AWS CLI avec l'option `--dry-run`). Dans ce cas, l'appel conclut le contrôle d'autorisation, mais non l'opération. Par exemple, vous pouvez vérifier si l'utilisateur peut terminer une instance particulière sans réellement l'achever. Si l'utilisateur a les autorisations requises, la demande retourne `DryRunOperation` ; sinon, elle retourne `UnauthorizedOperation`.

Si la stratégie n'accorde pas à l'utilisateur les autorisations que vous escomptiez, ou si elles sont trop excessives, vous pouvez ajuster la stratégie selon vos besoins et la tester à nouveau jusqu'à ce que vous obteniez les résultats souhaités.

> **Important**
>
> La propagation des modifications de la stratégie peut durer plusieurs minutes avant qu'elles ne prennent effet. Par conséquent, il est recommandé que vous laissiez s'écouler cinq minutes avant de tester les mises à jour de votre stratégie.

Si un contrôle d'autorisation échoue, la demande retourne un message codé avec les informations de diagnostic. Vous pouvez décoder le message à l'aide de l'action `DecodeAuthorizationMessage`. Pour

plus d'informations, consultez DecodeAuthorizationMessage dans le AWS Security Token Service API Reference, et decode-authorization-message dans le AWS CLI Command Reference.

Octroi de l'autorisation de baliser les ressources lors de la création

Certaines actions d'API Amazon EC2 de création de ressources vous permettent de spécifier des balises lorsque vous créez la ressource. Pour plus d'informations, consultez Balisage de vos ressources (p. 1310).

Pour permettre aux utilisateurs d'attribuer des balises aux ressources au moment de la création, ils doivent avoir les autorisations d'utiliser l'action qui crée la ressource (par exemple, `ec2:RunInstances` ou `ec2:CreateVolume`). Si les balises sont spécifiées dans l'action de création de ressources, Amazon effectue une autorisation supplémentaire sur l'action `ec2:CreateTags` pour vérifier si les utilisateurs sont autorisés à créer des balises. Par conséquent, les utilisateurs doivent également avoir des autorisations explicites d'utiliser l'action `ec2:CreateTags`.

Dans la définition de stratégie IAM de l'action `ec2:CreateTags`, utilisez l'élément `Condition` avec la clé de condition `ec2:CreateAction` pour accorder des autorisations de balisage à l'action qui crée la ressource.

L'exemple suivant illustre une stratégie qui permet aux utilisateurs de lancer des instances et d'appliquer des balises aux instances et aux volumes pendant le lancement. Les utilisateurs ne sont pas autorisés à attribuer des balises aux ressources existantes (ils ne peuvent pas appeler l'action `ec2:CreateTags` directement).

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
         "ec2:RunInstances"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:region:account:*/*",
      "Condition": {
         "StringEquals": {
             "ec2:CreateAction" : "RunInstances"
          }
       }
    }
  ]
}
```

De même, la stratégie suivante permet aux utilisateurs de créer des volumes et appliquer des balises à des volumes pendant la création de volume. Les utilisateurs ne sont pas autorisés à attribuer des balises aux ressources existantes (ils ne peuvent pas appeler l'action `ec2:CreateTags` directement).

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateVolume"
      ],
      "Resource": "*"
```

```
    },
    {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:region:account:*/*",
      "Condition": {
         "StringEquals": {
             "ec2:CreateAction" : "CreateVolume"
          }
       }
    }
  ]
}
```

L'action `ec2:CreateTags` est uniquement évaluée si les balises sont appliquées pendant l'action de création de ressources. Par conséquent, un utilisateur qui est autorisé à créer une ressource (en supposant qu'il n'existe aucune condition de balisage) n'a pas besoin des autorisations d'utiliser l'action `ec2:CreateTags` si aucune balise n'est spécifié dans la demande. Toutefois, si l'utilisateur essaie de créer une ressource avec des balises, la demande échoue s'il n'a pas les autorisations d'utiliser l'action `ec2:CreateTags`.

L'action `ec2:CreateTags` est également évaluée si des balises sont fournies dans un modèle de lancement. Pour un exemple de stratégie, consultez Balises dans un modèle de lancement (p. 1009).

Contrôle de l'accès à des balises spécifiques

Vous pouvez utiliser des conditions supplémentaires dans l'élément `Condition` de vos stratégies IAM pour contrôler les clés de balise et les valeurs qui peuvent être appliquées aux ressources.

Les clés de condition suivantes peuvent être utilisées avec les exemples de la section précédente :

- `aws:RequestTag` : Pour indiquer qu'une clé de balise ou une clé et valeur de balise particulière doit être présente dans une demande. D'autres balises peuvent également être spécifiées dans la demande.
  - Utilisez avec l'opérateur de condition `StringEquals` pour appliquer une combinaison de clé de balise et de valeur spécifique ; par exemple, pour appliquer la balise `cost-center=cc123` :

    ```
    "StringEquals": { "aws:RequestTag/cost-center": "cc123" }
    ```

  - A utiliser avec l'opération de condition `StringLike` pour appliquer une clé de balise spécifique dans la demande ; par exemple, pour appliquer la clé de balise `purpose`:

    ```
    "StringLike": { "aws:RequestTag/purpose": "*" }
    ```

- `aws:TagKeys` : Pour appliquer les clés de balise qui sont utilisées dans la demande.
  - A utiliser avec le modificateur `ForAllValues` pour appliquer des clés de balise spécifiques si celles-ci sont fournies dans la demande (si les balises sont spécifiées dans la demande, seules les clés de balise spécifiques sont autorisées ; aucune autre balise n'est autorisée). Par exemple, les clés de balise `environment` ou `cost-center` sont autorisées :

    ```
    "ForAllValues:StringEquals": { "aws:TagKeys": ["environment","cost-center"] }
    ```

  - A utiliser avec le modificateur `ForAnyValue` pour appliquer la présence d'au moins l'une des clés de balise spécifiées dans la demande. Par exemple, au moins l'une des clés de balise `environment` ou `webserver` doit être présente dans la demande :

    ```
    "ForAnyValue:StringEquals": { "aws:TagKeys": ["environment","webserver"] }
    ```

Ces clés de condition peuvent être appliqués aux actions de création de ressources qui prennent en charge le balisage ainsi qu'aux actions `ec2:CreateTags` et `ec2:DeleteTags`. Pour savoir si une action d'API Amazon EC2 prend en charge le balisage, veuillez consulter Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur.

Pour forcer les utilisateurs à spécifier des balises quand ils créent une ressource, vous devez utiliser la clé de condition `aws:RequestTag` ou la clé de condition `aws:TagKeys` avec le modificateur `ForAnyValue` sur l'action de création de ressources. L'action `ec2:CreateTags` n'est pas évaluée si un utilisateur ne spécifie pas de balises pour l'action de création de ressources.

Pour les conditions, la clé de condition n'est pas sensible à la casse et la valeur de la condition est sensible à la casse. Par conséquent pour forcer la sensibilité à la casse d'une clé de balise, utilisez la clé de condition `aws:TagKeys`, où la clé de balise est indiquée comme une valeur dans la condition.

Par exemple les stratégies IAM, consultez Exemples de stratégies à utiliser avec l'AWS CLI ou un kit SDK AWS (p. 987). Pour plus d'informations sur les conditions à valeur multiples, consultez Création d'une condition qui teste plusieurs valeurs de clés dans le IAM Guide de l'utilisateur.

Contrôle de l'accès aux ressources EC2 à l'aide des balises de ressources

Lorsque vous créez une stratégie IAM qui accorde aux utilisateurs IAM l'autorisation d'utiliser les ressources EC2, vous pouvez inclure des informations de balise dans l'élément `Condition` de la stratégie pour contrôler l'accès en fonction des balises. Vous bénéficiez ainsi d'un meilleur contrôle sur les ressources EC2 qu'un utilisateur peut modifier, utiliser ou supprimer.

Par exemple, vous pouvez créer une stratégie qui permet aux utilisateurs de résilier une instance, mais qui refuse l'action si l'instance possède la balise `environment=production`. Pour ce faire, vous utilisez la clé de condition `ec2:ResourceTag` pour autoriser ou refuser l'accès à la ressource en fonction des balises attachées à la ressource.

```
"StringEquals": { "ec2:ResourceTag/environment": "production" }
```

Pour savoir si une action d'API Amazon EC2 prend en charge le contrôle d'accès à l'aide de la clé de condition `ec2:ResourceTag`, veuillez consulter Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur. Notez que les actions `Describe` ne prennent pas en charge les autorisations au niveau des ressources. Vous devez donc les spécifier dans une instruction distincte sans condition.

Par exemple les stratégies IAM, consultez Exemples de stratégies à utiliser avec l'AWS CLI ou un kit SDK AWS (p. 987).

> **Note**
>
> Si vous autorisez ou refusez à des utilisateurs l'accès à des ressources en fonction de balises, vous devez envisager de refuser de manière explicite la possibilité pour les utilisateurs d'ajouter ces balises ou de les supprimer des mêmes ressources. Sinon, il sera possible pour un utilisateur de contourner vos restrictions et d'obtenir l'accès à une ressource en modifiant ses balises.

Exemples de stratégies à utiliser avec l'AWS CLI ou un kit SDK AWS

Les exemples suivants illustrent des déclarations de stratégie que vous pouvez utiliser pour contrôler les autorisations des utilisateurs IAM sur Amazon EC2. Ces stratégies sont destinées aux demandes formulées avec l'interface ligne de commande AWS CLI ou un kit SDK AWS. Pour obtenir des exemples

de stratégies à utiliser sur la console Amazon EC2, consultez Exemples de stratégies à utiliser sur la console Amazon EC2 (p. 1025). Pour obtenir des exemples de stratégies IAM spécifiques à Amazon VPC, consultez Identity and Access Management pour Amazon VPC.

Exemples

- Exemple : accès en lecture seule (p. 988)
- Exemple : Restriction de l'accès à une région spécifique (p. 988)
- Utilisation d'instances (p. 989)
- Utilisation des volumes (p. 991)
- Utilisation des instantanés (p. 993)
- Lancement d'instances (RunInstances) (p. 1000)
- Utilisation de Instances Spot (p. 1012)
- Exemple : Utilisation d'Instances réservées (p. 1017)
- Exemple : Balisage de vos ressources (p. 1018)
- Exemple : Utilisation des rôles IAM (p. 1020)
- Exemple : Utilisation des tables de routage (p. 1021)
- Exemple : Octroi à une instance spécifique de l'autorisation de visualiser des ressources dans d'autres services AWS (p. 1021)
- Exemple : Utilisation des modèles de lancement (p. 1022)
- Utilisation des métadonnées d'instance (p. 1022)

Exemple : accès en lecture seule

La stratégie suivante accorde aux utilisateurs les autorisations d'utiliser toutes les actions d'API Amazon EC2 dont les noms commencent par `Describe`. L'élément `Resource` utilise un caractère générique pour indiquer que les utilisateurs peuvent spécifier toutes les ressources avec ces actions d'API. Le caractère générique * est également nécessaire dans les cas où l'action d'API ne prend pas en charge les autorisations au niveau des ressources. Pour plus d'informations sur les ARN que vous pouvez utiliser avec les actions d'API Amazon EC2, consultez Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur.

Les utilisateurs n'ont pas l'autorisation d'effectuer la moindre action sur les ressources (à moins qu'une autre déclaration ne leur accorde l'autorisation de le faire), car, par défaut, l'autorisation d'utiliser les actions d'API leur est refusée.

```
{
   "Version": "2012-10-17",
   "Statement": [
    {
      "Effect": "Allow",
      "Action": "ec2:Describe*",
      "Resource": "*"
    }
   ]
}
```

Exemple : Restriction de l'accès à une région spécifique

La stratégie suivante refuse aux utilisateurs l'autorisation d'utiliser toutes les actions d'API Amazon EC2 à moins que la région soit Europe (Francfort). Elle utilise la clé de condition globale `aws:RequestedRegion` qui est prise en charge par toutes les actions d'API Amazon EC2.

```
{
  "Version":"2012-10-17",
  "Statement": [
       {
      "Effect": "Deny",
      "Action": "ec2:*",
      "Resource": "*",
      "Condition": {
        "StringNotEquals": {
          "aws:RequestedRegion": "eu-central-1"
        }
      }
    }
  ]
}
```

Sinon, vous pouvez utiliser la clé de condition `ec2:Region`, qui est spécifique à Amazon EC2 et qui est prise en charge par toutes les actions d'API Amazon EC2.

```
{
  "Version":"2012-10-17",
  "Statement": [
       {
      "Effect": "Deny",
      "Action": "ec2:*",
      "Resource": "*",
      "Condition": {
        "StringNotEquals": {
          "ec2:Region": "eu-central-1"
        }
      }
    }
  ]
}
```

Utilisation d'instances

Exemples

- Exemple : Décrire, lancer, arrêter, démarrer et résilier toutes les instances (p. 989)
- Exemple : Décrire toutes les instances, et arrêter, démarrer et résilier uniquement des instances particulières (p. 990)

Exemple : Décrire, lancer, arrêter, démarrer et résilier toutes les instances

La stratégie suivante autorise les utilisateurs à effectuer les actions d'API spécifiées dans l'élément `Action`. L'élément `Resource` utilise un caractère générique * pour indiquer que les utilisateurs peuvent spécifier toutes les ressources avec ces actions d'API. Le caractère générique * est également nécessaire dans les cas où l'action d'API ne prend pas en charge les autorisations au niveau des ressources. Pour plus d'informations sur les ARN que vous pouvez utiliser avec les actions d'API Amazon EC2, consultez Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur.

Les utilisateurs n'ont pas l'autorisation d'utiliser d'autres actions d'API (à moins qu'une autre déclaration ne leur accorde l'autorisation de le faire), car, par défaut, l'autorisation d'utiliser les actions d'API leur est refusée.

```
{
   "Version": "2012-10-17",
   "Statement": [
```

```
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeInstances",
        "ec2:DescribeImages",
        "ec2:DescribeKeyPairs",
        "ec2:DescribeSecurityGroups",
        "ec2:DescribeAvailabilityZones",
        "ec2:RunInstances",
        "ec2:TerminateInstances",
        "ec2:StopInstances",
        "ec2:StartInstances"
      ],
      "Resource": "*"
    }
   ]
}
```

Exemple : Décrire toutes les instances, et arrêter, démarrer et résilier uniquement des instances particulières

La stratégie suivante autorise les utilisateurs à décrire toutes les instances, à démarrer et à arrêter uniquement les instances i-1234567890abcdef0 et i-0598c7d356eba48d7, et à ne terminer que les instances de la région Région USA Est (Virginie du N.) (`us-east-1`) avec la balise de ressource "`purpose=test`".

La première déclaration utilise un caractère générique * pour l'élément `Resource` de façon à indiquer que les utilisateurs peuvent spécifier toutes les ressources avec l'action ; dans le cas présent, ils peuvent afficher toutes les instances. Le caractère générique * est également nécessaire dans les cas où l'action d'API ne prend pas en charge les autorisations au niveau des ressources (dans le cas présent, `ec2:DescribeInstances`). Pour plus d'informations sur les ARN que vous pouvez utiliser avec les actions d'API Amazon EC2, consultez Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur.

La deuxième déclaration utilise des permissions au niveau des ressources pour les actions `StopInstances` et `StartInstances`. Les instances spécifiques sont indiquées par leurs ARN dans l'élément `Resource`.

La troisième déclaration permet aux utilisateurs de terminer toutes les instances de la région Région USA Est (Virginie du N.) (`us-east-1`) qui appartiennent au compte AWS spécifié, mais uniquement lorsque l'instance a la balise `"purpose=test"`. L'élément `Condition` stipule quand la déclaration de stratégie est en vigueur.

```
{
   "Version": "2012-10-17",
   "Statement": [
   {
   "Effect": "Allow",
      "Action": "ec2:DescribeInstances",
      "Resource": "*"
   },
   {
      "Effect": "Allow",
      "Action": [
        "ec2:StopInstances",
        "ec2:StartInstances"
      ],
      "Resource": [
        "arn:aws:ec2:us-east-1:123456789012:instance/i-1234567890abcdef0",
        "arn:aws:ec2:us-east-1:123456789012:instance/i-0598c7d356eba48d7"
      ]
```

```
    },
    {
      "Effect": "Allow",
      "Action": "ec2:TerminateInstances",
      "Resource": "arn:aws:ec2:us-east-1:123456789012:instance/*",
      "Condition": {
         "StringEquals": {
            "ec2:ResourceTag/purpose": "test"
         }
      }
   }

   ]
}
```

Utilisation des volumes

Exemples

- Exemple : Attachement et détachement des volumes (p. 991)
- Exemple : Création d'un volume (p. 992)
- Exemple : Création d'un volume avec des balises (p. 992)

Exemple : Attachement et détachement des volumes

Quand une action d'API requiert qu'un mandataire spécifie plusieurs ressources, vous devez créer une déclaration de stratégie qui permet aux utilisateurs d'accéder à toutes les ressources requises. Si vous devez utiliser un élément `Condition` avec une ou plusieurs de ces ressources, vous devez créer plusieurs déclarations, comme dans l'exemple ci-dessous.

La stratégie suivante permet aux utilisateurs d'attacher des volumes avec la balise "`volume_user`=nom-utilisateur-iam" aux instances avec la balise "`department=dev`", et de détacher ces volumes de ces instances. Si vous attachez cette stratégie à un groupe IAM, la variable `aws:username` accorde à chaque utilisateur IAM du groupe l'autorisation d'attacher des volumes aux instances (ou de les en détacher) avec une balise nommée `volume_user` qui a son nom d'utilisateur IAM comme valeur.

```
{
   "Version": "2012-10-17",
   "Statement": [
       {
      "Effect": "Allow",
      "Action": [
        "ec2:AttachVolume",
        "ec2:DetachVolume"
      ],
      "Resource": "arn:aws:ec2:us-east-1:123456789012:instance/*",
      "Condition": {
        "StringEquals": {
          "ec2:ResourceTag/department": "dev"
        }
      }
   },
   {
      "Effect": "Allow",
      "Action": [
        "ec2:AttachVolume",
        "ec2:DetachVolume"
      ],
      "Resource": "arn:aws:ec2:us-east-1:123456789012:volume/*",
      "Condition": {
        "StringEquals": {
```

```
          "ec2:ResourceTag/volume_user": "${aws:username}"
        }
      }
    }
  ]
}
```

Exemple : Création d'un volume

La stratégie suivante permet aux utilisateurs d'utiliser l'action d'API CreateVolume. L'utilisateur est autorisé à créer un volume uniquement si le volume est chiffré et seulement si la taille du volume est inférieure à 20 Gio.

```
{
  "Version": "2012-10-17",
  "Statement": [
         {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateVolume"
      ],
      "Resource": "arn:aws:ec2:us-east-1:123456789012:volume/*",
      "Condition":{
         "NumericLessThan": {
             "ec2:VolumeSize" : "20"
          },
          "Bool":{
              "ec2:Encrypted" : "true"
          }
       }
    }
  ]
}
```

Exemple : Création d'un volume avec des balises

La stratégie suivante inclut la clé de condition `aws:RequestTag` qui exige aux utilisateurs d'attribuer des balises aux volumes qu'ils créent avec les balises `costcenter=115` et `stack=prod`. La clé de condition `aws:TagKeys` utilise le modificateur `ForAllValues` pour indiquer que seules les clés `costcenter` et `stack` sont autorisées dans la demande (aucune autre balise ne peut être spécifiée). Si les utilisateurs ne transmettent pas ces balises spécifiques ou s'ils ne spécifient pas du tout de balises, la demande échoue.

Pour les actions de création de ressources qui appliquent des balises, les utilisateurs doivent être autorisés à effectuer l'action `CreateTags`. La deuxième déclaration utilise la clé de condition `ec2:CreateAction` pour permettre aux utilisateurs de créer des balises uniquement dans le contexte de `CreateVolume`. Les utilisateurs ne peuvent pas attribuer des balises à des volumes existants ou à d'autres ressources. Pour plus d'informations, consultez Octroi de l'autorisation de baliser les ressources lors de la création (p. 985).

```
{
  "Version": "2012-10-17",
  "Statement": [
        {
      "Sid": "AllowCreateTaggedVolumes",
      "Effect": "Allow",
      "Action": "ec2:CreateVolume",
      "Resource": "arn:aws:ec2:us-east-1:123456789012:volume/*",
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/costcenter": "115",
          "aws:RequestTag/stack": "prod"
         },
```

```
          "ForAllValues:StringEquals": {
              "aws:TagKeys": ["costcenter","stack"]
          }
        }
      },
      {
        "Effect": "Allow",
        "Action": [
          "ec2:CreateTags"
        ],
        "Resource": "arn:aws:ec2:us-east-1:123456789012:volume/*",
        "Condition": {
          "StringEquals": {
              "ec2:CreateAction" : "CreateVolume"
         }
       }
     }
   ]
}
```

La stratégie suivante permet aux utilisateurs de créer un volume sans avoir à spécifier des balises. L'action `CreateTags` est uniquement évaluée si les balises sont spécifiées dans la demande `CreateVolume`. Si les utilisateurs spécifient une balise, elle doit être `purpose=test`. Aucune autre balise n'est autorisée dans la demande.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "ec2:CreateVolume",
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:us-east-1:1234567890:volume/*",
      "Condition": {
         "StringEquals": {
             "aws:RequestTag/purpose": "test",
             "ec2:CreateAction" : "CreateVolume"
          },
         "ForAllValues:StringEquals": {
             "aws:TagKeys": "purpose"
          }
       }
    }
  ]
}
```

Utilisation des instantanés

Cette section contient des exemples de stratégie pour `CreateSnapshot` (instantané à un instant donné d'un volume EBS) et `CreateSnapshots` (instantanés multi-volumes).

Exemples

- Exemple :Création d'un instantané (p. 994)
- Exemple : Création d'instantanés (p. 994)
- Exemple : Création d'un instantané avec des balises (p. 995)

- Exemple : Création d'instantanés avec des balises (p. 995)
- Exemple : Modification des paramètres d'autorisation pour des instantanés (p. 1000)

Exemple :Création d'un instantané

La stratégie suivante permet aux clients d'utiliser l'action d'API CreateSnapshot. Le client peut créer des instantanés uniquement si le volume est chiffré et seulement si la taille du volume est inférieure à 20 Gio.

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshot",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*"
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshot",
         "Resource":"arn:aws:ec2:us-east-1:123456789012:volume/*",
         "Condition":{
            "NumericLessThan":{
               "ec2:VolumeSize":"20"
            },
            "Bool":{
               "ec2:Encrypted":"true"
            }
         }
      }
   ]
}
```

Exemple : Création d'instantanés

La stratégie suivante permet aux clients d'utiliser l'action d'API CreateSnapshots. Le client peut créer des instantanés seulement si tous les volumes sur l'instance sont de type GP2.

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":[
"arn:aws:ec2:us-east-1::snapshot/*",
"arn:aws:ec2:*:*:instance/*"
   ]
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":"arn:aws:ec2:us-east-1:*:volume/*",
         "Condition":{
            "StringLikeIfExists":{
               "ec2:VolumeType":"gp2"
             }
     }

      }
   ]
}
```

Exemple : Création d'un instantané avec des balises

La stratégie suivante inclut la clé de condition `aws:RequestTag`, qui nécessite que le client applique les balises `costcenter=115` et `stack=prod` à tout nouvel instantané. La clé de condition `aws:TagKeys` utilise le modificateur `ForAllValues` pour indiquer que seules les clés `costcenter` et `stack` peuvent être spécifiées dans la demande. La demande échoue si l'une de ces conditions n'est pas remplie.

Pour les actions de création de ressources qui appliquent des balises, les clients doivent être autorisés à effectuer l'action `CreateTags`. La troisième déclaration utilise la clé de condition `ec2:CreateAction` pour permettre aux clients de créer des balises uniquement dans le contexte de `CreateSnapshot`. Les clients ne peuvent pas attribuer des balises à des volumes existants ou à d'autres ressources. Pour plus d'informations, consultez Octroi de l'autorisation de baliser les ressources lors de la création (p. 985).

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshot",
         "Resource":"arn:aws:ec2:us-east-1:123456789012:volume/*"
      },
      {
         "Sid":"AllowCreateTaggedSnapshots",
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshot",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "aws:RequestTag/costcenter":"115",
               "aws:RequestTag/stack":"prod"
            },
            "ForAllValues:StringEquals":{
               "aws:TagKeys":[
                  "costcenter",
                  "stack"
               ]
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateTags",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "ec2:CreateAction":"CreateSnapshot"
            }
         }
      }
   ]
}
```

Exemple : Création d'instantanés avec des balises

La stratégie suivante inclut la clé de condition `aws:RequestTag`, qui nécessite que le client applique les balises `costcenter=115` et `stack=prod` à tout nouvel instantané. La clé de condition `aws:TagKeys` utilise le modificateur `ForAllValues` pour indiquer que seules les clés `costcenter` et `stack` peuvent être spécifiées dans la demande. La demande échoue si l'une de ces conditions n'est pas remplie.

```
{
   "Version":"2012-10-17",
   "Statement": [
```

```
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":[
"arn:aws:ec2:us-east-1::snapshot/*",
"arn:aws:ec2:*:*:instance/*",
"arn:aws:ec2:*:*:volume/*"

   ]
      },
      {
         "Sid":"AllowCreateTaggedSnapshots",
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "aws:RequestTag/costcenter":"115",
               "aws:RequestTag/stack":"prod"
            },
            "ForAllValues:StringEquals":{
               "aws:TagKeys":[
                  "costcenter",
                  "stack"
               ]
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateTags",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "ec2:CreateAction":"CreateSnapshots"
            }
         }
      }
   ]
}
```

La stratégie suivante permet aux clients de créer un instantané sans avoir à spécifier des balises. L'action `CreateTags` est évaluée uniquement si des balises sont spécifiées dans la demande `CreateSnapshot` ou `CreateSnapshots`. Si une balise est spécifiée, elle doit être de type `purpose=test`. Aucune autre balise n'est autorisée dans la demande.

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshot",
         "Resource":"*"
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateTags",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "aws:RequestTag/purpose":"test",
               "ec2:CreateAction":"CreateSnapshot"
            },
            "ForAllValues:StringEquals":{
```

```
                "aws:TagKeys":"purpose"
            }
        }
    }
]
}
```

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":"*"
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateTags",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "aws:RequestTag/purpose":"test",
               "ec2:CreateAction":"CreateSnapshots"
            },
            "ForAllValues:StringEquals":{
               "aws:TagKeys":"purpose"
            }
         }
      }
   ]
}
```

La stratégie suivante permet de créer des instantanés uniquement si le volume source est balisé avec `User:`*`username`* pour le client et que l'instantané lui-même est balisé avec `Environment:Dev` et `User:`*`username`*. Le client peut ajouter des balises supplémentaires à l'instantané.

```
{
   "Version":"2012-10-17",
   "Statement": [
           {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshot",
         "Resource":"arn:aws:ec2:us-east-1:123456789012:volume/*",
         "Condition":{
            "StringEquals":{
               "ec2:ResourceTag/User":"${aws:username}"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshot",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "aws:RequestTag/Environment":"Dev",
               "aws:RequestTag/User":"${aws:username}"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateTags",
```

```
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*"
      }
   ]
}
```

La stratégie suivante pour `CreateSnapshots` permet de créer des instantanés uniquement si le volume source est balisé avec `User:`*`username`* pour le client et que l'instantané lui-même est balisé avec `Environment:Dev` et `User:`*`username`*.

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":"arn:aws:ec2:us-east-1:*:instance/*",
 },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":"arn:aws:ec2:us-east-1:123456789012:volume/*",
         "Condition":{
            "StringEquals":{
               "ec2:ResourceTag/User":"${aws:username}"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateSnapshots",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "aws:RequestTag/Environment":"Dev",
               "aws:RequestTag/User":"${aws:username}"
            }
         }
      },
      {
         "Effect":"Allow",
         "Action":"ec2:CreateTags",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*"
      }
   ]
}
```

La stratégie suivante permet de supprimer un instantané uniquement s'il est balisé à l'aide de User:username pour le client.

```
{
   "Version":"2012-10-17",
   "Statement": [
            {
         "Effect":"Allow",
         "Action":"ec2:DeleteSnapshot",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "ec2:ResourceTag/User":"${aws:username}"
            }
         }
      }
   ]
```

```
}
```

La stratégie suivante permet à un client de créer un instantané mais l'empêche d'exécuter cette action si l'instantané créé comporte une clé de balise `value=stack`.

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":[
            "ec2:CreateSnapshot",
            "ec2:CreateTags"
         ],
         "Resource":"*"
      },
      {
         "Effect":"Deny",
         "Action":"ec2:CreateSnapshot",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "ForAnyValue:StringEquals":{
               "aws:TagKeys":"stack"
            }
         }
      }
   ]
}
```

La stratégie suivante permet à un client de créer des instantanés mais l'empêche d'exécuter cette action si les instantanés créés comportent une clé de balise `value=stack`.

```
{
   "Version":"2012-10-17",
   "Statement": [
      {
         "Effect":"Allow",
         "Action":[
            "ec2:CreateSnapshots",
            "ec2:CreateTags"
         ],
         "Resource":"*"
      },
      {
         "Effect":"Deny",
         "Action":"ec2:CreateSnapshots",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "ForAnyValue:StringEquals":{
               "aws:TagKeys":"stack"
            }
         }
      }
   ]
}
```

La stratégie suivante vous permet d'associer plusieurs actions dans une même stratégie. Vous pouvez uniquement créer un instantané (dans le contexte de `CreateSnapshots`) lorsque l'instantané est créé dans la région `us-east-1`. Vous pouvez uniquement créer des instantanés (dans le contexte de `CreateSnapshots`) lorsque les instantanés sont créés dans la région `us-east-1` et que le type d'instance est `t2*`.

```
{
   "Version":"2012-10-17",
   "Statement": [
          {
         "Effect":"Allow",
         "Action":[
            "ec2:CreateSnapshots",
            "ec2:CreateSnapshot",
            "ec2:CreateTags"
         ],
         "Resource": [
            "arn:aws:ec2:*:*:instance/*",
            "arn:aws:ec2:*:*:snapshot/*",
            "arn:aws:ec2:*:*:volume/*"
         ],
         "Condition":{
            "StringEqualsIgnoreCase": {
              "ec2:Region": "us-east-1"
            },
            "StringLikeIfExists": {
              "ec2:InstanceType": ["t2.*"]
            }
         }
      }
   ]
}
```

Exemple : Modification des paramètres d'autorisation pour des instantanés

La stratégie suivante permet de modifier un instantané uniquement s'il est balisé avec `User:`*`username`*, où *`username`* est le nom d'utilisateur du compte AWS du client. La demande échoue si cette condition n'est pas respectée.

```
{
   "Version":"2012-10-17",
   "Statement": [
           {
         "Effect":"Allow",
         "Action":"ec2: ModifySnapshotAttribute",
         "Resource":"arn:aws:ec2:us-east-1::snapshot/*",
         "Condition":{
            "StringEquals":{
               "ec2:ResourceTag/user-name":"${aws:username}"
            }
         }
      }
   ]
}
```

Lancement d'instances (RunInstances)

L'action d'API RunInstances lance un ou plusieurs Instances à la demande ou un ou plusieurs Instances Spot. `RunInstances` nécessite une AMI et crée une instance. Les utilisateurs peuvent spécifier une paire de clés et un groupe de sécurité dans la demande. Le lancement dans un VPC nécessite un sous-réseau et crée une interface réseau. Le lancement à partir d'une AMI basée sur des volumes Amazon EBS crée un volume. Par conséquent, l'utilisateur doit être autorisé à utiliser ces ressources Amazon EC2. Vous pouvez créer une déclaration de stratégie qui requiert que les utilisateurs spécifient un paramètre facultatif sur `RunInstances`, ou limitent les utilisateurs à certaines valeurs pour tel ou tel paramètre.

Pour de plus amples informations sur les autorisations au niveau des ressources requises pour lancer une instance, veuillez consulter Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur.

Par défaut, les utilisateurs ne sont pas autorisés à décrire, démarrer, arrêter ni résilier les instances obtenues. Une solution pour accorder aux utilisateurs l'autorisation de gérer les instances obtenues consiste à créer une balise spécifique pour chaque instance, puis à créer une déclaration qui leur permet de gérer les instances avec cette balise. Pour plus d'informations, consultez Utilisation d'instances (p. 989).

Ressources

- AMI (p. 1001)
- Types d'instance (p. 1002)
- Sous-réseaux (p. 1003)
- Volumes EBS (p. 1004)
- Balises (p. 1005)
- Balises dans un modèle de lancement (p. 1009)
- GPU Elastic (p. 1009)
- Modèles de lancement (p. 1010)

AMI

La stratégie suivante permet aux utilisateurs de lancer les instances en n'utilisant que les AMI spécifiés, `ami-9e1670f7` et `ami-45cf5c3c`. Les utilisateurs ne peuvent pas lancer une instance à l'aide d'autres AMI (à moins qu'une autre déclaration n'accorde aux utilisateurs l'autorisation de le faire) ni sur un sous-réseau.

```
{
   "Version": "2012-10-17",
   "Statement": [
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
        "arn:aws:ec2:region::image/ami-9e1670f7",
        "arn:aws:ec2:region::image/ami-45cf5c3c",
        "arn:aws:ec2:region:account:instance/*",
        "arn:aws:ec2:region:account:volume/*",
        "arn:aws:ec2:region:account:key-pair/*",
        "arn:aws:ec2:region:account:security-group/*",
        "arn:aws:ec2:region:account:subnet/*",
        "arn:aws:ec2:region:account:network-interface/*"
      ]
    }
   ]
}
```

A titre d'autre solution, la stratégie suivante permet aux utilisateurs de lancer les instances à partir de tous les AMI dont Amazon est propriétaire. L'élément `Condition` de la première déclaration+ teste si `ec2:Owner` est `amazon`. Les utilisateurs ne peuvent pas lancer une instance à l'aide d'autres AMI (à moins qu'une autre déclaration n'accorde aux utilisateurs l'autorisation de le faire) ni sur un sous-réseau.

```
{
   "Version": "2012-10-17",
   "Statement": [
         {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region::image/ami-*"
      ],
```

```
      "Condition": {
         "StringEquals": {
            "ec2:Owner": "amazon"
         }
      }
   },
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region:account:instance/*",
         "arn:aws:ec2:region:account:subnet/*",
         "arn:aws:ec2:region:account:volume/*",
         "arn:aws:ec2:region:account:network-interface/*",
         "arn:aws:ec2:region:account:key-pair/*",
         "arn:aws:ec2:region:account:security-group/*"
         ]
      }
   ]
}
```

Types d'instance

La stratégie suivante permet aux utilisateurs de lancer des instances uniquement à l'aide du type d'instance `t2.micro` ou `t2.small`, ce que vous pourriez faire pour contrôler les coûts. Les utilisateurs ne peuvent pas lancer d'instances plus grandes parce que l'élément `Condition` de la première déclaration teste si `ec2:InstanceType` est `t2.micro` ou `t2.small`.

```
{
   "Version": "2012-10-17",
   "Statement": [
        {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region:account:instance/*"
      ],
      "Condition": {
         "StringEquals": {
            "ec2:InstanceType": ["t2.micro", "t2.small"]
         }
      }
   },
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region::image/ami-*",
         "arn:aws:ec2:region:account:subnet/*",
         "arn:aws:ec2:region:account:network-interface/*",
         "arn:aws:ec2:region:account:volume/*",
         "arn:aws:ec2:region:account:key-pair/*",
         "arn:aws:ec2:region:account:security-group/*"
         ]
      }
   ]
}
```

Vous pouvez également créer une stratégie qui refuse aux utilisateurs l'autorisation de lancer des instances, à l'exception des types d'instance `t2.micro` et `t2.small`.

```
{
```

```
   "Version": "2012-10-17",
   "Statement": [
        {
      "Effect": "Deny",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region:account:instance/*"
      ],
      "Condition": {
         "StringNotEquals": {
            "ec2:InstanceType": ["t2.micro", "t2.small"]
         }
      }
   },
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region::image/ami-*",
         "arn:aws:ec2:region:account:network-interface/*",
         "arn:aws:ec2:region:account:instance/*",
         "arn:aws:ec2:region:account:subnet/*",
         "arn:aws:ec2:region:account:volume/*",
         "arn:aws:ec2:region:account:key-pair/*",
         "arn:aws:ec2:region:account:security-group/*"
         ]
      }
   ]
}
```

Sous-réseaux

La stratégie suivante permet aux utilisateurs de lancer les instances en n'utilisant que le sous-réseau spécifié, `subnet-12345678`. Le groupe ne peut pas lancer d'instance sur un autre sous-réseau (à moins qu'une autre déclaration n'accorde aux utilisateurs l'autorisation de le faire).

```
{
   "Version": "2012-10-17",
   "Statement": [
    {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
        "arn:aws:ec2:region:account:subnet/subnet-12345678",
        "arn:aws:ec2:region:account:network-interface/*",
        "arn:aws:ec2:region:account:instance/*",
        "arn:aws:ec2:region:account:volume/*",
        "arn:aws:ec2:region::image/ami-*",
        "arn:aws:ec2:region:account:key-pair/*",
        "arn:aws:ec2:region:account:security-group/*"
      ]
    }
   ]
}
```

Vous pouvez également créer une stratégie qui refuse aux utilisateurs l'autorisation de lancer une instance sur un autre sous-réseau. La déclaration agit ainsi en refusant l'autorisation de créer une interface réseau, à l'exception de l'emplacement où le sous-réseau `subnet-12345678` est spécifié. Ce refus se substitue à toute autre stratégie créée pour autoriser le lancement d'instances sur d'autres sous-réseaux.

```
{
   "Version": "2012-10-17",
```

```
   "Statement": [
         {
      "Effect": "Deny",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region:account:network-interface/*"
      ],
      "Condition": {
         "ArnNotEquals": {
            "ec2:Subnet": "arn:aws:ec2:region:account:subnet/subnet-12345678"
         }
      }
   },
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:region::image/ami-*",
         "arn:aws:ec2:region:account:network-interface/*",
         "arn:aws:ec2:region:account:instance/*",
         "arn:aws:ec2:region:account:subnet/*",
         "arn:aws:ec2:region:account:volume/*",
         "arn:aws:ec2:region:account:key-pair/*",
         "arn:aws:ec2:region:account:security-group/*"
         ]
      }
   ]
}
```

Volumes EBS

La stratégie suivante permet aux utilisateurs de lancer des instances uniquement si les volumes EBS pour l'instance sont chiffrés. L'utilisateur doit lancer une instance à partir d'une AMI qui a été créée avec des instantanés chiffrés afin de garantir le chiffrement du volume racine. N'importe quel volume supplémentaire que l'utilisateur attache à l'instance pendant le lancement doit aussi être chiffré.

```
{
    "Version": "2012-10-17",
    "Statement": [
                {
            "Effect": "Allow",
            "Action": "ec2:RunInstances",
            "Resource": [
                "arn:aws:ec2:*:*:volume/*"
            ],
            "Condition": {
                "Bool": {
                    "ec2:Encrypted": "true"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": "ec2:RunInstances",
            "Resource": [
                "arn:aws:ec2:*::image/ami-*",
                "arn:aws:ec2:*:*:network-interface/*",
                "arn:aws:ec2:*:*:instance/*",
                "arn:aws:ec2:*:*:subnet/*",
                "arn:aws:ec2:*:*:key-pair/*",
                "arn:aws:ec2:*:*:security-group/*"
            ]
        }
    ]
```

```
}
```

Balises

Baliser les instances lors de la création

La stratégie suivante permet aux utilisateurs de lancer des instances et d'attribuer des balises aux instances lors de la création. Pour les actions de création de ressources qui appliquent des balises, les utilisateurs doivent être autorisés à effectuer l'action `CreateTags`. La deuxième déclaration utilise la clé de condition `ec2:CreateAction` pour permettre aux utilisateurs de créer des balises uniquement dans le cadre de `RunInstances` et uniquement pour des instances. Les utilisateurs ne peuvent pas attribuer de balises aux ressources existantes, et ils ne peuvent pas attribuer de balises aux volumes à l'aide de la demande `RunInstances`.

Pour plus d'informations, consultez Octroi de l'autorisation de baliser les ressources lors de la création (p. 985).

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
         "ec2:RunInstances"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:us-east-1:123456789012:instance/*",
      "Condition": {
         "StringEquals": {
             "ec2:CreateAction" : "RunInstances"
          }
       }
    }
  ]
}
```

Baliser des instances et des volumes lors de la création avec des balises spécifiques

La stratégie suivante inclut la clé de condition `aws:RequestTag` qui exige aux utilisateurs d'attribuer des balises aux instances et aux volumes créés par `RunInstances` avec les balises `environment=production` et `purpose=webserver`. La clé de condition `aws:TagKeys` utilise le modificateur `ForAllValues` pour indiquer que seules les clés `environment` et `purpose` sont autorisées dans la demande (aucune autre balise ne peut être spécifiée). Si aucune balise n'est spécifiée dans la demande, la demande échoue.

```
{
  "Version": "2012-10-17",
  "Statement": [
   {
      "Effect": "Allow",
      "Action": [
         "ec2:RunInstances"
      ],
      "Resource": [
```

```
          "arn:aws:ec2:region::image/*",
          "arn:aws:ec2:region:account:subnet/*",
          "arn:aws:ec2:region:account:network-interface/*",
          "arn:aws:ec2:region:account:security-group/*",
          "arn:aws:ec2:region:account:key-pair/*"
        ]
      },
      {
        "Effect": "Allow",
        "Action": [
          "ec2:RunInstances"
        ],
        "Resource": [
            "arn:aws:ec2:region:account:volume/*",
            "arn:aws:ec2:region:account:instance/*"
        ],
        "Condition": {
          "StringEquals": {
              "aws:RequestTag/environment": "production" ,
              "aws:RequestTag/purpose": "webserver"
            },
            "ForAllValues:StringEquals": {
                "aws:TagKeys": ["environment","purpose"]
            }
         }
      },
      {
        "Effect": "Allow",
        "Action": [
          "ec2:CreateTags"
        ],
        "Resource": "arn:aws:ec2:region:account:*/*",
        "Condition": {
          "StringEquals": {
              "ec2:CreateAction" : "RunInstances"
            }
         }
      }
    ]
}
```

Baliser des instances et des volumes lors de la création avec au moins une balise spécifique

La stratégie suivante utilise le modificateur `ForAnyValue` sur la condition `aws:TagKeys` pour indiquer qu'au moins une balise doit être spécifiée dans la demande, et elle doit comporter la clé `environment` ou `webserver`. La balise doit être appliquée à la fois aux instances et aux volumes. Toutes les valeurs de balise peuvent être spécifiées dans la demande.

```
{
  "Version": "2012-10-17",
  "Statement": [
   {
      "Effect": "Allow",
      "Action": [
         "ec2:RunInstances"
      ],
      "Resource": [
         "arn:aws:ec2:region::image/*",
         "arn:aws:ec2:region:account:subnet/*",
         "arn:aws:ec2:region:account:network-interface/*",
         "arn:aws:ec2:region:account:security-group/*",
         "arn:aws:ec2:region:account:key-pair/*"
      ]
    },
```

```
    {
      "Effect": "Allow",
      "Action": [
          "ec2:RunInstances"
      ],
      "Resource": [
          "arn:aws:ec2:region:account:volume/*",
          "arn:aws:ec2:region:account:instance/*"
      ],
      "Condition": {
          "ForAnyValue:StringEquals": {
              "aws:TagKeys": ["environment","webserver"]
          }
       }
    },
    {
      "Effect": "Allow",
      "Action": [
          "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:region:account:*/*",
      "Condition": {
          "StringEquals": {
              "ec2:CreateAction" : "RunInstances"
          }
       }
    }
  ]
}
```

Si les instances sont balisées lors de la création, elles doivent être balisées avec une balise spécifique

Dans la stratégie suivante, les utilisateurs ne doivent pas spécifier les balises dans la demande, mais s'ils le font, la balise doit être `purpose=test`. Aucune autre balise n'est autorisée. Les utilisateurs peuvent appliquer des balises à n'importe quelle ressource pouvant être balisée dans la demande `RunInstances`.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
         "ec2:RunInstances"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:region:account:*/*",
      "Condition": {
         "StringEquals": {
             "aws:RequestTag/purpose": "test",
             "ec2:CreateAction" : "RunInstances"
          },
          "ForAllValues:StringEquals": {
              "aws:TagKeys": "purpose"
          }
       }
    }
  ]
}
```

Pour interdire toute personne appelée balise lors de la création pour RunInstances

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*",
                "arn:aws:ec2:us-east-1:*:spot-instances-request/*"
            ]
        },
        {
            "Sid": "VisualEditor0",
            "Effect": "Deny",
            "Action": "ec2:CreateTags",
            "Resource": "*"
        }
    ]
}
```

Autoriser uniquement les balises spécifiques pour spot-instances-request. Incohérence surprise numéro 2 entre en jeu ici. Dans des circonstances normales, si vous ne spécifiez aucune balise, vous n'êtes pas authentifié. Dans le cas de spot-instances-request, cette stratégie ne sera pas évaluée s'il n'y a pas de balises spot-instances-request, donc une demande Spot à l'exécution sans balise réussira.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*",
            ]
        },
        {
            "Sid": "VisualEditor0",
            "Effect": "Allow",
            "Action": "ec2:RunInstances",
            "Resource": "arn:aws:ec2:us-east-1:*:spot-instances-request/*",
            "Condition": {
                "StringEquals": {
                    "aws:RequestTag/environment": "production"
                }
```

```
            }
        }
    ]
}
```

Balises dans un modèle de lancement

Dans l'exemple suivant, les utilisateurs peuvent lancer des instances, mais uniquement s'ils utilisent un modèle de lancement spécifique (`lt-09477bcd97b0d310e`). La clé de condition `ec2:IsLaunchTemplateResource` empêche les utilisateurs de remplacer les ressources spécifiées dans le modèle de lancement. La seconde partie de la déclaration permet aux utilisateurs de baliser les instances à la création. Cette partie de la déclaration est nécessaire si des balises sont spécifiées pour l'instance dans le modèle de lancement.

```
{
  "Version": "2012-10-17",
  "Statement": [
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": "*",
      "Condition": {
         "ArnLike": {
             "ec2:LaunchTemplate": "arn:aws:ec2:region:account:launch-template/
lt-09477bcd97b0d310e"
          },
          "Bool": {
             "ec2:IsLaunchTemplateResource": "true"
          }
       }
    },
    {
      "Effect": "Allow",
      "Action": [
         "ec2:CreateTags"
      ],
      "Resource": "arn:aws:ec2:region:account:instance/*",
      "Condition": {
         "StringEquals": {
             "ec2:CreateAction" : "RunInstances"
          }
       }
    }
  ]
}
```

GPU Elastic

Dans la stratégie suivante, les utilisateurs peuvent lancer une instance et spécifier un GPU Elastic à attacher à l'instance. Les utilisateurs peuvent lancer des instances dans n'importe quelle région, mais ils peuvent uniquement attacher un GPU Elastic lors d'un lancement dans la région `us-east-2`.

La clé de condition `ec2:ElasticGpuType` utilise le modificateur `ForAnyValue` pour indiquer que seules les types de GPU Elastic `eg1.medium` et `eg1.large` sont autorisés dans la demande.

```
{
    "Version": "2012-10-17",
    "Statement": [
             {
            "Effect": "Allow",
            "Action": [
```

```
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:*:account:elastic-gpu/*"
            ],
            "Condition": {
                "StringEquals": {
                    "ec2:Region": "us-east-2"
                },
                "ForAnyValue:StringLike": {
                    "ec2:ElasticGpuType": [
                        "eg1.medium",
                        "eg1.large"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": "ec2:RunInstances",
            "Resource": [
                "arn:aws:ec2:*::image/ami-*",
                "arn:aws:ec2:*:account:network-interface/*",
                "arn:aws:ec2:*:account:instance/*",
                "arn:aws:ec2:*:account:subnet/*",
                "arn:aws:ec2:*:account:volume/*",
                "arn:aws:ec2:*:account:key-pair/*",
                "arn:aws:ec2:*:account:security-group/*"
            ]
        }
    ]
}
```

Modèles de lancement

Dans l'exemple suivant, les utilisateurs peuvent lancer des instances, mais uniquement s'ils utilisent un modèle de lancement spécifique (`lt-09477bcd97b0d310e`). Les utilisateurs peuvent remplacer des paramètres dans le modèle de lancement en spécifiant dans l'action `RunInstances`.

```
{
  "Version": "2012-10-17",
  "Statement": [
         {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": "*",
      "Condition": {
         "ArnLike": {
             "ec2:LaunchTemplate": "arn:aws:ec2:region:account:launch-template/
lt-09477bcd97b0d310e"
          }
       }
    }
  ]
}
```

Dans cet exemple, les utilisateurs peuvent lancer des instances uniquement s'ils utilisent un modèle de lancement. La stratégie utilise la clé de condition `ec2:IsLaunchTemplateResource` pour empêcher les utilisateurs de remplacer les ARN préexistants dans le modèle de lancement.

```
{
  "Version": "2012-10-17",
  "Statement": [
```

```
        {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": "*",
      "Condition": {
         "ArnLike": {
             "ec2:LaunchTemplate": "arn:aws:ec2:region:account:launch-template/*"
          },
          "Bool": {
             "ec2:IsLaunchTemplateResource": "true"
          }
       }
    }
  ]
}
```

Dans l'exemple suivant, une stratégie permet aux utilisateurs de lancer des instances, mais uniquement s'ils utilisent un modèle de lancement. Les utilisateurs ne peuvent pas remplacer les paramètres du sous-réseau et de l'interface réseau dans la demande ; ceux-ci ne peuvent être spécifiés que dans le modèle de lancement. La première partie de la déclaration utilise l'élément NotResource pour autoriser toutes les autres ressources à l'exception des sous-réseaux et des interfaces réseau. La seconde partie de la déclaration autorise les ressources des sous-réseaux et des interfaces réseau, mais uniquement si elles proviennent du modèle de lancement.

```
{
  "Version": "2012-10-17",
  "Statement": [
        {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "NotResource": ["arn:aws:ec2:region:account:subnet/*",
                      "arn:aws:ec2:region:account:network-interface/*" ],
      "Condition": {
         "ArnLike": {
             "ec2:LaunchTemplate": "arn:aws:ec2:region:account:launch-template/*"
          }
       }
    },
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": ["arn:aws:ec2:region:account:subnet/*",
                   "arn:aws:ec2:region:account:network-interface/*" ],
      "Condition": {
         "ArnLike": {
             "ec2:LaunchTemplate": "arn:aws:ec2:region:account:launch-template/*"
          },
          "Bool": {
             "ec2:IsLaunchTemplateResource": "true"
          }
       }
    }
  ]
}
```

Dans l'exemple suivant, les utilisateurs sont autorisés à lancer des instances uniquement s'ils utilisent un modèle de lancement et seulement si celui-ci contient la balise `Purpose=Webservers`. Les utilisateurs ne peuvent pas remplacer les paramètres de modèle de lancement dans l'action `RunInstances`.

```
{
  "Version": "2012-10-17",
  "Statement": [
```

```
        {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "NotResource": "arn:aws:ec2:region:account:launch-template/*",
      "Condition": {
         "ArnLike": {
             "ec2:LaunchTemplate": "arn:aws:ec2:region:account:launch-template/*"
          },
         "Bool": {
             "ec2:IsLaunchTemplateResource": "true"
          }
       }
    },
    {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": "arn:aws:ec2:region:account:launch-template/*",
      "Condition": {
       "StringEquals": {
           "ec2:ResourceTag/Purpose": "Webservers"
        }
       }
     }
  ]
}
```

Utilisation de Instances Spot

Vous pouvez utiliser l'action RunInstances pour créer des demandes instance Spot et baliser les demandes instance Spot lors de la création. La ressource à spécifier pour RunInstances est `spot-instances-request`.

La ressource `spot-instances-request` est évaluée dans la stratégie IAM comme suit :

- Si vous ne balisez pas une demande instance Spot lors de la création, Amazon EC2 n'évalue pas la ressource `spot-instances-request` dans l'instruction RunInstances.
- Si vous balisez une demande instance Spot lors de la création, Amazon EC2 évalue la ressource `spot-instances-request` dans l'instruction RunInstances.

Par conséquent, pour la ressource `spot-instances-request`, les règles suivantes s'appliquent à la stratégie IAM :

- Si vous utilisez RunInstances pour créer une demande instance Spot et que vous n'avez pas l'intention de baliser la demande instance Spot lors de la création, vous n'avez pas besoin d'autoriser explicitement la ressource `spot-instances-request` ; l'appel réussira.
- Si vous utilisez RunInstances pour créer une demande instance Spot et que vous avez l'intention de baliser la demande instance Spot lors de la création, vous devez inclure la ressource `spot-instances-request` dans l'instruction RunInstances allow, sinon l'appel échouera.
- Si vous utilisez RunInstances pour créer une demande instance Spot et que vous avez l'intention de baliser la demande instance Spot lors de la création, vous devez spécifier la ressource `spot-instances-request` ou le caractère générique `*` dans l'instruction CreateTags allow, sinon l'appel échouera.

Vous pouvez demander Instances Spot en utilisant RunInstances ou RequestSpotInstances. L'exemple suivant de stratégies IAM s'applique uniquement lorsque vous demandez Instances Spot à l'aide de RunInstances.

Exemple : Demander Instances Spot en utilisant RunInstances

La stratégie suivante permet aux utilisateurs de demander Instances Spot à l'aide de l'action RunInstances. La ressource `spot-instances-request`, qui est créée par RunInstances, demande Instances Spot.

> **Note**
>
> Pour utiliser RunInstances pour créer des demandes instance Spot, vous pouvez omettre `spot-instances-request` de la liste `Resource` si vous n'avez pas l'intention de baliser les demandes instance Spot lors de la création. En effet, Amazon EC2 n'évalue pas la ressource `spot-instances-request` dans l'instruction RunInstances si la demande instance Spot n'est pas balisée lors de la création.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*",
                "arn:aws:ec2:us-east-1:*:spot-instances-request/*"
            ]
        }
    ]
}
```

> **Warning**
>
> NON PRIS EN CHARGE – Exemple : Refuser aux utilisateurs l'autorisation de demander Instances Spot à l'aide de RunInstances
> La stratégie suivante n'est pas prise en charge pour la ressource `spot-instances-request`. La stratégie suivante vise à donner aux utilisateurs l'autorisation de lancer Instances à la demande, mais à refuser aux utilisateurs l'autorisation de demander Instances Spot. La ressource `spot-instances-request`, qui est créée par RunInstances, est la ressource qui demande Instances Spot. La deuxième instruction est destinée à refuser l'action RunInstances pour la ressource `spot-instances-request`. Toutefois, cette condition n'est pas prise en charge car Amazon EC2 n'évalue pas la ressource `spot-instances-request` dans l'instruction RunInstances si la demande instance Spot n'est pas balisée lors de la création.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
```

```
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*"
            ]
        },
        {
            "Sid": "DenySpotInstancesRequests - NOT SUPPORTED - DO NOT USE!",
            "Effect": "Deny",
            "Action": "ec2:RunInstances",
            "Resource": "arn:aws:ec2:us-east-1:*:spot-instances-request/*"
        }
    ]
}
```

Exemple : Baliser les demandes instance Spot lors de la création

La stratégie suivante permet aux utilisateurs de baliser toutes les ressources créées lors du lancement de l'instance. La première instruction permet à RunInstances de créer les ressources répertoriées. La ressource `spot-instances-request`, qui est créée par RunInstances, est la ressource qui demande Instances Spot. La deuxième instruction fournit un caractère générique `*` pour permettre à toutes les ressources d'être balisées lorsqu'elles sont créées au lancement de l'instance.

> **Note**
>
> Si vous balisez une demande instance Spot lors de la création, Amazon EC2 évalue la ressource `spot-instances-request` dans l'instruction RunInstances. Par conséquent, vous devez explicitement autoriser la ressource `spot-instances-request` pour l'action RunInstances, sinon l'appel échouera.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*",
                "arn:aws:ec2:us-east-1:*:spot-instances-request/*"
            ]
        },
        {
            "Sid": "TagResources",
            "Effect": "Allow",
            "Action": "ec2:CreateTags",
            "Resource": "*"
        }
    ]
}
```

Exemple : Refuser la balise lors de la création pour les demandes instance Spot

La stratégie suivante refuse aux utilisateurs l'autorisation de baliser les ressources créées lors du lancement de l'instance.

La première instruction permet à RunInstances de créer les ressources répertoriées. La ressource `spot-instances-request`, qui est créée par RunInstances, est la ressource qui demande Instances Spot.
La deuxième instruction fournit un caractère générique `*` pour refuser toutes les ressources en cours de balisage lorsqu'elles sont créées au lancement de l'instance. Si `spot-instances-request` ou une autre ressource est balisée lors de la création, l'appel RunInstances échouera.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*",
                "arn:aws:ec2:us-east-1:*:spot-instances-request/*"
            ]
        },
        {
            "Sid": "DenyTagResources",
            "Effect": "Deny",
            "Action": "ec2:CreateTags",
            "Resource": "*"
        }
    ]
}
```

Warning

NON PRIS EN CHARGE – Exemple : Autoriser la création d'une demande instance Spot uniquement si une balise spécifique lui est attribuée
La stratégie suivante n'est pas prise en charge pour la ressource `spot-instances-request`.
La stratégie suivante vise à accorder à RunInstances l'autorisation de créer une demande instance Spot uniquement si la demande est balisée avec une balise spécifique.
La première instruction permet à RunInstances de créer les ressources répertoriées.
La deuxième instruction est destinée à accorder aux utilisateurs l'autorisation de créer une demande instance Spot uniquement si la demande a la balise `environment=production`.
Si cette condition est appliquée à d'autres ressources créées par RunInstances, la spécification d'aucune balise entraîne une erreur `Unauthenticated`. Toutefois, si aucune balise n'est spécifiée pour la demande instance Spot, Amazon EC2 n'évalue pas la ressource `spot-instances-request` dans l'instruction RunInstances, ce qui entraîne la création de demandes instance Spot non balisées par RunInstances.
Notez que la spécification d'une balise autre que `environment=production` entraîne une erreur `Unauthenticated`, car si un utilisateur marque une demande instance Spot, Amazon EC2 évalue la ressource `spot-instances-request` dans l'instruction RunInstances.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
```

```
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*"
            ]
        },
        {
            "Sid": "RequestSpotInstancesOnlyIfTagIs_environment=production - NOT
 SUPPORTED - DO NOT USE!",
            "Effect": "Allow",
            "Action": "ec2:RunInstances",
            "Resource": "arn:aws:ec2:us-east-1:*:spot-instances-request/*",
            "Condition": {
                "StringEquals": {
                    "aws:RequestTag/environment": "production"
                }
            }
        },
        {
            "Sid": "TagResources",
            "Effect": "Allow",
            "Action": "ec2:CreateTags",
            "Resource": "*"
        }

    ]
}
```

Exemple : Refuser la création d'une demande instance Spot si une balise spécifique lui est attribuée

La stratégie suivante refuse à RunInstances l'autorisation de créer une demande instance Spot si la demande est balisée avec `environment=production`.

La première instruction permet à RunInstances de créer les ressources répertoriées.

La deuxième instruction refuse aux utilisateurs l'autorisation de créer une demande instance Spot si la demande a la balise `environment=production`. La spécification `environment=production` en tant que balise entraîne une erreur `Unauthenticated`. La spécification d'autres balises ou l'absence de balises entraînera la création d'une demande instance Spot.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowRun",
            "Effect": "Allow",
            "Action": [
                "ec2:RunInstances"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1::image/*",
                "arn:aws:ec2:us-east-1:*:subnet/*",
                "arn:aws:ec2:us-east-1:*:network-interface/*",
                "arn:aws:ec2:us-east-1:*:security-group/*",
                "arn:aws:ec2:us-east-1:*:key-pair/*",
                "arn:aws:ec2:us-east-1:*:volume/*",
                "arn:aws:ec2:us-east-1:*:instance/*",
                "arn:aws:ec2:us-east-1:*:spot-instances-request/*"
```

```
            ]
        },
        {
            "Sid": "DenySpotInstancesRequests",
            "Effect": "Deny",
            "Action": "ec2:RunInstances",
            "Resource": "arn:aws:ec2:us-east-1:*:spot-instances-request/*",
            "Condition": {
                "StringEquals": {
                    "aws:RequestTag/environment": "production"
                }
            }
        },
        {
            "Sid": "TagResources",
            "Effect": "Allow",
            "Action": "ec2:CreateTags",
            "Resource": "*"
        }
    ]
}
```

Exemple : Utilisation d'Instances réservées

La stratégie suivante autorise les utilisateurs à afficher, modifier et acheter les Instances réservées de votre compte.

Il n'est pas possible de définir des autorisations au niveau des ressources pour les Instances réservées individuelles. Cette stratégie signifie que les utilisateurs ont accès à toutes les Instances réservées du compte.

L'élément `Resource` utilise un caractère générique * pour indiquer que les utilisateurs peuvent spécifier toutes les ressources avec l'action. Dans ce cas, ils peuvent afficher et modifier toutes les Instances réservées du compte. Ils peuvent aussi acheter des Instances réservées à l'aide des informations d'identification du compte. Le caractère générique * est également nécessaire dans les cas où l'action d'API ne prend pas en charge les autorisations au niveau des ressources.

```
{
   "Version": "2012-10-17",
   "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeReservedInstances",
        "ec2:ModifyReservedInstances",
        "ec2:PurchaseReservedInstancesOffering",
        "ec2:DescribeAvailabilityZones",
        "ec2:DescribeReservedInstancesOfferings"
      ],
      "Resource": "*"
    }
   ]
}
```

Pour permettre aux utilisateurs d'afficher et de modifier les Instances réservées de votre compte, mais pas d'acheter de nouvelles Instances réservées.

```
{
  "Version": "2012-10-17",
   "Statement": [
    {
```

```
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeReservedInstances",
        "ec2:ModifyReservedInstances",
        "ec2:DescribeAvailabilityZones"
      ],
      "Resource": "*"
    }
  ]
}
```

Exemple : Balisage de vos ressources

La stratégie suivante permet aux utilisateurs d'utiliser l'action `CreateTags` pour appliquer des balises à une instance uniquement si la balise contient la clé `environment` et la valeur `production`. Le modificateur `ForAllValues` est utilisé avec la clé de condition `aws:TagKeys` pour indiquer que seule la clé `environment` est autorisé dans la demande (aucune autre balise n'est autorisée). L'utilisateur ne peut pas attribuer de balises à d'autres types de ressource.

```
{
    "Version": "2012-10-17",
    "Statement": [
              {
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags"
            ],
            "Resource": "arn:aws:ec2:region:account:instance/*",
            "Condition": {
                "StringEquals": {
                    "aws:RequestTag/environment": "production"
                },
                "ForAllValues:StringEquals": {
                    "aws:TagKeys": [
                        "environment"
                    ]
                }
            }
        }
    ]
}
```

La stratégie suivante permet aux utilisateurs d'attribuer des balise à n'importe quelle ressource pouvant être balisée qui possède déjà une balise avec une clé de `owner` et une valeur du nom d'utilisateur IAM. En outre, les utilisateurs doivent spécifier une balise avec une clé de `anycompany:environment-type` et une valeur `test` ou `prod` dans la demande. Les utilisateurs peuvent spécifier des balises supplémentaires dans la demande.

```
{
    "Version": "2012-10-17",
    "Statement": [
               {
            "Effect": "Allow",
            "Action": [
                "ec2:CreateTags"
            ],
            "Resource": "arn:aws:ec2:region:account:*/*",
            "Condition": {
                "StringEquals": {
                    "aws:RequestTag/anycompany:environment-type": ["test","prod"],
```

```
                    "ec2:ResourceTag/owner": "${aws:username}"
                }
            }
        }
    ]
}
```

Vous pouvez créer une stratégie IAM qui permet aux utilisateurs de supprimer des balises spécifiques pour une ressource. Par exemple, la stratégie suivante permet aux utilisateurs de supprimer les balises pour un volume si les clés de balise spécifiées dans la demande sont `environment` ou `cost-center`. N'importe quelle valeur peut être spécifiée pour la balise, mais la clé de balise doit correspondre à l'une des clés spécifiées.

> **Note**
>
> Si vous supprimez une ressource, toutes les balises associées à celle-ci sont également supprimées. Les utilisateurs n'ont pas besoin d'être autorisés à effectuer l'action `ec2:DeleteTags` pour supprimer une ressource comportant des balises ; ils doivent seulement être autorisés à effectuer l'action de suppression.

```
{
  "Version": "2012-10-17",
  "Statement": [
       {
      "Effect": "Allow",
      "Action": "ec2:DeleteTags",
      "Resource": "arn:aws:ec2:us-east-1:123456789012:volume/*",
      "Condition": {
        "ForAllValues:StringEquals": {
          "aws:TagKeys": ["environment","cost-center"]
        }
      }
    }
  ]
}
```

Cette stratégie permet aux utilisateurs de supprimer uniquement la balise `environment=prod` sur n'importe quelle ressource et uniquement si la ressource porte déjà une balise avec une clé de `owner` et une valeur du nom d'utilisateur IAM. Les utilisateurs ne peuvent pas supprimer d'autres balises pour une ressource.

```
{
  "Version": "2012-10-17",
  "Statement": [
      {
      "Effect": "Allow",
      "Action": [
        "ec2:DeleteTags"
      ],
      "Resource": "arn:aws:ec2:region:account:*/*",
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/environment": "prod",
          "ec2:ResourceTag/owner": "${aws:username}"
        },
        "ForAllValues:StringEquals": {
          "aws:TagKeys": ["environment"]
        }
      }
    }
  ]
```

```
}
```

Exemple : Utilisation des rôles IAM

La stratégie suivante permet aux utilisateurs d'attacher, de remplacer et de détacher un rôle IAM pour les instances ayant la balise `department=test`. Le remplacement ou le détachement d'un rôle IAM nécessite un ID d'association. Par conséquent, la stratégie accorde également aux utilisateurs l'autorisation d'utiliser l'action `ec2:DescribeIamInstanceProfileAssociations`.

Les utilisateurs IAM doivent être autorisés à utiliser l'action `iam:PassRole` pour transmettre le rôle à l'instance.

```
{
  "Version": "2012-10-17",
  "Statement": [
       {
      "Effect": "Allow",
      "Action": [
         "ec2:AssociateIamInstanceProfile",
         "ec2:ReplaceIamInstanceProfileAssociation",
         "ec2:DisassociateIamInstanceProfile"
      ],
      "Resource": "arn:aws:ec2:region:account:instance/*",
      "Condition": {
        "StringEquals": {
          "ec2:ResourceTag/department":"test"
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": "ec2:DescribeIamInstanceProfileAssociations",
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": "iam:PassRole",
      "Resource": "*"
    }
  ]
}
```

La stratégie suivante permet aux utilisateurs d'attacher, de remplacer et de détacher un rôle IAM pour une instance. Les utilisateurs ne peuvent attacher ou remplacer que des rôles IAM dont les noms commencent par `TestRole-`. Pour l'action `iam:PassRole`, veillez à indiquer le nom du rôle IAM et non celui du profil d'instance (si ces noms ne sont pas identiques). Pour plus d'informations, consultez Profils d'instance (p. 1035).

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ec2:AssociateIamInstanceProfile",
                "ec2:ReplaceIamInstanceProfileAssociation"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
```

```
            "Action": "ec2:DescribeIamInstanceProfileAssociations",
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "iam:PassRole",
            "Resource": "arn:aws:iam::account:role/TestRole-*"
        }
    ]
}
```

Exemple : Utilisation des tables de routage

La stratégie suivante permet aux utilisateurs d'ajouter, de supprimer et de remplacer des routes pour les tables de routage associées au VPC `vpc-ec43eb89` uniquement. Pour spécifier un VPC pour la clé de condition `ec2:Vpc`, vous devez spécifier l'ARN complet du VPC.

```
{
    "Version": "2012-10-17",
    "Statement": [
              {
            "Effect": "Allow",
            "Action": [
                "ec2:DeleteRoute",
                "ec2:CreateRoute",
                "ec2:ReplaceRoute"
            ],
            "Resource": [
                "arn:aws:ec2:region:account:route-table/*"
            ],
            "Condition": {
                "StringEquals": {
                    "ec2:Vpc": "arn:aws:ec2:region:account:vpc/vpc-ec43eb89"
                }
            }
        }
    ]
}
```

Exemple : Octroi à une instance spécifique de l'autorisation de visualiser des ressources dans d'autres services AWS

Voici un exemple de stratégie que vous pouvez attacher à un rôle IAM. La stratégie permet à une instance de visualiser des ressources dans divers services AWS. Elle utilise la clé de condition `ec2:SourceInstanceARN` pour spécifier que l'instance dont émane la demande doit être l'instance `i-093452212644b0dd6`. Si le même rôle IAM est associé à une autre instance, l'autre instance ne peut effectuer aucune de ces actions.

La clé `ec2:SourceInstanceARN` est une clé de condition à l'échelle d'AWS. Par conséquent, elle peut être utilisée pour les actions d'autres services, et non pas seulement pour Amazon EC2.

```
{
    "Version": "2012-10-17",
    "Statement": [
              {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeVolumes",
                "s3:ListAllMyBuckets",
                "dynamodb:ListTables",
```

```
                "rds:DescribeDBInstances"
            ],
            "Resource": [
                "*"
            ],
            "Condition": {
                "ArnEquals": {
                    "ec2:SourceInstanceARN": "arn:aws:ec2:region:account:instance/
i-093452212644b0dd6"
                }
            }
        }
    ]
}
```

Exemple : Utilisation des modèles de lancement

La stratégie suivante permet aux utilisateurs de créer une version du modèle de lancement et de modifier un modèle de lancement, mais uniquement pour un modèle spécifique (`lt-09477bcd97b0d3abc`). Les utilisateurs ne peuvent pas utiliser d'autres modèles de lancement.

```
{
  "Version": "2012-10-17",
  "Statement": [
   {
      "Action": [
        "ec2:CreateLaunchTemplateVersion",
        "ec2:ModifyLaunchTemplate"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:ec2:region:account:launch-template/lt-09477bcd97b0d3abc"
    }
  ]
}
```

La stratégie suivante permet aux utilisateurs de supprimer un modèle de lancement et une version du modèle de lancement, sous réserve que le modèle de lancement contienne la balise `Purpose=Testing`.

```
{
  "Version": "2012-10-17",
  "Statement": [
       {
      "Action": [
        "ec2:DeleteLaunchTemplate",
        "ec2:DeleteLaunchTemplateVersions"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws:ec2:region:account:launch-template/*",
      "Condition": {
        "StringEquals": {
          "ec2:ResourceTag/Purpose": "Testing"
        }
      }
    }
  ]
}
```

Utilisation des métadonnées d'instance

Les stratégies suivantes garantissent que les utilisateurs peuvent uniquement récupérer les métadonnées d'instance (p. 701) à l'aide d'Service des métadonnées d'instance Version 2 (IMDSv2). Vous pouvez

combiner les quatre stratégies suivantes en une seule stratégie avec quatre instructions. Une fois combinée, vous pouvez l'utiliser en tant que stratégie de contrôle de service (SCP). Elle peut fonctionner aussi bien qu'une stratégie de refus que vous appliquez à une stratégie IAM existante (en retirant et en limitant les autorisations existantes), ou qu'une stratégie de contrôle de service appliquée globalement sur un compte, une unité organisationnelle ou l'ensemble d'une organisation.

Note

Les stratégies des options de métadonnées RunInstances suivantes doivent être utilisées avec une stratégie qui accorde au mandataire les autorisations pour lancer une instance avec RunInstances. Si le principal ne dispose pas également d'autorisations RunInstances, il ne pourra pas lancer une instance. Pour de plus amples informations, veuillez consulter les stratégies dans Utilisation d'instances (p. 989) et Lancement d'instances (RunInstances) (p. 1000).

Important

Si vous utilisez des groupes Auto Scaling et que vous devez exiger l'utilisation d'IMDSv2 sur toutes les nouvelles instances, vos groupes Auto Scaling doivent utiliser des modèles de lancement.
Lorsqu'un groupe Auto Scaling utilise un modèle de lancement, les autorisations `ec2:RunInstances` du mandataire IAM sont vérifiées lors de la création d'un nouveau groupe Auto Scaling. Elles sont également vérifiées lorsqu'un groupe Auto Scaling existant est mis à jour pour utiliser un nouveau modèle de lancement ou une nouvelle version d'un modèle de lancement. Les restrictions relatives à l'utilisation d'IMDSv1 sur les mandataires IAM pour `RunInstances` sont uniquement vérifiées lorsqu'un groupe Auto Scaling utilisant un modèle de lancement est créé ou mis à jour. Pour un groupe Auto Scaling configuré pour utiliser le modèle de lancement `Latest` ou `Default`, les autorisations ne sont pas vérifiées lors de la création d'une nouvelle version du modèle de lancement. Pour que les autorisations soient vérifiées, vous devez configurer le groupe Auto Scaling pour qu'il utilise une version spécifique du modèle de lancement.

Pour appliquer l'utilisation d'IMDSv2 sur des instances lancées par des groupes Auto Scaling, les étapes supplémentaires suivantes sont requises :

1. Désactivez l'utilisation des configurations de lancement pour tous les comptes de votre organisation à l'aide des stratégies de contrôle de service (SCP) ou des limites d'autorisations IAM pour les nouvelles entités créées. Pour les mandataires IAM existants disposant d'autorisations de groupe Auto Scaling, mettez à jour leurs stratégies associées avec cette clé de condition. Pour désactiver l'utilisation des configurations de lancement, créez ou modifiez la stratégie SCP, les limites d'autorisations ou la stratégie IAM avec la clé de condition `"autoscaling:LaunchConfigurationName"` avec la valeur spécifiée comme `null`.
2. Pour les nouveaux modèles de lancement, configurez les options de métadonnées d'instance dans le modèle de lancement. Pour les modèles de lancement existants, créez une nouvelle version du modèle de lancement et configurez les options de métadonnées d'instance dans la nouvelle version.
3. Dans la stratégie donnant à tout mandataire l'autorisation d'utiliser un modèle de lancement, restreignez l'association de `$latest` et de `$default` en spécifiant `"autoscaling:LaunchTemplateVersionSpecified": "true"`. En restreignant l'utilisation à une version spécifique d'un modèle de lancement, vous pouvez vous assurer que les nouvelles instances seront lancées à l'aide de la version dans laquelle les options de métadonnées d'instance sont configurées. Pour plus d'informations, consultez LaunchTemplateSpecification dans le Référence de l'API Amazon EC2 Auto Scaling, en particulier le paramètre `Version`.
4. Pour un groupe Auto Scaling qui utilise une configuration de lancement, remplacez la configuration de lancement par un modèle de lancement. Pour plus d'informations, voir Remplacement d'une configuration de lancement par un modèle de lancement dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.

5. Pour un groupe Auto Scaling qui utilise un modèle de lancement, assurez-vous qu'il utilise un nouveau modèle de lancement avec les options de métadonnées d'instance configurées ou qu'il utilise une nouvelle version du modèle de lancement actuel avec les options de métadonnées d'instance configurées. Pour plus d'informations, consultez update-auto-scaling-group dans le AWS CLI Command Reference.

Exemples

- Exigence d'utilisation d'IMDSv2 (p. 1024)
- Spécification d'une durée de vie (hop limit) maximale (p. 1024)
- Restriction des personnes habilitées à modifier les options de métadonnées d'instance (p. 1025)
- Exigence de récupération des informations d'identification de rôle à partir d'IMDSv2 (p. 1025)

Exigence d'utilisation d'IMDSv2

La stratégie suivante indique que vous ne pouvez pas appeler l'API RunInstances, sauf si l'instance est également inscrite pour exiger l'utilisation d'IMDSv2 (indiqué par `"ec2:MetadataHttpTokens": "required"`). Si vous n'indiquez pas que l'instance exige IMDSv2, vous obtenez une erreur `UnauthorizedOperation` lorsque vous appelez l'API RunInstances.

```
{
    "Version": "2012-10-17",
    "Statement": [
                {
            "Sid": "RequireImdsV2",
            "Effect": "Deny",
            "Action": "ec2:RunInstances",
            "Resource": "arn:aws:ec2:*:*:instance/*",
            "Condition": {
                "StringNotEquals": {
                    "ec2:MetadataHttpTokens": "required"
                }
            }
        }
    ]
}
```

Spécification d'une durée de vie (hop limit) maximale

La stratégie suivante indique que vous ne pouvez pas appeler l'API RunInstances, sauf si vous spécifiez également une durée de vie (hop limit) inférieure à 3. Si vous échouez, vous obtenez l'erreur `UnauthorizedOperation` lorsque vous appelez l'API RunInstances.

```
{
    "Version": "2012-10-17",
    "Statement": [
                {
            "Sid": "MaxImdsHopLimit",
            "Effect": "Deny",
            "Action": "ec2:RunInstances",
            "Resource": "arn:aws:ec2:*:*:instance/*",
            "Condition": {
                "NumericGreaterThan": {
                    "ec2:MetadataHttpPutResponseHopLimit": "3"
                }
            }
        }
    ]
```

```
}
```

Restriction des personnes habilitées à modifier les options de métadonnées d'instance

La stratégie suivante supprime la possibilité pour la population générale d'administrateurs de modifier les options de métadonnées d'instance et autorise uniquement les utilisateurs dotés du rôle `ec2-imds-admins` à apporter des modifications. Si un principal autre que le rôle `ec2-imds-admins` tente d'appeler l'API ModifyInstanceMetadataOptions, il obtient l'erreur `UnauthorizedOperation`. Cette instruction peut être utilisée pour contrôler l'utilisation de l'API ModifyInstanceMetadataOptions ; il n'existe actuellement aucun contrôle d'accès (conditions) précis pour l'API ModifyInstanceMetadataOptions.

```
{
    "Version": "2012-10-17",
    "Statement": [
                {
            "Sid": "AllowOnlyImdsAdminsToModifySettings",
            "Effect": "Deny",
            "Action": "ec2:ModifyInstanceMetadataOptions",
            "Resource": "*",
            "Condition": {
                "StringNotLike": {
                    "aws:PrincipalARN": "arn:aws:iam::*:role/ec2-imds-admins"
                }
            }
        }
    ]
}
```

Exigence de récupération des informations d'identification de rôle à partir d'IMDSv2

La stratégie suivante indique que si cette stratégie est appliquée à un rôle endossé par le service EC2, et que les informations d'identification obtenues sont utilisées pour signer une demande, la demande doit alors être signée par les informations d'identification de rôle EC2 extraites d'IMDSv2. Sinon, tous ses appels d'API recevront l'erreur `UnauthorizedOperation`. Cette instruction/stratégie peut être appliquée de manière générale car, si la demande n'est pas signée par les informations d'identification de rôle EC2, elle n'a aucun effet.

```
{
    "Version": "2012-10-17",
    "Statement": [
                {
            "Sid": "RequireAllEc2RolesToUseV2",
            "Effect": "Deny",
            "Action": "*",
            "Resource": "*",
            "Condition": {
                "NumericLessThan": {
                    "ec2:RoleDelivery": "2.0"
                }
            }
        }
    ]
}
```

Exemples de stratégies à utiliser sur la console Amazon EC2

Vous pouvez utiliser les stratégies IAM pour accorder aux utilisateurs les autorisations d'afficher et d'utiliser des ressources spécifiques sur la console Amazon EC2. Vous pouvez utiliser les exemples de stratégies de la section précédente. Toutefois, ces stratégies sont destinées aux demandes formulées avec

l'interface ligne de commande AWS CLI ou un kit SDK AWS. Puisque la console utilise des actions d'API supplémentaires pour ses fonctions, ces stratégies peuvent ne pas fonctionner comme escompté. Par exemple, un utilisateur n'ayant que l'autorisation d'utiliser l'action d'API `DescribeVolumes` rencontre une erreur s'il tente d'afficher les volumes sur la console. Cette section illustre les stratégies qui permettent aux utilisateurs d'utiliser des parties spécifiques de la console.

> Tip
>
> Pour vous aider à découvrir les actions d'API requises pour exécuter des tâches sur la console, vous pouvez utiliser un service tel que AWS CloudTrail. Pour plus d'informations, consultez le AWS CloudTrail User Guide. Si votre stratégie n'accorde pas l'autorisation de créer ou de modifier une ressource spécifique, la console affiche un message codé avec les informations de diagnostic. Vous pouvez décoder le message à l'aide de l'action d'API DecodeAuthorizationMessage pour AWS STS ou de la commande decode-authorization-message de l'AWS CLI.

Exemples

- Exemple : accès en lecture seule (p. 1026)
- Exemple : Utilisation de l'assistant de lancement d'EC2 (p. 1027)
- Exemple : Utilisation des volumes (p. 1030)
- Exemple : Utilisation des groupes de sécurité (p. 1031)
- Exemple : Utilisation d'adresses IP Elastic (p. 1033)
- Exemple : Utilisation d'Instances réservées (p. 1034)

Pour plus d'informations sur la création de stratégies pour la console Amazon EC2, consultez le billet suivant du blog sur la sécurité AWS : Granting Users Permission to Work in the Amazon EC2 Console.

Exemple : accès en lecture seule

Pour permettre aux utilisateurs d'afficher toutes les ressources sur la console Amazon EC2, vous pouvez utiliser la même stratégie que l'exemple suivant : Exemple : accès en lecture seule (p. 988). Les utilisateurs ne peuvent pas exécuter d'actions sur ces ressources ou créer des ressources, à moins qu'une autre déclaration ne leur accorde l'autorisation de le faire.

Afficher les instances, les AMI et les instantanés

Vous pouvez aussi fournir un accès en lecture seule à un sous-ensemble de ressources. Pour ce faire, remplacez le caractère générique * de l'action d'API `ec2:Describe` par les actions `ec2:Describe` spécifiques de chaque ressource. La stratégie suivante permet aux utilisateurs d'afficher l'ensemble des instances, AMI et instantanés sur la console Amazon EC2. L'action `ec2:DescribeTags` permet aux utilisateurs d'afficher les AMI publiques. La console nécessite les informations de balisage pour afficher les AMI publiques. Néanmoins, vous pouvez supprimer cette action pour permettre aux utilisateurs de voir uniquement les AMI privées.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
         "ec2:DescribeInstances",
         "ec2:DescribeImages",
         "ec2:DescribeTags",
         "ec2:DescribeSnapshots"
      ],
      "Resource": "*"
   }
   ]
}
```

> **Note**
>
> Comme les actions d'API Amazon EC2 `ec2:Describe*` ne prennent pas en charge les autorisations au niveau des ressources, vous ne pouvez pas contrôler les ressources individuelles que les utilisateurs peuvent afficher sur la console. Par conséquent, le caractère générique * est nécessaire dans l'élément `Resource` de la déclaration ci-dessus. Pour plus d'informations sur les ARN que vous pouvez utiliser avec les actions d'API Amazon EC2, consultez Actions, ressources et clés de condition pour Amazon EC2 dans le IAM Guide de l'utilisateur.

Afficher les instances et les métriques CloudWatch

La stratégie suivante permet aux utilisateurs d'afficher les instances sur la console Amazon EC2, ainsi que les alarmes et les métriques CloudWatch sous l'onglet Surveillance de la page Instances. Comme la console Amazon EC2 utilise l'API CloudWatch pour afficher les alarmes et les métriques, vous devez accorder aux utilisateurs l'autorisation d'utiliser les actions `cloudwatch:DescribeAlarms` et `cloudwatch:GetMetricStatistics`.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
         "ec2:DescribeInstances",
         "cloudwatch:DescribeAlarms",
         "cloudwatch:GetMetricStatistics"
      ],
      "Resource": "*"
   }
   ]
}
```

Exemple : Utilisation de l'assistant de lancement d'EC2

L'Assistant de lancement d'Amazon EC2 se compose d'un ensemble d'écrans contenant les options de configuration et de lancement d'une instance. Votre stratégie doit inclure l'autorisation d'utiliser les actions d'API qui permettent aux utilisateurs d'utiliser les options de l'Assistant. Si votre stratégie n'inclut pas l'autorisation d'utiliser ces actions, certains éléments de l'Assistant ne peuvent pas se charger correctement et les utilisateurs ne peuvent pas exécuter de lancement.

Accès de base à l'assistant de lancement

Pour exécuter un lancement avec succès, les utilisateurs doivent avoir l'autorisation d'utiliser l'action d'API `ec2:RunInstances`, ainsi qu'au moins les actions d'API suivantes :

- `ec2:DescribeImages` : afficher et sélectionner une AMI.
- `ec2:DescribeInstanceTypes` : afficher et sélectionner un type d'instance.
- `ec2:DescribeVpcs` : afficher les options réseau disponibles.
- `ec2:DescribeSubnets` : afficher tous les sous-réseaux disponibles pour le VPC choisi.
- `ec2:DescribeSecurityGroups` ou `ec2:CreateSecurityGroup` : pour afficher et sélectionner un groupe de sécurité existant, ou en créer un nouveau.
- `ec2:DescribeKeyPairs` ou `ec2:CreateKeyPair` : pour sélectionner une paire de clés existante ou en créer une nouvelle.
- `ec2:AuthorizeSecurityGroupIngress` : ajouter des règles entrantes.

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeInstances",
                "ec2:DescribeImages",
                "ec2:DescribeInstanceTypes",
                "ec2:DescribeKeyPairs",
                "ec2:DescribeVpcs",
                "ec2:DescribeSubnets",
                "ec2:DescribeSecurityGroups",
                "ec2:CreateSecurityGroup",
                "ec2:AuthorizeSecurityGroupIngress",
                "ec2:CreateKeyPair"
            ],
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": "ec2:RunInstances",
            "Resource": "*"
        }
    ]
}
```

Vous pouvez ajouter des actions d'API à votre stratégie pour fournir plus d'options pour les utilisateurs, par exemple :

- `ec2:DescribeAvailabilityZones` : afficher et sélectionner une zone de disponibilité spécifique.
- `ec2:DescribeNetworkInterfaces` : afficher et sélectionner les interfaces réseau existantes pour le sous-réseau sélectionné.
- Pour ajouter des règles sortantes à des groupes de sécurité VPC, les utilisateurs doivent recevoir l'autorisation d'utiliser l'action d'API `ec2:AuthorizeSecurityGroupEgress`. Pour modifier ou supprimer des règles existantes, les utilisateurs doivent recevoir l'autorisation d'utiliser l'action d'API `ec2:RevokeSecurityGroup*` appropriée.
- `ec2:CreateTags` : Pour attribuer des balises aux ressources qui sont créées par `RunInstances`. Pour plus d'informations, consultez Octroi de l'autorisation de baliser les ressources lors de la création (p. 985). Si les utilisateurs n'ont pas l'autorisation d'utiliser cette action et qu'ils essaient d'appliquer des balises sur la page de balisage de l'assistance de lancement, le lancement échoue.

  > **Important**
  >
  > Veillez à accorder aux utilisateurs l'autorisation d'utiliser l'action `ec2:CreateTags`, car cela limite votre capacité à utiliser la clé de condition `ec2:ResourceTag` pour restreindre leur utilisation d'autres ressources. Si vous accordez aux utilisateurs l'autorisation d'utiliser l'action `ec2:CreateTags`, ils peuvent modifier la balise d'une ressource afin de contourner ces restrictions. Pour de plus amples informations, veuillez consulter Contrôle de l'accès aux ressources EC2 à l'aide des balises de ressources (p. 987).
- Pour utiliser des paramètres Systems Manager lors de la sélection d'une AMI, vous devez ajouter `ssm:DescribeParameters` et `ssm:GetParameters` à votre stratégie. `ssm:DescribeParameters` accorde à vos utilisateurs IAM l'autorisation d'afficher et de sélectionner des paramètres Systems Manager. `ssm:GetParameters` accorde à vos utilisateurs IAM l'autorisation d'obtenir les valeurs des paramètres Systems Manager. Vous pouvez également restreindre l'accès à des paramètres Systems Manager spécifiques. Pour de plus amples informations, veuillez consulter Restreindre l'accès à des paramètres Systems Manager spécifiques plus loin dans cette section.

Comme les actions d'API Amazon EC2 `Describe*` ne prennent pas en charge les autorisations au niveau des ressources, vous ne pouvez pas limiter les ressources individuelles que les utilisateurs peuvent afficher dans l'Assistant de lancement. Cependant, vous pouvez appliquer les autorisations au niveau des ressources sur l'action d'API `ec2:RunInstances` pour limiter les ressources que les utilisateurs peuvent

employer pour lancer une instance. Le lancement échoue si les utilisateurs sélectionnent des options qu'ils ne sont pas autorisés à utiliser.

Limiter l'accès à un type d'instance, un sous-réseau et une région spécifiques

La stratégie suivante autorise les utilisateurs à lancer les instances `t2.micro` à l'aide des AMI dont Amazon est propriétaire, et uniquement sur un sous-réseau spécifique (`subnet-1a2b3c4d`). Les utilisateurs ne peuvent procéder à un lancement que dans la région sa-east-1. Si les utilisateurs sélectionnent une autre région ou un autre type d'instance, AMI ou sous-réseau dans l'Assistant, le lancement échoue.

La première déclaration accorde aux utilisateurs l'autorisation d'afficher les options dans l'Assistant de lancement, comme illustré dans l'exemple ci-dessus. La deuxième déclaration accorde aux utilisateurs l'autorisation d'utiliser les ressources interface réseau, volume, paire de clés, groupe de sécurité et sous-réseau pour l'action `ec2:RunInstances`, lesquelles sont requises pour lancer une instance sur un VPC. Pour plus d'informations sur l'utilisation de l'action `ec2:RunInstances`, consultez Lancement d'instances (RunInstances) (p. 1000). Les troisième et quatrième déclarations accordent aux utilisateurs l'autorisation d'utiliser, respectivement, les ressources de l'instance et les ressources de l'AMI, mais uniquement si l'instance est une instance `t2.micro`, et que l'AMI est la propriété d'Amazon.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
         "ec2:DescribeInstances",
         "ec2:DescribeImages",
         "ec2:DescribeInstanceTypes",
         "ec2:DescribeKeyPairs",
         "ec2:CreateKeyPair",
         "ec2:DescribeVpcs",
         "ec2:DescribeSubnets",
         "ec2:DescribeSecurityGroups",
         "ec2:CreateSecurityGroup",
         "ec2:AuthorizeSecurityGroupIngress"
   ],
   "Resource": "*"
   },
   {
      "Effect": "Allow",
      "Action":"ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:sa-east-1:111122223333:network-interface/*",
         "arn:aws:ec2:sa-east-1:111122223333:volume/*",
         "arn:aws:ec2:sa-east-1:111122223333:key-pair/*",
         "arn:aws:ec2:sa-east-1:111122223333:security-group/*",
         "arn:aws:ec2:sa-east-1:111122223333:subnet/subnet-1a2b3c4d"
      ]
   },
   {
      "Effect": "Allow",
      "Action": "ec2:RunInstances",
      "Resource": [
         "arn:aws:ec2:sa-east-1:111122223333:instance/*"
      ],
      "Condition": {
         "StringEquals": {
            "ec2:InstanceType": "t2.micro"
         }
      }
   },
   {
      "Effect": "Allow",
```

```
      "Action": "ec2:RunInstances",
      "Resource": [
            "arn:aws:ec2:sa-east-1::image/ami-*"
      ],
      "Condition": {
         "StringEquals": {
            "ec2:Owner": "amazon"
         }
      }
   }
   ]
}
```

Restreindre l'accès à des paramètres Systems Manager spécifiques

La stratégie suivante accorde l'accès à l'utilisation des paramètres Systems Manager avec un nom spécifique.

La première instruction accorde aux utilisateurs l'autorisation d'afficher les paramètres Systems Manager lors de la sélection d'une AMI dans l'assistant de lancement. La deuxième instruction accorde aux utilisateurs l'autorisation d'utiliser uniquement les paramètres nommés `prod-*`.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
         "ssm:DescribeParameters"
      ],
      "Resource": "*"
   },
   {
      "Effect": "Allow",
      "Action": [
      "ssm:GetParameters"
   ],
   "Resource": "arn:aws:ssm:us-east-2:123456123:parameter/prod-*"
   }
   ]
}
```

Exemple : Utilisation des volumes

La stratégie suivante accorde aux utilisateurs l'autorisation d'afficher et de créer des volumes, ainsi que d'en attacher à des instances spécifiques ou de les en détacher.

Les utilisateurs peuvent attacher un volume aux instances ayant la balise "`purpose=test`", ainsi que détacher des volumes de ces instances. Pour attacher un volume à l'aide de la console Amazon EC2, il est utile que les utilisateurs aient l'autorisation d'utiliser l'action `ec2:DescribeInstances`, car ils peuvent ainsi sélectionner une instance dans la liste préremplie de la boîte de dialogue Attacher un volume. Cependant, comme cela permet aussi aux utilisateurs d'afficher toutes les instances sur la page Instances de la console, vous pouvez ignorer cette action.

Dans la première déclaration, l'action `ec2:DescribeAvailabilityZones` est nécessaire pour garantir qu'un utilisateur puisse sélectionner une zone de disponibilité lors de la création d'un volume.

Les utilisateurs ne peuvent pas baliser les volumes qu'ils créent (pendant ou après la création de volume).

```
{
   "Version": "2012-10-17",
   "Statement": [{
```

```
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeVolumes",
        "ec2:DescribeAvailabilityZones",
        "ec2:CreateVolume",
        "ec2:DescribeInstances"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:AttachVolume",
        "ec2:DetachVolume"
      ],
      "Resource": "arn:aws:ec2:region:111122223333:instance/*",
      "Condition": {
        "StringEquals": {
          "ec2:ResourceTag/purpose": "test"
        }
     }
   },
   {
      "Effect": "Allow",
      "Action": [
        "ec2:AttachVolume",
        "ec2:DetachVolume"
      ],
      "Resource": "arn:aws:ec2:region:111122223333:volume/*"
   }
   ]
}
```

Exemple : Utilisation des groupes de sécurité

Afficher les groupes de sécurité et ajouter ou supprimer des règles

La stratégie suivante accorde aux utilisateurs l'autorisation d'afficher les groupes de sécurité dans la console Amazon EC2, d'ajouter et de supprimer des règles entrantes et sortantes et de modifier des descriptions de règles pour les groupes de sécurité existants ayant la balise `Department=Test`.

Dans la première déclaration, l'action `ec2:DescribeTags` permet aux utilisateurs d'afficher les balises sur la console, ce qui permet aux utilisateurs d'identifier plus facilement les groupes de sécurité qu'ils sont autorisés à modifier.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
         "ec2:DescribeSecurityGroups",
         "ec2:DescribeTags"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
         "ec2:AuthorizeSecurityGroupIngress",
         "ec2:RevokeSecurityGroupIngress",
         "ec2:AuthorizeSecurityGroupEgress",
         "ec2:RevokeSecurityGroupEgress",
         "ec2:UpdateSecurityGroupRuleDescriptionsIngress",
```

```
            "ec2:UpdateSecurityGroupRuleDescriptionsEgress"
         ],
         "Resource": [
            "arn:aws:ec2:region:111122223333:security-group/*"
         ],
         "Condition": {
            "StringEquals": {
               "ec2:ResourceTag/Department": "Test"
            }
         }
      }
      ]
}
```

Utilisation de la boîte de dialogue Créer un groupe de sécurité

Vous pouvez créer une stratégie qui permet aux utilisateurs d'utiliser la boîte de dialogue Créer un groupe de sécurité sur la console Amazon EC2. Pour utiliser cette boîte de dialogue, les utilisateurs doivent avoir l'autorisation d'utiliser au moins les actions d'API suivantes :

- `ec2:CreateSecurityGroup` : créer un groupe de sécurité.
- `ec2:DescribeVpcs` : afficher la liste des VPC existants dans la liste VPC.

Avec ces autorisations, les utilisateurs peuvent créer un groupe de sécurité avec succès, mais ne peuvent pas lui ajouter de règles. Pour utiliser les règles dans la boîte de dialogue Créer un groupe de sécurité, vous pouvez ajouter les actions d'API suivantes à votre stratégie :

- `ec2:AuthorizeSecurityGroupIngress` : ajouter des règles entrantes.
- `ec2:AuthorizeSecurityGroupEgress` : ajouter des règles sortantes aux groupes de sécurité VPC.
- `ec2:RevokeSecurityGroupIngress` : modifier ou supprimer des règles entrantes existantes. Cette règle est utile pour permettre aux utilisateurs d'utiliser la fonction Copier vers le nouveau sur la console. Cette fonction ouvre la boîte de dialogue Créer un groupe de sécurité et la complète avec les mêmes règles que le groupe de sécurité sélectionné.
- `ec2:RevokeSecurityGroupEgress` : modifier ou supprimer les règles sortantes pour les groupes de sécurité VPC. Cette règle permet aux utilisateurs de modifier ou de supprimer la règle sortante par défaut qui autorise tout le trafic sortant.
- `ec2:DeleteSecurityGroup` : répondre lorsque les règles non valides ne peuvent pas être enregistrées. La console commence par créer le groupe de sécurité et ajoute ensuite les règles spécifiées. Si les règles ne sont pas valides, l'action échoue et la console tente de supprimer le groupe de sécurité. Comme la boîte de dialogue Créer un groupe de sécurité reste affichée, l'utilisateur peut corriger la règle non valide et essayer de recréer le groupe de sécurité. Cette action d'API n'est pas obligatoire, mais si utilisateur n'a pas l'autorisation de l'utiliser et tente de créer un groupe de sécurité avec des règles non valides, le groupe de sécurité est créé sans aucune règle et l'utilisateur doit les ajouter après-coup.
- `ec2:UpdateSecurityGroupRuleDescriptionsIngress` : pour ajouter ou mettre à jour des descriptions des règles de trafic entrant pour les groupes de sécurité.
- `ec2:UpdateSecurityGroupRuleDescriptionsEgress` : pour ajouter ou mettre à jour des descriptions des règles de trafic sortant pour les groupes de sécurité.

Actuellement, l'action d'API `ec2:CreateSecurityGroup` ne prend pas en charge les autorisations au niveau des ressources ; cependant, vous pouvez appliquer ces autorisations aux actions `ec2:AuthorizeSecurityGroupIngress` et `ec2:AuthorizeSecurityGroupEgress` pour contrôler la façon dont les utilisateurs peuvent créer des règles.

La stratégie suivante accorde aux utilisateurs l'autorisation d'utiliser la boîte de dialogue Créer un groupe de sécurité, ainsi que de créer des règles entrantes et sortantes pour les groupes de sécurité associés à un

VPC spécifique (`vpc-1a2b3c4d`). Les utilisateurs peuvent créer des groupes de sécurité pour EC2-Classic ou autre VPC, mais ne peuvent pas leur ajouter de règles. De même, les utilisateurs ne peuvent pas ajouter de règles à un groupe de sécurité qui n'est pas associé au VPC `vpc-1a2b3c4d`. Les utilisateurs reçoivent aussi l'autorisation d'afficher tous les groupes de sécurité sur la console. Les utilisateurs peuvent ainsi identifier plus facilement les groupes de sécurité auxquels ils peuvent ajouter des règles entrantes. Cette stratégie accorde également aux utilisateurs l'autorisation de supprimer les groupes de sécurité associés au VPC `vpc-1a2b3c4d`.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeSecurityGroups",
        "ec2:CreateSecurityGroup",
        "ec2:DescribeVpcs"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DeleteSecurityGroup",
        "ec2:AuthorizeSecurityGroupIngress",
        "ec2:AuthorizeSecurityGroupEgress"
      ],
      "Resource": "arn:aws:ec2:region:111122223333:security-group/*",
      "Condition":{
         "ArnEquals": {
            "ec2:Vpc": "arn:aws:ec2:region:111122223333:vpc/vpc-1a2b3c4d"
         }
      }
    }
   ]
}
```

Exemple : Utilisation d'adresses IP Elastic

Pour autoriser les utilisateurs à afficher les adresses IP Elastic sur la console Amazon EC2, vous devez leur accorder l'autorisation d'utiliser l'action `ec2:DescribeAddresses`.

Pour autoriser les utilisateurs à utiliser les adresses IP Elastic, vous pouvez ajouter les actions suivantes à votre stratégie.

- `ec2:AllocateAddress` : allouer une adresse IP Elastic.
- `ec2:ReleaseAddress`: libérer une adresse IP Elastic.
- `ec2:AssociateAddress` : associer une adresse IP Elastic à une instance ou une interface réseau.
- `ec2:DescribeNetworkInterfaces` et `ec2:DescribeInstances` : utiliser l'écran Associer l'adresse. Cet écran affiche les instances ou interfaces réseau disponibles auxquelles vous pouvez associer une adresse IP Elastic.
- `ec2:DisassociateAddress` : dissocier une adresse IP Elastic d'une instance ou d'une interface réseau.

La stratégie suivante permet aux utilisateurs d'afficher, d'allouer et d'associer des adresses IP Elastic pour les instances. Les utilisateurs ne peuvent pas associer des adresses IP Elastic à des interfaces réseau, dissocier des adresses IP Elastic ou en libérer.

```
{
```

```
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ec2:DescribeAddresses",
                "ec2:AllocateAddress",
                "ec2:DescribeInstances",
                "ec2:AssociateAddress"
            ],
            "Resource": "*"
        }
    ]
}
```

Exemple : Utilisation d'Instances réservées

La stratégie suivante peut être attachée à un utilisateur IAM. Il donne à l'utilisateur l'accès pour afficher et modifier Instances réservées dans votre compte, ainsi que l'achat neuf Instances réservées dans le AWS Management Console.

Cette stratégie permet aux utilisateurs d'afficher tous les Instances réservées, ainsi que Instances à la demande, dans le compte. Il n'est pas possible de définir des autorisations au niveau des ressources pour les Instances réservées individuelles.

```
{
   "Version": "2012-10-17",
   "Statement": [{
      "Effect": "Allow",
      "Action": [
         "ec2:DescribeReservedInstances",
         "ec2:ModifyReservedInstances",
         "ec2:PurchaseReservedInstancesOffering",
         "ec2:DescribeInstances",
         "ec2:DescribeInstanceTypes",
         "ec2:DescribeAvailabilityZones",
         "ec2:DescribeReservedInstancesOfferings"
      ],
      "Resource": "*"
   }
   ]
}
```

L'action `ec2:DescribeAvailabilityZones` est nécessaire pour garantir que la console Amazon EC2 peut afficher des informations sur les zones de disponibilité dans lesquelles vous pouvez acheter des Instances réservées. L'action `ec2:DescribeInstances` n'est pas obligatoire, mais garantit que l'utilisateur peut afficher les instances du compte et acheter des réservations pour correspondre aux spécifications exactes.

Vous pouvez ajuster les actions d'API pour limiter l'accès utilisateur : par exemple, la suppression de `ec2:DescribeInstances` et `ec2:DescribeAvailabilityZones` signifie que l'utilisateur a l'accès en lecture seule.

Rôles IAM pour Amazon EC2

Les applications doivent signer leurs demandes d'API avec les informations d'identification AWS. Par conséquent, si vous êtes un développeur d'applications, vous avez besoin d'une stratégie pour gérer les informations d'identification de vos applications qui s'exécutent sur les instances EC2. Par exemple, vous pouvez distribuer en toute sécurité vos informations d'identification AWS aux instances, en permettant ainsi aux applications de ces instances d'utiliser vos informations d'identification pour signer des demandes,

tout en les protégeant des autres utilisateurs. Cependant, il est difficile de distribuer en toute sécurité les informations d'identification à chaque instance, particulièrement celles qu'AWS crée en votre nom, par exemple les instances Spot ou les instances des groupes Auto Scaling. Vous devez aussi pouvoir mettre à jour les informations d'identification de chaque instance lorsque vous faites tourner vos informations d'identification AWS.

Nous avons conçu les rôles IAM de telle sorte que vos applications puissent créer des demandes d'API en toute sécurité depuis vos instances, sans requérir que vous gériez les informations d'identification de sécurité que les applications utilisent. Au lieu de créer et de distribuer vos informations d'identification AWS, vous pouvez déléguer la permission pour créer des demandes d'API à l'aide des rôles IAM comme suit :

1. Créez un rôle IAM.
2. Définissez les comptes ou services AWS qui peuvent assumer le rôle.
3. Définissez les actions d'API et les ressources que l'application peut utiliser en assumant le rôle.
4. Spécifiez le rôle au lancement de votre instance ou attachez-le à une instance existante.
5. Demandez à l'application d'extraire un ensemble d'informations d'identification temporaires et utilisez-les.

Par exemple, vous pouvez utiliser des rôles IAM pour accorder l'autorisation aux applications de s'exécuter sur vos instances qui ont besoin d'utiliser un compartiment dans Amazon S3. Vous pouvez spécifier des permissions pour les rôles IAM en créant une politique au format JSON. Ces politiques sont similaires à celles que vous créez pour les utilisateurs IAM. Si vous modifiez un rôle, la modification est répercutée sur toutes les instances.

Lors de la création de rôles IAM, associez des stratégies IAM de moindres privilèges qui restreignent l'accès aux appels d'API spécifiques requis par l'application.

Vous ne pouvez pas attacher plusieurs rôles IAM à une seule instance, mais vous pouvez attacher un seul rôle IAM à plusieurs instances. Pour plus d'informations sur la création et l'utilisation des rôles IAM, consultez Rôles dans le IAM Guide de l'utilisateur.

Vous pouvez appliquer des autorisations au niveau des ressources à vos stratégies IAM pour contrôler la possibilité pour les utilisateurs d'attacher, de remplacer ou de détacher des rôles IAM pour une instance. Pour plus d'informations, consultez Autorisations au niveau des ressources prises en charge pour les opérations d'API Amazon Amazon EC2 (p. 981) et l'exemple suivant : Exemple : Utilisation des rôles IAM (p. 1020).

Sommaire

- Profils d'instance (p. 1035)
- Extraction des informations d'identification de sécurité à partir des métadonnées d'instance (p. 1036)
- Attribution d'une autorisation utilisateur IAM pour transmettre un rôle IAM à une instance (p. 1037)
- Utilisation des rôles IAM (p. 1037)

Profils d'instance

Amazon EC2 utilise un profil d'instance comme conteneur d'un rôle IAM. Lorsque vous créez un rôle IAM à l'aide de la console IAM, celle-ci crée automatiquement un profil d'instance et lui attribue le même nom qu'au rôle auquel il correspond. Si vous utilisez la console Amazon EC2 pour lancer une instance avec un rôle IAM ou pour attacher un rôle IAM à une instance, vous devez choisir le rôle en vous basant sur une liste de noms de profils d'instance.

Si vous utilisez l'AWS CLI, l'API ou un kit SDK AWS pour créer un rôle, vous devez créer le rôle et le profil d'instance sous la forme d'actions distinctes et leur attribuer éventuellement des noms différents. Si vous

utilisez ensuite l'AWS CLI, l'API ou un kit SDK AWS pour lancer une instance avec un rôle IAM ou pour attacher un rôle IAM à une instance, vous devez spécifier le nom du profil d'instance.

Un profil d'instance ne peut contenir qu'un seul rôle IAM. Cette limite ne peut pas être augmentée.

Pour plus d'informations, consultez Profils d'instance dans le IAM Guide de l'utilisateur.

Extraction des informations d'identification de sécurité à partir des métadonnées d'instance

Une application de l'instance extrait les informations d'identification de sécurité fournies par le rôle à partir de l'élément `iam/security-credentials/`nom-rôle des métadonnées d'instance. L'application reçoit les autorisations pour les actions et les ressources que vous avez définies pour le rôle via les informations d'identification de sécurité associées au rôle. Ces informations de sécurité sont temporaires et nous les faisons tourner automatiquement. Nous rendons disponibles de nouvelles informations d'identification au moins cinq minutes avant l'expiration des anciennes informations d'identification.

> **Warning**
>
> Si vous utilisez des services qui emploient les métadonnées d'instance avec les rôles IAM, assurez-vous de ne pas exposer vos informations d'identification quand les services effectuent des appels HTTP en votre nom. Les types de services qui peuvent exposer vos informations d'identification incluent les proxys HTTP, les services de validation HTML/CSS et les processeurs XML qui prennent en charge l'inclusion XML.

La commande suivante extrait les informations de sécurité pour un rôle IAM intitulé `s3access`.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/iam/security-credentials/s3access
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/iam/security-credentials/
s3access
```

Voici un exemple de sortie.

```
{
  "Code" : "Success",
  "LastUpdated" : "2012-04-26T16:39:16Z",
  "Type" : "AWS-HMAC",
  "AccessKeyId" : "ASIAIOSFODNN7EXAMPLE",
  "SecretAccessKey" : "wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY",
  "Token" : "token",
  "Expiration" : "2017-05-17T15:09:54Z"
}
```

Pour les applications, l'AWS CLI et les commandes Outils pour Windows PowerShell qui s'exécutent sur l'instance, vous n'avez pas à obtenir explicitement les informations d'identification de sécurité temporaires : les kits SDK AWS, l'AWS CLI et Outils pour Windows PowerShell obtiennent automatiquement les informations d'identification à partir du service des métadonnées d'instance EC2 et les utilisent. Pour effectuer un appel en dehors de l'instance à l'aide d'informations d'identification de sécurité temporaires

(par exemple, pour tester les stratégies IAM), vous devez fournir la clé d'accès, la clé secrète et le jeton de session. Pour plus d'informations, consultez Utilisation d'informations d'identification de sécurité temporaires pour demander l'accès aux ressources AWS dans le IAM Guide de l'utilisateur.

Pour obtenir plus d'informations sur les métadonnées d'instance, consultez Métadonnées d'instance et données utilisateur (p. 701). Pour de plus amples informations sur l'adresse IP des métadonnées d'instance, veuillez consulter Récupération des métadonnées d'instance (p. 708).

Attribution d'une autorisation utilisateur IAM pour transmettre un rôle IAM à une instance

Pour permettre à un utilisateur IAM de lancer une instance avec un rôle IAM ou d'attacher ou remplacer un rôle IAM pour une instance existante, vous devez accorder à l'utilisateur l'autorisation de transmettre le rôle à l'instance.

La stratégie IAM suivante accorde aux utilisateurs l'autorisation de lancer des instances (`ec2:RunInstances`) avec un rôle IAM ou d'attacher ou remplacer un rôle IAM pour une instance existante (`ec2:AssociateIamInstanceProfile` et `ec2:ReplaceIamInstanceProfileAssociation`).

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
         "ec2:RunInstances",
         "ec2:AssociateIamInstanceProfile",
         "ec2:ReplaceIamInstanceProfileAssociation"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": "iam:PassRole",
      "Resource": "*"
    }
  ]
}
```

Cette stratégie accorde aux utilisateurs IAM l'accès à l'ensemble de vos rôles en spécifiant la ressource sous la forme « * » dans la stratégie. Cependant, demandez-vous si les utilisateurs qui lancent des instances avec vos rôles (ceux qui existent ou que vous créez ultérieurement) peuvent se voir attribuer des autorisations dont ils n'ont pas besoin ou qu'ils ne devraient pas avoir.

Utilisation des rôles IAM

Vous pouvez créer un rôle IAM et l'attacher à une instance pendant ou après le lancement. Vous pouvez aussi remplacer ou détacher un rôle IAM pour une instance.

Sommaire

- Création d'un rôle IAM (p. 1038)
- Lancement d'une instance avec un rôle IAM (p. 1039)
- Attachement d'un rôle IAM à une instance (p. 1041)
- Remplacement d'un rôle IAM (p. 1042)
- Détachement d'un rôle IAM (p. 1043)

Création d'un rôle IAM

Vous devez créer un rôle IAM avant de pouvoir lancer une instance avec ce rôle ou attacher celui-ci à une instance.

Pour créer un rôle IAM avec la console IAM

1. Ouvrez la console IAM à l'adresse https://console.aws.amazon.com/iam/.
2. Dans le panneau de navigation, choisissez Rôles, puis Créer un rôle.
3. Sur la page Sélectionner un type de rôle, choisissez EC2 et le cas d'utilisation EC2. Sélectionnez Étape suivante : autorisations.
4. Sur la page Attacher une stratégie d'autorisations, choisissez une stratégie gérée AWS qui accorde à vos instances l'accès aux ressources dont elles ont besoin.
5. Sur la page Review (Vérifier), saisissez le nom du rôle et sélectionnez Create role (Créer un rôle).

Vous pouvez également utiliser l'AWS CLI pour créer un rôle IAM. L'exemple suivant crée un rôle IAM avec une stratégie qui lui permet d'utiliser un compartiment Amazon S3.

Pour créer un rôle IAM et un profil d'instance (AWS CLI)

1. Créez la stratégie d'approbation suivante et enregistrez-la dans un fichier texte intitulé `ec2-role-trust-policy.json`.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": { "Service": "ec2.amazonaws.com"},
      "Action": "sts:AssumeRole"
    }
  ]
}
```

2. Créez le rôle `s3access` et spécifiez la stratégie d'approbation que vous avez créée à l'aide de la commande create-role .

```
aws iam create-role --role-name s3access --assume-role-policy-document file://ec2-role-trust-policy.json
{
    "Role": {
        "AssumeRolePolicyDocument": {
            "Version": "2012-10-17",
            "Statement": [
                {
                    "Action": "sts:AssumeRole",
                    "Effect": "Allow",
                    "Principal": {
                        "Service": "ec2.amazonaws.com"
                    }
                }
            ]
        },
        "RoleId": "AROAIIZKPBKS2LEXAMPLE",
        "CreateDate": "2013-12-12T23:46:37.247Z",
        "RoleName": "s3access",
        "Path": "/",
        "Arn": "arn:aws:iam::123456789012:role/s3access"
    }
```

```
}
```

3. Créez une stratégie d'accès et enregistrez-la dans un fichier texte intitulé `ec2-role-access-policy.json`. Par exemple, cette stratégie d'accès accorde des permissions administratives pour Amazon S3 aux applications s'exécutant sur l'instance.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": ["s3:*"],
      "Resource": ["*"]
    }
  ]
}
```

4. Attachez la stratégie d'accès au rôle à l'aide de la commande put-role-policy.

```
aws iam put-role-policy --role-name s3access --policy-name S3-Permissions --policy-
document file://ec2-role-access-policy.json
```

5. Créez un profil d'instance nommé `s3access-profile` à l'aide de la commande create-instance-profile.

```
aws iam create-instance-profile --instance-profile-name s3access-profile
{
    "InstanceProfile": {
        "InstanceProfileId": "AIPAJTLBPJLEGREXAMPLE",
        "Roles": [],
        "CreateDate": "2013-12-12T23:53:34.093Z",
        "InstanceProfileName": "s3access-profile",
        "Path": "/",
        "Arn": "arn:aws:iam::123456789012:instance-profile/s3access-profile"
    }
}
```

6. Ajoutez le rôle `s3access` au profil d'instance `s3access-profile`.

```
aws iam add-role-to-instance-profile --instance-profile-name s3access-profile --role-
name s3access
```

Vous pouvez également utiliser les commandes Outils AWS pour Windows PowerShell suivantes :

- New-IAMRole
- Register-IAMRolePolicy
- New-IAMInstanceProfile

Lancement d'une instance avec un rôle IAM

Après avoir créé un rôle IAM, vous pouvez lancer une instance et associer le rôle à l'instance durant le lancement.

> **Important**
>
> Après que vous avez créé un rôle IAM, la propagation des permissions peut prendre plusieurs secondes. En cas d'échec de votre première tentative de lancer une instance avec un rôle,

attendez quelques secondes avant de recommencer. Pour plus d'informations, consultez la section relative à la résolution des problèmes liés à l'utilisation des rôles dans le IAM Guide de l'utilisateur.

Pour lancer une instance avec un rôle IAM (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sur le tableau de bord, choisissez Lancer une instance.
3. Sélectionnez une AMI et un type d'instance et choisissez ensuite Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
4. Sur la page Configurer les détails de l'instance, pour Rôle IAM, sélectionnez le rôle IAM que vous avez créé.

   Note

   La liste Rôle IAM affiche le nom du profil d'instance que vous avez créé lorsque vous avez créé votre rôle IAM. Si vous avez créé votre rôle IAM à l'aide de la console, le profil d'instance a été créé automatiquement et reçu le même nom que le rôle. Si vous avez créé votre rôle IAM à l'aide de l'AWS CLI, de l'API ou d'un kit SDK AWS, il se peut que vous ayez nommé votre profil d'instance différemment.
5. Configurez les autres détails, puis suivez les déclarations restantes de l'assistant, ou choisissez Vérifier et lancer pour accepter les paramètres par défaut et accéder directement à la page Examiner le lancement de l'instance.
6. Vérifiez vos paramètres, puis sélectionnez Lancer pour choisir une paire de clés et démarrer votre instance.
7. Si vous utilisez les actions d'API Amazon EC2 dans votre application, extrayez les informations d'identification de sécurité AWS rendues disponibles sur l'instance et utilisez-les pour signer les demandes. Le kit SDK AWS s'en charge pour vous.

   IMDSv2

   ```
   [ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-
   aws-ec2-metadata-token-ttl-seconds: 21600"` \
   && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/
   meta-data/iam/security-credentials/role_name
   ```

   IMDSv1

   ```
   [ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/iam/security-
   credentials/role_name
   ```

Vous pouvez aussi utiliser l'AWS CLI pour associer un rôle à une instance lors du lancement. Vous devez spécifier le profil d'instance dans la commande.

Pour lancer une instance avec un rôle IAM (AWS CLI)

1. Utilisez la commande run-instances pour lancer une instance à l'aide du profil d'instance. L'exemple suivant montre comment lancer une instance avec le profil d'instance.

```
aws ec2 run-instances \
    --image-id ami-11aa22bb \
    --iam-instance-profile Name="s3access-profile" \
    --key-name my-key-pair \
    --security-groups my-security-group \
    --subnet-id subnet-1a2b3c4d
```

Vous pouvez également utiliser la commande New-EC2Instance Outils pour Windows PowerShell.

2. Si vous utilisez les actions d'API Amazon EC2 dans votre application, extrayez les informations d'identification de sécurité AWS rendues disponibles sur l'instance et utilisez-les pour signer les demandes. Le kit SDK AWS s'en charge pour vous.

```
curl http://169.254.169.254/latest/meta-data/iam/security-credentials/role_name
```

Attachement d'un rôle IAM à une instance

Pour attacher un IAM à une instance qui n'a pas de rôle, l'instance doit être en état `stopped` ou `running`.

Nouvelle console

Pour attacher un rôle IAM à une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, choisissez Actions, Instance settings (Paramètres de l'instance), Modify IAM role (Modifier le rôle IAM).
4. Sélectionnez le rôle IAM à attacher à votre instance et choisissez Save (Enregistrer).

Ancienne console

Pour attacher un rôle IAM à une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, puis choisissez Actions, Paramètres de l'instance, Attacher/Remplacer le rôle IAM.
4. Sélectionnez le rôle IAM à attacher à votre instance et choisissez Appliquer.

Pour attacher un rôle IAM à une instance (AWS CLI)

1. Si nécessaire, décrivez vos instances pour obtenir l'ID de l'instance à laquelle attacher le rôle.

```
aws ec2 describe-instances
```

2. Utilisez la commande associate-iam-instance-profile pour attacher le rôle IAM à l'instance en spécifiant le profil d'instance. Vous pouvez utiliser l'Amazon Resource Name (ARN) du profil d'instance ou le nom du profil d'instance.

```
aws ec2 associate-iam-instance-profile \
    --instance-id i-1234567890abcdef0 \
    --iam-instance-profile Name="TestRole-1"

{
    "IamInstanceProfileAssociation": {
        "InstanceId": "i-1234567890abcdef0",
        "State": "associating",
        "AssociationId": "iip-assoc-0dbd8529a48294120",
        "IamInstanceProfile": {
            "Id": "AIPAJLNLDX3AMYZNWYYAY",
            "Arn": "arn:aws:iam::123456789012:instance-profile/TestRole-1"
```

```
        }
    }
}
```

Vous pouvez également utiliser les commandes Outils pour Windows PowerShell suivantes :

- Get-EC2Instance
- Register-EC2IamInstanceProfile

Remplacement d'un rôle IAM

Pour remplacer le rôle IAM sur une instance qui a déjà un rôle IAM, l'instance doit être en état `running`. Vous pouvez le faire si vous souhaitez modifier le rôle IAM pour une instance sans commencer par détacher le rôle existant. Pour exemple, vous pouvez le faire pour veiller à ce que les actions d'API effectuées par les applications exécutées sur l'instance ne soient pas interrompues.

Nouvelle console

Pour remplacer un rôle IAM pour une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, choisissez Actions, Instance settings (Paramètres de l'instance), Modify IAM role (Modifier le rôle IAM).
4. Sélectionnez le rôle IAM à attacher à votre instance et choisissez Save (Enregistrer).

Ancienne console

Pour remplacer un rôle IAM pour une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, puis choisissez Actions, Paramètres de l'instance, Attacher/Remplacer le rôle IAM.
4. Sélectionnez le rôle IAM à attacher à votre instance et choisissez Appliquer.

Pour remplacer un rôle IAM d'une instance (AWS CLI)

1. Si nécessaire, décrivez vos associations de profils d'instance IAM pour obtenir l'ID d'association du profil d'instance IAM à remplacer.

```
aws ec2 describe-iam-instance-profile-associations
```

2. Utilisez la commande replace-iam-instance-profile-association pour remplacer le profil d'instance IAM en spécifiant l'ID d'association du profil d'instance existant et l'ARN ou le nom du profil d'instance qui doit le remplacer.

```
aws ec2 replace-iam-instance-profile-association \
    --association-id iip-assoc-0044d817db6c0a4ba \
    --iam-instance-profile Name="TestRole-2"

{
    "IamInstanceProfileAssociation": {
```

```
        "InstanceId": "i-087711ddaf98f9489",
        "State": "associating",
        "AssociationId": "iip-assoc-09654be48e33b91e0",
        "IamInstanceProfile": {
            "Id": "AIPAJCJEDKX7QYHWYK7GS",
            "Arn": "arn:aws:iam::123456789012:instance-profile/TestRole-2"
        }
    }
}
```

Vous pouvez également utiliser les commandes Outils pour Windows PowerShell suivantes :

- Get-EC2IamInstanceProfileAssociation
- Set-EC2IamInstanceProfileAssociation

Détachement d'un rôle IAM

Vous ne pouvez pas détacher un rôle IAM d'une instance en cours d'exécution ou arrêtée.

Nouvelle console

Pour détacher un rôle IAM d'une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, choisissez Actions, Instance settings (Paramètres de l'instance), Modify IAM role (Modifier le rôle IAM).
4. Pour IAM role (Rôle IAM), choisissez No IAM Role (Aucun rôle IAM). Choisissez Save.
5. Dans la boîte de dialogue de confirmation, entrez Detach (Détacher), puis choisissez Detach (Détacher).

Ancienne console

Pour détacher un rôle IAM d'une instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance, puis choisissez Actions, Paramètres de l'instance, Attacher/Remplacer le rôle IAM.
4. Pour Rôle IAM, choisissez Aucun rôle. Choisissez Appliquer.
5. Dans la boîte de dialogue de confirmation, sélectionnez Oui, détacher.

Pour détacher un rôle IAM d'une instance (AWS CLI)

1. Si nécessaire, utilisez describe-iam-instance-profile-associations pour décrire vos associations de profils d'instance IAM et obtenir l'ID d'association pour le profil d'instance IAM à détacher.

```
aws ec2 describe-iam-instance-profile-associations

{
    "IamInstanceProfileAssociations": [
        {
            "InstanceId": "i-088ce778fbfeb4361",
```

```
            "State": "associated",
            "AssociationId": "iip-assoc-0044d817db6c0a4ba",
            "IamInstanceProfile": {
                "Id": "AIPAJEDNCAA64SSD265D6",
                "Arn": "arn:aws:iam::123456789012:instance-profile/TestRole-2"
            }
        }
    ]
}
```

2. Utilisez la commande disassociate-iam-instance-profile pour détacher le profil d'instance IAM en utilisant son ID d'association.

```
aws ec2 disassociate-iam-instance-profile --association-id iip-assoc-0044d817db6c0a4ba

{
    "IamInstanceProfileAssociation": {
        "InstanceId": "i-087711ddaf98f9489",
        "State": "disassociating",
        "AssociationId": "iip-assoc-0044d817db6c0a4ba",
        "IamInstanceProfile": {
            "Id": "AIPAJEDNCAA64SSD265D6",
            "Arn": "arn:aws:iam::123456789012:instance-profile/TestRole-2"
        }
    }
}
```

Vous pouvez également utiliser les commandes Outils pour Windows PowerShell suivantes :

- Get-EC2IamInstanceProfileAssociation
- Unregister-EC2IamInstanceProfile

Autorisation du trafic entrant pour vos instances Linux

Les groupes de sécurité vous permettent de contrôler le trafic vers votre instance, y compris celui qui peut atteindre votre instance. Par exemple, vous pouvez n'autoriser que les ordinateurs de votre réseau domestique à accéder à votre instance avec SSH. Si votre instance est un serveur web, vous pouvez autoriser toutes les adresses IP à accéder à votre instance via HTTP ou HTTPS, de telle sorte que les utilisateurs externes puissent parcourir le contenu de votre serveur web.

Vos groupes de sécurité par défaut et les groupes de sécurité nouvellement créés incluent les règles par défaut qui ne vous permettent pas d'accéder à votre instance depuis Internet. Pour plus d'informations, consultez Groupes de sécurité par défaut (p. 1066) et Custom security groups (p. 1066). Pour autoriser l'accès réseau à votre instance, vous devez autoriser le trafic entrant vers votre instance. Pour ouvrir un port pour le trafic entrant, ajoutez une règle au groupe de sécurité que vous avez associé à votre instance quand vous l'avez lancée.

Pour vous connecter à votre instance, vous devez configurer une règle pour autorise le trafic SSH à partir de l'adresse IPv4 publique de votre ordinateur. Pour autoriser le trafic SSH à partir des plages traditionnelles d'adresses IP, ajoutez une règle pour chaque plage que vous devez autoriser.

Si vous avez activé votre VPC pour IPv6 et lancé votre instance avec une adresse IPv6, vous pouvez vous connecter à l'instance à l'aide de son adresse IPv6 au lieu d'une adresse IPv4 publique. Votre ordinateur local doit avoir une adresse IPv6 et doit être configuré pour utiliser IPv6.

Si vous devez autoriser le trafic réseau vers une instance Windows, consultez Autorisation du trafic entrant pour vos instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Avant de commencer

Décidez qui nécessite un accès à votre instance. Il peut s'agir d'un hôte unique ou d'un réseau spécifique que vous approuvez. Par exemple, vous pouvez choisir l'adresse IPv4 publique de votre ordinateur local. L'éditeur de groupe de sécurité de la console Amazon EC2 peut détecter automatiquement l'adresse IPv4 publique de votre ordinateur local pour vous. Sinon, vous pouvez utiliser l'expression de recherche « quelle est mon adresse IP ? » dans un navigateur Internet, ou utiliser le service suivant : Check IP. Si votre connexion s'effectue via un ISP ou derrière un pare-feu sans adresse IP statique, vous devez déterminer la plage d'adresses IP utilisée par les ordinateurs clients.

> Warning
>
> Si vous utilisez `0.0.0.0/0`, vous permettez à toutes les adresses IPv4 d'accéder à votre instance avec SSH. Si vous utilisez `::/0`, vous permettez à toutes les adresses IPv6 d'accéder à votre instance. Cette solution est acceptable pour une brève durée dans un environnement de test, mais n'est pas sécurisée pour les environnements de production. Dans un environnement de production, vous autorisez uniquement l'accès à votre instance pour une adresse IP ou une plage d'adresses spécifiques.

Décidez si vous allez prendre en charge l'accès SSH à vos instances à l'aide de EC2 Instance Connect. Si vous n'avez pas l'intention d'utiliser EC2 Instance Connect, envisagez de le désinstaller ou de refuser l'action suivante dans vos stratégies IAM : `ec2-instance-connect:SendSSHPublicKey`. Pour plus d'informations, consultez Désinstallation d'EC2 Instance Connect (p. 613) et Configurer les autorisations IAM pour EC2 Instance Connect (p. 609).

Ajout d'une règle pour le trafic SSH entrant vers une instance Linux

Les groupes de sécurité font office de pare-feu pour les instances associées, en contrôlant le trafic entrant et le trafic sortant au niveau de l'instance. Vous devez ajouter des règles à un groupe de sécurité qui vous permettent de vous connecter à votre instance Linux à partir de votre adresse IP avec SSH.

Pour ajouter une règle à un groupe de sécurité pour le trafic SSH entrant sur IPv4 (console)

1. Dans le panneau de navigation de la console Amazon EC2, choisissez Instances. Sélectionnez votre instance et regardez l'onglet Description ; Groupes de sécurité répertorie les groupes de sécurité associés à l'instance. Choisissez afficher les règles pour afficher la liste des règles en vigueur pour l'instance.
2. Dans le panneau de navigation, sélectionnez Groupes de sécurité. Sélectionnez l'un des groupes de sécurité associés à votre instance.
3. Dans le volet des détails, sous l'onglet Entrant, choisissez Modifier. Dans la boîte de dialogue, choisissez Ajouter une règle, puis sélectionnez SSH dans la liste Type.
4. Dans le champ Source, choisissez Mon IP pour remplir automatiquement le champ avec l'adresse IPv4 publique de votre ordinateur local. Sinon, choisissez Personnalisé et spécifiez l'adresse IPv4 publique de votre ordinateur ou réseau en notation CIDR. Par exemple, si votre adresse IPv4 est `203.0.113.25`, spécifiez `203.0.113.25/32` pour afficher cette seule adresse IPv4 en notation CIDR. Si votre entreprise alloue des adresses à partir d'une plage, spécifiez la plage complète, telle que `203.0.113.0/24`.

   Pour plus d'informations sur la recherche de votre adresses IP, consultez Avant de commencer (p. 1045).
5. Choisissez Save.

Si vous avez lancé une instance avec une adresse IPv6 et que vous souhaitez vous connecter à votre instance à l'aide de son adresse IPv6, vous devez ajouter des règles qui autorisent le trafic IPv6 entrant sur SSH.

Pour ajouter une règle à un groupe de sécurité pour le trafic SSH entrant sur IPv6 (console)

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Groupes de sécurité. Sélectionnez le groupe de sécurité pour votre instance.
3. Choisissez Entrant, Modifier, puis Ajouter une règle.
4. Pour Type, choisissez SSH.
5. Dans le champ Source, spécifiez l'adresse IPv6 publique de votre ordinateur en notation CIDR. Par exemple, si votre adresse IPv6 est `2001:db8:1234:1a00:9691:9503:25ad:1761`, spécifiez `2001:db8:1234:1a00:9691:9503:25ad:1761/128` pour afficher la seule adresse IP en notation CIDR. Si votre entreprise alloue des adresses à partir d'une plage, spécifiez la plage complète, telle que `2001:db8:1234:1a00::/64`.
6. Choisissez Save.

Note

Veillez bien à exécuter les commandes suivantes sur votre système local, pas sur l'instance elle-même. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

Pour ajouter une règle à un groupe de sécurité à l'aide de la ligne de commande

1. Recherchez le groupe de sécurité associé à votre instance à l'aide de l'une des commandes suivantes :
   - describe-instance-attribute (AWS CLI)

     ```
     aws ec2 describe-instance-attribute --instance-id instance_id --attribute groupSet
     ```

   - Get-EC2InstanceAttribute (Outils AWS pour Windows PowerShell)

     ```
     PS C:\> (Get-EC2InstanceAttribute -InstanceId instance_id -Attribute groupSet).Groups
     ```

   Les deux commandes renvoient un ID de groupe de sécurité que vous utiliserez à l'étape suivante.
2. Ajoutez la règle au groupe de sécurité à l'aide de l'une des commandes suivantes :
   - authorize-security-group-ingress (AWS CLI)

     ```
     aws ec2 authorize-security-group-ingress --group-id security_group_id --protocol tcp
      --port 22 --cidr cidr_ip_range
     ```

   - Grant-EC2SecurityGroupIngress (Outils AWS pour Windows PowerShell)

     La commande `Grant-EC2SecurityGroupIngress` a besoin d'un paramètre `IpPermission` qui décrit le protocole, la plage de ports et la plage d'adresses IP à utiliser pour la règle de groupe de sécurité. La commande suivante crée le paramètre `IpPermission` :

     ```
     PS C:\> $ip1 = @{ IpProtocol="tcp"; FromPort="22"; ToPort="22";
      IpRanges="cidr_ip_range" }
     ```

     ```
     PS C:\> Grant-EC2SecurityGroupIngress -GroupId security_group_id -IpPermission
      @($ip1)
     ```

Affectation d'un groupe de sécurité à une instance

Vous pouvez assigner un groupe de sécurité à une instance lorsque vous lancez l'instance. Quand vous ajoutez ou supprimez des règles, ces modifications sont automatiquement appliquées à toutes les instances auxquelles vous avez affecté le groupe de sécurité.

Après avoir lancé une instance, vous pouvez modifier ses groupes de sécurité. Pour de plus amples informations, veuillez consulter Modification des groupes de sécurité d'une instance dans le Amazon VPC Guide de l'utilisateur.

Paires de clés Amazon EC2 et instances Linux

Une paire de clés, composée d'une clé privée et d'une clé publique, est un ensemble d'informations d'identification de sécurité que vous utilisez pour prouver votre identité lors de la connexion à une instance. Amazon EC2 stocke la clé publique et vous stockez la clé privée. Vous utilisez la clé privée, au lieu d'un mot de passe, pour accéder en toute sécurité à vos instances. Toute personne qui possède vos clés privées peut se connecter à vos instances, il est donc important que vous stockiez vos clés privées dans un endroit sécurisé.

Lorsque vous lancez une instance, vous êtes invité à entrer le nom d'une paire de clés (p. 536). Si vous prévoyez de vous connecter à l'instance en utilisant SSH, vous devez spécifier une paire de clés. Choisissez une paire de clés existante ou créez-en une. Lorsque votre instance démarre pour la première fois, le contenu de la clé publique que vous avez spécifiée au lancement est placé sur votre instance Linux dans une entrée dans `~/.ssh/authorized_keys`. Lorsque vous vous connectez à votre instance Linux à l'aide de SSH, vous devez spécifier la clé privée correspondant au contenu de la clé publique pour vous connecter. Pour plus d'informations sur la connexion à votre instance, consultez Connectez-vous à votre instance Linux (p. 598). Pour de plus amples informations sur les paires de clés et les instances Windows, veuillez consulter Paires de clés Amazon EC2 et instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows

Dans la mesure où Amazon EC2 ne conserve pas de copie de votre clé privée, si vous la perdez, il n'est pas possible de la récupérer. Cependant, il peut toujours y avoir un moyen de vous connecter aux instances pour lesquelles vous avez perdu la clé privée. Pour de plus amples informations, veuillez consulter Connexion à votre instance Linux en cas de perte de votre clé privée (p. 1056).

Les clés utilisées par Amazon EC2 sont des clés RSA SSH-2 2048 bits. Vous pouvez avoir jusqu'à 5 000 paires de clés par région.

Sommaire

- Création ou importation d'une paire de clés (p. 1047)
- Balisage d'une paire de clés (p. 1051)
- Récupération de la clé publique pour votre paire de clés sous Linux (p. 1053)
- Récupération de la clé publique de votre paire de clés via les métadonnées d'instance (p. 1053)
- Localisation de la clé publique sur une instance (p. 1054)
- Identification de la paire de clés spécifiée au lancement (p. 1055)
- Vérification de l'empreinte de votre paire de clés (p. 1055)
- Ajout ou remplacement d'une paire de clés pour votre instance (p. 1055)
- Connexion à votre instance Linux en cas de perte de votre clé privée (p. 1056)
- Suppression de votre paire de clés (p. 1061)

Création ou importation d'une paire de clés

Vous pouvez utiliser Amazon EC2 pour créer une paire de clés ou importer une paire de clés existante.

Options

- Option 1 : Créer une paire de clés à l'aide de Amazon EC2 (p. 1048)
- Option 2 : Importez votre propre clé publique dans Amazon EC2 (p. 1049)

Option 1 : Créer une paire de clés à l'aide de Amazon EC2

Vous pouvez créer une paire de clés à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour créer votre paire de clés

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous NETWORK & SECURITY, choisissez Key Pairs.
3. Choisissez Créer une paire de clés.
4. Dans Nom, entrez un nom descriptif pour la paire de clés. Amazon EC2 associe la clé publique au nom que vous spécifiez comme nom de clé. Le nom peut inclure jusqu'à 255 caractères ASCII. Il ne peut pas inclure d'espaces de début ou de fin.
5. Dans File format (Format de fichier), choisissez le format dans lequel enregistrer la clé privée. Pour enregistrer la clé privée dans un format qui peut être utilisé avec OpenSSH, choisissez pem. Pour enregistrer la clé privée dans un format qui peut être utilisé avec PuTTY, choisissez ppk.
6. Choisissez Créer une paire de clés.
7. Le fichier de clé privée est automatiquement téléchargé dans votre navigateur. Le nom de fichier de base est le nom que vous avez spécifié pour votre paire de clés, et l'extension de nom de fichier est déterminée par le format de fichier que vous avez choisi. Enregistrez le fichier de clé privée en lieu sûr.

   Important

   C'est votre seule occasion d'enregistrer le fichier de clé privée.
8. Si vous envisagez d'utiliser un client SSH sur un ordinateur macOS ou Linux pour vous connecter à votre instance Linux, utilisez la commande suivante pour définir les autorisations de votre fichier de clé privée afin d'être la seule personne autorisée à le lire.

   ```
   chmod 400 my-key-pair.pem
   ```

   Si vous ne définissez pas ces autorisations, vous ne pouvez pas vous connecter à votre instance à l'aide de cette paire de clés. Pour plus d'informations, consultez Erreur : fichier de clé privée non protégé (p. 1334).

Ancienne console

Pour créer votre paire de clés

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous NETWORK & SECURITY, choisissez Key Pairs.
3. Choisissez Create Key Pair.
4. Pour Nom de la paire de clés, entrez un nom descriptif pour la nouvelle paire de clés, puis choisissez Créer. Le nom peut inclure jusqu'à 255 caractères ASCII. Il ne peut pas inclure d'espaces de début ou de fin.
5. Le fichier de clé privée est automatiquement téléchargé dans votre navigateur. Le nom de fichier de base est le nom que vous avez spécifié comme nom de votre paire de clés et l'extension du nom de fichier est `.pem`. Enregistrez le fichier de clé privée en lieu sûr.

Important

C'est votre seule occasion d'enregistrer le fichier de clé privée.

6. Si vous envisagez d'utiliser un client SSH sur un ordinateur macOS ou Linux pour vous connecter à votre instance Linux, utilisez la commande suivante pour définir les autorisations de votre fichier de clé privée afin d'être la seule personne autorisée à le lire.

```
chmod 400 my-key-pair.pem
```

Si vous ne définissez pas ces autorisations, vous ne pouvez pas vous connecter à votre instance à l'aide de cette paire de clés. Pour plus d'informations, consultez Erreur : fichier de clé privée non protégé (p. 1334).

AWS CLI

Pour créer votre paire de clés

1. Utilisez la commande AWS CLI create-key-pair comme suit pour générer la clé et l'enregistrer dans un fichier `.pem`.

```
aws ec2 create-key-pair --key-name my-key-pair --query 'KeyMaterial' --output text
 > my-key-pair.pem
```

2. Si vous envisagez d'utiliser un client SSH sur un ordinateur macOS ou Linux pour vous connecter à votre instance Linux, utilisez la commande suivante pour définir les autorisations de votre fichier de clé privée afin d'être la seule personne autorisée à le lire.

```
chmod 400 my-key-pair.pem
```

Si vous ne définissez pas ces autorisations, vous ne pouvez pas vous connecter à votre instance à l'aide de cette paire de clés. Pour plus d'informations, consultez Erreur : fichier de clé privée non protégé (p. 1334).

PowerShell

Pour créer votre paire de clés

Utilisez la commande Outils AWS pour Windows PowerShell New-EC2KeyPair comme suit pour générer la clé et l'enregistrer dans un fichier `.pem`.

```
PS C:\> (New-EC2KeyPair -KeyName "my-key-pair").KeyMaterial | Out-File -Encoding ascii
 -FilePath C:\path\my-key-pair.pem
```

Option 2 : Importez votre propre clé publique dans Amazon EC2

Au lieu d'utiliser Amazon EC2 pour créer votre paire de clés, vous pouvez créer une paire de clés RSA à l'aide d'un outil tiers, puis importer la clé publique dans Amazon EC2. Par exemple, vous pouvez utiliser ssh-keygen (outil fourni avec l'installation OpenSSH standard) pour créer une paire de clés. Java, Ruby, Python, ainsi qu'un grand nombre d'autres langages de programmation fournissent également des bibliothèques standard pouvant être utilisées pour créer une paire de clés RSA.

Prérequis

- Les formats suivants sont pris en charge :

  - Format de clé publique OpenSSH (le format dans `~/.ssh/authorized_keys`). Si vous vous connectez avec SSH lorsque vous utilisez l'API EC2 Instance Connect, le format SSH2 est également pris en charge.
  - Le format DER codé en base64
  - Le format de fichier de clé publique SSH tel que spécifié dans RFC4716
  - Le format de fichier de clé privée SSH doit être PEM (par exemple, utilisez `ssh-keygen -m PEM` pour convertir la clé OpenSSH au format PEM)
- Créez une clé RSA. Amazon EC2 n'accepte pas les clés DSA.
- Les longueurs prises en charge sont 1024, 2048 et 4096. Si vous vous connectez avec SSH lorsque vous utilisez l'API EC2 Instance Connect, les longueurs prises en charge sont 2 048 et 4 096.

Pour créer une paire de clés à l'aide d'un outil tiers

1. Générez une paire de clés avec un outil tiers de votre choix.
2. Enregistrez la clé publique dans un fichier local. Par exemple, `~/.ssh/my-key-pair.pub` (Linux) ou `C:\keys\my-key-pair.pub` (Windows). L'extension du nom de fichier de ce fichier n'est pas importante.
3. Enregistrez la clé privée dans un autre fichier local ayant l'extension `.pem`. Par exemple, `~/.ssh/my-key-pair.pem` (Linux) ou `C:\keys\my-key-pair.pem` (Windows). Enregistrez le fichier de clé privée en lieu sûr. Vous devez fournir le nom de votre paire de clés quand vous lancez une instance, ainsi que la clé privée correspondante chaque fois que vous vous connectez à l'instance.

Après avoir créé la paire de clés, utilisez l'une des méthodes suivantes pour importer votre paire de clés vers Amazon EC2.

Nouvelle console

Pour importer la clé publique

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, cliquez sur Key Pairs.
3. Choisissez Import key pair (Importer une paire de clés).
4. Pour Name (Nom), entrez un nom descriptif pour la paire de clés. Le nom peut inclure jusqu'à 255 caractères ASCII. Il ne peut pas inclure d'espaces de début ou de fin.
5. Choisissez Browse (Parcourir) pour accéder à votre clé publique et la sélectionner, ou collez le contenu de votre clé publique dans le champ Public key contents (Contenu de la clé publique).
6. Choisissez Import key pair (Importer une paire de clés).
7. Vérifiez que la paire de clés importée apparaît dans la liste des paires de clés.

Ancienne console

Pour importer la clé publique

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sous NETWORK & SECURITY, choisissez Key Pairs.
3. Choisissez Importer une paire de clés.
4. Dans la boîte de dialogue Importer une paire de clés, choisissez Parcourir, puis sélectionnez le fichier de clé publique préalablement enregistré. Attribuez un nom à la paire de clés dans le champ Nom de la paire de clés, puis choisissez Importer. Le nom peut inclure jusqu'à 255 caractères ASCII. Il ne peut pas inclure d'espaces de début ou de fin.

5. Vérifiez que la paire de clés importée apparaît dans la liste des paires de clés.

AWS CLI

Pour importer la clé publique

Utilisez la commande de l'AWS CLI import-key-pair.

Pour vérifier que la paire de clés a été importée correctement

Utilisez la commande de l'AWS CLI describe-key-pairs.

PowerShell

Pour importer la clé publique

Utilisez la commande Outils AWS pour Windows PowerShell Import-EC2KeyPair.

Pour vérifier que la paire de clés a été importée correctement

Utilisez la commande Outils AWS pour Windows PowerShell Get-EC2KeyPair.

Balisage d'une paire de clés

Pour vous aider à classer et à gérer vos paires de clés existantes, vous pouvez les baliser avec des métadonnées personnalisées. Pour plus d'informations sur le fonctionnement des balises, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Vous pouvez afficher, ajouter et supprimer des balises à l'aide de la nouvelle console et des outils de ligne de commande.

Nouvelle console

Pour afficher, ajouter ou supprimer une balise pour une paire de clés existante

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, cliquez sur Key Pairs.
3. Sélectionnez une paire de clés, puis choisissez Actions, Gérer les balises.
4. La section Gérer les balises affiche toutes les balises affectées à la paire de clés.
   - Pour ajouter une balise, choisissez Ajouter la balise, puis entrez la clé et la valeur de la balise. Vous pouvez ajouter jusqu'à 50 balises par paire de clés. Pour de plus amples informations, veuillez consulter Restrictions liées aux balises (p. 1313).
   - Pour supprimer une balise, choisissez Supprimer en regard de la balise à supprimer.
5. Sélectionnez Save Changes.

AWS CLI

Pour afficher les balises de paires de clés

Utilisez la commande de l'AWS CLI describe-tags. Dans l'exemple suivant, vous décrivez les balises de toutes vos paires de clés.

```
$ aws ec2 describe-tags --filters "Name=resource-type,Values=key-pair"
```

```
{
    "Tags": [
    {
        "Key": "Environment",
        "ResourceId": "key-0123456789EXAMPLE",
        "ResourceType": "key-pair",
        "Value": "Production"
    },
    {
        "Key": "Environment",
        "ResourceId": "key-9876543210EXAMPLE",
        "ResourceType": "key-pair",
        "Value": "Production"
    }]
}
```

Pour décrire les balises d'une paire de clés spécifique

Utilisez la commande de l'AWS CLI describe-key-pairs.

```
$ aws ec2 describe-key-pairs --key-pair-ids key-0123456789EXAMPLE
```

```
{
    "KeyPairs": [
    {
        "KeyName": "MyKeyPair",
        "KeyFingerprint":
 "1f:51:ae:28:bf:89:e9:d8:1f:25:5d:37:2d:7d:b8:ca:9f:f5:f1:6f",
        "KeyPairId": "key-0123456789EXAMPLE",
        "Tags": [
        {
            "Key": "Environment",
            "Value": "Production"
        }]
    }]
}
```

Pour baliser une paire de clés existante

Utilisez la commande create-tags de l'AWS CLI. Dans l'exemple suivant, la paire de clés existante est balisée avec `Key=Cost-Center` et `Value=CC-123`.

```
$ aws ec2 create-tags --resources key-0123456789EXAMPLE --tags Key=Cost-
Center,Value=CC-123
```

Pour supprimer une balise d'une paire de clés

Utilisez la commande de l'AWS CLI delete-tags. Pour obtenir des exemples, reportez-vous à la section Exemples dans le document AWS CLI Command Reference.

PowerShell

Pour afficher les balises de paires de clés

Utilisez la commande Get-EC2Tag.

Pour décrire les balises d'une paire de clés spécifique

Utilisez la commande Get-EC2KeyPair.

Pour baliser une paire de clés existante

Utilisez la commande New-EC2Tag.

Pour supprimer une balise d'une paire de clés

Utilisez la commande Remove-EC2Tag.

Récupération de la clé publique pour votre paire de clés sous Linux

Sur votre ordinateur local Linux ou macOS, vous pouvez utiliser la commande ssh-keygen pour extraire la clé publique de votre paire de clés. Spécifiez le chemin où vous avez téléchargé votre clé privée (fichier `.pem`).

```
ssh-keygen -y -f /path_to_key_pair/my-key-pair.pem
```

La commande renvoie la clé publique, comme indiqué dans l'exemple suivant.

```
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQClKsfkNkuSevGj3eYhCe53pcjqP3maAhDFcvBS7O6V
hz2ItxCih+PnDSUaw+WNQn/mZphTk/a/gU8jEzoOWbkM4yxyb/wB96xbiFveSFJuOp/d6RJhJOI0iBXr
lsLnBItntckiJ7FbtxJMXLvvwJryDUilBMTjYtwB+QhYXUMOzce5Pjz5/i8SeJtjnV3iAoG/cQk+0FzZ
qaeJAAHco+CY/5WrUBkrHmFJr6HcXkvJdWPkYQS3xqC0+FmUZofz221CBt5IMucxXPkX4rWi+z7wB3Rb
BQoQzd8v7yeb7OzlPnWOyN0qFU0XA246RA8QFYiCNYwI3f05p6KLxEXAMPLE
```

Si la commande échoue, exécutez la commande suivante pour vous assurer d'avoir modifié les autorisations sur votre fichier de paire de clés afin d'être le seul à pouvoir l'afficher.

```
chmod 400 my-key-pair.pem
```

Récupération de la clé publique de votre paire de clés via les métadonnées d'instance

La clé publique spécifiée lorsque vous avez lancé une instance est également accessible via les métadonnées de l'instance. Pour afficher la clé publique que vous avez spécifiée lors du lancement de l'instance, utilisez la commande suivante à partir de votre instance :

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/public-keys/0/openssh-key
```

Voici un exemple de sortie.

```
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQClKsfkNkuSevGj3eYhCe53pcjqP3maAhDFcvBS7O6V
hz2ItxCih+PnDSUaw+WNQn/mZphTk/a/gU8jEzoOWbkM4yxyb/wB96xbiFveSFJuOp/d6RJhJOI0iBXr
lsLnBItntckiJ7FbtxJMXLvvwJryDUilBMTjYtwB+QhYXUMOzce5Pjz5/i8SeJtjnV3iAoG/cQk+0FzZ
qaeJAAHco+CY/5WrUBkrHmFJr6HcXkvJdWPkYQS3xqC0+FmUZofz221CBt5IMucxXPkX4rWi+z7wB3Rb
BQoQzd8v7yeb7OzlPnWOyN0qFU0XA246RA8QFYiCNYwI3f05p6KLxEXAMPLE my-key-pair
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/public-keys/0/openssh-key
```

Voici un exemple de sortie.

```
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQClKsfkNkuSevGj3eYhCe53pcjqP3maAhDFcvBS7O6V
hz2ItxCih+PnDSUaw+WNQn/mZphTk/a/gU8jEzoOWbkM4yxyb/wB96xbiFveSFJuOp/d6RJhJOI0iBXr
lsLnBItntckiJ7FbtxJMXLvvwJryDUilBMTjYtwB+QhYXUMOzce5Pjz5/i8SeJtjnV3iAoG/cQk+0FzZ
qaeJAAHco+CY/5WrUBkrHmFJr6HcXkvJdWPkYQS3xqC0+FmUZofz221CBt5IMucxXPkX4rWi+z7wB3Rb
BQoQzd8v7yeb7OzlPnWOyN0qFU0XA246RA8QFYiCNYwI3f05p6KLxEXAMPLE my-key-pair
```

Si vous modifiez la paire de clés que vous utilisez pour vous connecter à l'instance, nous ne mettons pas à jour les métadonnées de l'instance pour afficher la nouvelle clé publique. Au lieu de cela, les métadonnées d'instance continuent d'afficher la clé publique pour la paire de clés que vous avez spécifiée lors du lancement de l'instance. Pour plus d'informations, consultez Récupération des métadonnées d'instance (p. 708).

Sur une instance Linux, le contenu de la clé publique est placé dans une entrée du fichier `~/.ssh/authorized_keys`. Ouvrez le fichier dans un éditeur de texte. Voici un exemple d'entrée pour la paire de clés nommée **`my-key-pair`**. La clé publique est suivie du nom de la paire de clés.

```
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAABAQClKsfkNkuSevGj3eYhCe53pcjqP3maAhDFcvBS7O6V
hz2ItxCih+PnDSUaw+WNQn/mZphTk/a/gU8jEzoOWbkM4yxyb/wB96xbiFveSFJuOp/d6RJhJOI0iBXr
lsLnBItntckiJ7FbtxJMXLvvwJryDUilBMTjYtwB+QhYXUMOzce5Pjz5/i8SeJtjnV3iAoG/cQk+0FzZ
qaeJAAHco+CY/5WrUBkrHmFJr6HcXkvJdWPkYQS3xqC0+FmUZofz221CBt5IMucxXPkX4rWi+z7wB3Rb
BQoQzd8v7yeb7OzlPnWOyN0qFU0XA246RA8QFYiCNYwI3f05p6KLxEXAMPLE my-key-pair
```

Localisation de la clé publique sur une instance

Lorsque vous lancez une instance, vous êtes invité à entrer le nom d'une paire de clés (p. 536). Si vous prévoyez de vous connecter à l'instance en utilisant SSH, vous devez spécifier une paire de clés. Choisissez une paire de clés existante ou créez-en une. Lorsque votre instance démarre pour la première fois, le contenu de la clé publique que vous avez spécifiée au lancement est placé sur votre instance Linux dans une entrée dans `~/.ssh/authorized_keys`.

Pour localiser la clé publique sur une instance

1. Connectez-vous à votre instance. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Dans la fenêtre du terminal, ouvrez le fichier `authorized_keys` à l'aide de votre éditeur de texte préféré (tel que vim ou nano).

```
[ec2-user ~]$ nano ~/.ssh/authorized_keys
```

Le fichier `authorized_keys` s'ouvre, affichant la clé publique, comme le montre l'exemple suivant.

```
ssh-rsa
 AAAAB3NzaC1yc2EAAAADAQABAAABAQClKsfkNkuSevGj3eYhCe53pcjqP3maAhDFcvBS7O6Vhz2ItxCih
+PnDSUaw+WNQn/mZphTk/a/gU8jEzoOWbkM4yxyb/wB96xbiFveSFJuOp/
d6RJhJOI0iBXrlsLnBItntckiJ7FbtxJMXLvvwJryDUilBMTjYtwB+QhYXUMOzce5Pjz5/i8SeJtjnV3iAoG/
cQk+0FzZqaeJAAHco+CY/5WrUBkrHmFJr6HcXkvJdWPkYQS3xqC0+FmUZofz221CBt5IMucxXPkX4rWi
+z7wB3RbBQoQzd8v7yeb7OzlPnWOyN0qFU0XA246RA8QFYiCNYwI3f05p6KLxEXAMPLE
```

Identification de la paire de clés spécifiée au lancement

Lorsque vous lancez une instance, vous êtes invité à entrer le nom d'une paire de clés (p. 536). Si vous prévoyez de vous connecter à l'instance en utilisant SSH, vous devez spécifier une paire de clés.

Pour identifier la paire de clés spécifiée au lancement

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, puis sélectionnez votre instance.
3. Sous l'onglet Description, le champ Nom de la paire de clés affiche le nom de la paire de clés que vous avez spécifiée lors du lancement de l'instance. La valeur du nom de la paire de clés ne change pas même si vous modifiez la clé publique sur l'instance ou ajoutez des paires de clés.

Vérification de l'empreinte de votre paire de clés

Dans la page Paires de clés de la console Amazon EC2, la colonne Empreinte digitale affiche les empreintes digitales générées à partir de vos paires de clés. AWS calcule l'empreinte digitale différemment selon que la paire de clés a été générée par AWS ou un outil tiers. Si vous avez créé la paire de clés à l'aide d'AWS, l'empreinte est calculée à l'aide d'une fonction de hachage SHA-1. Si vous avez créé la paire de clés à l'aide d'un outil tiers et téléchargé la clé publique dans AWS, ou si vous avez généré une nouvelle clé publique à partir d'une clé privée créée avec AWS puis que vous l'avez téléchargée dans AWS, l'empreinte est calculée à partir d'une fonction de hachage MD5.

Vous pouvez utiliser l'empreinte SSH2 affichée sur la page Paires de clés pour vérifier que la clé privée présente sur votre ordinateur local correspond à la clé publique stockée dans AWS. À partir de l'ordinateur où vous avez téléchargé le fichier de clé privée, générez une empreinte SSH2 à partir du fichier de clé privée. La sortie doit correspondre à l'empreinte affichée dans la console.

Si vous avez créé la paire de clés avec AWS, vous pouvez utiliser les outils OpenSSL pour générer une empreinte, comme illustré dans l’exemple suivant :

```
$ openssl pkcs8 -in path_to_private_key -inform PEM -outform DER -topk8 -nocrypt | openssl
 sha1 -c
```

Si vous avez créé une paire de clés à l'aide d'un outil tiers puis téléchargé la clé publique dans AWS, vous pouvez utiliser les outils OpenSSL pour générer l'empreinte, comme illustré dans l’exemple suivant :

```
$ openssl rsa -in path_to_private_key -pubout -outform DER | openssl md5 -c
```

Si vous avez créé une paire de clés OpenSSH à l'aide d'OpenSSH 7.8 ou version ultérieure et téléchargé la clé publique dans AWS, vous pouvez utiliser ssh-keygen pour générer l'empreinte, comme illustré dans l’exemple suivant :

```
$ ssh-keygen -ef path_to_private_key -m PEM | openssl rsa -RSAPublicKey_in -outform DER |
 openssl md5 -c
```

Ajout ou remplacement d'une paire de clés pour votre instance

Vous pouvez modifier la paire de clés qui permet d'accéder au compte système par défaut de votre instance. Par exemple, si un utilisateur de votre organisation requiert l'accès au compte utilisateur système

à l'aide d'une paire de clés distincte, vous pouvez ajouter la paire de clés à votre instance. Ou, si quelqu'un possède une copie du fichier `.pem` et que vous voulez l'empêcher de se connecter à votre instance (par exemple, si la personne a quitté votre organisation), vous pouvez remplacer la paire de clés par la nouvelle.

Pour ajouter ou remplacer une paire de clés, vous devez pouvoir vous connecter à votre instance. Si vous avez perdu votre clé privée existante ou vous avez lancé votre instance sans paire de clés, vous ne pourrez pas vous connecter à votre instance et vous ne serez donc pas en mesure d'ajouter ou de remplacer une paire de clés. Si vous avez perdu votre clé privée existante, vous pourrez peut-être la récupérer. Pour plus d'informations, consultez Connexion à votre instance Linux en cas de perte de votre clé privée (p. 1056). Si vous lancez votre instance sans paire de clé, vous ne pourrez pas vous connecter à l'instance à moins de choisir une AMI configurée de façon à autoriser les utilisateurs à se connecter d'une autre façon.

> **Note**
>
> Ces procédures permettent de modifier la paire de clés pour le compte utilisateur par défaut, tel que `ec2-user`. Pour plus d'informations sur l'ajout de comptes d'utilisateur à votre instance, consultez Gestion des comptes d'utilisateur sur votre instance Amazon Linux (p. 660).

Pour ajouter ou remplacer une paire de clés

1. Créez une nouvelle paire de clés à l'aide de la console Amazon EC2 (p. 1048) ou d'un outil tiers (p. 1049).
2. Récupérez la clé publique de votre nouvelle paire de clés. Pour plus d'informations, consultez Récupération de la clé publique pour votre paire de clés sous Linux (p. 1053).
3. Connectez-vous à votre instance à l'aide de votre fichier de clé privée existant.
4. À l'aide d'un éditeur de texte de votre choix, ouvrez le fichier `.ssh/authorized_keys` sur l'instance. Collez les informations de clé publique depuis votre nouvelle paire de clés sous les informations existantes de clé publique. Sauvegardez le fichier.
5. Déconnectez-vous de votre instance et testez que vous pouvez vous connecter à votre instance à l'aide du nouveau fichier de clé privé.
6. (Facultatif) Si vous remplacez une paire de clés existante, connectez-vous à votre instance et supprimez les informations de clé publique de la paire de clés originale du fichier `.ssh/authorized_keys`.

> **Note**
>
> Si vous utilisez un groupe Auto Scaling, assurez-vous que la paire de clés que vous remplacez n'est pas spécifiée dans votre modèle ou votre configuration de lancement. Amazon EC2 Auto Scaling lance une instance de remplacement en cas de détection d'une instance défectueuse. Cependant, le lancement de l'instance échoue si la paire de clés est introuvable.

Connexion à votre instance Linux en cas de perte de votre clé privée

Si vous perdez la clé privée pour une instance basée sur des volumes EBS, vous pouvez à nouveau accéder à votre instance. Vous devez arrêter l'instance, détacher son volume racine et l'attacher à une autre instance en tant que volume de données, modifier le fichier `authorized_keys` avec une nouvelle clé publique, replacer le volume dans l'instance d'origine et redémarrer l'instance. Pour plus d'informations sur le lancement et l'arrêt des instances, ainsi que sur la connexion aux instances, consultez Cycle de vie d'une instance (p. 524).

Cette procédure n'est pas prise en charge pour les instances avec des volumes racine sauvegardés par le magasin d'instance. Pour déterminer le type d'appareil racine pour votre instance, ouvrez la console

Amazon EC2, choisissez Instances, sélectionnez l'instance et vérifiez la valeur de Type de périphérique racine dans le volet des détails. La valeur est `ebs` ou `instance store`. Si l'appareil racine est un volume de stockage d'instance, vous ne pouvez pas utiliser cette procédure pour rétablir l'accès à votre instance ; vous devez disposer de la clé privée pour vous connecter à l'instance.

Étapes de connexion à une instance basée sur des volumes EBS avec une paire de clés différente

- Étape 1 : Créer une nouvelle paire de clés (p. 1057)
- Étape 2 : Obtenir des informations sur l'instance d'origine et son volume racine (p. 1057)
- Étape 3 : Arrêter l'instance d'origine (p. 1058)
- Étape 4 : Lancer une instance temporaire (p. 1058)
- Étape 5 : Détacher le volume racine de l'instance d'origine et l'attacher à l'instance temporaire (p. 1058)
- Étape 6 : Ajouter la nouvelle clé publique authorized_keys sur le volume d'origine monté sur l'instance temporaire (p. 1059)
- Étape 7 : Démonter et détacher le volume d'origine de l'instance temporaire, puis le reconnecter à l'instance d'origine (p. 1060)
- Étape 8 : Se connecter à l'instance d'origine à l'aide de la nouvelle paire de clés (p. 1061)
- Étape 9 : nettoyer (p. 1061)

Étape 1 : Créer une nouvelle paire de clés

Créer une nouvelle paire de clés à l'aide de la console Amazon EC2 ou d'un outil tiers. Si vous souhaitez nommer votre nouvelle paire de clés exactement comme la clé privée perdue, vous devez commencer par supprimer la paire de clés existante. Pour de plus amples informations sur la création d'une paire de clés, veuillez consulter Option 1 : Créer une paire de clés à l'aide de Amazon EC2 (p. 1048) ou Option 2 : Importez votre propre clé publique dans Amazon EC2 (p. 1049).

Étape 2 : Obtenir des informations sur l'instance d'origine et son volume racine

Nouvelle console

Pour obtenir des informations sur votre instance d'origine

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instances dans le panneau de navigation, puis sélectionnez l'instance à laquelle vous souhaitez vous connecter. (Cette instance est qualifiée d'instance d'origine.)
3. Dans l'onglet Détails notez l'ID d'instance et l'ID d'AMI.
4. Sous l'onglet Mise en réseau notez la zone de disponibilité.
5. Sous l'onglet Stockage notez le nom de l'appareil pour le volume racine dans Nom de l'appareil racine (par exemple, `/dev/xvda`). Recherchez ce nom d'appareil sous Bloquer les appareils et notez l'ID de volume (par exemple, vol-0a1234b5678c910de).

Ancienne console

Pour obtenir des informations sur votre instance d'origine

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instances dans le panneau de navigation, puis sélectionnez l'instance à laquelle vous souhaitez vous connecter. (Cette instance est qualifiée d'instance d'origine.)

3. Dans l'onglet Description, enregistrez les informations suivantes qui vous seront nécessaires pour terminer cette procédure.

   - Notez l'ID de l'instance, l'ID d'AMI et la zone de disponibilité de l'instance d'origine.
   - Dans le champ Périphérique racine, notez le nom du périphérique du volume racine (par exemple, `/dev/sda1` ou `/dev/xvda`). Cliquez sur le lien et notez l'ID de volume dans le champ ID d'EBS (vol-xxxxxxxxxxxxxxxxx).

Étape 3 : Arrêter l'instance d'origine

Choisissez État de l'instance, Arrêter l'instance. Si cette option est désactivée, l'instance est déjà arrêtée ou son périphérique racine est un volume de stockage d'instance.

> **Warning**
>
> Lorsque vous arrêtez une instance, les données contenues sur les volumes de stockage d'instance sont effacées. Pour conserver les données des volumes de stockage d'instance, veillez à les sauvegarder sur un stockage permanent.

Étape 4 : Lancer une instance temporaire

Choisissez Lancer une instance, puis utilisez l'assistant de lancement pour lancer une instance temporaire avec les options suivantes :

- Dans la page Choisir une AMI, sélectionnez la même AMI que celle utilisée pour lancer l'instance d'origine. Si l'AMI n'est pas disponible, vous pouvez créer une AMI à utiliser depuis l'instance arrêtée. Pour de plus amples informations, veuillez consulter Création d'une AMI Linux basée sur Amazon EBS (p. 132).
- Sur la page Choisir un type d'instance, conservez le type d'instance par défaut sélectionné par l'assistant.
- Dans la page Configurer les détails de l'instance, spécifiez la même zone de disponibilité que l'instance d'origine. Si vous lancez une instance dans un VPC, sélectionnez un sous-réseau dans cette zone de disponibilité.
- Sur la page Ajouter des balises, ajoutez la balise `Name=Temporary` à l'instance pour indiquer qu'il s'agit d'une instance temporaire.
- Sur la page Vérification, choisissez Lancer. Créez une paire de clés, téléchargez-la dans un emplacement sûr de votre ordinateur, puis choisissez Lancer des instances.

Étape 5 : Détacher le volume racine de l'instance d'origine et l'attacher à l'instance temporaire

1. Dans le panneau de navigation, choisissez Volumes et sélectionnez le volume du périphérique racine pour l'instance d'origine (vous avez noté l'ID de volume au cours d'une étape précédente). Choisissez Actions, Detach Volume, puis sélectionnez Yes, Detach. Attendez que l'état du volume devienne `available`. (Vous devrez peut-être sélectionner l'icône Actualiser.)
2. Tandis que le volume est toujours sélectionné, choisissez Actions, puis sélectionnez Attacher un volume. Sélectionnez l'ID d'instance de l'instance temporaire, notez le nom de périphérique spécifié dans Device (par exemple, `//dev/sdf`), puis choisissez Attach.

   > **Note**
   >
   > Si vous avez lancé votre instance initiale à partir d'une AMI AWS Marketplace et que votre volume contient des codes AWS Marketplace, vous devez arrêter l'instance temporaire avant de pouvoir attacher le volume.

Étape 6 : Ajouter la nouvelle clé publique `authorized_keys` sur le volume d'origine monté sur l'instance temporaire

1. Connectez-vous à l'instance temporaire.
2. À partir de l'instance temporaire, montez le volume que vous avez attaché à l'instance afin de pouvoir accéder au système de fichiers. Par exemple, si le nom du périphérique est `/dev/sdf`, utilisez les commandes suivantes pour monter le volume en tant que `/mnt/tempvol`.

   Note

   Le nom du périphérique peut apparaître différemment sur votre instance. Par exemple, les périphériques montés en tant que `/dev/sdf` peuvent également s'afficher en tant que `/dev/xvdf` sur l'instance. Certaines versions de Red Hat (ou ses variantes, comme CentOS) peuvent même incrémenter la lettre finale de quatre caractères, et `/dev/sd`*`f`* devient `/dev/xvd`*`k`*.

   a. Utilisez la commande lsblk pour déterminer si le volume est divisé.

   ```
   [ec2-user ~]$ lsblk
   NAME    MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
   xvda    202:0    0    8G  0 disk
   ##xvda1 202:1    0    8G  0 part /
   xvdf    202:80   0  101G  0 disk
   ##xvdf1 202:81   0  101G  0 part
   xvdg    202:96   0   30G  0 disk
   ```

   Dans l'exemple précédent, `/dev/xvda` et `/dev/xvdf` sont des volumes partitionnés, mais `/dev/xvdg` ne l'est pas. Si votre volume est partitionné, vous montez la partition (`/dev/xvdf1`)) au lieu du périphérique brut (`/dev/xvdf`) au cours des étapes suivantes.

   b. Créez un répertoire temporaire pour monter le volume.

   ```
   [ec2-user ~]$ sudo mkdir /mnt/tempvol
   ```

   c. Montez le volume (ou la partition) sur le point de montage temporaire, en utilisant le nom du volume ou du périphérique que vous avez identifié plus tôt. La commande requise dépend du système de fichiers de votre système d'exploitation.

   - Amazon Linux, Ubuntu et Debian

   ```
   [ec2-user ~]$ sudo mount /dev/xvdf1 /mnt/tempvol
   ```

   - Amazon Linux 2, CentOS, SUSE Linux 12 et RHEL 7.x

   ```
   [ec2-user ~]$ sudo mount -o nouuid /dev/xvdf1 /mnt/tempvol
   ```

   Note

   Si vous obtenez une erreur indiquant que le système de fichiers est endommagé, exécutez la commande suivante pour utiliser l'utilitaire fsck afin de rechercher les erreurs dans votre système de fichiers et de les résoudre.

   ```
   [ec2-user ~]$ sudo fsck /dev/xvdf1
   ```

3. À partir de l'instance temporaire, utilisez la commande suivante pour mettre à jour `authorized_keys` sur le volume monté avec la nouvelle clé publique de `authorized_keys` pour l'instance temporaire.

Important

Les exemples suivants utilisent le nom d'utilisateur Amazon Linux `ec2-user`. Vous devrez peut-être modifier le nom d'utilisateur, par exemple, `ubuntu` pour les instances Ubuntu.

```
[ec2-user ~]$ cp .ssh/authorized_keys /mnt/tempvol/home/ec2-user/.ssh/authorized_keys
```

Une fois que cette étape est correctement effectuée, vous pouvez passer à l'étape suivante.

(Facultatif) Sinon, si vous n'êtes pas autorisé à modifier des fichiers dans `/mnt/tempvol`, vous devez mettre à jour le fichier à l'aide de la commande sudo, puis vérifier les autorisations sur le fichier afin de vous assurer que vous êtes en mesure de vous connecter à l'instance d'origine. Pour vérifier les autorisations sur le fichier, utilisez la commande suivante.

```
[ec2-user ~]$ sudo ls -l /mnt/tempvol/home/ec2-user/.ssh
total 4
-rw------- 1 222 500 398 Sep 13 22:54 authorized_keys
```

Dans cet exemple, *222* est l'ID d'utilisateur et *500* est l'ID de groupe. Utilisez ensuite la commande sudo pour ré-exécuter la commande copy ayant échoué.

```
[ec2-user ~]$ sudo cp .ssh/authorized_keys /mnt/tempvol/home/ec2-user/.ssh/
authorized_keys
```

Exécutez à nouveau la commande suivante pour déterminer si les autorisations ont été modifiées.

```
[ec2-user ~]$ sudo ls -l /mnt/tempvol/home/ec2-user/.ssh
```

Si l'ID d'utilisateur et l'ID de groupe ont été modifiés, utilisez la commande suivante pour les restaurer.

```
[ec2-user ~]$ sudo chown 222:500 /mnt/tempvol/home/ec2-user/.ssh/authorized_keys
```

Étape 7 : Démonter et détacher le volume d'origine de l'instance temporaire, puis le reconnecter à l'instance d'origine

1. À partir de l'instance temporaire, démontez le volume que vous avez attaché afin de pouvoir l'attacher à nouveau à l'instance d'origine. Par exemple, utilisez la commande suivante pour démonter le volume situé dans `/mnt/tempvol`.

   ```
   [ec2-user ~]$ sudo umount /mnt/tempvol
   ```

2. Détachez le volume de l'instance temporaire (vous l'avez démonté à l'étape précédente) : dans la console Amazon EC2, sélectionnez le volume de l'appareil racine de l'instance d'origine (vous avez noté l'ID de volume à l'étape précédente), choisissez Actions, Détacher le volume, puis sélectionnez Oui, Détacher. Attendez que l'état du volume devienne `available`. (Vous devrez peut-être sélectionner l'icône Actualiser.)
3. Rattachez le volume à l'instance d'origine : le volume étant toujours sélectionné, choisissez Actions, Attacher le volume. Sélectionnez l'ID d'instance de l'instance d'origine, précisez le nom de l'appareil que vous avez noté précédemment au cours de l'étape 2 (p. 1057) pour l'attachement de l'appareil racine d'origine (`/dev/sda1` ou `/dev/xvda`), puis choisissez Attacher.

Important

Si vous ne spécifiez pas le même nom de périphérique que pour l'attachement original, vous ne pourrez pas démarrer l'instance d'origine. Amazon EC2 s'attend à ce que le volume du périphérique racine soit à l'adresse `sda1` ou `/dev/xvda`.

Étape 8 : Se connecter à l'instance d'origine à l'aide de la nouvelle paire de clés

Sélectionnez l'instance d'origine, choisissez État de l'instance, Démarrer l'instance. Lorsque l'état de l'instance est `running`, vous pouvez vous y connecter à l'aide du fichier de clé privée de votre nouvelle paire de clés.

Note

Si le nom de votre paire de clés et du fichier de clé privée correspondant est différent du nom de la paire de clés initiale, veillez à spécifier le nom du nouveau fichier de clé privée lorsque vous vous connectez à votre instance.

Étape 9 : nettoyer

(Facultatif) Vous pouvez mettre fin à l'instance temporaire si vous n'en avez plus besoin. Sélectionnez l'instance temporaire, choisissez État de l'instance, Résilier l'instance.

Suppression de votre paire de clés

Lorsque vous supprimez une paire de clés, vous supprimez seulement la copie Amazon EC2 de la clé publique. La suppression d'une paire de clés n'affecte pas la clé privée présente sur votre ordinateur ni la clé publique sur les instances déjà lancées à l'aide de cette paire de clés. Vous ne pouvez pas lancer de nouvelle instance à l'aide d'une paire de clés supprimée, mais vous pouvez continuer à vous connecter à toutes les instances lancées à l'aide d'une paire de clés supprimée ultérieurement, tant que vous disposez du fichier de clé privée (`.pem`).

Si vous utilisez un groupe Auto Scaling (par exemple, dans un environnement Elastic Beanstalk), assurez-vous que la paire de clés que vous supprimez n'est pas spécifiée dans votre configuration de lancement. Amazon EC2 Auto Scaling lance une instance de remplacement en cas de détection d'une instance défectueuse. Cependant, le lancement de l'instance échoue si la paire de clés est introuvable.

Vous pouvez supprimer une paire de clés à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour supprimer votre paire de clés

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, cliquez sur Key Pairs.
3. Sélectionnez la paire de clés à supprimer et choisissez Delete (Supprimer).
4. Dans le champ de confirmation, entrez, `Delete` puis choisissez Delete (Supprimer).

Ancienne console

Pour supprimer votre paire de clés

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, sous NETWORK & SECURITY, choisissez Key Pairs.
3. Sélectionnez la paire de clés, puis choisissez Supprimer.
4. À l'invite, choisissez Oui.

AWS CLI

Pour supprimer votre paire de clés

Utilisez la commande de l'AWS CLI delete-key-pair.

PowerShell

Pour supprimer votre paire de clés

Utilisez la commande Outils AWS pour Windows PowerShell Remove-EC2KeypAir.

Si vous créez une AMI Linux à partir d'une instance, puis utilisez l'AMI pour lancer une nouvelle instance dans une autre région ou dans un compte différent, la nouvelle instance comprend la clé publique de l'instance d'origine. Vous pouvez ainsi vous connecter à la nouvelle instance à l'aide du même fichier de clé privée que celui utilisé pour l'instance d'origine. Vous pouvez supprimer cette clé publique de votre instance en supprimant son entrée du fichier `.ssh/authorized_keys` à l'aide de l'éditeur de texte de votre choix. Pour plus d'informations sur la gestion des utilisateurs sur votre instance et sur la fourniture d'un accès distant à l'aide d'une paire de clés spécifique, consultez Gestion des comptes d'utilisateur sur votre instance Amazon Linux (p. 660).

Groupes de sécurité Amazon EC2 pour les instances Linux

Un groupe de sécurité agit en tant que pare-feu virtuel pour vos instances EC2 afin de contrôler le trafic entrant et sortant. Les règles entrantes contrôlent le trafic entrant vers votre instance, et les règles sortantes contrôlent le trafic sortant de votre instance. Lorsque vous lancez une instance, vous pouvez spécifier un ou plusieurs groupes de sécurité. Si vous ne spécifiez pas de groupe de sécurité, Amazon EC2 utilise le groupe de sécurité par défaut. Vous pouvez ajouter des règles à chaque groupe de sécurité pour autoriser le trafic vers ou depuis ses instances associées. Vous pouvez modifier les règles pour un groupe de sécurité à la fois. Les nouvelles règles sont automatiquement appliquées à toutes les instances associées au groupe de sécurité. Lorsque Amazon EC2 décide d'autoriser ou non le trafic à atteindre une instance, toutes les règles issues de tous les groupes de sécurité associés à cette instance sont évaluées automatiquement.

Lorsque vous lancez une instance dans un VPC, vous devez spécifier un groupe de sécurité créé pour ce VPC. Après avoir lancé une instance, vous pouvez modifier ses groupes de sécurité. Les groupes de sécurité sont associés à des interfaces réseau. La modification des groupes de sécurité d'une instance change les groupes de sécurité associés à l'interface réseau principale (eth0). Pour de plus amples informations, veuillez consulter Modification des groupes de sécurité d'une instance dans le Amazon VPC Guide de l'utilisateur. Vous pouvez aussi modifier les groupes de sécurité associés à une autre interface réseau. Pour de plus amples informations, veuillez consulter Modification des attributs d'interface réseau (p. 857).

La sécurité est une responsabilité partagée entre AWS et vous-même. Pour plus d'informations, consultez Sécurité dans Amazon EC2 (p. 971). AWS fournit des groupes de sécurité comme un des outils permettant de sécuriser vos instances ; vous devez les configurer pour répondre à vos besoins en matière de sécurité. Si vous avez des exigences qui ne sont pas satisfaites par les groupes de sécurité, vous

pouvez maintenir votre propre pare-feu sur l'une de vos instances, quelle qu'elle soit, en plus de l'utilisation des groupes de sécurité.

Si vous devez autoriser le trafic vers une instance Windows, veuillez consulter Groupes de sécurité Amazon EC2 pour les instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Table des matières

- Règles des groupes de sécurité (p. 1063)
    - Suivi de la connexion (p. 1065)
- Groupes de sécurité par défaut (p. 1066)
- Custom security groups (p. 1066)
- Utilisation des groupes de sécurité (p. 1067)
    - Création d'un groupe de sécurité (p. 1067)
    - Copie d'un groupe de sécurité (p. 1068)
    - Affichage de vos groupes de sécurité (p. 1069)
    - Ajout de règles à un groupe de sécurité (p. 1069)
    - Mise à jour des règles des groupes de sécurité (p. 1072)
    - Suppression de règles d'un groupe de sécurité (p. 1073)
    - Suppression d'un groupe de sécurité (p. 1074)
- Référence des règles de groupe de sécurité (p. 1075)
    - Règles de serveur web (p. 1075)
    - Règles de serveur de base de données (p. 1076)
    - Règles pour la connexion à des instances à partir de votre ordinateur (p. 1077)
    - Règles pour la connexion à des instances à partir d'une instance avec le même groupe de sécurité (p. 1078)
    - Règles pour Ping/ICMP (p. 1078)
    - Règles de serveur DNS (p. 1079)
    - Règles Amazon EFS (p. 1079)
    - Règles Elastic Load Balancing (p. 1080)
    - Règles d'appairage de VPC (p. 1081)

Règles des groupes de sécurité

Les règles d'un groupe de sécurité contrôlent le trafic entrant autorisé à atteindre les instances associées au groupe de sécurité. Les règles contrôlent également le trafic sortant autorisé à les quitter.

Les caractéristiques des règles des groupes de sécurité sont les suivantes :

- Par défaut, les groupes de sécurité autorisent la totalité du trafic sortant.
- Les règles des groupes de sécurité sont toujours permissives ; vous ne pouvez pas créer de règles qui refusent l'accès.
- Les règles des groupes de sécurité vous permettent de filtrer le trafic en fonction des protocoles et des numéros de port.
- Les groupes de sécurité sont dynamiques. Si vous envoyez une demande à partir de votre instance, le trafic de la réponse à cette demande est autorisé, indépendamment des règles entrantes des groupes de sécurité. Pour les groupes de sécurité VPC, cela signifie aussi que les réponses au trafic entrant autorisé ont le droit d'être acheminées vers l'extérieur, indépendamment des règles sortantes. Pour plus d'informations, consultez Suivi de la connexion (p. 1065).

- Vous pouvez ajouter et supprimer des règles à tout moment. Vos modifications sont appliquées automatiquement aux instances associées au groupe de sécurité.

  L'effet de certaines modifications de règle peut dépendre de la manière dont le trafic est suivi. Pour plus d'informations, consultez Suivi de la connexion (p. 1065).
- Lorsque vous associez plusieurs groupes de sécurité à une instance, les règles de chaque groupe de sécurité sont effectivement agrégées pour créer un ensemble de règles. Amazon EC2 utilise cet ensemble de règles pour déterminer si l'accès doit être autorisé.

  Vous pouvez affecter plusieurs groupes de sécurité à une instance. Par conséquent, une instance peut avoir des centaines de règles qui s'appliquent. Cela peut entraîner des problèmes quand vous accédez à l'instance. Nous vous recommandons de condenser vos règles autant que possible.

Pour chaque règle, vous spécifiez les informations suivantes :

- Nom : nom du groupe de sécurité (par exemple, `my-security-group`).

  Un nom peut contenir jusqu'à 255 caractères. Les caractères autorisés sont : a-z, A-Z, 0-9, espaces et ._-:/()#,@[]+=;{}!$*. Lorsque le nom contient des espaces de fin, nous supprimons les espaces lorsque nous enregistrons le nom. Par exemple, si vous entrez « Test Security Group » pour le nom, nous le stockons comme « Test Security Group ».
- Protocole : le protocole à autoriser. Les protocoles les plus courants sont 6 (TCP) 17 (UDP) et 1 (ICMP).
- Port range (Plage de ports) : pour TCP, UDP ou un protocole personnalisé : la plage de ports autorisée. Vous pouvez spécifier un seul numéro de port (par exemple, `22`), ou une plage de numéros de port (par exemple, `7000-8000`).
- ICMP type and code (Type et code ICMP) : pour ICMP, le code et le type ICMP.
- Source or destination (Source ou destination) : la source (règles entrantes) ou la destination (règles sortante) pour le trafic. Spécifiez l'une des options suivantes :
  - Une adresse IPv4 individuelle. Vous devez utiliser la longueur de préfixe `/32`, par exemple `203.0.113.1/32`.
  - Une adresse IPv6 individuelle. Vous devez utiliser la longueur de préfixe `/128`, par exemple `2001:db8:1234:1a00::123/128`.
  - Plage d'adresses IPv4, en notation de bloc d'adresse CIDR : par exemple, `203.0.113.0/24`.
  - Plage d'adresses IPv6, en notation de bloc d'adresse CIDR : par exemple, `2001:db8:1234:1a00::/64`.
  - Un ID de liste de préfixes, par exemple `pl-1234abc1234abc123`. Pour de plus amples informations, veuillez consulter Listes de préfixes dans le Amazon VPC Guide de l'utilisateur.
  - Un autre groupe de sécurité. Les instances associées au groupe de sécurité spécifié peuvent ainsi accéder aux instances associées à ce groupe de sécurité. (Notez que cela n'ajoute pas de règles du groupe de sécurité source à ce groupe de sécurité.) Vous spécifiez l'un des groupes de sécurité suivants :
    - Groupe de sécurité en cours
    - Un groupe de sécurité différent pour le même VPC
    - Un groupe de sécurité différent pour un VPC homologue dans une connexion d'appairage de VPC
- (Facultatif) Description : vous pouvez ajouter une description pour la règle, par exemple, pour vous aider à l'identifier ultérieurement. Une description peut inclure jusqu'à 255 caractères. Les caractères autorisés sont : a-z, A-Z, 0-9, espaces et ._-:/()#,@[]+=;{}!$*.

Quand vous spécifiez un groupe de sécurité comme source ou destination d'une règle, celle-ci affecte toutes les instances associées au groupe de sécurité. Le trafic entrant est autorisé en fonction des adresses IP privées des instances associées au groupe de sécurité source (et non des adresses IP Elastic ou des adresses IP publiques). Pour plus d'informations sur les adresses IP, consultez Adressage IP des

instances Amazon EC2 (p. 808). Si votre règle de groupe de sécurité fait référence à un groupe de sécurité dans un VPC pair et si le groupe de sécurité référencé ou la connexion d'appairage de VPC est supprimée, la règle est marquée comme étant obsolète. Pour plus d'informations, consultez Utilisation de règles de groupes de sécurité obsolètes dans le Amazon VPC Peering Guide.

S'il existe plusieurs règles pour un port spécifique, Amazon EC2 applique la règle la lus permissive. Par exemple, si vous avez une règle qui autorise l'accès au port TCP 22 (SSH) à partir de l'adresse IP `203.0.113.1` et une autre règle qui autorise l'accès au port TCP 22 à partir de tous, chacun a accès au port TCP 22.

Suivi de la connexion

Vos groupes de sécurité utilisent le suivi de connexion pour suivre les informations sur le trafic en provenance ou à destination de l'instance. Les règles s'appliquent en fonction de l'état de connexion du trafic pour déterminer si le trafic est autorisé ou refusé. Cette approche permet aux groupes de sécurité d'être avec état. Les groupes de sécurité peuvent ainsi être avec état. Les réponses au trafic entrant sont autorisées à transiter en dehors de l'instance, indépendamment des règles sortantes des groupes de sécurité (et inversement). Par exemple, si vous déclenchez une commande ICMP `ping` sur votre instance à partir de votre ordinateur familial et que vos règles entrantes des groupes de sécurité autorisent le trafic ICMP, les informations sur la connexion (informations sur le port incluses) sont suivies. Le trafic de la réponse à partir de l'instance pour la commande `ping` n'est pas suivie comme nouvelle demande, mais plutôt comme connexion établie, et est autorisée à transiter en dehors de l'instance, même si vos règles sortantes des groupes de sécurité limitent le trafic ICMP sortant.

Certains flux de trafic ne sont pas suivis. Si une règle de groupe de sécurité autorise les flux TCP ou UDP pour la totalité du trafic (`0.0.0.0/0` ou `::/0`) et qu'il existe une règle correspondante dans l'autre sens qui autorise tout le trafic de la réponse (`0.0.0.0/0` ou `::/0`) pour tous les ports (0-65535), le flux du trafic n'est pas suivi. Par conséquent, le trafic de la réponse est autorisé à transiter en fonction de la règle entrante ou sortante qui autorise le trafic de la réponse, et non des informations de suivi.

Dans l'exemple suivant, le groupe de sécurité dispose de règles entrantes pour le trafic TCP et ICMP entrant, et de règles sortantes qui autorisent tout le trafic IPv4 et IPv6 sortant.

Règles entrantes		
Type de protocole	Numéro de port	IP Source
TCP	22 (SSH)	203.0.113.1/32
TCP	80 (HTTP)	0.0.0.0/0
TCP	80 (HTTP)	::/0
ICMP	Tous	0.0.0.0/0
Règles sortantes		
Type de protocole	Numéro de port	IP de destination
Tous	Tous	0.0.0.0/0
Tous	Tous	::/0

Le trafic TCP sur le port 22 (SSH) vers et depuis l'instance est suivi, car la règle entrante autorise le trafic depuis `203.0.113.1/32` uniquement, et pas toutes les adresses IP (`0.0.0.0/0`). Le trafic TCP sur le port 80 (HTTP) vers et depuis l'instance n'est pas suivi, car les règles entrantes et sortantes autorisent tout le trafic (`0.0.0.0/0` ou `::/0`). Le trafic ICMP est toujours suivi, quelles que soient les règles. Si vous

supprimez la règle sortante du groupe de sécurité, tout le trafic vers et depuis l'instance est suivi, y compris le trafic sur le port 80 (HTTP).

Un flux de trafic non suivi est immédiatement interrompu si la règle qui active le flux est supprimée ou modifiée. Par exemple, si vous disposez d'une règle sortante ouverte (0.0.0.0/0) et que vous supprimez une règle qui autorise tout le trafic SSH (port TCP 22) entrant (0.0.0.0/0) vers l'instance (ou que vous la modifiez de telle sorte que la connexion ne soit plus autorisée), vos connexions SSH existantes à l'instance sont immédiatement supprimées. La connexion n'était pas suivie auparavant, de sorte que la modification rompt la connexion. D'autre part, si vous avez une règle entrante plus étroite qui autorise initialement la connexion SSH (ce qui signifie que la connexion a été suivie), mais que vous modifiez cette règle pour ne plus autoriser de nouvelles connexions à partir de l'adresse du client SSH actuel, la connexion existante ne sera pas rompue en modifiant la règle.

Pour les protocoles autre que TCP, UDP ou ICMP, seuls l'adresse IP et le numéro de protocole sont suivis. Si votre instance envoie le trafic vers un autre hôte (hôte B) et que l'hôte B initie le même type de trafic vers votre instance dans une demande distincte dans un délai de 600 secondes après la demande ou réponse d'origine, votre instance l'accepte indépendamment des règles de groupe de sécurité entrantes. Votre instance l'accepte car elle est considérée comme un trafic de réponse.

Pour vous assurer que le trafic est immédiatement interrompu lorsque vous supprimez une règle de groupe de sécurité ou pour vous assurer que tout le trafic entrant est soumis à des règles de pare-feu, vous pouvez utiliser une liste ACL réseau pour votre sous-réseau. Les ACL réseau sont sans état et n'autorisent donc pas automatiquement le trafic de réponse. Pour plus d'informations, consultez ACL réseau dans le Amazon VPC Guide de l'utilisateur.

Groupes de sécurité par défaut

Votre compte AWS possède automatiquement un groupe de sécurité par défaut pour le VPC par défaut dans chaque région. Si vous ne spécifiez pas un groupe de sécurité lorsque vous lancez une instance, celle-ci est automatiquement associée au groupe de sécurité par défaut pour le VPC.

Un groupe de sécurité par défaut est appelé `default` et possède un ID attribué par AWS. Les règles par défaut pour chaque groupe de sécurité par défaut sont les suivantes :

- Autoriser le trafic entrant seulement à partir des autres instances associées au groupe de sécurité par défaut Le groupe de sécurité par défaut se spécifie lui-même comme groupe de sécurité source dans ses règles entrantes.
- Autorise la totalité du trafic sortant à partir de l'instance.

Vous pouvez ajouter ou supprimer des règles entrantes et sortantes pour n'importe quel groupe de sécurité par défaut.

Vous ne pouvez pas supprimer un groupe de sécurité par défaut. Si vous essayez de supprimer un groupe de sécurité par défaut, vous voyez l'erreur suivante : `Client.CannotDelete: the specified group: "sg-51530134" name: "default" cannot be deleted by a user`.

Custom security groups

Si vous ne voulez pas que vos instances utilisent le groupe de sécurité par défaut, vous pouvez créer vos propres groupes de sécurité et les spécifiez quand vous démarrez vos instances. Vous pouvez créer plusieurs groupes de sécurité pour refléter les différents rôles joués par vos instances ; par exemple, un serveur web ou un serveur de base de données.

Quand vous créez un groupe de sécurité, vous devez lui attribuer un nom et une description. Les noms et les descriptions des groupes de sécurité peuvent comporter jusqu'à 255 caractères de long et uniquement les caractères suivants :

a-z, A-Z, 0-9, espace et ._-:/()#,@[]+=&;{}!$*

Un nom de groupe de sécurité ne peut pas commencer par `sg-`. Un nom de groupe de sécurité doit être unique pour le VPC.

Les règles par défaut pour chaque groupe de sécurité que vous créez sont les suivantes :

- N'autorise aucun trafic entrant
- Autorise tout le trafic sortant

Après avoir créé un groupe de sécurité, vous pouvez modifier ses règles entrantes pour refléter le type de trafic entrant que vous voulez pour atteindre les instances associées. Vous pouvez aussi modifier ses règles sortantes.

Pour plus d'informations sur les règles que vous pouvez ajouter à un groupe de sécurité, consultez Référence des règles de groupe de sécurité (p. 1075).

Utilisation des groupes de sécurité

Vous pouvez assigner un groupe de sécurité à une instance lorsque vous lancez l'instance. Quand vous ajoutez ou supprimez des règles, ces modifications sont automatiquement appliquées à toutes les instances auxquelles vous avez affecté le groupe de sécurité.

Après avoir lancé une instance, vous pouvez modifier ses groupes de sécurité. Pour plus d'informations, consultez Modification des groupes de sécurité d'une instance dans le Amazon VPC Guide de l'utilisateur.

Vous pouvez créer, afficher, mettre à jour et supprimer des groupes de sécurité et des règles de groupe de sécurité à l'aide de la console Amazon EC2 et des outils de ligne de commande.

Tâches

- Création d'un groupe de sécurité (p. 1067)
- Copie d'un groupe de sécurité (p. 1068)
- Affichage de vos groupes de sécurité (p. 1069)
- Ajout de règles à un groupe de sécurité (p. 1069)
- Mise à jour des règles des groupes de sécurité (p. 1072)
- Suppression de règles d'un groupe de sécurité (p. 1073)
- Suppression d'un groupe de sécurité (p. 1074)

Création d'un groupe de sécurité

Vous pouvez créer un groupe de sécurité personnalisé à l'aide de l'une des méthodes suivantes. Vous devez spécifier le VPC pour lequel vous créez le groupe de sécurité.

Nouvelle console

Pour créer un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez Créer un groupe de sécurité.
4. Dans la section Basic details (Détails de base) procédez comme suit.

   a. Entrez un nom descriptif et une brève description pour le groupe de sécurité. Le nom et la description peuvent contenir jusqu'à 255 caractères, et ils peuvent inclure `a-z, A-Z, 0-9, spaces, and ._-:/()#,@[]+=&;{}!$*`.

   b. Pour VPC, choisissez le VPC dans lequel créer le groupe de sécurité. Le groupe de sécurité ne peut être utilisé que dans le VPC dans lequel il est créé.
5. Vous pouvez maintenant ajouter des règles de groupe de sécurité ou les ajouter à tout moment après avoir créé le groupe de sécurité. Pour de plus amples informations sur l'ajout de règles de groupe de sécurité, veuillez consulter Ajout de règles à un groupe de sécurité (p. 1069).
6. Sélectionnez Créer.

Ancienne console

Pour créer un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez Create Security Group.
4. Attribuez un nom et une description au groupe de sécurité.
5. Pour VPC, choisissez l'ID du VPC.
6. Vous pouvez commencer à ajouter des règles ou choisir Créer pour créer le groupe de sécurité maintenant (vous pourrez toujours ajouter des règles par la suite). Pour plus d'informations sur l'ajout de règles, consultez Ajout de règles à un groupe de sécurité (p. 1069).

Command line

Pour créer un groupe de sécurité

Utilisez l'une des commandes suivantes :

- create-security-group (AWS CLI)
- New-EC2SecurityGroup (Outils AWS pour Windows PowerShell)

Copie d'un groupe de sécurité

Vous pouvez créer un nouveau groupe de sécurité en créant la copie d'un groupe existant. Lorsque vous copiez un groupe de sécurité, la copie est créée avec les mêmes règles entrantes et sortantes que le groupe de sécurité d'origine. Si le groupe de sécurité d'origine se trouve dans un VPC, la copie est créée dans le même VPC, sauf si vous en spécifiez un autre.

La copie reçoit un nouvel ID de groupe de sécurité unique et vous devez lui donner un nom. Vous pouvez également ajouter une description.

Vous ne pouvez pas copier un groupe de sécurité d'une région vers une autre région.

Vous pouvez créer une copie d'un groupe de sécurité à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour créer un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez le groupe de sécurité à copier et choisissez Actions, Copy to new security group (Copier vers un nouveau groupe de sécurité).
4. Spécifiez un nom et une description facultative, puis modifiez les règles du VPC et du groupe de sécurité si nécessaire.
5. Sélectionnez Créer.

Ancienne console

Pour créer un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Groupes de sécurité.
3. Sélectionnez le groupe de sécurité que vous voulez copier, puis choisissez Actions, Copier vers le nouveau.
4. La boîte de dialogue Créer un groupe de sécurité s'ouvre. Elle contient les règles du groupe de sécurité existant. Attribuez un nom et une description à votre nouveau groupe de sécurité. Pour VPC, choisissez l'ID du VPC. Lorsque vous avez terminé, cliquez sur Créer.

Affichage de vos groupes de sécurité

Vous pouvez afficher des informations sur vos groupes de sécurité à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour afficher vos groupes de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Vos groupes de sécurité sont répertoriés. Pour afficher les détails d'un groupe de sécurité spécifique, y compris ses règles entrantes et sortantes, choisissez son ID dans la colonne Security group ID (ID du groupe de sécurité).

Ancienne console

Pour afficher vos groupes de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Groupes de sécurité.
3. (Facultatif) Sélectionnez ID de VPC dans la liste des filtres, puis choisissez l'ID du VPC.
4. Sélectionnez un groupe de sécurité. Les informations générales sont affichées dans l'onglet Description, les règles entrantes dans l'onglet Entrant, les règles sortantes dans l'onglet Sortant, et les balises dans l'onglet Balises.

Command line

Pour afficher vos groupes de sécurité

Utilisez l'une des commandes suivantes.

- describe-security-groups (AWS CLI)
- Get-EC2SecurityGroup (Outils AWS pour Windows PowerShell)

Ajout de règles à un groupe de sécurité

Lorsque vous ajoutez une règle à un groupe de sécurité, la nouvelle règle est automatiquement appliquée à toutes les instances associées au groupe de sécurité. Il peut y avoir un court délai avant l'application de la règle. Pour plus d'informations sur le choix de règles de groupe de sécurité pour des types d'accès

spécifiques, consultez Référence des règles de groupe de sécurité (p. 1075). Pour connaître les quotas de règles de groupe de sécurité, veuillez consulter les quotas Amazon VPC dans le Guide de l'utilisateur Amazon VPC.

Vous pouvez ajouter des règles à un groupe de sécurité à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour ajouter une règle d'entrée à un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Dans la liste, sélectionnez le groupe de sécurité et choisissez Actions, Modifier les règles entrantes.
4. Choisissez Add rule (Ajouter une règle) et procédez comme suit :
   a. Pour Type, choisissez le type de protocole à autoriser.
      - Si vous choisissez un protocole TCP ou UDP personnalisé, vous devez entrer manuellement la plage de ports à autoriser.
      - Si vous sélectionnez un protocole ICMP personnalisé, choisissez le nom de type ICMP dans Protocole et, le cas échéant, le nom de code dans Plage de ports.
      - Si vous choisissez un autre type, le protocole et la plage de ports sont configurés automatiquement.
   b. Pour Source, effectuez l'une des opérations suivantes.
      - Choisissez Personnalisé, puis entrez une adresse IP en notation CIDR, un bloc d'adresse CIDR, un autre groupe de sécurité ou une liste de préfixes à partir duquel autoriser le trafic entrant.
      - Choisissez Anywhere pour autoriser tout le trafic entrant du protocole spécifié à atteindre votre instance. Cette option ajoute automatiquement le bloc d'adresse CIDR IPv4 `0.0.0.0/0` en tant que source autorisée. Cette solution est acceptable pour une brève durée dans un environnement de test, mais n'est pas sécurisée pour les environnements de production. Dans un environnement de production, vous autorisez uniquement l'accès à votre instance pour une adresse IP ou une plage d'adresses spécifiques.

        Si votre groupe de sécurité se trouve dans un VPC activé pour IPv6, cette option ajoute automatiquement une deuxième règle pour le trafic IPv6 (`::/0`).
      - Choisissez My IP (Mon IP) pour autoriser le trafic entrant uniquement à partir de l'adresse IPv4 publique de votre ordinateur local.
   c. Dans Description, vous pouvez éventuellement spécifier une description de la règle.
5. Choisissez Prévisualiser les modifications, Enregistrer les règles.

Pour ajouter une règle sortante à un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Dans la liste, sélectionnez le groupe de sécurité et choisissez Actions, Edit outbound rules (Modifier les règles sortantes).
4. Choisissez Add rule (Ajouter une règle) et procédez comme suit :
   a. Pour Type, choisissez le type de protocole à autoriser.
      - Si vous choisissez un protocole TCP ou UDP personnalisé, vous devez entrer manuellement la plage de ports à autoriser.

- Si vous sélectionnez un protocole ICMP personnalisé, choisissez le nom de type ICMP dans Protocole et, le cas échéant, le nom de code dans Plage de ports.
- Si vous choisissez un autre type, le protocole et la plage de ports sont configurés automatiquement.

b. Pour Destination, effectuez l'une des opérations suivantes.

- Choisissez Personnalisé, puis entrez une adresse IP en notation CIDR, un bloc d'adresse CIDR ou un autre groupe de sécurité pour lequel autoriser le trafic sortant.
- Choisissez Anywhere pour autoriser le trafic sortant vers toutes les adresses IP. Cette option ajoute automatiquement le bloc d'adresse CIDR IPv4 `0.0.0.0/0` en tant que source autorisée.

  Si votre groupe de sécurité se trouve dans un VPC activé pour IPv6, cette option ajoute automatiquement une deuxième règle pour le trafic IPv6 (`::/0`).
- Choisissez My IP (Mon IP) pour autoriser le trafic sortant uniquement vers l'adresse IPv4 publique de votre ordinateur local.

c. Dans Description, vous pouvez éventuellement spécifier une description de la règle.

5. Choisissez Preview changes (Prévisualiser les modifications), Confirm (Confirmer).

Ancienne console

Pour ajouter des règles à un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Groupes de sécurité, puis sélectionnez le groupe de sécurité.
3. Dans l'onglet Entrant, choisissez Modifier.
4. Dans la boîte de dialogue, sélectionnez Ajouter une règle et effectuez les opérations suivantes :

   - Pour Type, sélectionnez le protocole.
   - Si vous sélectionnez un protocole TCP ou UDP personnalisé, spécifiez la plage de ports dans Plage de ports.
   - Si vous sélectionnez un protocole ICMP personnalisé, choisissez le nom de type ICMP dans Protocole et, le cas échéant, le nom de code dans Plage de ports.
   - Pour Source, choisissez l'une des options suivantes :
     - Personnalisé : dans le champ fourni, vous devez spécifier une adresse IP en notation CIDR, un bloc d'adresse CIDR ou un autre groupe de sécurité.
     - N'importe où : ajoute automatiquement le bloc d'adresse CIDR IPv4 `0.0.0.0/0`. Cette option permet à l'ensemble du trafic du type spécifié d'accéder à votre instance. Cette solution est acceptable pour une brève durée dans un environnement de test, mais n'est pas sécurisée pour les environnements de production. Dans un environnement de production, vous autorisez uniquement l'accès à votre instance pour une adresse IP ou une plage d'adresses spécifiques.

       Si votre groupe de sécurité est dans un VPC qui est activé pour IPv6, l'option N'importe où crée deux règles, l'une pour le trafic IPv4 (`0.0.0.0/0`) et l'autre pour le trafic IPv6 (`::/0`).
     - Mon IP : ajoute automatiquement l'adresse IPv4 publique de votre ordinateur local.
   - Pour Description, vous pouvez éventuellement spécifier une description pour la règle.

   Pour plus d'informations sur les types de règles que vous pouvez ajouter, consultez Référence des règles de groupe de sécurité (p. 1075).
5. Choisissez Save.

6. Vous pouvez aussi spécifier des règles sortantes. Dans l'onglet Sortant, choisissez Modifier, Ajouter une règle, puis procédez de la façon suivante :
   - Pour Type, sélectionnez le protocole.
   - Si vous sélectionnez un protocole TCP ou UDP personnalisé, spécifiez la plage de ports dans Plage de ports.
   - Si vous sélectionnez un protocole ICMP personnalisé, choisissez le nom de type ICMP dans Protocole et, le cas échéant, le nom de code dans Plage de ports.
   - Pour Destination, choisissez l'une des options suivantes :
     - Personnalisé : dans le champ fourni, vous devez spécifier une adresse IP en notation CIDR, un bloc d'adresse CIDR ou un autre groupe de sécurité.
     - N'importe où : ajoute automatiquement le bloc d'adresse CIDR IPv4 `0.0.0.0/0`. Cette option permet au trafic sortant d'accéder à toutes les adresses IP.

       Si votre groupe de sécurité est dans un VPC qui est activé pour IPv6, l'option N'importe où crée deux règles, l'une pour le trafic IPv4 (`0.0.0.0/0`) et l'autre pour le trafic IPv6 (`::/0`).
     - Mon IP : ajoute automatiquement l'adresse IP de votre ordinateur local.
   - Pour Description, vous pouvez éventuellement spécifier une description pour la règle.
7. Choisissez Save.

Command line

Pour ajouter des règles à un groupe de sécurité

Utilisez l'une des commandes suivantes.

- authorize-security-group-ingress (AWS CLI)
- Grant-EC2SecurityGroupIngress (Outils AWS pour Windows PowerShell)

Pour ajouter une ou plusieurs règles de trafic sortant à un groupe de sécurité

Utilisez l'une des commandes suivantes.

- authorize-security-group-egress (AWS CLI)
- Grant-EC2SecurityGroupEgress (Outils AWS pour Windows PowerShell)

Mise à jour des règles des groupes de sécurité

Vous pouvez mettre à jour une règle de groupe de sécurité à l'aide de l'une des méthodes suivantes.

Nouvelle console

Lorsque vous modifiez le protocole, la plage de ports, ou la source ou destination de la règle existante d'un groupe de sécurité à l'aide de la console, cette dernière supprime la règle existante et en ajoute une nouvelle.

Pour mettre à jour un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez le groupe de sécurité à mettre à jour, choisissez Actions, puis choisissez Edit inbound rules (Modifier les règles entrantes) pour mettre à jour une règle pour le trafic entrant

ou Edit outbound rules (Modifier les règles sortantes) pour mettre à jour une règle pour le trafic sortant.

4. Mettez à jour la règle selon vos besoins, puis choisissez Preview changes (Aperçu des modifications), Confirm (Confirmer).

Ancienne console

Lorsque vous modifiez le protocole, la plage de ports, ou la source ou destination de la règle existante d'un groupe de sécurité à l'aide de la console, cette dernière supprime la règle existante et en ajoute une nouvelle.

Pour mettre à jour un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez le groupe de sécurité à mettre à jour, puis choisissez l'onglet Règles entrantes afin de mettre à jour une règle pour le trafic entrant ou Règles sortantes afin de mettre à jour une règle pour le trafic sortant.
4. Choisissez Modifier.
5. Modifiez l'entrée de règle comme nécessaire, puis choisissez Enregistrer.

Command line

Vous ne pouvez pas modifier le protocole, la plage de ports, ou la source ou destination de la règle existante d'un groupe de sécurité à l'aide de l'API Amazon EC2 ou d'un outil de ligne de commande. Vous devez plutôt supprimer la règle existante, puis ajouter une nouvelle règle. Vous pouvez toutefois mettre à jour la description d'une règle existante.

Pour mettre à jour la description d'une règle entrante existante

Utilisez l'une des commandes suivantes.

- update-security-group-rule-descriptions-ingress (AWS CLI)
- Update-EC2SecurityGroupRuleIngressDescription (Outils AWS pour Windows PowerShell)

Pour mettre à jour la description d'une règle sortante existante

Utilisez l'une des commandes suivantes.

- update-security-group-rule-descriptions-egress (AWS CLI)
- Update-EC2SecurityGroupRuleEgressDescription (Outils AWS pour Windows PowerShell)

Suppression de règles d'un groupe de sécurité

Lorsque vous supprimez une règle d'un groupe de sécurité, la modification est automatiquement appliquée à toutes les instances associées au groupe de sécurité.

Vous pouvez supprimer des règles d'un groupe de sécurité à l'aide de l'une des méthodes suivantes.

Nouvelle console

Pour supprimer une règle de groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez le groupe de sécurité à mettre à jour, choisissez Actions, puis choisissez Edit inbound rules (Modifier les règles entrantes) pour supprimer une règle entrante ou Edit outbound rules (Modifier les règles sortantes) pour supprimer une règle sortante.
4. Cliquez sur le bouton Supprimer à droite de la règle à supprimer.
5. Choisissez Preview changes (Prévisualiser les modifications), Confirm (Confirmer).

Ancienne console

Pour supprimer une règle de groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Groupes de sécurité.
3. Sélectionnez un groupe de sécurité.
4. Dans l'onglet Entrant (pour les règles entrantes) ou Sortant (pour les règles sortantes), choisissez Modifier. Choisissez Supprimer (icône en forme de croix) à côté de chaque règle que vous devez supprimer.
5. Choisissez Save.

Command line

Supprimer une ou plusieurs règles de trafic entrant d'un groupe de sécurité

Utilisez l'une des commandes suivantes.

- revoke-security-group-ingress (AWS CLI)
- Revoke-EC2SecurityGroupIngress (Outils AWS pour Windows PowerShell)

Pour supprimer une ou plusieurs règles de trafic sortant d'un groupe de sécurité

Utilisez l'une des commandes suivantes.

- revoke-security-group-egress (AWS CLI)
- Revoke-EC2SecurityGroupEgress (Outils AWS pour Windows PowerShell)

Suppression d'un groupe de sécurité

Vous ne pouvez pas supprimer un groupe de sécurité associé à une instance. Vous ne pouvez pas supprimer le groupe de sécurité par défaut. Vous ne pouvez pas supprimer un groupe de sécurité référencé par un autre groupe de sécurité dans le même VPC. Si votre groupe de sécurité est référencé par l'une de ses propres règles, vous devez supprimer la règle avant de pouvoir supprimer le groupe de sécurité.

Nouvelle console

Pour supprimer un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Groupes de sécurité.
3. Sélectionnez le groupe de sécurité à supprimer et choisissez Actions, Delete security group (Supprimer le groupe de sécurité), et Delete (Supprimer).

Ancienne console

Pour supprimer un groupe de sécurité

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Groupes de sécurité.
3. Sélectionnez un groupe de sécurité, puis choisissez Actions, Supprimer le groupe de sécurité.
4. Sélectionnez Oui, supprimer.

Command line

Pour supprimer un groupe de sécurité

Utilisez l'une des commandes suivantes.

- delete-security-group (AWS CLI)
- Remove-EC2SecurityGroup (Outils AWS pour Windows PowerShell)

Référence des règles de groupe de sécurité

Vous pouvez créer un groupe de sécurité et ajouter des règles qui reflètent le rôle de l'instance qui est associée à ce groupe. Par exemple, une instance configurée en tant que serveur web nécessite des règles de groupe de sécurité qui autorisent l'accès HTTP et HTTPS entrant. De même, une instance de base de données a besoin de règles permettant l'accès au type de base de données, telles que l'accès via le port 3306 pour MySQL.

Voici des exemples de types de règles que vous pouvez ajouter à des groupes de sécurité pour des types d'accès spécifiques.

Exemples

- Règles de serveur web (p. 1075)
- Règles de serveur de base de données (p. 1076)
- Règles pour la connexion à des instances à partir de votre ordinateur (p. 1077)
- Règles pour la connexion à des instances à partir d'une instance avec le même groupe de sécurité (p. 1078)
- Règles pour Ping/ICMP (p. 1078)
- Règles de serveur DNS (p. 1079)
- Règles Amazon EFS (p. 1079)
- Règles Elastic Load Balancing (p. 1080)
- Règles d'appairage de VPC (p. 1081)

Règles de serveur web

Les règles entrantes suivantes autorisent l'accès HTTP et HTTPS à partir de n'importe quelle adresse IP. Si votre VPC est activé pour IPv6, vous pouvez ajouter des règles pour contrôler le trafic HTTP et HTTPS entrant à partir d'adresses IPv6.

Type de protocole	Numéro de protocole	Port	IP Source	Remarques
TCP	6	80 (HTTP)	0.0.0.0/0	Autorise l'accès HTTP entrant à

Type de protocole	Numéro de protocole	Port	IP Source	Remarques
				partir de n'importe quelle adresse IPv4
TCP	6	443 (HTTPS)	0.0.0.0/0	Autorise l'accès HTTPS entrant à partir de n'importe quelle adresse IPv4
TCP	6	80 (HTTP)	::/0	Autorise l'accès HTTP entrant à partir de n'importe quelle adresse IPv6.
TCP	6	443 (HTTPS)	::/0	Autorise l'accès HTTPS entrant à partir de n'importe quelle adresse IPv6.

Règles de serveur de base de données

Les règles entrantes suivantes sont des exemples de règles que vous pouvez ajouter pour un accès à une base de données selon le type de base de données que vous exécutez sur votre instance. Pour plus d'informations sur les instances Amazon RDS, consultez le Amazon RDS Guide de l'utilisateur.

Pour l'adresse IP source, spécifiez l'une des options suivantes :

- Une adresse IP spécifique ou une plage d'adresses IP (en notation de bloc CIDR) de votre réseau local
- Un ID de groupe de sécurité pour un groupe d'instances qui accèdent à la base de données

Type de protocole	Numéro de protocole	Port	Remarques
TCP	6	1433 (MS SQL)	Port par défaut pour accéder à une base de données Microsoft SQL Server, par exemple, sur une instance Amazon RDS
TCP	6	3306 (MYSQL/Aurora)	Port par défaut pour accéder à une base MySQL ou Aurora, par exemple, sur une instance Amazon RDS
TCP	6	5439 (Redshift)	Port par défaut pour accéder à une base de données de cluster Amazon Redshift.

Type de protocole	Numéro de protocole	Port	Remarques
TCP	6	5432 (PostgreSQL)	Port par défaut pour accéder à une base de données PostgreSQL, par exemple, sur une instance Amazon RDS
TCP	6	1521 (Oracle)	Port par défaut pour accéder à une base de données Oracle, par exemple, sur une instance Amazon RDS

Vous pouvez éventuellement restreindre le trafic sortant de vos serveurs de base de données. Par exemple, vous pouvez autoriser l'accès à Internet pour les mises à jour logicielles, mais limiter tous les autres types de trafic. Vous devez d'abord supprimer la règle sortante par défaut qui autorise tout le trafic sortant.

Type de protocole	Numéro de protocole	Port	IP de destination	Remarques
TCP	6	80 (HTTP)	0.0.0.0/0	Autorise l'accès HTTP sortant vers toute adresse IPv4
TCP	6	443 (HTTPS)	0.0.0.0/0	Autorise l'accès HTTPS sortant vers toute adresse IPv4
TCP	6	80 (HTTP)	::/0	(VPC activé pour IPv6 uniquement) Autorise l'accès HTTP sortant vers toute adresse IPv6
TCP	6	443 (HTTPS)	::/0	(VPC activé pour IPv6 uniquement) Autorise l'accès HTTPS sortant vers toute adresse IPv6

Règles pour la connexion à des instances à partir de votre ordinateur

Pour se connecter à votre instance, votre groupe de sécurité doit avoir des règles entrantes qui autorisent l'accès SSH (pour les instances Linux) ou l'accès RDP (pour les instances Windows).

Type de protocole	Numéro de protocole	Port	IP Source
TCP	6	22 (SSH)	Adresse IPv4 publique de votre ordinateur ou

Type de protocole	Numéro de protocole	Port	IP Source
			plage d'adresses IP (en notation de bloc CIDR) de votre réseau local. Si votre VPC est activé pour IPv6 et que votre instance a une adresse IPv6, vous pouvez entrer une adresse ou une plage d'adresses IPv6.
TCP	6	3389 (RDP)	Adresse IPv4 publique de votre ordinateur ou plage d'adresses IP (en notation de bloc CIDR) de votre réseau local. Si votre VPC est activé pour IPv6 et que votre instance a une adresse IPv6, vous pouvez entrer une adresse ou une plage d'adresses IPv6.

Règles pour la connexion à des instances à partir d'une instance avec le même groupe de sécurité

Pour autoriser les instances associées au même groupe de sécurité à communiquer les unes avec les autres, vous devez à cette fin ajouter des règles explicitement.

Le tableau suivant décrit la règle entrante pour un groupe de sécurité qui permet aux instances associées de communiquer les unes avec les autres. La règle autorise tous les types de trafic.

Type de protocole	Numéro de protocole	Ports	IP Source
-1 (Tout)	-1 (Tout)	-1 (Tout)	ID du groupe de sécurité

Règles pour Ping/ICMP

La commande `ping` est un type de trafic ICMP. Pour envoyer une commande ping à votre instance, vous devez ajouter la règle ICMP entrante suivante.

Type de protocole	Numéro de protocole	ICMP type	Code ICMP	IP Source
ICMP	1	8 (écho)	N/A	Adresse IPv4 publique de votre ordinateur ou plage d'adresses IPv4 (en notation

Type de protocole	Numéro de protocole	ICMP type	Code ICMP	IP Source
				de bloc CIDR) de votre réseau local

Pour utiliser la commande `ping6` afin d'effectuer un test ping sur l'adresse IPv6 pour votre instance, vous devez ajouter la règle ICMPv6 entrante suivante.

Type de protocole	Numéro de protocole	ICMP type	Code ICMP	IP Source
ICMPv6	58	128 (écho)	0	Adresse IPv6 publique de votre ordinateur ou plage d'adresses IPv6 (en notation de bloc CIDR) de votre réseau local

Règles de serveur DNS

Si vous avez configuré votre instance EC2 en tant que serveur DNS, vous devez vous assurer que le trafic TCP et UDP peut accéder à votre serveur DNS via le port 53.

Pour l'adresse IP source, spécifiez l'une des options suivantes :

- Adresse IP ou plage d'adresses IP (en notation de bloc CIDR) d'un réseau
- L'ID d'un groupe de sécurité pour l'ensemble d'instances de votre réseau devant d'accéder au serveur DNS

Type de protocole	Numéro de protocole	Port
TCP	6	53
UDP	17	53

Règles Amazon EFS

Si vous utilisez un système de fichiers Amazon EFS avec vos instances Amazon EC2, le groupe de sécurité que vous associez à vos cibles de montage Amazon EFS doit autoriser le trafic via le protocole NFS.

Type de protocole	Numéro de protocole	Ports	IP Source	Remarques
TCP	6	2049 (NFS)	ID du groupe de sécurité.	Autorise l'accès NFS entrant à partir des ressources (y compris la cible

Type de protocole	Numéro de protocole	Ports	IP Source	Remarques
				de montage) associées à ce groupe de sécurité.

Pour monter un système de fichiers Amazon EFS sur votre instance Amazon EC2 vous devez vous connecter à votre instance. Par conséquent, le groupe de sécurité associé à votre instance doit avoir des règles qui autorisent le trafic SSH entrant à partir de votre ordinateur local ou de votre réseau local.

Type de protocole	Numéro de protocole	Ports	IP Source	Remarques
TCP	6	22 (SSH)	Plage d'adresses IP de votre ordinateur local ou plage d'adresses IP (en notation de bloc CIDR) de votre réseau.	Autorise l'accès SSH entrant depuis votre ordinateur local.

Règles Elastic Load Balancing

Si vous utilisez un équilibreur de charge, le groupe de sécurité associé à celui-ci doit avoir des règles qui autorisent la communication avec vos instances ou cibles.

Entrant				
Type de protocole	Numéro de protocole	Port	IP Source	Remarques
TCP	6	Port d'écoute	Pour un équilibreur de charge accessible sur Internet : 0.0.0.0/0 (toutes les adresses IPv4) Pour un équilibreur de charge interne : le bloc d'adresse CIDR IPv4 du VPC	Autoriser le trafic entrant sur le port d'écoute de l'équilibreur de charge.
Sortant				
Type de protocole	Numéro de protocole	Port	IP de destination	Remarques
TCP	6	Port d'écoute de l'instance.	ID du groupe de sécurité de l'instance.	Autoriser le trafic sortant vers les instances sur le port d'écoute des instances.

TCP	6	Port de vérification de l'état.	ID du groupe de sécurité de l'instance.	Autoriser le trafic sortant vers les instances sur le port de vérification de l'état.

Les règles du groupe de sécurité de vos instances doivent autoriser l'équilibreur de charge à communiquer avec vos instances sur le port d'écoute et sur le port de vérification de l'état.

Entrant				
Type de protocole	Numéro de protocole	Port	IP Source	Remarques
TCP	6	Port d'écoute de l'instance.	ID du groupe de sécurité de l'équilibreur de charge.	Autoriser le trafic depuis l'équilibreur de charge sur le port d'écoute des instances.
TCP	6	Port de vérification de l'état.	ID du groupe de sécurité de l'équilibreur de charge.	Autoriser le trafic depuis l'équilibreur de charge sur le port de vérification de l'état.

Pour de plus amples informations, veuillez consulter Configurer des groupes de sécurité pour votre Equilibreur de charge classique dans le Guide de l'utilisateur pour les Equilibreurs de charge classiques et Groupes de sécurité pour votre Equilibreur de charge d'application dans le Guide de l'utilisateur pour les Application Load Balancers.

Règles d'appairage de VPC

Vous pouvez mettre à jour les règles entrantes ou sortantes pour les groupes de sécurité de votre VPC pour référencer des groupes de sécurité dans le VPC appairé. Cette étape autorise la circulation du trafic vers et depuis les instances associées au groupe de sécurité référencé dans le VPC appairé. Pour de plus amples informations sur la configuration des groupes de sécurité pour l'appairage de VPC, veuillez consulter Mise à jour de vos groupes de sécurité pour référencer les groupes de VPC pairs.

Gestion des mises à jour dans Amazon EC2

Nous vous recommandons d’appliquer les correctifs, de procéder aux mises à jours et de sécuriser le système d'exploitation et les applications sur vos instances EC2 régulièrement. Vous pouvez utiliser le Gestionnaire de correctifs AWS Systems Manager pour automatiser le processus d'installation des mises à jour de sécurité pour le système d'exploitation et les applications. Vous pouvez aussi utiliser n'importe quel service de mise à jour automatique ou processus recommandé pour l'installation des mises à jour fourni par le fournisseur de l'application.

Validation de la conformité pour Amazon EC2

Les auditeurs tiers évaluent la sécurité et la conformité de Amazon EC2 dans le cadre de plusieurs programmes de conformité AWS. Il s'agit notamment des certifications SOC, PCI, FedRAMP, HIPAA et autres.

Pour obtenir la liste des services AWS relevant de programmes de conformité spécifiques, veuillez consulter Services AWS dans Scope by Compliance Program. Pour obtenir des renseignements généraux, consultez Programmes de conformitéAWS .

Vous pouvez télécharger les rapports de l'audit externe avec AWS Artifact. Pour de plus amples informations, veuillez consulter Téléchargement de rapports dans AWS Artifact.

Votre responsabilité en matière de conformité lors de l'utilisation d'Amazon EC2 est déterminée par la sensibilité de vos données, les objectifs de conformité de l'entreprise, ainsi que la législation et la réglementation applicables. AWS fournit les ressources suivantes pour faciliter le respect de la conformité :

- Guides de démarrage rapide de la sécurité et de la conformité – Ces guides de déploiement proposent des considérations architecturales et fournissent des étapes pour déployer des environnements de référence centrés sur la sécurité et la conformité sur AWS.
- Livre blanc sur l'architecture pour la sécurité et la conformité HIPAA – Le livre blanc décrit comment les entreprises peuvent utiliser AWS pour créer des applications conformes à la loi HIPAA.
- Ressources de conformité AWS – Cet ensemble de manuels et de guides peut s'appliquer à votre secteur et à votre emplacement.
- Évaluation des ressources à l'aide de règles dans le AWS Config Developer Guide – AWS Config évalue dans quelle mesure vos configurations de ressources sont conformes aux pratiques internes, aux directives sectorielles et aux réglementations.
- AWS Security Hub – Ce service AWS fournit une vue complète de votre état de sécurité au sein d'AWS qui vous permet de vérifier votre conformité aux normes du secteur et aux bonnes pratiques de sécurité.

Stockage

Amazon EC2 vous offre des options de stockage de données flexibles, économiques et simples d'utilisation pour vos instances. Chaque option est une combinaison unique de performance et de durabilité. Ces options de stockage peuvent être utilisées seules ou associées pour satisfaire vos besoins.

Après avoir lu cette section, vous devriez avoir une bonne compréhension de la façon d'utiliser les options de stockage de données prises en charge par Amazon EC2 afin de satisfaire vos besoins spécifiques. Parmi ces options de stockage, on trouve :

- Amazon Elastic Block Store (p. 1084)
- Stockage d'instance Amazon EC2 (p. 1264)
- Utilisation de Amazon EFS avec Amazon EC2 (p. 1283)
- Utilisation de Amazon S3 avec Amazon EC2 (p. 1281)

L'illustration suivante représente la relation entre ces options de stockage et votre instance.

Amazon EBS

Amazon EBS fournit des volumes de stockage de niveau bloc que vous pouvez attacher à une instance en cours d'exécution. Vous pouvez utiliser Amazon EBS comme périphérique de stockage principal pour les données nécessitant des mises à jour fréquentes et précises. Par exemple, Amazon EBS est l'option de stockage recommandée lorsque vous exécutez une base de données sur une instance.

Un volume EBS se comporte comme un périphérique de stockage en mode bloc externe brut et non formaté que vous pouvez attacher à une instance unique. Le volume persiste indépendamment de la durée d'exécution d'une instance. Une fois qu'un volume EBS est attaché à une instance, vous pouvez l'utiliser comme n'importe quel autre disque dur physique. Comme le montre l'illustration précédente, plusieurs volumes peuvent être attachés à une instance. Vous pouvez également détacher un volume EBS d'une instance et l'attacher à une autre instance. Vous pouvez modifier dynamiquement la configuration d'un

volume attaché à une instance. Les volumes EBS peuvent aussi être créés comme des volumes chiffrés en utilisant la fonction Chiffrement Amazon EBS. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Pour conserver une copie de sauvegarde de vos données, vous pouvez créer un instantané d'un volume EBS qui est stocké dans Amazon S3. Vous pouvez créer un volume EBS à partir d'un instantané, puis l'attacher à une autre instance. Pour plus d'informations, consultez Amazon Elastic Block Store (p. 1084).

Stockage d'instance Amazon EC2

De nombreuses instances peuvent accéder au stockage sur des disques physiquement attachés à l'ordinateur hôte. Ce disque de stockage est appelé stockage d'instance. Le stockage d'instance fournit un stockage temporaire de niveau bloc pour les instances. Les données sur un volume de stockage d'instance persistent uniquement pendant la vie de l'instance associée. Si vous arrêtez, mettez en veille prolongée ou résiliez une instance, toutes les données sur les volumes de stockage d'instance sont perdues. Pour plus d'informations, consultez Stockage d'instance Amazon EC2 (p. 1264).

Système de fichiers Amazon EFS

Amazon EFS offre un stockage de fichiers scalable, destiné à être utilisé avec Amazon EC2. Vous pouvez créer un système de fichiers EFS et configurer vos instances afin de l'installer. Vous pouvez utiliser un système de fichiers EFS comme source de données commune aux charges de travail et applications exécutées sur plusieurs instances. Pour plus d'informations, consultez Utilisation de Amazon EFS avec Amazon EC2 (p. 1283).

Amazon S3

Amazon S3 fournit un accès à une infrastructure de stockage de données fiable et économique. Cet outil est conçu pour faciliter l'accès aux ressources informatiques à l'échelle du Web en vous permettant de stocker et de récupérer à tout moment n'importe quelle quantité de données, depuis Amazon EC2, ou depuis n'importe quel emplacement sur le Web. Par exemple, vous pouvez utiliser Amazon S3 pour stocker des copies de sauvegarde de vos données et applications. Amazon EC2 utilise Amazon S3 pour stocker des instantanés EBS et des AMI basées sur le stockage d'instance. Pour plus d'informations, consultez Utilisation de Amazon S3 avec Amazon EC2 (p. 1281).

Ajouter du stockage

A chaque fois que vous lancez une instance depuis une AMI, un périphérique de stockage racine est créé pour cette instance. Le périphérique de stockage racine contient toutes les informations nécessaires pour démarrer l'instance. Vous pouvez spécifier les volumes de stockage en plus du volume du périphérique racine lorsque vous créez une AMI ou lancez une instance en utilisant un mappage de périphérique de stockage en mode bloc. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289).

Vous pouvez également attacher des volumes EBS à une instance en cours d'exécution. Pour plus d'informations, consultez Attacher un volume Amazon EBS à une instance (p. 1109).

Tarification du stockage

Pour de plus amples informations sur la tarification du stockage, ouvrez Tarification AWS, faites défiler jusqu'à Tarification des services, choisissez Stockage, puis choisissez l'option de stockage pour ouvrir la page de tarification de cette option de stockage. Pour de plus amples informations sur l'estimation du coût du stockage, veuillez consulter le Calculateur de tarification AWS.

Amazon Elastic Block Store (Amazon EBS)

Amazon Elastic Block Store (Amazon EBS) fournit des volumes de stockage niveau bloc à utiliser avec les instances EC2. Les volumes EBS se comportent comme des périphériques de stockage en mode bloc bruts non formatés. Vous pouvez monter ces volumes en tant qu'appareils sur vos instances. Les

volumes EBS qui sont attachés à une instance sont exposés en tant que volumes de stockage qui sont conservés indépendamment du cycle de vie de l'instance. Vous pouvez créer un système de fichiers au-dessus de ces volumes ou les utiliser comme vous utiliseriez un périphérique de stockage en mode bloc (comme un disque dur). Vous pouvez modifier dynamiquement la configuration d'un volume attaché à une instance.

Amazon EBS est conseillé lorsque les données doivent être rapidement accessibles et qu'elles nécessitent une persistance à long terme. Les volumes EBS sont particulièrement adaptés à une utilisation en tant que stockage principal pour des fichiers système, des bases de données ou pour toutes les applications ayant besoin de mises à jour précises et d'accès à un stockage niveau bloc brut et non formaté. Amazon EBS convient particulièrement aux applications de type base de données qui impliquent de nombreuses lectures et écritures aléatoires, ainsi qu'aux applications à débit élevé qui effectuent des lectures et écritures longues et continues.

Avec Amazon EBS, vous ne payez que ce que vous utilisez. Pour plus d'informations sur la tarification Amazon EBS, veuillez consulter la section consacrée à la planification des coût sur la page Amazon Elastic Block Store.

Sommaire

- Fonctions d'Amazon EBS (p. 1085)
- Volumes Amazon EBS (p. 1086)
- Instantanés Amazon EBS (p. 1129)
- Services de données Amazon EBS (p. 1181)
- Amazon EBS et NVMe sur des instances Linux (p. 1209)
- Amazon EBS–Instances optimisées (p. 1213)
- Performances des volumes Amazon EBS pour les instances Linux (p. 1230)
- Métriques Amazon CloudWatch pour Amazon EBS (p. 1247)
- Amazon CloudWatch Events pour Amazon EBS (p. 1253)
- Quotas Amazon EBS (p. 1264)

Fonctions d'Amazon EBS

- Les volumes EBS sont créés dans une zone de disponibilité spécifique et peuvent être attachés à toute instance de cette même Zone. Pour qu'un volume soit disponible en dehors de la zone de disponibilité, vous pouvez créer un instantané et restaurer celui-ci sur un nouveau volume n'importe où dans cette région. Vous pouvez également copier les instantanés dans d'autres régions, puis les restaurer dans de nouveaux volumes de ces régions, ce qui vous permet d'utiliser plus facilement de nombreuses régions AWS à des fins d'expansion géographique, de migration des centres de données et de reprise après sinistre.
- Amazon EBS fournit les types de volume suivants : SSD à usage général (`gp2`), IOPS provisionnés SSD (`io1` and `io2`), HDD à débit optimisé (`st1`) et Cold HDD (`sc1`). Voici un résumé des performances et des cas d'utilisation pour chaque type de volume.
  - Avec les volumes SSD à usage général, vous pouvez bénéficier de performances de base de 3 E/S par seconde par Gio, avec la capacité de passer à 3 000 E/S par seconde pour des durées étendues. Les volumes &GP2; conviennent parfaitement à un grand nombre de cas d'utilisation, notamment pour les volumes de démarrage, les bases de données de moyenne et de petite tailles, ainsi que les environnements de développement et de test. Pour de plus amples informations, veuillez consulter Volumes SSD à usage général (`gp2`) (p. 1091).
  - Les volumes IOPS provisionnés SSD prennent en charge jusqu'à 64,000 IOPS et 1,000 Mio/s de débit. Vous pouvez ainsi évoluer naturellement vers des dizaines de milliers d'IOPS par instance EC2. Pour plus d'informations, consultez Volumes IOPS provisionnés SSD (`io1` and `io2`) (p. 1095).
  - Les volumes HDD à débit optimisé offrent un stockage magnétique économique qui définit les performances en termes de débit plutôt que d'IOPS. Ce type de volume convient aux charges de

travail séquentielles et volumineuses comme Amazon EMR, ETL, les entrepôts de données et le traitement des journaux. Pour de plus amples informations, veuillez consulter Volumes HDD à débit optimisé (`st1`) (p. 1096).

- Les volumes Cold HDD offrent un stockage magnétique économique qui définit les performances en termes de débit plutôt que d'IOPS. Ces volumes conviennent à des charges de travail volumineuses et séquentielles dont les données sont légères. Si vous n'avez pas besoin d'accéder souvent à vos données et si vous cherchez à réaliser des économies, ces volumes fournissent un stockage de bloc économique. Pour de plus amples informations, veuillez consulter Volumes Cold HDD (`sc1`) (p. 1099).

- Vous pouvez créer des volumes EBS chiffrés pour satisfaire à un grand nombre d'exigences en matière de chiffrement de données au repos pour les données et applications réglementées et auditées. Lorsque vous créez un volume EBS chiffré et que vous l'attachez à un type d'instance pris en charge, les données stockées au repos sur le volume, les I/O de disque et les instantanés créés à partir du volume sont tous chiffrés. Le chiffrement est effectué sur les serveurs hébergeant les instances EC2, assurant ainsi le chiffrement des données qui se déplacent entre les instances EC2 et le stockage EBS. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).
- Vous pouvez créer des instantanés ponctuels des volumes EBS qui sont conservés dans Amazon S3. Les instantanés protègent les données à long terme et peuvent être utilisés comme point de départ des nouveaux volumes EBS. Le même instantané peut être utilisé pour instancier autant de volumes que vous le souhaitez. Les instantanés peuvent être copiés entre différentes régions AWS. Pour plus d'informations, consultez Instantanés Amazon EBS (p. 1129).
- Les métriques de performances, par exemple la bande passante, le débit, la latence et la longueur moyenne de file d'attente, sont disponibles sur l'AWS Management Console. Ces métriques, fournies par Amazon CloudWatch, vous permettent de surveiller les performances de vos volumes afin de veiller à fournir des performances suffisamment élevées pour vos applications sans avoir à payer les ressources dont vous n'avez pas besoin. Pour plus d'informations, consultez Performances des volumes Amazon EBS pour les instances Linux (p. 1230).

Volumes Amazon EBS

Un volume Amazon EBS est un dispositif de stockage durable au niveau bloc que vous pouvez attacher à vos instances. Après avoir attaché à une instance, vous pouvez l'utiliser comme n'importe quel autre disque dur physique. Les volumes EBS sont flexibles. Pour des volumes de génération actuelle attachés à des types d'instances de génération actuelle, vous pouvez augmenter dynamiquement la taille, modifier la capacité IOPS provisionnée et changer le type des volumes de production en direct.

Vous pouvez utiliser des volumes EBS comme stockage principal pour des données nécessitant des mises à jour fréquentes, telles que le lecteur système pour une instance ou le stockage pour une application de base de données. Vous pouvez également les utiliser pour les applications à débit élevé qui effectuent des analyses continues du disque. Les volumes EBS sont permanents indépendamment de la durée d'exécution d'une instance EC2.

Vous pouvez également attacher plusieurs volumes EBS à une seule instance. Le volume et l'instance doivent être dans la même zone de disponibilité. En fonction du volume et des types d'instances, vous pouvez utiliser Multi-Attach (p. 1110) pour monter un volume sur plusieurs instances en même temps.

Amazon EBS fournit les types de volumes suivants : SSD à usage général (`gp2`), IOPS provisionnés SSD (`io1` et `io2`) HDD à débit optimisé (`st1`) Cold HDD (`sc1`) et magnétique (`standard`, un type de génération précédente). Ils se distinguent par leurs caractéristiques de performance et leurs tarifs, ce qui vous permet d'adapter vos performances de stockage et vos coûts en fonction des besoins de vos applications. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).

Votre compte est limité en ce qui concerne le nombre de volumes EBS que vous pouvez utiliser et le stockage total dont vous disposez. Pour plus d'informations sur ces limites et pour savoir comment demander leur augmentation, veuillez consulter Quotas de service Amazon EC2 (p. 1322).

Pour plus d'informations sur la tarification, consultez Tarification Amazon EBS.

Sommaire

- Avantages offerts par l'utilisation de volumes EBS (p. 1087)
- Types de volume Amazon EBS (p. 1089)
- Contraintes sur la taille et la configuration d'un volume EBS (p. 1104)
- Création d'un volume Amazon EBS (p. 1107)
- Attacher un volume Amazon EBS à une instance (p. 1109)
- Attacher un volume à plusieurs instances à l'aide d'Amazon EBS Multi-Attach (p. 1110)
- Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113)
- Affichage d'informations sur un volume Amazon EBS (p. 1117)
- Remplacement d'un volume Amazon EBS à l'aide d'un instantané précédent (p. 1118)
- Surveillance du statut de vos volumes (p. 1119)
- Détacher un volume Amazon EBS d'une instance Linux (p. 1127)
- Suppression d'un volume Amazon EBS (p. 1129)

Avantages offerts par l'utilisation de volumes EBS

Les volumes EBS offrent des avantages supplémentaires par rapport aux volumes de stockage d'instance.

Disponibilité des données

Lorsque vous créez un volume EBS, il est automatiquement répliqué au sein la zone de disponibilité pour empêcher toute perte de données consécutive à la défaillance d'un composant matériel. Vous pouvez attacher un volume EBS à n'importe quelle instance EC2 dans la même zone de disponibilité. Une fois qu'un volume est attaché, il se présente comme un périphérique de stockage en mode bloc natif similaire à un disque dur ou à un autre périphérique physique. À ce stade, l'instance peut interagir avec le volume de la même façon qu'avec un périphérique local. Vous pouvez vous connecter à l'instance et formater le volume EBS avec un système de fichiers, tel que ext3, puis installer des applications.

Si vous attachez plusieurs volumes à un périphérique que vous avez nommé, vous pouvez agréger les données par bandes entre ces volumes pour de meilleures performances E/S et en matière de débit.

Vous pouvez attacher un volume EBS `io1` à un maximum de 16 instances basées sur Nitro. Pour de plus amples informations, veuillez consulter Attacher un volume à plusieurs instances à l'aide d'Amazon EBS Multi-Attach (p. 1110). Sinon, vous pouvez attacher un volume EBS à une seule instance.

Vous pouvez obtenir des données de surveillance pour vos volumes EBS, y compris les données pour les volumes du périphérique racine des instances basées sur EBS, sans coût supplémentaire. Pour plus d'informations sur la surveillance des métriques, consultez Métriques Amazon CloudWatch pour Amazon EBS (p. 1247). Pour de plus amples informations sur le suivi de l'état de vos volumes, veuillez consulter Amazon CloudWatch Events pour Amazon EBS (p. 1253).

Persistance des données

Un volume EBS est un stockage hors instance qui peut persister indépendamment de la vie d'une instance. Vous continuez à payer pour l'utilisation du volume tant que les données persistent.

Les volumes EBS attachés à une instance en cours d'exécution peuvent se détacher automatiquement de l'instance avec leurs données intactes lorsque l'instance est résiliée, si vous cochez la case Supprimer lors de la résiliation lorsque vous configurez les volumes EBS pour votre instance sur la console EC2. Le volume peut être attaché à une nouvelle instance, ce qui permet une récupération rapide. Si la case

Supprimer lors de la résiliation est cochée, le ou les volumes seront supprimés lors de la résiliation de l'instance EC2. Si vous utilisez une instance basée sur EBS, vous pouvez arrêter et redémarrer l'instance sans affecter les données stockées dans le volume attaché. Le volume reste attaché pendant le cycle d'arrêt-démarrage. Cela vous permet de traiter et de stocker indéfiniment les données sur votre volume, en utilisant les ressources de traitement et de stockage uniquement lorsque cela est nécessaire. Les données persistent sur le volume jusqu'à ce que ce volume soit explicitement supprimé. Le stockage par bloc physique utilisé par les volumes EBS supprimés est remplacé par des zéros avant d'être alloué à un autre compte. Si vous travaillez avec des données sensibles, nous vous recommandons de chiffrer vos données manuellement ou de les stocker sur un volume protégé par Chiffrement Amazon EBS. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Par défaut, le volume EBS racine qui est créé et attaché à une instance au moment du lancement est supprimé lorsque cette instance prend fin. Vous pouvez modifier ce comportement en changeant la valeur de l'indicateur `DeleteOnTermination` en `false` lorsque vous lancez l'instance. En modifiant cette valeur, le volume persiste même après que l'instance ait pris fin, ce qui vous permet de l'attacher à une autre instance.

Par défaut, les volumes EBS supplémentaires qui sont créés et attachés à une instance au moment du lancement ne sont pas supprimés lorsque cette instance prend fin. Vous pouvez modifier ce comportement en changeant la valeur de l'indicateur `DeleteOnTermination` en `true` lorsque vous lancez l'instance. Cette valeur modifiée entraîne la suppression des volumes lorsque l'instance prend fin.

Chiffrement des données

Pour simplifier le chiffrement des données, vous pouvez créer des volumes EBS chiffrés avec la fonction Chiffrement Amazon EBS. Tous les types de volume EBS prennent en charge le chiffrement. Vous pouvez utiliser des volumes EBS chiffrés pour satisfaire un grand nombre d'exigences en matière de chiffrement de données au repos pour les données et applications réglementées et auditées. Chiffrement Amazon EBS utilise les algorithmes Advanced Encryption Standard 256-bit (AES-256) et une infrastructure à clés gérée par Amazon. Le chiffrement est effectué sur le serveur qui héberge l'instance EC2, assurant ainsi le chiffrement des données qui se déplacent entre l'instance EC2 et le stockage Amazon EBS. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Chiffrement Amazon EBS utilise les clés principales AWS Key Management Service (AWS KMS) lors de la création de volumes chiffrés et d'instantanés créés à partir de vos volumes chiffrés. La première fois que vous créez un volume EBS chiffré dans une région, une clé principale par défaut est créée automatiquement. Cette clé est utilisée pour Chiffrement Amazon EBS, à moins que vous ne sélectionniez une clé principale client (CMK, Customer Master Key) que vous avez créée séparément avec AWS KMS. La création de votre propre clé principale client vous donne davantage de flexibilité, en vous permettant notamment de créer, effectuer la rotation, désactiver et définir les contrôles d'accès, ainsi que de contrôler les clés de chiffrement utilisées pour protéger vos données. Pour plus d'informations, consultez le AWS Key Management Service Developer Guide.

Instantanés

Amazon EBS offre la possibilité de créer des instantanés (sauvegardes) d'un volume EBS et de copier les données dans le volume sur Amazon S3, où elles sont stockées de façon redondante dans plusieurs zones de disponibilité. Le volume n'a pas besoin d'être attaché à une instance en cours d'exécution pour pouvoir créer un instantané. Alors que vous continuez à écrire des données sur un volume, vous pouvez créer régulièrement un instantané de ce dernier afin de l'utiliser comme base pour de nouveaux volumes. Ces instantanés peuvent être utilisés pour créer plusieurs volumes EBS ou déplacer des volumes entre les zones de disponibilité. Les instantanés de volumes EBS chiffrés sont chiffrés automatiquement.

Lorsque vous créez un volume à partir d'un instantané, celui-ci est une copie exacte du volume initial au moment où l'instantané a été créé. Les volumes qui sont créés à partir d'instantanés chiffrés sont automatiquement chiffrés. En spécifiant éventuellement une zone de disponibilité différente, vous pouvez utiliser cette fonctionnalité pour dupliquer un volume dans cette zone. Les instantanés peuvent être

partagés avec des comptes AWS spécifiques ou rendus publics. Lorsque vous créez des instantanés, vous êtes facturé dans Amazon S3 en fonction de la taille totale du volume. En cas de nouvel instantané du volume, vous n'êtes facturé que pour les données additionnelles excédant la taille originale du volume.

Les instantanés sont des sauvegardes incrémentielles, ce qui signifie que seuls les blocs du volume qui ont changé depuis l'instantané le plus récent sont enregistrés. Si vous avez un volume de 100 Gio de données mais que seulement 5 Gio ont changé depuis votre dernier instantané, seuls ces 5 Gio de données modifiées sont écrits sur Amazon S3. Bien que les instantanés soient enregistrés de manière incrémentielle, le processus de suppression de l'instantané prévoit que vous avez uniquement besoin de conserver l'instantané le plus récent.

Pour vous aider à classer et à gérer vos volumes et instantanés, vous pouvez les étiqueter avec les métadonnées de votre choix. Pour plus d'informations, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Pour sauvegarder automatiquement vos volumes, vous pouvez utiliser Amazon Data Lifecycle Manager (p. 1169) ou AWS Backup.

Flexibilité

Les volumes EBS acceptent les modifications de configuration en direct en cours de production. Vous pouvez modifier le type de volume, la taille du volume et la capacité IOPS sans interruption de service. Pour plus d'informations, consultez Amazon EBS Elastic Volumes (p. 1182).

Types de volume Amazon EBS

Amazon EBS fournit les types de volume suivants, qui ont des caractéristiques de performances et des prix différents, ce qui vous permet d'adapter vos performances de stockage et vos coûts en fonction des besoins de vos applications. Les types de volume entrent dans ces catégories :

- Disques SSD (Solid State Drives) (p. 1089) : optimisés pour les charges de travail transactionnelles impliquant des opérations fréquentes en lecture/écriture avec des E/S de petite taille, où les entrées et sorties par seconde sont l'attribut de performances principal.
- Disques durs (HDD) (p. 1090) : optimisés pour les charges de travail importantes en streaming où l'attribut de performance dominant est le débit.
- Génération précédente (p. 1091) : disques durs pouvant être utilisés pour des charges de travail comportant des ensembles de données réduits où l'accès aux données est rare et où les performances n'ont pas une importance primordiale. Nous vous recommandons de privilégier plutôt un type de volume de génération actuelle.

Plusieurs facteurs peuvent affecter les performances des volumes EBS, tels que la configuration d'instance, les caractéristiques E/S et la demande en matière de charge de travail. Pour plus d'informations sur la façon d'exploiter au mieux vos volumes EBS, consultez Performances des volumes Amazon EBS pour les instances Linux (p. 1230).

Pour plus d'informations sur la tarification, consultez Tarification Amazon EBS.

Disques SSD

Les volumes SSD fournis par Amazon EBS entrent dans les catégories suivantes :

- SSD à usage général : fournit un équilibre entre le prix et la performance. Nous recommandons ces volumes pour la plupart des charges de travail.
- IOPS provisionnés SSD : fournit des performances élevées pour les charges de travail critiques à faible latence ou à débit élevé.

	SSD à usage général	IOPS provisionnés SSD	
Type de volume	`gp2`	`io2`	`io1`
Durabilité	Durabilité de 99,8 % - 99,9 % (taux de défaillance annuel de 0,1 % - 0,2 %)	Durabilité de 99,999 % (taux de défaillance annuel de 0,001 %)	Durabilité de 99,8 % - 99,9 % (taux de défaillance annuel de 0,1 % - 0,2 %)
Cas d'utilisation	• Volumes de démarrage • Applications interactives à faible latence • Environnements de développement et de test	• Charges de travail nécessitant des performances IOPS soutenues ou supérieures à 16,000 IOPS ou 250 Mio/s de débit par volume • Charges de travail de base de données à fort taux d'E/S.	
Taille du volume	1 Gio - 16 Tio	4 Gio - 16 Tio	
Nombre maximal d'E/S par seconde par volume (16 KiO d'E/S)	16,000 *	64,000 †	
Débit maximal par volume	250 Mio/s *	1,000 Mio/s †	
Multi-Attach Amazon EBS	Non pris en charge	Non pris en charge	Pris en charge

Disques durs (HDD)

Les volumes de disque dur fournis par Amazon EBS entrent dans les catégories suivantes :

- HDD à débit optimisé : Disque dur économique conçu pour les charges de travail à débit élevé fréquemment consultées.
- Cold HDD : disque dur le plus abordable pour les charges de travail moins fréquemment consultées.

	HDD à débit optimisé	Cold HDD
Type de volume	`st1`	`sc1`
Durabilité	Durabilité de 99,8 % - 99,9 % (taux de défaillance annuel de 0,1 % - 0,2 %)	Durabilité de 99,8 % - 99,9 % (taux de défaillance annuel de 0,1 % - 0,2 %)
Cas d'utilisation	• Big Data • Entrepôts de données • Traitement de fichiers journaux	• Stockage axé sur le débit pour les données consultées de manière occasionnelle • Scénarios dans lesquels il est important que le coût de stockage soit le plus bas possible

	HDD à débit optimisé	Cold HDD
Taille du volume	500 Gio - 16 Tio	500 Gio - 16 Tio
Nombre maximal d'E/S par seconde par volume (1 MiO d'E/S)	500	250
Débit maximal par volume	500 Mio/s	250 Mio/s
Multi-Attach Amazon EBS	Non pris en charge	Non pris en charge

Types de volume de la génération précédente

Le tableau suivant décrit les types de volume EBS de la génération précédente. Si vous avez besoin de performances supérieures ou plus homogènes que les volumes de la génération précédente, nous vous recommandons d'utiliser SSD à usage général (`gp2`) ou d'autres types de volume actuels. Pour plus d'informations, consultez Volumes de la génération précédente.

	magnétique
Type de volume	`standard`
Cas d'utilisation	Charges de travail où l'accès aux données est occasionnel
Taille du volume	1 Gio - 1 Tio
IOPS maximum par volume	40–200
Débit maximal par volume	40–90 Mio/s
IOPS maximum par instance	80 000
Débit maximal par instance	1,750 MB/s

Volumes SSD à usage général (`gp2`)

Les volumes SSD à usage général (`gp2`) offrent un stockage économique idéal pour un large éventail de charges de travail. Ces volumes fournissent des latences inférieures à 10 millisecondes, avec la capacité d'émettre en rafale jusqu'à 3 000 E/S par seconde pendant de longues périodes de temps. Entre un minimum de 100 IOPS par seconde (à 33,33 Gio ou moins) et un maximum de 16,000 IOPS par seconde (à 5 334 Gio ou plus), les performances de base évoluent de manière linéaire à 3 IOPS par seconde par Gio de la taille du volume. AWS conçoit des volumes `gp2` afin de fournir 99 % des performances prévues 99 % du temps. La taille des volumes `gp2` peut aller de 1 Gio à 16 Tio.

Crédits E/S et performances en rafale

Les performances des volumes `gp2` sont régies par la taille des volumes. Cette taille dicte le niveau de performance de base du volume et la vitesse à laquelle il accumule des crédits E/S. Les volumes les plus gros ont un niveau de performance de base plus élevé et ils accumulent des crédits E/S plus vite. Les crédits E/S représentent la bande passante disponible que le volume `gp2` peut utiliser pour émettre en rafale de grandes quantités d'E/S lorsqu'il est nécessaire de dépasser les performances de base. Plus

votre volume dispose de crédits pour les E/S, plus il peut émettre en rafale au-delà de son niveau de performances de base, et plus il est performant quand cela est nécessaire. Le schéma suivant illustre le comportement du compartiment en rafales pour `gp2`.

Chaque volume reçoit un solde de crédit d'E/S initial de 5,4 millions de crédits d'E/S, ce qui est suffisant pour maintenir la performance de rafale maximale de 3 000 E/S par seconde pendant au moins 30 minutes. Le solde de crédits initial est conçu pour fournir un cycle de démarrage initial rapide pour les volumes de démarrage ainsi qu'une bonne expérience d'action d'amorçage (bootstrap) pour les autres applications. Les volumes gagnent des crédits E/S, au rythme de performances de base de 3 IOPS par Gio de la taille du volume. Par exemple, un volume `gp2` de 100 Gio a une performance de base de 300 E/S par seconde.

Lorsque votre volume a besoin d'un niveau de performances E/S plus élevé que le niveau de base, il utilise simplement les crédits E/S du solde de crédits pour émettre en rafales, jusqu'à un niveau de performances maximum de 3 000 E/S par seconde. Lorsque votre volume utilise moins de crédits E/S qu'il n'en gagne en une seconde, les crédits E/S non utilisés sont ajoutés au solde de crédits E/S. Le solde de crédits E/S maximum pour un volume est égal au solde de crédits initial (5 400 000 crédits E/S).

Lorsque les performances de base d'un volume sont supérieures aux performances en rafale maximales, les crédits d'E/S ne sont jamais dépensés. Si le volume est attaché à une instance construite sur le Système Nitro (p. 218), l'équilibre en rafale n'est pas signalé. Dans d'autres cas, l'équilibre en rafale déclaré est de 100 %.

La durée de rafale d'un volume dépend de sa taille, des E/S par seconde de rafale nécessaires et du solde de crédits au début de la rafale. Cela est représenté dans l'équation ci-après :

```
                              (Credit balance)
Burst duration  =  ------------------------------------
                   (Burst IOPS) - 3(Volume size in GiB)
```

Le tableau ci-après énumère les tailles de volume et les performances de base associées du volume (qui correspondent également au rythme auquel il accumule les crédits E/S), la durée d'émission en rafales

à un maximum de 3 000 E/S par seconde (en commençant avec un solde créditeur plein) et la durée en secondes qu'il faut au volume pour remplir un solde de crédits vide.

Taille du volume (Gio)	Performances de base (E/S par seconde)	Durée de transmission en rafales avec 3 000 IOPS (seconde) soutenues	Secondes nécessaires pour remplir un solde de crédits vide lorsqu'il n'y a pas d'IO
1	100	1 802	54 000
100	300	2 000	18 000
250	750	2 400	7 200
334 (taille minimale pour débit maximal)	1 002	2 703	5 389
500	1 500	3 600	3 600
750	2 250	7 200	2 400
1 000	3 000	Non applicable*	Non applicable*
5 334 (taille minimale pour IOPS maximum)	16,000	Non applicable*	Non applicable*
16 384 (16 Tio, taille maximale du volume)	16,000	Non applicable*	Non applicable*

* Les performances de base du volume dépassent les performances en rafale maximales.

Que se passe-t-il si j'utilise tout mon solde de crédits E/S ?

Si votre volume `gp2` utilise tout son solde de crédits E/S, les performances d'E/S par seconde maximum du volume restent au niveau de performance d'E/S par seconde de base (le taux auquel votre volume gagne des crédits), et le débit maximal du volume est réduit au niveau d'E/S par seconde de base multiplié par la taille d'E/S maximale. Le débit ne peut jamais dépasser 250 Mio/s. Lorsque la demande en I/O tombe en dessous du niveau de référence et que les crédits non utilisés sont ajoutés au solde de crédits I/O, les performances d'IOPS maximum du volume dépassent à nouveau le niveau de référence. Par exemple, un volume `gp2` de 100 Gio avec un solde de crédits nul a une performance de base de 300 E/S par seconde et un débit limite de 75 Mio/s (300 opérations d'E/S par seconde * 256 KiO par opération d'E/S = 75 Mio/s). Plus le volume est gros, plus la performance de base est grande et plus la vitesse à laquelle le volume réapprovisionne son solde de crédits est grande. Pour de plus amples informations sur la façon dont les E/S par seconde sont mesurées, veuillez consulter la section Caractéristiques d'E/S et surveillance (p. 1233).

Si vous constatez que les performances du volume sont fréquemment limitées au niveau de base (en raison d'un solde de crédits E/S nul), envisagez d'utiliser un volume `gp2` plus gros (avec un niveau de performances de base plus élevé) ou d'utiliser un volume `io1` ou `io2` pour des charges de travail nécessitant des performances IOPS soutenues supérieures à 16,000 IOPS.

Pour plus d'informations sur l'utilisation des métriques et alarmes CloudWatch pour surveiller l'équilibre du compartiment en rafales, consultez Surveillance de l'équilibre du compartiment en rafales pour les volumes `gp2`, `st1` et `sc1` (p. 1104).

Performances de débit

Le débit d'un volume `gp2` peut être calculé à l'aide de la formule suivante, jusqu'à la limite de débit de 250 Mio/s :

```
Throughput in MiB/s = ((Volume size in GiB) × (IOPS per GiB) × (I/O size in KiB))
```

En supposant que V = taille du volume, I = taille des I/O, R = débit d'I/O et T = débit, la formule peut être simplifiée en :

```
T = VIR
```

La plus petite taille de volume qui atteint le débit maximal est donnée par :

```
        T
V  =  -----
       I R

            250 MiB/s
   =  ---------------------
      (256 KiB)(3 IOPS/GiB)


               [(250)(2^20)(Bytes)]/s
   =  -----------------------------------------
      (256)(2^10)(Bytes)([3 IOP/s]/[(2^30)(Bytes)])


      (250)(2^20)(2^30)(Bytes)
   =  ------------------------
           (256)(2^10)(3)

   =  357,913,941,333 Bytes

   =  333# GiB (334 GiB in practice because volumes are provisioned in whole gibibytes)
```

Volumes IOPS provisionnés SSD (io1 and io2)

Les volumes IOPS provisionnés SSD (io1 and io2) sont conçus pour satisfaire les besoins des charges de travail gourmandes en I/O, notamment les charges de travail de base de données qui sont sensibles aux performances de stockage et à l'homogénéité du débit d'I/O. Contrairement à gp2 qui utilise un modèle de compartiment et de crédit pour calculer les performances, les volumes io1 and io2 vous permettent de spécifier un taux d'I/O par seconde cohérent lorsque vous créez les volumes, et Amazon EBS fournit les performances provisionnées 99,9 % du temps.

Les volumes io1 sont conçus pour offrir une durabilité de 99,8 à 99,9 % avec un taux de défaillance annuel (AFR) ne dépassant pas 0,2 %, ce qui se traduit par un maximum de deux défaillances de volume pour 1 000 volumes de fonctionnement sur une période d'un an. Les volumes io2 sont conçus pour offrir une durabilité de 99,999 % avec un AFR ne dépassant pas 0,001 %, ce qui se traduit par une défaillance de volume unique pour 100 000 volumes en cours sur une période d'un an.

La taille des volumes io1 and io2 peut aller de 4 Gio à 16 Tio. Vous pouvez provisionner entre 100 IOPS et 64,000 IOPS par volume sur les instances du Instances reposant sur le système Nitro (p. 218) et jusqu'à 32,000 sur les autres instances. Le rapport maximal entre les volumes IOPS provisionnés et le volume demandé (en Gio) est 50:1 pour les volumes io1 et 500:1 pour les volumes io2. Par exemple, un volume io1 de 100 Gio peut être provisionné avec jusqu'à 5 000 IOPS, tandis qu'un volume io2 de 100 Gio peut être provisionné avec jusqu'à 50 000 IOPS. Sur un type d'instance pris en charge, les tailles de volume suivantes permettent le provisionnement jusqu'au maximum de 64,000 IOPS:

- Volume io1 de 1 280 Gio ou plus (50 × 1 280 Gio = 64 000 IOPS)
- Volume io2 de 128 Gio ou plus (500 × 128 Gio = 64 000 IOPS)

Les volumes `io1` and `io2` mis en service avec 32 000 IOPS maximum prennent en charge une taille maximale d'E/S de 256 Kio et génèrent jusqu'à 500 Mio/s de débit. Avec la taille d'E/S au maximum, le débit de pointe est atteint à 2 000 IOPS. Un volume mis en service avec plus de 32,000 IOPS (jusqu'à 64 000 IOPS maximum) prend en charge une taille maximale d'E/S de 16 KiO et génère jusqu'à 1 000 MiO/s de débit. Le graphique suivant illustre ces performances :

La latence subie par E/S dépend des IOPS mis en service et de votre profil de charge de travail. Pour bénéficier de la meilleure expérience de latence d'E/S, assurez-vous que vous provisionnez des E/S par seconde afin de respecter le profil d'E/S de votre charge de travail.

> **Note**
>
> Certains comptes AWS créés avant 2012 peuvent avoir accès à des zones de disponibilité dans les régions us-west-1 ou ap-northeast-1 où les volumes IOPS provisionnés SSD (`io1`) ne sont pas pris en charge. Si vous ne pouvez pas créer un volume `io1` (ou lancer une instance avec un volume `io1` dans son mappage de périphérique de stockage en mode bloc) dans une de ces régions, essayez une autre zone de disponibilité dans la région. Vous pouvez vérifier que la zone de disponibilité prend en charge les volumes `io1` en créant un volume `io1` de 4 Gio dans cette zone.

Volumes HDD à débit optimisé (`st1`)

Les volumes HDD à débit optimisé (`st1`) offrent un stockage magnétique économique qui définit les performances en termes de débit plutôt que d'IOPS. Ce type de volume convient aux charges de travail séquentielles et volumineuses comme Amazon EMR, ETL, les entrepôts de données et le traitement des journaux. Les volumes `st1` démarrables ne sont pas pris en charge.

Les volumes HDD à débit optimisé (`st1`) sont similaires aux volumes Cold HDD (`sc1`), mais ils sont conçus pour prendre en charge les données fréquemment consultées.

Ce type de volume est optimisé pour les charges de travail impliquant des E/S séquentielles volumineuses, et nous recommandons aux clients dont les charges de travail exécutent des E/S aléatoires de petite taille d'utiliser `gp2`. Pour plus d'informations, consultez Manque d'efficacité des lectures/écritures de petite taille sur disque dur (p. 1104).

Crédits de débit et performances en rafale

À l'instar de `gp2`, `st1` utilise un modèle de transmission de compartiment en rafales pour assurer les performances. La taille du volume détermine le débit de base du volume, qui correspond à la vitesse à laquelle le volume accumule des crédits de débit. La taille du volume détermine également le débit de transmission en rafales du volume, qui correspond à la vitesse à laquelle vous pouvez utiliser des crédits lorsqu'ils sont disponibles. Les gros volumes ont un débit de base et de transmission en rafales plus élevé. Plus votre volume a de crédits, plus longtemps il est en mesure d'assurer la transmission des E/S en rafales.

Le schéma suivant illustre le comportement du compartiment en rafales pour `st1`.

Sous réserve de la limite de débit et de crédits, le débit disponible d'un volume `st1` est exprimé par la formule suivante :

```
(Volume size) x (Credit accumulation rate per TiB) = Throughput
```

Pour un volume `st1` de 1 Tio, le débit de transmission en rafales est limité à 250 Mio/s, le compartiment se remplit de crédits à 40 Mio/s et il peut contenir jusqu'à 1 Tio de crédits.

Les volumes de plus grande taille mettent à l'échelle ces limites de manière linéaire, avec un débit maximum de 500 Mio/s. Une fois que le compartiment est épuisé, le débit est limité à la valeurs de référence (40 Mio/s par Tio).

Avec les volumes dont la taille est comprise entre 0,5 et 16 Tio, le débit de base va de 20 Mio/s à 500 Mio/s (limite maximale), ce qui est atteint à 12,5 Tio comme suit :

```
             40 MiB/s
12.5 TiB x ---------- = 500 MiB/s
              1 TiB
```

Le débit de transmission en rafales va de 125 Mio/s à 500 Mio/s (limite maximale), ce qui est atteint à 2 Tio comme suit :

```
          250 MiB/s
```

```
2 TiB x ---------- = 500 MiB/s
          1 TiB
```

Le tableau suivant indique l'ensemble des valeurs de base en matière de débit et de transmission en rafales pour `st1` :

Taille du volume (Tio)	Débit de base ST1 (Mio/s)	Débit de transmission en rafales ST1 (Mio/s)
0,5	20	125
1	40	250
2	80	500
3	120	500
4	160	500
5	200	500
6	240	500
7	280	500
8	320	500
9	360	500
10	400	500
11	440	500
12	480	500
12,5	500	500
13	500	500
14	500	500
15	500	500
16	500	500

Le schéma suivant illustre le tableau de valeurs sous forme de tracé :

Note

Lorsque vous créez un instantané d'un volume HDD à débit optimisé (`st1`), les performances peuvent diminuer jusqu'à la valeur de référence du volume alors que l'instantané est en cours de création.

Pour plus d'informations sur l'utilisation des métriques et alarmes CloudWatch pour surveiller l'équilibre du compartiment en rafales, consultez Surveillance de l'équilibre du compartiment en rafales pour les volumes `gp2`, `st1` et `sc1` (p. 1104).

Volumes Cold HDD (`sc1`)

Les volumes Cold HDD (`sc1`) offrent un stockage magnétique économique qui définit les performances en termes de débit plutôt que d'IOPS. Avec une limite de débit inférieure à celle des volumes `st1`, `sc1` convient aux charges de travail séquentielles et volumineuses dont les données sont légères. Si vous n'avez pas besoin d'accéder souvent à vos données et si vous cherchez à réaliser des économies, `sc1` fournit un stockage de bloc économique. Les volumes `sc1` démarrables ne sont pas pris en charge.

Les volumes Cold HDD (`sc1`) sont similaires aux volumes HDD à débit optimisé (`st1`), mais ils sont conçus pour prendre en charge les données consultées de manière occasionnelle.

Note

Ce type de volume est optimisé pour les charges de travail impliquant des E/S séquentielles volumineuses, et nous recommandons aux clients dont les charges de travail exécutent des E/S aléatoires de petite taille d'utiliser `gp2`. Pour plus d'informations, consultez Manque d'efficacité des lectures/écritures de petite taille sur disque dur (p. 1104).

Crédits de débit et performances en rafale

À l'instar de `gp2`, `sc1` utilise un modèle de transmission de compartiment en rafales pour assurer les performances. La taille du volume détermine le débit de base du volume, qui correspond à la vitesse à laquelle le volume accumule des crédits de débit. La taille du volume détermine également le débit de transmission en rafales du volume, qui correspond à la vitesse à laquelle vous pouvez utiliser des crédits lorsqu'ils sont disponibles. Les gros volumes ont un débit de base et de transmission en rafales plus élevé. Plus votre volume a de crédits, plus longtemps il est en mesure d'assurer la transmission des E/S en rafales.

Sous réserve de la limite de débit et de crédits, le débit disponible d'un volume `sc1` est exprimé par la formule suivante :

```
(Volume size) x (Credit accumulation rate per TiB) = Throughput
```

Pour un volume `sc1` de 1 Tio, le débit de transmission en rafales est limité à 80 Mio/s, le compartiment se remplit de crédits à 12 Mio/s et il peut contenir jusqu'à 1 Tio de crédits.

Les volumes de plus grande taille mettent à l'échelle ces limites de manière linéaire, avec un débit maximum de 250 Mio/s. Une fois que le compartiment est épuisé, le débit est limité à la valeurs de référence (12 Mio/s par Tio).

Avec les volumes dont la taille est comprise entre 0,5 et 16 Tio, le débit de base va de 6 Mio/s à 192 Mio/s (limite maximale), ce qui est atteint à 16 Tio comme suit :

```
           12 MiB/s
16 TiB x ---------- = 192 MiB/s
            1 TiB
```

Le débit de transmission en rafales va de 40 Mio/s à 250 Mio/s (limite maximale), ce qui est atteint à 3,125 Tio comme suit :

```
             80 MiB/s
3.125 TiB x ----------- = 250 MiB/s
              1 TiB
```

Le tableau suivant indique l'ensemble des valeurs de base en matière de débit et de transmission en rafales pour `sc1` :

Taille du volume (Tio)	Débit de base SC1 (Mio/s)	Débit de transmission en rafales SC1 (Mio/s)
0,5	6	40
1	12	80
2	24	160
3	36	240

Taille du volume (Tio)	Débit de base SC1 (Mio/s)	Débit de transmission en rafales SC1 (Mio/s)
3,125	37,5	250
4	48	250
5	60	250
6	72	250
7	84	250
8	96	250
9	108	250
10	120	250
11	132	250
12	144	250
13	156	250
14	168	250
15	180	250
16	192	250

Le schéma suivant illustre le tableau de valeurs sous forme de tracé :

Note

Lorsque vous créez un instantané d'un volume Cold HDD (`sc1`), les performances peuvent diminuer jusqu'à la valeur de référence du volume alors que l'instantané est en cours de création.

Pour plus d'informations sur l'utilisation des métriques et alarmes CloudWatch pour surveiller l'équilibre du compartiment en rafales, consultez Surveillance de l'équilibre du compartiment en rafales pour les volumes `gp2`, `st1` et `sc1` (p. 1104).

magnétique (`standard`)

Les volumes magnétique sont basés sur des disques magnétiques et conviennent aux charges de travail où l'accès aux données est rare et aux scénarios qui nécessitent un faible coût de stockage pour les volumes de petite taille. Ces volumes fournissent en moyenne 100 E/S par seconde, avec la possibilité d'émettre en rafale jusqu'à des centaines d'E/S par seconde. Leur taille varie entre 1 Gio et 1 Tio.

> **Note**
>
> magnétique est un volume de la génération précédente. Pour les nouvelles applications, nous recommandons d'utiliser l'un des types de volumes plus récents. Pour de plus amples informations, veuillez consulter Volumes de la génération précédente.

Pour plus d'informations sur l'utilisation des métriques et alarmes CloudWatch pour surveiller l'équilibre du compartiment en rafales, consultez Surveillance de l'équilibre du compartiment en rafales pour les volumes `gp2`, `st1` et `sc1` (p. 1104).

Considérations relatives aux performances lors de l'utilisation de volumes HDD

Pour des performances de débit optimales avec les volumes HDD, planifiez vos charges de travail en gardant à l'esprit les éléments suivants.

HDD à débit optimisé et Cold HDD

Les tailles de compartiment `st1` et `sc1` varient selon la taille du volume, et un compartiment complet contient assez de jetons pour une analyse complète du volume. Cependant, l'analyse des volumes `st1` et `sc1` de plus grande taille est plus longue en raison des limites de débit par instance et par volume. Les volumes attachés à des instances plus petites sont limités par le débit par instance plutôt que par les limites de débit de `st1` ou `sc1`.

`st1` et `sc1` sont conçus pour assurer l'homogénéité des performances de 90 % du débit de transmission en rafales 99 % du temps. Les périodes non conformes sont assez uniformément réparties, en ciblant 99 % du débit total attendu chaque heure.

Le tableau suivant illustre les durées d'analyse idéales pour les volumes de différentes tailles, en supposant que les compartiments sont complets et que le débit d'instance est suffisant.

En général, les durées d'analyse sont exprimées par cette formule :

```
 Volume size
------------- = Scan time
 Throughput
```

Par exemple, en prenant en compte les garanties en matière de cohérence des performances et les autres optimisations, un client `st1` avec un volume de 5 Tio effectue généralement une analyse complète du volume en 2,91 à 3,27 heures.

```
   5 TiB            5 TiB
----------- = -------------------- = 10,486 s = 2.91 hours (optimal)
 500 MiB/s     0.00047684 TiB/s


                2.91 hours
```

```
2.91 hours + -------------- = 3.27 hours (minimum expected)
              (0.90)(0.99) <-- From expected performance of 90% of burst 99% of the time
```

De même, un client `sc1` avec un volume de 5 Tio effectue généralement une analyse complète du volume en 5,83 à 6,54 heures.

```
      5 TiB
------------------- = 20972 s = 5.83 hours (optimal)
 0.000238418 TiB/s


  5.83 hours
-------------- = 6.54 hours (minimum expected)
 (0.90)(0.99)
```

Taille du volume (Tio)	Durée d'analyse avec transmission en rafales pour ST1 (en heures)*	Durée d'analyse avec transmission en rafales pour SC1 (en heures)*
1	1,17	3,64
2	1,17	3,64
3	1,75	3,64
4	2,33	4,66
5	2,91	5,83
6	3,50	6,99
7	4,08	8,16
8	4,66	9,32
9	5,24	10,49
10	5,83	11,65
11	6,41	12,82
12	6,99	13,98
13	7,57	15,15
14	8,16	16,31
15	8,74	17,48
16	9,32	18,64

* Ces durées d'analyse supposent une profondeur de file d'attente moyenne (arrondie au nombre entier le plus proche) de quatre éléments ou plus lors de l'exécution de 1 Mio d'E/S séquentielles.

Par conséquent, si vous avez une charge de travail axée sur le débit qui doit effectuer des analyses rapidement (jusqu'à 500 Mo/s) ou qui nécessite plusieurs analyses complètes de volume par jour, utilisez `st1`. Si vous cherchez à optimiser la rentabilité, si vous accédez à vos données de manière occasionnelle et si vous n'avez besoin de performances d'analyse de plus de 250 Mio/s, utilisez `sc1`.

Manque d'efficacité des lectures/écritures de petite taille sur disque dur

Le modèle de performances des volumes `st1` et `sc1` est optimisé pour les I/O séquentielles. Il favorise les charges de travail à haut débit et offre des performances acceptables avec les charges de travail dont les IOPS et le débit varient, tout en décourageant les charges de travail avec des I/O aléatoires de petite taille.

Par exemple, une requête d'E/S de 1 Mio ou moins correspond à un crédit d'E/S de 1 Mio. Toutefois, si les E/S sont séquentielles, elles sont fusionnées dans des blocs d'E/S de 1 Mio et correspondent uniquement à un crédit d'E/S de 1 Mio.

Restrictions de débit par instance

Le débit des volumes `st1` et `sc1` est toujours déterminé par la limite suivante la plus faible :

- Limites de débit du volume
- Limites de débit de l'instance

Comme pour tous les volumes Amazon EBS, nous vous recommandons de sélectionner une instance EC2 optimisées EBS appropriée afin d'éviter tout goulot d'étranglement du réseau. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Surveillance de l'équilibre du compartiment en rafales pour les volumes `gp2`, `st1` et `sc1`

Vous pouvez surveiller le niveau du compartiment en rafales des volumes `gp2`, `st1` et `sc1` à l'aide de la métrique EBS `BurstBalance` disponible dans Amazon CloudWatch. Cette métrique indique le pourcentage de crédits d'I/O (pour `gp2`) ou de crédits de débit (pour `st1` et `sc1`) restant dans le compartiment en rafales. Pour de plus amples informations sur la métrique `BurstBalance` et les autres métriques d'I/O, veuillez consulter Caractéristiques d'E/S et surveillance (p. 1233). CloudWatch vous permet également de définir une alarme qui vous avertit lorsque la valeur `BurstBalance` passe en dessous d'un certain niveau. Pour plus d'informations, consultez Création d'alarmes Amazon CloudWatch.

Contraintes sur la taille et la configuration d'un volume EBS

La taille d'un volume Amazon EBS est limitée par la physique et l'arithmétique du stockage de données en bloc, ainsi que par les décisions d'implémentation des concepteurs du système d'exploitation et du système de fichiers. AWS impose des limites supplémentaires à la taille du volume afin de préserver la fiabilité de ses services.

Les sections suivantes décrivent les facteurs les plus importants qui limitent la taille utilisable d'un volume EBS et fournissent des recommandations pour configurer vos volumes EBS.

Sommaire

- Capacité de stockage (p. 1104)
- Limitations de service (p. 1105)
- Schémas de partitionnement (p. 1105)
- Tailles des blocs de données (p. 1106)

Capacité de stockage

Le tableau suivant résume les capacités de stockage théoriques et implémentées des systèmes de fichiers les plus courants sur Amazon EBS, en supposant une taille de bloc de 4 096 octets.

Schéma de partitionnem	Nombre max de blocs adressables	Taille max théorique (blocs x taille de blocs)	Taille max implémentée Ext4*	Taille max implémentée XFS**	Taille max implémentée NTFS	Nombre max pris en charge par EBS
MBR	2^{32}	2 TiO	2 TiO	2 TiO	2 TiO	2 TiO
GPT	2^{64}	64 ZiO	1 Eio = 1024^2 Tio (50 Tio certifiés sur RHEL7)	500 TiO (certifiés sur RHEL7)	256 TiO	16 TiO

* https://ext4.wiki.kernel.org/index.php/Ext4_Howto et https://access.redhat.com/solutions/1532

** https://access.redhat.com/solutions/1532

Limitations de service

Amazon EBS extrait le stockage distribué massivement d'un centre de données sur des disques durs virtuels. Pour un système d'exploitation installé sur une instance EC2, un volume EBS attaché semble être un disque dur physique contenant des secteurs disque de 512 octets. Le système d'exploitation gère l'allocation des blocs de données (ou clusters) sur ces secteurs virtuels au moyen de ses utilitaires de gestion de stockage. L'allocation est conforme à un schéma de partitionnement de volume, comme un MBR (enregistrement de démarrage principal) ou GPT (table de partition GUID), et dans les capacités du système de fichiers installé (ext4, NTFS, etc.).

EBS n'est pas conscient des données contenues dans ses secteurs disque virtuels ; il s'assure uniquement de l'intégrité des secteurs. Cela signifie que les actions d'AWS et du système d'exploitation sont indépendantes les unes des autres. Lorsque vous sélectionnez une taille de volume, soyez conscient des capacités et des limites de chacune, comme dans les cas suivants :

- EBS prend actuellement en charge une taille de volume maximale de 16 Tio. Cela signifie que vous pouvez créer un volume EBS pouvant atteindre 16 Tio. Toutefois, le fait que le système d'exploitation reconnaîtra ou non l'ensemble de cette capacité dépend de ses propres caractéristiques de conception et de la façon dont le volume est partitionné.
- Les volumes de démarrage Linux peuvent utiliser le schéma de partitionnement MBR ou GPT. MBR prend en charge les volumes de démarrage jusqu'à 2 047 Gio (2 TiO - 1 GiO). GPT avec GRUB 2 prend en charge les volumes de démarrage de 2 Tio ou plus. Si votre AMI Linux utilise MBR, votre volume de démarrage est limité à 2047 Gio, mais cette limite ne s'applique pas aux volumes autres que ceux de démarrage. Pour plus d'informations, consultez Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113).

Schémas de partitionnement

Parmi les autres impacts, le schéma de partitionnement détermine le nombre de blocs de données logiques pouvant être traités de manière unique sur un seul volume. Pour plus d'informations, consultez Tailles des blocs de données (p. 1106). Les schémas de partitionnement communs utilisés sont MBR (enregistrement de démarrage principal) et GPT (table de partition GUID). Les différences importantes entre ces schémas peuvent être résumées comme suit.

MBR

MBR utilise une structure de données 32 bits pour stocker les adresses de blocs. Autrement, chaque bloc de données est mappé à l'un des 2^{32} entiers possibles. La taille maximale adressable d'un volume est fournie par :

```
(2^32 - 1) × Block size = Number of addressable blocks
```

La taille des blocs des volumes MBR est limitée par convention à 512 octets. Par conséquent :

```
(2^32 - 1) × 512 bytes = 2 TiB - 512 bytes
```

Les solutions d'ingénierie visant à augmenter cette limite de 2 Tio pour les volumes MBR n'ont pas été adoptées largement dans le secteur. Par conséquent, Linux et Windows ne détectent jamais un volume MBR comme étant supérieur à 2 Tio, même si AWS indique que sa taille est supérieure.

GPT

GPT utilise une structure de données 64 bits pour stocker les adresses de blocs. Autrement, chaque bloc de données est mappé à l'un des 2^{64} entiers possibles. La taille maximale adressable d'un volume est fournie par :

```
(2^64 - 1) × Block size = Number of addressable blocks
```

La taille des blocs des volumes GPT est limitée communément à 4 096 octets. Par conséquent :

```
(2^64 - 1) × 4,096 bytes
   = 2^64 x 4,096 bytes - 1 x 4,096 bytes
   = 2^64 x 2^12 bytes - 4,096 bytes
   = 2^70 x 2^6 bytes - 4,096 bytes
   = 64 ZiB - 4,096 bytes
```

Les systèmes informatiques réels ne prennent rien en charge qui atteigne ce maximum théorique. La taille du système de fichiers implémenté est actuellement limitée à 50 Tio pour ext4 et à 256 Tio pou NTFS, qui excèdent tous deux la limite de 16 Tio imposée par AWS.

Tailles des blocs de données

Le stockage de données sur un disque dur moderne est géré via l'adressage par blocs logiques, une couche d'abstraction qui permet au système d'exploitation de lire et d'écrire des données dans des blocs logiques sans bien connaître le matériel sous-jacent. Le système d'exploitation s'appuie sur le périphérique de stockage pour mapper les blocs à ses secteurs physiques. EBS publie ses secteurs de 512 octets sur le système d'exploitation, qui lit et écrit les données sur le disque à l'aide de blocs de données qui sont un multiple de la taille du secteur.

La taille par défaut des blocs de données logiques dans l'informatique est actuellement de 4 096 octets (4 Kio). Du fait que certaines charges de travail bénéficient d'une taille de taille inférieure ou supérieure, les systèmes de fichiers prennent en charge des tailles de blocs autres que par défaut et spécifiées au moment du formatage. Cette rubrique ne comporte pas de scénarios dans lesquels des tailles de blocs de données autres que par défaut sont utilisés, mais le choix de la taille des blocs a des conséquences sur la capacité de stockage du volume. Le tableau suivant indique la capacité de stockage en fonction de la taille des blocs :

Taille du bloc	Taille maximale du volume
4 Kio (par défaut)	16 Tio
8 Kio	32 Tio
16 Kio	64 Tio
32 Kio	128 Tio

Taille du bloc	Taille maximale du volume
64 Kio (maximum)	256 Tio

La limite imposée à EBS concernant la taille du volume (16 Tio) est actuellement égale à la taille maximale permise par les blocs de données de 4 Kio.

Création d'un volume Amazon EBS

Vous pouvez créer un volume Amazon EBS, puis l'attacher à n'importe quelle instance EC2 dans la même zone de disponibilité. Si vous créez un volume EBS chiffré, vous ne pouvez l'attacher qu'aux types d'instance pris en charge. Pour plus d'informations, consultez Types d'instance pris en charge (p. 1195).

Si vous créez un volume pour un scénario de stockage à hautes performances, vous devriez vous assurer d'utiliser un volume IOPS provisionnés SSD (`io1` ou `io2`) et l'attacher à une instance disposant de suffisamment de bande passante pour prendre en charge votre application, telle qu'une instance optimisée EBS ou une instance avec une connectivité réseau 10 gigabits. Le même conseil est valable pour les volumes HDD à débit optimisé (`st1`) et Cold HDD (`sc1`). Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Les volumes EBS vides reçoivent leurs performances maximum au moment où ils sont disponibles et ne nécessitent pas d'initialisation (anciennement préchauffage). Toutefois, les blocs de stockage des volumes ayant été créés à partir d'instantanés doivent être initialisés (extraits d'Amazon S3 et écrits sur le volume) avant de pouvoir accéder à ces blocs. Cette action préalable prend du temps et peut causer une hausse significative de la latence d'une opération E/S lors du premier accès à chaque bloc. Les performances du volume sont obtenues une fois que tous les blocs ont été téléchargés et écrits sur le volume. Pour la plupart des applications, l'amortissement de ce coût sur la durée de vie du volume est acceptable. Pour éviter cet impact initial sur les performances dans un environnement de production, vous pouvez forcer l'initialisation immédiate de l'intégralité du volume ou activer la restauration rapide des instantanés. Pour plus d'informations, consultez Initialisation des volumes Amazon EBS (p. 1235).

Méthodes de création d'un volume

- Créez et attachez des volumes EBS lorsque vous lancez des instances en spécifiant un mappage d'appareil de stockage en mode bloc. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530) et Mappage de périphérique de stockage en mode bloc (p. 1289).
- Créez un volume EBS et attachez-le à une instance en cours d'exécution. Pour plus d'informations, consultez Création d'un volume vide (p. 1107) ci-dessous.
- Créez un volume EBS à partir d'un instantané créé précédemment et attachez-le à une instance en cours d'exécution. Pour plus d'informations, consultez Création d'un volume à partir d'un instantané (p. 1108) ci-dessous.

Création d'un volume vide

Les volumes vides reçoivent leurs performances maximales au moment où ils sont disponibles et ne nécessitent pas d'initialisation.

Pour créer un volume EBS vide à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région dans laquelle vous souhaitez créer votre volume. Ce choix est important car certaines ressources Amazon EC2 peuvent être partagées entre plusieurs régions, contrairement à d'autres ressources. Pour plus d'informations, consultez Emplacements des ressources (p. 1301).

3. Dans le panneau de navigation, choisissez ELASTIC BLOCK STORE, Volumes.
4. Choisissez Créer un volume.
5. Pour Type de volume, choisissez un type de volume. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).
6. Pour Taille (Gio), entrez la taille du volume. Pour plus d'informations, consultez Contraintes sur la taille et la configuration d'un volume EBS (p. 1104).
7. Avec un volume IOPS provisionnés SSD pour IOPS, saisissez le nombre maximal d'opérations d'I/O par seconde (IOPS) que le volume doit prendre en charge.
8. Pour Zone de disponibilité, choisissez la zone de disponibilité dans laquelle créer le volume. Les volumes EBS ne peuvent être attachés qu'à des instances EC2 de la même zone de disponibilité.
9. (Facultatif) Si le type d'instance prend en charge le chiffrement EBS et que vous voulez chiffrer le volume, sélectionnez Chiffrer ce volume et choisissez une clé CMK. Si le chiffrement par défaut est activé dans cette région, le chiffrement EBS est activé et la clé CMK par défaut de chiffrement EBS est sélectionnée. Vous pouvez choisir une autre clé CMK dans Clé principale ou coller l'ARN complet de n'importe quelle clé à laquelle vous pouvez accéder. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).
10. (Facultatif) Choisissez Créer des balises supplémentaires pour ajouter des balises au volume. Pour chaque balise, indiquez une clé de balise et une valeur de balise. Pour plus d'informations, consultez Balisage de vos ressources Amazon EC2 (p. 1309).
11. Choisissez Créer un volume. Le volume est prêt à l'emploi lorsque son état indique Disponible.
12. Pour utiliser votre nouveau volume, attachez-le à une instance, formatez-le et montez-le. Pour plus d'informations, consultez Attacher un volume Amazon EBS à une instance (p. 1109).

Pour créer un volume EBS vide à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-volumeAWS CLI
- New-EC2VolumeOutils AWS pour Windows PowerShell

Création d'un volume à partir d'un instantané

Les volumes créés à partir d'instantanés se chargent lentement en arrière-plan. Cela signifie qu'il n'est pas nécessaire d'attendre que toutes les données soient transférées depuis Amazon S3 vers votre volume EBS avant que l'instance puisse commencer à accéder à un volume attaché et à toutes ses données. Si votre instance accède à des données qui n'ont pas encore été chargées, le volume télécharge immédiatement les données demandées depuis Amazon S3, puis continue à charger le reste des données de volume en arrière-plan. Les performances du volume sont obtenues une fois que tous les blocs ont été téléchargés et écrits sur le volume. Pour éviter l'impact initial sur les performances dans un environnement de production, consultez Initialisation des volumes Amazon EBS (p. 1235).

Les nouveaux volumes qui sont créés à partir d'instantanés chiffrés sont automatiquement chiffrés. Vous pouvez également chiffrer un volume à la volée pendant sa restauration à partir d'un instantané non chiffré. Les volumes chiffrés peuvent uniquement être attachés aux types d'instances compatibles avec le chiffrement EBS. Pour plus d'informations, consultez Types d'instance pris en charge (p. 1195).

Utilisez la procédure suivante pour créer un volume à partir d'un instantané.

Pour restaurer un volume EBS à partir d'un instantané à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région dans laquelle votre instantané se trouve.

Pour utiliser l'instantané afin de créer un volume dans une autre région, copiez votre instantané dans la nouvelle région, puis utilisez-le pour créer un volume dans cette région. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).

3. Dans le panneau de navigation, choisissez ELASTIC BLOCK STORE, Volumes.
4. Choisissez Créer un volume.
5. Pour Type de volume, choisissez un type de volume. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).
6. Dans le champ ID d'instantané, commencez à saisir l'ID ou la description de l'instantané à partir duquel vous restaurez le volume et sélectionnez-le dans la liste des options suggérées.
7. (Facultatif) La sélection de l'option Encrypted (Chiffré) vous permet de modifier l'état de chiffrement de votre volume. Ceci est facultatif si le chiffrement par défaut (p. 1196) est activé. Sélectionnez une CMK dans Clé principale pour spécifier une CMK autre que celle par défaut pour le chiffement EBS.
8. Dans la zone Taille (Gio), entrez la taille du volume ou vérifiez que la taille par défaut de l'instantané est correcte.

   Si vous spécifiez à la fois une taille de volume et un instantané, la taille doit être égale ou supérieure à la taille de l'instantané. Lorsque vous sélectionnez un type de volume et un instantané, les tailles minimale et maximale du volume s'affichent à côté de Taille. Pour plus d'informations, consultez Contraintes sur la taille et la configuration d'un volume EBS (p. 1104).
9. Avec un volume IOPS provisionnés SSD pour IOPS, saisissez le nombre maximal d'opérations d'I/O par seconde (IOPS) que le volume doit prendre en charge.
10. Pour Zone de disponibilité, choisissez la zone de disponibilité dans laquelle créer le volume. Les volumes EBS ne peuvent être attachés qu'aux instances EC2 de la même zone de disponibilité.
11. (Facultatif) Choisissez Créer des balises supplémentaires pour ajouter des balises au volume. Pour chaque balise, indiquez une clé de balise et une valeur de balise.
12. Choisissez Créer un volume.
13. Pour utiliser votre nouveau volume, attachez-le à une instance et montez-le. Pour plus d'informations, consultez Attacher un volume Amazon EBS à une instance (p. 1109).
14. Si vous avez créé un volume de taille supérieure à l'instantané, vous devez étendre le système de fichiers du volume pour profiter de l'espace supplémentaire. Pour plus d'informations, consultez Amazon EBS Elastic Volumes (p. 1182).

Pour créer un volume EBS à partir d'un instantané en utilisant la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-volumeAWS CLI
- New-EC2VolumeOutils AWS pour Windows PowerShell

Attacher un volume Amazon EBS à une instance

Vous pouvez attacher un volume EBS disponible à l'une de vos instances se trouvant dans la même zone de disponibilité que le volume.

Conditions préalables requises

- Déterminez combien de volumes vous pouvez attacher à votre instance. Pour plus d'informations, consultez Limites de volume d'instance (p. 1286).
- Déterminez si vous pouvez attacher votre volume à plusieurs instances et activer Multi-Attach. Pour plus d'informations, consultez Attacher un volume à plusieurs instances à l'aide d'Amazon EBS Multi-Attach (p. 1110).

- Si un volume est chiffré, il ne peut être attaché qu'à une instance prenant en charge Chiffrement Amazon EBS. Pour plus d'informations, consultez Types d'instance pris en charge (p. 1195).
- Si un volume a un code produit AWS Marketplace :
  - Le volume ne peut être attaché qu'à une instance arrêtée.
  - Vous devez être abonné au code AWS Marketplace qui se trouve sur le volume.
  - La configuration (type d'instance, système d'exploitation) de l'instance doit prendre en charge ce code AWS Marketplace spécifique. Par exemple, vous ne pouvez pas prendre un volume sur une instance Windows et l'attacher à une instance Linux.
  - Les codes produit AWS Marketplace sont copiés du volume vers l'instance.

Pour attacher un volume EBS à une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Elastic Block Store, Volumes.
3. Sélectionnez le volume et choisissez Actions, puis Attacher un volume.
4. Dans le champ Instance, saisissez le nom ou l'ID de l'instance. Sélectionnez l'instance dans la liste des options (seules les instances qui sont dans la même zone de disponibilité que le volume sont affichées).
5. Pour Dispositif, vous pouvez conserver le nom d'appareil suggéré ou entrer un autre nom d'appareil pris en charge. Pour plus d'informations, consultez Noms d'appareil pour les instances Linux (p. 1287).
6. Choisissez Attacher.
7. Connectez-vous à votre instance et montez le volume. Pour plus d'informations, consultez Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113).

Pour attacher un volume EBS à une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- attach-volume (AWS CLI)
- Add-EC2Volume (Outils AWS pour Windows PowerShell)

Attacher un volume à plusieurs instances à l'aide d'Amazon EBS Multi-Attach

Amazon EBS Multi-Attach vous permet d'attacher un volume SSD d'IOPS provisionnés (io1) à un maximum de 16 instances basées sur Nitro situées dans la même zone de disponibilité. Vous pouvez attacher plusieurs volumes activés pour Multi-Attach à une instance ou à un ensemble d'instances. Chaque instance à laquelle le volume est attaché dispose d'une autorisation complète en lecture et en écriture sur le volume partagé. Multi-Attach permet de bénéficier d'une disponibilité d'application plus importante dans les applications Linux en cluster qui gèrent des opérations d'écriture simultanée.

Sommaire

- Considérations et restrictions (p. 322)
- Performances (p. 1111)
- Utilisation de Multi-Attach (p. 1111)
- Surveillance (p. 1113)
- Tarification et facturation (p. 1113)

Considérations et restrictions

- Les volumes Multi-Attach peuvent être attachés à un maximum de 16 instances Linux construites sur le Système Nitro (p. 218) qui se trouvent dans la même zone de disponibilité. Vous pouvez attacher un volume Multi-Attach activé à des instances Windows, mais le système d'exploitation ne reconnaît pas les données sur le volume partagé entre les instances.
- Multi-Attach est pris en charge exclusivement sur les volumes SSD d'IOPS provisionnés (io1) (p. 1095). Il n'est pas pris en charge sur les volumes SSD d'IOPS provisionnés (io2).
- Multi-Attach est uniquement disponible dans les régions `us-east-1`, `us-west-2`, `eu-west-1` et `ap-northeast-2`.
- Les volumes activés pour Multi-Attach ne prennent pas en charge l'isolation d'E/S. Les protocoles d'isolation d'E/S contrôlent l'accès en écriture dans un environnement de stockage partagé afin de maintenir la cohérence des données. Vos applications doivent fournir un ordre d'écriture pour les instances attachées afin de maintenir la cohérence des données.
- Les volumes activés pour Multi-Attach ne peuvent pas être créés en tant que volumes de démarrage.
- Les volumes activés pour Multi-Attach peuvent être attachés à un mappage de périphérique de stockage en mode bloc par instance.
- Vous ne pouvez pas activer ou désactiver Multi-Attach après la création du volume.
- Vous ne pouvez pas modifier le type de volume, la taille ou les IOPS provisionnés d'un volume activé pour Multi-Attach.
- Multi-Attach ne peut pas être activé lors du lancement de l'instance via la console Amazon EC2 ou l'API RunInstances.
- Les volumes activés pour Multi-Attach présentant un problème au niveau de la couche d'infrastructure Amazon EBS sont indisponibles pour toutes les instances attachées. Des problèmes au niveau de la couche Amazon EC2 ou de mise en réseau peuvent affecter seulement certaines instances attachées.

Performances

Chaque instance attachée est capable de piloter ses performances IOPS maximales jusqu'aux performances provisionnées maximales du volume. Toutefois, les performances agrégées de toutes les instances attachées ne peuvent pas dépasser les performances provisionnées maximales du volume. Si la demande d'IOPS des instances attachées est supérieure aux IOPS provisionnées du volume, le volume ne dépassera pas ses performances provisionnées.

Par exemple, supposons que vous créez un volume `io1` activé pour Multi-Attach avec `50,000` IOPS provisionnées et que vous l'attachez à une instance `m5.8xlarge` et à une instance `c5.12xlarge`. Les instances `m5.8xlarge` et `c5.12xlarge` prennent respectivement en charge `30,000` et `40,000` IOPS. Chaque instance peut gérer ses IOPS maximum car la valeur est inférieure aux IOPS provisionnés (`50,000`) du volume. Toutefois, si les deux instances conduisent simultanément des E/S vers le volume, leurs E/S par seconde combinées ne peuvent pas dépasser les performances provisionnées du volume de `50,000 IOPS`. Le volume ne dépassera pas `50,000`IOPS.

Pour obtenir des performances cohérentes, il est recommandé d'équilibrer les E/S basées sur les instances attachées parmi les secteurs d'un volume activé pour Multi-Attach.

Utilisation de Multi-Attach

Les volumes activés pour Multi-Attach peuvent être gérés de la même manière que n'importe quel autre volume Amazon EBS. Toutefois, pour utiliser la fonctionnalité Multi-Attach, vous devez l'activer pour le volume. Lorsque vous créez un volume, Multi-Attach est désactivé par défaut.

Table des matières

- Activation de Multi-Attach (p. 1112)
- Attacher un volume à une instance (p. 1112)

- Suppression à la résiliation (p. 1112)

Activation de Multi-Attach

Vous pouvez activer Multi-Attach pour un volume Amazon EBS uniquement lors de la création.

Utilisez l'une des méthodes suivantes pour activer Multi-Attach pour un volume Amazon EBS lors de la création.

Console

Pour activer Multi-Attach lors de la création du volume

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Volumes.
3. Choisissez Créer un volume.
4. Pour Volume Type (Type de volume), choisissez Provisioned IOPS SSD (io1).
5. Pour Size (Taille) et IOPS, choisissez la taille de volume requise et le nombre d'E/S par seconde à provisionner.
6. Pour Availability Zone (Zone de disponibilité), choisissez la même zone de disponibilité que celle dans laquelle se trouvent les instances.
7. Pour Multi-Attach, choisissez Enable (Activer).
8. Choisissez Créer un volume.

Command line

Pour activer Multi-Attach lors de la création du volume

Utilisez la commande create-volume et spécifiez le paramètre `--multi-attach-enabled`.

```
$ aws ec2 create-volume --volume-type io1 --multi-attach-enabled --size 100 --iops 2000
 --region us-west-2 --availability-zone us-west-2b
```

Attacher un volume à une instance

Vous attachez un volume activé pour Multi-Attach à une instance de la même manière que vous attachez tout autre volume EBS. Pour plus d'informations, consultez Attacher un volume Amazon EBS à une instance (p. 1109).

Suppression à la résiliation

Les volumes activés pour Multi-Attach sont supprimés lors de la résiliation de l'instance si la dernière instance attachée est interrompue et si cette instance est configurée pour supprimer le volume lors de la résiliation. Si le volume est attaché à plusieurs instances présentant des paramètres de suppression à la résiliation différents dans leurs mappages de périphérique de stockage en mode bloc, le paramètre de mappage de périphériques de bloc de la dernière instance attachée détermine le comportement de suppression à la résiliation.

Pour garantir un comportement prévisible en matière de suppression à la résiliation, activez ou désactivez la suppression à la résiliation pour toutes les instances auxquelles le volume est attaché.

Par défaut, lorsqu'un volume est attaché à une instance, le paramètre de suppression à la résiliation pour le mappage de périphérique de stockage en mode bloc est défini sur false. Si vous souhaitez activer la suppression à la résiliation pour un volume activé pour Multi-Attach, modifiez le mappage de périphérique de stockage en mode bloc.

Si vous souhaitez que le volume soit supprimé lorsque les instances attachées sont résiliées, activez la suppression à la résiliation dans le mappage de périphérique de stockage en mode bloc pour toutes les instances attachées. Si vous souhaitez conserver le volume une fois que les instances attachées ont été résiliées, désactivez la suppression à la résiliation dans le mappage de périphérique de stockage en mode bloc pour toutes les instances attachées. Pour plus d'informations, consultez Conserver les volumes Amazon EBS lors de la résiliation d'une instance (p. 649).

Vous pouvez modifier le paramètre de suppression à la résiliation d'une instance lors de son lancement ou après son lancement. Si vous activez ou désactivez la suppression à la résiliation pendant le lancement de l'instance, les paramètres s'appliquent uniquement aux volumes attachés lors du lancement. Si vous attachez un volume à une instance après le lancement, vous devez définir explicitement le comportement de suppression à la résiliation pour ce volume.

Vous pouvez modifier le paramètre de suppression à la résiliation d'une instance à l'aide des outils de ligne de commande uniquement.

Pour modifier le paramètre de suppression à la résiliation pour une instance existante

Utilisez la commande modify-instance-attribute et spécifiez l'attribut `DeleteOnTermination` dans `--block-device-mappings option`.

```
aws ec2 modify-instance-attribute --instance-id i-1234567890abcdef0 --block-device-mappings
 file://mapping.json
```

Spécifiez les éléments suivants dans `mapping.json`.

```
[
  {
    "DeviceName": "/dev/sdf",
    "Ebs": {
      "DeleteOnTermination": true|false
    }
  }
]
```

Surveillance

Vous pouvez surveiller un volume activé pour Multi-Attach à l'aide des métriques CloudWatch pour des volumes Amazon EBS. Pour de plus amples informations, consultez Métriques Amazon CloudWatch pour Amazon EBS (p. 1247).

Les données sont agrégées dans toutes les instances attachées. Vous ne pouvez pas surveiller les métriques pour des instances individuelles attachées.

Tarification et facturation

L'utilisation d'Amazon EBS Multi-Attach est disponible sans frais supplémentaires. Vous êtes facturé selon les frais standard qui s'appliquent aux volumes SSD IOPS provisionnés ((io1). Pour de plus amples informations, veuillez consulter Tarification Amazon EBS.

Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux

Une fois un volume Amazon EBS attaché à votre instance, il est exposé en tant que périphérique de stockage en mode bloc. Vous pouvez formater le volume avec n'importe quel système de fichiers puis le monter. Après avoir rendu le volume EBS disponible à l'utilisation, vous pouvez y accéder de la même

façon que n'importe quel volume. Toutes les données écrites sur ce système de fichiers sont écrites sur le volume EBS et sont transparentes pour les applications utilisant cet appareil.

Vous pouvez prendre des instantanés de votre volume EBS à des fins de sauvegarde ou pour servir de base à la création d'un autre volume. Pour de plus amples informations, veuillez consulter Instantanés Amazon EBS (p. 1129).

Vous pouvez obtenir des instructions pour les volumes sur une instance Windows dans Rendre le volume disponible sur Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Formatage et montage d'un volume attaché

Supposons que vous ayez une instance EC2 avec un volume EBS pour le périphérique racine, `/dev/xvda`, et que vous venez d'attacher un volume EBS vide à l'instance en utilisant `/dev/sdf`. Utilisez la procédure suivante pour mettre le volume nouvellement attaché à disposition.

Pour formater et monter un volume EBS sous Linux

1. Connectez-vous à votre instance à l'aide de SSH. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Le périphérique peut être attaché à l'instance avec un nom de périphérique différent de celui que vous avez spécifié dans le mappage de périphérique de stockage en mode bloc. Pour plus d'informations, consultez Noms d'appareil pour les instances Linux (p. 1287). Utilisez la commande lsblk pour voir vos périphériques de disques disponibles et leurs points de montage (le cas échéant) pour vous aider à déterminer quel nom d'appareil utiliser. Le résultat de lsblk supprime le préfixe `/dev/` des chemins d'accès complets à l'appareil.

   Voici un exemple de sortie pour une instance construite sur le Système Nitro (p. 218), qui expose les volumes EBS en tant que périphériques de bloc NVMe. Le périphérique racine est `/dev/nvme0n1` Le volume attaché est `/dev/nvme1n1`, qui n'est pas encore monté.

   ```
   [ec2-user ~]$ lsblk
   NAME          MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
   nvme1n1       259:0    0  10G  0 disk
   nvme0n1       259:1    0   8G  0 disk
   -nvme0n1p1    259:2    0   8G  0 part /
   -nvme0n1p128  259:3    0   1M  0 part
   ```

   L'exemple ci-dessous représente la sortie pour une instance T2. Le périphérique racine est `/dev/xvda` Le volume attaché est `/dev/xvdf`, qui n'est pas encore monté.

   ```
   [ec2-user ~]$ lsblk
   NAME    MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
   xvda    202:0    0    8G  0 disk
   -xvda1  202:1    0    8G  0 part /
   xvdf    202:80   0   10G  0 disk
   ```

3. Déterminez s'il y a un système de fichiers sur le volume. Les nouveaux volumes sont des périphériques de stockage en mode bloc bruts et vous devez créer un système de fichiers sur ces volumes avant de pouvoir les monter et les utiliser. Les volumes créés à partir d'instantanés disposent probablement déjà d'un système de fichiers. Si vous créez un autre système de fichiers par-dessus le système de fichiers existant, l'opération remplace vos données.

   Utilisez la commande file -s pour obtenir les informations sur un périphérique, telles que son type de système de fichiers. Si le résultat de la commande précédente est simplement `data`, comme dans l'exemple de sortie suivant, il n'y a pas de système de fichiers sur l'appareil et vous devez en créer un.

   ```
   [ec2-user ~]$ sudo file -s /dev/xvdf
   ```

```
/dev/xvdf: data
```

Si l'appareil possède un système de fichiers, la commande affiche des informations sur le type de système de fichiers. Par exemple, la sortie suivante montre un périphérique racine avec le système de fichiers XFS.

```
[ec2-user ~]$ sudo file -s /dev/xvda1
/dev/xvda1: SGI XFS filesystem data (blksz 4096, inosz 512, v2 dirs)
```

4. (Condition) Si vous avez découvert qu'il y a un système de fichiers sur le périphérique à l'étape précédente, ignorez cette étape. Si vous avez un volume vide, utilisez la commande mkfs -t pour créer un système de fichiers sur le volume.

   Warning

   N'utilisez pas cette commande si vous montez un volume qui contient déjà des données (par exemple, un volume qui a été créé à partir d'un instantané). Sinon, vous formaterez le volume et supprimerez les données existantes.

```
[ec2-user ~]$ sudo mkfs -t xfs /dev/xvdf
```

   Si vous obtenez une erreur indiquant que `mkfs.xfs` est introuvable, utilisez la commande suivante pour installer les outils XFS, puis répétez la commande précédente :

```
[ec2-user ~]$ sudo yum install xfsprogs
```

5. Utilisez la commande mkdir pour créer un répertoire de point de montage pour le volume. Le point de montage est l'endroit où se trouve le volume dans l'arborescence du système de fichiers et où vous lisez et écrivez des fichiers après avoir monté le volume. L'exemple suivant crée un répertoire nommé `/data`.

```
[ec2-user ~]$ sudo mkdir /data
```

6. Utilisez la commande suivante pour monter le volume dans le répertoire que vous avez créé à l'étape précédente.

```
[ec2-user ~]$ sudo mount /dev/xvdf /data
```

7. Vérifiez les autorisations sur les fichiers de votre nouveau montage de volume pour vous assurer que les utilisateurs et les applications peuvent écrire sur le volume. Pour plus d'informations sur les autorisations sur les fichiers, consultez File security dans Le projet de documentation Linux.
8. Le point de montage n'est pas automatiquement préservé après le redémarrage de votre instance. Pour monter automatiquement ce volume EBS après le redémarrage, consultez Monter automatiquement un volume attaché après le redémarrage (p. 1115).

Monter automatiquement un volume attaché après le redémarrage

Pour monter un volume EBS attaché à chaque redémarrage du système, ajoutez une entrée pour l'appareil dans le fichier `/etc/fstab`.

Vous pouvez utiliser le nom du périphérique, comme `/dev/xvdf`, dans `/etc/fstab`, mais nous recommandons d'utiliser plutôt l'identificateur universel unique (UUID) de 128 bits de l'appareil. Les noms de périphériques peuvent changer, mais l'UUID persiste pendant toute la durée de vie de la partition. En utilisant l'UUID, vous réduisez les risques que le système devienne impossible à démarrer après une reconfiguration du matériel. Pour plus d'informations, consultez Identification du périphérique EBS (p. 1211).

Pour monter automatiquement un volume attaché après le redémarrage

1. (Facultatif) Créez une sauvegarde de votre fichier `/etc/fstab` que vous pouvez utiliser si vous détruisez ou supprimez accidentellement ce fichier en l'éditant.

```
[ec2-user ~]$ sudo cp /etc/fstab /etc/fstab.orig
```

2. Utilisez la commande blkid pour trouver l'UUID du périphérique.

```
[ec2-user ~]$ sudo blkid
/dev/xvda1: LABEL="/" UUID="ca774df7-756d-4261-a3f1-76038323e572" TYPE="xfs"
 PARTLABEL="Linux" PARTUUID="02dcd367-e87c-4f2e-9a72-a3cf8f299c10"
/dev/xvdf: UUID="aebf131c-6957-451e-8d34-ec978d9581ae" TYPE="xfs"
```

   Pour Ubuntu 18.04, utilisez la commande lsblk.

```
[ec2-user ~]$ sudo lsblk -o +UUID
```

3. Ouvrez le fichier `/etc/fstab` avec un éditeur de texte tel que nano ou vim.

```
[ec2-user ~]$ sudo vim /etc/fstab
```

4. Ajoutez l'entrée suivante à `/etc/fstab` pour monter le périphérique au point de montage spécifié. Les champs sont la valeur UUID renvoyée par blkid (ou lsblk pour Ubuntu 18.04), le point de montage, le système de fichiers et les options de montage recommandées. Pour de plus amples informations, veuillez consulter la page du manuel pour fstab (exécutez man fstab).

```
UUID=aebf131c-6957-451e-8d34-ec978d9581ae  /data  xfs  defaults,nofail  0  2
```

   Note

   Si jamais vous démarrez votre instance sans ce volume attaché (par exemple, après avoir déplacé ce volume sur une autre instance), l'option de montage `nofail` permet à l'instance de démarrer même si des erreurs se produisent lors du montage du volume. Les dérivés Debian, y compris les versions Ubuntu antérieures à 16.04, doivent également ajouter l'option de montage `nobootwait`.

5. Pour vérifier que votre entrée fonctionne, exécutez les commandes suivantes pour démonter le périphérique, puis montez tous les systèmes de fichiers dans `/etc/fstab`. S'il n'y a pas d'erreur, le fichier `/etc/fstab` est correct et votre système de fichiers sera monté automatiquement après avoir été redémarré.

```
[ec2-user ~]$ sudo umount /data
[ec2-user ~]$ sudo mount -a
```

   Si vous recevez un message d'erreur, traitez les erreurs dans le fichier.

   Warning

   Des erreurs dans le fichier `/etc/fstab` peuvent rendre un système impossible à démarrer. N'arrêtez pas un système dont le fichier `/etc/fstab` contient des erreurs.

   Si vous n'êtes pas sûr de savoir comment corriger des erreurs dans `/etc/fstab` et que vous avez créé un fichier de sauvegarde lors de la première étape de la procédure, vous avez toujours la possibilité de restaurer votre fichier depuis votre fichier de sauvegarde avec la commande suivante.

```
[ec2-user ~]$ sudo mv /etc/fstab.orig /etc/fstab
```

Affichage d'informations sur un volume Amazon EBS

Vous pouvez visualiser les informations descriptives relatives à vos volumes EBS. Par exemple, vous pouvez afficher des informations sur tous les volumes d'une région spécifique ou afficher des informations détaillées sur un seul volume (parmi lesquelles la taille, le type de volume, le chiffrement ou non du volume, la clé principale utilisée pour chiffrer le volume et l'instance spécifique à laquelle le volume est attaché).

Vous pouvez obtenir des informations supplémentaires sur vos volumes EBS, telles que l'espace disque disponible, à partir du système d'exploitation sur l'instance.

Affichage des informations sur un volume

Pour obtenir des informations sur un volume EBS à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Volumes.
3. (Facultatif) Utilisez les options de filtre de la barre de recherche pour afficher uniquement les volumes qui vous intéressent. Par exemple, si vous connaissez l'ID d'instance, choisissez Instance ID (ID d'instance) dans le menu du champ de recherche, puis choisissez l'ID d'instance dans la liste fournie. Pour supprimer un filtre, sélectionnez-le à nouveau.
4. Sélectionnez le volume.
5. Le volet des détails permet d'inspecter les informations fournies sur le volume. Les informations de pièce jointe indiquent l'ID d'instance auquel ce volume est attaché et le nom d'appareil sous lequel il est attaché.
6. (Facultatif) Cliquez sur le lien Informations sur la pièce jointe pour afficher des détails supplémentaires sur l'instance.

Pour afficher les volumes EBS qui sont attachés à une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez l'instance.
4. Dans l'onglet Stockage, affichez les informations fournies sur les appareils racine et bloc.
5. (Facultatif) Choisissez un lien dans la colonne ID de volume pour afficher des détails supplémentaires sur le volume.

Pour obtenir des informations sur un volume EBS à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes pour afficher les attributs de volume. Pour plus d'informations, consultez Accès à Amazon EC2 (p. 3).

- describe-volumes (AWS CLI)
- Get-EC2Volume (Outils AWS pour Windows PowerShell)

État du volume

L'état du volume décrit la disponibilité d'un volume Amazon EBS. Vous pouvez afficher l'état du volume dans la colonne State (État) de la page Volumes de la console ou à l'aide de la commande describe-volumes de l'AWS CLI.

Les états de volume possibles sont les suivants :

creating

Le volume est en cours de création.

available

Le volume n'est pas attaché à une instance.

in-use

Le volume est attaché à une instance.

deleting

Le volume est en cours de suppression.

deleted

Le volume est supprimé.

error

Le matériel sous-jacent associé à votre volume EBS a échoué et les données associées au volume ne peuvent pas être récupérées. Pour de plus amples informations sur la restauration du volume ou la récupération des données sur le volume, veuillez consulter Mon volume EBS est associé au statut « error » (erreur).

Affichage des métriques d'un volume

Vous pouvez obtenir des informations supplémentaires sur vos volumes EBS à partir de Amazon CloudWatch. Pour plus d'informations, consultez Métriques Amazon CloudWatch pour Amazon EBS (p. 1247).

Affichage de l'espace disque disponible

Vous pouvez obtenir des informations supplémentaires sur vos volumes EBS, telles que l'espace disque disponible, à partir du système d'exploitation Linux sur l'instance. Par exemple, utilisez la commande suivante :

```
[ec2-user ~]$ df -hT /dev/xvda1
Filesystem     Type      Size  Used Avail Use% Mounted on
/dev/xvda1     xfs       8.0G  1.2G  6.9G  15% /
```

Remplacement d'un volume Amazon EBS à l'aide d'un instantané précédent

Les instantanés Amazon EBS constituent l'outil de sauvegarde idéal sur Amazon EC2 en raison de leur vitesse, de leur commodité et de leur coût. Lors de la création d'un volume à partir d'un instantané, vous recréez son état à un moment précis du passé avec toutes les données intactes. En attachant un volume créé à partir d'un instantané à une instance, vous pouvez dupliquer des données dans plusieurs régions, créer des environnements de test, remplacer un volume de production endommagé ou corrompu dans son intégralité ou récupérer des fichiers et des répertoires spécifiques et les transférer vers un autre volume attaché. Pour de plus amples informations, veuillez consulter Instantanés Amazon EBS (p. 1129).

Vous pouvez suivre la procédure ci-après pour remplacer un volume EBS par un autre volume créé à partir d'un instantané antérieur du volume. Vous devez détacher le volume actuel, puis attacher le nouveau volume.

Notez que les volumes EBS ne peuvent être attachés qu'aux instances EC2 de la même zone de disponibilité.

Pour remplacer un volume

1. Créez un volume à partir de l'instantané et notez l'ID du nouveau volume. Pour plus d'informations, consultez Création d'un volume à partir d'un instantané (p. 1108).
2. Sur la page Volumes, activez la case à cocher du volume à remplacer. Sous l'onglet Description, recherchez les informations de pièce jointe et notez le nom du périphérique du volume (par exemple, `/dev/sda1` ou `/dev/xvda` pour un volume racine, ou `/dev/sdb` ou `xvdb`) et l'ID de l'instance.
3. (Facultatif) Avant de pouvoir détacher le volume racine d'une instance, vous devez l'arrêter. Si vous ne remplacez pas le volume racine, vous pouvez passer à l'étape suivante sans arrêter l'instance. Sinon, pour arrêter l'instance, à partir des informations de pièce jointe, passez la souris sur l'ID de l'instance, cliquez avec le bouton droit de la souris et ouvrez l'instance dans un nouvel onglet de navigateur. Choisissez État de l'instance, Arrêter l'instance. Laissez l'onglet avec la page des instances ouverte et revenez à l'onglet du navigateur avec la page des volumes.
4. Avec le volume toujours sélectionné, choisissez Actions, Détacher le volume. Lorsque vous êtes invité à confirmer l'opération, choisissez Oui, détacher. Désactivez la case à cocher pour ce volume.
5. Activez la case à cocher correspondant au nouveau volume que vous avez créé à l'étape 1. Sélectionnez Actions, puis Attacher un volume. Entrez l'ID d'instance et le nom de périphérique que vous avez noté à l'étape 2, puis choisissez Attacher.
6. (Facultatif) Si vous avez arrêté l'instance, vous devez la redémarrer. Revenez à l'onglet du navigateur avec la page des instances et choisissez État de l'instance, Démarrer l'instance.
7. Connectez-vous à votre instance et montez le volume. Pour plus d'informations, consultez Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113).

Surveillance du statut de vos volumes

Amazon Web Services (AWS) fournit automatiquement les données que vous pouvez utiliser pour surveiller vos volumes Amazon Elastic Block Store (Amazon EBS).

Table des matières

- Vérifications du statut du volume EBS (p. 1119)
- Événements de volume EBS (p. 1121)
- Utilisation d'un volume dégradé (p. 1123)
- Utilisation de l'attribut de volume Auto-Enabled IO (Activation automatique des E/S) (p. 1125)

Pour de plus amples informations, veuillez consulter Métriques Amazon CloudWatch pour Amazon EBS (p. 1247) et Amazon CloudWatch Events pour Amazon EBS (p. 1253).

Vérifications du statut du volume EBS

Les contrôles de statut de volume vous permettent de mieux comprendre, suivre et gérer les incohérences potentielles des données d'un volume Amazon EBS. Ils sont destinés à vous fournir les informations dont vous avez besoin pour déterminer si vos volumes Amazon EBS rencontrent des problèmes et pour vous aider à contrôler comment un volume potentiellement incohérent est géré.

Les contrôles de statut de volume sont exécutés automatiquement toutes les cinq minutes et renvoie un statut de réussite ou d'échec. Si tous les contrôles réussissent, le statut du volume est `ok`. Si un contrôle échoue, le statut du volume est `impaired`. Si le statut est `insufficient-data`, il se peut que les contrôles soient toujours en cours sur le volume. Vous pouvez afficher les résultats des contrôles de statut de volume pour identifier les volumes confrontés à des problèmes et prendre les actions nécessaires.

Quand Amazon EBS détermine que les données d'un volume sont potentiellement incohérentes, par défaut, les E/S sur le volume sont désactivées à partir des instances EC2 attachées, ce qui permet d'empêcher la corruption des données. Une fois que les E/S ont été désactivées, le contrôle de statut de volume suivant échoue et le statut du volume est `impaired`. De plus, vous remarquerez un événement qui vous permet de savoir que les E/S sont désactivées, et que vous pouvez résoudre le statut de défaillance du volume en activant les E/S sur le volume. Nous attendons jusqu'à ce que vous activiez les I/O pour vous offrir la possibilité de décider si vous laissez vos instances continuer à utiliser le volume ou d'exécuter un contrôle de cohérence à l'aide d'une commande, telle que fsck, avant de procéder ainsi.

> **Note**
>
> Le statut du volume s'appuie sur les vérifications de statut du volume et ne reflète pas l'état du volume. Par conséquent, le statut du volume n'indique pas de volumes avec l'état `error` (par exemple, lorsqu'un volume est incapable d'accepter l'E/S.) Pour obtenir des informations sur l'état du volume, veuillez consulter État du volume (p. 1117).

Si la cohérence d'un volume particulier ne constitue pas un problème et que vous préfèreriez que le volume soit rendu disponible immédiatement s'il rencontre des problèmes, vous pouvez remplacer le comportement par défaut en configurant le volume de façon à activer automatiquement les I/O. Si vous activez l'attribut de volume Auto-Enable IO (Activation automatique des E/S) (`autoEnableIO` dans l'API), le contrôle du statut de volume continue à s'effectuer avec succès. De plus, vous remarquerez un événement qui vous permet de savoir que le volume a été déterminé pour être potentiellement incohérent, mais que ses E/S ont été automatiquement activées. Cela vous permet de vérifier la cohérence du volume ou de le remplacer ultérieurement.

Le contrôle de statut des performances d'E/S compare les performances réelles du volume aux performances attendues, et vous prévient si le volume se comporte en-deçà des attentes. Cette vérification d'état n'est disponible que pour les volumes IOPS provisionnés SSD (`io1` and `io2`) attachés à une instance. Il n'est pas valide pour les volumes SSD à usage général (`gp2`), HDD à débit optimisé (`st1`), Cold HDD (`sc1`) ou magnétique (`standard`). Comme le contrôle de statut des performances d'I/O s'effectue toutes les minutes et que CloudWatch collecte les données toutes les 5 minutes, il peut s'écouler jusqu'à 5 minutes entre le moment où vous attachez un volume `io1` ou `io2` à une instance et celui où le contrôle rapporte le statut des performances d'I/O.

> **Important**
>
> Lors de l'initialisation des volumes `io1` and `io2` restaurés à partir des instantanés, les performances du volume peuvent chuter jusqu'à plus de 50 % en dessous du niveau attendu, ce qui entraîne l'affichage par le volume d'un état `warning` dans le contrôle de statut Performances des E/S. Cette situation est attendue et vous pouvez ignorer l'état `warning` des volumes `io1` and `io2` lorsque vous les initialisez. Pour plus d'informations, consultez Initialisation des volumes Amazon EBS (p. 1235).

Le tableau suivant répertorie les statuts des volumes Amazon EBS.

Statut du volume	Statut d'activation des E/S	Statut des performances d'E/S (uniquement disponible pour les volumes IOPS provisionnés)
`ok`	Activé (E/S activées ou E/S activées automatiquement)	Normal (performances du volume telles qu'attendues)
`warning`	Activé (E/S activées ou E/S activées automatiquement)	Dégradé (performances du volume inférieures aux attentes) Profondément dégradé (performances du volume bien inférieures aux attentes)

Statut du volume	Statut d'activation des E/S	Statut des performances d'E/S (uniquement disponible pour les volumes IOPS provisionnés)
`impaired`	Activé (E/S activées ou E/S activées automatiquement) Désactivé (volume hors connexion et récupération en attente, ou en attente d'activation par l'utilisateur des E/S)	Interrompu (performances du volume profondément impactées) Non disponible (impossible de déterminer les performances d'E/S parce que les E/S sont désactivées)
`insufficient-data`	Activé (E/S activées ou E/S activées automatiquement) Données insuffisantes	Données insuffisantes

Pour afficher et utiliser les contrôles de statut, vous pouvez utiliser la console Amazon EC2 l'API ou l'interface ligne de commande.

Pour afficher les contrôles de statut dans la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Volumes. La colonne Volume Status (Statut du volume) affiche le statut opérationnel de chaque volume.
3. Pour afficher les détails du statut d'un volume spécifique, sélectionnez ce volume, puis Contrôles des statuts.

4. Si vous avez un volume avec un contrôle de statut ayant échoué (le statut est impaired (dégradé)), consultez Utilisation d'un volume dégradé (p. 1123).

Vous pouvez aussi choisir Événements dans le navigateur pour afficher tous les événements de vos instances et volumes. Pour plus d'informations, consultez Événements de volume EBS (p. 1121).

Pour afficher les informations de statut de volume avec la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes pour afficher le statut de vos volumes Amazon EBS. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-volume-status (AWS CLI)
- Get-EC2VolumeStatus (Outils AWS pour Windows PowerShell)

Événements de volume EBS

Quand Amazon EBS détermine que les données d'un volume sont potentiellement incohérentes, par défaut, les E/S sur le volume sont désactivées à partir des instances EC2 attachées. Il s'ensuit que le

contrôle du statut du volume échoue et qu'un événement de statut de volume est créé indiquant la raison de l'échec.

Pour activer automatiquement les E/S sur un volume avec des incohérences de données potentielles, changez le paramètre de l'attribut de volume Auto-Enabled IO (Activation automatique des E/S) (`autoEnableIO` dans l'API). Pour plus d'informations sur la modification de cet attribut, consultez Utilisation d'un volume dégradé (p. 1123).

Chaque événement inclut une heure de début qui indique l'heure à laquelle l'événement s'est produit, ainsi qu'une durée qui spécifie combien de temps les E/S du volume ont été désactivées. L'heure de fin est ajoutée à l'événement quand les E/S du volume sont activées.

Les événements de statut de volume incluent l'une des descriptions suivantes :

Action en attente : activer E/S

Les données du volume sont potentiellement incohérentes. Les E/S sont désactivées pour le volume jusqu'à ce que vous les activiez explicitement. La description de l'événement devient IO Enabled après que vous avez explicitement activé les I/O.

Activation des E/S

Les opérations d'E/S ont été explicitement activées pour ce volume.

Activation automatique des E/S

Les opérations d'E/S ont été automatiquement activées sur ce volume après qu'un événement s'est produit. Nous vous recommandons de vérifier leurs incohérences avant de continuer à utiliser les données.

Normal

Pour les volumes `io1` and `io2` uniquement. Performances du volume telles qu'attendues.

Dégradé

Pour les volumes `io1` and `io2` uniquement. Performances du volume inférieures aux attentes.

Profondément dégradé

Pour les volumes `io1` and `io2` uniquement. Performances du volume bien inférieures aux attentes.

Interrompu

Pour les volumes `io1` and `io2` uniquement. Performances du volume profondément impactées.

Vous pouvez afficher les événements de vos volumes à l'aide de la console Amazon EC2, de l'API ou de l'interface ligne de commande.

Pour afficher les événements de vos volumes dans la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, sélectionnez Events. Tous les volumes et instances ayant des événements sont affichés.
3. Vous pouvez filtrer par volume pour n'afficher que le statut de volume. Vous pouvez aussi filtrer sur des types de statut spécifiques.
4. Sélectionnez un volume pour afficher son événement spécifique.

Si vous avez un volume où les E/S sont désactivées, consultez Utilisation d'un volume dégradé (p. 1123). Si vous avez un volume où les performances des E/S sont inférieures à la normale, il peut s'agir d'une condition temporaire due à une action que vous avez prise (par exemple, création d'un instantané d'un volume lors d'une utilisation de pointe, exécution du volume sur une instance qui ne peut pas prendre en charge la bande passante d'E/S requise ou premier accès aux données du volume).

Pour afficher les événements de vos volumes avec la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes pour afficher les informations d'événement de vos volumes Amazon EBS. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-volume-status (AWS CLI)
- Get-EC2VolumeStatus (Outils AWS pour Windows PowerShell)

Utilisation d'un volume dégradé

Cette section présente vos options si un volume est dégradé parce que ses données sont potentiellement incohérentes.

Options

- Option 1 : exécuter un contrôle de cohérence sur le volume attaché à son instance (p. 1123)
- Option 2 : exécuter un contrôle de cohérence sur le volume à l'aide d'une autre instance (p. 1124)
- Option 3 : supprimer le volume si vous n'en avez plus besoin (p. 1125)

Option 1 : exécuter un contrôle de cohérence sur le volume attaché à son instance

L'option la plus simple consiste à activer les E/S, puis à exécuter un contrôle de cohérences des données sur le volume, pendant que celui-ci est toujours attaché à son instance Amazon EC2.

Pour exécuter un contrôle de cohérence sur un volume attaché

1. Arrêtez l'utilisation du volume par les applications.
2. Activez les E/S sur le volume.

   a. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

b. Dans le panneau de navigation, choisissez Volumes.

c. Sélectionnez le volume sur lequel vous souhaitez activer les opérations d'E/S.

d. Dans le volet des détails, sélectionnez Activation des E/S du volume, puis choisissez Yes, Enable (Oui, activer).

3. Vérifiez les données du volume.

   a. Exécutez la commande fsck.

   b. (Facultatif) Recherchez dans les journaux des applications journaux système disponibles les messages d'erreur appropriés.

   c. Si le volume a été dégradé pendant plus de 20 minutes, vous pouvez contacter le Centre de support AWS. Sélectionnez Dépannage puis, dans la boîte de dialogue Dépanner les contrôles de statut, sélectionnez Contactez Support pour soumettre une demande de support.

Pour activer les E/S pour un volume avec la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes pour afficher les informations d'événement de vos volumes Amazon EBS. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- enable-volume-io (AWS CLI)
- Enable-EC2VolumeIO (Outils AWS pour Windows PowerShell)

Option 2 : exécuter un contrôle de cohérence sur le volume à l'aide d'une autre instance

Utilisez la procédure suivante pour vérifier le volume en dehors de votre environnement de production.

> **Important**
>
> Cette procédure peut entraîner la perte d'E/S en écriture suspendues quand les E/S du volume ont été désactivées.

Pour exécuter un contrôle de cohérence sur un volume isolé

1. Arrêtez l'utilisation du volume par les applications.
2. Détachez le volume de l'instance.

   a. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

   b. Dans le panneau de navigation, choisissez Volumes.

   c. Sélectionnez le volume à détacher.

   d. Sélectionnez Actions, puis Force Detach Volume (Forcer le détachement de volume). Vous serez invité à confirmer votre choix.
3. Activez les E/S sur le volume.

   a. Dans le panneau de navigation, choisissez Volumes.

   b. Sélectionnez le volume que vous avez détaché à l'étape précédente.

c. Dans le volet des détails, sélectionnez Activation des E/S du volume, puis choisissez Yes, Enable (Oui, activer).

4. Attachez le volume à une autre instance. Pour plus d'informations, consultez Lancer votre instance (p. 529) et Attacher un volume Amazon EBS à une instance (p. 1109).
5. Vérifiez les données du volume.

   a. Exécutez la commande fsck.
   b. (Facultatif) Recherchez dans les journaux des applications journaux système disponibles les messages d'erreur appropriés.
   c. Si le volume a été dégradé pendant plus de 20 minutes, vous pouvez contacter le Centre de support AWS. Sélectionnez Dépannage puis, dans la boîte de dialogue de dépannage, sélectionnez Contactez Support pour soumettre une demande de support.

Pour activer les E/S pour un volume avec la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes pour afficher les informations d'événement de vos volumes Amazon EBS. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- enable-volume-io (AWS CLI)
- Enable-EC2VolumeIO (Outils AWS pour Windows PowerShell)

Option 3 : supprimer le volume si vous n'en avez plus besoin

Si vous voulez supprimer le volume de votre environnement, supprimez-le simplement. Pour plus d'informations sur la suppression d'un volume, consultez Suppression d'un volume Amazon EBS (p. 1129).

Si vous avez un instantané récent qui sauvegarde les données sur le volume, vous pouvez créer un volume à partir de l'instantané. Pour plus d'informations, consultez Création d'un volume à partir d'un instantané (p. 1108).

Utilisation de l'attribut de volume Auto-Enabled IO (Activation automatique des E/S)

Quand Amazon EBS détermine que les données d'un volume sont potentiellement incohérentes, par défaut, les E/S sur le volume sont désactivées à partir des instances EC2 attachées. Il s'ensuit que le contrôle du statut du volume échoue et qu'un événement de statut de volume est créé indiquant la raison de l'échec. Si la cohérence d'un volume particulier ne constitue pas un problème et que vous préfèrerez que le volume soit rendu disponible immédiatement s'il affiche impaired (dégradé), vous pouvez remplacer le comportement par défaut en configurant le volume de façon à activer automatiquement les I/O. Si vous activez l'attribut de volume Auto-Enabled IO (Activation automatique des E/S) (`autoEnableIO` dans l'API), les E/S entre le volume et l'instance sont automatiquement réactivées et le contrôle de statut du volume aboutit. De plus, vous remarquerez un événement qui vous permet de savoir que le volume se trouvait dans un état potentiellement incohérent, mais que ses E/S ont été automatiquement activées. Quand cet événement se produit, vous devez vérifier la cohérence du volume et le remplacer si nécessaire. Pour de plus amples informations, veuillez consulter Événements de volume EBS (p. 1121).

Cette procédure explique comment afficher et modifier l'attribut Auto-Enabled IO d'un volume.

Pour afficher l'attribut Auto-Enabled IO (Activation automatique des E/S) d'un volume dans la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Volumes.
3. Sélectionnez le volume et choisissez Status Checks (Contrôles de statut). L'attribut Auto-Enabled IO (Activation automatique des E/S) affiche le paramètre actuel (Enabled (Activé) ou Disabled (Désactivé)) pour votre volume.

Pour modifier l'attribut Auto-Enabled IO (Activation automatique des E/S) d'un volume dans la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Volumes.
3. Sélectionnez le volume, puis choisissez Actions, Change Auto-Enable IO Setting (Modifier le paramètre d'activation automatique des E/S). Vous pouvez également cliquer sur l'onglet Status Checks (Contrôles de statut) et pour Auto-Enabled IO (Activation automatique des E/S), choisir Edit (Modifier).

4. Cochez la case Auto-Enable Volume IO (Activer automatiquement les E/S du volume) afin d'activer automatiquement les E/S d'un volume dégradé. Pour désactiver la fonction, décochez la case.

5. Choisissez Save.

Pour afficher ou modifier l'attribut AutoEnableIO d'un volume avec la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes pour afficher l'attribut `autoEnableIO` de vos volumes Amazon EBS. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-volume-attribute (AWS CLI)
- Get-EC2VolumeAttribute (Outils AWS pour Windows PowerShell)

Pour modifier l'attribut `autoEnableIO` d'un volume, vous pouvez utiliser l'une des commandes ci-après.

- modify-volume-attribute (AWS CLI)
- Edit-EC2VolumeAttribute (Outils AWS pour Windows PowerShell)

Détacher un volume Amazon EBS d'une instance Linux

Détacher un volume Amazon EBS d'une instance entraîne le fait que ce volume peut être attaché à une autre instance ou supprimé. Le détachement d'un volume n'affecte pas les données du volume.

Considérations

- Vous pouvez détacher un volume Amazon EBS d'une instance explicitement ou en mettant fin à l'instance. Toutefois, si l'instance est en cours d'exécution, vous devez d'abord démonter le volume à partir de l'instance.
- Si un volume EBS est le volume racine d'une instance, vous devez arrêter l'instance avant de pouvoir détacher le volume.
- Vous pouvez rattacher un volume que vous avez détaché (sans l'avoir démonté), mais celui-ci n'aura peut-être pas le même point de montage. S'il y avait des écritures en cours sur le volume au moment où il a été détaché, les données sur le volume peuvent ne pas être synchronisées.
- Après avoir détaché un volume, vous continuez à payer le stockage de volume tant que la quantité de stockage dépasse la limite du niveau d'offre gratuite AWS. Vous devez supprimer un volume pour éviter de générer des frais supplémentaires. Pour plus d'informations, consultez Suppression d'un volume Amazon EBS (p. 1129).

Vous pouvez obtenir des directions pour les volumes sur une instance Windows à partir de Détacher un volume d'une instance Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Démonter et détacher un volume

Utilisez la procédure suivante pour démonter et détacher un volume d'une instance. Cela peut être utile lorsque vous devez attacher le volume à une autre instance.

Pour détacher un volume EBS à l'aide de la console

1. À partir de votre instance Linux, utilisez la commande suivante pour démonter l'unité `/dev/sdh`.

```
[ec2-user ~]$ umount -d /dev/sdh
```

2. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
3. Dans le panneau de navigation, choisissez Volumes.

4. Sélectionnez un volume et choisissez Actions, puis Detach Volume (Détacher un volume).
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Oui, détacher.

Pour détacher un volume EBS d'une instance à l'aide de la ligne de commande

Après avoir démonté le volume, vous pouvez utiliser l'une des commandes suivantes pour le détacher. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- detach-volume (AWS CLI)
- Dismount-EC2Volume (Outils AWS pour Windows PowerShell)

Dépannage

Voici des problèmes courants rencontrés lors du détachement de volumes, ainsi que la façon de les résoudre.

> **Note**
>
> Pour vous prémunir contre la possibilité de perte de données, prenez un instantané de votre volume avant d'essayer de le démonter. Le détachement forcé d'un volume bloqué peut endommager le système de fichiers ou les données qu'il contient ou entraîner une incapacité d'attacher un volume à l'aide du même nom de périphérique, sauf si vous redémarrez l'instance.

- Si vous rencontrez des problèmes lors du détachement d'un volume à l'aide de la console Amazon EC2, il peut être utile d'utiliser la commande describe-volumes de l'interface de ligne de commande pour diagnostiquer le problème. Pour plus d'informations, consultez describe-volumes.
- Si votre volume reste à l'état `detaching`, vous pouvez forcer le détachement en cliquant sur Force Detach (Forcer le détachement). Utilisez cette option uniquement comme dernier recours pour détacher un volume d'une instance en échec, ou si vous détachez un volume avec l'intention de le supprimer. L'instance n'a pas la possibilité de vider les caches du système de fichiers ou les métadonnées du système de fichiers. Si vous utilisez cette option, vous devez effectuer un contrôle du système de fichiers et des procédures de réparation.
- Si vous avez tenté à plusieurs reprises et pendant plusieurs minutes de détacher le volume et que celui-ci reste à l'état `detaching`, vous pouvez publier une demande d'aide sur le Amazon EC2 forum. Pour aider à accélérer la résolution d'un problème, incluez l'ID du volume et décrivez les étapes que vous avez déjà effectuées.
- Lorsque vous essayez de détacher un volume qui est toujours monté, le volume peut se bloquer dans l'état `busy` lorsque vous tentez de le détacher. La sortie suivante de la commande describe-volumes présente un exemple de cette condition :

```
"Volumes": [
    {
        "AvailabilityZone": "us-west-2b",
        "Attachments": [
            {
                "AttachTime": "2016-07-21T23:44:52.000Z",
                "InstanceId": "i-fedc9876",
                "VolumeId": "vol-1234abcd",
                "State": "busy",
                "DeleteOnTermination": false,
                "Device": "/dev/sdf"
            }
        ...
    }
]
```

Lorsque vous rencontrez cet état, le détachement peut être retardé indéfiniment jusqu'à ce que vous démontiez le volume, forciez le détachement, redémarriez l'instance ou les trois.

Suppression d'un volume Amazon EBS

Lorsque vous n'avez plus besoin d'un volume Amazon EBS, vous pouvez le supprimer. Une fois le volume supprimé, ses données sont perdues et il ne peut être attaché à aucune instance. Toutefois, avant sa suppression, vous pouvez stocker un instantané du volume que vous pouvez utiliser ultérieurement pour recréer le volume.

> **Note**
>
> Vous ne pouvez pas supprimer un volume si celui-ci est attaché à une instance. Pour supprimer un volume, vous devez d'abord le détacher. Pour plus d'informations, consultez Détacher un volume Amazon EBS d'une instance Linux (p. 1127).
> Vous pouvez vérifier si un volume est attaché à une instance. Dans la console, sur la page Volumes ,vous pouvez afficher l'état de vos volumes.
>
> - Si un volume est attaché à une instance, son état est `in-use`.
> - Si un volume est détaché d'une instance, son état est `available`. Vous pouvez supprimer ce volume.

Pour supprimer un volume EBS à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Volumes.
3. Sélectionnez un volume et choisissez Actions, puis Delete Volume (Supprimer un volume). Si Delete Volume (Supprimer le volume) est grisé, le volume est attaché à une instance.
4. Dans la boîte de dialogue de confirmation, choisissez Oui, supprimer.

Pour supprimer un volume EBS à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- delete-volume (AWS CLI)
- Remove-EC2Volume (Outils AWS pour Windows PowerShell)

Instantanés Amazon EBS

Vous pouvez sauvegarder les données des volumes Amazon EBS sur Amazon S3 en créant des instantanés à un instant donné. Les instantanés sont des sauvegardes incrémentielles, ce qui signifie que seuls les blocs de l'appareil qui ont changé depuis l'instantané le plus récent sont enregistrés. Cela réduit le temps nécessaire pour créer l'instantané, ainsi que les coûts de stockage en ne dupliquant pas les données. Chaque instantané contient toutes les informations nécessaires à la restauration de vos données (à partir du moment où l'instantané a été pris) sur un nouveau volume EBS.

Lorsque vous créez un volume EBS basé sur un instantané, au départ, le nouveau volume est donc une copie fidèle du volume initial qui a été utilisé pour créer l'instantané. Le volume répliqué charge les données en arrière-plan afin que vous puissiez commencer à les utiliser immédiatement. Si vous avez besoin d'accéder à des données qui n'ont pas encore été chargées, le volume télécharge immédiatement les données demandées depuis Amazon S3, puis continue à charger le reste des données du volume

en arrière-plan. Pour de plus amples informations, veuillez consulter Création d'instantanés Amazon EBS (p. 1132).

Lorsque vous supprimez un instantané, seules les données figurant uniquement dans cet instantané sont supprimées. Pour plus d'informations, consultez Suppression d'un instantané Amazon EBS (p. 1136).

Événements d'instantané

Vous pouvez suivre l'état de vos instantanés EBS via CloudWatch Events. Pour de plus amples informations, veuillez consulter Événements d'instantané EBS (p. 1257).

Instantanés multi-volumes

Des instantanés peuvent être utilisés pour créer une sauvegarde de charges de travail essentielles, telles qu'une grande base de données ou un système de fichiers qui s'étend sur plusieurs volumes EBS. Les instantanés multi-volumes vous permettent de prendre des instantanés en lien avec la panne, aux données coordonnées et à un instant donné exact, sur plusieurs volumes EBS attachés à une instance EC2. Vous n'avez plus besoin d'arrêter votre instance ni de la coordonner entre les volumes pour assurer un lien avec la panne, car les instantanés sont automatiquement pris sur plusieurs volumes EBS. Pour de plus amples informations, veuillez consulter les étapes de la création d'un instantané EBS multi-volume, dans Création d'instantanés Amazon EBS (p. 1132).

Tarification des instantanés

Les frais pour vos instantanés sont basés sur la quantité de données stockées. Étant donné que les instantanés sont incrémentiels, la suppression d'un instantané risque de ne pas réduire vos coûts de stockage des données. Les données référencées exclusivement par un instantané sont supprimées lorsque cet instantané est supprimé, mais les données référencées par d'autres instantanés sont conservées. Pour de plus amples informations, veuillez consulter Volumes et instantanés Amazon Elastic Block Store dans le AWS Billing and Cost Management Guide de l'utilisateur.

Sommaire

- Fonctionnement des instantanés incrémentiels (p. 1130)
- Copie et partage d'instantanés (p. 1132)
- Prise en charge du chiffrement pour les instantanés (p. 1132)
- Création d'instantanés Amazon EBS (p. 1132)
- Suppression d'un instantané Amazon EBS (p. 1136)
- Copie d'un instantané Amazon EBS (p. 1138)
- Affichage d'informations sur un instantané Amazon EBS (p. 1142)
- Partage d'un instantané Amazon EBS (p. 1143)
- Utilisation d'API directes EBS pour accéder au contenu d'un instantané EBS (p. 1147)
- Automatisation du cycle de vie des instantanés Amazon EBS (p. 1169)

Fonctionnement des instantanés incrémentiels

Cette section fournit des illustrations de la façon dont un instantané EBS capture l'état d'un volume à un moment donné et dont des instantanés successifs d'un volume modifié crée un historique de ces modifications.

Dans le graphique ci-dessous, Volume 1 est affiché à trois moments différents. Un instantané de chacun de ces trois états du volume est pris.

- A l'état 1, le volume contient 10 Gio de données. Comme Snap A est le premier instantané pris du volume, la totalité des 10 Gio de données doit être copiée.

- A l'état 2, le volume contient toujours 10 Gio de données, mais 4 Gio de données ont été modifiés. Snap B a besoin d'une copie et ne stocke que les 4 Gio qui ont été modifiés après que l'instantané Snap A ait été pris. Les autres 6 Gio de données non modifiées qui avaient déjà été copiés et stockés dans Snap A, sont référencés par Snap B plutôt que copiés (à nouveau). Ceci est indiqué par la flèche en pointillé.
- A l'état 3, 2 Gio de données ont été ajoutés au volume, pour un total de 12 Gio. Snap C a besoin d'une copie des 2 Gio qui ont été ajoutés après que l'instantané Snap B ait été pris. Comme indiqué par les flèches en pointillé, Snap C fait également référence aux 4 Gio de données stockés dans Snap B, et aux 6 Gio de données stockés dans Snap A.
- L'espace de stockage total nécessaire pour les trois instantanés est de 16 Gio.

Relations entre plusieurs instantanés d'un volume

Note

Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Pour plus d'informations sur la façon dont les données sont gérées lorsque vous supprimez un instantané, consultez Suppression d'un instantané Amazon EBS (p. 1136).

Copie et partage d'instantanés

Vous pouvez partager un instantané entre plusieurs comptes AWS en modifiant ses autorisations d'accès. Vous pouvez faire des copies de vos propres instantanés, ainsi que de ceux qui ont été partagés avec vous. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

Un instantané est limité à la région AWS dans laquelle il a été créé. Après avoir créé l'instantané d'un volume EBS, vous pouvez l'utiliser afin de créer d'autres volumes dans la même région. Pour plus d'informations, consultez Création d'un volume à partir d'un instantané (p. 1108). Vous pouvez également copier les instantanés d'une région à une autre, ce qui vous permet d'utiliser plusieurs régions à des fins d'expansion géographique, de migration des centres de données et de reprise après sinistre. Vous pouvez copier n'importe quel instantané accessible ayant l'état `completed`. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).

Prise en charge du chiffrement pour les instantanés

Les instantanés EBS prennent complètement en charge le chiffrement EBS.

- Les instantanés des volumes chiffrés sont chiffrés automatiquement.
- Les volumes que vous créez à partir d'instantanés chiffrés sont automatiquement chiffrés.
- Les volumes que vous créez à partir d'un instantané non chiffré que vous possédez ou auquel vous avez accès peuvent être chiffrés à la volée.
- Lorsque vous copiez un instantané non chiffré qui vous appartient, vous pouvez le chiffrer pendant le processus de copie.
- Lorsque vous copiez un instantané chiffré qui vous appartient ou auquel vous avez accès, vous pouvez le chiffrer à nouveau avec une autre clé pendant le processus de copie.
- Le premier instantané que vous faites d'un volume chiffré qui a été créé à partir d'un instantané non chiffré est toujours un instantané complet.
- Le premier instantané que vous faites d'un volume rechiffré, dont la CMK diffère de celle de l'instantané source, est toujours un instantané complet.

> **Note**
>
> Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Une documentation complète des scénarios possibles de chiffrement d'instantanés est fournie dans Création d'instantanés Amazon EBS (p. 1132) et dans Copie d'un instantané Amazon EBS (p. 1138).

Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Création d'instantanés Amazon EBS

Vous pouvez créer un instantané à un instant donné d'un volume EBS et utiliser celui-ci comme base pour les nouveaux volumes ou pour la sauvegarde des données. Si vous créez régulièrement des instantanés d'un volume, ces instantanés sont incrémentiels : le nouvel instantané enregistre seulement les blocs qui ont changé depuis le dernier instantané.

Les instantanés sont créés de façon asynchrone, ce qui signifie que l'instantané à un instant donné est créé immédiatement, mais qu'il conserve l'état `pending` jusqu'à ce qu'il soit terminé (une fois que tous

les blocs modifiés ont été transférés vers Amazon S3), ce qui peut prendre plusieurs heures pour les premiers instantanés volumineux ou pour les instantanés suivants dans lesquels de nombreux blocs ont été modifiés. Tant qu'il n'est pas finalisé, un instantané en cours de création n'est pas affecté par les lectures et écritures qui se produisent sur le volume.

Vous pouvez prendre un instantané d'un volume attaché en cours d'utilisation. Toutefois, les instantanés capturent uniquement les données qui ont été écrites dans votre volume Amazon EBS au moment de l'émission de la commande d'instantané. Cela peut exclure les données mises en cache par toutes les applications ou le système d'exploitation. Si vous pouvez interrompre une écriture de fichier sur le volume assez longtemps pour prendre un instantané, celui-ci devrait être complet. Toutefois, si vous ne pouvez pas interrompre toutes les écritures de fichier sur le volume, vous devez démonter le volume à partir de l'instance, émettre la commande d'instantané, puis remonter le volume afin de garantir un instantané cohérent et complet. Vous pouvez remonter et utiliser votre volume alors que le statut de l'instantané est `pending`.

Pour faciliter la gestion des instantanés, vous pouvez les baliser lors de leur création ou ajouter des balises ultérieurement. Par exemple, vous pouvez appliquer des balises décrivant le volume d'origine à partir duquel l'instantané a été créé, ou le nom du périphérique qui a été utilisé pour attacher le volume d'origine à une instance. Pour plus d'informations, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Chiffrement des instantanés

Les instantanés créés à partir de volumes chiffrés sont chiffrés automatiquement. Les volumes qui sont créés à partir d'instantanés chiffrés sont également chiffrés automatiquement. Les données de vos volumes chiffrés et les instantanés associés sont protégés qu'ils soient mobiles ou sédentaires. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).

Par défaut, seul le propriétaire des instantanés peut s'en servir pour créer des volumes. Toutefois, vous pouvez partager vos instantanés non chiffrés avec des comptes AWS spécifiques, ou même avec toute la communauté AWS en les rendant publics. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

Vous ne pouvez partager un instantané chiffré qu'avec des comptes AWS spécifiques. Pour permettre aux autres d'utiliser votre instantané chiffré partagé, vous devez également partager la clé CMK qui a été utilisée pour le chiffrer. Les utilisateurs ayant accès à votre instantané chiffré doivent créer leur propre copie personnelle de cette clé, puis utiliser cette copie. Vous pouvez également chiffrer à nouveau votre copie d'un instantané chiffré partagé avec une autre clé. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

> **Note**
>
> Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Instantanés multi-volumes

Vous pouvez créer des instantanés multi-volumes, qui sont des instantanés à un instant donné pour tous les volumes EBS attachés à une instance EC2. Vous pouvez également créer des stratégies de cycle de vie pour automatiser la création et la conservation des instantanés multi-volumes. Pour plus d'informations, consultez Automatisation du cycle de vie des instantanés Amazon EBS (p. 1169).

Une fois les instantanés créés, chaque instantané est traité en tant qu'instantané individuel. Vous pouvez effectuer toutes les opérations d'instantanés, telles que la restauration, la suppression et la copie entre régions ou comptes, tout comme vous le feriez avec un instantané d'un seul volume. Vous pouvez également baliser vos instantanés multi-volumes comme vous le feriez pour un instantané d'un seul volume. Nous vous recommandons de baliser vos instantanés multi-volumes afin de les gérer collectivement lors de leur restauration, copie ou conservation.

Les instantanés multi-volumes en lien avec la panne sont généralement restaurés sous la forme d'un ensemble. Il est utile d'identifier les instantanés qui figurent dans un ensemble en lien avec la panne en balisant cet ensemble avec l'ID d'instance, le nom ou d'autres détails pertinents. Vous pouvez également choisir de copier les balises à partir du volume source dans les instantanés correspondants. Ceci vous aide à définir les métadonnées des instantanés, telles que les stratégies d'accès, les informations de pièce jointe et l'allocation des coûts, pour qu'elles correspondent au volume source.

Une fois que vos instantanés sont créés, ils apparaissent dans votre console EC2 créée à l'instant donné exact. Les instantanés sont gérés collectivement. Par conséquent, si un instantané pour le volume défini échoue, tous les autres instantanés affichent un statut d'erreur.

Amazon Data Lifecycle Manager

Vous pouvez créer, conserver et supprimer les instantanés manuellement, ou vous pouvez utiliser Amazon Data Lifecycle Manager pour les gérer à votre place. Pour de plus amples informations, consultez Automatisation des instantanés (p. 1169).

Considérations

Les considérations suivantes s'appliquent à la création des instantanés :

- Lors de la création d'un instantané pour un volume EBS qui sert de périphérique racine, vous devez arrêter l'instance avant de prendre l'instantané.
- Vous ne pouvez pas créer d'instantanés à partir d'instances pour lesquelles la mise en veille prolongée est activée.
- Vous ne pouvez pas créer d'instantanés à partir d'instances mises en veille prolongée.
- Vous pouvez prendre un instantané d'un volume pendant qu'un instantané précédent du même volume a le statut `pending`, mais le fait d'avoir plusieurs instantanés `pending` d'un même volume peut réduire les performances du volume tant que les instantanés ne sont pas terminés.
- Il existe une limite de cinq instantanés `pending` pour un seul volume `gp2`, `io1`, `io2` ou magnétique et d'un instantané `pending` pour un seul volume `st1` ou `sc1`. Si vous recevez un message d'erreur `ConcurrentSnapshotLimitExceeded` lorsque vous cherchez à créer plusieurs instantanés simultanés du même volume, attendez qu'un ou plusieurs instantanés `pending` soient terminés avant de créer un autre instantané de ce volume.
- Lorsqu'un instantané est créé à partir d'un volume avec un code produit AWS Marketplace, le code produit est propagé à l'instantané.

Création d'un instantané

Utilisez la procédure suivante pour créer un instantané à partir du volume spécifié.

Pour créer un instantané avec la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instantanés sous Elastic Block Store.
3. Choisissez Create Snapshot (Créer un instantané).
4. Pour Sélectionner le type de ressource, choisissez Volume.
5. Pour Volume, sélectionnez le volume.
6. (Facultatif) Entrez une description pour l'instantané.
7. (Facultatif) Choisissez Ajouter une balise pour ajouter des balises à votre instantané. Pour chaque balise, indiquez une clé de balise et une valeur de balise.
8. Choisissez Create Snapshot (Créer un instantané).

Pour créer un instantané à partir de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-snapshot (AWS CLI)
- New-EC2Snapshot (Outils AWS pour Windows PowerShell)

Création d'un instantané multi-volume

Utilisez la procédure suivante pour créer un instantané à partir des volumes d'une instance.

Pour créer des instantanés multi-volumes à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instantanés sous Elastic Block Store.
3. Choisissez Create Snapshot (Créer un instantané).
4. Pour Sélectionner le type de ressource, choisissez Instance.
5. Sélectionnez l'ID de l'instance pour laquelle vous voulez créer des sauvegardes simultanées pour tous les volumes EBS attachés. Les instantanés multi-volumes prennent en charge jusqu'à 40 volumes EBS par instance.
6. (Facultatif) Définissez Exclude root volume (Exclure le volume racine).
7. (Facultatif) Définissez l'indicateur Copy tags from volume (Copier les balises à partir du volume) de manière à copier automatiquement les balises à partir du volume source dans les instantanés correspondants. Ceci définit les métadonnées des instantanés (telles que les stratégies d'accès, les informations de pièce jointe et l'allocation des coûts) pour qu'elles correspondent au volume source.
8. (Facultatif) Choisissez Add tag (Ajouter une balise). Pour chaque balise, indiquez une clé de balise et une valeur de balise.
9. Choisissez Create Snapshot (Créer un instantané).

   Au cours de la création des instantanés, les instantanés sont gérés ensemble. Si l'un des instantanés figurant dans le volume défini échoue, les autres instantanés obtiennent le statut d'erreur pour le volume défini. Vous pouvez surveiller l'avancement de vos instantanés à l'aide de CloudWatch Events. Une fois que le processus de création d'instantanés s'est terminé, CloudWatch génère un événement qui contient le statut et tous les détails pertinents des instantanés pour l'instance affectée.

Pour créer des instantanés multi-volumes à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-snapshots (AWS CLI)
- New-EC2SnapshotBatch (Outils AWS pour Windows PowerShell)

Utilisation des instantanés EBS

Vous pouvez copier des instantanés, les partager et créer des volumes à partir d'instantanés. Pour plus d'informations, consultez les ressources suivantes :

- Copie d'un instantané Amazon EBS (p. 1138)
- Partage d'un instantané Amazon EBS (p. 1143)

- Création d'un volume à partir d'un instantané (p. 1108)

Suppression d'un instantané Amazon EBS

Une fois que vous n'avez plus besoin d'un instantané Amazon EBS d'un volume, vous pouvez le supprimer. La suppression d'un instantané n'a aucun effet sur le volume. La suppression d'un volume n'a aucun effet sur les instantanés créés à partir de celui-ci.

Suppression d'instantané incrémentielle

Si vous effectuez régulièrement des instantanés d'un volume, les instantanés sont incrémentiels. Cela signifie que seuls les blocs qui ont changé sur l'appareil depuis le dernier instantané sont enregistrés dans le nouvel instantané. Bien que les instantanés soient enregistrés de manière incrémentielle, le processus de suppression de l'instantané prévoit que vous ayez uniquement besoin de conserver l'instantané le plus récent pour créer le volume. Les données qui étaient présentes sur un volume, stocké dans un instantané ou une série d'instantanés précédents, qui sont supprimées ultérieurement de ce volume sont toujours considérées comme des données uniques d'instantanés antérieurs. Ces données uniques ne sont pas supprimées de la séquence d'instantanés sauf si tous les instantanés qui référencent les données uniques sont supprimés.

Lorsque vous supprimez un instantané, seules les données référencées exclusivement par cet instantané sont supprimées. Les données uniques ne seront pas supprimées sauf si tous les instantanés qui référencent ces données sont supprimés. La suppression d'instantanés précédents d'un volume n'a aucune répercussion sur votre capacité à créer des volumes à partir d'instantanés ultérieurs de ce même volume.

La suppression d'un instantané ne réduit pas les coûts de stockage des données de votre organisation. D'autres instantanés peuvent faire référence aux données de cet instantané et les données référencées sont toujours conservées. Si vous supprimez un instantané contenant des données utilisées par un instantané ultérieur, les coûts associés aux données référencées sont alloués à l'instantané ultérieur. Pour plus d'informations sur la façon dont les instantanés stockent les données, veuillez consulter Fonctionnement des instantanés incrémentiels (p. 1130) et l'exemple ci-dessous.

Dans le graphique suivant, Volume 1 est affiché à trois moments différents. Un instantané a capturé chacun des deux premiers états, et dans le troisième, un instantané a été supprimé.

- A l'état 1, le volume contient 10 Gio de données. Comme Snap A est le premier instantané pris du volume, la totalité des 10 Gio de données doit être copiée.
- A l'état 2, le volume contient toujours 10 Gio de données, mais 4 Gio de données ont été modifiés. Snap B a besoin d'une copie et ne stocke que les 4 Gio qui ont été modifiés après que l'instantané Snap A ait été pris. Les autres 6 Gio de données non modifiées qui avaient déjà été copiés et stockés dans Snap A, sont référencés par Snap B plutôt que copiés (à nouveau). Ceci est indiqué par la flèche en pointillé.
- À l'état 3, le volume n'a pas changé depuis l'état 2, mais l'instantané Snapshot A a été supprimé. Les 6 Gio de données stockés dans Snapshot A qui étaient référencées par Snapshot B ont été transférées vers Snapshot B, comme indiqué par la flèche pleine. Par conséquent, vous êtes encore facturé pour le stockage de 10 Gio de données : 6 Gio de données non modifiées conservées de Snap A, et 4 Gio de données modifiées de Snap B.

Suppression d'un instantané avec certaines de ses données référencées par un autre instantané

Considérations

Les considérations suivantes s'appliquent à la suppression des instantanés :

- Vous ne pouvez pas supprimer un instantané de l'appareil racine d'un volume EBS utilisé par une AMI enregistrée. Vous devez commencer par annuler l'inscription de l'AMI avant de pouvoir supprimer l'instantané. Pour plus d'informations, consultez Annulation de l'inscription de votre AMI Linux (p. 182).
- Vous ne pouvez pas supprimer un instantané géré par le service AWS Backup avec Amazon EC2. Utilisez plutôt AWS Backup pour supprimer les points de récupération correspondants dans le coffre-fort de sauvegarde.
- Vous pouvez créer, conserver et supprimer les instantanés manuellement, ou vous pouvez utiliser Amazon Data Lifecycle Manager pour les gérer à votre place. Pour de plus amples informations, consultez Automatisation des instantanés (p. 1169).
- Même si vous pouvez supprimer un instantané qui est toujours en cours, l'instantané doit être terminé avant que la suppression prenne effet. Cela pourrait prendre beaucoup de temps. En outre, si vous avez atteint votre limite d'instantanés simultanés et que vous tenez de prendre un instantané supplémentaire, vous pouvez obtenir une erreur `ConcurrentSnapshotLimitExceeded`. Pour de plus amples informations, veuillez consulter Quotas du service pour Amazon EBS dans la Référence générale d'Amazon Web Services.

Suppression d'un instantané

Procédez comme indiqué ci-dessous pour supprimer un instantané.

Pour supprimer un instantané avec la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instantanés dans le panneau de navigation.
3. Sélectionnez un instantané, puis choisissez Supprimer dans la liste Actions.
4. Sélectionnez Oui, supprimer.

Pour supprimer un instantané à partir de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- delete-snapshot (AWS CLI)
- Remove-EC2Snapshot (Outils AWS pour Windows PowerShell)

Suppression d’un instantané multi-volume

Pour supprimer des instantanés multi-volumes, récupérez tous les instantanés de votre groupe multi-volume en utilisant la balise que vous avez appliquée au groupe lorsque vous avez créé les instantanés. Ensuite, supprimez individuellement les instantanés. Vous ne vous verrez pas empêché de supprimer des instantanés individuels dans le groupe d'instantanés multi-volumes.

Copie d'un instantané Amazon EBS

Amazon EBS vous permet de créer des instantanés à un instant donné de volumes que nous stockons pour vous dans Amazon S3. Une fois que vous avez créé un instantané et que sa copie sur Amazon S3 est terminée (lorsque le statut de l'instantané est `completed`), vous pouvez le copier d'une région AWS vers une autre, ou au sein de la même région. Le chiffrement côté serveur Amazon S3 (AES 256 bits) protège les données en transit d'un instantané pendant une opération de copie. La copie de l'instantané reçoit un ID différent de celui de l'instantané d'origine.

Pour copier des instantanés multi-volumes vers une autre région AWS, récupérez les instantanés en utilisant la balise que vous avez appliquée au groupe d'instantanés multi-volumes lorsque vous l'avez créé. Ensuite, copiez individuellement les instantanés vers une autre région.

Pour plus d'informations sur la copie d'un instantané Amazon RDS, consultez Copie d'un instantané de base de données dans le Amazon RDS Guide de l'utilisateur.

Si vous souhaitez qu'un autre compte puisse copier votre instantané, vous devez modifier les autorisations d'instantané de façon à permettre l'accès à ce compte ou rendre l'instantané public afin que tous les comptes AWS puissent le copier. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

Pour de plus amples informations sur la tarification de la copie d'instantanés entre régions AWS et comptes, veuillez consulter Tarification Amazon EBS. Notez que les opérations de copie d'instantané au sein d'un même compte et d'une même région ne copient aucune donnée réelle et sont donc gratuites aussi longtemps que le statut de chiffrement de la copie d'instantané ne change pas.

> Note
>
> Si vous copiez un instantané dans une nouvelle région, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.
>
> Note
>
> Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Cas d'utilisation

- Expansion géographique : lancez vos applications dans une nouvelle région AWS.
- Migration : transférez une application vers une nouvelle région afin de permettre une disponibilité accrue et de réduire les coûts.
- Reprise après sinistre : sauvegardez vos données et fichiers journaux dans différents emplacements géographiques à intervalles réguliers. En cas de sinistre, vous pouvez restaurer vos applications à l'aide des sauvegardes ponctuelles stockées dans la région secondaire. Cela permet de limiter les pertes de données et la durée de récupération.
- Chiffrement : chiffrez un instantané qui n'a pas encore été chiffré, modifiez la clé de chiffrement de l'instantané ou créez une copie vous appartenant afin de restaurer un volume à partir de celle-ci (pour les instantanés chiffrés qui ont été partagés avec vous).
- Rétention des données et exigences en matière d'audit : copiez vos instantanés EBS chiffrés d'un compte AWS vers un autre pour conserver les journaux de données ou d'autres fichiers aux fins d'audit ou de rétention des données. L'utilisation d'un autre compte permet d'éviter les suppressions accidentelles d'instantanés et vous protège si votre compte AWS principal est compromis.

Prérequis

- Vous pouvez copier les instantanés accessibles ayant un statut `completed`, y compris les instantanés partagés et ceux que vous avez créés.
- Vous pouvez copier les instantanés AWS Marketplace, VM Import/Export et AWS Storage Gateway, mais vous devez vérifier au préalable que l'instantané est pris en charge dans la région cible.

Restrictions

- Chaque compte peut avoir jusqu'à vingt demandes de copie d'instantané simultanées vers une même région de destination.
- Les balises définies par l'utilisateur ne sont pas copiées de l'instantané source vers le nouvel instantané. Vous pouvez ajouter des balises définies par l'utilisateur pendant ou après l'opération de copie. Pour plus d'informations, consultez Balisage de vos ressources Amazon EC2 (p. 1309).
- Les instantanés créés par l'action `CopySnapshot` ont un ID de volume arbitraire qui ne doit être utilisé en aucun cas.

Copie d'instantané incrémentielle

L'aspect incrémentiel d'une copie d'instantané est déterminé par la dernière copie d'instantané effectuée. Lorsque vous copiez un instantané entre des régions ou des comptes, la copie est une copie incrémentielle si les conditions suivantes sont remplies :

- L'instantané a été préalablement copié dans la région ou le compte de destination.
- La dernière copie d'instantané existe toujours dans la région ou le compte de destination.
- Toutes les copies de l'instantané dans la région ou dans le compte de destination sont soit non chiffrées, soit chiffrées avec la même clé CMK.

Si la dernière copie d'instantané a été supprimée, la copie suivante est une copie complète, non une copie incrémentielle. Si une copie est toujours en attente lorsque vous démarrez une autre copie, la deuxième copie ne démarre qu'une fois la première copie terminée.

Nous vous recommandons de baliser vos instantanés avec l'ID de volume et l'heure de création afin de pouvoir conserver une trace de la dernière copie d'instantané d'un volume dans la région ou dans le compte de destination.

Pour savoir si vos copies d'instantané sont incrémentielles, consultez l'événement CloudWatch copySnapshot (p. 1259).

Chiffrement et copie d'instantanés

Lorsque vous copiez un instantané, vous pouvez chiffrer la copie ou vous pouvez spécifier une clé CMK différente de l'originale et l'instantané copié utilise cette nouvelle clé CMK. Toutefois, le changement de l'état de chiffrement d'un instantané au cours d'une opération de copie entraîne toujours une copie complète (non incrémentielle), ce qui peut impliquer des frais plus importants de transfert et de stockage de données.

Pour copier un instantané chiffré partagé à partir d'un autre compte AWS, vous devez disposer des autorisations d'utilisation de cet instantané et vous devez être autorisé à utiliser la clé principale client (CMK) qui a été utilisée pour chiffrer l'instantané d'origine. En cas d'utilisation d'un instantané chiffré qui a été partagé avec vous, nous vous recommandons de chiffrer à nouveau l'instantané en le copiant à l'aide d'un clé CMK en votre possession. Cela vous protège si la clé CMK d'origine est compromise ou si le propriétaire la révoque, ce qui pourrait vous faire perdre l'accès aux volumes chiffrés que vous avez créés en utilisant l'instantané. Pour plus d'informations, consultez Partage d'un instantané Amazon EBS (p. 1143).

Vous appliquez le chiffrement aux instantanés EBS en définissant le paramètre `Encrypted` avec la valeur `true`. (Le paramètre `Encrypted` est facultatif si le chiffrement par défaut (p. 1196) est activé).

Vous pouvez aussi utiliser `KmsKeyId` pour spécifier une clé personnalisée pour chiffrer la copie de l'instantané. (Le paramètre `Encrypted` doit aussi être défini avec la valeur `true`, même si le chiffrement par défaut est activé.) Si `KmsKeyId` n'est pas spécifié, la clé utilisée pour le chiffrement dépend de l'état de chiffrement de l'instantané source et de son propriétaire.

Le tableau suivant décrit le résultat de chiffrement pour chaque combinaison possible de paramètres.

Résultats du chiffrement : copie d'un instantané

Le paramètre **`Encrypted`** est-il défini ?	Le chiffrement est-il défini par défaut ?	Instantané source	Par défaut (aucun KmsKeyId spécifié)	Personnalisé (KmsKeyId spécifié)
Non	Non	Instantané non chiffré que vous possédez	Non chiffré	S/O
Non	Non	Instantané chiffré que vous possédez	Chiffré par la même clé	
Non	Non	Instantané non chiffré qui est partagé avec vous	Non chiffré	
Non	Non	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut*	
Oui	Non	Instantané non chiffré que vous possédez	Chiffré par la clé CMK par défaut	Chiffré par une clé CMK spécifiée**
Oui	Non	Instantané chiffré que vous possédez	Chiffré par la même clé	
Oui	Non	Instantané non chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	

Le paramètre **Encrypted** est-il défini ?	Le chiffrement est-il défini par défaut ?	Instantané source	Par défaut (aucun KmsKeyId spécifié)	Personnalisé (KmsKeyId spécifié)
Oui	Non	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Non	Oui	Instantané non chiffré que vous possédez	Chiffré par la clé CMK par défaut	S/O
Non	Oui	Instantané chiffré que vous possédez	Chiffré par la même clé	
Non	Oui	Instantané non chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Non	Oui	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Oui	Oui	Instantané non chiffré que vous possédez	Chiffré par la clé CMK par défaut	Chiffré par une clé CMK spécifiée
Oui	Oui	Instantané chiffré que vous possédez	Chiffré par la même clé	
Oui	Oui	Instantané non chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Oui	Oui	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	

*Clé CMK par défaut utilisée pour le chiffrement EBS pour le compte et la région AWS. Il s'agit par défaut d'une clé CMK unique gérée par AWS pour EBS, ou vous pouvez spécifier une clé CMK gérée par le client. Pour plus d'informations, consultez Clé par défaut pour le chiffrement EBS (p. 1196).

**Clé CMK gérée par le client, spécifiée pour l'action de copie. Cette clé CMK est utilisée à la place de la clé CMK par défaut pour le compte et la région AWS.

Copie d'un instantané

Utilisez la procédure suivante pour copier un instantané à l'aide de la console Amazon EC2.

Pour copier un instantané à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Snapshots.
3. Sélectionnez l'instantané à copier, puis choisissez Copier dans la liste Actions.
4. Dans la boîte de dialogue Copier un instantané, mettez à jour les informations suivantes si nécessaire :
   - Région de destination : sélectionnez la région vers laquelle vous souhaitez transférer la copie de l'instantané.
   - Description : par défaut, la description inclut des informations sur l'instantané source afin que vous puissiez identifier une copie à partir de l'original. Vous pouvez modifier cette description si nécessaire.

- Chiffrement : si l'instantané source n'est pas chiffré, vous pouvez choisir de chiffrer la copie. Si vous avez activé le chiffrement par défaut (p. 1196), l'option Encryption (Chiffrement) est activée et ne peut pas être désactivée depuis la console de l'instantané. Si l'option Encryption (Chiffrement) est activée, vous pouvez choisir d'effectuer le chiffrement sous une clé CMK gérée par le client en sélectionnant celle-ci dans le champ, comme décrit ci-dessous.

  Vous ne pouvez pas supprimer le chiffrement d'un instantané chiffré.

  > **Note**
  >
  > Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

- Clé principale : il s'agit de la clé principale du client (CMK) qui sera utilisée pour chiffrer cet instantané. La clé par défaut pour votre compte est affichée au début. Toutefois, vous pouvez sélectionner l'une des clés principales dans votre compte ou saisir/coller l'ARN d'une clé issue d'un autre compte. Vous pouvez créer de nouvelles clés principales de chiffrement dans la console IAM https://console.aws.amazon.com/iam/.

5. Choisissez Copy.
6. Dans la boîte de dialogue de confirmation Copier un instantané, choisissez Instantanés afin d'accéder à la page Instantanés dans la région spécifiée, ou choisissez Fermer.

   Pour consulter l'évolution du processus de copie, basculez vers la région de destination, puis actualisez la page Instantanés. Les copies en cours sont répertoriées en haut de la page.

Pour déterminer la cause d'une erreur

Si vous essayez de copier un instantané chiffré sans disposer des autorisations d'utilisation de la clé de chiffrement, l'opération échoue silencieusement. L'état d'erreur ne s'affiche pas sur la console tant que vous n'avez pas actualisé la page. Vous pouvez également vérifier l'état de l'instantané à partir de la ligne de commande, comme dans l'exemple suivant.

```
aws ec2 describe-snapshots --snapshot-id snap-0123abcd
```

Si la copie échoue en raison d'autorisations insuffisantes d'utilisation de la clé, vous verrez le message : "StateMessage": "Given key ID is not accessible".

Lors de la copie d'un instantané chiffré, vous devez disposer des autorisations `DescribeKey` sur la clé CMK par défaut. Le refus explicite de ces autorisations entraîne l'échec de la copie. Pour plus d'informations sur la gestion des clés CMK, consultez Contrôle de l'accès aux clés principales du client.

Pour copier un instantané à partir de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- copy-snapshot (AWS CLI)
- Copy-EC2Snapshot (Outils AWS pour Windows PowerShell)

Affichage d'informations sur un instantané Amazon EBS

Vous pouvez afficher des informations détaillées sur vos instantanés.

Pour afficher des informations sur un instantané à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Choisissez Instantanés dans le panneau de navigation.
3. Pour réduire la liste, choisissez une option dans la liste Filtrer. Par exemple, pour afficher uniquement vos instantanés, choisissez M'appartenant. Vous pouvez également filtrer vos instantanés à l'aide de balises et d'attributs d'instantané. Choisissez la barre de recherche pour afficher les balises et attributs disponibles.
4. Pour afficher des informations supplémentaires sur un instantané, sélectionnez-le.

Pour afficher des informations sur un instantané à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- describe-snapshots (AWS CLI)
- Get-EC2Snapshot (Outils AWS pour Windows PowerShell)

Example Exemple : filtre basé sur des balises

La commande suivante décrit les instantanés avec la balise Stack=production.

```
aws ec2 describe-snapshots --filters Name=tag:Stack,Values=production
```

Example Exemple : filtre basé sur le volume

La commande suivante décrit les instantanés créés à partir du volume spécifié.

```
aws ec2 describe-snapshots --filters Name=volume-id,Values=vol-049df61146c4d7901
```

Example Exemple : filtrer basé sur l'ancienneté des instantanés

Avec l'AWS CLI, vous pouvez utiliser JMesPath pour filtrer les résultats à l'aide d'expressions. Par exemple, la commande suivante affiche les ID de tous les instantanés créés par votre compte AWS (représenté par `123456789012`) avant la date spécifiée (représentée par `2020-03-31`). Si vous ne spécifiez pas le propriétaire, les résultats incluent tous les instantanés publics.

```
aws ec2 describe-snapshots --filters Name=owner-id,Values=123456789012 --query "Snapshots[?
(StartTime<=`2020-03-31`)].[SnapshotId]" --output text
```

La commande suivante affiche les ID de tous les instantanés créés dans la plage de dates spécifiée.

```
aws ec2 describe-snapshots --filters Name=owner-id,Values=123456789012 --query "Snapshots[?
(StartTime>=`2019-01-01`) && (StartTime<=`2019-12-31`)].[SnapshotId]" --output text
```

Partage d'un instantané Amazon EBS

En modifiant les autorisations d'un instantané, vous pouvez le partager avec les comptes AWS que vous spécifiez. Les utilisateurs qui bénéficient de votre autorisation peuvent utiliser les instantanés que vous partagez comme base de création de leurs propres volumes EBS, pendant que votre instantané d'origine reste inchangé.

Vous pouvez choisir de rendre vos instantanés non chiffrés publiquement disponibles pour tous les utilisateurs d'AWS. Vous ne pouvez pas rendre vos instantanés chiffrés publiquement disponibles.

Lorsque vous partagez un instantané chiffré, vous devez également partager la clé CMK gérée par le client qui a servi à chiffrer l'instantané. Vous pouvez appliquer des autorisations entre comptes à une clé CMK gérée par le client lors de sa création ou ultérieurement.

> **Important**
>
> Lorsque vous partagez un instantané, vous autorisez d'autres personnes à accéder à toutes les données de l'instantané. Partagez vos instantanés uniquement avec les personnes avec lesquelles vous souhaitez partager toutes vos données d'instantané.

Considérations

Les considérations suivantes s'appliquent au partage des instantanés :

- Les instantanés sont limités à la région dans laquelle ils ont été créés. Pour partager un instantané avec une autre région, copiez l'instantané dans cette région. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).
- AWS vous empêche de partager les instantanés qui ont été chiffrés avec votre clé CMK par défaut. Les instantanés que vous avez l'intention de partager doivent donc être chiffrés avec une clé CMK gérée par le client. Pour plus d'informations, consultez Création de clés dans le AWS Key Management Service Developer Guide.
- Les utilisateurs de votre clé CMK partagée qui accèdent aux instantanés chiffrés doivent recevoir les autorisations permettant d'exécuter les actions suivantes sur la clé : `kms:DescribeKey`, `kms:CreateGrant`, `GenerateDataKey` et `kms:ReEncrypt`. Pour plus d'informations, consultez Contrôle de l'accès aux clés principales du client dans le AWS Key Management Service Developer Guide.

Partage d'un instantané non chiffré à l'aide de la console

Pour partager un instantané à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instantanés dans le panneau de navigation.
3. Sélectionnez l'instantané, puis choisissez Actions, Modifier des autorisations.
4. Rendez l'instantané public ou partagez-le avec des comptes AWS spécifiques comme suit :
   - Pour rendre l'instantané public, choisissez Public.

     Cette option n'est pas valide pour les instantanés chiffrés ou les instantanés avec un code produit AWS Marketplace.
   - Pour partager l'instantané avec un ou plusieurs comptes AWS, choisissez Privé, entrez l'ID du compte AWS (sans traits d'union) dans Numéro de compte AWS, puis choisissez Ajouter autorisation. Répétez cette opération pour les éventuels comptes AWS supplémentaires.
5. Choisissez Save.

Pour utiliser un instantané non chiffré qui a été partagé en privé avec vous

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instantanés dans le panneau de navigation.
3. Choisissez le filtre Instantanés privés.
4. Localisez l'instantané par son ID ou sa description. Vous pouvez utiliser cet instantané comme vous le feriez avec n'importe quel autre. Par exemple, vous pouvez créer un volume à partir de l'instantané ou le copier dans une autre région.

Partage d'un instantané chiffré à l'aide de la console

Pour partager un instantané chiffré en utilisant la console

1. Ouvrez la console AWS KMS à l'adresse https://console.aws.amazon.com/kms.
2. Pour changer de région AWS, utilisez le sélecteur de région dans l'angle supérieur droit de la page.
3. Choisissez Customer managed keys (Clés gérées par le client) dans le volet de navigation.
4. Dans la colonne Alias choisissez l'alias (lien texte) de la clé gérée par le client que vous avez utilisée pour chiffrer l'instantané. Les détails de la clé s'ouvrent dans une nouvelle page.
5. Dans la section Key policy (Stratégie de clé) s'affiche soit la vue de la stratégie soit la vue par défaut. La vue de la stratégie affiche le document de la stratégie de clé. La vue par défaut affiche les sections Key administrators (Administrateurs de clé), Key deletion (Suppression de clé), Key Use (Utilisation de clé) et Other AWS accounts (Autres comptes AWS). L'affichage par défaut s'affiche si vous avez créé la stratégie dans la console et que vous ne l'avez pas personnalisée. Si l'affichage par défaut n'est pas disponible, vous devez modifier manuellement la stratégie dans l'affichage de stratégie. Pour de plus amples informations, veuillez consulter Affichage d'une stratégie de clé (console) dans le AWS Key Management Service Developer Guide.

   Utilisez la vue de la stratégie ou la vue par défaut, en fonction de l'affichage auquel vous pouvez accéder, pour ajouter un ou plusieurs ID de compte AWS à la stratégie, comme suit :

   - (Vue de la stratégie) Choisissez Edit (Modifier). Ajoutez un ou plusieurs ID de compte AWS aux instructions suivantes : `"Allow use of the key"` et `"Allow attachment of persistent resources"`. Sélectionnez Save Changes (Enregistrer les modifications). Dans l'exemple suivant, l'ID de compte AWS `444455556666` est ajouté à la stratégie.

```
{
  "Sid": "Allow use of the key",
  "Effect": "Allow",
  "Principal": {"AWS": [
    "arn:aws:iam::111122223333:user/CMKUser",
    "arn:aws:iam::444455556666:root"
  ]},
  "Action": [
    "kms:Encrypt",
    "kms:Decrypt",
    "kms:ReEncrypt*",
    "kms:GenerateDataKey*",
    "kms:DescribeKey"
  ],
  "Resource": "*"
},
{
  "Sid": "Allow attachment of persistent resources",
  "Effect": "Allow",
  "Principal": {"AWS": [
    "arn:aws:iam::111122223333:user/CMKUser",
    "arn:aws:iam::444455556666:root"
  ]},
  "Action": [
    "kms:CreateGrant",
    "kms:ListGrants",
    "kms:RevokeGrant"
  ],
  "Resource": "*",
  "Condition": {"Bool": {"kms:GrantIsForAWSResource": true}}
}
```

   - (Vue par défaut) Faites défiler la page jusqu'à Other AWS accounts (Autres comptes AWS). Choisissez Add other AWS accounts (Ajouter d'autres comptes AWS) et entrez l'ID du compte AWS à l'invite. Pour ajouter un autre compte, choisissez Add another AWS account (Ajouter un

autre compte) et entre l'ID du compte AWS. Une fois que vous avez ajouté tous les comptes AWS, choisissez Enregistrer les modifications.

6. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
7. Choisissez Instantanés dans le panneau de navigation.
8. Sélectionnez l'instantané, puis choisissez Actions, Modifier des autorisations.
9. Pour chaque compte AWS, entrez l'ID du compte AWS dans Numéro de compte AWS, puis choisissez Ajouter autorisation. Une fois que vous avez ajouté tous les comptes AWS, choisissez Enregistrer.

Pour utiliser un instantané chiffré qui a été partagé avec vous

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Instantanés dans le panneau de navigation.
3. Choisissez le filtre Instantanés privés. Vous pouvez éventuellement ajouter le filtre Chiffré.
4. Localisez l'instantané par son ID ou sa description.
5. Sélectionnez l'instantané, puis choisissez Actions, Copier.
6. (Facultatif) Sélectionnez une région de destination.
7. La copie de l'instantané est chiffrée avec la clé affichée dans le champ Clé principale. Par défaut, la clé sélectionnée est la clé CMK par défaut de votre compte. Pour sélectionner une clé CMK gérée par le client, cliquez dans le champ de saisie pour voir la liste des clés disponibles.
8. Choisissez Copy.

Partage d'un instantané à l'aide de la ligne de commande

Les autorisations pour un instantané sont spécifiées à l'aide de l'attribut `createVolumePermission` de l'instantané. Pour qu'un instantané devienne public, définissez le groupe sur `all`. Pour partager un instantané avec un compte AWS spécifique, définissez l'utilisateur sur l'ID du compte AWS.

Pour modifier les autorisations d'instantané à l'aide de la ligne de commande

Utilisez l'une des commandes suivantes :

- modify-snapshot-attribute (AWS CLI)
- Edit-EC2SnapshotAttribute (Outils AWS pour Windows PowerShell)

Pour afficher les autorisations d'instantané à l'aide de la ligne de commande

Utilisez l'une des commandes suivantes :

- describe-snapshot-attribute (AWS CLI)
- Get-EC2SnapshotAttribute (Outils AWS pour Windows PowerShell)

Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

Détermination de l'utilisation des instantanés partagés

Vous pouvez utiliser AWS CloudTrail pour vérifier si un instantané que vous avez partagé avec d'autres utilisateurs est copié ou utilisé pour créer un volume. Les événements suivants sont connectés dans CloudTrail :

- SharedSnapshotCopyInitiated — Un instantané partagé est en cours de copie.

- SharedSnapshotVolumeCreated — Un instantané partagé est utilisé pour créer un volume.

Pour plus d'informations sur l'utilisation de CloudTrail, consultez Journalisation des appels d'API Amazon EC2 et Amazon EBS avec AWS CloudTrail (p. 804).

Utilisation d'API directes EBS pour accéder au contenu d'un instantané EBS

Vous pouvez utiliser les API Amazon Elastic Block Store (Amazon EBS) directes pour créer des instantanés EBS, écrire des données directement sur vos instantanés, lire des données sur vos instantanés et identifier les différences ou les modifications entre deux instantanés. Si vous êtes un fournisseur indépendant de logiciel (FIL) qui offre des services de sauvegarde pour Amazon EBS, il est plus rentable et efficace d'utiliser le API directes EBS pour de suivre les modifications incrémentielles sur vos volumes EBS via des instantanés EBS à l'aide des instantanés. Pour ce faire, vous n'avez pas besoin de créer des volumes à partir d'instantanés, puis d'utiliser des instances Amazon Elastic Compute Cloud (Amazon EC2) pour comparer leurs différences.

Vous pouvez créer des instantanés incrémentiels directement à partir de données locales dans des volumes EBS et dans le cloud afin de les utiliser pour une reprise après sinistre rapide. Avec la possibilité d'écrire et de lire des instantanés, vous pouvez écrire vos données locales dans un instantané EBS lors d'un sinistre. Ensuite, après la restauration, vous pouvez le restaurer sur AWS ou sur site à partir de l'instantané. Vous n'avez plus besoin de créer et de gérer des mécanismes complexes pour copier des données depuis et vers Amazon EBS.

Le présent guide de l'utilisateur fournit une vue d'ensemble des éléments qui composent les API directes EBS, ainsi que des exemples d'utilisation efficace de ces éléments. Pour de plus amples informations sur les actions, les types de données, les paramètres et les erreurs des API, veuillez consulter la référence des API directes EBS. Pour de plus amples informations sur les régions, les points de terminaison et les quotas de service AWS pris en charge pour les API directes EBS, veuillez consulter Points de terminaison et quotas Amazon EBS dans le document AWS General Reference.

Sommaire

- Présentation des API directes EBS (p. 1147)
- Autorisations pour les utilisateurs IAM (p. 1150)
- Utilisation du chiffrement (p. 1154)
- Utilisation de la signature de la version 4 (p. 1155)
- Utilisation des sommes de contrôle (p. 1155)
- Utilisation des API directes EBS à l'aide de l'API ou des kits AWS SDK (p. 1156)
- Utilisation de API directes EBS à l'aide de la ligne de commande (p. 1160)
- Optimisation des performances (p. 1163)
- Questions fréquentes (FAQ) (p. 1164)
- Journalisation des appels d'API pour le API directes EBS avec AWS CloudTrail (p. 1165)
- API directes EBS et points de terminaison de VPC d'interface (p. 1167)
- Idempotence pour l'API StartSnapshot (p. 1168)

Présentation des API directes EBS

Vous devez bien comprendre les éléments clés suivants avant d'utiliser les API directes EBS.

Tarification

Le prix que vous payez pour utiliser le API directes EBS dépend des demandes que vous faites. Pour de plus amples informations, veuillez consulter Tarification Amazon EBS.

Instantanés

Les instantanés représentent le principal moyen de sauvegarde des données de vos volumes EBS. Avec le API directes EBS, vous pouvez également sauvegarder des données de vos disques locaux vers des instantanés. Afin d'économiser les frais de stockage, les instantanés successifs sont incrémentiels ; ils contiennent uniquement les données du volume ayant changé depuis l'instantané précédent. Pour de plus amples informations, veuillez consulter Instantanés Amazon EBS (p. 1129).

> Note
>
> Les instantanés publics ne sont pas pris en charge par les API directes EBS.

Blocs

Un bloc est un fragment de données au sein d'un instantané. Chaque instantané peut contenir des milliers de blocs. Tous les blocs d'un instantané sont de taille fixe.

Index de bloc

Un index de bloc correspond à la position de décalage d'un bloc dans un instantané. Il est utilisé pour identifier le bloc. Multipliez la valeur blockIndex par la valeur blockSize (blockIndex * blockSize) pour identifier le décalage logique des données dans le volume logique.

Jetons de bloc

Un jeton de bloc est le hachage d'identification d'un bloc dans un instantané. Il est utilisé pour localiser les données de bloc. Les jetons de bloc renvoyés par les API directes EBS sont temporaires. Ils changent sur l'horodatage d'expiration spécifié pour eux, ou si vous exécutez une autre requête ListSnapshotBlocks ou ListChangedBlocks pour le même instantané.

Total de contrôle

Une somme de contrôle est une référence de petite taille dérivée d'un bloc de données dans le but de détecter les erreurs introduites lors de sa transmission ou de son stockage. Les API directes EBS utilisent les sommes de contrôle pour valider l'intégrité des données. Lorsque vous lisez des données à partir d'un instantané EBS, le service fournit des sommes de contrôle SHA256 codées en Base64 pour chaque bloc de données transmis, que vous pouvez utiliser pour la validation. Lorsque vous écrivez des données dans un instantané EBS, vous devez fournir une somme de contrôle SHA256 codée en Base64 pour chaque bloc de données transmis. Le service valide les données reçues à l'aide de la somme de contrôle fournie. Pour de plus amples informations, veuillez consulter Utilisation des sommes de contrôle (p. 1155) plus loin dans ce guide.

Chiffrement

Le chiffrement protège vos données en les convertissant en code illisible qui ne peut être déchiffré que par les personnes ayant accès à la clé utilisée pour les chiffrer. Vous pouvez utiliser les API directes EBS pour lire et écrire des instantanés chiffrés, mais il existe certaines limitations. Pour de plus amples informations, veuillez consulter Utilisation du chiffrement (p. 1154) plus loin dans ce guide.

Actions d'API

Les API directes EBS se composent de six actions. Il y a trois actions de lecture et trois actions d'écriture. Les actions de lecture sont ListSnapshotBlocks, ListChangedBlocks et GetSnapshotBlock. Les actions d'écriture sont StartSnapshot, PutSnapshotBlock et CompleteSnapshot. Ces opérations sont décrites dans les sections suivantes.

Répertorier les blocs d'instantané

L'action ListSnapshotBlocks renvoie les index de bloc et les jetons de bloc des blocs dans l'instantané spécifié.

Répertorier les blocs modifiés

L'action ListChangedBlocks renvoie les index de bloc et les jetons de bloc pour les blocs qui sont différents entre deux instantanés spécifiés de la même lignée de volume/d'instantané.

Obtenir des blocs d'instantanés

L'action GetSnapshotBlock renvoie les données d'un bloc pour l'ID d'instantané, l'index de bloc et le jeton de bloc spécifiés.

Démarrer l'instantané

L'action StartSnapshot lance un instantané, soit un instantané incrémentiel, soit un nouvel instantané. L'instantané démarré reste en attente jusqu'à sa terminaison par l'action CompleteSnapshot.

Ajouter le bloc d'instantanés

L'action PutSnapshotBlock ajoute des données à un instantané démarré sous la forme de blocs individuels. Vous devez spécifier un total de contrôle SHA256 codé en Base64 pour le bloc de données transmises. Le service valide le total de contrôle après la fin de la transmission des données. La requête échoue lorsque le total de contrôle calculé par le service ne correspond pas à la valeur que vous avez spécifiée.

Terminer l'instantané

L'action CompleteSnapshot termine un instantané démarré qui est en attente. L'instantané passe alors à l'état terminé.

Utilisation des API directes EBS pour lire les instantanés

Les étapes suivantes décrivent comment utiliser les API directes EBS pour lire des instantanés :

1. Utilisez l'action ListSnapshotBlocks pour afficher tous les index de bloc et les jetons de bloc des blocs dans un instantané. Vous pouvez également utiliser l'action ListChangedBlocks pour afficher uniquement les index de blocs et les jetons de blocs de blocs différents entre deux instantanés du même volume et de la même lignée d'instantanés. Ces actions vous aident à identifier les jetons de bloc et les index de bloc des blocs pour lesquels vous pouvez obtenir des données.
2. Utilisez l'action GetSnapshotBlock et spécifiez l'index de bloc et le jeton de bloc du bloc pour lequel vous souhaitez obtenir des données.

Pour obtenir des exemples d'exécution de ces actions, veuillez consulter les sections Utilisation des API directes EBS à l'aide de l'API ou des kits AWS SDK (p. 1156) et Utilisation de API directes EBS à l'aide de la ligne de commande (p. 1160) plus loin dans ce guide.

Utilisation des API directes EBS pour écrire des instantanés incrémentiels

Les étapes suivantes décrivent comment utiliser des API directes EBS pour écrire des instantanés incrémentiels :

1. Utilisez l'action StartSnapshot et spécifiez un ID d'instantané parent pour démarrer un instantané en tant qu'instantané incrémentiel d'un instantané existant, ou omettez l'ID d'instantané parent pour démarrer un nouvel instantané. Cette action renvoie le nouvel ID d'instantané, qui est en attente.
2. Utilisez l'action PutSnapshotBlock et spécifiez l'ID de l'instantané en attente pour y ajouter des données sous la forme de blocs individuels. Vous devez spécifier un total de contrôle SHA256 codé en Base64 pour le bloc de données transmises. Le service calcule la somme de contrôle des données reçues et la valide avec la somme de contrôle que vous avez spécifiée. L'action échoue si les sommes de contrôle ne correspondent pas.
3. Lorsque vous avez terminé d'ajouter des données à l'instantané en attente, utilisez l'action CompleteSnapshot pour démarrer un flux de travail asynchrone qui scelle l'instantané et le déplace vers un état terminé.

Répétez ces étapes pour créer un nouvel instantané incrémentiel à l'aide de l'instantané précédemment créé en tant que parent.

Par exemple, dans le diagramme suivant, l'instantané A est le premier nouvel instantané démarré. L'instantané A est utilisé comme instantané parent pour démarrer l'instantané B. L'instantané B est utilisé comme instantané parent pour démarrer et créer l'instantané C. Les instantanés A, B et C sont des instantanés incrémentiels. L'instantané A est utilisé pour créer le volume EBS 1. L'instantané D est créé à partir du volume EBS 1. L'instantané D est un instantané incrémentiel de A ; et non un instantané incrémentiel de B ou C.

Pour obtenir des exemples d'exécution de ces actions, veuillez consulter les sections Utilisation des API directes EBS à l'aide de l'API ou des kits AWS SDK (p. 1156) et Utilisation de API directes EBS à l'aide de la ligne de commande (p. 1160) plus loin dans ce guide.

Autorisations pour les utilisateurs IAM

Un utilisateur AWS Identity and Access Management (IAM) doit disposer des stratégies suivantes pour pouvoir utiliser les API directes EBS. Pour de plus amples informations, veuillez consulter Modification des autorisations pour un utilisateurIAM.

Soyez prudent lorsque vous affectez les stratégies suivantes aux utilisateurs IAM. En affectant ces stratégies, vous risquez de donner accès à un utilisateur auquel l'accès à la même ressource est refusé via les API Amazon EC2, telles que les actions CopySnapshot ou CreateVolume.

Autorisations de lire des instantanés

La stratégie suivante permet d'utiliser la lecture API directes EBS sur tous les instantanés d'une région AWS spécifique. Dans la stratégie, remplacez `<Region>` par la région de l'instantané.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ebs:ListSnapshotBlocks",
                "ebs:ListChangedBlocks",
                "ebs:GetSnapshotBlock"
            ],
```

```
            "Resource": "arn:aws:ec2:<Region>::snapshot/*"
        }
    ]
}
```

La stratégie suivante permet d'utiliser la lecture API directes EBS sur les instantanés avec une balise clé-valeur spécifique. Dans la stratégie, remplacez *<Key>* par la valeur de clé de la balise et *<Value>* par la valeur de la balise.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ebs:ListSnapshotBlocks",
                "ebs:ListChangedBlocks",
                "ebs:GetSnapshotBlock"
            ],
            "Resource": "arn:aws:ec2:*::snapshot/*",
            "Condition": {
                "StringEqualsIgnoreCase": {
                    "aws:ResourceTag/<Key>": "<Value>"
                }
            }
        }
    ]
}
```

La stratégie suivante permet d'utiliser toutes les API directes EBS de lecture sur tous les instantanés du compte uniquement dans une plage de temps spécifique. Cette stratégie autorise l'utilisation du API directes EBS basé sur la clé de condition `aws:CurrentTime` globale. Dans la stratégie, veillez à remplacer la plage de dates et d'heures affichée par la plage de dates et d'heures de votre stratégie.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ebs:ListSnapshotBlocks",
                "ebs:ListChangedBlocks",
                "ebs:GetSnapshotBlock"
            ],
            "Resource": "arn:aws:ec2:*::snapshot/*",
            "Condition": {
                "DateGreaterThan": {
                    "aws:CurrentTime": "2018-05-29T00:00:00Z"
                },
                "DateLessThan": {
                    "aws:CurrentTime": "2020-05-29T23:59:59Z"
                }
            }
        }
    ]
}
```

La stratégie suivante permet de déchiffrer un instantané chiffré à l'aide d'un ID de clé spécifique issu d'AWS Key Management Service (AWS KMS). Elle permet de chiffrer de nouveaux instantanés à l'aide de l'ID de clé AWS KMS par défaut pour les instantanés EBS. Elle permet également de déterminer si le chiffrement par défaut est activé sur le compte. Dans la stratégie, remplacez *<Region>* par la région de

la clé AWS KMS, *<AccountId>* l'ID du compte AWS de la clé et *<KeyId>* par l'ID de la clé utilisée pour chiffrer l'instantané souhaité à lire avec le API directes EBS.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "VisualEditor0",
            "Effect": "Allow",
            "Action": [
                "kms:Encrypt",
                "kms:Decrypt",
                "kms:GenerateDataKey",
                "kms:GenerateDataKeyWithoutPlaintext",
                "kms:ReEncrypt*",
                "kms:CreateGrant",
                "ec2:CreateTags",
                "kms:DescribeKey",
                "ec2:GetEbsDefaultKmsKeyId",
                "ec2:GetEbsEncryptionByDefault"
            ],
            "Resource": "arn:aws:kms:<Region>:<AccountId>:key/<KeyId>"
        }
    ]
}
```

Pour de plus amples informations, veuillez consulter Modification des autorisations pour un utilisateur IAM dans le IAM Guide de l'utilisateur.

Autorisations d'écrire des instantanés

La stratégie suivante permet d'utiliser les API directes EBS d'écriture sur tous les instantanés d'une région AWS spécifique. Dans la stratégie, remplacez *<Region>* par la région de l'instantané.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ebs:StartSnapshot",
                "ebs:PutSnapshotBlock",
                "ebs:CompleteSnapshot"
            ],
            "Resource": "arn:aws:ec2:<Region>::snapshot/*"
        }
    ]
}
```

La stratégie suivante permet d'utiliser les API directes EBS d'écriture sur les instantanés avec une balise clé-valeur spécifique. Dans la stratégie, remplacez *<Key>* par la valeur de clé de la balise et *<Value>* par la valeur de la balise.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ebs:StartSnapshot",
                "ebs:PutSnapshotBlock",
                "ebs:CompleteSnapshot"
```

```
            ],
            "Resource": "arn:aws:ec2:*::snapshot/*",
            "Condition": {
                "StringEqualsIgnoreCase": {
                    "aws:ResourceTag/<Key>": "<Value>"
                }
            }
        }
    ]
}
```

La stratégie suivante permet l'utilisation de toutes les API directes EBS. Elle n'autorise également l'action `StartSnapshot` que si un ID d'instantané parent est spécifié. Par conséquent, cette stratégie bloque la possibilité de démarrer de nouveaux instantanés sans utiliser un instantané parent.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "ebs:*",
            "Resource": "*",
            "Condition": {
                "StringEquals": {
                    "ebs:ParentSnapshot": "arn:aws:ec2:*::snapshot/*"
                }
            }
        }
    ]
}
```

La stratégie suivante permet l'utilisation de toutes les API directes EBS. Il permet également de créer uniquement la clé de balise `user` pour un nouvel instantané. Cette stratégie garantit également que l'utilisateur dispose de l'accès approprié pour créer des balises. L'action `StartSnapshot` est la seule action qui peut spécifier des balises.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "ebs:*",
            "Resource": "*",
            "Condition": {
                "ForAllValues:StringEquals": {
                    "aws:TagKeys": "user"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": "ec2:CreateTags",
            "Resource": "*"
        }
    ]
}
```

La stratégie suivante permet d'utiliser toutes les API directes EBS d'écriture sur tous les instantanés du compte uniquement dans une plage de temps spécifique. Cette stratégie autorise l'utilisation du API directes EBS basé sur la clé de condition `aws:CurrentTime` globale. Dans la stratégie, veillez à remplacer la plage de dates et d'heures affichée par la plage de dates et d'heures de votre stratégie.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "ebs:StartSnapshot",
                "ebs:PutSnapshotBlock",
                "ebs:CompleteSnapshot"
            ],
            "Resource": "arn:aws:ec2:*::snapshot/*",
            "Condition": {
                "DateGreaterThan": {
                    "aws:CurrentTime": "2018-05-29T00:00:00Z"
                },
                "DateLessThan": {
                    "aws:CurrentTime": "2020-05-29T23:59:59Z"
                }
            }
        }
    ]
}
```

La stratégie suivante permet de déchiffrer un instantané chiffré à l'aide d'un ID de clé spécifique issu d'AWS Key Management Service (AWS KMS). Elle permet de chiffrer de nouveaux instantanés à l'aide de l'ID de clé AWS KMS par défaut pour les instantanés EBS. Elle permet également de déterminer si le chiffrement par défaut est activé sur le compte. Dans la stratégie, remplacez *<Region>* par la région de la clé AWS KMS, *<AccountId>* l'ID du compte AWS de la clé et *<KeyId>* par l'ID de la clé utilisée pour chiffrer l'instantané souhaité à lire avec le API directes EBS.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "VisualEditor0",
            "Effect": "Allow",
            "Action": [
                "kms:Encrypt",
                "kms:Decrypt",
                "kms:GenerateDataKey",
                "kms:GenerateDataKeyWithoutPlaintext",
                "kms:ReEncrypt*",
                "kms:CreateGrant",
                "ec2:CreateTags",
                "kms:DescribeKey",
                "ec2:GetEbsDefaultKmsKeyId",
                "ec2:GetEbsEncryptionByDefault"
            ],
            "Resource": "arn:aws:kms:<Region>:<AccountId>:key/<KeyId>"
        }
    ]
}
```

Pour de plus amples informations, veuillez consulter Modification des autorisations pour un utilisateur IAM dans le IAM Guide de l'utilisateur.

Utilisation du chiffrement

Si le chiffrement Amazon EBS est activé par défaut sur votre compteAWS , vous ne pouvez pas démarrer un nouvel instantané à l'aide d'un instantané parent non chiffré. Vous devez d'abord chiffrer l'instantané parent en le copiant. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138) et Chiffrement par défaut (p. 1196).

Pour démarrer un instantané chiffré, spécifiez le nom de ressource Amazon (ARN) d'une clé AWS KMS ou spécifiez un instantané parent chiffré dans votre demande StartSnapshot. Si rien n'est spécifié et que le chiffrement Amazon EBS par défaut est activé sur le compte, la clé CMK par défaut du compte est utilisée. Si aucune clé CMK par défaut n'a été spécifiée pour le compte, la clé CMK gérée par AWS est utilisée.

> **Important**
>
> Par défaut, tous les principaux du compte ont accès à la clé CMK gérée par AWS par défaut, et ils peuvent l'utiliser pour les opérations de chiffrement et de déchiffrement EBS. Pour plus d'informations, consultez Clé par défaut pour le chiffrement EBS (p. 1196).

Vous pouvez avoir besoin d'autorisations IAM supplémentaires pour utiliser le API directes EBS avec le chiffrement. Pour de plus amples informations, veuillez consulter la section Autorisations pour les utilisateurs IAM (p. 1150) de ce guide.

Utilisation de la signature de la version 4

Signature Version 4 est le processus permettant d'ajouter des informations d'authentification à des demandes AWS par HTTP. Pour des raisons de sécurité, la plupart des demandes à AWS doivent être signées avec une clé d'accès, qui se compose d'un ID de clé d'accès et de la clé d'accès secrète. Ces deux clés sont généralement appelées informations d'identification de sécurité. Pour de plus amples informations sur la façon d'obtenir des informations d'identification pour votre compte, veuillez consulter Comprendre et obtenir vos informations d'identification.

Si vous avez l'intention de créer manuellement des requêtes HTTP, vous devez apprendre à les signer. Lorsque vous utilisez l'AWS Command Line Interface (AWS CLI) ou un des kits SDK AWS pour effectuer des demandes à AWS;, ces outils signent automatiquement les demandes avec la clé d'accès que vous spécifiez lors de la configuration des outils. Lorsque vous utilisez ces derniers, vous n'avez pas besoin d'apprendre à signer vous-même les demandes.

Pour de plus amples informations, veuillez consulter Signature des demandes AWS avec Signature Version 4 dans le AWS General Reference.

Utilisation des sommes de contrôle

L'action GetSnapshotBlock renvoie des données qui se trouvent dans un bloc d'un instantané, et l'action PutSnapshotBlock ajoute des données à un bloc d'un instantané. Les données de bloc transmises ne sont pas signées dans le cadre du processus de signature de la version 4. Par conséquent, les sommes de contrôle sont utilisées pour valider l'intégrité des données comme suit :

- Lorsque vous utilisez l'action GetSnapshotBlock, la réponse fournit une somme de contrôle SHA256 codée en Base64 pour les données de bloc à l'aide de l'en-tête x-amz-Checksum, et l'algorithme checksum à l'aide de l'en-tête x-amz-Checksum-Algorithm. Utilisez la somme de contrôle renvoyée pour valider l'intégrité des données. Si la somme de contrôle que vous générez ne correspond pas à celle fournie par Amazon EBS, vous devez considérer les données non valides et réessayer votre demande.
- Lorsque vous utilisez l'action PutSnapshotBlock, votre demande doit fournir une somme de contrôle SHA256 codée en Base64 pour les données de bloc à l'aide de l'en-tête x-amz-Checksum, et l'algorithme checksum à l'aide de l'en-tête x-amz-Checksum-Algorithm. La somme de contrôle que vous fournissez est validée par rapport à une somme de contrôle générée par Amazon EBS pour valider l'intégrité des données. Si les sommes de contrôle ne correspondent pas, la demande échoue.
- Lorsque vous utilisez l'action CompleteSnapshot, votre demande peut éventuellement fournir une somme de contrôle SHA256 codée en Base64 agrégée pour l'ensemble complet des données ajoutées à l'instantané. Fournissez la somme de contrôle à l'aide de l'en-tête x-amz-Checksum, l'algorithme de somme de contrôle à l'aide de l'en-tête x-amz-Checksum-Algorithm et la méthode d'agrégation de somme de contrôle à l'aide de l'en-tête x-amz-Checksum-Aggregation-Method. Pour générer la somme de contrôle agrégée à l'aide de la méthode d'agrégation linéaire, organisez les sommes de contrôle pour chaque bloc écrit dans l'ordre croissant de leur index de bloc, concaténez-les pour former une seule chaîne, puis générez la somme de contrôle sur la chaîne entière à l'aide de l'algorithme SHA256.

Les sommes de contrôle de ces actions font partie du processus de signature de la version 4.

Utilisation des API directes EBS à l'aide de l'API ou des kits AWS SDK

La référence des API directes EBS fournit des descriptions et la syntaxe pour chacune des actions et chacun des types de données du service. Vous pouvez également utiliser l'un des kits SDK AWS pour accéder à une API adaptée au langage de programmation ou à la plateforme que vous utilisez. Pour de plus amples informations, veuillez consulter SDK AWS.

Les API directes EBS exigent une signature AWS Signature Version 4. Pour de plus amples informations, veuillez consulter Utilisation de la signature de la version 4 (p. 1155).

Utilisation de l'API pour lire des instantanés

Liste des blocs dans un instantané

L'exemple de requête ListChangedBlocks suivant renvoie les index de bloc et les jetons de bloc des blocs qui sont dans l'instantané `snap-0acEXAMPLEcf41648`. Le paramètre `startingBlockIndex` limite les résultats aux index de blocs supérieurs à `1000`, et le paramètre `maxResults` limite les résultats aux premiers blocs `100`.

```
GET /snapshots/snap-0acEXAMPLEcf41648/blocks?maxResults=100&startingBlockIndex=0 HTTP/1.1
Host: ebs.us-east-2.amazonaws.com
Accept-Encoding: identity
User-Agent: <User agent parameter>
X-Amz-Date: 20200617T231953Z
Authorization: <Authentication parameter>
```

L'exemple de réponse suivant pour la demande précédente répertorie les index de bloc et les jetons de bloc dans l'instantané. Utilisez l'action GetSnapshotBlock et spécifiez l'index de bloc et le jeton de bloc du bloc pour lequel vous souhaitez obtenir des données. Les jetons de bloc sont valides jusqu'au délai d'expiration indiqué.

```
HTTP/1.1 200 OK
x-amzn-RequestId: d6e5017c-70a8-4539-8830-57f5557f3f27
Content-Type: application/json
Content-Length: 2472
Date: Wed, 17 Jun 2020 23:19:56 GMT
Connection: keep-alive

{
    "BlockSize": 524288,
    "Blocks": [
        {
            "BlockIndex": 0,
            "BlockToken": "AAUBAcuWqOCnDNuKle11s7IIX6jp6FYcC/q8oT93913HhvLvA+3JRrSybp/0"
        },
        {
            "BlockIndex": 1536,
            "BlockToken": "AAUBAWudwfmofcrQhGVlLwuRKm2b8ZXPiyrgoykTRC6IU1NbxKWDY1pPjvnV"
        },
        {
            "BlockIndex": 3072,
            "BlockToken": "AAUBAV7p6pC5fKAC7TokoNCtAnZhqq27u6YEXZ3MwRevBkDjmMx6iuA6tsBt"
        },
        {
            "BlockIndex": 3073,
            "BlockToken": "AAUBAbqt9zpqBUEvtO2HINAfFaWToOwlPjbIsQOlx6JUN/0+iMQl0NtNbnX4"
        },
        …
    ],
```

```
    "ExpiryTime": 1.59298379649E9,
    "VolumeSize": 3
}
```

Liste des blocs qui sont différents entre deux instantanés

L'exemple de requête ListChangedBlocks suivant renvoie les index de bloc et les jetons de bloc des blocs qui sont différents entre les instantanés `snap-0acEXAMPLEcf41648` et `snap-0c9EXAMPLE1b30e2f`. Le paramètre `startingBlockIndex` limite les résultats aux index de blocs supérieurs à `0`, et le paramètre `maxResults` limite les résultats aux premiers blocs `500`.

```
GET /snapshots/snap-0c9EXAMPLE1b30e2f/changedblocks?
firstSnapshotId=snap-0acEXAMPLEcf41648&maxResults=500&startingBlockIndex=0 HTTP/1.1
Host: ebs.us-east-2.amazonaws.com
Accept-Encoding: identity
User-Agent: <User agent parameter>
X-Amz-Date: 20200617T232546Z
Authorization: <Authentication parameter>
```

L'exemple de réponse suivant pour la requête précédente montre que les index de bloc `0`, `3072`, `6002` et `6003` sont différents entre les deux instantanés. De plus, les index de bloc `6002` et `6003` existent uniquement dans le premier ID d'instantané spécifié, et pas dans le second ID d'instantané car la réponse ne répertorie aucun second jeton de bloc.

Utilisez l'action `GetSnapshotBlock` et spécifiez l'index de bloc et le jeton de bloc du bloc pour lequel vous souhaitez obtenir des données. Les jetons de bloc sont valides jusqu'au délai d'expiration indiqué.

```
HTTP/1.1 200 OK
x-amzn-RequestId: fb0f6743-6d81-4be8-afbe-db11a5bb8a1f
Content-Type: application/json
Content-Length: 1456
Date: Wed, 17 Jun 2020 23:25:47 GMT
Connection: keep-alive

{
    "BlockSize": 524288,
    "ChangedBlocks": [
        {
            "BlockIndex": 0,
            "FirstBlockToken": "AAUBAVaWqOCnDNuKle11s7IIX6jp6FYcC/tJuVT1GgP23AuLntwiMdJ
+OJkL",
            "SecondBlockToken": "AAUBASxzy0Y0b33JVRLoYm3NOresCxn5RO+HVFzXW3Y/
RwfFaPX2Edx8QHCh"
        },
        {
            "BlockIndex": 3072,
            "FirstBlockToken": "AAUBAcHp6pC5fKAC7TokoNCtAnZhqq27u6fxRfZOLEmeXLmHBf2R/
Yb24MaS",
            "SecondBlockToken":
 "AAUBARGCaufCqBRZC8tEkPYGGkSv3vqvOjJ2xKDi3ljDFiytUxBLXYgTmkid"
        },
        {
            "BlockIndex": 6002,
            "FirstBlockToken": "AAABASqX4/
NWjvNceoyMUljcRd0DnwbSwNnes1UkoP62CrQXvn47BY5435aw"
        },
        {
            "BlockIndex": 6003,
            "FirstBlockToken":
 "AAABASmJ0O5JxAOce25rF4P1sdRtyIDsX12tFEDunnePYUKOf4PBROuICb2A"
        },
        ...
```

```
    ],
    "ExpiryTime": 1.592976647009E9,
    "VolumeSize": 3
}
```

Obtenir des données de bloc à partir d'un instantané

L'exemple de requête GetSnapshotBlock suivant renvoie les données de l'index de bloc `3072` avec un jeton de bloc `AAUBARGCaufCqBRZC8tEkPYGGkSv3vqvOjJ2xKDi3ljDFiytUxBLXYgTmkid`, dans un instantané `snap-0c9EXAMPLE1b30e2f`.

```
GET /snapshots/snap-0c9EXAMPLE1b30e2f/blocks/3072?
blockToken=AAUBARGCaufCqBRZC8tEkPYGGkSv3vqvOjJ2xKDi3ljDFiytUxBLXYgTmkid HTTP/1.1
Host: ebs.us-east-2.amazonaws.com
Accept-Encoding: identity
User-Agent: <User agent parameter>
X-Amz-Date: 20200617T232838Z
Authorization: <Authentication parameter>
```

L'exemple de réponse suivant pour la demande précédente montre la taille des données renvoyées, la somme de contrôle pour valider les données et l'algorithme utilisé pour générer la somme de contrôle. Les données binaires sont transmises dans le corps de la réponse et sont représentées comme *`BlockData`* dans l'exemple suivant.

```
HTTP/1.1 200 OK
x-amzn-RequestId: 2d0db2fb-bd88-474d-a137-81c4e57d7b9f
x-amz-Data-Length: 524288
x-amz-Checksum: Vc0yY2j3qg8bUL9I6GQuI2orTudrQRBDMIhcy7bdEsw=
x-amz-Checksum-Algorithm: SHA256
Content-Type: application/octet-stream
Content-Length: 524288
Date: Wed, 17 Jun 2020 23:28:38 GMT
Connection: keep-alive

BlockData
```

Utilisation de l'API pour écrire des instantanés incrémentiels

Démarrer un instantané

L'exemple de demande StartSnapshot suivant démarre un instantané `8` Gio en utilisant l'instantané `snap-123EXAMPLE1234567` comme instantané parent. Le nouvel instantané sera un instantané incrémentiel de l'instantané parent. L'instantané passe à un état d'erreur s'il n'y a pas de demande d'ajout ou d'exécution pour l'instantané pendant la période de `60` minutes spécifiée. Le jeton client `550e8400-e29b-41d4-a716-446655440000` garantit l'idempotence pour la demande. Si le jeton client est omis, le kit AWS SDK en génère automatiquement un pour vous. Pour de plus amples informations sur l'idempotence, veuillez consulter Idempotence pour l'API StartSnapshot (p. 1168).

```
POST /snapshots HTTP/1.1
Host: ebs.us-east-2.amazonaws.com
Accept-Encoding: identity
User-Agent: <User agent parameter>
X-Amz-Date: 20200618T040724Z
Authorization: <Authentication parameter>

{
    "VolumeSize": 8,
    "ParentSnapshot": snap-123EXAMPLE1234567,
    "ClientToken": "550e8400-e29b-41d4-a716-446655440000",
    "Timeout": 60
```

```
}
```

L'exemple de réponse suivant pour la demande précédente indique l'ID d'instantané, l'ID de compte AWS, l'état, la taille du volume en Gio et la taille des blocs dans l'instantané. L'instantané est démarré dans un état en attente. Spécifiez l'ID d'instantané dans une demande `PutSnapshotBlocks` ultérieure d'écriture de données dans l'instantané .

```
HTTP/1.1 201 Created
x-amzn-RequestId: 929e6eb9-7183-405a-9502-5b7da37c1b18
Content-Type: application/json
Content-Length: 181
Date: Thu, 18 Jun 2020 04:07:29 GMT
Connection: keep-alive

{
    "BlockSize": 524288,
    "Description": null,
    "OwnerId": "138695307491",
    "Progress": null,
    "SnapshotId": "snap-052EXAMPLEc85d8dd",
    "StartTime": null,
    "Status": "pending",
    "Tags": null,
    "VolumeSize": 8
}
```

Ajouter des données dans un instantané

L'exemple de requête PutSnapshot suivant écrit `524288` octets de données pour bloquer l'index `1000` sur l'instantané `snap-052EXAMPLEc85d8dd`. La somme de contrôle `QOD3gmEQOXATfJx2Aa34W4FU2nZGyXfqtsUuktOw8DM=` codée en Base64 a été générée à l'aide de l'algorithme `SHA256`. Les données sont transmises dans le corps de la requête et sont représentées comme *`BlockData`* dans l'exemple suivant.

```
PUT /snapshots/snap-052EXAMPLEc85d8dd/blocks/1000 HTTP/1.1
Host: ebs.us-east-2.amazonaws.com
Accept-Encoding: identity
x-amz-Data-Length: 524288
x-amz-Checksum: QOD3gmEQOXATfJx2Aa34W4FU2nZGyXfqtsUuktOw8DM=
x-amz-Checksum-Algorithm: SHA256
User-Agent: <User agent parameter>
X-Amz-Date: 20200618T042215Z
X-Amz-Content-SHA256: UNSIGNED-PAYLOAD
Authorization: <Authentication parameter>

          BlockData
```

Voici un exemple de réponse pour la demande précédente avec confirmation de la longueur des données, la somme de contrôle et l'algorithme de somme de contrôle pour les données reçues par le service.

```
HTTP/1.1 201 Created
x-amzn-RequestId: 643ac797-7e0c-4ad0-8417-97b77b43c57b
x-amz-Checksum: QOD3gmEQOXATfJx2Aa34W4FU2nZGyXfqtsUuktOw8DM=
x-amz-Checksum-Algorithm: SHA256
Content-Type: application/json
Content-Length: 2
Date: Thu, 18 Jun 2020 04:22:12 GMT
Connection: keep-alive

{}
```

Terminer un instantané

L'exemple de demande CompleteSnapshot suivant termine l'instantané `snap-052EXAMPLEc85d8dd`. La commande spécifie que les blocs `5` ont été écrits dans l'instantané. La somme de contrôle `6D3nmwi5f2F0wlh7xX8QprrJBFzDX8aacdOcA3KCM3c=` représente la somme de contrôle de l'ensemble complet des données écrites dans un instantané.

```
POST /snapshots/completion/snap-052EXAMPLEc85d8dd HTTP/1.1
Host: ebs.us-east-2.amazonaws.com
Accept-Encoding: identity
x-amz-ChangedBlocksCount: 5
x-amz-Checksum: 6D3nmwi5f2F0wlh7xX8QprrJBFzDX8aacdOcA3KCM3c=
x-amz-Checksum-Algorithm: SHA256
x-amz-Checksum-Aggregation-Method: LINEAR
User-Agent: <User agent parameter>
X-Amz-Date: 20200618T043158Z
Authorization: <Authentication parameter>
```

Voici un exemple de réponse pour la demande précédente.

```
HTTP/1.1 202 Accepted
x-amzn-RequestId: 06cba5b5-b731-49de-af40-80333ac3a117
Content-Type: application/json
Content-Length: 20
Date: Thu, 18 Jun 2020 04:31:50 GMT
Connection: keep-alive

{"Status":"pending"}
```

Utilisation de API directes EBS à l'aide de la ligne de commande

Les exemples suivants montrent comment utiliser les API directes EBS à l'aide d'AWS Command Line Interface (AWS CLI). Pour de plus amples informations sur l'installation et la configuration d'AWS CLI, veuillez consulter Installation de l'AWS CLI version 1 et Configuration rapide de l'AWS CLI.

Utilisation des AWS CLI pour lire des instantanés

Liste des blocs dans un instantané

L'exemple de commande list-snapshot-blocks suivant renvoie les index de bloc et les jetons de bloc des blocs qui sont dans l'instantané `snap-0987654321`. Le paramètre `--starting-block-index` limite les résultats aux index de blocs supérieurs à `1000`, et le paramètre `--max-results` limite les résultats aux premiers blocs `100`.

```
aws ebs list-snapshot-blocks --snapshot-id snap-0987654321 --starting-block-index 1000 --
max-results 100
```

L'exemple de réponse suivant pour la commande précédente répertorie les index de bloc et les jetons de bloc dans l'instantané. Utilisez la commande `get-snapshot-block` et spécifiez l'index de bloc et le jeton de bloc du bloc pour lequel vous souhaitez obtenir des données. Les jetons de bloc sont valides jusqu'au délai d'expiration indiqué.

```
{
    "Blocks": [
        {
            "BlockIndex": 1001,
            "BlockToken": "AAABAV3/PNhXOynVdMYHUpPsetaSvjLB1dtIGfbJv5OJ0sX855EzGTWos4a4"
        },
        {
            "BlockIndex": 1002,
```

```
            "BlockToken": "AAABATGQIgwr0WwIuqIMjCA/Sy7e/YoQFZsHejzGNvjKauzNgzeI13YHBfQB"
        },
        {
            "BlockIndex": 1007,
            "BlockToken": "AAABAZ9CTuQtUvp/dXqRWw4d07eOgTZ3jvn6hiW30W9duM8MiMw6yQayzF2c"
        },
        {
            "BlockIndex": 1012,
            "BlockToken": "AAABAQdzxhw0rVV6PNmsfo/YRIxo9JPR85XxPf1BLjg0Hec6pygYr6laE1p0"
        },
        {
            "BlockIndex": 1030,
            "BlockToken": "AAABAaYvPax6mv+iGWLdTUjQtFWouQ7Dqz6nSD9L+CbXnvpkswA6iDID523d"
        },
        {
            "BlockIndex": 1031,
            "BlockToken": "AAABATgWZC0XcFwUKvTJbUXMiSPg59KVxJGL+BWBClkw6spzCxJVqDVaTskJ"
        },
        ...
    ],
    "ExpiryTime": 1576287332.806,
    "VolumeSize": 32212254720,
    "BlockSize": 524288
}
```

Liste des blocs qui sont différents entre deux instantanés

L'exemple de commande list-changed-blocks suivant renvoie les index de bloc et les jetons de bloc des blocs qui sont différents entre les instantanés `snap-1234567890` et `snap-0987654321`. Le paramètre `--starting-block-index` limite les résultats aux index de blocs supérieurs à `0`, et le paramètre `--max-results` limite les résultats aux premiers blocs `500`.

```
aws ebs list-changed-blocks --first-snapshot-id snap-1234567890 --second-snapshot-
id snap-0987654321 --starting-block-index 0 --max-results 500
```

L'exemple de réponse suivant pour la commande précédente montre que les index de bloc 0, 6000, 6001, 6002 et 6003 sont différents entre les deux instantanés. De plus, les index de bloc 6001, 6002 et 6003 existent uniquement dans le premier ID d'instantané spécifié, et pas dans le second ID d'instantané car la réponse ne répertorie aucun second jeton de bloc.

Utilisez la commande `get-snapshot-block` et spécifiez l'index de bloc et le jeton de bloc du bloc pour lequel vous souhaitez obtenir des données. Les jetons de bloc sont valides jusqu'au délai d'expiration indiqué.

```
{
    "ChangedBlocks": [
        {
            "BlockIndex": 0,
            "FirstBlockToken": "AAABAVahm9SO60Dyi0ORySzn2ZjGjW/
KN3uygGlS0QOYWesbzBbDnX2dGpmC",
            "SecondBlockToken":
 "AAABAf8o0o6UFi1rDbSZGIRaCEdDyBu9TlvtCQxxoKV8qrUPQP7vcM6iWGSr"
        },
        {
            "BlockIndex": 6000,
            "FirstBlockToken": "AAABAbYSiZvJ0/
R9tz8suI8dSzecLjN4kkazK8inFXVintPkdaVFLfCMQsKe",
            "SecondBlockToken":
 "AAABAZnqTdzFmKRpsaMAsDxviVqEI/3jJzI2crq2eFDCgHmyNf777elD9oVR"
        },
        {
            "BlockIndex": 6001,
```

```
            "FirstBlockToken": "AAABASBpSJ2UAD3PLxJnCt6zun4/
T4sU25Bnb8jB5Q6FRXHFqAIAqE04hJoR"
        },
        {
            "BlockIndex": 6002,
            "FirstBlockToken": "AAABASqX4/
NWjvNceoyMUljcRd0DnwbSwNnes1UkoP62CrQXvn47BY5435aw"
        },
        {
            "BlockIndex": 6003,
            "FirstBlockToken":
 "AAABASmJ0O5JxAOce25rF4P1sdRtyIDsX12tFEDunnePYUKOf4PBROuICb2A"
        },
        ...
    ],
    "ExpiryTime": 1576308931.973,
    "VolumeSize": 32212254720,
    "BlockSize": 524288,
    "NextToken": "AAADARqElNng/sV98CYk/bJDCXeLJmLJHnNSkHvLzVaO0zsPH/QM3Bi3zF//O6Mdi/
BbJarBnp8h"
}
```

Obtenir des données de bloc à partir d'un instantané

L'exemple de commande get-snapshot-block suivant renvoie les données de l'index de bloc `6001` avec le jeton de bloc `AAABASBpSJ2UAD3PLxJnCt6zun4/T4sU25Bnb8jB5Q6FRXHFqAIAqE04hJoR`, dans l'instantané `snap-1234567890`. Les données binaires seront générées dans le fichier `data` dans le répertoire `C:\Temp` sur un ordinateur Windows. Si vous exécutez la commande sur un ordinateur Linux ou Unix, remplacez le chemin de sortie par `/tmp/data` pour générer les données dans le fichier `data` du répertoire `/tmp`.

```
aws ebs get-snapshot-block --snapshot-id snap-1234567890 --block-index 6001 --block-
token AAABASBpSJ2UAD3PLxJnCt6zun4/T4sU25Bnb8jB5Q6FRXHFqAIAqE04hJoR C:/Temp/data
```

L'exemple de réponse suivant pour la commande précédente montre la taille des données renvoyées, la somme de contrôle pour valider les données et l'algorithme de la somme de contrôle. Les données binaires sont automatiquement enregistrées dans le répertoire et le fichier que vous avez spécifiés dans la commande de demande.

```
{
    "DataLength": "524288",
    "Checksum": "cf0Y6/Fn0oFa4VyjQPOa/iD0zhTflPTKzxGv2OKowXc=",
    "ChecksumAlgorithm": "SHA256"
}
```

Utilisation des AWS CLI pour écrire des instantanés incrémentiels

Démarrer un instantané

L'exemple de commande start-snapshot suivant démarre un instantané `8` Gio en utilisant l'instantané `snap-123EXAMPLE1234567` comme instantané parent. Le nouvel instantané sera un instantané incrémentiel de l'instantané parent. L'instantané passe à un état d'erreur s'il n'y a pas de demande d'ajout ou d'exécution pour l'instantané pendant la période de `60` minutes spécifiée. Le jeton client `550e8400-e29b-41d4-a716-446655440000` garantit l'idempotence pour la demande. Si le jeton client est omis, le kit AWS SDK en génère automatiquement un pour vous. Pour de plus amples informations sur l'idempotence, veuillez consulter Idempotence pour l'API StartSnapshot (p. 1168).

```
aws ebs start-snapshot --volume-size 8 --parent-snapshot snap-123EXAMPLE1234567 --
timeout 60 --client-token 550e8400-e29b-41d4-a716-446655440000
```

L'exemple de réponse suivant pour la commande précédente indique l'ID du snapshot, l'ID du compte AWS, l'état, la taille du volume en Gio et la taille des blocs dans l'instantané. L'instantané est démarré dans un état `pending`. Spécifiez l'ID d'instantané dans les commandes `put-snapshot-block` suivantes pour écrire des données dans l'instantané, puis utilisez la commande `complete-snapshot` pour terminer l'instantané et modifier son état sur `completed`.

```
{
    "SnapshotId": "snap-0aaEXAMPLEe306d62",
    "OwnerId": "111122223333",
    "Status": "pending",
    "VolumeSize": 8,
    "BlockSize": 524288
}
```

Ajouter des données dans un instantané

L'exemple de commande put-snapshot suivant écrit les `524288` octets de données pour bloquer l'index `1000` sur l'instantané `snap-0aaEXAMPLEe306d62`. La somme de contrôle `QOD3gmEQOXATfJx2Aa34W4FU2nZGyXfqtsUuktOw8DM=` codée en Base64 a été générée à l'aide de l'algorithme `SHA256`. Les données transmises se trouvent dans le fichier `/tmp/data`.

```
aws ebs put-snapshot-block --snapshot-id snap-0aaEXAMPLEe306d62
 --block-index 1 --data-length 524288 --block-data /tmp/data --
checksum QOD3gmEQOXATfJx2Aa34W4FU2nZGyXfqtsUuktOw8DM= --checksum-algorithm SHA256
```

L'exemple de réponse suivant pour la commande précédente confirme la longueur des données, la somme de contrôle et l'algorithme de somme de contrôle pour les données reçues par le service.

```
{
    "DataLength": "524288",
    "Checksum": "QOD3gmEQOXATfJx2Aa34W4FU2nZGyXfqtsUuktOw8DM=",
    "ChecksumAlgorithm": "SHA256"
}
```

Terminer un instantané

L'exemple de commande complete-snapshot suivant termine l'instantané `snap-0aaEXAMPLEe306d62`. La commande spécifie que les blocs `5` ont été écrits dans l'instantané. La somme de contrôle `6D3nmwi5f2F0wlh7xX8QprrJBFzDX8aacdOcA3KCM3c=` représente la somme de contrôle de l'ensemble complet des données écrites dans un instantané. Pour de plus amples informations sur les sommes de contrôle, veuillez consulter Utilisation des sommes de contrôle (p. 1155) plus haut dans ce guide.

```
aws ebs complete-snapshot --snapshot-id snap-0aaEXAMPLEe306d62 --changed-blocks-count 5
 --checksum 6D3nmwi5f2F0wlh7xX8QprrJBFzDX8aacdOcA3KCM3c= --checksum-algorithm SHA256 --
checksum-aggregation-method LINEAR
```

Voici un exemple de réponse pour la commande précédente.

```
{
    "Status": "pending"
}
```

Optimisation des performances

Vous pouvez exécuter des demandes d'API simultanément. En supposant que la latence de PutSnapshotBlock est de 100 ms, un thread peut traiter 10 demandes en une seconde. En outre, en supposant que votre application cliente crée plusieurs threads et connexions (par exemple,

100 connexions), elle peut faire 1000 (10 * 100) demandes par seconde au total. Cela correspondra à un débit d'environ 500 Mo par seconde.

La liste suivante contient quelques éléments à rechercher dans votre application :

- Chaque thread utilise-t-il une connexion séparée ? Si les connexions sont limitées sur l'application, plusieurs threads attendront que la connexion soit disponible et vous remarquerez un débit inférieur.
- Y a-t-il un temps d'attente dans l'application entre deux demandes d'ajout ? Cela réduira le débit effectif d'un thread.
- Limite de bande passante sur l'instance : si la bande passante sur l'instance est partagée par d'autres applications, elle pourrait limiter le débit disponible pour les requêtes PutSnapshotBlock.

Veillez à prendre note des autres charges de travail qui peuvent être exécutées dans le compte pour éviter les goulots d'étranglement. Vous devez également créer des mécanismes de nouvelle tentative dans vos flux de travail API directes EBS afin de gérer la limitation, les délais d'attente et l'indisponibilité du service.

Passez en revue les quotas de service API directes EBS pour déterminer le maximum de demandes d'API que vous pouvez exécuter par seconde. Pour de plus amples informations, veuillez consulter Points de terminaison et quotas Amazon Elastic Block Store dans la Référence générale AWS.

Questions fréquentes (FAQ)

Est-il possible d'accéder à un instantané à l'aide des API directes EBS si son statut est en attente ?

Non. L'instantané n'est accessible que si son statut est terminé.

Les index de bloc sont-ils renvoyés par les API directes EBS dans l'ordre numérique ?

Oui. Les index de bloc renvoyés sont uniques et classés par ordre numérique.

Puis-je soumettre une demande avec une valeur de paramètre MaxResults inférieure à 100 ?

Non. La valeur minimale du paramètre MaxResult que vous pouvez utiliser est 100. Si vous soumettez une demande avec une valeur de paramètre MaxResult inférieure à 100 et que l'instantané comporte plus de 100 blocs, l'API renvoie au moins 100 résultats.

Puis-je exécuter des demandes d'API simultanément ?

Vous pouvez exécuter des demandes d'API simultanément. Veillez à prendre note des autres charges de travail qui peuvent être exécutées dans le compte pour éviter les goulots d'étranglement. Vous devez également créer des mécanismes de nouvelle tentative dans vos flux de travail API directes EBS afin de gérer la limitation, les délais d'attente et l'indisponibilité du service. Pour de plus amples informations, veuillez consulter Optimisation des performances (p. 1163).

Passez en revue les quotas de service API directes EBS pour déterminer les demandes d'API que vous pouvez exécuter par seconde. Pour de plus amples informations, veuillez consulter Points de terminaison et quotas Amazon Elastic Block Store dans la Référence générale AWS.

Lors de l'exécution de l'action ListChangedBlocks, est-il possible d'obtenir une réponse vide même si l'instantané comporte des blocs ?

Oui. Si les blocs modifiés sont rares dans l'instantané, la réponse peut être vide, mais l'API renverra une valeur de jeton de page suivante. Utilisez la valeur de jeton de page suivante pour passer à la page suivante des résultats. Vous pouvez confirmer que vous avez atteint la dernière page de résultats lorsque l'API renvoie une valeur de jeton de page suivante nulle.

Si le paramètre NextToken est spécifié avec un paramètre StartingBlockIndex, lequel des deux est utilisé ?

Le NextToken est utilisé et le StartingBlockIndex est ignoré.

Quelle est la durée de validité des jetons de bloc et des jetons suivants ?

Les jetons de bloc sont valides pendant sept jours et les jetons suivants sont valides pendant 60 minutes.

Les instantanés chiffrés sont-ils pris en charge ?

Oui. Les instantanés chiffrés sont accessibles à l'aide des API directes EBS.

Pour accéder à un instantané chiffré, l'utilisateur doit avoir accès à la clé utilisée pour chiffrer l'instantané et à l'action de déchiffrement AWS KMS. Pour de plus amples informations sur la stratégie AWS KMS à affecter à un utilisateur, veuillez consulter la section Autorisations pour les utilisateurs IAM (p. 1150) plus haut dans le présent guide.

Les instantanés publics sont-ils pris en charge ?

Les instantanés publics ne sont pas pris en charge.

L'opération ListSnapshotBlocks renvoie-t-elle tous les index de bloc et tous les jetons de bloc d'un instantané, ou seulement ceux dans lesquels des données ont été écrites ?

Elle renvoie uniquement les index de bloc et les jetons de bloc dans lesquels des données ont été écrites.

Puis-je obtenir un historique de tous les appels d'API SNS effectués par les API directes EBS sur mon compte à des fins d'analyse de sécurité et de résolution des problèmes opérationnels ?

Oui. Pour recevoir un historique des appels d'API API directes EBS effectués sur votre compte, activez l'option AWS CloudTrail dans AWS Management Console. Pour de plus amples informations, veuillez consulter Journalisation des appels d'API pour le API directes EBS avec AWS CloudTrail (p. 1165).

Journalisation des appels d'API pour le API directes EBS avec AWS CloudTrail

Le service API directes EBS est intégré à AWS CloudTrail. CloudTrail est un service qui fournit un enregistrement des actions réalisées par un utilisateur, un rôle ou un service AWS dans API directes EBS. CloudTrail capture tous les appels d'API pour StartSnapshot et CompleteSnapshot pour API directes EBS en tant qu'événements. Si vous créez un journal de suivi, vous pouvez activer la diffusion en continu des événements CloudTrail sur un compartiment Amazon Simple Storage Service (Amazon S3), y compris les événements pour le API directes EBS. Si vous ne configurez pas de journal de suivi, vous pouvez toujours afficher les événements les plus récents dans la console CloudTrail dans Event history (Historique des événements). Vous pouvez utiliser les informations collectées par CloudTrail pour déterminer la demande qui a été envoyée à API directes EBS, l'adresse IP source à partir de laquelle la demande a été effectuée, l'auteur de la demande et la date de la demande, ainsi que d'autres informations.

Pour de plus amples informations sur CloudTrail, veuillez consultez le AWS CloudTrail User Guide.

Informations API directes EBS dans CloudTrail

CloudTrail est activé sur votre compte AWS lorsque vous créez le compte. Quand une activité d'événement prise en charge a lieu dans API directes EBS, elle est enregistrée dans un événement CloudTrail avec d'autres événements de services AWS dans Historique des événements. Vous pouvez afficher, rechercher et télécharger les événements récents dans votre compte AWS. Pour plus d'informations, consultez Affichage des événements avec l'historique des événements CloudTrail.

Pour un enregistrement continu des événements dans votre compte AWS, y compris les événements pour API directes EBS, créez un journal de suivi. Un journal de suivi permet à CloudTrail de livrer les fichiers journaux dans un compartiment S3. Par défaut, lorsque vous créez un journal de suivi dans la console, il s'applique à toutes les régions AWS. Le journal de suivi consigne les événements de toutes les régions dans la partition AWS et livre les fichiers journaux dans le compartiment S3 de votre choix. En outre, vous pouvez configurer d'autres services AWS pour analyser plus en profondeur les données d'événement collectées dans les journaux CloudTrail et agir sur celles-ci. Pour plus d'informations, consultez les ressources suivantes :

- Présentation de la création d'un journal de suivi
- Intégrations et services pris en charge par CloudTrail

- Configuration des Notifications de Amazon SNS pour CloudTrail
- Réception de fichiers journaux CloudTrail de plusieurs régions et Réception de fichiers journaux CloudTrail de plusieurs comptes

Actions d'API prises en charge

Les actions d'API suivantes prennent en charge la journalisation en tant qu'événements dans les fichiers journaux CloudTrail :

- StartSnapshot
- CompleteSnapshot

Informations relatives à l'identité

Chaque événement ou entrée du journal contient des informations sur la personne qui a généré la demande. Les informations relatives à l'identité permettent de déterminer les éléments suivants :

- Si la demande a été effectuée avec les informations d'identification racine ou utilisateur AWS Identity and Access Management (IAM).
- Si la demande a été effectuée avec des informations d'identification de sécurité temporaires pour un rôle ou un utilisateur fédéré.
- Si la requête a été effectuée par un autre service AWS.

Pour de plus amples informations, veuillez consulter Élément userIdentity CloudTrail.

Présentation des entrées des fichiers journaux API directes EBS

Un journal de suivi est une configuration qui active la livraison d'événements en tant que fichiers journaux à un compartiment S3 que vous spécifiez. Les fichiers journaux CloudTrail contiennent une ou plusieurs entrées de journal. Un événement représente une demande individuelle à partir d'une source quelconque et comprend des informations sur l'action demandée, sur tous les paramètres, les paramètres de la demande, etc. Les fichiers journaux CloudTrail ne sont pas des séries ordonnées retraçant les appels d'API publics. Ils ne suivent aucun ordre précis.

Les exemples suivants montrent des entrées de journal CloudTrail qui illustrent les actions `StartSnapshot` et `CompleteSnapshot`.

Exemple avec StartSnapshot :

```
{
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "123456789012",
        "arn": "arn:aws:iam::123456789012:root",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "userName": "user"
    },
    "eventTime": "2020-07-03T23:27:26Z",
    "eventSource": "ebs.amazonaws.com",
    "eventName": "StartSnapshot",
    "awsRegion": "eu-west-1",
    "sourceIPAddress": "192.0.2.0",
    "userAgent": "PostmanRuntime/7.25.0",
    "requestParameters": {
        "volumeSize": 8,
        "clientToken": "token",
```

```
        "encrypted": true
    },
    "responseElements": {
        "snapshotId": "snap-123456789012",
        "ownerId": "123456789012",
        "status": "pending",
        "startTime": "Jul 3, 2020 11:27:26 PM",
        "volumeSize": 8,
        "blockSize": 524288,
        "kmsKeyArn": "HIDDEN_DUE_TO_SECURITY_REASONS"
    },
    "requestID": "be112233-1ba5-4ae0-8e2b-1c302EXAMPLE",
    "eventID": "6e12345-2a4e-417c-aa78-7594fEXAMPLE",
    "eventType": "AwsApiCall",
    "recipientAccountId": "123456789012"
}
```

Exemple avec CompleteSnapshot :

```
{
    "eventVersion": "1.05",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "123456789012",
        "arn": "arn:aws:iam::123456789012:root",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "userName": "user"
    },
    "eventTime": "2020-07-03T23:28:24Z",
    "eventSource": "ebs.amazonaws.com",
    "eventName": "CompleteSnapshot",
    "awsRegion": "eu-west-1",
    "sourceIPAddress": "192.0.2.0",
    "userAgent": "PostmanRuntime/7.25.0",
    "requestParameters": {
        "snapshotId": "snap-123456789012",
        "changedBlocksCount": 5
    },
    "responseElements": {
        "status": "completed"
    },
    "requestID": "be112233-1ba5-4ae0-8e2b-1c302EXAMPLE",
    "eventID": "6e12345-2a4e-417c-aa78-7594fEXAMPLE",
    "eventType": "AwsApiCall",
    "recipientAccountId": "123456789012"
}
```

API directes EBS et points de terminaison de VPC d'interface

Vous pouvez établir une connexion privée entre votre VPC et API directes EBS en créant un point de terminaison de VPC d'interface. Les points de terminaison d'interface sont alimentés par AWS PrivateLink, une technologie qui vous permet d'accéder en privé à API directes EBS sans passerelle Internet, périphérique NAT, connexion VPN ou connexion AWS Direct Connect. Les instances de votre VPC ne nécessitent pas d'adresses IP publiques pour communiquer avec API directes EBS. Le trafic entre votre VPC et API directes EBS ne quitte pas le réseau Amazon.

Chaque point de terminaison d'interface est représenté par une ou plusieurs interfaces réseau Elastic dans vos sous-réseaux.

Pour de plus amples informations, veuillez consulter Points de terminaison d'un VPC d'interface (AWS PrivateLink) dans le Guide de l'utilisateur Amazon VPC.

Considérations relatives aux points de terminaison de VPC API directes EBS

Avant de configurer un point de terminaison de VPC d'interface pour API directes EBS, assurez-vous de vérifier les propriétés et limitations du point de terminaison d'interface dans le Amazon VPC Guide de l'utilisateur.

Les stratégies de point de terminaison de VPC sont prises en charge pour API directes EBS. Par défaut, l'accès complet à API directes EBS est autorisé via le point de terminaison. Toutefois, vous pouvez contrôler l'accès au point de terminaison de l'interface à l'aide de groupes de sécurité. Pour de plus amples informations, veuillez consulter Contrôle de l'accès aux services avec points de terminaison d'un VPC dans le Amazon VPC Guide de l'utilisateur.

Création d'un point de terminaison de VPC d'interface pour API directes EBS

Vous pouvez créer un point de terminaison de VPC pour le service API directes EBS à l'aide de la console Amazon VPC ou de l'AWS Command Line Interface (AWS CLI). Pour de plus amples informations, veuillez consulter Création d'un point de terminaison d'interface dans le Amazon VPC Guide de l'utilisateur.

Pour créer un point de terminaison de VPC pour API directes EBS, utilisez le nom de service suivant :

- `com.amazonaws.`*`region`*`.ebs`

Si vous activez le DNS privé pour le point de terminaison, vous pouvez faire des demandes d'API à API directes EBS en utilisant son nom DNS par défaut pour la région, par exemple `ebs.us-east-1.amazonaws.com`. Pour de plus amples informations, veuillez consulter Accès à un service via un point de terminaison d'interface dans le Amazon VPC Guide de l'utilisateur.

Idempotence pour l'API StartSnapshot

L'idempotence garantit qu'une requête API n'est exécutée qu'une seule fois. Avec une demande idempotente, si la demande d'origine se termine correctement. Les tentatives suivantes renvoient le résultat de la demande d'origine réussie et elles n'ont aucun effet supplémentaire.

L'API StartSnapshot prend en charge l'idempotence en utilisant un jeton client. Un jeton client est une chaîne unique que vous spécifiez lorsque vous effectuez une demande d'API. Si vous réessayez une demande d'API avec le même jeton client et les mêmes paramètres de requête une fois qu'elle est terminée correctement, le résultat de la demande d'origine est renvoyé. Si vous réessayez une demande avec le même jeton client, mais que vous modifiez un ou plusieurs paramètres de requête, l'erreur `ConflictException` est renvoyée.

Si vous ne spécifiez pas votre propre jeton client, les kits AWS SDK génèrent automatiquement un jeton client pour la requête afin de s'assurer qu'il est idempotent.

Un jeton client peut être n'importe quelle chaîne qui comprend jusqu'à 64 caractères ASCII. Vous ne devez pas réutiliser les mêmes jetons client pour différentes demandes.

Pour faire une demande StartSnapshot idempotente avec votre propre jeton client à l'aide de l'API

Spécifiez le paramètre de demande `ClientToken`.

```
POST /snapshots HTTP/1.1
Host: ebs.us-east-2.amazonaws.com
Accept-Encoding: identity
User-Agent: <User agent parameter>
X-Amz-Date: 20200618T040724Z
Authorization: <Authentication parameter>

{
    "VolumeSize": 8,
    "ParentSnapshot": snap-123EXAMPLE1234567,
    "ClientToken": "550e8400-e29b-41d4-a716-446655440000",
```

```
    "Timeout": 60
}
```

Pour effectuer une demande StartSnapshot idempotente avec votre propre jeton client à l'aide de la commande AWS CLI

Spécifiez le paramètre de demande `client-token`.

```
$ aws ebs start-snapshot --region us-east-2 --volume-size 8 --parent-snapshot
 snap-123EXAMPLE1234567 --timeout 60 --client-token 550e8400-e29b-41d4-a716-446655440000
```

Automatisation du cycle de vie des instantanés Amazon EBS

Vous pouvez utiliser Amazon Data Lifecycle Manager pour automatiser la création, la conservation et la suppression des instantanés utilisés pour sauvegarder vos volumes Amazon EBS. Lorsque vous automatisez la gestion des instantanés, cela vous aide à :

- protéger les données importantes en appliquant un planning de sauvegarde régulière ;
- conserver des sauvegardes comme exigé par les auditeurs ou les réglementations internes ;
- réduire les frais de stockage en supprimant les sauvegardes périmées.

Associé aux fonctions de surveillance d'Amazon CloudWatch Events et d'AWS CloudTrail, Amazon Data Lifecycle Manager constitue une solution de sauvegarde complète pour les volumes EBS sans frais supplémentaires.

Sommaire

- Fonctionnement de Amazon Data Lifecycle Manager (p. 1169)
- Planifications de stratégie (p. 1170)
- Considérations relatives à Amazon Data Lifecycle Manager (p. 1171)
- Prérequis (p. 1172)
- Gérer les sauvegardes à l'aide de la console (p. 1174)
- Gérer des sauvegardes à l'aide de l'AWS CLI (p. 1177)
- Gérer les sauvegardes à l'aide de l'API (p. 1180)
- Surveiller le cycle de vie des instantanés (p. 1180)

Fonctionnement de Amazon Data Lifecycle Manager

Voici les éléments de base de Amazon Data Lifecycle Manager.

Eléments

- Instantanés (p. 1169)
- Target resource tags (Balises de ressource cibles) (p. 1170)
- Balises d'instantané (p. 1170)
- Stratégies de cycle de vie (p. 1170)

Instantanés

Les instantanés représentent le principal moyen de sauvegarde des données de vos volumes EBS. Afin d'économiser les frais de stockage, les instantanés successifs sont incrémentiels ; ils contiennent uniquement les données du volume ayant changé depuis l'instantané précédent. Lorsque vous supprimez un instantané dans une série d'instantanés d'un volume, seules les données figurant uniquement dans cet instantané sont supprimées. Le reste de l'historique de capture du volume est conservé.

Pour plus d'informations, consultez Instantanés Amazon EBS (p. 1129).

Target resource tags (Balises de ressource cibles)

Amazon Data Lifecycle Manager utilise des balises de ressource pour identifier les volumes EBS à sauvegarder. Les balises sont des métadonnées personnalisables que vous pouvez affecter à vos ressources AWS (y compris les volumes et instantanés EBS). Une stratégie Amazon Data Lifecycle Manager (décrite ci-après) cible un volume pour la sauvegarde en utilisant une seule balise. Il est possible d'affecter plusieurs balises à un volume si vous voulez exécuter plusieurs stratégies sur celui-ci.

Vous ne pouvez pas utiliser un caractère « \ » ou « = » dans une clé de balise.

Pour plus d'informations, consultez Balisage de vos ressources Amazon EC2 (p. 1309).

Balises d'instantané

Amazon Data Lifecycle Manager applique les balises suivantes à tous les instantanés créés par une stratégie, afin de les distinguer des instantanés créés par d'autres moyens :

- `aws:dlm:lifecycle-policy-id`
- `aws:dlm:lifecycle-schedule-name`
- `aws:dlm:expirationTime`
- `dlm:managed`

Vous pouvez également spécifier des balises personnalisées à appliquer aux instantanés au moment de la création.

Vous ne pouvez pas utiliser un caractère « \ » ou « = » dans une clé de balise.

Les balises cible utilisées par Amazon Data Lifecycle Manager pour associer des volumes avec une stratégie peuvent éventuellement être appliquées aux instantanés créés par la stratégie.

Stratégies de cycle de vie

Une stratégie de cycle de vie se compose des paramètres de base suivants :

- Type de ressource : définit le type de ressource AWS gérée par la stratégie. Utilisez `VOLUME` pour créer des instantanés de volumes individuels ou `INSTANCE` pour créer des instantanés multi-volumes à partir des volumes d'une instance. Pour de plus amples informations, veuillez consulter Instantanés multi-volumes (p. 1133).
- Balises cible : définit les balises qui doivent être associées à un volume EBS ou à une instance Amazon EC2 pour être gérées par la stratégie.
- Planifications : heures de début et intervalles de création d'instantanés. Le premier instantané est créé par une stratégie dans l'heure suivant l'heure de début spécifiée. Les instantanés suivants sont créés dans l'heure suivant leur heure prévue. Une stratégie peut comporter jusqu'à quatre plannings ; un planning obligatoire et jusqu'à trois plannings facultatifs. Pour de plus amples informations, veuillez consulter Planifications de stratégie (p. 1170).
- Conservation :spécifie la façon dont les instantanés sont conservés. Vous pouvez conserver des instantanés en fonction du nombre total d'instantanés ou de l'âge de chaque instantané.

Par exemple, vous pouvez créer une stratégie qui gère tous les volumes EBS dotés d'une balise avec une clé `account` et une valeur `finance`, crée des instantanés toutes les 24 heures à 9h00 UTC et conserve les cinq instantanés les plus récents. La création d'instantanés commencerait à 0959 chaque jour.

Planifications de stratégie

Les planifications de stratégie définissent le moment où les instantanés sont créés par la stratégie. Les stratégies peuvent comporter jusqu'à quatre planifications, une obligatoire et jusqu'à trois facultatives.

L'ajout de plusieurs planifications pour une seule stratégie vous permet de créer des instantanés à différentes fréquences à l'aide de la même stratégie. Par exemple, vous pouvez créer une stratégie unique qui crée des instantanés quotidiens, hebdomadaires, mensuels et annuels. Cela vous évite de devoir gérer plusieurs stratégies.

Pour chaque planification, vous pouvez définir la fréquence, les paramètres de restauration d'instantané rapide, les règles de copie entre régions et les balises. Les balises affectées à une planification sont automatiquement affectées aux instantanés créés lors du déclenchement de la planification. En outre, Amazon Data Lifecycle Manager attribue automatiquement une balise générée par le système à chaque instantané en fonction de la fréquence de la planification.

Chaque planification est déclenchée individuellement en fonction de sa fréquence. Si plusieurs planifications sont déclenchées en même temps, Amazon Data Lifecycle Manager crée un seul instantané et utilise les paramètres de conservation d'instantané de la planification dont la période de conservation est la plus élevée. Les balises de toutes les planifications déclenchées sont appliquées à l'instantané.

- Si la restauration d'instantané rapide est activée pour plusieurs planifications déclenchées, l'instantané est activé pour la restauration d'instantané rapide dans toutes les zones de disponibilité spécifiées parmi toutes les planifications déclenchées et les paramètres de conservation les plus élevés des planifications déclenchées sont utilisés pour chaque zone de disponibilité.
- Si la copie entre régions est activée pour plusieurs planifications déclenchées, l'instantané est copié dans toutes les régions spécifiées parmi les planifications déclenchées et la période de conservation la plus élevée des planifications déclenchées est utilisée.

Considérations relatives à Amazon Data Lifecycle Manager

Les quotas de votre compte AWS concernant Amazon Data Lifecycle Manager sont les suivantes :

- Vous pouvez créer jusqu'à 100 stratégies de cycle de vie par région.
- Vous pouvez ajouter jusqu'à 45 balises par ressource.
- Vous pouvez créer jusqu'à quatre planifications par stratégie de cycle de vie.

Les considérations suivantes s'appliquent aux stratégies de cycle de vie :

- Une stratégie ne commence pas la création des instantanés tant que vous n'avez pas défini son statut d'activation sur activé. Vous pouvez configurer une stratégie pour être activée dès sa création.
- Le premier instantané est créé par une stratégie dans l'heure suivant l'heure de début spécifiée. Les instantanés suivants sont créés dans l'heure suivant leur heure prévue.
- Si vous modifiez une stratégie en supprimant ou en modifiant sa balise cible, les volumes EBS comportant cette balise ne sont plus affectés par la stratégie.
- Si vous modifiez le nom du planning d'une stratégie, les instantanés créés sous l'ancien nom de planning ne sont plus affectés par la stratégie.
- Si vous modifiez un programme de conservation (aussi appelé programme de rétention) en fonction du délai au-delà duquel utiliser un nouvel intervalle, le nouvel intervalle est uniquement utilisé pour les nouveaux instantanés. Le nouveau programme n'affecte pas le programme de conservation des instantanés existants créés par cette stratégie.
- Vous ne pouvez pas modifier le programme de conservation d'une stratégie pour passer du nombre d'instantanés à l'âge de chaque instantané. Pour pouvoir effectuer ce changement, vous devez créer une nouvelle stratégie.
- Si vous désactivez une stratégie dont le programme de conservation se base sur l'âge de chaque instantané, les instantanés dont les périodes de conservation expirent alors que la stratégie est désactivée sont conservés indéfiniment. Vous devez supprimer ces instantanés manuellement. Lorsque vous réactivez la stratégie, Amazon Data Lifecycle Manager reprend la suppression des instantanés à mesure que leurs périodes de conservation expirent.

- Si vous supprimez la ressource à laquelle une stratégie avec rétention basée sur un nombre s'applique, cette dernière ne gère plus les instantanés créés précédemment. Vous devez supprimer les instantanés manuellement lorsqu'ils ne sont plus nécessaires.
- Si vous supprimez la ressource à laquelle s'applique une stratégie de rétention basée sur l'âge, la stratégie continue de supprimer des instantanés sur la planification définie, jusqu'au dernier instantané. Vous devez supprimer manuellement le dernier instantané s'il n'est plus nécessaire.
- Vous pouvez créer plusieurs stratégies pour sauvegarder un volume EBS ou une instance Amazon EC2. Par exemple, si un volume EBS comporte deux balises, la balise A étant la cible de la stratégie A qui permet de créer un instantané toutes les 12 heures et la balise B étant la cible de la stratégie B qui permet de créer un instantané toutes les 24 heures, Amazon Data Lifecycle Manager crée des instantanés en fonction des plannings des deux stratégies.

Les considérations suivantes s'appliquent aux stratégies de cycle de vie et à la restauration d'instantané rapide (p. 1205) :

- Un instantané qui est activé pour la restauration d'instantané rapide reste activé, même si vous supprimez ou désactivez la stratégie de cycle de vie, ou si vous désactivez la restauration d'instantané rapide pour la stratégie de cycle de vie ou pour la zone de disponibilité. Vous pouvez désactiver manuellement la restauration d'instantané rapide pour ces instantanés.
- Si vous activez la restauration d'instantané rapide et que vous dépassez le nombre maximum d'instantanés pouvant être activés pour la restauration d'instantané rapide, Amazon Data Lifecycle Manager crée des instantanés comme prévu, mais ne les active pas pour la restauration d'instantané rapide. Une fois qu'un instantané activé pour la restauration d'instantané rapide est supprimé, l'instantané suivant créé par Amazon Data Lifecycle Manager est activé pour la restauration rapide d'instantané.
- Lorsque vous activez la restauration d'instantané rapide pour un instantané, l'optimisation de ce dernier dure 60 minutes par Tio. Nous vous recommandons de créer un programme qui assure l'optimisation complète de chaque instantané avant que Amazon Data Lifecycle Manager ne crée l'instantané suivant.
- Vous êtes facturé pour chaque minute pendant laquelle la restauration d'instantané rapide est activée pour un instantané dans une zone de disponibilité particulière. Les frais sont calculés au prorata avec un minimum d'une heure. Pour plus d'informations, consultez Tarification et facturation (p. 1209).

  Note

  Selon la configuration de vos stratégies de cycle de vie, plusieurs instantanés peuvent être activés pour une restauration rapide simultanée.

Les considérations suivantes s'appliquent aux stratégies de cycle de vie et aux volumes activés pour Multi-Attach (p. 1110) :

- Lors de la création d'une stratégie de cycle de vie basée sur des balises d'instance pour les instantanés multi-volumes, Amazon Data Lifecycle Manager lance un instantané du volume pour chaque instance attachée. Utilisez la balise timestamp pour identifier l'ensemble d'instantanés temporels créés à partir des instances attachées.

Prérequis

Les conditions préalables suivantes sont requises par Amazon Data Lifecycle Manager.

Conditions préalables requises

- Autorisations d'Amazon Data Lifecycle Manager (p. 1173)
- Autorisations pour les utilisateurs IAM (p. 1174)
- Autorisations pour les instantanés chiffrés (p. 1174)

Autorisations d'Amazon Data Lifecycle Manager

Amazon Data Lifecycle Manager utilise un rôle IAM pour obtenir les autorisations requises pour gérer les instantanés en votre nom. Amazon Data Lifecycle Manager crée AWSDataLifecycleManagerDefaultRole la première fois que vous créez une stratégie de cycle de vie avec AWS Management Console. Vous pouvez aussi créer ce rôle en utilisant la commande create-default-role suivante.

```
aws dlm create-default-role
```

Sinon, vous pouvez créer un rôle IAM personnalisé avec les autorisations requises et le sélectionner lors de la création d'une stratégie de cycle de vie.

Pour créer un rôle IAM personnalisé

1. Créez un rôle avec les autorisations suivantes.

```
{
   "Version": "2012-10-17",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": [
            "ec2:CreateSnapshot",
            "ec2:CreateSnapshots",
            "ec2:DeleteSnapshot",
            "ec2:DescribeVolumes",
            "ec2:DescribeInstances",
            "ec2:DescribeSnapshots"
         ],
         "Resource": "*"
      },
      {
         "Effect": "Allow",
         "Action": [
            "ec2:CreateTags"
         ],
         "Resource": "arn:aws:ec2:*::snapshot/*"
      }
   ]
}
```

   Pour plus d'informations, consultez Création d'un rôle dans le IAM Guide de l'utilisateur.

2. Ajoutez une relation d'approbation au rôle.

   a. Dans la console IAM, choisissez Rôles.
   b. Sélectionnez le rôle que vous venez de créer, puis choisissez Relations d'approbation.
   c. Choisissez Modifier la relation d'approbation, ajoutez la stratégie suivante, puis choisissez Mettre à jour la stratégie d'approbation.

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

```
}
```

Autorisations pour les utilisateurs IAM

Un utilisateur IAM doit avoir les autorisations suivantes pour pouvoir utiliser Amazon Data Lifecycle Manager.

```
{
   "Version": "2012-10-17",
   "Statement": [
      {
         "Effect": "Allow",
         "Action": ["iam:PassRole", "iam:ListRoles"],
         "Resource": "arn:aws:iam::123456789012:role/AWSDataLifecycleManagerDefaultRole"
      },
      {
         "Effect": "Allow",
         "Action": "dlm:*",
         "Resource": "*"
      }
   ]
}
```

Pour plus d'informations, consultez Modification des autorisations pour un utilisateur IAM dans le IAM Guide de l'utilisateur.

Autorisations pour les instantanés chiffrés

Pour copier un instantané chiffré entre les régions, vous devez avoir accès à la clé principale du client source et de destination (CMK) à partir de AWS Key Management Service (AWS KMS).

Si le volume source est chiffré, assurez-vous que AWSDataLifecycleManagerDefaultRole est autorisé à utiliser la clé CMK utilisée pour chiffrer le volume. Si vous activez la copie inter-région et choisissez de chiffrer l'instantané copié, assurez-vous que AWSDataLifeCycleManagerDefaultRole est autorisé à utiliser la clé CMK nécessaire pour chiffrer l'instantané dans la région de destination. Pour plus d'informations, consultez Gestion de l'accès aux clés CMK AWS KMS dans le AWS Key Management Service Developer Guide.

Gérer les sauvegardes à l'aide de la console

Les exemples suivants montrent comment utiliser Amazon Data Lifecycle Manager pour gérer les sauvegardes de vos volumes EBS à l'aide de l'AWS Management Console.

Tâches

- Créer une stratégie de cycle de vie (p. 1174)
- Afficher une stratégie de cycle de vie (p. 1176)
- Modifier une stratégie de cycle de vie (p. 1176)
- Supprimer une stratégie de cycle de vie (p. 1176)

Créer une stratégie de cycle de vie

Utilisez la procédure suivante pour créer une stratégie de cycle de vie.

Pour créer une stratégie de cycle de vie

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, choisissez Elastic Block Store, Lifecycle Manager (Gestionnaire de cycle de vie), puis Create snapshot lifecycle policy (Créer une stratégie de cycle de vie d'instantané).
3. Fournissez si nécessaire les informations suivantes pour votre stratégie :
   - Description – Description de la stratégie.
   - Resource type (Type de ressource) – Type de ressource à sauvegarder. Utilisez `Volume` pour créer des instantanés de volumes individuels ou `Instance` pour créer des instantanés multi-volumes à partir des volumes d'une instance.
   - Target with tags (Cible avec balises) – Balises de ressource qui identifient les volumes ou les instances à sauvegarder.
   - Lifecycle policy tags (Balises de stratégie de cycle de vie) – Balises de la stratégie de cycle de vie.
4. Pour IAM role (Rôle IAM), choisissez le rôle IAM qui possède les autorisations de créer, supprimer et décrire les instantanés, et de décrire les volumes. AWS fournit un rôle par défaut, AWSDataLifecycleManagerDefaultRole. Vous pouvez aussi créer un rôle IAM personnalisé.
5. Ajoutez les planifications de stratégie. La planification 1 est obligatoire. Les planifications 2, 3 et 4 sont facultatives. Pour chaque planification de stratégie, spécifiez les informations suivantes :
   - Schedule name (Nom du programme) – Nom du programme.
   - Fréquence – Intervalle entre les cycles de la stratégie. Vous pouvez configurer les exécutions de stratégie selon une planification quotidienne, hebdomadaire, mensuelle ou annuelle. Vous pouvez également choisir Expression cron personnalisée pour spécifier un intervalle allant jusqu'à 1 an. Pour de plus amples informations, veuillez consulter Expressions Cron dans le Guide de l'utilisateur Amazon CloudWatch Events.
   - Démarrage à hh:mm UTC – Heure de démarrage programmée pour les cycles de la stratégie. La première exécution de stratégie commence dans l'heure qui suit l'heure programmée.
   - Type de conservation – Vous pouvez conserver des instantanés en fonction du nombre total d'instantanés ou de l'âge de chaque instantané. Pour la conservation basée sur le nombre d'instantanés, la plage est comprise entre 1 et 1 000. Une fois le nombre maximum atteint, l'instantané le plus ancien est supprimé lorsqu'un nouvel instantané est créé. Pour la conservation basée sur l'âge, la plage est comprise entre 1 jour et 100 ans. Une fois la période de conservation de chaque instantané expirée, ce dernier est supprimé. La période de conservation doit être supérieure ou égale à l'intervalle de création.

     Note

     Toutes les planifications doivent avoir le même type de conservation. Vous pouvez spécifier le type de conservation pour la planification 1 uniquement. Les planifications 2, 3 et 4 héritent du type de conservation de la planification 1. Chaque planification peut avoir son propre nombre ou sa propre période de conservation.
   - Tagging Information (Balisage des informations) – Indiquez si vous souhaitez copier toutes les balises définies par l'utilisateur sur un volume source vers des instantanés du volume créés par cette stratégie. Vous pouvez également spécifier des balises supplémentaires pour les instantanés, en plus des balises appliquées par Amazon Data Lifecycle Manager. Si le type de ressource est `INSTANCE`, vous pouvez choisir de baliser automatiquement vos instantanés avec les balises de variable suivantes : `instance-id` et `timestamp`. Les valeurs des balises de variable sont déterminées lorsque les balises sont ajoutées.
   - Restauration d'instantané rapide – Indiquez si vous souhaitez activer la restauration d'instantané rapide pour tous les instantanés créés par cette stratégie. Si vous activez la restauration d'instantané rapide, vous devez choisir les zones de disponibilité dans lesquelles le faire. Vous êtes facturé pour chaque minute pendant laquelle la restauration d'instantané rapide est activée pour un instantané dans une zone de disponibilité particulière. Les frais sont calculés au prorata avec un minimum d'une heure. Vous pouvez également spécifier le nombre maximum d'instantanés pouvant être activés pour la restauration d'instantané rapide.
   - Copie inter-régions – Vous pouvez copier chaque instantané sur trois régions supplémentaires maximum. Vous devez vous assurer que vous ne dépassez pas le nombre de copies d'instantanés

simultanées par région. Pour chaque région, vous pouvez choisir différentes stratégies de conservation et indiquer s'il convient de copier toutes les balises ou de n'en copier aucune. Si l'instantané source est chiffré ou si le chiffrement est activé par défaut, les copies d'instantané sont chiffrées. Si l'instantané source n'est pas chiffré, vous pouvez activer le chiffrement. Si vous ne spécifiez pas de clé CMK, les instantanés sont chiffrés à l'aide de la clé par défaut pour le chiffrement EBS dans chaque région de destination. Si vous spécifiez une clé CMK pour la région de destination, vous devez avoir accès à la clé CMK.

6. Pour Policy status after creation (Statut de la stratégie après création), choisissez Enable policy (Activer la stratégie) pour lancer les exécutions de stratégie lors de la prochaine heure planifiée ou Disable policy (Désactiver la stratégie) pour empêcher l'exécution de la stratégie.
7. Choisissez Créer une stratégie.

Afficher une stratégie de cycle de vie

Utilisez la procédure suivante pour afficher une stratégie de cycle de vie.

Pour afficher une stratégie de cycle de vie

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Elastic Block Store, Lifecycle Manager (Gestionnaire de cycle de vie).
3. Sélectionnez une stratégie de cycle de vie dans la liste. L'onglet Details (Détails) affiche des informations sur la stratégie.

Modifier une stratégie de cycle de vie

Utilisez la procédure suivante pour modifier une stratégie de cycle de vie.

Pour modifier une stratégie de cycle de vie

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Elastic Block Store, Lifecycle Manager (Gestionnaire de cycle de vie).
3. Sélectionnez une stratégie de cycle de vie dans la liste.
4. Choisissez Actions, Modify Snapshot Lifecycle Policy (Modifier la stratégie de cycle de vie de l'instantané).
5. Modifiez les paramètres de stratégie selon vos besoins. Par exemple, vous pouvez modifier le programme, ajouter ou supprimer des balises, ou encore activer ou désactiver la stratégie.
6. Choisissez Mettre à jour une stratégie.

Supprimer une stratégie de cycle de vie

Suivez la procédure suivante pour supprimer une stratégie de cycle de vie.

> **Note**
>
> Vous pouvez supprimer les instantanés créés uniquement par Amazon Data Lifecycle Manager.

Pour supprimer une stratégie de cycle de vie

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Elastic Block Store, Lifecycle Manager (Gestionnaire de cycle de vie).

3. Sélectionnez une stratégie de cycle de vie dans la liste.
4. Choisissez Actions, Delete Snapshot Lifecycle Policy (Supprimer la stratégie de cycle de vie de l'instantané).
5. Lorsque vous êtes invité à confirmer l'opération, choisissez Delete Snapshot Lifecycle Policy (Supprimer la stratégie de cycle de vie de l'instantané).

Gérer des sauvegardes à l'aide de l'AWS CLI

Les exemples suivants montrent comment utiliser Amazon Data Lifecycle Manager pour gérer les sauvegardes de vos volumes EBS à l'aide de l'AWS CLI.

Exemples

- Créer une stratégie de cycle de vie (p. 1177)
- Afficher une stratégie de cycle de vie (p. 1178)
- Modifier une stratégie de cycle de vie (p. 1179)
- Supprimer une stratégie de cycle de vie (p. 1180)

Créer une stratégie de cycle de vie

Utilisez la commande create-lifecycle-policy pour créer une stratégie de cycle de vie. Pour simplifier la syntaxe, cet exemple utilise un fichier JSON, `policyDetails.json`, qui inclut les détails de la stratégie.

Cet exemple utilise le type de ressource `VOLUME` pour créer des instantanés de tous les volumes avec les balises cibles spécifiées. Pour créer des instantanés de tous les volumes pour toutes les instances avec les balises cibles spécifiées, utilisez plutôt un type de ressource `INSTANCE`. La stratégie comprend deux planifications. La première planification crée un instantané tous les jours à 3h00 UTC. La deuxième planification crée un instantané hebdomadaire tous les vendredis à 17h00 UTC.

```
aws dlm create-lifecycle-policy --description "My volume policy" --state ENABLED --
execution-role-arn arn:aws:iam::12345678910:role/AWSDataLifecycleManagerDefaultRole --
policy-details file://policyDetails.json
```

Voici un exemple du fichier `policyDetails.json`.

```
{
 "ResourceTypes": [
  "VOLUME"
 ],
 "TargetTags": [{
  "Key": "costcenter",
  "Value": "115"
 }],
 "Schedules": [{
   "Name": "DailySnapshots",
   "TagsToAdd": [{
    "Key": "type",
    "Value": "myDailySnapshot"
   }],
   "CreateRule": {
    "Interval": 24,
    "IntervalUnit": "HOURS",
    "Times": [
     "03:00"
    ]
   },
   "RetainRule": {
```

```
    "Count": 5
   },
   "CopyTags": false
  },
  {
   "Name": "WeeklySnapshots",
   "TagsToAdd": [{
    "Key": "type",
    "Value": "myWeeklySnapshot"
   }],
   "CreateRule": {
    "CronExpression": "cron(0 0 17 ? * FRI *)"
   },
   "RetainRule": {
    "Count": 5
   },
   "CopyTags": false
  }
 ]
}
```

En cas de réussite, la commande renvoie l'ID de la nouvelle stratégie créée. Voici un exemple de sortie.

```
{
   "PolicyId": "policy-0123456789abcdef0"
}
```

Afficher une stratégie de cycle de vie

Utilisez la commande get-lifecycle-policy pour afficher des informations sur une stratégie de cycle de vie.

```
aws dlm get-lifecycle-policy --policy-id policy-0123456789abcdef0
```

Voici un exemple de sortie. Elle contient les informations que vous avez spécifiées, ainsi que les métadonnées insérées par AWS.

```
{
   "Policy":{
      "Description": "My first policy",
      "DateCreated": "2018-05-15T00:16:21+0000",
      "State": "ENABLED",
      "ExecutionRoleArn":
 "arn:aws:iam::210774411744:role/AWSDataLifecycleManagerDefaultRole",
      "PolicyId": "policy-0123456789abcdef0",
      "DateModified": "2018-05-15T00:16:22+0000",
      "PolicyDetails": {
        "PolicyType":"EBS_SNAPSHOT_MANAGEMENT",
         "ResourceTypes": [
            "VOLUME"
         ],
         "TargetTags": [
            {
               "Value": "115",
               "Key": "costcenter"
            }
         ],
         "Schedules": [
            {
               "TagsToAdd": [
                  {
                     "Value": "myDailySnapshot",
```

```
                    "Key": "type"
                }
            ],
            "RetainRule": {
                "Count": 5
            },
            "CopyTags": false,
            "CreateRule": {
                "Interval": 24,
                "IntervalUnit": "HOURS",
                "Times": [
                    "03:00"
                ]
            },
            "Name": "DailySnapshots"
        }
    ]
  }
 }
}
```

Modifier une stratégie de cycle de vie

Utilisez la commande update-lifecycle-policy pour modifier les informations dans une stratégie de cycle de vie. Pour simplifier la syntaxe, cet exemple fait référence à un fichier JSON, `policyDetailsUpdated.json`, qui inclut les détails de la stratégie.

```
aws dlm update-lifecycle-policy --state DISABLED --execution-role-arn
 arn:aws:iam::12345678910:role/AWSDataLifecycleManagerDefaultRole" --policy-details
 file://policyDetailsUpdated.json
```

Voici un exemple du fichier `policyDetailsUpdated.json`.

```
{
   "ResourceTypes":[
      "VOLUME"
   ],
   "TargetTags":[
      {
         "Key": "costcenter",
         "Value": "120"
      }
   ],
   "Schedules":[
      {
         "Name": "DailySnapshots",
         "TagsToAdd": [
            {
               "Key": "type",
               "Value": "myDailySnapshot"
            }
         ],
         "CreateRule": {
            "Interval": 12,
            "IntervalUnit": "HOURS",
            "Times": [
               "15:00"
            ]
         },
         "RetainRule": {
            "Count" :5
         },
         "CopyTags": false
```

```
        }
    ]
}
```

Pour afficher la stratégie mise à jour, utilisez la commande `get-lifecycle-policy`. Vous pouvez voir que l'état, la valeur de la balise, l'intervalle de prise d'instantané et l'heure de début de la prise d'instantané ont été modifiés.

Supprimer une stratégie de cycle de vie

Utilisez la commande delete-lifecycle-policy pour supprimer une stratégie de cycle de vie et libérer les balises cible spécifiées dans la stratégie afin de pouvoir les réutiliser.

> **Note**
>
> Vous pouvez supprimer les instantanés créés uniquement par Amazon Data Lifecycle Manager.

```
aws dlm delete-lifecycle-policy --policy-id policy-0123456789abcdef0
```

Gérer les sauvegardes à l'aide de l'API

Le manuel Référence d'API Amazon Data Lifecycle Manager contient des descriptions et la syntaxe de chacune des actions et chacun des types de données de l'API de requête Amazon Data Lifecycle Manager.

Vous pouvez aussi utiliser l'un des kits SDK AWS pour accéder à l'API d'une façon adaptée au langage de programmation ou à la plateforme que vous utilisez. Pour de plus amples informations, veuillez consulter SDK AWS.

Surveiller le cycle de vie des instantanés

Vous pouvez utiliser les fonctions suivantes pour surveiller le cycle de vie de vos instantanés.

Fonctions

- Console et AWS CLI (p. 1180)
- CloudWatch Events (p. 1180)
- AWS CloudTrail (p. 1181)

Console et AWS CLI

Vous pouvez afficher vos stratégies de cycle de vie à l'aide de la console Amazon EC2 ou de l'AWS CLI. Chaque instantané créé par une stratégie possède un horodatage et des balises liées à la stratégie. Vous pouvez filtrer les instantanés à l'aide des balises afin de vérifier que vos sauvegardes sont créées comme vous l'avez souhaité. Pour plus d'informations sur l'affichage des stratégies de cycle de vie à l'aide de la console, consultez Afficher une stratégie de cycle de vie (p. 1176). Pour plus d'informations sur l'affichage d'informations sur les stratégies de cycle de vie à l'aide de l'interface de ligne de commande, consultez Afficher une stratégie de cycle de vie (p. 1178).

CloudWatch Events

Amazon EBS et Amazon Data Lifecycle Manager émettent des événements liés aux actions de la stratégie de cycle de vie. Vous pouvez utiliser AWS Lambda et Amazon CloudWatch Events pour gérer par programmation les notifications d'événement. Pour plus d'informations, consultez le Guide de l'utilisateur Amazon CloudWatch Events.

Les événements suivants sont disponibles :

- `createSnapshot` — Événement Amazon EBS émis en cas de réussite ou d'échec d'une action `CreateSnapshot`. Pour de plus amples informations, veuillez consulter Amazon CloudWatch Events pour Amazon EBS (p. 1253).
- `DLM Policy State Change` — Événement Amazon Data Lifecycle Manager émis lorsqu'une stratégie de cycle de vie passe en mode erreur. L'événement contient une description de la cause de l'erreur. Vous trouverez ci-après un exemple d'événement émis lorsque les autorisations accordées par le rôle IAM sont insuffisantes.

```
{
   "version": "0",
   "id": "01234567-0123-0123-0123-0123456789ab",
   "detail-type": "DLM Policy State Change",
   "source": "aws.dlm",
   "account": "123456789012",
   "time": "2018-05-25T13:12:22Z",
   "region": "us-east-1",
   "resources": [
      "arn:aws:dlm:us-east-1:123456789012:policy/policy-0123456789abcdef"
   ],
   "detail": {
      "state": "ERROR",
      "cause": "Role provided does not have sufficient permissions",
      "policy_id": "arn:aws:dlm:us-east-1:123456789012:policy/policy-0123456789abcdef"
   }
}
```

Voici un exemple d'événement émis lorsqu'une limite est dépassée.

```
{
   "version": "0",
   "id": "01234567-0123-0123-0123-0123456789ab",
   "detail-type": "DLM Policy State Change",
   "source": "aws.dlm",
   "account": "123456789012",
   "time": "2018-05-25T13:12:22Z",
   "region": "us-east-1",
   "resources": [
      "arn:aws:dlm:us-east-1:123456789012:policy/policy-0123456789abcdef"
   ],
   "detail":{
      "state": "ERROR",
      "cause": "Maximum allowed active snapshot limit exceeded",
      "policy_id": "arn:aws:dlm:us-east-1:123456789012:policy/policy-0123456789abcdef"
   }
}
```

AWS CloudTrail

AWS CloudTrail vous permet de suivre l'activité utilisateur et l'utilisation de l'API afin d'apporter la preuve de la conformité des stratégies internes et de la réglementation en vigueur. Pour plus d'informations, consultez la AWS CloudTrail User Guide.

Services de données Amazon EBS

Amazon EBS fournit les services de données suivants :

Services de données

- Amazon EBS Elastic Volumes (p. 1182)

- Chiffrement Amazon EBS (p. 1194)
- Restauration d'instantané rapide Amazon EBS (p. 1205)

Amazon EBS Elastic Volumes

Amazon EBS Elastic Volumes vous permet d'augmenter la taille du volume, de changer le type de volume ou d'ajuster les performances de vos volumes EBS. Si votre instance prend en charge Elastic Volumes, vous pouvez procéder sans détacher le volume ni redémarrer l'instance. Cela vous permet de continuer à utiliser votre application pendant que les modifications prennent effet.

Aucuns frais supplémentaires ne sont facturés pour modifier la configuration d'un volume. La configuration du nouveau volume vous est facturée une fois que la modification du volume a commencé. Pour de plus amples informations, veuillez consulter la page Tarification Amazon EBS.

Sommaire

- Exigences liées à la modification des volumes (p. 1182)
- Demande de modification de vos volumes EBS (p. 1184)
- Surveillance de la progression des modifications de volume (p. 1187)
- Extension d'un système de fichiers Linux après redimensionnement d'un volume (p. 1190)

Exigences liées à la modification des volumes

Les exigences et les limites suivantes s'appliquent lorsque vous modifiez un volume Amazon EBS. Pour en savoir plus sur les exigences générales des volumes EBS, consultez Contraintes sur la taille et la configuration d'un volume EBS (p. 1104).

Types d'instance pris en charge

Elastic Volumes est pris en charge sur les instances suivantes :

- Toutes les instances de la génération actuelle (p. 214)
- Les instances de génération précédente suivantes : C1, C3, CC2, CR1, G2, I2, M1, M3 et R3

Si votre type d'instance ne prend pas en charge Elastic Volumes, consultez Modification d'un volume EBS si Elastic Volumes n'est pas pris en charge (p. 1186).

Exigences pour les volumes Linux

Les AMI Linux exigent une table de partition GPT GUID et GRUB 2 pour les volumes de démarrage de 2 Tio (2 048 Gio) ou plus. Beaucoup d'AMI Linux utilisent encore le schéma de partitionnement MBR, qui prend en charge seulement les volumes de démarrage de 2 Tio maximum. Si votre instance ne démarre pas avec un volume de démarrage supérieur à 2 Tio, l'AMI que vous utilisez peut être limitée à une taille de volume de démarrage inférieure à 2 Tio. Les volumes autres que ceux de démarrage ne sont pas soumis à cette restriction sur les instances Linux. Pour connaître les exigences qui affectent les volumes Windows, consultez Exigences pour les volumes Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Avant de redimensionner un volume de démarrage avec une capacité de plus de 2 Tio, vous pouvez déterminer si le volume utilise un partitionnement MBR ou GPT en exécutant la commande suivante sur votre instance :

```
[ec2-user ~]$ sudo gdisk -l /dev/xvda
```

Une instance Amazon Linux avec un partitionnement GPT renvoie les informations suivantes :

```
GPT fdisk (gdisk) version 0.8.10

  Partition table scan:
    MBR: protective
    BSD: not present
    APM: not present
    GPT: present

  Found valid GPT with protective MBR; using GPT.
```

Une instance SUSE avec un partitionnement MBR renvoie les informations suivantes :

```
GPT fdisk (gdisk) version 0.8.8

  Partition table scan:
    MBR: MBR only
    BSD: not present
    APM: not present
    GPT: not present
```

Limitations

- Les opérations Elastic Volume ne sont pas prises en charge sur les volumes Amazon EBS activés pour la fonctionnalité Multi-Attach.
- La taille d'un nouveau volume ne peut pas dépasser la capacité de volume prise en charge. Pour plus d'informations, consultez Contraintes sur la taille et la configuration d'un volume EBS (p. 1104).
- Si le volume a été attaché avant le 3 novembre 2016, vous devez initialiser la prise en charge d'Elastic Volumes. Pour plus d'informations, consultez Initialisation de la prise en charge d'Elastic Volumes (p. 1185).
- Si vous utilisez un type d'instance de la génération précédente non pris en charge ou si vous rencontrez une erreur en tentant de modifier un volume, consultez Modification d'un volume EBS si Elastic Volumes n'est pas pris en charge (p. 1186).
- Il n'est pas possible de transformer un volume `gp2` qui est attaché à une instance en tant que volume racine en volume `st1` ou `sc1`. S'il est détaché et transformé en volume `st1` ou `sc1`, il n'est pas possible de l'attacher à une instance en tant que volume racine.
- Il n'est pas possible de transformer un volume `gp2` en volume `st1` ou `sc1` si la taille du volume demandé est inférieure à la taille minimale pour les volumes `st1` et `sc1`.
- Dans certains cas, vous devez détacher le volume ou arrêter l'instance pour effectuer la modification. Si vous recevez un message d'erreur lorsque vous tentez de modifier un volume EBS ou que vous êtes sur le point de modifier un volume EBS attaché à un type d'instance de la génération précédente, effectuez l'une des actions suivantes :
  - Pour un volume non-racine, détachez le volume de l'instance, appliquez les modifications, puis attachez à nouveau le volume.
  - Pour un volume racine (démarrage), arrêtez l'instance, appliquez les modifications, puis redémarrez l'instance.
- Après avoir alloué plus de 32,000 IOPS sur un volume `io1` ou `io2` existant, vous devrez peut-être effectuer l'une des opérations suivantes pour bénéficier d'une amélioration complète des performances :
  - Détacher et attacher le volume.
  - Redémarrez l'instance.
- La réduction de la taille d'un volume EBS n'est pas prise en charge. Toutefois, vous pouvez créer un volume plus petit et y migrer ensuite vos données à l'aide d'un outil de niveau application, tel que rsync.
- Le temps de modification est augmenté si vous modifiez un volume qui n'a pas été totalement initialisé. Pour de plus amples informations, veuillez consulter Initialisation des volumes Amazon EBS (p. 1235).

- Après avoir modifié un volume, attendez au moins six heures et assurez-vous que le volume est à l'état `in-use` ou `available` avant de lui appliquer des modifications supplémentaires.
- Si les instances `m3.medium` prennent totalement en charge la modification des volumes, il est possible que les instances `m3.large`, `m3.xlarge` et `m3.2xlarge` ne prennent pas en charge toutes les fonctions de modification des volumes.

Demande de modification de vos volumes EBS

Avec Elastic Volumes, vous pouvez modifier dynamiquement la taille, les performances et le type de volume de vos volumes Amazon EBS sans les détacher.

Utilisez le processus suivant lors de la modification d'un volume :

1. (Facultatif) Avant de modifier un volume contenant des données importantes, une bonne pratique consiste à créer un instantané du volume au cas où vous auriez besoin d'annuler vos modifications. Pour plus d'informations, consultez Création d'instantanés Amazon EBS (p. 1132).
2. Demandez la modification du volume.
3. Surveillez la progression de la modification du volume. Pour plus d'informations, consultez Surveillance de la progression des modifications de volume (p. 1187).
4. Si la taille du volume a été modifiée, étendez le système de fichiers du volume pour tirer parti de la capacité de stockage accrue. Pour plus d'informations, consultez Extension d'un système de fichiers Linux après redimensionnement d'un volume (p. 1190).

Sommaire

- Modification d'un volume EBS à l'aide d'Elastic Volumes (console) (p. 1184)
- Modification d'un volume EBS à l'aide d'Elastic Volumes (AWS CLI) (p. 1185)
- Initialisation de la prise en charge d'Elastic Volumes (si nécessaire) (p. 1185)
- Modification d'un volume EBS si Elastic Volumes n'est pas pris en charge (p. 1186)

Modification d'un volume EBS à l'aide d'Elastic Volumes (console)

Utilisez la procédure suivante pour modifier un volume EBS.

Pour modifier un volume EBS à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Volumes, sélectionnez le volume à modifier, puis choisissez Actions, Modifier le volume.
3. La fenêtre Modifier le volume affiche l'ID du volume et la configuration actuelle du volume, notamment le type, la taille et les IOPS. Vous pouvez modifier certains paramètres ou tous dans une même action. Définissez les nouvelles valeurs de configuration comme suit :
   - Pour modifier le type, choisissez une valeur pour Type de volume.
   - Pour modifier la taille, indiquez une valeur entière autorisée pour Taille.
   - Si vous avez choisi SSD IOPS provisionnés SSD (io1) ou SSD IOPS provisionnés (io2) comme type de volume, entrez une valeur entière autorisée pour IOPS.
4. Une fois que vous avez fini de modifier les paramètres du volume, choisissez Modifier. Lorsque vous êtes invité à confirmer l'opération, choisissez Oui.
5. La modification de la taille du volume n'a pas d'effet pratique tant que vous n'étendez pas le système de fichiers du volume en vue d'utiliser la nouvelle capacité de stockage. Pour plus d'informations, consultez Extension d'un système de fichiers Linux après redimensionnement d'un volume (p. 1190).

Modification d'un volume EBS à l'aide d'Elastic Volumes (AWS CLI)

Utilisez la commande modify-volume pour modifier un ou plusieurs paramètres de configuration d'un volume. Par exemple, si vous avez un volume du type `gp2` d'une taille de 100 Gio, la commande suivante modifie sa configuration en un volume de type `io1` avec 10 000 IOPS et une taille de 200 Gio.

```
aws ec2 modify-volume --volume-type io1 --iops 10000 --size 200 --volume-
id vol-11111111111111111
```

Voici un exemple de sortie :

```
{
    "VolumeModification": {
        "TargetSize": 200,
        "TargetVolumeType": "io1",
        "ModificationState": "modifying",
        "VolumeId": "vol-11111111111111111",
        "TargetIops": 10000,
        "StartTime": "2017-01-19T22:21:02.959Z",
        "Progress": 0,
        "OriginalVolumeType": "gp2",
        "OriginalIops": 300,
        "OriginalSize": 100
    }
}
```

La modification de la taille du volume n'a pas d'effet pratique tant que vous n'étendez pas le système de fichiers du volume en vue d'utiliser la nouvelle capacité de stockage. Pour plus d'informations, consultez Extension d'un système de fichiers Linux après redimensionnement d'un volume (p. 1190).

Initialisation de la prise en charge d'Elastic Volumes (si nécessaire)

Avant de pouvoir modifier un volume attaché à une instance avant le 3 novembre 2016 à 23 h 40 UTC, vous devez initialiser la prise en charge de modification des volumes par l'une des actions suivantes :

- Détacher et attacher le volume
- Arrêter et démarrer l'instance

Utilisez l'une des procédures suivantes pour déterminer si vos instances sont prêtes pour la modification de volume.

Nouvelle console

Pour déterminer si vos instances sont prêtes à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez Instances.
3. Choisissez l'icône Afficher / Masquer les colonnes (icône d'engrenage). Sélectionnez la colonne d'attribut Heure de lancement, puis choisissez Confirmer.
4. Triez la liste d'instances par colonne d'Heure de lancement. Pour chaque instance démarrée avant la date limite, choisissez l'onglet Stockage et cochez la colonne Heure des pièces jointes pour voir quand ses volumes ont été attachés.

Ancienne console

Pour déterminer si vos instances sont prêtes à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le volet de navigation, choisissez Instances.
3. Choisissez l'icône Afficher / Masquer les colonnes (icône d'engrenage). Sélectionnez les attributs Heure de lancement et Périphériques de stockage en mode bloc, puis choisissez Fermer.
4. Triez la liste d'instances par colonne d'Heure de lancement. Pour les instances qui ont commencé avant la date de coupure, vérifiez quand les appareils ont été attachés. Dans l'exemple suivant, vous devez initialiser la modification des volumes pour la première instance car elle a commencé avant la date de coupure et son volume de racine a été attaché avant la date de coupure. Les autres instances sont prêtes car elles ont été démarrées après la date de coupure.

Instance ID	Launch Time	Block Devices
i-e905622e	February 25, 2016 at 1:49:35 PM UTC-8	/dev/xvda=vol-e6b46410:attached:2016-02-25T21:49:35.000Z:true
i-719f99a8	December 8, 2016 at 2:21:51 PM UTC-8	/dev/xvda=vol-bad60e7a:attached:2016-01-15T18:36:12.000Z:true
i-006b02c1b78381e57	May 17, 2017 at 1:52:52 PM UTC-7	/dev/sda1=vol-0de9250441c73024c:attached:2017-05-17T20:52:53.000Z:true, xvdb=vol-0863a86c393496d3d:attached:2017-05-17T20:52:53.000Z:false
i-e3d172ed	May 17, 2017 at 2:48:54 PM UTC-7	/dev/sda1=vol-04c34d0b:attached:2015-01-21T21:19:46.000Z:true

Pour déterminer si vos instances sont prêtes à l'aide de l'interface de ligne de commande

Utilisez la commande describe-instances suivante pour déterminer si le volume a été attaché avant le 3 novembre 2016 à 23 h 40 UTC.

```
aws ec2 describe-instances --query "Reservations[*].Instances[*].
[InstanceId,LaunchTime<='2016-11-01',BlockDeviceMappings[*][Ebs.AttachTime<='2016-11-01']]"
 --output text
```

Pour chaque instance, la première ligne de la sortie montre son ID et si elle a été démarrée avant la date de coupure (vrai ou faux). La première ligne est suivie d'une ou plusieurs lignes qui montrent si chaque volume EBS a été attaché avant la date de coupure (vrai ou faux). Dans la sortie de l'exemple suivant, vous devez initialiser la modification des volumes pour la première instance car elle a commencé avant la date de coupure et son volume de racine a été attaché avant la date de coupure. Les autres instances sont prêtes car elles ont été démarrées après la date de coupure.

```
i-e905622e              True
True
i-719f99a8              False
True
i-006b02c1b78381e57     False
False
False
i-e3d172ed              False
True
```

Modification d'un volume EBS si Elastic Volumes n'est pas pris en charge

Si vous utilisez un type d'instance pris en charge, vous pouvez utiliser Elastic Volumes pour modifier dynamiquement la taille, les performances et le type de volume de vos volumes Amazon EBS sans les détacher.

Si vous ne pouvez pas utiliser Elastic Volumes mais que vous devez modifier le volume racine (de démarrage), vous devez arrêter l'instance, modifier le volume, puis redémarrer l'instance.

Une fois que l'instance a démarré, vous pouvez vérifier la taille du système de fichiers pour vérifier que votre instance reconnaît l'espace de volume agrandi. Sur Linux, utilisez la commande df -h pour vérifier la taille du système de fichiers.

```
[ec2-user ~]$ df -h
Filesystem            Size  Used Avail Use% Mounted on
/dev/xvda1            7.9G  943M  6.9G  12% /
tmpfs                 1.9G     0  1.9G   0% /dev/shm
```

Si la taille ne reflète pas votre volume nouvellement étendu, vous devez étendre le système de fichiers de votre périphérique pour permettre à votre instance d'utiliser le nouvel espace. Pour plus d'informations, consultez Extension d'un système de fichiers Linux après redimensionnement d'un volume (p. 1190).

Surveillance de la progression des modifications de volume

Lorsque vous modifiez un volume EBS, il passe par une suite d'états. Le volume passe à l'état `modifying`, à l'état `optimizing` et enfin à l'état `completed`. A ce stade, le volume est prêt à recevoir d'autres modifications.

> **Note**
>
> Une erreur AWS transitoire peut générer l'état `failed` dans de rares cas. Ceci n'indique pas la santé du volume, mais uniquement l'échec de modification du volume. Si cela se produit, réessayez de modifier le volume.

Lorsque le volume a l'état `optimizing`, ses performances se situent entre les spécifications de configuration source et les spécifications de configuration cible. Les performances de volume transitoires ne seront jamais inférieures aux performances de volume source. Si vous mettez à niveau les opérations d'E/S par seconde, les performances de volume transitoires ne seront jamais inférieures aux performances de volume cible.

Les changements des modifications du volume prennent effet comme suit :

- Les modifications de taille prennent normalement quelques secondes et sont effectives lorsqu'un volume passe à l'état `Optimizing`.
- Les modifications de performances (opérations d'E/S par seconde) peuvent prendre quelques minutes à quelques heures et dépendent de la modification de configuration effectuée.
- La prise en compte d'une nouvelle configuration peut prendre jusqu'à 24 heures, parfois plus, par exemple lorsque le volume n'a pas été entièrement initialisé. En général, un volume d'1 Tio pleinement utilisé met environ 6 heures à migrer vers une nouvelle configuration de performances.

Utilisez l'une des méthodes suivantes pour surveiller la progression des modifications d'un volume.

Sommaire

- Surveillance de la progression des modifications d'un volume (console) (p. 1187)
- Surveillance de la progression des modifications d'un volume (AWS CLI) (p. 1188)
- Surveillance de la progression des modifications d'un volume (CloudWatch Events) (p. 1189)

Surveillance de la progression des modifications d'un volume (console)

Utilisez la procédure suivante pour afficher la progression d'une ou de plusieurs modifications d'un volume.

Pour surveiller la progression d'une modification à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Volumes.
3. Sélectionnez le volume.
4. La colonne État et le champ État du volet d'informations contiennent des informations au format suivant : volume-state - modification-state (progress%). Les états de volume possibles sont la création, la disponibilité, l'utilisation, la suppression en cours, la suppression terminée et l'erreur. Les états de modification possibles sont en cours de modification, d'optimisation, et d'achèvement. Peu de temps après la modification du volume, nous supprimons l'état de modification et la progression, ne laissant que l'état du volume.

   Dans cet exemple, l'état de modification du volume sélectionné est en cours d'optimisation. L'état de modification du volume suivant est en cours de modification.

5. Choisissez le texte dans le champ État du volet d'informations pour afficher des informations sur l'action de modification la plus récente, comme indiqué à l'étape précédente.

Surveillance de la progression des modifications d'un volume (AWS CLI)

Utilisez la commande describe-volumes-modifications pour afficher la progression d'une ou de plusieurs modifications d'un volume. L'exemple suivant décrit les modifications de volume de deux volumes.

```
aws ec2 describe-volumes-modifications --volume-
id vol-1111111111111111 vol-2222222222222222
```

Dans l'exemple de sortie suivant, les modifications de volume sont encore à l'état `modifying`. La progression est présentée en pourcentage.

```
{
    "VolumesModifications": [
        {
            "TargetSize": 200,
            "TargetVolumeType": "io1",
            "ModificationState": "modifying",
            "VolumeId": "vol-1111111111111111",
            "TargetIops": 10000,
            "StartTime": "2017-01-19T22:21:02.959Z",
            "Progress": 0,
            "OriginalVolumeType": "gp2",
            "OriginalIops": 300,
            "OriginalSize": 100
        },
        {
            "TargetSize": 2000,
            "TargetVolumeType": "sc1",
            "ModificationState": "modifying",
            "VolumeId": "vol-2222222222222222",
            "StartTime": "2017-01-19T22:23:22.158Z",
            "Progress": 0,
            "OriginalVolumeType": "gp2",
            "OriginalIops": 300,
            "OriginalSize": 1000
        }
    ]
```

```
}
```

L'exemple suivant décrit tous les volumes dont l'état de modification est `optimizing` ou `completed`, puis filtre et formate les résultats pour n'afficher que les modifications initiées le 1er février 2017 ou après cette date :

```
aws ec2 describe-volumes-modifications --filters Name=modification-
state,Values="optimizing","completed" --query "VolumesModifications[?
StartTime>='2017-02-01'].{ID:VolumeId,STATE:ModificationState}"
```

Voici un exemple de sortie avec des informations sur deux volumes :

```
[
    {
        "STATE": "optimizing",
        "ID": "vol-06397e7a0eEXAMPLE"
    },
    {
        "STATE": "completed",
        "ID": "vol-ba74e18c2aEXAMPLE"
    }
]
```

Surveillance de la progression des modifications d'un volume (CloudWatch Events)

Avec CloudWatch Events, vous pouvez créer une règle de notification pour les événements de modification de volume. Vous pouvez utiliser votre règle pour générer un message de notification avec Amazon SNS ou appeler une fonction Lambda en réponse aux événements correspondants.

Pour surveiller la progression d'une modification avec CloudWatch Events

1. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
2. Choisissez Événements, Créer une règle.
3. Pour Créer un modèle d'événement correspondant aux événements par service, choisissez Un modèle d'événement personnalisé.
4. Pour Créer un modèle d'événement personnalisé, remplacez le contenu par ce qui suit et choisissez Enregistrer.

```
{
  "source": [
    "aws.ec2"
  ],
  "detail-type": [
    "EBS Volume Notification"
  ],
  "detail": {
    "event": [
      "modifyVolume"
    ]
  }
}
```

Voici un exemple de données d'événement :

```
{
   "version": "0",
   "id": "01234567-0123-0123-0123-012345678901",
   "detail-type": "EBS Volume Notification",
   "source": "aws.ec2",
```

```
    "account": "012345678901",
    "time": "2017-01-12T21:09:07Z",
    "region": "us-east-1",
    "resources": [
       "arn:aws:ec2:us-east-1:012345678901:volume/vol-03a55cf56513fa1b6"
    ],
    "detail": {
       "result": "optimizing",
       "cause": "",
       "event": "modifyVolume",
       "request-id": "01234567-0123-0123-0123-0123456789ab"
    }
}
```

Extension d'un système de fichiers Linux après redimensionnement d'un volume

Après avoir augmenté la taille d'un volume EBS (p. 1184), vous devez utiliser des commandes spécifiques au système de fichiers pour étendre le système de fichiers à cette plus grande taille. Vous pouvez redimensionner le système de fichiers dès que le volume passe à l'état `optimizing`.

> **Important**
>
> Avant d'étendre un système de fichiers qui contient des données critiques, une bonne pratique consiste à créer un instantané du volume, au cas où vous auriez besoin d'annuler vos modifications. Pour plus d'informations, consultez Création d'instantanés Amazon EBS (p. 1132). Si votre AMI Linux utilise un schéma de partitionnement MBR, la taille de votre volume de démarrage est limitée à 2 TiO. Pour de plus amples informations, veuillez consulter Exigences pour les volumes Linux (p. 1182) et Contraintes sur la taille et la configuration d'un volume EBS (p. 1104).

Le processus d'extension d'un système de fichiers sous Linux est le suivant :

1. Votre volume EBS peut avoir une partition contenant le système de fichiers et les données. L'augmentation de la taille d'un volume n'augmente pas la taille de la partition. Avant d'étendre le système de fichiers sur un volume redimensionné, vérifiez si le volume possède une partition qui doit être étendue à la nouvelle taille du volume.
2. Utilisez une commande propre au système de fichiers pour redimensionner chaque système de fichiers en indiquant la nouvelle capacité de volume.

Pour obtenir des informations sur l'extension d'un système de fichiers Windows, consultez Extension d'un système de fichiers Windows après redimensionnement d'un volume dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Les exemples suivants vous montrent tout au long du processus d'extension d'un système de fichiers Linux. Pour les systèmes de fichiers et les schémas de partitionnement autres que ceux présentés ici, veuillez consulter la documentation de ces systèmes de fichiers et schémas de partitionnement pour obtenir des instructions.

Exemples

- Exemple : extension du système de fichiers des volumes EBS NVMe (p. 1190)
- Exemple : extension du système de fichiers des volumes EBS (p. 1192)

Exemple : extension du système de fichiers des volumes EBS NVMe

Pour cet exemple, supposons que vous ayez une instance construite sur le système Nitro (p. 218), telle qu'une instance M5. Vous avez redimensionné le volume de démarrage de 8 Go à 16 Go et un volume supplémentaire de 8 Go à 30 Go. Utilisez la procédure suivante pour étendre le système de fichiers des volumes redimensionnés.

Pour étendre le système de fichiers des volumes EBS NVMe

1. Connectez-vous à votre instance (p. 598).
2. Pour vérifier le système de fichiers de chaque volume, utilisez la commande df -hT.

```
[ec2-user ~]$ df -hT
```

L'exemple suivant montre une instance qui possède un volume de démarrage doté d'un système de fichiers XFS et un volume supplémentaire doté d'un système de fichiers XFS. La convention d'attribution de noms `/dev/nvme[0-26]n1` indique que les volumes sont exposés en tant qu'appareils de blocs NVMe.

```
[ec2-user ~]$ df -hT
Filesystem      Type  Size  Used Avail Use% Mounted on
/dev/nvme0n1p1  xfs   8.0G  1.6G  6.5G  20% /
/dev/nvme1n1    xfs   8.0G   33M  8.0G   1% /data
...
```

3. Pour vérifier si le volume possède une partition qui doit être étendue, utilisez la commande lsblk pour afficher des informations sur les appareils de bloc NVMe attachés à votre instance.

```
[ec2-user ~]$ lsblk
NAME          MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
nvme1n1       259:0    0  30G  0 disk /data
nvme0n1       259:1    0  16G  0 disk
##nvme0n1p1   259:2    0   8G  0 part /
##nvme0n1p128 259:3    0   1M  0 part
```

Cet exemple de sortie montre ce qui suit :

- Le volume racine, `/dev/nvme0n1`, possède une partition, `/dev/nvme0n1p1`. La taille du volume racine reflète la nouvelle taille, 16 Go, mais la taille de la partition reflète la taille d'origine, 8 Go, et doit être étendue avant que vous puissiez étendre le système de fichiers.
- Le volume `/dev/nvme1n1` n'a pas de partitions. La taille du volume reflète la nouvelle taille, 30 Go.

4. Pour étendre la partition sur le volume racine, utilisez la commande growpart suivante. Notez qu'un espace figure entre le nom du périphérique et le numéro de partition.

```
[ec2-user ~]$ sudo growpart /dev/nvme0n1 1
```

5. Vous pouvez vérifier que la partition reflète la taille augmentée du volume en utilisant de nouveau la commande lsblk.

```
[ec2-user ~]$ lsblk
NAME          MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
nvme1n1       259:0    0  30G  0 disk /data
nvme0n1       259:1    0  16G  0 disk
##nvme0n1p1   259:2    0  16G  0 part /
##nvme0n1p128 259:3    0   1M  0 part
```

6. Utilisez la commande df -h pour vérifier la taille du système de fichiers pour chaque volume. Dans cet exemple de sortie, les deux systèmes de fichiers reflètent la taille du volume d'origine, 8 Go.

```
[ec2-user ~]$ df -h
Filesystem       Size  Used Avail Use% Mounted on
/dev/nvme0n1p1   8.0G  1.6G  6.5G  20% /
/dev/nvme1n1     8.0G   33M  8.0G   1% /data
...
```

7. [XFS file system] Utilisez la commande xfs_growfs pour étendre le système de fichiers XFS sur chaque volume. Dans cet exemple, `/` et `/data` sont les points de montage de volume indiqués dans la sortie pour df -h.

```
[ec2-user ~]$ sudo xfs_growfs -d /
[ec2-user ~]$ sudo xfs_growfs -d /data
```

   Si les outils XFS ne sont pas déjà installés, vous pouvez les installer comme suit.

```
[ec2-user ~]$ sudo yum install xfsprogs
```

8. [ext4 file system] Utilisez la commande resize2fs pour étendre le système de fichiers sur chaque volume.

```
[ec2-user ~]$ sudo resize2fs /dev/nvme0n1p1
[ec2-user ~]$ sudo resize2fs /dev/nvme1n1
```

9. [Autre système de fichiers] Reportez-vous à la documentation de votre système de fichiers pour obtenir des instructions.
10. (Facultatif) Utilisez à nouveau la commande df -h pour vérifier que chaque système de fichiers reflète l'augmentation de la taille du volume.

```
[ec2-user ~]$ df -h
Filesystem       Size  Used Avail Use% Mounted on
/dev/nvme0n1p1    16G  1.6G   15G  10% /
/dev/nvme1n1      30G   33M   30G   1% /data
...
```

Exemple : extension du système de fichiers des volumes EBS

Dans cet exemple, supposons que vous ayez redimensionné le volume de démarrage d'une instance, telle qu'une instance T2, de 8 Go à 16 Go et un volume supplémentaire de 8 Go à 30 Go. Utilisez la procédure suivante pour étendre le système de fichiers des volumes redimensionnés.

Pour étendre le système de fichiers des volumes EBS

1. Connectez-vous à votre instance (p. 598).
2. Pour vérifier le système de fichiers utilisé pour chaque volume, utilisez la commande df -hT.

```
[ec2-user ~]$ df -hT
```

   L'exemple suivant montre une instance ext4 qui possède un volume de démarrage doté d'un système de fichiers ext4 et un volume supplémentaire doté d'un système de fichiers XFS.

```
[ec2-user ~]$ df -hT
Filesystem      Type  Size  Used Avail Use% Mounted on
/dev/xvda1      ext4  8.0G  1.9G  6.2G  24% /
/dev/xvdf1      xfs   8.0G   45M  8.0G   1% /data
...
```

3. Pour vérifier si le volume possède une partition qui doit être étendue, utilisez la commande lsblk pour afficher des informations sur les appareils de bloc attachés à votre instance.

```
[ec2-user ~]$ lsblk
NAME    MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
xvda    202:0    0  16G  0 disk
```

```
##xvda1 202:1    0   8G  0 part /
xvdf    202:80   0  30G  0 disk
##xvdf1 202:81   0   8G  0 part /data
```

Cet exemple de sortie montre ce qui suit :

- Le volume racine, `/dev/xvda`, possède une partition, `/dev/xvda1`. La taille du volume est de 16 Go, tandis que la taille de la partition est encore de 8 Go et doit être étendue.
- Le volume `/dev/xvdf` possède une partition, `/dev/xvdf1`. La taille du volume est de 30 Go, tandis que la taille de la partition est encore de 8 Go et doit être étendue.

4. Pour étendre la partition sur chaque volume, utilisez les commandes growpart suivantes. Notez qu'un espace figure entre le nom du périphérique et le numéro de partition.

```
[ec2-user ~]$ sudo growpart /dev/xvda 1
[ec2-user ~]$ sudo growpart /dev/xvdf 1
```

5. (Facultatif) Utilisez à nouveau la commande lsblk pour vérifier que les partitions reflètent l'augmentation de la taille du volume.

```
[ec2-user ~]$ lsblk
NAME    MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
xvda    202:0    0  16G  0 disk
##xvda1 202:1    0  16G  0 part /
xvdf    202:80   0  30G  0 disk
##xvdf1 202:81   0  30G  0 part /data
```

6. Utilisez la commande df -h pour vérifier la taille du système de fichiers pour chaque volume. Dans cet exemple de sortie, les deux systèmes de fichiers reflètent la taille du volume d'origine, 8 Go.

```
[ec2-user ~]$ df -h
Filesystem       Size  Used Avail Use% Mounted on
/dev/xvda1       8.0G  1.9G  6.2G  24% /
/dev/xvdf1       8.0G   45M  8.0G   1% /data
...
```

7. [XFS volumes] Utilisez la commande xfs_growfs pour étendre le système de fichiers sur chaque volume. Dans cet exemple, `/` et `/data` sont les points de montage de volume indiqués dans la sortie pour df -h.

```
[ec2-user ~]$ sudo xfs_growfs -d /
[ec2-user ~]$ sudo xfs_growfs -d /data
```

Si les outils XFS ne sont pas déjà installés, vous pouvez les installer comme suit.

```
[ec2-user ~]$ sudo yum install xfsprogs
```

8. [ext4 volumes] Utilisez la commande resize2fs pour étendre le système de fichiers sur chaque volume.

```
[ec2-user ~]$ sudo resize2fs /dev/xvda1
[ec2-user ~]$ sudo resize2fs /dev/xvdf1
```

9. [Autre système de fichiers] Reportez-vous à la documentation de votre système de fichiers pour obtenir des instructions.
10. (Facultatif) Utilisez à nouveau la commande df -h pour vérifier que chaque système de fichiers reflète l'augmentation de la taille du volume.

```
[ec2-user ~]$ df -h
```

```
Filesystem       Size  Used Avail Use% Mounted on
/dev/xvda1        16G  1.9G  14G  12% /
/dev/xvdf1        30G   45M  30G   1% /data
...
```

Chiffrement Amazon EBS

Utilisez Chiffrement Amazon EBS comme solution de chiffrement simple pour vos ressources EBS associées à vos instances EC2. Avec Chiffrement Amazon EBS, vous n'êtes pas tenu de créer, de maintenir et de sécuriser votre propre infrastructure de gestion des clés. Chiffrement Amazon EBS utilise les clés principales du client (CMK) AWS Key Management Service (AWS KMS) lors de la création de volumes chiffrés et d'instantanés.

Les opérations de chiffrement se produisent sur les serveurs qui hébergent des instances EC2, assurant la sécurité des données au repos et des données en transit entre une instance et le stockage EBS associé.

Vous pouvez attacher simultanément des volumes chiffrés et des volumes non chiffrés à une instance.

Sommaire

- Fonctionnement du chiffrement EBS (p. 1194)
- Prérequis (p. 1195)
- Clé par défaut pour le chiffrement EBS (p. 1196)
- Chiffrement par défaut (p. 1196)
- Chiffrement des ressources EBS (p. 1197)
- Scénarios de chiffrement (p. 1198)
- Définition des valeurs par défaut de chiffrement avec l'API et l'interface de ligne de commande (p. 1204)

Fonctionnement du chiffrement EBS

Vous pouvez chiffrer à la fois les volumes de démarrage et de données d'une instance EC2. Lorsque vous créez un volume EBS chiffré et l'attachez à un type d'instance pris en charge, les types de données suivants sont chiffrés :

- Données au repos à l'intérieur du volume
- Toutes les données circulant entre le volume et l'instance
- Tous les instantanés créés à partir du volume
- Tous les volumes créés à partir de ces instantanés

EBS chiffre votre volume avec une clé de données à l'aide de l'algorithme AES-256 standard. Votre clé de données est stockée sur le disque avec vos données chiffrées, mais seulement après qu'EBS l'a chiffrée avec votre clé CMK. Votre clé de données n'apparaît jamais sur le disque en texte brut. Cette même clé de données est partagée par les instantanés du volume et de tous les volumes suivants créés à partir de ces instantanés. Pour de plus amples informations, veuillez consulter Clés de données dans le AWS Key Management Service Developer Guide.

Amazon EBS fonctionne avec AWS KMS pour chiffrer et déchiffrer vos volumes EBS comme suit :

1. Amazon EBS envoie une demande GenerateDataKeyWithoutPlainText à AWS KMS, en spécifiant la CMK que vous avez choisie pour le chiffrement du volume.
2. AWS KMS génère une nouvelle clé de données, la chiffre sous la clé CMK que vous avez choisie pour le chiffrement de volume et envoie la clé de données chiffrée à Amazon EBS pour qu'elle soit stockée avec les métadonnées de volume.

3. Lorsque vous attachez un volume chiffré à une instance, Amazon EC2 envoie une demande Decrypt à AWS KMS, en spécifiant la clé de données chiffrée.
4. Amazon EBS envoie une demande CreateGrant à AWS KMS, afin qu'il puisse déchiffrer la clé de données.
5. AWS KMS déchiffre la clé de données chiffrée et envoie la clé de données déchiffrée à Amazon EC2.
6. Amazon EC2 utilise la clé de données en texte brut dans la mémoire de l'hyperviseur pour chiffrer les E/S de disque sur le volume. La clé de données en texte brut est conservée en mémoire tant que le volume est attaché à l'instance.

Pour de plus amples informations, veuillez consulter Comment Amazon Elastic Block Store (Amazon EBS) utilise AWS KMS et Amazon EC2 exemple deux dans le AWS Key Management Service Developer Guide.

Prérequis

Avant de commencer, vérifiez que les conditions requises suivantes sont respectées :

Types de volume pris en charge

Le chiffrement est pris en charge par tous les types de volume EBS. Les mêmes performances IOPS sont à prévoir sur les volumes chiffrés que sur les volumes non chiffrés, avec des conséquences minimes sur la latence. Vous pouvez accéder à des volumes chiffrés de la même façon qu'à des volumes non chiffrés. Le chiffrement et le déchiffrement sont gérés de façon transparente et ne nécessitent aucune action supplémentaire de votre part ou de vos applications.

Types d'instance pris en charge

Chiffrement Amazon EBS est disponible sur tous les types d'instance de la génération actuelle (p. 214) et sur les types d'instance de la génération précédente (p. 217) suivants : C3 `cr1.8xlarge`, G2, I2, M3 et R3.

Autorisations pour les utilisateurs IAM

Lorsque vous configurez une clé CMK comme clé par défaut pour le chiffrement EBS, la stratégie de clé par défaut permet à tout utilisateur IAM ayant accès aux actions KMS requises d'utiliser cette clé pour chiffrer ou déchiffrer des ressources EBS. Vous devez accorder aux utilisateurs IAM l'autorisation d'appeler les actions suivantes afin d'utiliser le chiffrement EBS :

- `kms:CreateGrant`
- `kms:Decrypt`
- `kms:DescribeKey`
- `kms:GenerateDataKeyWithoutPlainText`
- `kms:ReEncrypt`

Pour suivre le principe du moindre privilège, n'autorisez pas l'accès complet à `kms:CreateGrant`. Au lieu de cela, autorisez l'utilisateur à créer des octrois sur la clé CMK uniquement lorsque l'octroi est créé pour le compte de l'utilisateur par un service AWS, comme illustré dans l'exemple suivant :

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": "kms:CreateGrant",
            "Resource": [
                "arn:aws:kms:us-east-2:123456789012:key/abcd1234-a123-456d-a12b-
a123b4cd56ef"
            ],
```

```
            "Condition": {
                "Bool": {
                    "kms:GrantIsForAWSResource": true
                }
            }
        }
    ]
}
```

Pour de plus amples informations, veuillez consulter Autorise l'accès au compte AWS et active les stratégies IAM dans la section Stratégie de clé par défaut du AWS Key Management Service Developer Guide.

Clé par défaut pour le chiffrement EBS

Amazon EBS crée automatiquement une clé CMK unique gérée par AWS dans chaque région où vous stockez des ressources AWS. Cette clé possède l'alias `alias/aws/ebs`. Par défaut, Amazon EBS utilise cette clé pour le chiffrement. Vous pouvez également spécifier une clé CMK symétrique gérée par le client que vous avez créée comme clé par défaut pour le chiffrement EBS. L'utilisation de votre propre clé CMK vous donne plus de flexibilité dans la mesure où elle vous permet de créer des clés, de les modifier ou de les désactiver à votre convenance.

> **Important**
>
> Amazon EBS ne prend pas en charge les clés CMK asymétriques. Pour de plus amples informations, veuillez consulter Utilisation des clés symétriques et asymétriques dans le Manuel du développeur AWS Key Management Service.

Nouvelle console

Pour configurer la clé par défaut de chiffrement EBS pour une région

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région.
3. Dans le volet de navigation, sélectionnez Tableau de bord EC2.
4. En haut à droite de la page, choisissez Attributs du compte, Chiffrement EBS.
5. Choisissez Gérer.
6. Pour Clé de chiffrement par défaut, choisissez une CMK symétrique gérée par le client.
7. Choisissez Mettre à jour le chiffrement EBS.

Ancienne console

Pour configurer la clé par défaut de chiffrement EBS pour une région

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région.
3. Dans le volet de navigation, sélectionnez Tableau de bord EC2.
4. En haut à droite de la page, choisissez Attributs du compte, Paramètres.
5. Choisissez Change the default key (Modifier la clé par défaut), puis choisissez une clé disponible.
6. Choisissez Save settings (Enregistrer les paramètres).

Chiffrement par défaut

Vous pouvez configurer votre compte AWS pour imposer le chiffrement des nouveaux volumes EBS et des copies d'instantané que vous créez. Par exemple, Amazon EBS chiffre les volumes EBS créés lorsque vous lancez une instance et les instantanés que vous copiez à partir d'un instantané non chiffré. Pour

voir des exemples de transition de ressources EBS non chiffrées à chiffrées, consultez Chiffrement de ressources non chiffrées (p. 1198).

Le chiffrement par défaut n'a aucun effet sur les instantanés ni les volumes EBS existants.

Considérations

- Le chiffrement par défaut est un paramètre spécifique à une région. Si vous l'activez pour une région, vous ne pouvez pas le désactiver pour certains volumes ou instantanés spécifiques dans cette région.
- Lorsque vous activez le chiffrement par défaut, vous ne pouvez lancer une instance que si le type d'instance prend en charge le chiffrement EBS. Pour plus d'informations, consultez Types d'instance pris en charge (p. 1195).
- Lors de la migration des serveurs avec AWS Server Migration Service (SMS), n'activez pas le chiffrement par défaut. Si le chiffrement par défaut est déjà activé et que vous rencontrez des échecs de réplication delta, désactivez cette fonction. Activez plutôt un chiffrement AMI lorsque vous créez la tâche de réplication.

Nouvelle console

Pour activer le chiffrement par défaut pour une région

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région.
3. Dans le volet de navigation, sélectionnez Tableau de bord EC2.
4. En haut à droite de la page, choisissez Attributs du compte, Chiffrement EBS.
5. Choisissez Gérer.
6. Sélectionnez Activer. Vous conservez la CMK gérée par AWS avec l'alias `alias/aws/ebs` créé en votre nom comme clé de chiffrement par défaut, ou choisissez une CMK gérée par le client symétrique.
7. Choisissez Mettre à jour le chiffrement EBS.

Ancienne console

Pour activer le chiffrement par défaut pour une région

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez la région.
3. Dans le volet de navigation, sélectionnez Tableau de bord EC2.
4. En haut à droite de la page, choisissez Attributs du compte, Paramètres.
5. Sous EBS Storage (Stockage EBS), sélectionnez Always encrypt new EBS volumes (Toujours chiffrer les nouveaux volumes EBS).
6. Choisissez Save settings (Enregistrer les paramètres).

Vous ne pouvez pas modifier la clé CMK associée à un volume chiffré ou un instantané existant. Toutefois, vous pouvez associer une autre clé CMK pendant une opération de copie d'instantané, de sorte que l'instantané copié obtenu soit chiffré par cette nouvelle clé CMK.

Chiffrement des ressources EBS

Vous chiffrez les volumes EBS en activant le chiffrement, soit en utilisant le chiffrement par défaut (p. 1196), soit en activant le chiffrement lorsque vous créez un volume que vous souhaitez chiffrer.

Lorsque vous chiffrez un volume, vous pouvez spécifier la clé CMK symétrique à utiliser pour ce faire. Si vous ne spécifiez pas de clé CMK, la clé utilisée pour le chiffrement dépend de l'état de chiffrement de

l'instantané source et de son propriétaire. Pour de plus amples informations, veuillez consulter le tableau des résultats de chiffrement (p. 1203).

> **Note**
>
> Si vous utilisez l'API ou l'AWS CLI pour spécifier une clé CMK, sachez qu'AWS authentifie la clé CMK de manière asynchrone. Si vous spécifiez un ID, un alias ou un ARN qui n'est pas valide, l'action peut sembler se terminer mais finalement échouer.

Vous ne pouvez pas modifier la clé CMK associée à un volume ou un instantané existant. Toutefois, vous pouvez associer une autre clé CMK pendant une opération de copie d'instantané, de sorte que l'instantané copié obtenu soit chiffré par cette nouvelle clé CMK.

Chiffrement d'un volume vide lors de la création

Lorsque vous créez un volume EBS vide, vous pouvez le chiffrer en activant le chiffrement pour l'opération de création du volume spécifique. Si vous avez activé le chiffrement EBS par défaut, le volume est automatiquement chiffré. Par défaut, le volume est chiffré avec votre clé par défaut de chiffrement EBS. Sinon, vous pouvez spécifier une autre clé CMK symétrique pour l'opération de création du volume spécifique. Le volume est chiffré dès sa mise à disposition afin que vos données soient toujours sécurisées. Pour connaître les procédures détaillées, consultez Création d'un volume Amazon EBS (p. 1107).

Par défaut, la clé CMK que vous avez sélectionnée lors de la création d'un volume chiffre les instantanés que vous créez à partir de ce volume et les volumes que vous restaurez à partir de ces instantanés chiffrés. Vous ne pouvez pas supprimer le chiffrement d'un volume ou d'un instantané chiffré, ce qui signifie qu'un volume restauré à partir d'un instantané chiffré, ou une copie d'un instantané chiffré, reste toujours chiffré(e).

Les instantanés publics de volumes chiffrés ne sont pas pris en charge. Vous pouvez cependant partager un instantané chiffré avec certains comptes. Pour obtenir des instructions complètes, consultez Partage d'un instantané Amazon EBS (p. 1143).

Chiffrement de ressources non chiffrées

Bien qu'il n'existe aucun moyen direct de chiffrer un volume ou un instantané non chiffré existant, vous pouvez les chiffrer en créant un volume ou un instantané. Si vous avez activé le chiffrement par défaut, Amazon EBS chiffre le nouveau volume ou instantané obtenu à l'aide de votre clé par défaut de chiffrement EBS. Même si vous n'avez pas activé le chiffrement par défaut, vous pouvez activer le chiffrement lorsque vous créez un volume ou un instantané spécifique. Que vous activiez le chiffrement par défaut ou dans des opérations de création spécifiques, vous pouvez remplacer la clé par défaut de chiffrement EBS et sélectionner une clé CMK symétrique gérée par le client. Pour plus d'informations, consultez Création d'un volume Amazon EBS (p. 1107) et Copie d'un instantané Amazon EBS (p. 1138).

Pour chiffrer la copie de l'instantané avec une clé CMK gérée par le client, vous devez activer le chiffrement et spécifier la clé, comme illustré dans Copie d'un instantané non chiffré (chiffrement par défaut non activé) (p. 1200).

> **Important**
>
> Amazon EBS ne prend pas en charge les clés CMK asymétriques. Pour de plus amples informations, veuillez consulter Utilisation des clés symétriques et asymétriques dans le Manuel du développeur AWS Key Management Service.

Vous pouvez également appliquer de nouveaux états de chiffrement au moment de lancer une instance à partir d'une AMI basée sur EBS. En effet, les AMI basées sur EBS incluent des instantanés des volumes EBS qui peuvent être chiffrés comme décrit. Pour plus d'informations, consultez Utiliser le chiffrement avec des AMI basées sur EBS (p. 168).

Scénarios de chiffrement

Lorsque vous créez une ressource EBS chiffrée, elle est chiffrée par la clé par défaut de chiffrement EBS de votre compte, sauf si vous spécifiez une autre clé CMK gérée par le client dans les paramètres de

création de volume ou le mappage de périphérique de stockage en mode bloc pour l'AMI ou l'instance. Pour plus d'informations, consultez Clé par défaut pour le chiffrement EBS (p. 1196).

Les exemples suivants montrent comment vous pouvez gérer l'état de chiffrement de vos volumes et instantanés. Pour obtenir une liste complète des cas de chiffrement, consultez le tableau des résultats de chiffrement (p. 1203).

Exemples

- Restauration d'un volume non chiffré (chiffrement par défaut non activé) (p. 1199)
- Restauration d'un volume non chiffré (chiffrement par défaut activé) (p. 1199)
- Copie d'un instantané non chiffré (chiffrement par défaut non activé) (p. 1200)
- Copie d'un instantané non chiffré (chiffrement par défaut activé) (p. 1201)
- Rechiffrement d'un volume chiffré (p. 1201)
- Rechiffrement d'un instantané chiffré (p. 1202)
- Migration des données entre les volumes chiffrés et non chiffrés (p. 1203)
- Résultats du chiffrement (p. 1203)

Restauration d'un volume non chiffré (chiffrement par défaut non activé)

Sans le chiffrement par défaut activé, un volume restauré à partir d'un instantané non chiffré est non chiffré par défaut. Cependant, vous pouvez chiffrer le volume créé en définissant le paramètre `Encrypted` et, éventuellement, le paramètre `KmsKeyId`. Le schéma suivant illustre le processus.

Si vous ne spécifiez pas le paramètre `KmsKeyId`, le volume obtenu est chiffré à l'aide de votre clé par défaut de chiffrement EBS. Vous devez fournir un ID de clé pour chiffrer le volume avec une autre clé CMK.

Pour plus d'informations, consultez Création d'un volume à partir d'un instantané (p. 1108).

Restauration d'un volume non chiffré (chiffrement par défaut activé)

Lorsque vous avez activé le chiffrement par défaut, le chiffrement est obligatoire pour les volumes restaurés à partir d'instantanés non chiffrés et aucun paramètre de chiffrement n'est requis pour utiliser votre clé CMK par défaut. Le schéma suivant illustre ce cas simple par défaut :

Si vous souhaitez chiffrer le volume restauré avec une clé CMK symétrique gérée par le client, vous devez fournir les paramètres `Encrypted` et `KmsKeyId`, comme illustré dans Restauration d'un volume non chiffré (chiffrement par défaut non activé) (p. 1199).

Copie d'un instantané non chiffré (chiffrement par défaut non activé)

Sans le chiffrement par défaut activé, une copie d'un instantané non chiffré est non chiffrée par défaut. Cependant, vous pouvez chiffrer l'instantané créé en définissant le paramètre `Encrypted` et, éventuellement, le paramètre `KmsKeyId`. Si vous omettez le paramètre `KmsKeyId`, l'instantané obtenu est chiffré par votre clé CMK par défaut. Vous devez fournir un ID de clé pour chiffrer le volume avec une autre clé CMK symétrique.

Le schéma suivant illustre le processus.

Note

Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Vous pouvez chiffrer un volume EBS en copiant un instantané non chiffré sur un instantané chiffré, puis en créant un volume à partir de l'instantané chiffré. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).

Copie d'un instantané non chiffré (chiffrement par défaut activé)

Lorsque vous avez activé le chiffrement par défaut, le chiffrement est obligatoire pour les copies d'instantanés non chiffrés et aucun paramètre de chiffrement n'est requis si votre clé CMK par défaut est utilisée. Le schéma suivant illustre ce scénario par défaut :

Note

Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Rechiffrement d'un volume chiffré

Lorsque l'action `CreateVolume` est effectuée sur un instantané chiffré, vous avez l'option de rechiffrer celui-ci avec une autre clé CMK. Le schéma suivant illustre le processus. Dans cet exemple, vous possédez deux clés CMK, la clé CMK A et la clé CMK B. L'instantané source est chiffré avec la clé CMK A. Pendant la création de volume, avec l'ID de clé de la clé CMK B spécifié comme paramètre, les données sources sont automatiquement déchiffrées, puis rechiffrées avec la clé CMK B.

Note

Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Pour plus d'informations, consultez Création d'un volume à partir d'un instantané (p. 1108).

Rechiffrement d'un instantané chiffré

La possibilité de chiffrer un instantané pendant la copie vous permet d'appliquer une nouvelle clé CMK symétrique à un instantané déjà chiffré que vous possédez. Les volumes restaurés à partir de la copie qui en résulte sont uniquement accessibles à l'aide de la nouvelle clé CMK. Le schéma suivant illustre le processus. Dans cet exemple, vous possédez deux clés CMK, la clé CMK A et la clé CMK B. L'instantané source est chiffré avec la clé CMK A. Pendant la copie, avec l'ID de clé de la clé CMK B spécifié comme paramètre, les données sources sont automatiquement rechiffrées avec la clé CMK B.

> **Note**
>
> Si vous copiez et chiffrez un instantané sur une nouvelle clé CMK, une copie complète (non incrémentielle) est toujours créée, ce qui entraîne un délai et des coûts de stockage supplémentaires.

Dans un scénario similaire, vous pouvez choisir d'appliquer de nouveaux paramètres de chiffrement à une copie d'instantané partagée avec vous. Par défaut, la copie est chiffrée avec la clé principale client partagée par le propriétaire de l'instantané. Toutefois, nous vous recommandons de créer une copie de l'instantané partagé sous une clé CMK différente de celle dont vous avez le contrôle. Cela protège votre accès au volume si la clé principale client d'origine est compromise ou si le propriétaire la révoque pour quelque raison que ce soit. Pour plus d'informations, consultez Chiffrement et copie d'instantanés (p. 1140).

Migration des données entre les volumes chiffrés et non chiffrés

Lorsque vous avez accès à la fois à un volume chiffré et non chiffré, vous pouvez librement transférer des données entre eux. EC2 effectue les opérations de chiffrement et de déchiffrement en toute transparence.

Par exemple, utilisez la commande rsync pour copier les données. Dans l'exemple suivant, les données source se trouvent à l'emplacement `/mnt/source` et le volume de destination est monté à l'emplacement `/mnt/destination`.

```
[ec2-user ~]$ sudo rsync -avh --progress /mnt/source/ /mnt/destination/
```

Résultats du chiffrement

Le tableau suivant décrit le résultat de chiffrement pour chaque combinaison possible de paramètres.

Le chiffrement est-il activé ?	Le chiffrement est-il activé par défaut ?	Source du volume	Cas par défaut (aucune clé CMK spécifiée)	Cas personnalisé (clé CMK spécifiée)
Non	Non	Nouveau volume (vide)	Non chiffré	S/O
Non	Non	Instantané non chiffré que vous possédez	Non chiffré	
Non	Non	Instantané chiffré que vous possédez	Chiffré par la même clé	
Non	Non	Instantané non chiffré qui est partagé avec vous	Non chiffré	
Non	Non	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut*	
Oui	Non	Nouveau volume	Chiffré par la clé CMK par défaut	Chiffré par une clé CMK spécifiée**
Oui	Non	Instantané non chiffré que vous possédez	Chiffré par la clé CMK par défaut	
Oui	Non	Instantané chiffré que vous possédez	Chiffré par la même clé	
Oui	Non	Instantané non chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	

Le chiffrement est-il activé ?	Le chiffrement est-il activé par défaut ?	Source du volume	Cas par défaut (aucune clé CMK spécifiée)	Cas personnalisé (clé CMK spécifiée)
Oui	Non	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Non	Oui	Nouveau volume (vide)	Chiffré par la clé CMK par défaut	S/O
Non	Oui	Instantané non chiffré que vous possédez	Chiffré par la clé CMK par défaut	
Non	Oui	Instantané chiffré que vous possédez	Chiffré par la même clé	
Non	Oui	Instantané non chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Non	Oui	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Oui	Oui	Nouveau volume	Chiffré par la clé CMK par défaut	Chiffré par une clé CMK spécifiée
Oui	Oui	Instantané non chiffré que vous possédez	Chiffré par la clé CMK par défaut	
Oui	Oui	Instantané chiffré que vous possédez	Chiffré par la même clé	
Oui	Oui	Instantané non chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	
Oui	Oui	Instantané chiffré qui est partagé avec vous	Chiffré par la clé CMK par défaut	

*Clé CMK par défaut utilisée pour le chiffrement EBS pour le compte et la région AWS. Il s'agit par défaut d'une clé CMK unique gérée par AWS pour EBS, ou vous pouvez spécifier une clé CMK gérée par le client. Pour plus d'informations, consultez Clé par défaut pour le chiffrement EBS (p. 1196).

**Clé CMK gérée par le client, spécifiée pour le volume lors du lancement. Cette clé CMK est utilisée à la place de la clé CMK par défaut pour le compte et la région AWS.

Définition des valeurs par défaut de chiffrement avec l'API et l'interface de ligne de commande

Vous pouvez gérer le chiffrement par défaut et la clé principale client (clé CMK) par défaut en utilisant les actions d'API et les commandes CLI suivantes.

Action d'API	Commande de l'interface de ligne de commande	Description
DisableEbsEncryptionByDefault	disable-ebs-encryption-by-default	Désactive le chiffrement par défaut.
EnableEbsEncryptionByDefault	enable-ebs-encryption-by-default	Active le chiffrement par défaut.

Action d'API	Commande de l'interface de ligne de commande	Description
GetEbsDefaultKmsKeyId	get-ebs-default-kms-key-id	Décrit la clé CMK par défaut.
GetEbsEncryptionByDefault	get-ebs-encryption-by-default	Indique si le chiffrement par défaut est activé.
ModifyEbsDefaultKmsKeyId	modify-ebs-default-kms-key-id	Modifie la clé CMK par défaut utilisée pour chiffrer les volumes EBS.
ResetEbsDefaultKmsKeyId	reset-ebs-default-kms-key-id	Réinitialise la clé CMK par défaut gérée par AWS en tant que clé CMK par défaut utilisée pour chiffrer les volumes EBS.

Restauration d'instantané rapide Amazon EBS

La restauration d'instantané rapide Amazon EBS vous permet de créer un volume à partir d'un instantané entièrement initialisé à la création. Elle élimine les temps de latence liés aux opérations d'E/S sur un bloc lors du premier accès à ce dernier. Les volumes créés avec la restauration rapide d'instantané fournissent instantanément la totalité des performances allouées.

Pour commencer, activez la restauration d'instantané rapide pour des instantanés spécifiques dans des zones de disponibilité déterminées. Chaque paire d'instantanés/zones de disponibilité fait référence à une seule restauration d'instantané rapide. Lorsque vous créez un volume à partir d'un de ces instantanés dans l'une de ses zones de disponibilité activées, le volume est restauré à l'aide de la restauration d'instantané rapide.

Vous pouvez activer la restauration d'instantané rapide pour les instantanés que vous possédez et pour les instantanés publics et privés qui sont partagés avec vous.

Sommaire

- Quotas de la fonction de restauration d'instantané rapide (p. 1205)
- États de la fonction de restauration d'instantané rapide (p. 1206)
- Crédits de création de volume (p. 1206)
- Gestion de la restauration d'instantané rapide (p. 1206)
- Affichage d'instantanés avec la restauration d'instantané rapide activée (p. 1207)
- Affichage des volumes restaurés à l'aide de la restauration d'instantané rapide (p. 1208)
- Suivi de la restauration d'instantané rapide (p. 1209)
- Tarification et facturation (p. 1209)

Quotas de la fonction de restauration d'instantané rapide

Vous pouvez activer jusqu'à 50 instantanés pour une restauration d'instantané rapide par région. Le quota s'applique aux instantanés que vous possédez et aux instantanés qui sont partagés avec vous. Si vous activez la restauration d'instantané rapide pour un instantané partagé avec vous, elle est comptabilisée dans votre quota de restauration d'instantané rapide. Elle n'est pas comptabilisée dans le quota de restauration rapide du propriétaire de l'instantané.

États de la fonction de restauration d'instantané rapide

Lorsque vous activez la fonction de restauration d'instantané rapide pour un instantané, elle peut être dans l'un des états suivants.

- `enabling` — Une demande d'activation de la fonction de restauration d'instantané rapide a été effectuée.
- `optimizing` — La fonction de restauration d'instantané rapide est en cours d'activation. L'optimisation d'un instantané prend 60 minutes par Tio.
- `enabled` — La fonction de restauration d'instantané rapide est activée.
- `disabling` — Une demande de désactivation de la fonction de restauration d'instantané rapide a été faite ou une demande d'activation de la fonction de restauration d'instantané rapide a échoué.
- `disabled` — La fonction de restauration d'instantané rapide est désactivée. Vous pouvez réactiver la fonction de restauration d'instantané rapide lorsque vous le souhaitez.

Crédits de création de volume

Le nombre de volumes qui reçoivent la totalité des bénéfices en matière de performances de la fonction de restauration d'instantané rapide est déterminé par les crédits de création de volume associés à l'instantané. Il y a un compartiment de crédits par instantané et par zone de disponibilité. Chaque volume que vous créez à partir d'un instantané pour lequel la fonction de restauration d'instantané rapide est activée consomme un crédit du compartiment de crédits.

Lorsque vous activez la restauration d'instantané rapide pour un instantané partagé avec vous, vous obtenez un compartiment de crédit distinct pour l'instantané partagé dans votre compte. Si vous créez des volumes à partir de l'instantané partagé, les crédits sont consommés à partir de votre compartiment de crédit ; ils ne sont pas consommés à partir du compartiment de crédit du propriétaire de l'instantané.

La taille du compartiment de crédits dépend de celle de l'instantané, ce qui n'est pas le cas de la taille des volumes créés à partir de l'instantané. La taille du compartiment de crédits de chaque instantané est calculée comme suit :

```
MAX (1, MIN (10, FLOOR(1024/snapshot_size_gib)))
```

Au fur et à mesure que vous consommez des crédits, le compartiment de crédits est rechargé. Le taux de rechargement de chaque compartiment de crédits est calculé comme suit :

```
MIN (10, 1024/snapshot_size_gib)
```

Par exemple, si vous activez la fonction de restauration d'instantané rapide pour un instantané ayant une taille de 100 Gio, la taille maximale de son compartiment de crédits est de 10 crédits et son taux de rechargement est de 10 crédits par heure. Lorsque le compartiment de crédits est plein, vous pouvez créer simultanément 10 volumes initialisés à partir de cet instantané.

Vous pouvez utiliser les métriques Cloudwatch pour surveiller la taille de vos compartiments de crédits et le nombre de crédits disponibles dans chaque compartiment. Pour plus d'informations, consultez Métriques de la fonction de restauration d'instantané rapide (p. 1252).

Après que vous avez créé un volume à partir d'un instantané avec la fonction de restauration d'instantané rapide activée, vous pouvez décrire le volume à l'aide de describe-volumes et vérifier le champ `fastRestored` e la sortie pour déterminer si le volume a été créé comme volume initialisé à l'aide de la restauration d'instantané rapide.

Gestion de la restauration d'instantané rapide

La restauration d'instantané rapide est désactivée par défaut pour un instantané. Vous pouvez activer ou désactiver la restauration d'instantané rapide pour les instantanés que vous possédez et pour ceux qui

sont partagés avec vous. Lorsque vous activez ou désactivez la restauration d'instantané rapide pour un instantané, les modifications s'appliquent uniquement à votre compte.

> **Note**
>
> Lorsque vous activez la restauration d'instantané rapide pour un instantané, votre compte est facturé pour chaque minute pendant laquelle la restauration d'instantané rapide est activée dans une zone de disponibilité particulière. Les frais sont calculés au prorata avec un minimum d'une heure.

Lorsque vous supprimez un instantané que vous possédez, la restauration d'instantané rapide est automatiquement désactivée pour cet instantané dans votre compte. Si vous avez activé la restauration d'instantané rapide pour un instantané partagé avec vous et que le propriétaire de l'instantané le supprime ou l'annule, la restauration d'instantané rapide est automatiquement désactivée pour l'instantané partagé dans votre compte.

Si vous avez activé la restauration d'instantané rapide pour un instantané partagé avec vous et qu'il est chiffré à l'aide d'une clé CMK personnalisée, la restauration d'instantané rapide n'est pas automatiquement désactivée pour l'instantané lorsque le propriétaire de ce dernier révoque votre accès à la clé CMK personnalisée. Vous devez désactiver manuellement la restauration d'instantané rapide pour cet instantané.

Utilisez la procédure suivante pour activer ou désactiver la restauration d'instantané rapide pour un instantané que vous possédez ou pour un instantané partagé avec vous.

Pour activer ou désactiver la restauration d'instantané rapide

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Snapshots (Instantanés).
3. Sélectionnez l'instantané.
4. Choisissez Actions, Manage Fast Snapshot Restore (Gérer la restauration d'instantané rapide).
5. Sélectionnez ou désélectionnez les zones de disponibilité, puis choisissez Save (Enregistrer).
6. Pour suivre l'état de la fonction de restauration d'instantané rapide lorsqu'elle est activée, consultez Fast Snapshot Restore (Restauration d'instantané rapide) sous l'onglet Description.

Pour gérer la restauration d'instantané rapide à l'aide de l'AWS CLI

- enable-fast-snapshot-restores
- disable-fast-snapshot-restores
- describe-fast-snapshot-restores

Affichage d'instantanés avec la restauration d'instantané rapide activée

Suivez la procédure ci-dessous pour afficher l'état de la restauration d'instantané rapide que vous possédez ou d'un instantané partagé avec vous.

Pour afficher l'état de la restauration d'instantané rapide à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Snapshots (Instantanés).
3. Sélectionnez l'instantané.
4. Sous l'onglet Description, examinez Fast Snapshot Restore (Restauration d'instantané rapide), qui indique l'état de la restauration d'instantané rapide. Par exemple, il peut afficher un état de « Optimisation de 2 zones de disponibilité » ou « 2 zones de disponibilité activées ».

Pour afficher les instantanés avec la restauration d'instantané rapide activée à l'aide de l'AWS CLI

Utilisez la commande describe-fast-snapshot-restores pour décrire les instantanés activés pour la restauration d'instantané rapide.

```
aws ec2 describe-fast-snapshot-restores --filters Name=state,Values=enabled
```

Voici un exemple de sortie.

```
{
    "FastSnapshotRestores": [
        {
            "SnapshotId": "snap-0e946653493cb0447",
            "AvailabilityZone": "us-east-2a",
            "State": "enabled",
            "StateTransitionReason": "Client.UserInitiated - Lifecycle state transition",
            "OwnerId": "123456789012",
            "EnablingTime": "2020-01-25T23:57:49.596Z",
            "OptimizingTime": "2020-01-25T23:58:25.573Z",
            "EnabledTime": "2020-01-25T23:59:29.852Z"
        },
        {
            "SnapshotId": "snap-0e946653493cb0447",
            "AvailabilityZone": "us-east-2b",
            "State": "enabled",
            "StateTransitionReason": "Client.UserInitiated - Lifecycle state transition",
            "OwnerId": "123456789012",
            "EnablingTime": "2020-01-25T23:57:49.596Z",
            "OptimizingTime": "2020-01-25T23:58:25.573Z",
            "EnabledTime": "2020-01-25T23:59:29.852Z"
        }
    ]
}
```

Affichage des volumes restaurés à l'aide de la restauration d'instantané rapide

Lorsque vous créez un volume à partir d'un instantané qui est activé pour la restauration d'instantané rapide dans la zone de disponibilité du volume, il est restauré à l'aide de la restauration d'instantané rapide.

Utilisez la commande describe-volumes pour afficher les volumes qui ont été créés à partir d'un instantané activé pour la restauration d'instantané rapide.

```
aws ec2 describe-volumes --filters Name=fast-restored,Values=true
```

Voici un exemple de sortie.

```
{
    "Volumes": [
        {
            "Attachments": [],
            "AvailabilityZone": "us-east-2a",
            "CreateTime": "2020-01-26T00:34:11.093Z",
            "Encrypted": true,
            "KmsKeyId": "arn:aws:kms:us-west-2:123456789012:key/8c5b2c63-b9bc-45a3-
a87a-5513e232e843",
            "Size": 20,
            "SnapshotId": "snap-0e946653493cb0447",
            "State": "available",
            "VolumeId": "vol-0d371921d4ca797b0",
```

```
            "Iops": 100,
            "VolumeType": "gp2",
            "FastRestored": true
        }
    ]
}
```

Suivi de la restauration d'instantané rapide

Amazon EBS émet des événements Amazon CloudWatch lorsque l'état de restauration d'instantané rapide change. Pour de plus amples informations, veuillez consulter Événements de restauration d'instantané rapide EBS (p. 1261).

Tarification et facturation

Vous êtes facturé pour chaque minute pendant laquelle la restauration d'instantané rapide est activée pour un instantané dans une zone de disponibilité particulière. Les frais sont calculés au prorata avec un minimum d'une heure.

Par exemple, si vous activez la restauration d'instantané rapide pour un instantané dans `US-East-1a` pendant un mois (30 jours), vous êtes facturé 540 $ (`1` instantané x `1` AZ x `720` heures x `$0.75` par heure). Si vous activez la restauration d'instantané rapide pour deux instantanés dans `us-east-1a`, `us-east-1b` et `us-east-1c` pour la même période, vous êtes facturé 3 240 $ (`2` instantané x `3` AZs x `720` heures x `$0.75` par heure).

Si vous activez la restauration d'instantané rapide pour un instantané public ou privé partagé avec vous, votre compte est facturé ; le propriétaire de l'instantané ne l'est pas. Lorsqu'un instantané partagé avec vous est supprimé ou non partagé par son propriétaire, la restauration d'instantané rapide est désactivée pour l'instantané dans votre compte et la facturation est arrêtée.

Pour de plus amples informations, veuillez consulter Tarification Amazon EBS.

Amazon EBS et NVMe sur des instances Linux

Les volumes EBS sont exposés sous forme de blocs NVMe sur des instances construites sur le Système Nitro (p. 218). Les noms de périphériques sont `/dev/nvme0n1`, `/dev/nvme1n1`, etc. Les noms du périphérique que vous spécifiez dans un mappage de périphérique de stockage en mode bloc sont modifiés par les noms du périphérique NVMe (`/dev/nvme[0-26]n1`). Le pilote du périphérique de stockage en mode bloc peut attribuer les noms du périphérique NVMe dans un autre ordre que celui que vous avez spécifié pour les volumes dans le mappage de périphérique de stockage en mode bloc.

Les garanties de performance EBS définies sur la page Description détaillée d'Amazon EBS s'appliquent quelle que soit l'interface du périphérique de stockage en mode bloc.

Sommaire

- Installation ou mise à niveau du pilote NVMe (p. 1209)
- Identification du périphérique EBS (p. 1211)
- Utilisation des volumes EBS NVMe (p. 1212)
- Expiration de l'intégration des E/S (p. 1212)

Installation ou mise à niveau du pilote NVMe

Pour accéder aux volumes NVMe, les pilotes NVMe doivent être installés. Les instances peuvent prendre en charges les volumes EBS NVMe, les volumes de stockage d'instance NVMe, les deux types de volumes

NVMe ou aucun volume NVMe. Pour plus d'informations, consultez Résumé des fonctions de réseautage et de stockage (p. 220).

Les AMI suivantes incluent les pilotes NVMe requis :

- Amazon Linux 2
- Amazon Linux AMI 2018.03
- Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure
- Red Hat Enterprise Linux 7.4 ou version ultérieure
- SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
- CentOS 7.4.1708 ou version ultérieure
- FreeBSD 11.1 ou version ultérieure
- Debian GNU/Linux 9 ou version ultérieure

Pour plus d'informations sur les pilotes NVMe des instances Windows, consultez Amazon EBS et NVMe sur des instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Pour vérifier que votre instance a le pilote NVMe

Vous pouvez contrôler que votre instance a le pilote NVMe et vérifier la version du pilote à l'aide de la commande suivante. Si l'instance a le pilote NVMe, la commande renvoie des informations sur le pilote.

```
$ modinfo nvme
```

Pour mettre à jour le pilote NVMe

Si votre instance a le pilote NVMe, vous pouvez le mettre à jour vers la dernière version à l'aide de la procédure suivante.

1. Connectez-vous à votre instance.
2. Mettez à jour le cache de votre package pour obtenir les mises à jour de packages nécessaires, comme suit.

   - Pour Amazon Linux 2, Amazon Linux, CentOS et Red Hat Enterprise Linux :

     ```
     [ec2-user ~]$ sudo yum update -y
     ```

   - Pour Ubuntu et Debian :

     ```
     [ec2-user ~]$ sudo apt-get update -y
     ```

3. Ubuntu 16.04 et versions ultérieures incluent le package `linux-aws`, qui contient les pilotes NVMe et ENA requis par les instances basées sur Nitro. Mettez à niveau le package `linux-aws` pour recevoir la version la plus récente, comme suit :

   ```
   [ec2-user ~]$ sudo apt-get install --only-upgrade -y linux-aws
   ```

   Pour Ubuntu 14.04, vous pouvez installer le package `linux-aws` le plus récent, comme suit :

   ```
   [ec2-user ~]$ sudo apt-get install linux-aws
   ```

4. Redémarrez votre instance pour charger la dernière version du noyau.

```
sudo reboot
```

5. Reconnectez-vous à votre instance après son redémarrage.

Identification du périphérique EBS

EBS utilise la virtualisation d'E/S d'une racine unique (SR-IOV) afin de fournir des volumes attachés sur les instances basées sur Nitro à l'aide de la spécification NVMe. Ces périphériques dépendent des pilotes NVMe standard du système d'exploitation. Habituellement, ces pilotes détectent les périphériques attachés en analysant le bus PCI au démarrage de l'instance, puis créent des nœuds de périphériques selon l'ordre dans lequel les périphériques répondent, et non selon la spécification des périphériques dans le mappage de périphérique de stockage en mode bloc. Sous Linux, les périphériques NVMe sont nommés selon le modèle `/dev/nvme<x>n<y>`, où **<x>** correspond à l'ordre d'énumération et, pour EBS, **<y>** correspond à 1. Lors de démarrages consécutifs de l'instance, il arrive que les périphériques répondent à la détection dans un ordre différent, d'où un changement de nom des périphériques.

Nous vous recommandons d'utiliser des identificateurs stables pour les volumes EBS au sein de votre instance, par exemple :

- Pour les instances basées sur Nitro, les mappages de périphériques de stockage en mode bloc spécifiés dans la console Amazon EC2 lorsque vous attachez un volume EBS ou durant les appels de l'API `AttachVolume` ou `RunInstances` sont capturés dans le champ de données propre au fournisseur de l'identification du contrôleur NVMe. Avec les AMI Amazon Linux ultérieures à la version 2017.09.01, nous fournissons une règle `udev` qui lit ces données et crée un lien symbolique vers le mappage de périphérique de stockage en mode bloc.
- Les volumes EBS NVMe ont l'ID de volume EBS comme numéro de série dans l'identification du périphérique. Utilisez la commande `lsblk -o +SERIAL` pour répertorier le numéro de série.
- Lors du formatage d'un périphérique, un UUID est généré, qui persiste pendant toute la durée de vie du système de fichiers. Il est possible de spécifier une étiquette de périphérique au même moment. Pour de plus amples informations,veuillez consulter Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113) et Démarrage à partir du mauvais volume (p. 1363).

AMI Amazon Linux

Avec l'AMI Amazon Linux 2017.09.01 ou ultérieure (y compris Amazon Linux 2), vous pouvez exécuter la commande ebsnvme-id comme suit afin de mapper le nom de périphérique NVMe à un ID de volume et un nom de périphérique :

```
[ec2-user ~]$ sudo /sbin/ebsnvme-id /dev/nvme1n1
Volume ID: vol-01324f611e2463981
/dev/sdf
```

Amazon Linux crée également un lien symbolique entre le nom du périphérique du mappage de périphérique de stockage en mode bloc (par exemple, `/dev/sdf`) et le nom du périphérique NVMe.

AMI FreeBSD

À partir de FreeBSD 12.2-RELEASE, vous pouvez exécuter la commande ebsnvme-id comme indiqué ci-dessus. Transmettez le nom du périphérique NVMe (par exemple, `nvme0`) ou du périphérique de disque (par exemple, `nvd0` ou `nda0`). FreeBSD crée également des liens symboliques vers les périphériques de disque (par exemple `/dev/aws/disk/ebs/`*`volume_id`*).

Autres AMI Linux

Avec la version 4.2 ou une version ultérieure du noyau, vous pouvez exécuter la commande nvme id-ctrl comme suit pour mapper un périphérique NVMe à un ID de volume. Commencez par installer le package de ligne de commande NVMe, `nvme-cli`, à l'aide des outils de gestion du package pour votre distribution Linux. Pour obtenir des instructions de téléchargement et d'installation pour d'autres distributions, reportez-vous à la documentation correspondante.

L'exemple suivant permet d'obtenir l'ID de volume et le nom de périphérique. Le nom de périphérique est disponible via une extension du contrôleur NVMe spécifique au fournisseur (octets 384:4095 de l'identification du contrôleur) :

```
[ec2-user ~]$ sudo nvme id-ctrl -v /dev/nvme1n1
NVME Identify Controller:
vid     : 0x1d0f
ssvid   : 0x1d0f
sn      : vol01234567890abcdef
mn      : Amazon Elastic Block Store
...
0000: 2f 64 65 76 2f 73 64 6a 20 20 20 20 20 20 20 20 "/dev/sdf..."
```

La commande lsblk répertorie les périphériques disponibles ainsi que leurs points de montage (le cas échéant). Vous pouvez ainsi déterminer quel nom de périphérique utiliser. Dans cet exemple, `/dev/nvme0n1p1` est monté comme périphérique racine et `/dev/nvme1n1` est attaché mais pas monté.

```
[ec2-user ~]$ lsblk
NAME          MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
nvme1n1       259:3   0  100G  0 disk
nvme0n1       259:0   0    8G  0 disk
  nvme0n1p1   259:1   0    8G  0 part /
  nvme0n1p128 259:2   0    1M  0 part
```

Utilisation des volumes EBS NVMe

Pour formater et monter un volume EBS NVMe, consultez Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113).

Si vous utilisez la version 4.2 du noyau Linux ou une version ultérieure, tout changement apporté à la taille d'un volume EBS NVMe sera automatiquement appliqué à l'instance. Pour des noyaux Linux plus anciens, il se peut que vous ayez besoin de détacher et de rattacher le volume EBS ou de redémarrer l'instance afin que le changement de taille soit effectif. Avec une version 3.19 ou ultérieure du noyau Linux, vous pouvez exécuter la commande hdparm comme suit pour forcer une nouvelle analyse du périphérique NVMe :

```
[ec2-user ~]$ sudo hdparm -z /dev/nvme1n1
```

Lorsque vous détachez un volume EBS NVMe, l'instance n'a pas la possibilité de vider les caches du système de fichiers ou les métadonnées avant le détachement du volume. Ainsi, avant de détacher un volume EBS NVMe, veuillez d'abord le synchroniser et le démonter. Si le volume ne se détache pas, vous pouvez tenter une commande `force-detach`, tel que décrit dans Détacher un volume Amazon EBS d'une instance Linux (p. 1127).

Expiration de l'intégration des E/S

Les volumes EBS associés à des instances basées sur Nitro utilisent le pilote NVMe par défaut fourni par le système d'exploitation. La plupart des systèmes d'exploitation spécifient un délai d'attente pour les opérations d'E/S soumises aux périphériques NVMe. Le délai d'attente par défaut est de 30 secondes. Il peut être modifié à l'aide du paramètre de démarrage `nvme_core.io_timeout`. Pour la plupart des noyaux Linux antérieurs à la version 4.6, ce paramètre est `nvme.io_timeout`.

Si la latence d'E/S dépasse la valeur de ce paramètre de délai d'attente, le pilote NVMe Linux fait échouer l'E/S et renvoie une erreur dans le système de fichiers ou l'application. Selon l'opération d'E/S, le système de fichiers ou l'application peut retenter l'erreur. Dans certains cas, il est possible de remonter le système de fichiers en lecture seule.

Pour bénéficier d'une expérience similaire à celles des volumes EBS attachés aux instances Xen, nous vous recommandons de définir `nvme_core.io_timeout` sur la valeur la plus élevée possible. Pour les noyaux actuels, le maximum est 4294967295, alors que pour les noyaux précédents, le maximum est 255. Selon la version de Linux, il se peut que la temporisation soit déjà réglée à la valeur maximale prise en charge. Par exemple, la temporisation est réglée sur 4294967295 par défaut pour les AMI Linux Amazon 2017.09.01 et ultérieures.

Vous pouvez vérifier la valeur maximale pour votre distribution de Linux en écrivant une valeur plus élevée que la valeur maximale suggérée dans `/sys/module/nvme_core/parameters/io_timeout` et en recherchant l'erreur Numerical result out of range au moment d'enregistrer le fichier.

Amazon EBS–Instances optimisées

Une instance optimisée Amazon EBS– a recours à une pile de configuration optimisée et fournit une capacité supplémentaire dédiée pour les E/S Amazon EBS. Cette optimisation offre les meilleures performances pour vos volumes EBS en réduisant les conflits entre les E/S Amazon EBS et le trafic restant de votre instance.

Les instances optimisées par EBS fournissent une bande passante dédiée vers Amazon EBS. Lorsqu'ils sont attachés à une instance optimisée EBS, les volumes SSD à usage général (`gp2`) sont conçus pour garantir leurs performances de base et de leurs performances en rafale pendant 99 % du temps, tandis que les volumes IOPS provisionnés SSD (`io1` and `io2`) sont conçus pour garantir leurs performances provisionnées pendant 99,9 % du temps. Les HDD à débit optimisé (`st1`) et les Cold HDD (`sc1`) garantissent l'homogénéité des performances de 90 % du débit de transmission en paquets pendant 99 % du temps. Les périodes non conformes sont assez uniformément réparties, en ciblant 99 % du débit total attendu chaque heure. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).

Sommaire

- Types d'instance pris en charge (p. 1213)
- Obtention de performances maximales (p. 1228)
- Activation de l'optimisation EBS au lancement (p. 1229)
- Activer l'optimisation EBS pour une instance existante (p. 1229)

Types d'instance pris en charge

Les tableaux suivants présentent les types d'instance qui prennent en charge l'optimisation pour EBS. Ils comprennent la bande passante dédiée à Amazon EBS, le débit agrégé maximum type qui peut être atteint sur cette connexion avec une charge de travail de diffusion en continu et une taille d'E/S de 128 Kio, ainsi que le nombre maximal d'E/S par seconde que l'instance peut prendre en charge si vous utilisez une taille d'E/S de 16 Kio. Choisissez une instance optimisée pour EBS qui fournit un débit Amazon EBS dédié supérieur aux besoins de votre application. Sinon, la connexion entre Amazon EBS et Amazon EC2 peut devenir un goulot d'étranglement.

Optimisée pour EBS par défaut

Le tableau suivant présente les types d'instance qui prennent en charge l'optimisation EBS par défaut et pour lesquels cette optimisation est activée par défaut. Il n'y a aucune nécessité d'activer l'optimisation EBS. Désactiver celle-ci n'a aucun effet.

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`a1.medium` *	3 500	437,5	20 000
`a1.large` *	3 500	437,5	20 000
`a1.xlarge` *	3 500	437,5	20 000
`a1.2xlarge` *	3 500	437,5	20 000
`a1.4xlarge`	3 500	437,5	20 000
`a1.metal`	3 500	437,5	20 000
`c4.large`	500	62,5	4 000
`c4.xlarge`	750	93,75	6 000
`c4.2xlarge`	1 000	125	8 000
`c4.4xlarge`	2 000	250	16 000
`c4.8xlarge`	4 000	500	32 000
`c5.large` *	4 750	593,75	20 000
`c5.xlarge` *	4 750	593,75	20 000
`c5.2xlarge` *	4 750	593,75	20 000
`c5.4xlarge`	4 750	593,75	20 000
`c5.9xlarge`	9 500	1 187,5	40 000
`c5.12xlarge`	9 500	1 187,5	40 000
`c5.18xlarge`	19 000	2 375	80 000
`c5.24xlarge`	19 000	2 375	80 000
`c5.metal`	19 000	2 375	80 000
`c5a.large` *	3 170	396	13 300
`c5a.xlarge` *	3 170	396	13 300
`c5a.2xlarge` *	3 170	396	13 300
`c5a.4xlarge` *	3 170	396	13 300
`c5a.8xlarge`	3 170	396	13 300
`c5a.12xlarge`	4 750	594	20 000
`c5a.16xlarge`	6 300	788	26 700
`c5a.24xlarge`	9 500	1 188	40 000
`c5ad.large` *	3 170	396	13 300
`c5ad.xlarge` *	3 170	396	13 300

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`c5ad.2xlarge` *	3 170	396	13 300
`c5ad.4xlarge` *	3 170	396	13 300
`c5ad.8xlarge`	3 170	396	13 300
`c5ad.12xlarge`	4 750	594	20 000
`c5ad.16xlarge`	6 300	788	26 700
`c5ad.24xlarge`	9 500	1 188	40 000
`c5d.large` *	4 750	593,75	20 000
`c5d.xlarge` *	4 750	593,75	20 000
`c5d.2xlarge` *	4 750	593,75	20 000
`c5d.4xlarge`	4 750	593,75	20 000
`c5d.9xlarge`	9 500	1 187,5	40 000
`c5d.12xlarge`	9 500	1 187,5	40 000
`c5d.18xlarge`	19 000	2 375	80 000
`c5d.24xlarge`	19 000	2 375	80 000
`c5d.metal`	19 000	2 375	80 000
`c5n.large` *	4 750	593,75	20 000
`c5n.xlarge` *	4 750	593,75	20 000
`c5n.2xlarge` *	4 750	593,75	20 000
`c5n.4xlarge`	4 750	593,75	20 000
`c5n.9xlarge`	9 500	1 187,5	40 000
`c5n.18xlarge`	19 000	2 375	80 000
`c5n.metal`	19 000	2 375	80 000
`c6g.medium` *	4 750	593,75	20 000
`c6g.large` *	4 750	593,75	20 000
`c6g.xlarge` *	4 750	593,75	20 000
`c6g.2xlarge` *	4 750	593,75	20 000
`c6g.4xlarge`	4 750	593,75	20 000
`c6g.8xlarge`	9 500	1 187,5	40 000
`c6g.12xlarge`	14 250	1 781,25	50 000
`c6g.16xlarge`	19 000	2 375	80 000

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`c6g.metal`	19 000	2 375	80 000
`c6gd.medium` *	4 750	593,75	20 000
`c6gd.large` *	4 750	593,75	20 000
`c6gd.xlarge` *	4 750	593,75	20 000
`c6gd.2xlarge` *	4 750	593,75	20 000
`c6gd.4xlarge`	4 750	593,75	20 000
`c6gd.8xlarge`	9 500	1 187,5	40 000
`c6gd.12xlarge`	14 250	1 781,25	50 000
`c6gd.16xlarge`	19 000	2 375	80 000
`c6gd.metal`	19 000	2 375	80 000
`d2.xlarge`	750	93,75	6 000
`d2.2xlarge`	1 000	125	8 000
`d2.4xlarge`	2 000	250	16 000
`d2.8xlarge`	4 000	500	32 000
`f1.2xlarge`	1 700	212,5	12 000
`f1.4xlarge`	3 500	437,5	44 000
`f1.16xlarge`	14 000	1 750	75 000
`g3s.xlarge`	850	106,25	5 000
`g3.4xlarge`	3 500	437,5	20 000
`g3.8xlarge`	7 000	875	40 000
`g3.16xlarge`	14 000	1 750	80 000
`g4dn.xlarge` *	3 500	437,5	20 000
`g4dn.2xlarge` *	3 500	437,5	20 000
`g4dn.4xlarge`	4 750	593,75	20 000
`g4dn.8xlarge`	9 500	1 187,5	40 000
`g4dn.12xlarge`	9 500	1 187,5	40 000
`g4dn.16xlarge`	9 500	1 187,5	40 000
`g4dn.metal`	19 000	2 375	80 000
`h1.2xlarge`	1 750	218,75	12 000
`h1.4xlarge`	3 500	437,5	20 000

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`h1.8xlarge`	7 000	875	40 000
`h1.16xlarge`	14 000	1 750	80 000
`i3.large`	425	53,13	3 000
`i3.xlarge`	850	106,25	6 000
`i3.2xlarge`	1 700	212,5	12 000
`i3.4xlarge`	3 500	437,5	16 000
`i3.8xlarge`	7 000	875	32 500
`i3.16xlarge`	14 000	1 750	65 000
`i3.metal`	19 000	2 375	80 000
`i3en.large` *	4 750	593,75	20 000
`i3en.xlarge` *	4 750	593,75	20 000
`i3en.2xlarge` *	4 750	593,75	20 000
`i3en.3xlarge` *	4 750	593,75	20 000
`i3en.6xlarge`	4 750	593,75	20 000
`i3en.12xlarge`	9 500	1 187,5	40 000
`i3en.24xlarge`	19 000	2 375	80 000
`i3en.metal`	19 000	2 375	80 000
`inf1.xlarge` *	4 750	593,75	20 000
`inf1.2xlarge` *	4 750	593,75	20 000
`inf1.6xlarge`	4 750	593,75	20 000
`inf1.24xlarge`	19 000	2 375	80 000
`m4.large`	450	56,25	3 600
`m4.xlarge`	750	93,75	6 000
`m4.2xlarge`	1 000	125	8 000
`m4.4xlarge`	2 000	250	16 000
`m4.10xlarge`	4 000	500	32 000
`m4.16xlarge`	10 000	1 250	65 000
`m5.large` *	4 750	593,75	18 750
`m5.xlarge` *	4 750	593,75	18 750
`m5.2xlarge` *	4 750	593,75	18 750

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`m5.4xlarge`	4 750	593,75	18 750
`m5.8xlarge`	6 800	850	30 000
`m5.12xlarge`	9 500	1 187,5	40 000
`m5.16xlarge`	13 600	1 700	60 000
`m5.24xlarge`	19 000	2 375	80 000
`m5.metal`	19 000	2 375	80 000
`m5a.large` *	2 880	360	16 000
`m5a.xlarge` *	2 880	360	16 000
`m5a.2xlarge` *	2 880	360	16 000
`m5a.4xlarge`	2 880	360	16 000
`m5a.8xlarge`	4 750	593,75	20 000
`m5a.12xlarge`	6 780	847,5	30 000
`m5a.16xlarge`	9 500	1 187,50	40 000
`m5a.24xlarge`	13 570	1 696,25	60 000
`m5ad.large` *	2 880	360	16 000
`m5ad.xlarge` *	2 880	360	16 000
`m5ad.2xlarge` *	2 880	360	16 000
`m5ad.4xlarge`	2 880	360	16 000
`m5ad.8xlarge`	4 750	593,75	20 000
`m5ad.12xlarge`	6 780	847,5	30 000
`m5ad.16xlarge`	9 500	1 187,5	40 000
`m5ad.24xlarge`	13 570	1 696,25	60 000
`m5d.large` *	4 750	593,75	18 750
`m5d.xlarge` *	4 750	593,75	18 750
`m5d.2xlarge` *	4 750	593,75	18 750
`m5d.4xlarge`	4 750	593,75	18 750
`m5d.8xlarge`	6 800	850	30 000
`m5d.12xlarge`	9 500	1 187,5	40 000
`m5d.16xlarge`	13 600	1 700	60 000
`m5d.24xlarge`	19 000	2 375	80 000

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`m5d.metal`	19 000	2 375	80 000
`m5dn.large` *	4 750	593,75	18 750
`m5dn.xlarge` *	4 750	593,75	18 750
`m5dn.2xlarge` *	4 750	593,75	18 750
`m5dn.4xlarge`	4 750	593,75	18 750
`m5dn.8xlarge`	6 800	850	30 000
`m5dn.12xlarge`	9 500	1 187,5	40 000
`m5dn.16xlarge`	13 600	1 700	60 000
`m5dn.24xlarge`	19 000	2 375	80 000
`m5n.large` *	4 750	593,75	18 750
`m5n.xlarge` *	4 750	593,75	18 750
`m5n.2xlarge` *	4 750	593,75	18 750
`m5n.4xlarge`	4 750	593,75	18 750
`m5n.8xlarge`	6 800	850	30 000
`m5n.12xlarge`	9 500	1 187,5	40 000
`m5n.16xlarge`	13 600	1 700	60 000
`m5n.24xlarge`	19 000	2 375	80 000
`m6g.medium` *	4 750	593,75	20 000
`m6g.large` *	4 750	593,75	20 000
`m6g.xlarge` *	4 750	593,75	20 000
`m6g.2xlarge` *	4 750	593,75	20 000
`m6g.4xlarge`	4 750	593,75	20 000
`m6g.8xlarge`	9 500	1 187,5	40 000
`m6g.12xlarge`	14 250	1 781,25	50 000
`m6g.16xlarge`	19 000	2 375	80 000
`m6g.metal`	19 000	2 375	80 000
`m6gd.medium` *	4 750	593,75	20 000
`m6gd.large` *	4 750	593,75	20 000
`m6gd.xlarge` *	4 750	593,75	20 000
`m6gd.2xlarge` *	4 750	593,75	20 000

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`m6gd.4xlarge`	4 750	593,75	20 000
`m6gd.8xlarge`	9 500	1 187,5	40 000
`m6gd.12xlarge`	14 250	1 781,25	50 000
`m6gd.16xlarge`	19 000	2 375	80 000
`m6gd.metal`	19 000	2 375	80 000
`p2.xlarge`	750	93,75	6 000
`p2.8xlarge`	5 000	625	32 500
`p2.16xlarge`	10 000	1 250	65 000
`p3.2xlarge`	1 750	218,75	10 000
`p3.8xlarge`	7 000	875	40 000
`p3.16xlarge`	14 000	1 750	80 000
`p3dn.24xlarge`	19 000	2 375	80 000
`r4.large`	425	53,13	3 000
`r4.xlarge`	850	106,25	6 000
`r4.2xlarge`	1 700	212,5	12 000
`r4.4xlarge`	3 500	437,5	18 750
`r4.8xlarge`	7 000	875	37 500
`r4.16xlarge`	14 000	1 750	75 000
`r5.large` *	4 750	593,75	18 750
`r5.xlarge` *	4 750	593,75	18 750
`r5.2xlarge` *	4 750	593,75	18 750
`r5.4xlarge`	4 750	593,75	18 750
`r5.8xlarge`	6 800	850	30 000
`r5.12xlarge`	9 500	1 187,5	40 000
`r5.16xlarge`	13 600	1 700	60 000
`r5.24xlarge`	19 000	2 375	80 000
`r5.metal`	19 000	2 375	80 000
`r5a.large` *	2 880	360	16 000
`r5a.xlarge` *	2 880	360	16 000
`r5a.2xlarge` *	2 880	360	16 000

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`r5a.4xlarge`	2 880	360	16 000
`r5a.8xlarge`	4 750	593,75	20 000
`r5a.12xlarge`	6 780	847,5	30 000
`r5a.16xlarge`	9 500	1 187,5	40 000
`r5a.24xlarge`	13 570	1 696,25	60 000
`r5ad.large` *	2 880	360	16 000
`r5ad.xlarge` *	2 880	360	16 000
`r5ad.2xlarge` *	2 880	360	16 000
`r5ad.4xlarge`	2 880	360	16 000
`r5ad.8xlarge`	4 750	593,75	20 000
`r5ad.12xlarge`	6 780	847,5	30 000
`r5ad.16xlarge`	9 500	1 187,5	40 000
`r5ad.24xlarge`	13 570	1 696,25	60 000
`r5d.large` *	4 750	593,75	18 750
`r5d.xlarge` *	4 750	593,75	18 750
`r5d.2xlarge` *	4 750	593,75	18 750
`r5d.4xlarge`	4 750	593,75	18 750
`r5d.8xlarge`	6 800	850	30 000
`r5d.12xlarge`	9 500	1 187,5	40 000
`r5d.16xlarge`	13 600	1 700	60 000
`r5d.24xlarge`	19 000	2 375	80 000
`r5d.metal`	19 000	2 375	80 000
`r5dn.large` *	4 750	593,75	18 750
`r5dn.xlarge` *	4 750	593,75	18 750
`r5dn.2xlarge` *	4 750	593,75	18 750
`r5dn.4xlarge`	4 750	593,75	18 750
`r5dn.8xlarge`	6 800	850	30 000
`r5dn.12xlarge`	9 500	1 187,5	40 000
`r5dn.16xlarge`	13 600	1 700	60 000
`r5dn.24xlarge`	19 000	2 375	80 000

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`r5n.large` *	4 750	593,75	18 750
`r5n.xlarge` *	4 750	593,75	18 750
`r5n.2xlarge` *	4 750	593,75	18 750
`r5n.4xlarge`	4 750	593,75	18 750
`r5n.8xlarge`	6 800	850	30 000
`r5n.12xlarge`	9 500	1 187,5	40 000
`r5n.16xlarge`	13 600	1 700	60 000
`r5n.24xlarge`	19 000	2 375	80 000
`r6g.medium` *	4 750	593,75	20 000
`r6g.large` *	4 750	593,75	20 000
`r6g.xlarge` *	4 750	593,75	20 000
`r6g.2xlarge` *	4 750	593,75	20 000
`r6g.4xlarge`	4 750	593,75	20 000
`r6g.8xlarge`	9 500	1 187,5	40 000
`r6g.12xlarge`	14 250	1 781,25	50 000
`r6g.16xlarge`	19 000	2 375	80 000
`r6g.metal`	19 000	2 375	80 000
`r6gd.medium` *	4 750	593,75	20 000
`r6gd.large` *	4 750	593,75	20 000
`r6gd.xlarge` *	4 750	593,75	20 000
`r6gd.2xlarge` *	4 750	593,75	20 000
`r6gd.4xlarge`	4 750	593,75	20 000
`r6gd.8xlarge`	9 500	1 187,5	40 000
`r6gd.12xlarge`	14 250	1 781,25	50 000
`r6gd.16xlarge`	19 000	2 375	80 000
`r6gd.metal`	19 000	2 375	80 000
`t3.nano` *	2 085	260,57	11 800
`t3.micro` *	2 085	260,57	11 800
`t3.small` *	2 085	260,57	11 800
`t3.medium` *	2 085	260,57	11 800

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`t3.large` *	2 780	347,5	15 700
`t3.xlarge` *	2 780	347,5	15 700
`t3.2xlarge` *	2 780	347,5	15 700
`t3a.nano` *	2 085	260,57	11 800
`t3a.micro` *	2 085	260,57	11 800
`t3a.small` *	2 085	260,57	11 800
`t3a.medium` *	2 085	260,57	11 800
`t3a.large` *	2 780	347,5	15 700
`t3a.xlarge` *	2 780	347,5	15 700
`t3a.2xlarge` *	2 780	347,5	15 700
`u-6tb1.metal`	38 000	4 750	160 000
`u-9tb1.metal`	38 000	4 750	160 000
`u-12tb1.metal`	38 000	4 750	160 000
`u-18tb1.metal`	38 000	4 750	160 000
`u-24tb1.metal`	38 000	4 750	160 000
`x1.16xlarge`	7 000	875	40 000
`x1.32xlarge`	14 000	1 750	80 000
`x1e.xlarge`	500	62,5	3 700
`x1e.2xlarge`	1 000	125	7 400
`x1e.4xlarge`	1 750	218,75	10 000
`x1e.8xlarge`	3 500	437,5	20 000
`x1e.16xlarge`	7 000	875	40 000
`x1e.32xlarge`	14 000	1 750	80 000
`z1d.large` *	3 170	396,25	13 333
`z1d.xlarge` *	3 170	396,25	13 333
`z1d.2xlarge`	3 170	396,25	13 333
`z1d.3xlarge`	4 750	593,75	20 000
`z1d.6xlarge`	9 500	1 187,5	40 000
`z1d.12xlarge`	19 000	2 375	80 000
`z1d.metal`	19 000	2 375	80 000

* Ces types d'instance peuvent prendre en charge des performances maximales pendant 30 minutes au moins une fois toutes les 24 heures. Si vous avez une charge de travail exigeant des performances maximales soutenues pendant plus de 30 minutes, sélectionnez un type d'instance selon les performances de référence, comme illustré dans le tableau suivant.

Taille d'instance	Bande passante de référence (Mbit/s)	Débit de référence (Mbit/s, E/S de 128 Kio)	IOPS de référence (E/S de 16 Kio)
`a1.medium`	300	37,5	2 500
`a1.large`	525	65,625	4 000
`a1.xlarge`	800	100	6 000
`a1.2xlarge`	1 750	218,75	10 000
`c5.large`	650	81,25	4 000
`c5.xlarge`	1 150	143,75	6 000
`c5.2xlarge`	2 300	287,5	10 000
`c5a.large`	200	25	800
`c5a.xlarge`	400	50	1 600
`c5a.2xlarge`	800	100	3 200
`c5a.4xlarge`	1,580	198	6 600
`c5ad.large`	200	25	800
`c5ad.xlarge`	400	50	1 600
`c5ad.2xlarge`	800	100	3 200
`c5ad.4xlarge`	1,580	198	6 600
`c5d.large`	650	81,25	4 000
`c5d.xlarge`	1 150	143,75	6 000
`c5d.2xlarge`	2 300	287,5	10 000
`c5n.large`	650	81,25	4 000
`c5n.xlarge`	1 150	143,75	6 000
`c5n.2xlarge`	2 300	287,5	10 000
`c6g.medium`	315	39,375	2 500
`c6g.large`	630	78,75	3 600
`c6g.xlarge`	1 188	148,5	6 000
`c6g.2xlarge`	2 375	296,875	12 000
`c6gd.medium`	315	39,375	2 500
`c6gd.large`	630	78,75	3 600

Taille d'instance	Bande passante de référence (Mbit/s)	Débit de référence (Mbit/s, E/S de 128 Kio)	IOPS de référence (E/S de 16 Kio)
`c6gd.xlarge`	1 188	148,5	6 000
`c6gd.2xlarge`	2 375	296,875	12 000
`g4dn.xlarge`	950	118.75	3 000
`g4dn.2xlarge`	1 150	143,75	6 000
`i3en.large`	577	72.1	3 000
`i3en.xlarge`	1,154	144.2	6 000
`i3en.2xlarge`	2,307	288.39	12 000
`i3en.3xlarge`	3,800	475	15 000
`inf1.xlarge`	1,190	148.75	4 000
`inf1.2xlarge`	1,190	148.75	6 000
`m5.large`	650	81,25	3 600
`m5.xlarge`	1 150	143,75	6 000
`m5.2xlarge`	2 300	287,5	12 000
`m5a.large`	650	81,25	3 600
`m5a.xlarge`	1,085	135.63	6 000
`m5a.2xlarge`	1,580	197.5	8 333
`m5ad.large`	650	81,25	3 600
`m5ad.xlarge`	1,085	135.63	6 000
`m5ad.2xlarge`	1,580	197.5	8 333
`m5d.large`	650	81,25	3 600
`m5d.xlarge`	1 150	143,75	6 000
`m5d.2xlarge`	2 300	287,5	12 000
`m5dn.large`	650	81,25	3 600
`m5dn.xlarge`	1 150	143,75	6 000
`m5dn.2xlarge`	2 300	287,5	12 000
`m5n.large`	650	81,25	3 600
`m5n.xlarge`	1 150	143,75	6 000
`m5n.2xlarge`	2 300	287,5	12 000
`m6g.medium`	315	39,375	2 500
`m6g.large`	630	78,75	3 600

Taille d'instance	Bande passante de référence (Mbit/s)	Débit de référence (Mbit/s, E/S de 128 Kio)	IOPS de référence (E/S de 16 Kio)
`m6g.xlarge`	1 188	148,5	6 000
`m6g.2xlarge`	2 375	296,875	12 000
`m6gd.medium`	315	39,375	2 500
`m6gd.large`	630	78,75	3 600
`m6gd.xlarge`	1 188	148,5	6 000
`m6gd.2xlarge`	2 375	296,875	12 000
`r5.large`	650	81,25	3 600
`r5.xlarge`	1 150	143,75	6 000
`r5.2xlarge`	2 300	287,5	12 000
`r5a.large`	650	81,25	3 600
`r5a.xlarge`	1,085	135.63	6 000
`r5a.2xlarge`	1,580	197.5	8 333
`r5ad.large`	650	81,25	3 600
`r5ad.xlarge`	1,085	135.63	6 000
`r5ad.2xlarge`	1,580	197.5	8 333
`r5d.large`	650	81,25	3 600
`r5d.xlarge`	1 150	143,75	6 000
`r5d.2xlarge`	2 300	287,5	12 000
`r5dn.large`	650	81,25	3 600
`r5dn.xlarge`	1 150	143,75	6 000
`r5dn.2xlarge`	2 300	287,5	12 000
`r5n.large`	650	81,25	3 600
`r5n.xlarge`	1 150	143,75	6 000
`r5n.2xlarge`	2 300	287,5	12 000
`r6g.medium`	315	39,375	2 500
`r6g.large`	630	78,75	3 600
`r6g.xlarge`	1 188	148,5	6 000
`r6g.2xlarge`	2 375	296,875	12 000
`r6gd.medium` *	315	39,375	2 500
`r6gd.large` *	630	78,75	3 600

Taille d'instance	Bande passante de référence (Mbit/s)	Débit de référence (Mbit/s, E/S de 128 Kio)	IOPS de référence (E/S de 16 Kio)
`r6gd.xlarge` *	1 188	148,5	6 000
`r6gd.2xlarge` *	2 375	296,875	12 000
`t3.nano`	43	5.43	250
`t3.micro`	87	10.86	500
`t3.small`	174	21.71	1 000
`t3.medium`	347	43.43	2 000
`t3.large`	695	86.86	4 000
`t3.xlarge`	695	86.86	4 000
`t3.2xlarge`	695	86.86	4 000
`t3a.nano`	45	5.63	250
`t3a.micro`	90	11.25	500
`t3a.small`	175	21.88	1 000
`t3a.medium`	350	43.75	2 000
`t3a.large`	695	86.86	4 000
`t3a.xlarge`	695	86.86	4 000
`t3a.2xlarge`	695	86.86	4 000
`z1d.large`	800	100	3 333
`z1d.xlarge`	1,580	197.5	6 667

Optimisation EBS prise en charge

Le tableau suivant présente les types d’instance qui prennent en charge l’optimisation EBS par défaut, mais pour lesquels cette optimisation n’est pas activée par défaut. Vous pouvez activer l’optimisation EBS lorsque vous lancez ces instances ou lorsqu’elles sont en cours d’exécution. L'optimisation EBS doit être activée sur les instances afin d'atteindre le niveau de performance décrit. Lorsque vous activez l'optimisation EBS pour une instance qui n'est pas optimisée EBS par défaut, vous payez un droit horaire supplémentaire peu élevé pour la capacité dédiée. Pour de plus amples informations sur la tarification, veuillez consulter Instances optimisées EBS sur la page Tarification Amazon EC2, Tarification à la demande.

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`c1.xlarge`	1 000	125	8 000
`c3.xlarge`	500	62,5	4 000
`c3.2xlarge`	1 000	125	8 000
`c3.4xlarge`	2 000	250	16 000

Taille d'instance	Bande passante maximum (Mbit/s)	Débit maximal (Mbit/s, E/S de 128 Kio)	IOPS maximum (E/S de 16 Kio)
`g2.2xlarge`	1 000	125	8 000
`i2.xlarge`	500	62,5	4 000
`i2.2xlarge`	1 000	125	8 000
`i2.4xlarge`	2 000	250	16 000
`m1.large`	500	62,5	4 000
`m1.xlarge`	1 000	125	8 000
`m2.2xlarge`	500	62,5	4 000
`m2.4xlarge`	1 000	125	8 000
`m3.xlarge`	500	62,5	4 000
`m3.2xlarge`	1 000	125	8 000
`r3.xlarge`	500	62,5	4 000
`r3.2xlarge`	1 000	125	8 000
`r3.4xlarge`	2 000	250	16 000

Les instances `i2.8xlarge`, `c3.8xlarge` et `r3.8xlarge` ne disposent pas de bande passante EBS dédiée et n'offrent donc pas d'optimisation EBS. Sur ces instances, le trafic réseau et le trafic Amazon EBS partagent la même interface réseau 10 gigabits.

Obtention de performances maximales

Vous pouvez utiliser les métriques `EBSIOBalance%` et `EBSByteBalance%` pour déterminer si vos instances sont dimensionnées correctement. Vous pouvez afficher ces métriques dans la console CloudWatch et définir une alarme qui est déclenchée en fonction du seuil spécifié. Ces métriques sont exprimées sous forme de pourcentage. Les instances avec un pourcentage d'équilibre constamment faible sont candidates pour une augmentation de leur taille. Les instances pour lesquelles le pourcentage d'équilibre ne descend jamais sous 100 % sont candidates pour une diminution de leur taille. Pour plus d'informations, consultez Surveillance de vos instances avec CloudWatch (p. 761).

Les instances à mémoire élevée sont conçues pour exécuter d'importantes bases de données en mémoire, notamment des déploiements en production de la base de données en mémoire SAP HANA, dans le Cloud. Pour optimiser les performances EBS, utilisez des instances à mémoire élevée avec un nombre pair de volumes `io1` ou `io2` avec des performances provisionnées identiques. Par exemple, pour les charges de travail lourdes d'E/S par seconde, utilisez quatre volumes `io1` ou `io2` avec 40 000 E/S par seconde provisionnées pour obtenir le maximum de 160 000 E/S par seconde d'instance. De même, pour les charges de travail lourdes à débit, utilisez six volumes `io1` ou `io2` avec 48 000 E/S par seconde provisionnées pour obtenir le débit maximal de 4 750 Mo/s. Pour plus de recommandations, consultez Configuration du stockage pour SAP HANA.

Considérations

- Les instances G4, i3en Inf1, M5a, M5ad, R5a, R5ad, T3, T3a et Z1d lancées après le 26 février 2020 fournissent les performances maximales indiquées dans le tableau ci-dessus. Pour obtenir les performances maximales d'une instance lancée avant le 26 février 2020, arrêtez-la et démarrez-la.

- Les instances C5, C5d, C5n, M5, M5d, M5n, M5dn, R5, R5d, R5n, R5dn et P3dn lancées après le 3 décembre 2019 fournissent les performances maximales indiquées dans le tableau ci-dessus. Pour obtenir les performances maximales d'une instance lancée avant le 3 décembre 2019, arrêtez-la et démarrez-la.
- Les instances `u-6tb1.metal`, `u-9tb1.metal` et `u-12tb1.metal` lancées après le 12 mars 2020 fournissent les performances indiquées dans le tableau ci-dessus. Les instances de ce type lancées avant le 12 mars 2020 sont susceptibles de fournir des performances inférieures. Pour obtenir les performances maximales d'une instance lancée avant le 12 mars 2020, contactez votre équipe de compte pour mettre à niveau l'instance sans frais supplémentaires.

Activation de l'optimisation EBS au lancement

Vous pouvez activer cette optimisation pour une instance en définissant son attribut pour l'optimisation EBS.

Pour activer l'optimisation Amazon EBS lors du lancement d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Dans Step 1: Choose an Amazon Machine Image (AMI), sélectionnez une AMI.
4. Dans Step 2: Choose an Instance Type, sélectionnez un type d'instance affiché comme prenant en charge l'optimisation Amazon EBS.
5. Dans Step 3: Configure Instance Details, renseignez les champs nécessaires et choisissez Launch as EBS-optimized instance. Si le type d'instance que vous avez sélectionné dans l'étape précédente ne prend pas en charge l'optimisation Amazon EBS, cette option n'est pas présente. Si le type d'instance que vous avez sélectionné est optimisé Amazon EBS par défaut, cette option est activée et vous ne pouvez pas la désactiver.
6. Suivez les instructions pour terminer l'Assistant et lancer votre instance.

Pour activer l'optimisation EBS lors du lancement d'une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes avec l'option correspondante. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- run-instances avec `--ebs-optimized` (AWS CLI)
- New-EC2Instance avec `-EbsOptimized` (Outils AWS pour Windows PowerShell)

Activer l'optimisation EBS pour une instance existante

Vous pouvez activer ou désactiver l'optimisation EBS pour une instance en cours d'exécution en modifiant son attribut d'instance optimisée Amazon EBS –. Si l'instance est en cours d'exécution, vous devez d'abord l'arrêter.

> **Warning**
>
> Lorsque vous arrêtez une instance, les données contenues sur les volumes de stockage d'instance sont effacées. Pour conserver les données des volumes de stockage d'instance, veillez à les sauvegarder sur un stockage permanent.

Pour activer l'optimisation EBS d'une instance existante à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez l'instance.

3. Sélectionnez l'instance et choisissez Actions, État de l'instance, Arrêter l'instance. L'arrêt de l'instance peut prendre quelques minutes.
4. Tandis que l'instance est toujours sélectionnée, choisissez Actions, Paramètres de l'instance, puis Changer le type d'instance.
5. Pour Modifier le type d'instance, effectuez l'une des opérations suivantes :
   - Si le type de votre instance est optimisé Amazon EBS par défaut, l'option Optimisé pour EBS est activée et vous ne pouvez pas la modifier. Vous pouvez choisir Annuler, car l'optimisation Amazon EBS est déjà activée pour l'instance.
   - Si le type de votre instance prend en charge l'optimisation Amazon EBS, choisissez Optimisé pour EBS, puis choisissez Appliquer.
   - Si le type de votre instance ne prend pas en charge l'optimisation Amazon EBS, vous ne pouvez pas choisir l'option Optimisé pour EBS. Vous pouvez sélectionner un type d'instance à partir de Type d'instance qui prend en charge l'optimisation Amazon EBS, puis choisir Optimisé pour EBS et Appliquer.
6. Choisissez État de l'instance, Démarrer l'instance.

Pour activer l'optimisation EBS d'une instance en cours d'exécution à l'aide de la ligne de commande

1. Si l'instance est en cours d'exécution, utilisez l'une des commandes suivantes pour l'arrêter :
   - stop-instances (AWS CLI)
   - Stop-EC2Instance (Outils AWS pour Windows PowerShell)
2. Pour activer l'optimisation EBS, utilisez l'une des commandes suivantes avec l'option correspondante :
   - modify-instance-attribute avec `--ebs-optimized` (AWS CLI)
   - Edit-EC2InstanceAttribute avec `-EbsOptimized` (Outils AWS pour Windows PowerShell)

Performances des volumes Amazon EBS pour les instances Linux

Plusieurs facteurs, dont les caractéristiques d'E/S et la configuration de vos instances et volumes, peuvent avoir des répercussions sur les performances d'Amazon EBS. Les clients qui suivent les recommandations de nos pages de description détaillée des produits Amazon EBS et Amazon EC2 obtiennent généralement de bons résultats dès la première utilisation. Toutefois, dans certains cas, vous pouvez être amené à effectuer quelques réglages afin d'enregistrer des performances optimales sur la plateforme. Cette rubrique décrit les bonnes pratiques d'ordre général et les ajustements de performances propres à certains cas d'utilisation. Nous vous recommandons d'optimiser les performances à l'aide des informations provenant de votre charge de travail réelle, en plus des comparaisons, afin de déterminer votre configuration optimale. Maintenant que vous maîtrisez les bases de l'utilisation des volumes EBS, nous allons examiner les performances d'E/S dont vous avez besoin et les options qui vous permettront d'améliorer les performances d'Amazon EBS afin de répondre à ces besoins.

Il se peut que les mises à jour AWS apportées aux types de volume EBS sur le plan des performances ne s'appliquent pas immédiatement à vos volumes existants. Pour bénéficier de performances optimales sur un ancien volume, vous devrez peut-être d'abord effectuer une action `ModifyVolume` sur celui-ci. Pour plus d'informations, consultez Modification de la taille, la capacité d'IOPS ou le type d'un volume EBS sur Linux.

Sommaire

- Conseils sur les performances Amazon EBS (p. 1231)

- Caractéristiques d'E/S et surveillance (p. 1233)
- Initialisation des volumes Amazon EBS (p. 1235)
- Configuration RAID sur Linux (p. 1237)
- Comparer les volumes EBS (p. 1242)

Conseils sur les performances Amazon EBS

Ces conseils constituent des bonnes pratiques à appliquer pour obtenir des performances optimales à partir de vos volumes EBS, dans différents scénarios d'utilisation.

Utiliser les instances optimisées pour EBS

Sur les instances sans prise en charge d'un débit optimisé pour EBS, le trafic réseau peut se heurter au trafic entre votre instance et vos volumes EBS. Sur les instances optimisées pour EBS, les deux types de trafic sont séparés. Certaines configurations d'instance optimisées pour EBS entraînent des frais supplémentaires (par exemple, C3, R3 et M3), tandis que d'autres sont optimisées pour EBS sans frais supplémentaires (par exemple, M4, C4, C5 et D2). Pour plus d'informations, consultez Amazon EBS–Instances optimisées (p. 1213).

Comprendre comment les performances sont calculées

Lorsque vous mesurez les performances de vos volumes EBS, il est important de comprendre les unités de mesure impliquées et la méthode de calcul des performances. Pour plus d'informations, consultez Caractéristiques d'E/S et surveillance (p. 1233).

Comprendre votre charge de travail

Il existe un lien entre les performances maximales de vos volumes EBS, la taille et le nombre d'opérations d'E/S, et le temps nécessaire pour effectuer chaque action. Chacun de ces critères (performances, E/S et latence) a un impact sur les autres, et chaque application est plus ou moins sensible à un critère ou à un autre. Pour plus d'informations, consultez Comparer les volumes EBS (p. 1242).

Être conscient des pertes de performances lors de l'initialisation des volumes à partir d'instantanés

La latence augmente considérablement la première fois que vous accédez à chaque bloc de données sur un nouveau volume EBS créé à partir d'un instantané. Vous pouvez éviter cette baisse de performances à l'aide de l'une des solutions suivantes :

- Accédez à chaque bloc avant de placer le volume en production. Ce processus est appelé initialisation (anciennement « préchauffage »). Pour plus d'informations, consultez Initialisation des volumes Amazon EBS (p. 1235).
- Activez la restauration d'instantané rapide sur un instantané pour vous assurer que les volumes EBS créés à partir de l'instantané sont entièrement initialisés à la création et fournissent instantanément la totalité des performances allouées. Pour plus d'informations, consultez Restauration d'instantané rapide Amazon EBS (p. 1205).

Facteurs qui peuvent dégrader les performances des volumes HDD

Lorsque vous créez un instantané d'un volume HDD à débit optimisé (`st1`) ou Cold HDD (`sc1`), les performances peuvent diminuer jusqu'à la valeur de référence du volume alors que l'instantané est en cours de création. Ce comportement est propre à ces types de volume. Voici d'autres facteurs qui peuvent limiter les performances : débit généré supérieur à celui que l'instance peut accepter, pertes de performance lors de l'initialisation des volumes créés à partir d'un instantané, et quantité excessive d'E/S aléatoires de petite taille sur le volume. Pour de plus amples informations sur le calcul du débit des volumes HDD, veuillez consulter Types de volume Amazon EBS (p. 1089).

Vos performances peuvent également être affectées si votre application n'envoie pas suffisamment de demandes d'E/S. Il est possible de contrôler ce phénomène en examinant la longueur de file d'attente et la taille d'E/S de votre volume. La longueur de la file d'attente est le nombre de demandes d'E/S en attente, en provenance de votre application et à destination de votre volume. Pour une cohérence optimale, les volumes basés sur HDD doivent conserver une longueur de file d'attente de 4 ou plus (arrondie au nombre entier le plus proche) lors de l'exécution d'I/O séquentielles d'1 Mio. Pour plus d'informations sur la façon d'assurer la cohérence des performances de vos volumes, consultez Caractéristiques d'E/S et surveillance (p. 1233).

Accroître la lecture anticipée pour les charges de travail à forte densité de lectures et à haut débit sur `st1` et `sc1`

Certaines charges de travail impliquent une forte densité de lecture et accèdent au périphérique de stockage en mode bloc via le cache d'une page du système d'exploitation (par exemple, à partir d'un système de fichiers). Dans ce cas, afin d'obtenir un débit optimal, nous vous recommandons de configurer le paramètre de lecture anticipée sur 1 Mio. Il s'agit d'un paramètre par périphérique de stockage en mode bloc, qui ne s'applique qu'à vos volumes HDD.

Afin d'examiner la valeur actuelle de lecture anticipée pour vos périphériques de stockage en mode bloc, utilisez la commande suivante :

```
[ec2-user ~]$ sudo blockdev --report /dev/<device>
```

Les informations sur les périphériques de stockage en mode bloc s'affichent au format suivant :

```
RO    RA   SSZ   BSZ   StartSec            Size   Device
rw   256   512  4096       4096      8587820544   /dev/<device>
```

Le périphérique affiché indique une valeur de lecture anticipée de 256 (la valeur par défaut). Multipliez ce nombre par la taille du secteur (512 octets) afin d'obtenir la taille de la mémoire tampon de lecture anticipée (128 Kio ici). Pour définir la valeur de la mémoire tampon sur 1 Mio, utilisez la commande suivante :

```
[ec2-user ~]$ sudo blockdev --setra 2048 /dev/<device>
```

Pour vérifier que le paramètre de lecture anticipée affiche maintenant 2 048, exécutez de nouveau la première commande.

N'utilisez ce paramètre que lorsque votre charge de travail se compose d'E/S séquentielles de grande taille. Si elle se compose essentiellement d'E/S aléatoires de petite taille, ce paramètre va dégrader vos performances. En général, si votre charge de travail se compose principalement d'I/O aléatoires ou de petite taille, vous devez envisager d'utiliser un volume SSD à usage général (`gp2`) plutôt que `st1` ou `sc1`.

Utiliser un noyau Linux récent

Utilisez un noyau Linux récent avec une prise en charge des descripteurs indirects. Tous les noyaux Linux version 3.8 et supérieures les prennent en charge, ainsi que toute instance EC2 de la génération actuelle. Si votre taille moyenne d'E/S atteint 44 Kio ou s'en rapproche, il est possible que vous utilisiez une instance ou un noyau qui ne prend pas en charge les descripteurs indirects. Pour plus d'informations sur la façon d'obtenir la taille moyenne d'I/O à partir des métriques Amazon CloudWatch, consultez Caractéristiques d'E/S et surveillance (p. 1233).

Pour obtenir un débit optimal sur les volumes `st1` ou `sc1`, nous vous recommandons d'appliquer la valeur 256 au paramètre `xen_blkfront.max` (pour les versions de noyau Linux antérieures à la 4.6) ou au paramètre `xen_blkfront.max_indirect_segments` (pour un noyau Linux version 4.6 et supérieures). Le paramètre approprié peut être défini dans la ligne de commande de démarrage de votre système d'exploitation.

Par exemple, dans une AMI Amazon Linux avec un noyau antérieur, vous pouvez l'ajouter à la fin de la ligne du noyau, dans la configuration GRUB disponible dans `/boot/grub/menu.lst` :

```
kernel /boot/vmlinuz-4.4.5-15.26.amzn1.x86_64 root=LABEL=/ console=ttyS0
 xen_blkfront.max=256
```

Pour un noyau plus récent, la commande serait semblable à ce qui suit :

```
kernel /boot/vmlinuz-4.9.20-11.31.amzn1.x86_64 root=LABEL=/ console=tty1 console=ttyS0
 xen_blkfront.max_indirect_segments=256
```

Redémarrez votre instance pour que ce paramètre prenne effet.

Pour de plus amples informations, veuillez consulter Configuration de GRUB (p. 207). D'autres distributions Linux, en particulier celles qui n'utilisent pas le programme d'amorçage GRUB, peuvent nécessiter une approche différente pour le réglage des paramètres du noyau.

Pour plus d'informations sur les caractéristiques d'I/O EBS, consultez la présentation re:Invent à ce sujet, intitulée Amazon EBS: Designing for Performance.

Utiliser RAID 0 pour optimiser l'utilisation des ressources d'instance

Certains types d'instance peuvent générer un débit d'E/S supérieur à celui que vous pouvez provisionner pour un seul volume EBS. Vous pouvez associer plusieurs volumes `gp2`, `io1`, `io2`, `st1` ou `sc1` dans une configuration RAID 0 afin d'utiliser la bande passante disponible pour ces instances. Pour de plus amples informations, veuillez consulter Configuration RAID sur Linux (p. 1237).

Suivi des performances à l'aide d'Amazon CloudWatch

Amazon Web Services offre des métriques de performances pour Amazon EBS que vous pouvez analyser et consulter avec Amazon CloudWatch et des contrôles d'état que vous pouvez utiliser afin de surveiller l'intégrité de vos volumes. Pour plus d'informations, consultez Surveillance du statut de vos volumes (p. 1119).

Caractéristiques d'E/S et surveillance

Sur une configuration de volume donnée, certaines caractéristiques d'E/S déterminent les performances pour vos volumes EBS. Les volumes basés sur SSD (SSD à usage général (`gp2`) et IOPS provisionnés SSD (`io1` and `io2`)) offrent des performances constantes, que l'opération d'I/O soit aléatoire ou séquentielle. Les volumes basés sur HDD (HDD à débit optimisé (`st1`) et Cold HDD (`sc1`)) offrent des performances optimales seulement lorsque les opérations d'I/O sont volumineuses et séquentielles. Afin de comprendre comment les volumes SSD et HDD se comporteront dans votre application, il est important de comprendre la connexion entre la demande sur le volume, le nombre d'IOPS disponibles pour ce dernier, le temps nécessaire pour effectuer une opération d'E/S et les limites de débit du volume.

IOPS

Les IOPS constituent une unité de mesure qui correspond aux opérations d'E/S par seconde. Les opérations sont mesurées en Kio, et la technologie de disque sous-jacent détermine la quantité maximale de données qu'un type de volume comptabilise comme une seule I/O. La taille d'une I/O est limitée à 256 Kio pour les volumes SSD et à 1 024 Kio pour les volumes HDD, car les volumes SSD gèrent les I/O aléatoires ou de petite taille beaucoup plus efficacement que les volumes HDD.

Lorsque des opérations d'E/S de petite taille sont physiquement contiguës, Amazon EBS tente de les fusionner dans une seule opération d'E/S, sans dépasser la taille maximale. Par exemple, pour les volumes SSD, une seule opération d'E/S de 1 024 Kio est comptabilisée comme 4 opérations (1 024÷256 = 4), tandis que 8 opérations d'E/S contiguës de 32 Kio chacune représentent une seule opération (8×32 = 256). Toutefois, 8 opérations d'E/S aléatoires non contiguës de 32 Kio chacune sont

comptabilisées comme 8 opérations. Dans ce cas, chaque opération d'E/S inférieure à 32 Kio compte comme une seule opération.

De même, pour les volumes basés sur HDD, une opération d'E/S de 1 024 Kio et huit opérations séquentielles de 128 Kio seraient comptabilisées comme une seule opération. Toutefois, 8 opérations d'E/S aléatoires de 128 Kio seraient comptabilisées comme 8 opérations.

Par conséquent, lorsque vous créez un volume basé sur SSD prenant en charge 3 000 IOPS (soit en provisionnant un volume `io1` ou `io2` à 3 000 IOPS ou en dimensionnant un volume `gp2` à 1 000 Gio), et que vous l'attachez à une instance optimisée pour EBS capable de fournir la bande passante nécessaire, vous pouvez transférer jusqu'à 3 000 I/O de données par seconde, le débit étant déterminé par la taille d'I/O.

Latence et longueur de file d'attente d'un volume

La longueur de file d'attente d'un volume correspond au nombre de demandes d'E/S pour un appareil. La latence correspond à la véritable durée client complète d'une opération d'E/S. En d'autres termes, il s'agit du temps qui s'écoule entre l'envoi d'une E/S à EBS et la réception de la confirmation d'EBS indiquant que la lecture ou l'écriture de l'E/S est terminée. La longueur de la file d'attente doit être correctement calibrée avec la taille et la latence d'E/S, pour éviter de créer des goulots d'étranglement sur le système d'exploitation « invité » ou sur le lien réseau vers EBS.

La longueur de la file d'attente optimale varie en fonction des charges de travail, selon la sensibilité de votre application à la latence et à l'IOPS. Si votre charge de travail ne fournit pas suffisamment de demandes d'E/S pour tirer pleinement parti des performances disponibles dans votre volume EBS, il est possible que le volume ne donne pas les IOPS ou le débit que vous avez provisionnés.

Les applications qui génèrent de nombreuses transactions sont sensibles à une latence d'I/O accrue et sont adaptées à des volumes `io1`, `io2` et `gp2` basés sur SSD. Vous pouvez conserver des IOPS élevées et une latence faible grâce à une longueur de file d'attente moyenne réduite et à un nombre élevé d'IOPS disponibles pour le volume. Si vous envoyez vers un volume un nombre d'IOPS supérieur à la quantité qu'il peut contenir, vous risquez d'accroître la latence d'E/S.

Les applications qui génèrent des débits élevés sont moins sensibles à une latence d'I/O accrue et sont adaptées à des volumes `st1` et `sc1` basés sur HDD. Vous pouvez conserver un débit élevé vers les volumes basés sur HDD grâce à une longueur de file d'attente élevée lors de l'exécution d'E/S séquentielles volumineuses.

Taille des E/S et limites de débit par volume

Pour les volumes basés sur SSD, si votre taille d'E/S est très volumineuse, vous aurez peut-être un nombre inférieur d'IOPS par rapport aux IOPS provisionnées, dans la mesure où vous aurez atteint le débit limite pour le volume. Par exemple, un volume `gp2` inférieur à 1 000 Gio avec un solde de crédits par rafales disponible a une limite de 3 000 IOPS et une limite de débit pour le volume de 250 Mio/s. Si vous utilisez une taille d'E/S de 256 Kio, votre volume atteint sa limite de débit à 1 000 IOPS (1 000 x 256 Kio = 250 Mio). Pour des tailles d'E/S inférieures (par exemple, 16 Kio), ce même volume peut contenir 3 000 IOPS dans la mesure où le débit est nettement inférieur à 250 Mio/s. (Ces exemples supposent que l'I/O de votre volume n'atteint pas les limites de débit de l'instance.) Pour plus d'informations sur les limites de débit pour chaque type de volume EBS, consultez Types de volume Amazon EBS (p. 1089).

Pour les opérations d'E/S moins volumineuses, vous verrez parfois une valeur d'IOPS supérieure à celle provisionnée (en cas de mesure depuis votre instance). Cela se produit lorsque le système d'exploitation de l'instance fusionne les petites opérations d'E/S dans une opération de plus grande taille avant de les transmettre à Amazon EBS.

Si votre charge de travail utilise des I/O séquentielles sur des volumes `st1` et `sc1` basés sur HDD, vous risquez d'obtenir un nombre d'IOPS plus élevé que prévu (mesuré depuis votre instance). Cela se produit lorsque le système d'exploitation de l'instance fusionne des E/S séquentielles et les comptabilise dans des unités de 1 024 Kio. Si votre charge de travail utilise des E/S de petite taille ou aléatoires, vous risquez

d'obtenir un débit moins élevé que prévu. En effet, nous comptabilisons chaque E/S aléatoire et non séquentielle par rapport au nombre total d'IOPS, ce qui peut vous conduire à atteindre la limite d'IOPS du volume plus tôt que prévu.

Quel que soit votre type de volume EBS, si vous n'obtenez pas les IOPS ou le débit prévus dans votre configuration, veillez à ce que la bande passante de votre instance EC2 ne soit pas à l'origine de la limite. Pour obtenir des performances optimales, vous devez toujours utiliser une instance optimisée pour EBS de la génération actuelle (ou qui inclut une connectivité réseau de 10 Go/s). Pour plus d'informations, consultez Amazon EBS–Instances optimisées (p. 1213). Si vous ne voyez pas les IOPS attendues, le problème peut également être dû au manque d'E/S sur les volumes EBS.

Surveiller les caractéristiques d'E/S à l'aide de CloudWatch

Vous pouvez surveiller ces caractéristiques d'E/S grâce aux métriques CloudWatch de chaque volume (p. 1248). Les métriques importantes à prendre en compte sont les suivantes :

- `BurstBalance`
- `VolumeReadBytes`
- `VolumeWriteBytes`
- `VolumeReadOps`
- `VolumeWriteOps`
- `VolumeQueueLength`

`BurstBalance` affiche l'équilibre du compartiment en rafales des volumes `gp2`, `st1` et `sc1` sous forme de pourcentage du solde restant. Lorsque votre compartiment en rafales est épuisé, le I/O du volume (pour volumes `gp2`) ou le débit de volume (pour les volumes `st1` et `sc1`) est limité au niveau de référence. Vérifiez la valeur `BurstBalance` pour déterminer si votre volume est limité pour cette raison.

Les volumes `st1` et `sc1` basés sur HDD sont conçus pour générer de meilleures performances avec des charges de travail qui tirent parti de la taille d'I/O maximale de 1 024 Kio. Pour déterminer la taille d'E/S moyenne de votre volume, divisez `VolumeWriteBytes` par `VolumeWriteOps`. Le même calcul s'applique pour les opérations de lecture. Si la taille d'I/O moyenne est inférieure à 64 Kio, vous devriez pouvoir améliorer les performances en augmentant la taille des opérations d'I/O envoyées à un volume `st1` ou `sc1`.

> **Note**
>
> Si la taille moyenne d'E/S atteint 44 Kio ou s'en rapproche, il est possible que vous utilisiez une instance ou un noyau qui ne prend pas en charge les descripteurs indirects. Tous les noyaux Linux version 3.8 et supérieures les prennent en charge, ainsi que toute instance de la génération actuelle.

Si votre latence d'E/S est supérieure à vos besoins, vérifiez `VolumeQueueLength` afin de vous assurer que votre application ne cherche pas à gérer plus d'IOPS que celles qui ont été provisionnées. Si votre application a besoin d'un nombre d'IOPS supérieur à ce que votre volume peut fournir, envisagez l'utilisation d'un volume `gp2` plus important avec un niveau de performance de base plus élevé, ou un volume `io1` ou `io2` avec plus d'IOPS provisionnées afin de bénéficier de latences plus rapides.

Ressources connexes

Pour plus d'informations sur les caractéristiques d'E/S Amazon EBS, consultez la présentation re:Invent suivante : Amazon EBS: Designing for Performance.

Initialisation des volumes Amazon EBS

Les volumes EBS vides reçoivent leurs performances maximum au moment où ils sont créés et ne nécessitent pas d'initialisation (anciennement préchauffage).

Concernant les volumes ayant été créés à partir d'instantanés, les blocs de stockage doivent être extraits d'Amazon S3 et écrits sur le volume pour que vous puissiez y accéder. Cette action préalable prend du temps et peut causer une hausse significative de la latence des opérations d'E/S lors du premier accès à chaque bloc. Les performances du volume sont obtenues une fois que tous les blocs ont été téléchargés et écrits sur le volume.

> **Important**
>
> Lors de l'initialisation des volumes `io1` and `io2` créés à partir d'instantanés, les performances du volume peuvent chuter jusqu'à plus de 50 % en dessous du niveau attendu, ce qui entraîne l'affichage par le volume d'un état `warning` dans le contrôle de statut Performances des E/S. Cette situation est attendue et vous pouvez ignorer l'état `warning` des volumes `io1` and `io2` lorsque vous les initialisez. Pour plus d'informations, consultez Vérifications du statut du volume EBS (p. 1119).

Pour la plupart des applications, l'amortissement du coût d'initialisation sur la durée de vie du volume est acceptable. Pour éviter cette baisse de performances initiale dans un environnement de production, vous pouvez utiliser l'une des solutions suivantes :

- Forcez l'initialisation immédiate de la totalité du volume. Pour plus d'informations, consultez Initialisation des volumes Amazon EBS sur Linux (p. 1236).
- Activez la restauration d'instantané rapide sur un instantané pour vous assurer que les volumes EBS créés à partir de l'instantané sont entièrement initialisés à la création et fournissent instantanément la totalité des performances allouées. Pour plus d'informations, consultez Restauration d'instantané rapide Amazon EBS (p. 1205).

Initialisation des volumes Amazon EBS sur Linux

Les volumes EBS vides reçoivent leurs performances maximum au moment où ils sont disponibles et ne nécessitent pas d'initialisation (anciennement préchauffage). Pour les volumes qui ont été créés à partir d'instantanés, optez pour les utilitaires dd ou fio afin de lire les données de tous les blocs d'un volume. Toutes les données existantes du volume seront conservées.

Pour de plus amples informations sur l'initialisation de volumes Amazon EBS sous Windows, veuillez consulter Initialisation de volumes Amazon EBS sous Windows.

Pour initialiser un volume créé à partir d'un instantané sur Linux

1. Attachez le volume qui vient d'être restauré à votre instance Linux.
2. Utilisez la commande lsblk pour afficher les périphériques de stockage en mode bloc attachés à votre instance.

   ```
   [ec2-user ~]$ lsblk
   NAME  MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
   xvdf  202:80   0  30G  0 disk
   xvda1 202:1    0   8G  0 disk /
   ```

   Ici, vous pouvez voir que le nouveau volume, `/dev/xvdf`, est attaché, mais pas monté (car aucun chemin n'est répertorié sous la colonne `MOUNTPOINT`).
3. Utilisez les utilitaires dd ou fio pour lire tous les blocs de l'appareil. La commande dd est installée par défaut sur les systèmes Linux, mais la commande fio est nettement plus rapide dans la mesure où elle permet les lectures multithreads.

   > **Note**
   >
   > Cette étape peut prendre de plusieurs minutes à plusieurs heures selon la bande passante de votre instance EC2, les IOPS fournies pour le volume et la taille du volume.

[dd] Le paramètre `if` (fichier en entrée) doit être défini sur le lecteur que vous souhaitez initialiser. Le paramètre `of` (fichier de sortie) doit être défini sur l'appareil virtuel null Linux, `/dev/null`. Le paramètre `bs` définit la taille de bloc de l'opération de lecture. Pour des performances optimales, il doit être défini sur 1 Mo.

> **Important**
>
> L'utilisation incorrecte de la commande dd peut facilement entraîner la destruction des données d'un volume. Veillez à suivre précisément l'exemple de commande ci-dessous. Seul le paramètre `if=/dev/`*`xvdf`* varie en fonction du nom de l'appareil que vous lisez.

```
[ec2-user ~]$ sudo dd if=/dev/xvdf of=/dev/null bs=1M
```

[fio] Si la commande fio est installée sur votre système, utilisez la commande suivante pour initialiser votre volume. Le paramètre `--filename` (fichier en entrée) doit être défini sur le lecteur que vous souhaitez initialiser.

```
[ec2-user ~]$ sudo fio --filename=/dev/xvdf --rw=read --bs=128k --iodepth=32 --
ioengine=libaio --direct=1 --name=volume-initialize
```

Pour installer la commande fio sur Amazon Linux, utilisez la commande suivante :

```
sudo yum install -y fio
```

Pour installer la commande fio sur Ubuntu, utilisez la commande suivante :

```
sudo apt-get install -y fio
```

Une fois l'opération terminée, un rapport s'affiche au sujet de l'opération de lecture. Votre volume est maintenant prêt à être utilisé. Pour plus d'informations, consultez Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113).

Configuration RAID sur Linux

Avec Amazon EBS, vous pouvez utiliser toute configuration RAID standard que vous pourriez utiliser avec un serveur bare metal traditionnel, dans la mesure où cette configuration RAID est prise en charge par le système d'exploitation de votre instance. Cela est dû au fait que l'ensemble de la configuration RAID est mise en œuvre au niveau logiciel. Pour de meilleures performances E/S que celles pouvant être obtenues avec un seul volume, RAID 0 peut agréger plusieurs volumes par bandes ; pour une redondance sur les instances, RAID 1 peut mettre en miroir deux volumes.

Les données de volume Amazon EBS sont répliquées sur plusieurs serveurs dans une zone de disponibilité pour éviter la perte de données résultant de la défaillance d'un seul composant. Cette réplication rend les volumes Amazon EBS dix fois plus fiables que les disques durs classiques. Pour plus d'informations, consultez la section relative à la disponibilité et la durabilité d'Amazon EBS dans les pages de description détaillée du produit Amazon EBS.

> **Note**
>
> Vous devez éviter de démarrer à partir d'un volume RAID. Grub est généralement installé sur un seul périphérique dans une grappe RAID, et si l'un des périphériques en miroir est défaillant, vous risquez de ne pas pouvoir démarrer le système d'exploitation.

Si vous devez créer une grappe RAID sur une instance Windows, consultez Configuration RAID sur Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Table des matières

- Options de configuration RAID (p. 1238)
- Création d'une grappe RAID sous Linux (p. 1239)
- Création d'instantanés de volumes d'une grappe RAID (p. 1242)

Options de configuration RAID

Le tableau suivant compare les options RAID 0 et RAID 1 courantes.

Configuration	Utilisez	Avantages	Inconvénients
RAID 0	Lorsque les performances d'E/S ont plus d'importance que la tolérance aux pannes, par exemple, dans une base de données très utilisée (où la réplication de données est déjà configurée séparément).	Les E/S sont réparties entre les volumes dans un agrégat par bandes. Si vous ajoutez un volume, du débit et des IOPS sont ajoutés directement.	Les performances de l'agrégat par bandes sont limitées au volume du jeu dont les performances sont les plus mauvaises. La perte d'un seul volume se traduit par une perte complète de données pour la grappe.
RAID 1	Lorsque la tolérance aux pannes a plus d'importance que les performances d'E/S, par exemple, dans une application critique.	Plus sûr en matière de durabilité des données.	N'offre pas d'amélioration des performances en écriture ; nécessite une plus grande bande passante entre Amazon EC2 et Amazon EBS que les configurations non-RAID, car les données sont écrites simultanément sur plusieurs volumes.

> **Important**
>
> RAID 5 et RAID 6 ne sont pas recommandés pour Amazon EBS, car les opérations d'écritures de parité de ces modes RAID consomment certaines des E/S par seconde (IOPS) disponibles pour vos volumes. En fonction de la configuration de votre grappe RAID, ces modes RAID fournissent de 20 à 30 % d'E/S par seconde utilisables en moins qu'une configuration RAID 0. Le coût accru est également un facteur à prendre en compte avec ces modes RAID ; avec l'utilisation de tailles et de vitesses de volume identiques, une grappe RAID 0 à 2 volumes peut offrir de meilleures performances qu'une grappe RAID 6 à 4 volumes dont le coût est deux fois plus élevé.

La création d'une grappe RAID 0 vous permet d'obtenir un niveau de performance plus élevé pour un système de fichiers que vous pouvez mettre en service sur un volume Amazon EBS unique. Une grappe RAID 1 fournit un « miroir » de vos données pour offrir une redondance supplémentaire. Avant d'effectuer cette procédure, vous devez décider de la taille souhaitée pour votre grappe RAID et du nombre d'IOPS à mettre en service.

La taille résultante d'une grappe RAID 0 est la somme des tailles des volumes contenues par celle-ci, et la bande passante correspond au total de bande passante disponible des volumes de la grappe. La taille et la bande passante résultantes d'une grappe RAID 1 sont égales à la taille et à la bande passante des volumes de la grappe. Par exemple, deux volumes `io1` Amazon EBS de 500 Gio avec 4 000 IOPS provisionnées pour chacun créeront une grappe RAID 0 de 1000 Gio avec une bande passante disponible de 8 000 IOPS et 1 000 Mio/s de débit ou une grappe RAID 1 de 500 Gio avec une bande passante disponible de 4 000 IOPS et 500 Mio/s de débit.

Cette documentation fournit des exemples de configuration RAID de base. Pour plus d'informations sur la configuration RAID, les performances et la récupération, consultez le Wiki concernant la technologie RAID Linux à l'adresse https://raid.wiki.kernel.org/index.php/Linux_Raid.

Création d'une grappe RAID sous Linux

Utilisez la procédure suivante pour créer la grappe RAID. Notez que vous pouvez obtenir des instructions pour les instances Windows dans la section Création d'une grappe RAID sous Windows dans Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Pour créer une grappe RAID sous Linux

1. Créez les volumes Amazon EBS pour votre grappe. Pour plus d'informations, consultez Création d'un volume Amazon EBS (p. 1107).

   Important

   Créez des volumes avec des tailles et des valeurs de performances d'E/S par seconde (IOPS) identiques pour votre grappe. Veillez à ne pas créer une grappe qui dépasse la bande passante disponible de votre instance EC2. Pour plus d'informations, consultez Amazon EBS–Instances optimisées (p. 1213).

2. Attachez les volumes Amazon EBS à l'instance devant héberger la grappe. Pour plus d'informations, consultez Attacher un volume Amazon EBS à une instance (p. 1109).
3. Utilisez la commande mdadm pour créer une unité RAID logique à partir des volumes Amazon EBS nouvellement attachés. Remplacez le nombre de volumes de votre grappe par *number_of_volumes* et les noms de périphérique de chaque volume de la grappe (tels que `/dev/xvdf`) par *device_name*. Vous pouvez également remplacer *MY_RAID* par votre propre nom unique pour la grappe.

   Note

   Vous pouvez afficher la liste des unités de votre instance avec la commande lsblk pour trouver les noms d'unité.

   (RAID 0 uniquement) Pour créer une grappe RAID 0, exécutez la commande suivante (notez l'option `--level=0` pour agréger la grappe) :

   ```
   [ec2-user ~]$ sudo mdadm --create --verbose /dev/md0 --level=0 --name=MY_RAID --raid-
   devices=number_of_volumes device_name1 device_name2
   ```

   (RAID 1 uniquement) Pour créer une grappe RAID 1, exécutez la commande suivante (notez l'option `--level=1` pour mettre la grappe en miroir) :

   ```
   [ec2-user ~]$ sudo mdadm --create --verbose /dev/md0 --level=1 --name=MY_RAID --raid-
   devices=number_of_volumes device_name1 device_name2
   ```

4. Laissez à la grappe RAID le temps de s'initialiser et se synchroniser. Vous pouvez suivre la progression de ces opérations avec la commande suivante :

   ```
   [ec2-user ~]$ sudo cat /proc/mdstat
   ```

   Voici un exemple de sortie :

   ```
   Personalities : [raid1]
   md0 : active raid1 xvdg[1] xvdf[0]
         20955008 blocks super 1.2 [2/2] [UU]
         [=========>...........]  resync = 46.8% (9826112/20955008) finish=2.9min
    speed=63016K/sec
   ```

En général, vous pouvez afficher des informations détaillées sur votre grappe RAID avec la commande suivante :

```
[ec2-user ~]$ sudo mdadm --detail /dev/md0
```

Voici un exemple de sortie :

```
/dev/md0:
        Version : 1.2
  Creation Time : Mon Jun 27 11:31:28 2016
     Raid Level : raid1
     Array Size : 20955008 (19.98 GiB 21.46 GB)
  Used Dev Size : 20955008 (19.98 GiB 21.46 GB)
   Raid Devices : 2
  Total Devices : 2
    Persistence : Superblock is persistent

    Update Time : Mon Jun 27 11:37:02 2016
          State : clean
...
...
...

    Number   Major   Minor   RaidDevice State
       0     202       80        0      active sync   /dev/sdf
       1     202       96        1      active sync   /dev/sdg
```

5. Créez un système de fichiers sur votre grappe RAID et attribuez-lui une étiquette à utiliser quand vous le monterez ultérieurement. Par exemple, pour créer un système de fichiers ext4 avec l'étiquette *MY_RAID*, exécutez la commande suivante :

```
[ec2-user ~]$ sudo mkfs.ext4 -L MY_RAID /dev/md0
```

Selon les exigences de votre application ou les restrictions de votre système d'exploitation, vous pouvez utiliser un autre système de fichiers, comme ext3 ou XFS (consultez la documentation relative à votre système de fichiers pour trouver la commande de création de système de fichiers correspondante).

6. Pour vous assurer que la grappe RAID est réassemblée automatiquement au démarrage, créez un fichier de configuration qui contient les informations RAID :

```
[ec2-user ~]$ sudo mdadm --detail --scan | sudo tee -a /etc/mdadm.conf
```

> **Note**
>
> Si vous utilisez une distribution Linux autre qu'Amazon Linux, ce fichier doit être placé dans un autre emplacement. Pour plus d'informations, consultez man mdadm.conf sur votre système Linux.

7. Créez une nouvelle une image ramdisk pour précharger correctement les modules de périphérique de stockage en mode bloc pour votre nouvelle configuration RAID :

```
[ec2-user ~]$ sudo dracut -H -f /boot/initramfs-$(uname -r).img $(uname -r)
```

8. Créez un point de montage (mount) pour votre grappe RAID.

```
[ec2-user ~]$ sudo mkdir -p /mnt/raid
```

9. Enfin, montez l'unité RAID sur le point de montage que vous avez créé :

```
[ec2-user ~]$ sudo mount LABEL=MY_RAID /mnt/raid
```

Votre unité RAID est maintenant prête à être utilisée.

10. (Facultatif) Pour monter ce volume Amazon EBS à chaque redémarrage du système, ajoutez une entrée pour l'appareil dans le fichier `/etc/fstab`.

    a. Créez une sauvegarde de votre fichier `/etc/fstab` que vous pouvez utiliser si vous détruisez ou supprimez accidentellement ce fichier en l'éditant.

    ```
    [ec2-user ~]$ sudo cp /etc/fstab /etc/fstab.orig
    ```

    b. Ouvrez le fichier `/etc/fstab` avec votre éditeur de texte préféré (comme nano ou vim).

    c. Placez en commentaires les lignes commençant par « `UUID=` » et, à la fin du fichier, ajoutez une nouvelle ligne pour votre volume RAID à l'aide du format suivant :

    ```
    device_label mount_point file_system_type fs_mntops fs_freq fs_passno
    ```

    Les trois derniers champs de cette ligne correspondent aux options de montage du système de fichiers, à la fréquence de vidage du système de fichiers et à l'ordre des contrôles de système de fichiers au démarrage. Si vous ne savez pas quelles valeurs indiquer, utilisez les valeurs de l'exemple ci-dessous pour ces champs (`defaults,nofail 0 2`). Pour plus d'informations sur les entrées `/etc/fstab` , consultez la page du manuel fstab (en entrant man fstab sur la ligne de commande). Par exemple, pour monter le système de fichiers ext4 sur l'unité avec l'étiquette MY_RAID au point de montage `/mnt/raid`, ajoutez l'entrée suivante à `/etc/fstab`.

    > **Note**
    >
    > Si jamais vous prévoyez de démarrer votre instance sans ce volume attaché (par exemple, pour que ce volume puisse basculer entre différentes instances), vous devez ajouter l'option de montage `nofail` qui permet à l'instance de démarrer même si des erreurs se produisent lors du montage du volume. Les dérivés Debian, comme Ubuntu, doivent également ajouter l'option de montage `nobootwait`.

    ```
    LABEL=MY_RAID       /mnt/raid   ext4    defaults,nofail        0       2
    ```

    d. Après avoir ajouté la nouvelle entrée à `/etc/fstab`, vous devez vérifier que celle-ci fonctionne. Exécutez la commande sudo mount -a pour monter tous les systèmes de fichiers dans `/etc/fstab`.

    ```
    [ec2-user ~]$ sudo mount -a
    ```

    Si la commande précédente ne génère pas d'erreur, votre fichier `/etc/fstab` est correct et votre système de fichiers sera monté automatiquement au prochain démarrage. Si la commande génère des erreurs, examinez celles-ci et essayez de corriger votre fichier `/etc/fstab`.

    > **Warning**
    >
    > Des erreurs dans le fichier `/etc/fstab` peuvent rendre un système impossible à démarrer. N'arrêtez pas un système dont le fichier `/etc/fstab` contient des erreurs.

    e. (Facultatif) Si vous n'êtes pas sûr de savoir comment corriger des erreurs dans `/etc/fstab`, vous avez toujours la possibilité de restaurer votre fichier `/etc/fstab` de sauvegarde avec la commande suivante.

```
[ec2-user ~]$ sudo mv /etc/fstab.orig /etc/fstab
```

Création d'instantanés de volumes d'une grappe RAID

Si vous voulez sauvegarder les données sur les volumes EBS d'une grappe RAID à l'aide d'instantanés, vous devez vous assurer que les instantanés sont cohérents. La raison en est que les instantanés de ces volumes sont créés indépendamment. La restauration de volumes EBS d'une grappe RAID à partir d'instantanés non synchronisés peut dégrader l'intégrité de la grappe.

Pour créer un ensemble cohérent d'instantanés pour votre grappe RAID, utilisez les instantanés multi-volume EBS. Les instantanés multi-volumes vous permettent de prendre des instantanés en lien avec la panne, aux données coordonnées et à un instant donné exact, sur plusieurs volumes EBS attachés à une instance EC2. Vous n'avez pas à arrêter votre instance pour la coordonner entre les volumes et assurer la régularité, car les instantanés sont automatiquement pris sur plusieurs volumes EBS. Pour de plus amples informations, veuillez consulter les étapes de la création d'instantanés multi-volume dans Création d'instantanés Amazon EBS.

Comparer les volumes EBS

Vous pouvez tester les performances des volumes Amazon EBS en simulant des charges de travail d'E/S. Procédez comme suit :

1. Lancez une instance optimisée EBS.
2. Créez des volumes EBS.
3. Attachez les volumes à votre instance optimisée pour EBS.
4. Configurez et installez le périphérique de stockage en mode bloc.
5. Installez un outil permettant de comparer les performances d'E/S.
6. Comparez les performances d'E/S de vos volumes.
7. Supprimez vos volumes et mettez l'instance hors service pour éviter de générer des frais.

> **Important**
>
> Certaines des procédures entraîneront la destruction des données existantes sur les volumes EBS que vous comparez. Les procédures de comparaison sont conçues pour être utilisées sur des volumes créés spécialement à des fins de tests, et pas sur des volumes de production.

Configurer votre instance

Afin d'obtenir des performances optimales des volumes EBS, nous vous recommandons d'utiliser une instance optimisée pour EBS. Les instances optimisées pour EBS offrent un débit supplémentaire et dédié entre Amazon EC2 et Amazon EBS, avec l'instance. Les instances optimisées pour EBS délivrent une bande passante dédiée entre Amazon EC2 et Amazon EBS, avec des spécifications en fonction du type d'instance utilisé. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Afin de créer une instance optimisée pour EBS, choisissez Launch as an EBS-Optimized instance (Lancer en tant qu'instance optimisée pour EBS) lorsque vous lancez l'instance à l'aide de la console Amazon EC2, ou spécifiez --ebs-optimized lorsque vous utilisez la ligne de commande. Veillez à lancer une instance de la génération actuelle, prenant en charge cette option. Pour de plus amples informations, veuillez consulter Amazon EBS–Instances optimisées (p. 1213).

Configuration des volumes IOPS provisionnés SSD (io1 and io2)

Pour créer un volume io1 ou io2, choisissez IOPS provisionnés SSD (**io1**) ou IOPS provisionnés SSD(**io2**) lorsque vous créez le volume via la console Amazon EC2 ou, dans la ligne de commande, spécifiez --volume-type io1|io2 --iopsn, où n correspond à un nombre entier compris entre 100 et 64,000. Pour connaître les spécifications plus détaillées des volumes EBS, veuillez consulter Types de volume Amazon EBS (p. 1089). Pour plus d'informations sur la création d'un volume EBS, consultez Création d'un volume Amazon EBS (p. 1107). Pour plus d'informations sur l'attachement d'un volume à une instance, consultez Attacher un volume Amazon EBS à une instance (p. 1109).

Pour les exemples de test, nous vous recommandons de créer une grappe RAID avec 6 volumes afin de bénéficier d'un niveau de performance élevé. Dans la mesure où vous êtes facturé en fonction des gigaoctets provisionnés (et du nombre d'IOPS provisionnés pour les volumes io1 and io2), et non du nombre de volumes, aucun coût supplémentaire ne sera appliqué pour la création de plusieurs volumes de plus petite taille, puis pour leur utilisation afin de créer un agrégat par bandes. Si vous utilisez Oracle Orion afin de comparer vos volumes, vous pouvez effectuer une simulation de l'agrégation par bandes comme avec Oracle ASM. C'est pourquoi nous vous recommandons de laisser Orion se charger de l'agrégation par bandes. Si vous utilisez un outil de comparaison différent, vous devez effectuer vous-même l'agrégation des volumes par bandes.

Pour créer un agrégat par bandes sur six volumes sous Amazon Linux, utilisez une commande similaire à la suivante :

```
[ec2-user ~]$ sudo mdadm --create /dev/md0 --level=0 --chunk=64 --raid-devices=6 /dev/sdf /
dev/sdg /dev/sdh /dev/sdi /dev/sdj /dev/sdk
```

Dans cet exemple, nous utilisons le système de fichiers XFS, mais vous devez utiliser celui qui correspond à vos besoins. Utilisez la commande suivante pour installer la prise en charge du système de fichiers XFS :

```
[ec2-user ~]$ sudo yum install -y xfsprogs
```

Ensuite, utilisez ces commandes pour créer, installer et affecter la propriété au système de fichiers XFS :

```
[ec2-user ~]$ sudo mkdir -p /mnt/p_iops_vol0 && sudo mkfs.xfs /dev/md0
[ec2-user ~]$ sudo mount -t xfs /dev/md0 /mnt/p_iops_vol0
[ec2-user ~]$ sudo chown ec2-user:ec2-user /mnt/p_iops_vol0/
```

Configuration des volumes HDD à débit optimisé (st1) ou Cold HDD (sc1)

Pour créer un volume st1, choisissez HDD à débit optimisé lorsque vous créez le volume via la console Amazon EC2 ou spécifiez --type **st1** si vous utilisez la ligne de commande. Pour créer un volume sc1, choisissez Cold HDD lorsque vous créez le volume via la console Amazon EC2 ou spécifiez --type **sc1** si vous utilisez la ligne de commande. Pour plus d'informations sur la création de volumes EBS, consultez Création d'un volume Amazon EBS (p. 1107). Pour plus d'informations sur la liaison de ces volumes à votre instance, consultez Attacher un volume Amazon EBS à une instance (p. 1109).

AWS fournit un modèle JSON à utiliser avec AWS CloudFormation, qui simplifie cette procédure de configuration. Accédez au modèle et enregistrez-le en tant que fichier JSON. AWS CloudFormation vous permet de configurer vos propres clés SSH et offre une méthode plus simple pour configurer un environnement de test de performances afin d'évaluer les volumes st1. Le modèle crée une instance de la génération actuelle et un volume st1 de 2 Tio, et attache ce dernier à l'instance dans /dev/xvdf.

Pour créer un volume HDD avec le modèle

1. Ouvrez la console AWS CloudFormation à l'adresse https://console.aws.amazon.com/cloudformation.
2. Choisissez Create Stack.

3. Choisissez Télécharger un modèle sur Amazon S3 et sélectionnez le modèle JSON que vous avez obtenu précédemment.
4. Attribuez un nom à votre stack (comme "test-perf-ebs"), puis sélectionnez un type d'instance (par défaut : r3.8xlarge) et une clé SSH.
5. Choisissez Suivant à deux reprises, puis sélectionnez Créer une pile.
6. Lorsque l'état de votre nouvelle pile passe de CREATE_IN_PROGRESS à COMPLETE, choisissez Sorties afin d'obtenir l'entrée DNS publique de votre nouvelle instance, qui sera attachée à un volume `st1` de 2 Tio.
7. Connectez-vous à votre nouvelle pile via SSH en tant qu'utilisateur **`ec2-user`**, avec le nom d'hôte obtenu à partir de l'entrée DNS lors de l'étape précédente.
8. Passez à l'Installer les outils d'évaluation (p. 1244).

Installer les outils d'évaluation

Le tableau suivant répertorie certains des outils que vous pouvez utiliser pour comparer les performances des volumes EBS.

Outil	Description

fio	Pour comparer les performances d'E/S. (Notez que la commande fio a une dépendance sur `libaio-devel`.)
	
	Pour installer fio sur Amazon Linux, exécutez la commande suivante :
	
	`[ec2-user ~]$ sudo yum install -y fio`
	
	Pour installer fio sur Ubuntu, exécutez la commande suivante :
	
	`sudo apt-get install -y fio`

Outil de calibrage Oracle Orion	Pour calibrer les performances d'E/S des systèmes de stockage à utiliser avec les bases de données Oracle.

Ces outils de comparaison prennent en charge un large éventail de paramètres de test. Vous devez utiliser des commandes proches des charges de travail que vos volumes devront prendre en charge. Les commandes ci-dessous sont proposées à titre d'exemple pour vous permettre de débuter.

Choisir la longueur de file d'attente d'un volume

Choisissez la meilleure longueur de file d'attente du volume en fonction de votre charge de travail et du type de volume.

Longueur de la file d'attente sur les volumes basés sur SSD

Afin de déterminer la longueur moyenne optimale de file d'attente pour votre charge de travail sur des volumes basés sur SSD, nous vous recommandons de cibler une longueur de file d'attente d'1 toutes les 1 000 IOPS disponibles (quantité de référence pour les volumes `gp2` et quantité provisionnée pour les volumes `io1` and `io2`). Vous pouvez ensuite contrôler les performances de votre application et ajuster cette valeur en fonction des exigences de votre application.

L'augmentation de la longueur de file d'attente offre un avantage jusqu'à ce que vous atteigniez le nombre d'IOPS provisionnés, le débit ou la valeur optimale de la longueur de file d'attente du système, actuellement définie sur 32. Par exemple, un volume avec 3 000 IOPS provisionnés doit cibler une longueur de file

d'attente de 3. Vous devez essayer d'augmenter ou de diminuer ces valeurs afin de déterminer ce qui fonctionne le mieux pour votre application.

Longueur de la file d'attente sur les volumes basés sur HDD

Afin de déterminer la longueur moyenne optimale de file d'attente pour votre charge de travail sur des volumes basés sur HDD, nous vous recommandons de cibler une longueur de file d'attente de 4 tout en exécutant des E/S séquentielles d'1 Mio. Vous pouvez ensuite contrôler les performances de votre application et ajuster cette valeur en fonction des exigences de votre application. Par exemple, un volume `st1` de 2 Tio avec un débit de transmission en rafales de 500 Mio/s et des IOPS de 500 doit cibler une longueur de file d'attente de 4, 8 ou 16 lors de l'exécution d'E/S séquentielles de 1 024 Kio, 512 Kio ou 256 Kio respectivement. Vous devez essayer d'augmenter ou de diminuer ces valeurs afin de déterminer ce qui fonctionne le mieux pour votre application.

Désactivation des états « C-state »

Avant de procéder à des comparaisons, vous devez désactiver les états « C-state » du processeur. Les cœurs temporairement inutilisés dans une UC prise en charge peuvent passer à l'état « C-state » pour économiser de l'énergie. Lorsque le cœur est appelé afin de reprendre le traitement, un certain laps de temps est nécessaire avant que le cœur soit à nouveau entièrement opérationnel. Cette latence peut interférer avec les routines de comparaison du processeur. Pour plus d'informations sur les états « C-state » et les types d'instance EC2 qui les prennent en charge, consultez Contrôle des états du processeur pour votre instance EC2.

Désactivation des états C sous Linux

Vous pouvez désactiver les états « C-state » sur Amazon Linux, RHEL et CentOS de la manière suivante :

1. Identifiez le nombre d'états « C-state ».

```
$ cpupower idle-info | grep "Number of idle states:"
```

2. Désactivez les états « C-state » de c1 à cN. Idéalement, l'état des cœurs doit être c0.

```
$ for i in `seq 1 $((N-1))`; do cpupower idle-set -d $i; done
```

Effectuer la comparaison

Les procédures suivantes décrivent les commandes de comparaison pour différents types de volume EBS.

Exécutez les commandes suivantes sur une instance optimisée pour EBS avec les volumes EBS attachés. Si les volumes EBS ont été créés à partir d'instantanés, veillez à les initialiser avant d'effectuer la comparaison. Pour plus d'informations, consultez Initialisation des volumes Amazon EBS (p. 1235).

Une fois que vous avez terminé de tester les volumes, consultez les rubriques suivantes pour apprendre à les nettoyer : Suppression d'un volume Amazon EBS (p. 1129) et Résilier une instance (p. 645).

Comparaison des volumes `io1` and `io2`

Exécutez la commande fio sur l'ensemble de bandes que vous avez créé.

La commande suivante effectue des opérations d'écriture aléatoires 16 Ko.

```
[ec2-user ~]$ sudo fio --directory=/mnt/p_iops_vol0 --name fio_test_file --direct=1 --
rw=randwrite --bs=16k --size=1G --numjobs=16 --time_based --runtime=180 --group_reporting
 --norandommap
```

La commande suivante effectue des opérations de lecture aléatoires 16 Ko.

```
[ec2-user ~]$ sudo fio --directory=/mnt/p_iops_vol0 --name fio_test_file --direct=1 --
rw=randread --bs=16k --size=1G --numjobs=16 --time_based --runtime=180 --group_reporting --
norandommap
```

Pour plus d'informations sur l'interprétation des résultats, consultez le didacticiel Inspecting disk IO performance with fio.

Comparaison des volumes `st1` et `sc1`

Exécutez la commande fio sur votre volume `st1` ou `sc1`.

> **Note**
>
> Avant d'exécuter ces tests, définissez les E/S mises en mémoire tampon sur votre instance, comme indiqué dans Accroître la lecture anticipée pour les charges de travail à forte densité de lectures et à haut débit sur `st1` et `sc1` (p. 1232).

La commande suivante exécute des opérations de lecture séquentielle d'1 Mio sur un périphérique de stockage en mode bloc `st1` attaché (par exemple, `/dev/xvdf`) :

```
[ec2-user ~]$ sudo fio --filename=/dev/<device> --direct=1 --rw=read --randrepeat=0
 --ioengine=libaio --bs=1024k --iodepth=8 --time_based=1 --runtime=180 --
name=fio_direct_read_test
```

La commande suivante exécute des opérations d'écriture séquentielle d'1 Mio sur un périphérique de stockage en mode bloc `st1` attaché :

```
[ec2-user ~]$ sudo fio --filename=/dev/<device> --direct=1 --rw=write --randrepeat=0
 --ioengine=libaio --bs=1024k --iodepth=8 --time_based=1 --runtime=180 --
name=fio_direct_write_test
```

Certaines charges de travail exécutent une combinaison de lectures séquentielles et d'écritures séquentielles dans différentes parties du périphérique de stockage en mode bloc. Pour évaluer une telle charge de travail, nous vous recommandons d'utiliser des tâches fio distinctes et simultanées pour les lectures et les écritures, et d'utiliser l'option fio `offset_increment` pour cibler différents emplacements du périphérique de stockage en mode bloc pour chaque tâche.

L'exécution de cette charge de travail est un peu plus compliquée qu'une charge de travail d'écriture séquentielle ou de lecture séquentielle. Utilisez un éditeur de texte pour créer un fichier de tâche fio, appelé `fio_rw_mix.cfg` dans cet exemple, contenant les éléments suivants :

```
[global]
clocksource=clock_gettime
randrepeat=0
runtime=180
offset_increment=100g

[sequential-write]
bs=1M
ioengine=libaio
direct=1
iodepth=8
filename=/dev/<device>
do_verify=0
rw=write
rwmixread=0
rwmixwrite=100

[sequential-read]
bs=1M
```

```
ioengine=libaio
direct=1
iodepth=8
filename=/dev/<device>
do_verify=0
rw=read
rwmixread=100
rwmixwrite=0
```

Ensuite, exécutez la commande suivante :

```
[ec2-user ~]$ sudo fio fio_rw_mix.cfg
```

Pour plus d'informations sur l'interprétation des résultats, consultez le didacticiel Inspecting disk IO performance with fio.

Plusieurs tâches fio pour l'I/O directe, même en cas d'utilisation d'opérations de lecture ou d'écriture séquentielle, peuvent se traduire par un débit inférieur à celui attendu pour les volumes `st1` et `sc1`. Nous vous recommandons d'utiliser une tâche d'E/S directe et le paramètre `iodepth` pour contrôler le nombre d'opérations d'E/S simultanées.

Métriques Amazon CloudWatch pour Amazon EBS

Les métriques CloudWatch sont des données statistiques que vous pouvez utiliser pour afficher, analyser et définir des alarmes sur le comportement opérationnel de vos volumes.

Le tableau suivant décrit les types de surveillance des données disponibles pour vos volumes Amazon EBS.

Type	Description
Base	Les données sont disponibles automatiquement toutes les 5 minutes sans coût aucun. Elles incluent celles des volumes du périphérique racine des instances basées sur les volumes EBS.
Détaillé :	Les volumes IOPS provisionnés SSD (`io1` and `io2`) envoient automatiquement les métriques toutes les minutes à CloudWatch.

Lorsque vous obtenez les données de CloudWatch, vous pouvez inclure un paramètre de demande `Period` pour spécifier la granularité des données retournées. Cette option est différente de la période que nous utilisons quand nous collectons les données (périodes de 5 minutes). Il est recommandé de spécifier une période dans votre demande qui soit égale ou supérieure à la période de collection pour garantir que les données retournées sont valides.

Vous pouvez obtenir les données à l'aide de l'API CloudWatch ou de la console Amazon EC2. La console prend les données brutes de l'API CloudWatch et affiche une série de graphiques basés sur ces données. En fonction de vos besoins, vous pouvez préfèrerez utiliser les données de l'API ou les graphiques de la console.

Métriques Amazon EBS

Amazon Elastic Block Store (Amazon EBS) envoie les points de données à CloudWatch pour plusieurs métriques. Amazon EBS Les volumes à usage général SSD (gp2), HDD à débit optimisé (st1), HDD à froid (sc1) et les volumes magnétiques (standard) envoient automatiquement des métriques toutes les cinq minutes à CloudWatch. Les volumes SSD IOPS provisionnés (`io1` and `io2`) envoient automatiquement des métriques toutes les minutes à CloudWatch. Les données sont uniquement présentées à CloudWatch lorsque le volume est attaché à une instance.

Certaines de ces métriques présentent des différences sur les instances basées sur Nitro. Pour obtenir la liste des types d'instances basées sur le système Nitro, consultez Instances reposant sur le système Nitro (p. 218).

L'espace de noms `AWS/EBS` inclut les métriques suivantes.

Métriques

- Métriques de volume (p. 1248)
- Métriques de la fonction de restauration d'instantané rapide (p. 1252)

Métriques de volume

L'espace de noms `AWS/EBS` inclut les métriques suivantes pour les volumes EBS. Pour obtenir des informations sur l'espace disque disponible à partir du système d'exploitation sur une instance, veuillez consulter Affichage de l'espace disque disponible (p. 1118).

Métrique	Description
`VolumeReadBytes`	Fournit des informations sur les opérations de lecture au cours d'une période donnée. Les rapports statistiques `Sum` indiquent le nombre total d'octets transférés pendant la période. Les rapports statistiques `Average` indiquent la taille moyenne de chaque opération de lecture durant la période, sauf pour les volumes attachés à une instance basée sur Nitro, pour lesquels la moyenne représente la moyenne sur la période spécifiée. La statistique `SampleCount` indique le nombre total d'opérations de lecture durant la période, sauf pour les volumes attachés à une instance basée sur Nitro, pour lesquels le nombre d'échantillons représente le nombre de points de données utilisés pour le calcul statistique. Pour les instances Xen, les données sont présentées uniquement lorsqu'une activité de lecture se produit sur le volume. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : octets
`VolumeWriteBytes`	Fournit des informations sur les opérations d'écriture au cours d'une période donnée. Les rapports statistiques `Sum` indiquent le nombre total d'octets transférés pendant la période. Les rapports statistiques `Average` indiquent la taille moyenne de chaque opération d'écriture durant la période, sauf pour les volumes attachés à une instance basée sur Nitro, pour lesquels la moyenne représente la moyenne sur la période spécifiée. La statistique `SampleCount` indique le nombre total d'opérations d'écriture durant la période, sauf pour les volumes attachés à une instance basée sur Nitro, pour lesquels le nombre d'échantillons représente le nombre de points de données utilisés pour le calcul statistique. Pour les instances Xen, les données sont présentées uniquement lorsqu'une activité d'écriture se produit sur le volume. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : octets

Métrique	Description
`VolumeReadOps`	Nombre total d'opérations de lecture au cours d'une période donnée. Pour calculer la moyenne d'opérations de lecture d'E/S par seconde (IOPS en lecture) pour la période, divisez le nombre total d'opérations de lecture de la période par le nombre de secondes de la période. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : nombre
`VolumeWriteOps`	Nombre total d'opérations d'écriture au cours d'une période donnée. Pour calculer la moyenne d'opérations d'écriture d'E/S par seconde (IOPS en écriture) pour la période, divisez le nombre total d'opérations d'écriture de la période par le nombre de secondes de la période. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : nombre
`VolumeTotalReadTime`	Note Cette métrique n'est pas prise en charge avec les volumes activés pour attachement multiple. Nombre total de secondes passées par toutes les opérations de lecture terminées, au cours d'une période donnée. Si plusieurs demandes sont soumises en même temps, ce total peut être supérieur à la durée de la période. Par exemple, pour une période de 5 minutes (300 secondes) : si 700 opérations ont été réalisées au cours de cette période et que chaque opération a pris une seconde, la valeur serait 700 secondes. Pour les instances Xen, les données sont présentées uniquement lorsqu'une activité de lecture se produit sur le volume. La statistique `Average` sur cette métrique n'est pas pertinente pour les volumes attachés à des instances basées sur Nitro. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : secondes

Métrique	Description
`VolumeTotalWriteTime`	**Note** Cette métrique n'est pas prise en charge avec les volumes activés pour attachement multiple. Nombre total de secondes passées par toutes les opérations d'écriture terminées, au cours d'une période donnée. Si plusieurs demandes sont soumises en même temps, ce total peut être supérieur à la durée de la période. Par exemple, pour une période de 5 minutes (300 secondes) : si 700 opérations ont été réalisées au cours de cette période et que chaque opération a pris une seconde, la valeur serait 700 secondes. Pour les instances Xen, les données sont présentées uniquement lorsqu'une activité d'écriture se produit sur le volume. La statistique `Average` sur cette métrique n'est pas pertinente pour les volumes attachés à des instances basées sur Nitro. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : secondes
`VolumeIdleTime`	**Note** Cette métrique n'est pas prise en charge avec les volumes activés pour attachement multiple. Nombre total de secondes dans une période données, alors qu'aucune opération de lecture ou écriture n'a été soumise. La statistique `Average` sur cette métrique n'est pas pertinente pour les volumes attachés à des instances basées sur Nitro. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : secondes
`VolumeQueueLength`	Nombre de demandes d'opérations de lecture et d'écriture en attente de réalisation au cours d'une période donnée. La statistique `Sum` sur cette métrique n'est pas pertinente pour les volumes attachés à des instances basées sur Nitro. Les statistiques `Minimum` et `Maximum` sur cette métrique sont uniquement prises en charge par des volumes attachés à des instances basées sur Nitro. Unités : nombre

Métrique	Description
`VolumeThroughputPercentage`	**Note** Cette métrique n'est pas prise en charge avec les volumes activés pour attachement multiple. Utilisé uniquement avec les volumes IOPS provisionnés SSD. Pourcentage d'opérations d'E/S par seconde (IOPS) fournies par rapport au total IOPS provisionné pour un volume Amazon EBS. Les volumes IOPS provisionnés SSD fournissent leurs performances IOPS provisionnées 99,9 % du temps. Pendant une écriture, s'il n'y a aucune autre requête d'E/S en suspens en une minute, la valeur de la métrique est 100 %. De plus, les performances d'E/S d'un volume peuvent chuter temporairement en raison d'une action que vous avez effectuée (par exemple, création de l'instantané d'un volume lors de l'utilisation au cours d'une période de pointe, exécution du volume sur une instance EBS non optimisée ou premier accès aux données sur le volume). Unités : pourcentage
`VolumeConsumedReadWriteOps`	Utilisé uniquement avec les volumes IOPS provisionnés SSD. Nombre total d'opérations de lecture et d'écriture (normalisé selon les unités de capacité 256 K) utilisées au cours d'une période donnée. Les opérations d'E/S inférieures à 256 Ko chacune comptent comme 1 IOPS consommé. Les opérations d'E/S supérieures à 256 K sont comptées dans les unités de capacité de 256 K. Par exemple, une E/S de 1 024 K compte comme 4 IOPS consommées. Unités : nombre
`BurstBalance`	Utilisé avec les volumes SSD à usage général (`gp2`), HDD à débit optimisé (`st1`) et Cold HDD (`sc1`) uniquement. Fournit des informations concernant le pourcentage de crédits d'E/S (pour `gp2`) ou de crédits de débit (pour `st1` et `sc1`) restant dans le compartiment en rafales. Les données sont présentées à CloudWatch uniquement lorsque le volume est actif. Si le volume n'est pas attaché, aucune donnée n'est présentée. La statistique `Sum` sur cette métrique n'est pas pertinente pour les volumes attachés à des instances basées sur Nitro. Si les performances de base du volume dépassent les performances en rafale maximales, les crédits ne sont jamais dépensés. Si le volume est attaché à une instance basée sur Nitro, le solde de la rafale n'est pas présenté. Pour une instance non basée sur Nitro, le solde de la rafale présenté est de 100 %. Pour plus d'informations, consultez Crédits E/S et performances en rafale (p. 1091). Unités : pourcentage

Métriques de la fonction de restauration d'instantané rapide

L'espace de noms `AWS/EBS` inclut les métriques suivantes pour une restauration d'instantané rapide (p. 1205).

Métrique	Description
`FastSnapshotRestoreCreditsBucketSize`	Nombre maximum de crédits de création de volume pouvant être accumulés. Cette métrique est signalée par instantané et par zone de disponibilité. La statistique la plus significative est `Average`. Les résultats des statistiques `Minimum` et `Maximum` sont les mêmes que ceux de `Average` et peuvent être utilisés indifféremment.
`FastSnapshotRestoreCreditsBalance`	Nombre de crédits de création de volume disponibles. Cette métrique est signalée par instantané et par zone de disponibilité. La statistique la plus significative est `Average`. Les résultats des statistiques `Minimum` et `Maximum` sont les mêmes que ceux de `Average` et peuvent être utilisés indifféremment.

Dimensions pour les métriques Amazon EBS

La dimension prise en charge est l'ID de volume (`VolumeId`). Toutes les statistiques disponibles sont filtrées par ID de volume.

Pour les métriques de volume (p. 1248), la dimension prise en charge est l'ID de volume (`VolumeId`). Toutes les statistiques disponibles sont filtrées par ID de volume.

Pour les métriques de restauration d'instantané rapide (p. 1252), les dimensions prises en charge sont l'ID d'instantané (`SnapshotId`) et la zone de disponibilité (`AvailabilityZone`).

Graphiques de la console Amazon EC2

Après avoir créé un volume, vous pouvez afficher les graphiques de surveillance du volume dans la console Amazon EC2. Sélectionnez un volume dans la page Volumes de la console, puis sélectionnez Surveillance. Le tableau ci-après répertorie les graphiques affichés. La colonne de droite décrit l'utilisation des métriques de données brutes de l'API CloudWatch pour produire chaque graphique. La période de tous les graphiques est de 5 minutes.

Graphique	Description de l'utilisation des métriques brutes
Bande passante de lecture (Kbits/s)	`Sum(VolumeReadBytes) / Period / 1024`
Bande passante d'écriture (Kbits/s)	`Sum(VolumeWriteBytes) / Period / 1024`
Débit de lecture (IOPS)	`Sum(VolumeReadOps) / Period`
Débit d'écriture (IOPS)	`Sum(VolumeWriteOps) / Period`
Longueur moyenne de file d'attente (opérations)	`Avg(VolumeQueueLength)`
% temps inactif	`Sum(VolumeIdleTime) / Period × 100`

Graphique	Description de l'utilisation des métriques brutes

Taille de lecture moyenne (Kio/opération)	`Avg(VolumeReadBytes) / 1024`
	
	Pour les instances basées sur Nitro, la formule suivante permet de déduire la taille de lecture moyenne en utilisant les mathématiques appliquées aux métriques CloudWatch :
	
	`(Sum(VolumeReadBytes) / Sum(VolumeReadOps)) / 1024`
	
	Les métriques `VolumeReadBytes` et `VolumeReadOps` sont disponibles dans la console EBS CloudWatch.

Taille d'écriture moyenne (Kio/opération)	`Avg(VolumeWriteBytes) / 1024`
	
	Pour les instances basées sur Nitro, la formule suivante permet de déduire la taille d'écriture moyenne en utilisant les mathématiques appliquées aux métriques CloudWatch :
	
	`(Sum(VolumeWriteBytes) / Sum(VolumeWriteOps)) / 1024`
	
	Les métriques `VolumeWriteBytes` et `VolumeWriteOps` sont disponibles dans la console EBS CloudWatch.

Latence de lecture moyenne (ms/opération)	`Avg(VolumeTotalReadTime) × 1000`
	
	Pour les instances basées sur Nitro, la formule suivante permet de déduire la latence de lecture moyenne en utilisant les mathématiques appliquées aux métriques CloudWatch :
	
	`(Sum(VolumeTotalReadTime) / Sum(VolumeReadOps)) × 1000`
	
	Les métriques `VolumeTotalReadTime` et `VolumeReadOps` sont disponibles dans la console EBS CloudWatch.

Latence d'écriture moyenne (ms/opération)	`Avg(VolumeTotalWriteTime) × 1000`
	
	Pour les instances basées sur Nitro, la formule suivante permet de déduire la latence d'écriture moyenne en utilisant les mathématiques appliquées aux métriques CloudWatch :
	
	`(Sum(VolumeTotalWriteTime) / Sum(VolumeWriteOps)) * 1000`
	
	Les métriques `VolumeTotalWriteTime` et `VolumeWriteOps` sont disponibles dans la console EBS CloudWatch.

Pour les graphiques de latence moyenne et ceux de taille moyenne, la moyenne est calculée par rapport au nombre total d'opérations (lecture ou écriture, quel que soit celui applicable au graphe) complétées durant la période.

Amazon CloudWatch Events pour Amazon EBS

Amazon EBS émet des notifications basées sur Amazon CloudWatch Events pour diverses modifications d'état de volume, d'instantané et de chiffrement. Avec CloudWatch Events, vous pouvez établir des

règles qui déclenchent des actions par programmation en réponse à une modification d'état du volume, de l'instantané ou de la clé de chiffrement. Par exemple, lorsqu'un instantané est créé, vous pouvez déclencher une fonction AWS Lambda pour partager l'instantané terminé avec un autre compte ou le copier vers une autre région à des fins de reprise après sinistre.

Les événements d'CloudWatch sont représentés comme des objets JSON. Les champs spécifiques à l'événement figurent dans la section « détail » de l'objet JSON. Le champ « événement » contient le nom de l'événement. Le champ « résultat » contient l'état terminé de l'action qui déclenche l'événement. Pour plus d'informations, consultez Modèles d'événements dans CloudWatch Events dans le Guide de l'utilisateur Amazon CloudWatch Events.

Pour plus d'informations, consultez Utilisation d'événements dans le Guide de l'utilisateur Amazon CloudWatch.

Table des matières

- Événements de volume EBS (p. 1254)
- Événements d'instantané EBS (p. 1257)
- Événements de modification de volume EBS (p. 1261)
- Événements de restauration d'instantané rapide EBS (p. 1261)
- Utilisation de AWS Lambda pour gérer les événements CloudWatch (p. 1262)

Événements de volume EBS

Amazon EBS envoie des événements à CloudWatch Events lorsque les événements de volume suivants se produisent.

Événements

- Créer un volume (createVolume) (p. 1254)
- Supprimer le volume (deleteVolume) (p. 1255)
- Attacher ou réattacher le volume (attachVolume, reattachVolume) (p. 1256)

Créer un volume (createVolume)

L'événement `createVolume` est envoyé à votre compte AWS lorsqu'une action de création d'un volume est terminée. Toutefois, il n'est pas enregistré, consigné ou archivé. Cet événement peut avoir le résultat `available` ou `failed`. La création échoue si une clé KMS non valide est fournie, comme illustré dans les exemples ci-dessous.

Données d'événement

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS pour un événement `createVolume` réussi.

```
{
   "version": "0",
   "id": "01234567-0123-0123-0123-012345678901",
   "detail-type": "EBS Volume Notification",
   "source": "aws.ec2",
   "account": "012345678901",
   "time": "yyyy-mm-ddThh:mm:ssZ",
   "region": "us-east-1",
   "resources": [
      "arn:aws:ec2:us-east-1:012345678901:volume/vol-01234567"
   ],
   "detail": {
      "result": "available",
```

```
      "cause": "",
      "event": "createVolume",
      "request-id": "01234567-0123-0123-0123-0123456789ab"
   }
}
```

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS après un événement `createVolume` ayant échoué. La cause de l'échec est attribuée à une clé KMS désactivée.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-0123456789ab",
  "detail-type": "EBS Volume Notification",
  "source": "aws.ec2",
  "account": "012345678901",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "sa-east-1",
  "resources": [
    "arn:aws:ec2:sa-east-1:0123456789ab:volume/vol-01234567",
  ],
  "detail": {
    "event": "createVolume",
    "result": "failed",
    "cause": "arn:aws:kms:sa-east-1:0123456789ab:key/01234567-0123-0123-0123-0123456789ab
 is disabled.",
    "request-id": "01234567-0123-0123-0123-0123456789ab",
  }
}
```

Voici un exemple d'objet JSON émis par EBS après un événement `createVolume` ayant échoué. La cause de l'échec est attribuée à l'importation en attente d'une clé CMK.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-0123456789ab",
  "detail-type": "EBS Volume Notification",
  "source": "aws.ec2",
  "account": "012345678901",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "sa-east-1",
  "resources": [
    "arn:aws:ec2:sa-east-1:0123456789ab:volume/vol-01234567",
  ],
  "detail": {
    "event": "createVolume",
    "result": "failed",
    "cause": "arn:aws:kms:sa-east-1:0123456789ab:key/01234567-0123-0123-0123-0123456789ab
 is pending import.",
    "request-id": "01234567-0123-0123-0123-0123456789ab",
  }
}
```

Supprimer le volume (deleteVolume)

L'événement `deleteVolume` est envoyé à votre compte AWS lorsqu'une action de suppression d'un volume est terminée. Toutefois, il n’est pas enregistré, consigné ou archivé. Le résultat de cet événement est `deleted`. Si la suppression ne se termine pas, l'événement n'est pas envoyé.

Données d'événement

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS pour un événement `deleteVolume` réussi.

```
{
   "version": "0",
   "id": "01234567-0123-0123-0123-012345678901",
   "detail-type": "EBS Volume Notification",
   "source": "aws.ec2",
   "account": "012345678901",
   "time": "yyyy-mm-ddThh:mm:ssZ",
   "region": "us-east-1",
   "resources": [
      "arn:aws:ec2:us-east-1:012345678901:volume/vol-01234567"
   ],
   "detail": {
      "result": "deleted",
      "cause": "",
      "event": "deleteVolume",
      "request-id": "01234567-0123-0123-0123-0123456789ab"
   }
}
```

Attacher ou réattacher le volume (attachVolume, reattachVolume)

L'événement `attachVolume` ou `reattachVolume` est envoyé à votre compte AWS si un volume ne parvient pas à s'attacher ou à se rattacher à une instance. Toutefois, il n'est pas enregistré, consigné ou archivé. Si vous utilisez une clé KMS pour chiffrer un volume EBS et que la clé devient non valide, EBS émet un événement si cette clé est utilisée ultérieurement pour l'attachement ou le rattachement d'un volume, comme illustré dans les exemples ci-dessous.

Données d'événement

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS après un événement `attachVolume` ayant échoué. La cause de l'échec est attribuée à la suppression en attente d'une clé CMK.

> **Note**
>
> AWS peut tenter le rattachement d'un volume après la maintenance habituelle des serveurs.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-0123456789ab",
  "detail-type": "EBS Volume Notification",
  "source": "aws.ec2",
  "account": "012345678901",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
  "arn:aws:ec2:us-east-1:0123456789ab:volume/vol-01234567",
  "arn:aws:kms:us-east-1:0123456789ab:key/01234567-0123-0123-0123-0123456789ab"
  ],
  "detail": {
    "event": "attachVolume",
    "result": "failed",
    "cause": "arn:aws:kms:us-east-1:0123456789ab:key/01234567-0123-0123-0123-0123456789ab
 is pending deletion.",
    "request-id": ""
  }
}
```

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS après un événement `reattachVolume` ayant échoué. La cause de l'échec est attribuée à la suppression en attente d'une clé CMK.

```
{
```

```
  "version": "0",
  "id": "01234567-0123-0123-0123-0123456789ab",
  "detail-type": "EBS Volume Notification",
  "source": "aws.ec2",
  "account": "012345678901",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
  "arn:aws:ec2:us-east-1:0123456789ab:volume/vol-01234567",
  "arn:aws:kms:us-east-1:0123456789ab:key/01234567-0123-0123-0123-0123456789ab"
  ],
  "detail": {
    "event": "reattachVolume",
    "result": "failed",
    "cause": "arn:aws:kms:us-east-1:0123456789ab:key/01234567-0123-0123-0123-0123456789ab
 is pending deletion.",
    "request-id": ""
  }
}
```

Événements d'instantané EBS

Amazon EBS envoie des événements à CloudWatch Events lorsque les événements de volume suivants se produisent.

Événements

- Créer un instantané (createSnapshot) (p. 1257)
- Créer des instantanés (createSnapshots) (p. 1258)
- Copier un instantané (copySnapshot) (p. 1259)
- Partager un instantané (shareSnapshot) (p. 1260)

Créer un instantané (createSnapshot)

L'événement `createSnapshot` est envoyé à votre compte AWS lorsqu'une action de création d'un instantané est terminée. Toutefois, il n'est pas enregistré, consigné ou archivé. Cet événement peut avoir le résultat `succeeded` ou `failed`.

Données d'événement

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS pour un événement `createSnapshot` réussi. Dans la section `detail`, le champ `source` contient l'ARN du volume source. Les champs `startTime` et `endTime` indiquent le moment où la création de l'instantané a commencé et est terminée.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "EBS Snapshot Notification",
  "source": "aws.ec2",
  "account": "012345678901",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
     "arn:aws:ec2:us-west-2::snapshot/snap-01234567"
  ],
  "detail": {
    "event": "createSnapshot",
    "result": "succeeded",
    "cause": "",
    "request-id": "",
    "snapshot_id": "arn:aws:ec2:us-west-2::snapshot/snap-01234567",
```

```
    "source": "arn:aws:ec2:us-west-2::volume/vol-01234567",
    "startTime": "yyyy-mm-ddThh:mm:ssZ",
    "endTime": "yyyy-mm-ddThh:mm:ssZ"  }
}
```

Créer des instantanés (createSnapshots)

L'événement `createSnapshots` est envoyé à votre compte AWS lorsqu'une action de création d'un instantané multi-volume est terminée. Cet événement peut avoir le résultat `succeeded` ou `failed`.

Données d'événement

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS pour un événement `createSnapshots` réussi. Dans la section `detail`, le champ `source` contient les ARN des volumes sources de l'ensemble d'instantanés multi-volumes. Les champs `startTime` et `endTime` indiquent le moment où la création de l'instantané a commencé et est terminée.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "EBS Multi-Volume Snapshots Completion Status",
  "source": "aws.ec2",
  "account": "012345678901",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
    "arn:aws:ec2::us-east-1:snapshot/snap-01234567",
    "arn:aws:ec2::us-east-1:snapshot/snap-012345678"
  ],
  "detail": {
    "event": "createSnapshots",
    "result": "succeeded",
    "cause": "",
    "request-id": "",
    "startTime": "yyyy-mm-ddThh:mm:ssZ",
    "endTime": "yyyy-mm-ddThh:mm:ssZ",
    "snapshots": [
      {
        "snapshot_id": "arn:aws:ec2::us-east-1:snapshot/snap-01234567",
        "source": "arn:aws:ec2::us-east-1:volume/vol-01234567",
        "status": "completed"
      },
      {
        "snapshot_id": "arn:aws:ec2::us-east-1:snapshot/snap-012345678",
        "source": "arn:aws:ec2::us-east-1:volume/vol-012345678",
        "status": "completed"
      }
    ]
  }
}
```

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS après un événement `createSnapshots` ayant échoué. La cause de l'échec correspondait à un ou plusieurs instantanés qui n'ont pas pu aboutir. Les valeurs de `snapshot_id` sont les ARN des instantanés qui ont échoué. `startTime` et `endTime` représentent les instants où l'action de création de l'instantané a commencé et s'est terminée.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "EBS Multi-Volume Snapshots Completion Status",
  "source": "aws.ec2",
  "account": "012345678901",
```

```
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
    "arn:aws:ec2::us-east-1:snapshot/snap-01234567",
    "arn:aws:ec2::us-east-1:snapshot/snap-012345678"
  ],
 "detail": {
    "event": "createSnapshots",
    "result": "failed",
    "cause": "Snapshot snap-01234567 is in status deleted",
    "request-id": "",
    "startTime": "yyyy-mm-ddThh:mm:ssZ",
    "endTime": "yyyy-mm-ddThh:mm:ssZ",
    "snapshots": [
      {
        "snapshot_id": "arn:aws:ec2::us-east-1:snapshot/snap-01234567",
        "source": "arn:aws:ec2::us-east-1:volume/vol-01234567",
        "status": "error"
      },
      {
        "snapshot_id": "arn:aws:ec2::us-east-1:snapshot/snap-012345678",
        "source": "arn:aws:ec2::us-east-1:volume/vol-012345678",
        "status": "deleted"
      }
    ]
  }
}
```

Copier un instantané (copySnapshot)

L'événement `copySnapshot` est envoyé à votre compte AWS lorsqu'une action de copie d'un instantané est terminée. Toutefois, il n'est pas enregistré, consigné ou archivé. Cet événement peut avoir le résultat `succeeded` ou `failed`.

Données d'événement

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS après un événement `copySnapshot` réussi. La valeur `snapshot_id` correspond à l'ARN de l'instantané nouvellement créé. Dans la section `detail`, la valeur `source` correspond à l'ARN de l'instantané source. `startTime` et `endTime` représentent le moment où l'action de copie de l'instantané a démarré et s'est terminée.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "EBS Snapshot Notification",
  "source": "aws.ec2",
  "account": "123456789012",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
    "arn:aws:ec2:us-west-2::snapshot/snap-01234567"
  ],
  "detail": {
    "event": "copySnapshot",
    "result": "succeeded",
    "cause": "",
    "request-id": "",
    "snapshot_id": "arn:aws:ec2:us-west-2::snapshot/snap-01234567",
    "source": "arn:aws:ec2:eu-west-1::snapshot/snap-76543210",
    "startTime": "yyyy-mm-ddThh:mm:ssZ",
    "endTime": "yyyy-mm-ddThh:mm:ssZ",
    "Incremental": "True"
  }
```

```
}
```

La liste ci-dessous est l'exemple d'un objet JSON émis par EBS après un événement `copySnapshot` ayant échoué. La cause de l'échec est attribuée à un ID d'instantané source non valide. La valeur `snapshot_id` correspond à l'ARN de l'instantané ayant échoué. Dans la section `detail`, la valeur `source` correspond à l'ARN de l'instantané source. `startTime` et `endTime` représentent le moment où l'action de copie de l'instantané a démarré et s'est terminée.

```
{
  "version": "0",
  "id": "01234567-0123-0123-0123-012345678901",
  "detail-type": "EBS Snapshot Notification",
  "source": "aws.ec2",
  "account": "123456789012",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
    "arn:aws:ec2:us-west-2::snapshot/snap-01234567"
  ],
  "detail": {
    "event": "copySnapshot",
    "result": "failed",
    "cause": "Source snapshot ID is not valid",
    "request-id": "",
    "snapshot_id": "arn:aws:ec2:us-west-2::snapshot/snap-01234567",
    "source": "arn:aws:ec2:eu-west-1::snapshot/snap-76543210",
    "startTime": "yyyy-mm-ddThh:mm:ssZ",
    "endTime": "yyyy-mm-ddThh:mm:ssZ"
  }
}
```

Partager un instantané (shareSnapshot)

L'événement `shareSnapshot` est envoyé à votre compte AWS lorsqu'un autre compte partage un instantané avec lui. Toutefois, il n'est pas enregistré, consigné ou archivé. Le résultat est toujours `succeeded`.

Données d'événement

Ce qui suit est un exemple d'objet JSON émis par EBS après un événement `shareSnapshot` terminé. Dans la section `detail`, la valeur `source` correspond au numéro de compte AWS de l'utilisateur qui a partagé l'instantané avec vous. `startTime` et `endTime` représentent le moment où l'action de partage de l'instantané a démarré et s'est terminée. L'événement `shareSnapshot` est émis uniquement lorsqu'un instantané privé est partagé avec un autre utilisateur. Le partage d'un instantané public ne déclenche pas l'événement.

```
{
  "version": "0",
  "id": "01234567-01234-0123-0123-012345678901",
  "detail-type": "EBS Snapshot Notification",
  "source": "aws.ec2",
  "account": "012345678901",
  "time": "yyyy-mm-ddThh:mm:ssZ",
  "region": "us-east-1",
  "resources": [
    "arn:aws:ec2:us-west-2::snapshot/snap-01234567"
  ],
  "detail": {
    "event": "shareSnapshot",
    "result": "succeeded",
    "cause": "",
```

```
    "request-id": "",
    "snapshot_id": "arn:aws:ec2:us-west-2::snapshot/snap-01234567",
    "source": 012345678901,
    "startTime": "yyyy-mm-ddThh:mm:ssZ",
    "endTime": "yyyy-mm-ddThh:mm:ssZ"
  }
}
```

Événements de modification de volume EBS

Amazon EBS envoie des événements `modifyVolume` à CloudWatch Events lorsqu'un volume est modifié. Toutefois, il n'est pas enregistré, consigné ou archivé.

```
{
   "version": "0",
   "id": "01234567-0123-0123-0123-012345678901",
   "detail-type": "EBS Volume Notification",
   "source": "aws.ec2",
   "account": "012345678901",
   "time": "yyyy-mm-ddThh:mm:ssZ",
   "region": "us-east-1",
   "resources": [
      "arn:aws:ec2:us-east-1:012345678901:volume/vol-03a55cf56513fa1b6"
   ],
   "detail": {
      "result": "optimizing",
      "cause": "",
      "event": "modifyVolume",
      "request-id": "01234567-0123-0123-0123-0123456789ab"
   }
}
```

Événements de restauration d'instantané rapide EBS

Amazon EBS envoie des événements à CloudWatch Events en cas de variation de l'état de la fonction de restauration d'instantané rapide pour un instantané.

Voici un exemple de données pour cet événement.

```
{
   "version": "0",
   "id": "01234567-0123-0123-0123-012345678901",
   "detail-type": "EBS Fast Snapshot Restore State-change Notification",
   "source": "aws.ec2",
   "account": "123456789012",
   "time": "yyyy-mm-ddThh:mm:ssZ",
   "region": "us-east-1",
   "resources": [
      "arn:aws:ec2:us-east-1::snapshot/snap-03a55cf56513fa1b6"
   ],
   "detail": {
      "snapshot-id": "snap-1234567890abcdef0",
      "state": "optimizing",
      "zone": "us-east-1a",
      "message": "Client.UserInitiated - Lifecycle state transition",
   }
}
```

Les valeurs possibles pour `state` sont `enabling`, `optimizing`, `enabled`, `disabling` et `disabled`.

Les valeurs possibles pour `message` sont les suivantes :

`Client.InvalidSnapshot.InvalidState - The requested snapshot transitioned to an invalid state (Error)`

Une demande d'activation de la fonction de restauration d'instantané rapide a échoué et l'état est passé à `disabling` ou `disabled`. La fonction de restauration d'instantané rapide ne peut pas être activée pour cet instantané.

`Client.UserInitiated`

L'état est passé avec succès à `enabling` ou `disabling`.

`Client.UserInitiated - Lifecycle state transition`

L'état est passé avec succès à `optimizing`, `enabled` ou `disabled`.

`Server.InsufficientCapacity - There was insufficient capacity available to satisfy the request`

Une demande d'activation de la fonction de restauration d'instantané rapide a échoué en raison d'une capacité insuffisante et l'état est passé à `disabling` ou `disabled`. Attendez, puis recommencez.

`Server.InternalError - An internal error caused the operation to fail`

Une demande d'activation de la fonction de restauration d'instantané rapide a échoué en raison d'une erreur interne et l'état est passé à `disabling` ou `disabled`. Attendez, puis recommencez.

`Client.InvalidSnapshot.InvalidState - The requested snapshot was deleted or access permissions were revoked`

L'état de restauration d'instantané rapide est passé à `disabling` ou `disabled` parce que l'instantané a été supprimé ou non partagé par son propriétaire. La restauration d'instantané rapide ne peut pas être activée pour un instantané qui a été supprimé ou qui n'est plus partagé avec vous.

Utilisation de AWS Lambda pour gérer les événements CloudWatch

Vous pouvez utiliser Amazon EBS et CloudWatch Events pour automatiser votre flux de travail de sauvegarde des données. Cela vous permet de créer une stratégie IAM, une fonction AWS Lambda pour gérer l'événement, et une règle Amazon CloudWatch Events qui fait correspondre les événements entrants et les achemine vers la fonction Lambda.

La procédure suivante utilise l'événement `createSnapshot` pour copier automatiquement un instantané terminé vers une autre région pour la reprise après sinistre.

Pour copier un instantané terminé vers une autre région

1. Créez une stratégie IAM, telle que celle indiquée dans l'exemple suivant, pour fournir des autorisations d'exécution d'une action `CopySnapshot` et d'écriture dans le journal CloudWatch Events. Attribuez la stratégie à l'utilisateur IAM qui gérera l'événement CloudWatch.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "logs:CreateLogGroup",
        "logs:CreateLogStream",
        "logs:PutLogEvents"
      ],
      "Resource": "arn:aws:logs:*:*:*"
    },
    {
```

```
      "Effect": "Allow",
      "Action": [
        "ec2:CopySnapshot"
      ],
      "Resource": "*"
    }
  ]
}
```

2. Définissez une fonction dans Lambda qui sera disponible à partir de la console CloudWatch. L'exemple de fonction Lambda ci-dessous, écrite dans Node.js, est appelée par CloudWatch lorsqu'un événement `createSnapshot` correspondant est émis par Amazon EBS (ce qui signifie qu'un instantané est terminé). Lorsqu'elle est appelée, la fonction copie l'instantané de `us-east-2` vers `us-east-1`.

```
// Sample Lambda function to copy an EBS snapshot to a different region

var AWS = require('aws-sdk');
var ec2 = new AWS.EC2();

// define variables
var destinationRegion = 'us-east-1';
var sourceRegion = 'us-east-2';
console.log ('Loading function');

//main function
exports.handler = (event, context, callback) => {

    // Get the EBS snapshot ID from the CloudWatch event details
    var snapshotArn = event.detail.snapshot_id.split('/');
    const snapshotId = snapshotArn[1];
    const description = `Snapshot copy from ${snapshotId} in ${sourceRegion}.`;
    console.log ("snapshotId:", snapshotId);

    // Load EC2 class and update the configuration to use destination Region to
 initiate the snapshot.
    AWS.config.update({region: destinationRegion});
    var ec2 = new AWS.EC2();

    // Prepare variables for ec2.modifySnapshotAttribute call
    const copySnapshotParams = {
        Description: description,
        DestinationRegion: destinationRegion,
        SourceRegion: sourceRegion,
        SourceSnapshotId: snapshotId
    };

    // Execute the copy snapshot and log any errors
    ec2.copySnapshot(copySnapshotParams, (err, data) => {
        if (err) {
            const errorMessage = `Error copying snapshot ${snapshotId} to Region
 ${destinationRegion}.`;
            console.log(errorMessage);
            console.log(err);
            callback(errorMessage);
        } else {
            const successMessage = `Successfully started copy of snapshot ${snapshotId}
 to Region ${destinationRegion}.`;
            console.log(successMessage);
            console.log(data);
            callback(null, successMessage);
        }
    });
};
```

Pour vous assurer que votre fonction Lambda est disponible à partir de la console CloudWatch, créez-la dans la région où l'événement CloudWatch doit se produire. Pour plus d'informations, consultez le Manuel du développeur AWS Lambda.

3. Ouvrez la console CloudWatch à l'adresse https://console.aws.amazon.com/cloudwatch/.
4. Choisissez Events, Create rule, Select event source et Amazon EBS Snapshots.
5. Dans le champ Specific Event(s), choisissez createSnapshot et dans le champ Specific Result(s), choisissez succeeded.
6. Dans le champ Rule target, trouvez et choisissez l'exemple de fonction que vous avez créé précédemment.
7. Choisissez Target, puis Add Target.
8. Dans le champ Lambda function, sélectionnez la fonction Lambda que vous avez créée précédemment, puis choisissez Configure details.
9. Sur la page Configure rule details, tapez les valeurs de Name et Description. Sélectionnez la case à cocher State pour activer la fonction (en la définissant sur Enabled).
10. Choisissez Create rule.

Votre règle doit désormais apparaître sur l'onglet Rules. Dans l'exemple présenté, l'événement que vous avez configuré doit être émis par EBS la prochaine fois que vous copiez un instantané.

Quotas Amazon EBS

Pour afficher les quotas pour vos ressources Amazon EBS, ouvrez la console Quotas de service à l'adresse https://console.aws.amazon.com/servicequotas/. Dans le volet de navigation, choisissez AWS services (Services AWS) et sélectionnez Amazon Elastic Block Store (Amazon EBS).

Pour obtenir la liste des quotas de service Amazon EBS, veuillez consulter Points de terminaison et quotas Amazon Elastic Block Store dans le AWS General Reference.

Stockage d'instance Amazon EC2

Un stockage d'instance fournit un stockage temporaire de niveau bloc pour votre instance. Le stockage réside sur les disques physiquement attachés à l'ordinateur hôte. Le stockage d'instance est particulièrement adapté pour le stockage temporaire d'informations qui changent fréquemment, telles que les tampons, les caches, les données temporaires et autres contenus provisoires, ou pour les données répliquées sur une flotte d'instances, telle qu'un pool à charge équilibrée de serveurs web.

Un stockage d'instance se compose d'un ou de plusieurs stockages d'instance exposés comme périphériques de stockage en mode bloc. La taille d'un stockage d'instance ainsi que le nombre de périphériques disponibles varient en fonction du type d'instance.

Les périphériques virtuels des volumes de stockage d'instance sont `ephemeral[0-23]`. Les types d'instance qui prennent en charge un seul volume de stockage d'instance ont `ephemeral0`. Les types d'instance qui prennent en charge deux volumes de stockage d'instance ont `ephemeral0` et `ephemeral1`, et ainsi de suite.

Sommaire

- Durée de vie d'un stockage d'instance (p. 1265)
- Volumes de stockage d'instance (p. 1266)
- Ajouter des volumes de stockage d'instance à votre instance EC2 (p. 1272)
- Volumes de stockage d'instance SSD (p. 1276)
- Volumes d'échange de stockage d'instance (p. 1278)
- Optimisation des performances disque des volumes de stockage d'instance (p. 1280)

Durée de vie d'un stockage d'instance

Vous ne pouvez spécifier les volumes de stockage d'une instance que lors de son lancement. Vous ne pouvez pas détacher un volume de stockage d'instance à partir d'une instance et l'attacher à une autre instance.

Les données d'un stockage d'instance ne persistent que pendant la durée de vie de son instance associée. Si une instance redémarre (intentionnellement ou accidentellement), les données du stockage d'instance persistent. Cependant, les données du stockage d'instance sont perdues dans les cas suivants :

- Défaillance du disque dur sous-jacent
- Arrêt de l'instance
- Mise en veille prolongée de l'instance
- Terminaison de l'instance

Par conséquent, ne vous fiez pas au stockage d'instance pour les données précieuses et à long terme. Utilisez plutôt un stockage de données plus durable comme Amazon S3, Amazon EBS ou Amazon EFS.

Lorsque vous arrêtez, mettez en veille prolongée ou résiliez une instance, chaque bloc de stockage du stockage d'instance est réinitialisé. Par conséquent, vos données ne sont pas accessibles via le stockage d'instance d'une autre instance.

Si vous créez une AMI à partir d'une instance, les données de ses volumes de stockage d'instance ne sont pas conservées et ne sont pas présentes sur les volumes de stockage d'instance que vous lancez depuis l'AMI.

Si vous modifiez le type d'instance, un magasin d'instances ne sera pas attaché au nouveau type d'instance. Pour plus d'informations, consultez Modification du type d'instance (p. 313).

Volumes de stockage d'instance

Le type d'instance détermine la taille du stockage d'instance disponible, ainsi que le type de matériel utilisé pour les volumes de stockage d'instance. Les volumes de stockage d'instance sont inclus dans le coût d'utilisation de l'instance. Vous devez spécifier les volumes de stockage d'instance que vous souhaitez utiliser lorsque vous lancez l'instance (à l'exception des volumes de stockage d'instance NVMe, qui sont disponibles par défaut). Formatez et montez ensuite les volumes de stockage d'instance avant de les utiliser. Vous ne pouvez pas rendre disponible un volume de stockage d'instance après l'avoir lancé. Pour plus d'informations, consultez Ajouter des volumes de stockage d'instance à votre instance EC2 (p. 1272).

Certains types d'instance utilisent les disques SSD (Solid State Drive) NVMe ou SATA pour fournir des performances d'E/S aléatoires élevées. C'est une bonne option lorsque vous avez besoin d'un stockage ayant une latence très basse, mais qu'il n'est pas nécessaire que les données persistent quand l'instance est mise hors service ou que vous pouvez tirer parti des architectures tolérantes aux pannes. Pour plus d'informations, consultez Volumes de stockage d'instance SSD (p. 1276).

Le tableau suivant indique la quantité, la taille, le type et les optimisations des performances des volumes de stockage d'instance disponibles pour chaque type d'instance pris en charge. Pour obtenir la liste complète des types d'instance, y compris les types d'EBS uniquement, consultez Types d'instances Amazon EC2.

Type d'instance	Volumes de stockage d'instance	Type	Initialisation requise*	Commande Trim**
`c1.medium`	1 x 350 Go†	HDD	✓	
`c1.xlarge`	4 x 420 Go (1,6 To)	HDD	✓	
`c3.large`	2 x 16 Go (32 Go)	SSD	✓	
`c3.xlarge`	2 x 40 Go (80 Go)	SSD	✓	
`c3.2xlarge`	2 x 80 Go (160 Go)	SSD	✓	
`c3.4xlarge`	2 x 160 Go (320 Go)	SSD	✓	
`c3.8xlarge`	2 x 320 Go (640 Go)	SSD	✓	
`c5ad.large`	1 x 75 Go	SSD NVMe		✓
`c5ad.xlarge`	1 x 150 Go	SSD NVMe		✓
`c5ad.2xlarge`	1 x 300 Go	SSD NVMe		✓
`c5ad.4xlarge`	2 x 300 Go (600 Go)	SSD NVMe		✓
`c5ad.8xlarge`	2 x 600 Go (1,2 To)	SSD NVMe		✓
`c5ad.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓

Type d'instance	Volumes de stockage d'instance	Type	Initialisation requise*	Commande Trim**
`c5ad.16xlarge`	2 x 1,2 To (2,4 To)	SSD NVMe		✓
`c5ad.24xlarge`	2 x 1,9 To (3,8 To)	SSD NVMe		✓
`c5d.large`	1 x 50 Go	SSD NVMe		✓
`c5d.xlarge`	1 x 100 Go	SSD NVMe		✓
`c5d.2xlarge`	1 x 200 Go	SSD NVMe		✓
`c5d.4xlarge`	1 x 400 Go	SSD NVMe		✓
`c5d.9xlarge`	1 x 900 Go	SSD NVMe		✓
`c5d.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`c5d.18xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`c5d.24xlarge`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`c5d.metal`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`c6gd.medium`	1 x 59 Go	SSD NVMe		✓
`c6gd.large`	1 x 118 Go	SSD NVMe		✓
`c6gd.xlarge`	1 x 237 Go	SSD NVMe		✓
`c6gd.2xlarge`	1 x 474 Go	SSD NVMe		✓
`c6gd.4xlarge`	1 x 950 Go	SSD NVMe		✓
`c6gd.8xlarge`	1 x 1 900 Go	SSD NVMe		✓
`c6gd.12xlarge`	2 x 1 425 Go (2,85 To)	SSD NVMe		✓
`c6gd.16xlarge`	2 x 1 900 Go (3,8 To)	SSD NVMe		✓
`c6gd.metal`	2 x 1 900 Go (3,8 To)	SSD NVMe		✓
`cc2.8xlarge`	4 x 840 Go (3,36 To)	HDD	✓	
`cr1.8xlarge`	2 x 120 Go (240 Go)	SSD	✓	
`d2.xlarge`	3 x 2 000 Go (6 To)	HDD		
`d2.2xlarge`	6 x 2 000 Go (12 To)	HDD		
`d2.4xlarge`	12 x 2 000 Go (24 To)	HDD		
`d2.8xlarge`	24 x 2 000 Go (48 To)	HDD		
`f1.2xlarge`	1 x 470 Go	SSD NVMe		✓
`f1.4xlarge`	1 x 940 Go	SSD NVMe		✓
`f1.16xlarge`	4 x 940 Go (3,76 To)	SSD NVMe		✓
`g2.2xlarge`	1 x 60 Go	SSD	✓	

Type d'instance	Volumes de stockage d'instance	Type	Initialisation requise*	Commande Trim**
`g2.8xlarge`	2 x 120 Go (240 Go)	SSD	✓	
`g4dn.xlarge`	1 x 125 Go	SSD NVMe		✓
`g4dn.2xlarge`	1 x 225 Go	SSD NVMe		✓
`g4dn.4xlarge`	1 x 225 Go	SSD NVMe		✓
`g4dn.8xlarge`	1 x 900 Go	SSD NVMe		✓
`g4dn.12xlarge`	1 x 900 Go	SSD NVMe		✓
`g4dn.16xlarge`	1 x 900 Go	SSD NVMe		✓
`g4dn.metal`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`h1.2xlarge`	1 x 2 000 Go (2 To)	HDD		
`h1.4xlarge`	2 x 2 000 Go (4 To)	HDD		
`h1.8xlarge`	4 x 2 000 Go (8 To)	HDD		
`h1.16xlarge`	8 x 2 000 Go (16 To)	HDD		
`hs1.8xlarge`	24 x 2 000 Go (48 To)	HDD	✓	
`i2.xlarge`	1 x 800 Go	SSD		✓
`i2.2xlarge`	2 x 800 Go (1,6 To)	SSD		✓
`i2.4xlarge`	4 x 800 Go (3,2 To)	SSD		✓
`i2.8xlarge`	8 x 800 Go (6,4 To)	SSD		✓
`i3.large`	1 x 475 Go	SSD NVMe		✓
`i3.xlarge`	1 x 950 Go	SSD NVMe		✓
`i3.2xlarge`	1 x 1,900 Go	SSD NVMe		✓
`i3.4xlarge`	2 x 1 900 Go (3,8 To)	SSD NVMe		✓
`i3.8xlarge`	4 x 1 900 Go (7,6 To)	SSD NVMe		✓
`i3.16xlarge`	8 x 1 900 Go (15,2 To)	SSD NVMe		✓
`i3.metal`	8 x 1 900 Go (15,2 To)	SSD NVMe		✓
`i3en.large`	1 x 1,250 Go	SSD NVMe		✓
`i3en.xlarge`	1 x 2,500 Go	SSD NVMe		✓
`i3en.2xlarge`	2 x 2,500 Go (5 To)	SSD NVMe		✓
`i3en.3xlarge`	1 x 7,500 Go	SSD NVMe		✓
`i3en.6xlarge`	2 x 7,500 Go (15 To)	SSD NVMe		✓
`i3en.12xlarge`	4 x 7,500 Go (30 To)	SSD NVMe		✓

Type d'instance	Volumes de stockage d'instance	Type	Initialisation requise*	Commande Trim**
`i3en.24xlarge`	8 x 7 500 Go (60 To)	SSD NVMe		✓
`i3en.metal`	8 x 7 500 Go (60 To)	SSD NVMe		✓
`m1.small`	1 x 160 Go†	HDD	✓	
`m1.medium`	1 x 410 Go	HDD	✓	
`m1.large`	2 x 420 Go (840 Go)	HDD	✓	
`m1.xlarge`	4 x 420 Go (1,6 To)	HDD	✓	
`m2.xlarge`	1 x 420 Go	HDD	✓	
`m2.2xlarge`	1 x 850 Go	HDD	✓	
`m2.4xlarge`	2 x 840 Go (1,68 To)	HDD	✓	
`m3.medium`	1 x 4 Go	SSD	✓	
`m3.large`	1 x 32 Go	SSD	✓	
`m3.xlarge`	2 x 40 Go (80 Go)	SSD	✓	
`m3.2xlarge`	2 x 80 Go (160 Go)	SSD	✓	
`m5ad.large`	1 x 75 Go	SSD NVMe		✓
`m5ad.xlarge`	1 x 150 Go	SSD NVMe		✓
`m5ad.2xlarge`	1 x 300 Go	SSD NVMe		✓
`m5ad.4xlarge`	2 x 300 Go (600 Go)	SSD NVMe		✓
`m5ad.8xlarge`	2 x 600 Go (1,2 To)	SSD NVMe		✓
`m5ad.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`m5ad.16xlarge`	4 x 600 Go (2,4 To)	SSD NVMe		✓
`m5ad.24xlarge`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`m5d.large`	1 x 75 Go	SSD NVMe		✓
`m5d.xlarge`	1 x 150 Go	SSD NVMe		✓
`m5d.2xlarge`	1 x 300 Go	SSD NVMe		✓
`m5d.4xlarge`	2 x 300 Go (600 Go)	SSD NVMe		✓
`m5d.8xlarge`	2 x 600 Go (1,2 To)	SSD NVMe		✓
`m5d.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`m5d.16xlarge`	4 x 600 Go (2,4 To)	SSD NVMe		✓
`m5d.24xlarge`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`m5d.metal`	4 x 900 Go (3,6 To)	SSD NVMe		✓

Type d'instance	Volumes de stockage d'instance	Type	Initialisation requise*	Commande Trim**
`m5dn.large`	1 x 75 Go	SSD NVMe		✓
`m5dn.xlarge`	1 x 150 Go	SSD NVMe		✓
`m5dn.2xlarge`	1 x 300 Go	SSD NVMe		✓
`m5dn.4xlarge`	2 x 300 Go (600 Go)	SSD NVMe		✓
`m5dn.8xlarge`	2 x 600 Go (1,2 To)	SSD NVMe		✓
`m5dn.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`m5dn.16xlarge`	4 x 600 Go (2,4 To)	SSD NVMe		✓
`m5dn.24xlarge`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`m6gd.medium`	1 x 59 Go	SSD NVMe		✓
`m6gd.large`	1 x 118 Go	SSD NVMe		✓
`m6gd.xlarge`	1 x 237 Go	SSD NVMe		✓
`m6gd.2xlarge`	1 x 474 Go	SSD NVMe		✓
`m6gd.4xlarge`	1 x 950 Go	SSD NVMe		✓
`m6gd.8xlarge`	1 x 1 900 Go	SSD NVMe		✓
`m6gd.12xlarge`	2 x 1 425 Go (2,85 To)	SSD NVMe		✓
`m6gd.16xlarge`	2 x 1 900 Go (3,8 To)	SSD NVMe		✓
`m6gd.metal`	2 x 1 900 Go (3,8 To)	SSD NVMe		✓
`p3dn.24xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`r3.large`	1 x 32 Go	SSD		✓
`r3.xlarge`	1 x 80 Go	SSD		✓
`r3.2xlarge`	1 x 160 Go	SSD		✓
`r3.4xlarge`	1 x 320 Go	SSD		✓
`r3.8xlarge`	2 x 320 Go (640 Go)	SSD		✓
`r5ad.large`	1 x 75 Go	SSD NVMe		✓
`r5ad.xlarge`	1 x 150 Go	SSD NVMe		✓
`r5ad.2xlarge`	1 x 300 Go	SSD NVMe		✓
`r5ad.4xlarge`	2 x 300 Go (600 Go)	SSD NVMe		✓
`r5ad.8xlarge`	2 x 600 Go (1,2 To)	SSD NVMe		✓
`r5ad.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`r5ad.16xlarge`	4 x 600 Go (2,4 To)	SSD NVMe		✓

Type d'instance	Volumes de stockage d'instance	Type	Initialisation requise*	Commande Trim**
`r5ad.24xlarge`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`r5d.large`	1 x 75 Go	SSD NVMe		✓
`r5d.xlarge`	1 x 150 Go	SSD NVMe		✓
`r5d.2xlarge`	1 x 300 Go	SSD NVMe		✓
`r5d.4xlarge`	2 x 300 Go (600 Go)	SSD NVMe		✓
`r5d.8xlarge`	2 x 600 Go (1,2 To)	SSD NVMe		✓
`r5d.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`r5d.16xlarge`	4 x 600 Go (2,4 To)	SSD NVMe		✓
`r5d.24xlarge`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`r5d.metal`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`r5dn.large`	1 x 75 Go	SSD NVMe		✓
`r5dn.xlarge`	1 x 150 Go	SSD NVMe		✓
`r5dn.2xlarge`	1 x 300 Go	SSD NVMe		✓
`r5dn.4xlarge`	2 x 300 Go (600 Go)	SSD NVMe		✓
`r5dn.8xlarge`	2 x 600 Go (1,2 To)	SSD NVMe		✓
`r5dn.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`r5dn.16xlarge`	4 x 600 Go (2,4 To)	SSD NVMe		✓
`r5dn.24xlarge`	4 x 900 Go (3,6 To)	SSD NVMe		✓
`r6gd.medium`	1 x 59 Go	SSD NVMe		✓
`r6gd.large`	1 x 118 Go	SSD NVMe		✓
`r6gd.xlarge`	1 x 237 Go	SSD NVMe		✓
`r6gd.2xlarge`	1 x 474 Go	SSD NVMe		✓
`r6gd.4xlarge`	1 x 950 Go	SSD NVMe		✓
`r6gd.8xlarge`	1 x 1 900 Go	SSD NVMe		✓
`r6gd.12xlarge`	2 x 1 425 Go (2,85 To)	SSD NVMe		✓
`r6gd.16xlarge`	2 x 1 900 Go (3,8 To)	SSD NVMe		✓
`r6gd.metal`	2 x 1 900 Go (3,8 To)	SSD NVMe		✓
`x1.16xlarge`	1 x 1,920 Go	SSD		
`x1.32xlarge`	2 x 1 920 Go (3,84 To)	SSD		
`x1e.xlarge`	1 x 120 Go	SSD		

Type d'instance	Volumes de stockage d'instance	Type	Initialisation requise*	Commande Trim**
`x1e.2xlarge`	1 x 240 Go	SSD		
`x1e.4xlarge`	1 x 480 Go	SSD		
`x1e.8xlarge`	1 x 960 Go	SSD		
`x1e.16xlarge`	1 x 1,920 Go	SSD		
`x1e.32xlarge`	2 x 1 920 Go (3,84 To)	SSD		
`z1d.large`	1 x 75 Go	SSD NVMe		✓
`z1d.xlarge`	1 x 150 Go	SSD NVMe		✓
`z1d.2xlarge`	1 x 300 Go	SSD NVMe		✓
`z1d.3xlarge`	1 x 450 Go	SSD NVMe		✓
`z1d.6xlarge`	1 x 900 Go	SSD NVMe		✓
`z1d.12xlarge`	2 x 900 Go (1,8 To)	SSD NVMe		✓
`z1d.metal`	2 x 900 Go (1,8 To)	SSD NVMe		✓

* Les volumes attachés à certaines instances subissent des pertes de performance lors de la première écriture s'ils ne sont pas initialisés. Pour plus d'informations, consultez Optimisation des performances disque des volumes de stockage d'instance (p. 1280).

** Pour plus d'informations, consultez Prise en charge de TRIM sur les volumes de stockage d'instance (p. 1278).

† Les types d'instance `c1.medium` et `m1.small` incluent aussi un volume d'échange de stockage d'instance de 900 Mo, qui ne peut pas être activé automatiquement au moment du démarrage. Pour plus d'informations, consultez Volumes d'échange de stockage d'instance (p. 1278).

Ajouter des volumes de stockage d'instance à votre instance EC2

Vous spécifiez les volumes EBS et les volumes de stockage d'instance de votre instance à l'aide d'un mappage de périphérique de stockage en mode bloc. Chaque entrée d'un mappage de périphérique de stockage en mode bloc inclut un nom de périphérique et le volume sur lequel il est mappé. Le mappage de périphérique de stockage en mode bloc par défaut est spécifiée par l'AMI que vous utilisez. Vous pouvez également spécifier un mappage de périphérique de stockage en mode bloc pour l'instance lors de son lancement.

Tous les volumes de stockage d'instance NVMe pris en charge par un type d'instance sont automatiquement énumérés et un nom de périphérique leur est automatiquement attribué au lancement de l'instance. Le fait de les ajouter dans le mappage de périphérique de stockage en mode bloc pour l'AMI ou l'instance n'a aucun effet. Pour plus d'informations, consultez Mappage de périphérique de stockage en mode bloc (p. 1289).

Un mappage de périphérique de stockage en mode bloc spécifie toujours le volume racine de l'instance. Le volume racine est un volume Amazon EBS ou un volume de stockage d'instance. Pour plus d'informations, consultez Stockage pour le périphérique racine (p. 106). Le volume racine est monté automatiquement.

Dans le cas des instances ayant un volume de stockage d'instance pour le volume racine, la taille de ce volume varie en fonction de l'AMI, mais la taille maximale est de 10 Go.

Vous pouvez utiliser un mappage de périphérique de stockage en mode bloc pour spécifier des volumes EBS supplémentaires quand vous lancez votre instance ou en attacher une fois que votre instance est en cours d'exécution. Pour plus d'informations, consultez Volumes Amazon EBS (p. 1086).

Vous ne pouvez spécifier les volumes de stockage d'instance de votre instance que lors de son lancement. Vous ne pouvez pas attacher des volumes de stockage d'instance à une instance après l'avoir lancée.

Si vous modifiez le type d'instance, un magasin d'instances ne sera pas attaché au nouveau type d'instance. Pour plus d'informations, consultez Modification du type d'instance (p. 313).

Le nombre et la taille de volumes de stockage d'instance disponibles pour votre instance varient par type d'instance. Certains types d'instance ne prennent pas en charge les volumes de stockage d'instance. Si le nombre de volumes de stockage d'instances dans un mappage d'appareils en bloc dépasse le nombre de volumes de stockage d'instances disponibles pour une instance, les volumes supplémentaires sont ignorés. Pour de plus amples informations sur la prise en charge des volumes de stockage d'instance par chaque type d'instance, veuillez consulter Volumes de stockage d'instance (p. 1266).

Si le type d'instance que vous avez choisie pour votre instance prend en charge les volumes de stockage d'instance non NVMe, vous devez les ajouter au mappage de périphérique de stockage en mode bloc de l'instance lorsque vous la lancez. Les volumes de stockage d'instances NVMe sont disponibles par défaut. Après le lancement d'une instance, vous devez vous assurer que les volumes de stockage d'instance de votre instance sont formatés et montés avant que vous ne puissiez les utiliser. Le volume racine d'une instance basée sur le stockage d'instance est monté automatiquement.

Sommaire

- Ajout de volumes de stockage d'instance à une AMI (p. 1273)
- Ajout de volumes de stockage d'instance à une instance (p. 1274)
- Rendre disponibles les volumes de stockage d'instance sur votre instance (p. 1275)

Ajout de volumes de stockage d'instance à une AMI

Vous pouvez créer une AMI avec un mappage de périphérique de stockage en mode bloc incluant des volumes de stockage d'instance. Si vous lancez une instance avec un type qui prend en charge les volumes de stockage d'instances et une AMI qui spécifie les volumes de stockage d'instances dans son mappage d'appareils en bloc, l'instance inclut ces volumes de stockage d'instances. Si le nombre de volumes de stockage d'instances dans le mappage d'appareils en bloc dépasse le nombre de volumes de stockage d'instances disponibles pour l'instance, les volumes supplémentaires sont ignorés.

Considérations

- Pour les instances M3, spécifiez les volumes de stockage d'instance dans le mappage de périphérique de stockage en mode bloc de l'instance, pas l'AMI. Amazon EC2 pourrait ignorer les volumes de stockage d'instance qui sont spécifiés uniquement dans le mappage de périphérique de stockage en mode bloc de l'AMI.
- Lors du lancement d'une instance, vous pouvez omettre les volumes de stockage d'instances non NVMe spécifiés dans le mappage d'appareils de stockage en mode bloc de l'AMI et ajouter de nouveaux volumes de stockage d'instances.

Pour ajouter des volumes de stockage d'instance à une AMI basée sur des volumes Amazon EBS à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.

2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez l'instance.
3. Sélectionnez Actions, Image, Créer une image.
4. Dans la boîte de dialogue Créer une image, attribuez à votre image un nom et une description significatifs.
5. Pour chaque volume de stockage d'instance à ajouter, choisissez Ajouter un nouveau volume, et sélectionnez un volume de stockage d'instance dans Type de volume puis un nom de périphérique dans Dispositif. (Pour plus d'informations, consultez Noms d'appareil pour les instances Linux (p. 1287).) Le nombre de volumes de stockage d'instance disponibles dépend du type d'instance. Pour les instances avec volumes de stockage d'instance NVMe, le mappage de périphérique de ces volumes dépend de l'ordre dans lequel le système d'exploitation énumère les volumes.
6. Choisissez Créer une image.

Pour ajouter des volumes de stockage d'instance à une AMI à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- create-image ou register-image (AWS CLI)
- New-EC2Image et Register-EC2Image (Outils AWS pour Windows PowerShell)

Ajout de volumes de stockage d'instance à une instance

Lors du lancement d'une instance, le mappage de périphérique de stockage en mode bloc par défaut est fourni par l'AMI spécifié. Si vous avez besoin de volumes de stockage d'instance supplémentaires, vous devez les ajouter à l'instance lors de son lancement. Vous pouvez aussi omettre les périphériques spécifiés dans le mappage de périphérique de stockage en mode bloc de l'AMI.

Considérations

- Pour les instances M3, il se peut que vous receviez les volumes de stockage d'instance, même si vous ne les spécifiez pas dans le mappage de périphérique de stockage en mode bloc de l'instance.
- Pour les instances HS1, quel que soit le nombre de volumes de stockage d'instance que vous spécifiez dans le mappage de périphérique de stockage en mode bloc d'une AMI, le mappage de périphérique de stockage en mode bloc d'une instance lancée depuis l'AMI inclut automatiquement le nombre maximal de volumes de stockage d'instance pris en charge. Vous devez supprimer les volumes de stockage d'instance dont vous ne voulez pas du mappage de périphérique de stockage en mode bloc de l'instance avant le lancement de celle-ci.

Pour mettre à jour le mappage de périphérique de stockage en mode bloc d'une instance à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le tableau de bord, choisissez Lancer une instance.
3. Dans Étape 1 : Sélection d'une Amazon Machine Image (AMI), sélectionnez l'AMI à utiliser et choisissez Sélectionner.
4. Suivez l'Assistant pour exécuter les étapes Étape 1 : Sélection d'une Amazon Machine Image (AMI), Étape 2 : Choisir un type d'instance et Étape 3 : Configurer les détails de l'instance.
5. Dans Étape 4 : Ajouter le stockage, modifiez les entrées existantes selon vos besoins. Pour chaque volume de stockage d'instance à ajouter, choisissez Ajouter un nouveau volume, et sélectionnez un volume de stockage d'instance dans Type de volume puis un nom de périphérique dans Dispositif. Le nombre de volumes de stockage d'instance disponibles dépend du type d'instance.

6. Exécutez l'assistant et lancez l'instance.
7. (Facultatif) Pour afficher les volumes de stockage d'instances disponibles sur votre instance, exécutez la commande lsblk.

Pour mettre à jour le mappage de périphérique de stockage en mode bloc d'une instance à l'aide de la ligne de commande

Vous pouvez utiliser l'une des commandes d'option suivantes avec la commande correspondante. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- `--block-device-mappings` avec run-instances (AWS CLI)
- `-BlockDeviceMapping` avec New-EC2Instance (Outils AWS pour Windows PowerShell)

Rendre disponibles les volumes de stockage d'instance sur votre instance

Après le lancement, les volumes de stockage d'instance sont accessibles à l'instance, mais vous ne pouvez pas y accéder tant qu'ils ne sont pas montés. Pour les instances Linux, le type d'instance détermine quels sont les volumes de stockage d'instance qui sont montés automatiquement et lesquels sont disponibles pour que vous les montiez vous-même. Pour les instances Windows, le service EC2Config monte les volumes de stockage d'instance pour une instance. Le pilote du périphérique de stockage en mode bloc de l'instance attribue le nom réel du volume au montage de celui-ci et le nom affecté peut être différent de celui recommandé par Amazon EC2.

La plupart des volumes de stockage d'instance sont préformatés avec le système de fichiers ext3. Les volumes de stockage d'instance SSD prenant en charge l'instruction TRIM ne sont pas préformatés avec un système de fichiers. Cependant, vous pouvez formater les volumes avec le système de fichiers de votre choix après avoir lancé votre instance. Pour plus d'informations, consultez Prise en charge de TRIM sur les volumes de stockage d'instance (p. 1278). Pour les instances Windows, le service EC2Config reformate les volumes de stockage d'instance avec le système de fichiers NTFS.

Vous pouvez confirmer que les périphériques de stockage d'instance sont disponibles depuis l'instance elle-même à l'aide des métadonnées d'instance. Pour plus d'informations, consultez Affichage du mappage de périphérique de stockage en mode bloc d'une instance pour les volumes de stockage d'instances (p. 1299).

Pour les instances Windows, vous pouvez aussi afficher les volumes de stockage d'instance avec la gestion de disques Windows. Pour de plus amples informations, veuillez consulter Liste des disques utilisant le gestionnaire de disques Windows.

Pour les instances Linux, vous pouvez afficher et monter les volumes de stockage d'instance, comme décrit dans la procédure suivante.

Pour rendre disponible un volume de stockage d'instance sur Linux

1. Connectez-vous à l'instance à l'aide d'un client SSH. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Utilisez la commande `df -h` pour afficher les volumes qui sont formatés et montés.

```
[ec2-user ~]$ df -h
Filesystem      Size  Used Avail Use% Mounted on
devtmpfs        3.8G   72K  3.8G   1% /dev
tmpfs           3.8G     0  3.8G   0% /dev/shm
```

```
/dev/nvme0n1p1  7.9G  1.2G  6.6G  15% /
```

3. Utilisez la commande `lsblk` pour afficher les volumes qui ont été mappés au lancement, mais ne sont ni formatés ni montés.

```
[ec2-user ~]$ lsblk
NAME          MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
nvme0n1       259:1    0    8G  0 disk
##nvme0n1p1   259:2    0    8G  0 part /
##nvme0n1p128 259:3    0    1M  0 part
nvme1n1       259:0    0 69.9G  0 disk
```

4. Pour formater et monter un volume de stockage d'instance qui a seulement été mappé, procédez comme suit :

   a. Créez un système de fichiers sur le périphérique avec la commande `mkfs`.

   ```
   [ec2-user ~]$ sudo mkfs -t xfs /dev/nvme1n1
   ```

   b. Créez un répertoire sur lequel monter le périphérique avec la commande `mkdir`.

   ```
   [ec2-user ~]$ sudo mkdir /data
   ```

   c. Montez le périphérique sur le répertoire nouvellement créé à l'aide de la commande `mount`.

   ```
   [ec2-user ~]$ sudo mount /dev/nvme1n1 /data
   ```

Pour obtenir des instructions sur le montage automatique d'un volume attaché après le redémarrage, veuillez consulter Monter automatiquement un volume attaché après le redémarrage (p. 1115).

Volumes de stockage d'instance SSD

Pour garantir les meilleures performances d'IOPS à partir de vos volumes de stockage d'instance SSD sur Linux, nous vous recommandons d'utiliser la version la plus récente d'Amazon Linux ou d'une autre AMI Linux avec la version 3.8 du noyau ou version ultérieure. Si vous n'utilisez pas une AMI Linux dotée de la version 3.8 ou ultérieure du noyau, votre instance n'atteindra pas les performances IOPS maximales pour ces types d'instance.

Comme pour tout autre volume de stockage d'instance, vous devez mapper les volumes de stockage d'instance SSD de votre instance lorsque cette dernière est lancée. Les données d'un volume d'instance SSD ne persistent que pendant la vie de son instance associée. Pour plus d'informations, consultez Ajouter des volumes de stockage d'instance à votre instance EC2 (p. 1272).

Volumes SSD NVMe

Certaines instances offrent des volumes de stockage d'instance SSD NVMe (Non-Volatile Memory Express). Pour de plus amples informations sur le type de volume de stockage d'instance pris en charge par chaque type d'instance, veuillez consulter Volumes de stockage d'instance (p. 1266).

Pour accéder aux volumes NVMe, les pilotes NVMe (p. 1209) doivent être installés. Les AMI suivantes satisfont cette exigence :

- Amazon Linux 2
- Amazon Linux AMI 2018.03
- Ubuntu 14.04 (avec le noyau `linux-aws`) ou version ultérieure

- Red Hat Enterprise Linux 7.4 ou version ultérieure
- SUSE Linux Enterprise Server 12 SP2 ou version ultérieure
- CentOS 7.4.1708 ou version ultérieure
- FreeBSD 11.1 ou version ultérieure
- Debian GNU/Linux 9 ou version ultérieure

Après avoir connecté votre instance, vous pouvez répertorier les périphériques NVMe à l'aide de la commande lspci. Voici un exemple de sortie d'une instance `i3.8xlarge`, qui prend en charge quatre périphériques NVMe.

```
[ec2-user ~]$ lspci
00:00.0 Host bridge: Intel Corporation 440FX - 82441FX PMC [Natoma] (rev 02)
00:01.0 ISA bridge: Intel Corporation 82371SB PIIX3 ISA [Natoma/Triton II]
00:01.1 IDE interface: Intel Corporation 82371SB PIIX3 IDE [Natoma/Triton II]
00:01.3 Bridge: Intel Corporation 82371AB/EB/MB PIIX4 ACPI (rev 01)
00:02.0 VGA compatible controller: Cirrus Logic GD 5446
00:03.0 Ethernet controller: Device 1d0f:ec20
00:17.0 Non-Volatile memory controller: Device 1d0f:cd01
00:18.0 Non-Volatile memory controller: Device 1d0f:cd01
00:19.0 Non-Volatile memory controller: Device 1d0f:cd01
00:1a.0 Non-Volatile memory controller: Device 1d0f:cd01
00:1f.0 Unassigned class [ff80]: XenSource, Inc. Xen Platform Device (rev 01)
```

Si vous utilisez un système d'exploitation pris en charge, mais que les périphériques NVMe ne sont pas visibles, vérifiez que le module NVMe est chargé à l'aide de la commande suivante.

- Amazon Linux, Amazon Linux 2, Ubuntu 14/16, Red Hat Enterprise Linux, SUSE Linux Enterprise Server, CentOS 7

```
$ lsmod | grep nvme
nvme           48813  0
```

- Ubuntu 18

```
$ cat /lib/modules/$(uname -r)/modules.builtin | grep nvme
s/nvme/host/nvme-core.ko
kernel/drivers/nvme/host/nvme.ko
kernel/drivers/nvmem/nvmem_core.ko
```

Les volumes NVMe sont conformes à la spécification NVMe 1.0e. Vous pouvez utiliser les commandes NVMe avec vos volumes NVMe. Avec Amazon Linux, vous pouvez installer le package `nvme-cli` à partir du référentiel à l'aide de la commande yum install. Avec d'autres versions de Linux prises en charge, vous pouvez télécharger le package `nvme-cli` s'il n'est pas disponible dans l'image.

Les données sur le stockage d'instance NVMe sont chiffrées à l'aide d'un chiffrement par blocs XTS-AES-256 implémenté dans un module matériel sur l'instance. Les clés de chiffrement sont générées à l'aide du module matériel et sont uniques pour chaque périphérique de stockage d'instance NVMe. Toutes les clés de chiffrement sont détruites lorsque l'instance est arrêtée ou résiliée et ne peuvent pas être récupérées. Vous ne pouvez pas désactiver le chiffrement et vous ne pouvez pas fournir votre propre clé de chiffrement.

Volumes SSD non NVMe

Les instances suivantes prennent en charge les volumes de stockage d'instance qui utilisent les disques SSD (Solid State Drive) NVMe pour fournir des performances d'E/S aléatoires élevées : C, G2, 2, I3, M3,

R3 et X1. Pour de plus amples informations sur la prise en charge des volumes de stockage d'instance par chaque type d'instance, veuillez consulter Volumes de stockage d'instance (p. 1266).

Prise en charge de TRIM sur les volumes de stockage d'instance

Certains types d'instance prennent en charge les volumes SSD avec TRIM. Pour plus d'informations, consultez Volumes de stockage d'instance (p. 1266).

Les volumes de stockage d'instance qui prennent en charge TRIM sont intégralement soumis à l'instruction TRIM avant d'être alloués à votre instance. Comme ces volumes ne sont pas formatés avec un système de fichiers au lancement de l'instance, vous devez les formater avant qu'ils ne puissent être montés et utilisés. Pour un accès plus rapide à ces volumes, vous devez ignorer l'opération TRIM lorsque vous les formatez.

Avec les volumes de stockage d'instance qui prennent en charge TRIM, vous pouvez utiliser la commande TRIM pour informer le contrôleur SSD du moment où vous n'avez plus besoin des données que vous avez écrites. Cela fournit au contrôleur plus d'espace disponible, ce qui peut réduire l'amplification d'écriture et augmenter les performances. Sous Linux, utilisez la commande `fstrim` pour activer le TRIM périodique.

Volumes d'échange de stockage d'instance

L'espace d'échange de Linux peut être utilisé quand un système nécessite plus de mémoire que celle qui lui a été allouée physiquement. Quand l'espace d'échange est activé, les systèmes peuvent échanger exceptionnellement les pages mémoire utilisées entre la mémoire physique et l'espace d'échange (partition dédiée ou fichier d'échange d'un système de fichiers existant) et libérer cet espace pour les pages mémoire qui nécessitent un accès à haute vitesse.

> **Note**
>
> L'utilisation de l'espace d'échange pour la pagination mémoire n'est pas aussi rapide ou efficace que celle de la RAM. Si votre charge de travail pagine régulièrement la mémoire dans l'espace d'échange, envisagez de migrer vers un type d'instance plus grand avec plus de RAM. Pour plus d'informations, consultez Modification du type d'instance (p. 313).

Les types d'instance `c1.medium` et `m1.small` ont une quantité limitée de mémoire physique à utiliser et bénéficient d'un volume d'échange de 900 Mio au moment du lancement qui fait office de mémoire virtuelle pour les AMIs Linux. Même si le noyau Linux considère cet espace d'échange comme une partition du périphérique racine, il s'agit réellement d'un volume de stockage d'instance distinct, quel que soit votre type de périphérique racine.

Amazon Linux active et utilise automatiquement cet espace d'échange et l'utilisent, mais il se peut que votre AMI nécessite quelques étapes supplémentaires pour reconnaître et utiliser cet espace d'échange. Pour vérifier si votre instance utilise un espace d'échange, vous pouvez utiliser la commande swapon -s.

```
[ec2-user ~]$ swapon -s
Filename                                Type            Size    Used    Priority
/dev/xvda3                              partition       917500  0       -1
```

L'instance ci-dessus possède un volume d'échange de 900 Mio attaché et activé. Si vous ne voyez aucun volume d'échange apparaître avec cette commande, vous devez peut-être activer l'espace d'échange pour le périphérique. Vérifiez vos disques disponibles à l'aide de la commande lsblk.

```
[ec2-user ~]$ lsblk
NAME  MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
xvda1 202:1    0    8G  0 disk /
xvda3 202:3    0  896M  0 disk
```

Ici, le volume d'échange `xvda3` est accessible à l'instance, mais il n'est pas activé (notez que le champ `MOUNTPOINT` est vide). Vous pouvez activer le volume d'échange avec la commande swapon.

Note

Vous devez préfixer par `/dev/` le nom du périphérique affiché par lsblk. Votre périphérique peut avoir un nom différent, tel que `sda3`, `sde3` ou `xvde3`. Utilisez le nom de périphérique pour votre système dans la commande ci-après.

```
[ec2-user ~]$ sudo swapon /dev/xvda3
```

L'espace d'échange apparaît désormais dans la sortie lsblk et swapon -s.

```
[ec2-user ~]$ lsblk
NAME  MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
xvda1 202:1    0    8G  0 disk /
xvda3 202:3    0  896M  0 disk [SWAP]
[ec2-user ~]$ swapon -s
Filename                                Type            Size    Used    Priority
/dev/xvda3                              partition       917500  0       -1
```

Vous devez aussi modifier votre fichier `/etc/fstab` de telle sorte que cet espace d'échange soit automatiquement activé à chaque démarrage système.

```
[ec2-user ~]$ sudo vim /etc/fstab
```

Ajoutez la ligne suivante à votre fichier `/etc/fstab` (à l'aide du nom du périphérique d'échange de votre système) :

```
/dev/xvda3       none    swap    sw  0       0
```

Pour utiliser un volume de stockage d'instance comme espace d'échange

Tout volume de stockage d'instance peut être utilisé comme espace d'échange. Par exemple, le type d'instance `m3.medium` inclut un volume de stockage d'instance SSD de 4 Go, qui convient à l'espace d'échange. Si votre volume de stockage d'instance est beaucoup plus grand (par exemple, 350 Go), vous pouvez envisager de partitionner le volume avec une partition d'échange plus petite de 4 à 8 Go, le reste étant affecté à un volume de données.

Note

Cette procédure s'applique uniquement aux types d'instance prenant en charge ce stockage d'instance. Pour obtenir la liste des types d'instances, consultez Volumes de stockage d'instance (p. 1266).

1. Affichez les périphériques de stockage en mode bloc attachés à votre instance pour obtenir le nom de périphérique de votre volume de stockage d'instance.

   ```
   [ec2-user ~]$ lsblk -p
   NAME       MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
   /dev/xvdb  202:16   0   4G  0 disk /media/ephemeral0
   /dev/xvda1 202:1    0   8G  0 disk /
   ```

   Dans cet exemple, le volume de stockage d'instance est `/dev/xdvb`. Comme il s'agit d'une instance Amazon Linux, le volume de stockage d'instance est formaté et monté à `/media/ephemeral0` ; certains systèmes d'exploitation Linux n'opèrent pas ainsi automatiquement.
2. (Facultatif) Si votre volume de stockage d'instance est monté (il affiche un `MOUNTPOINT` dans la sortie de la commande lsblk), vous devez le démonter à l'aide de la commande suivante.

   ```
   [ec2-user ~]$ sudo umount /dev/xvdb
   ```

3. Configurez une zone d'échange Linux sur le périphérique avec la commande mkswap.

```
[ec2-user ~]$ sudo mkswap /dev/xvdb
mkswap: /dev/xvdb: warning: wiping old ext3 signature.
Setting up swapspace version 1, size = 4188668 KiB
no label, UUID=b4f63d28-67ed-46f0-b5e5-6928319e620b
```

4. Activez le nouvel espace d'échange.

```
[ec2-user ~]$ sudo swapon /dev/xvdb
```

5. Vérifiez que le nouvel espace d'échange est en cours d'utilisation.

```
[ec2-user ~]$ swapon -s
Filename    Type  Size Used Priority
/dev/xvdb                               partition 4188668 0 -1
```

6. Modifiez votre fichier `/etc/fstab` de telle sorte que cet espace d'échange soit automatiquement activé à chaque démarrage système.

```
[ec2-user ~]$ sudo vim /etc/fstab
```

Si votre fichier `/etc/fstab` a une entrée pour `/dev/xvdb` (ou `/dev/sdb`), modifiez-la pour qu'elle corresponde à la ligne ci-dessous ; dans le cas contraire, ajoutez la ligne suivante à votre fichier `/etc/fstab` (à l'aide du nom de périphérique d'échange de votre système) :

```
/dev/xvdb       none    swap    sw  0       0
```

> **Important**
>
> Les données du volume de stockage d'instance sont perdues quand une instance est arrêtée ou mise en veille prolongée. Cela inclut également le formatage de l'espace d'échange du volume de stockage créé dans Step 3 (p. 1280). Si vous arrêtez et redémarrez une instance qui a été configurée pour utiliser un espace d'échange de stockage d'instance, vous devez répéter Step 1 (p. 1279) via Step 5 (p. 1280) sur le nouveau volume de stockage d'instance.

Optimisation des performances disque des volumes de stockage d'instance

En raison de la façon dont Amazon EC2 virtualise les disques, la première écriture sur tout emplacement de certains volumes de stockage d'instance s'effectue plus lentement que les écritures suivantes. Pour la plupart des applications, l'amortissement de ce coût sur la durée de vie de l'instance est acceptable. Cependant, si vous exigez des performances disque élevées, il est recommandé que vous initialisiez vos disques en écrivant une fois sur chaque emplacement disque avant l'utilisation en production.

> **Note**
>
> Certains types d'instance dotés de disques SSD et de la prise en charge de la commande TRIM offrent des performances maximales au lancement, sans initialisation. Pour plus d'informations sur le stockage d'instance pour chaque type d'instance, consultez Volumes de stockage d'instance (p. 1266).

Si vous exigez une plus grande souplesse en termes de latence ou de débit, nous vous recommandons l'utilisation d'Amazon EBS.

Pour initialiser les volumes de stockage d'instance, utilisez les commandes `dd` suivantes, en fonction du stockage à initialiser (par exemple, `/dev/sdb` ou `/dev/nvme1n1`).

> **Note**
>
> Veillez bien à démonter le disque avant d'exécuter cette commande.
> L'initialisation peut durer longtemps (8 heures environ pour une grande instance supplémentaire).

Pour initialiser les volumes de stockage d'instance, utilisez les commandes suivantes sur les types d'instance `m1.large`, `m1.xlarge`, `c1.xlarge`, `m2.xlarge`, `m2.2xlarge` et `m2.4xlarge` :

```
dd if=/dev/zero of=/dev/sdb bs=1M
dd if=/dev/zero of=/dev/sdc bs=1M
dd if=/dev/zero of=/dev/sdd bs=1M
dd if=/dev/zero of=/dev/sde bs=1M
```

Pour initialiser simultanément tous les volumes de stockage d'instance, utilisez la commande suivante :

```
dd if=/dev/zero bs=1M|tee /dev/sdb|tee /dev/sdc|tee /dev/sde > /dev/sdd
```

La configuration des disques d'un système RAID les initialise en écrivant sur chaque emplacement disque. Lors de la configuration d'un système RAID basé sur un logiciel, assurez-vous de modifier la vitesse de reconstruction minimale :

```
echo $((30*1024)) > /proc/sys/dev/raid/speed_limit_min
```

Stockage de fichiers

Le stockage de fichiers dans le cloud est une méthode de stockage des données dans le cloud qui permet aux serveurs et aux applications d'accéder aux données via des systèmes de fichiers partagés. Cette compatibilité rend le stockage de fichiers dans le cloud idéal pour les charges de travail reposant sur des systèmes de fichiers partagés et offre une intégration simple sans modification de code.

Les solutions de stockage de fichiers sont nombreuses : serveur de fichiers à un seul nœud sur une instance de calcul utilisant le stockage en mode bloc comme fondement sans scalabilité ou avec peu de redondances afin de protéger les données, solution en cluster à créer vous-même, ou encore solution entièrement gérée. Le contenu suivant présente certains des services de stockage fournis par AWS pour une utilisation avec Linux.

Table des matières

- Utilisation de Amazon S3 avec Amazon EC2 (p. 1281)
- Utilisation de Amazon EFS avec Amazon EC2 (p. 1283)

Utilisation de Amazon S3 avec Amazon EC2

Amazon S3 est un référentiel pour les données Internet. Amazon S3 fournit un accès à une infrastructure de stockage de données fiable, rapide et économique. Cet outil est conçu pour faciliter l'accès aux ressources informatiques à l'échelle du web en vous permettant de stocker et de récupérer n'importe quelle quantité de données, à tout moment, depuis Amazon EC2 ou depuis n'importe quel emplacement sur le web. Amazon S3 stocke des objets de données de manière redondante sur plusieurs périphériques et installations, et permet un accès simultané en lecture et en écriture à ces objets de données par un grand nombre de clients ou de threads d'application distincts. Vous pouvez utiliser les données redondantes stockées dans Amazon S3 pour récupérer rapidement et en toute fiabilité en cas de défaillances d'une instance ou d'une application.

Amazon EC2 utilise Amazon S3 pour stocker les Amazon Machine Images (AMI). Vous utilisez les AMI pour lancer des instances EC2. En cas de défaillance d'une instance, vous pouvez utiliser l'AMI stockée afin de lancer immédiatement une autre instance, ce qui permet une récupération rapide et une continuité des activités.

Amazon EC2 utilise également Amazon S3 pour stocker des instantanés (copies de sauvegarde) des volumes de données. Vous pouvez utiliser des instantanés (snapshots) pour récupérer des données rapidement et en toute fiabilité en cas de défaillances d'une application ou du système. Vous pouvez également utiliser des instantanés (snapshots) comme base pour créer plusieurs volumes de données, pour augmenter la taille d'un volume de données existant ou pour déplacer des volumes de données entre plusieurs zones de disponibilité (AZ), ce qui rend l'utilisation de vos données hautement évolutive. Pour plus d'informations sur l'utilisation des volumes de données et des instantanés (snapshots), consultez Amazon Elastic Block Store (p. 1084).

Les objets sont les entités fondamentales stockées dans Amazon S3. Chaque objet stocké dans Amazon S3 se trouve dans un compartiment. Les compartiments organisent l'espace de noms Amazon S3 au plus haut niveau et identifient le compte qui assure ce stockage. Les compartiments Amazon S3 sont similaires aux noms de domaines Internet. Les objets stockés dans les compartiments ont une valeur de clé unique et sont récupérés à l'aide d'une URL. Par exemple, si un objet avec la valeur de clé `/photos/mygarden.jpg` est stocké dans le compartiment *`DOC-EXAMPLE-BUCKET1`*, il est adressable à l'aide de l'URL `https://`*`DOC-EXAMPLE-BUCKET1`*`.s3.amazonaws.com/photos/mygarden.jpg`.

Pour plus d'informations sur les fonctions de Amazon S3, consultez la page produit Amazon S3.

Exemples d'utilisation :

Au vu des avantages qu'offre Amazon S3 pour le stockage, vous pouvez décider d'utiliser ce service pour stocker des fichiers et des ensembles de données à utiliser avec des instances EC2. Vous pouvez déplacer des données entre Amazon S3 et vos instances de différentes façons. En plus des exemples présentés ci-après, vous pouvez utiliser de nombreux outils conçus par des utilisateurs pour accéder à vos données dans Amazon S3 depuis votre ordinateur ou votre instance. Les plus courants sont présentés dans les forums AWS.

Si vous y êtes autorisé, vous pouvez copier un fichier vers ou depuis Amazon S3 et votre instance en utilisant l'une des méthodes suivantes.

GET ou wget

L'utilitaire wget est un client HTTP et FTP qui vous permet de télécharger des objets publics depuis Amazon S3. Il est installé par défaut dans Amazon Linux et la plupart des autres distributions, et est disponible en téléchargement sur Windows. Pour télécharger un objet Amazon S3, utilisez la commande suivante, en remplaçant l'URL par celle de l'objet à télécharger.

```
[ec2-user ~]$ wget https://my_bucket.s3.amazonaws.com/path-to-file
```

Cette méthode nécessite que l'objet que vous demandez soit public. Si l'objet n'est pas public, vous recevrez un message indiquant « ERREUR 403 : accès refusé/interdit ». Si vous recevez cette erreur, ouvrez la console Amazon S3 et modifiez les autorisations de l'objet pour le définir comme public. Pour plus d'informations, consultez le Amazon Simple Storage Service Manuel du développeur.

AWS Command Line Interface

AWS Command Line Interface (AWS CLI) est un outil unifié qui permet de gérer vos services AWS. L'AWS CLI permet aux utilisateurs de s'authentifier et de télécharger des éléments restreints depuis Amazon S3, et également de charger des éléments. Pour plus d'informations notamment sur l'installation et la configuration des outils, consultez la page détaillée sur l'AWS Command Line Interface.

La commande aws s3 cp est similaire à la commande Unix cp. Vous pouvez copier des fichiers depuis Amazon S3 vers votre instance, copier des fichiers depuis votre instance vers Amazon S3 et même copier des fichiers d'un emplacement Amazon S3 vers un autre.

Utilisez la commande suivante pour copier un objet depuis Amazon S3 vers votre instance.

```
[ec2-user ~]$ aws s3 cp s3://my_bucket/my_folder/my_file.ext my_copied_file.ext
```

Utilisez la commande suivante pour copier un objet depuis votre instance vers Amazon S3.

```
[ec2-user ~]$ aws s3 cp my_copied_file.ext s3://my_bucket/my_folder/my_file.ext
```

La commande aws s3 sync peut synchroniser un compartiment Amazon S3 entier vers un répertoire local. Cela peut être utile pour télécharger un jeu de données et maintenir la copie locale à jour par rapport au jeu à distance. Si vous disposez des autorisations adéquates sur le compartiment Amazon S3, vous pouvez renvoyer votre répertoire local sur le cloud lorsque vous avez terminé, en inversant les emplacements source et de destination dans la commande.

Utilisez la commande suivante pour télécharger un bucket Amazon S3 entier vers un répertoire local sur votre instance.

```
[ec2-user ~]$ aws s3 sync s3://remote_S3_bucket local_directory
```

API Amazon S3

Si vous êtes un développeur, vous pouvez utiliser une API pour accéder aux données dans Amazon S3. Pour plus d'informations, consultez le Amazon Simple Storage Service Manuel du développeur. Vous pouvez utiliser cette API et ses exemples pour développer votre application et l'intégrer avec d'autres API et SDK, tels que l'interface Python Boto.

Utilisation de Amazon EFS avec Amazon EC2

Amazon EFS offre un stockage de fichiers scalable, destiné à être utilisé avec Amazon EC2. Vous pouvez créer un système de fichiers EFS et configurer vos instances afin de l'installer. Vous pouvez utiliser un système de fichiers EFS comme source de données commune aux charges de travail et applications exécutées sur plusieurs instances. Pour plus d'informations, consultez la page produit d'Amazon Elastic File System.

Dans ce didacticiel, vous créez un système de fichiers EFS et deux instances Linux qui peuvent partager des données à l'aide du système de fichiers.

> **Important**
>
> Amazon EFS n'est pas pris en charge par les instances Windows.

Tâches

- Prérequis (p. 1283)
- Étape 1 : Créer un système de fichiers EFS (p. 1284)
- Étape 2 : Monter le système de fichiers (p. 1284)
- Étape 3 : Tester le système de fichiers (p. 1285)
- Étape 4 : Nettoyer (p. 1286)

Prérequis

- Créez un groupe de sécurité (par exemple, efs-sg) à associé aux instances EC2 et à la cible de montage EFS, puis ajoutez les règles suivantes :
  - Autoriser les connexions SSH entrantes aux instances EC2 en provenance de votre ordinateur (avec le bloc d'adresse CIDR de votre réseau comme source).

- Autoriser les connexions NFS entrantes au système de fichiers via la cible de montage EFS depuis les instances EC2 associées à ce groupe de sécurité (avec le groupe de sécurité lui-même en tant que source). Pour plus d'informations, consultez Règles Amazon EFS (p. 1079) et Création de groupes de sécurité dans le Amazon Elastic File System Guide de l'utilisateur.
- Créer une paire de clé. Vous devez indiquer une paire de clés lorsque vous configurez vos instances, à défaut de quoi vous ne pourrez vous y connecter. Pour plus d'informations, consultez Création d'une paire de clés (p. 28).

Étape 1 : Créer un système de fichiers EFS

Amazon EFS vous permet de créer un système de fichiers que plusieurs instances peuvent monter et auquel elles peuvent simultanément accéder. Pour plus d'informations, consultez Création de ressources pour Amazon EFS dans le Amazon Elastic File System Guide de l'utilisateur.

Pour créer un système de fichiers

1. Ouvrez la console Amazon Elastic File System à l'adresse https://console.aws.amazon.com/efs/.
2. Choisissez Create file system.
3. (Facultatif) Pour Nom, entrez un nom pour le système de fichiers. Une balise ayant la clé Nom et le nom du système de fichiers comme valeur est alors créée.
4. Pour Virtual Private Cloud (VPC), sélectionnez le VPC à utiliser pour vos instances.
5. Sélectionnez Créer.
6. Une fois le système de fichiers créé, notez l'ID du système de fichiers. Celui-ci est utilisé plus tard dans ce didacticiel.
7. Choisissez l'ID du système de fichiers.
8. Sur la page des systèmes de fichiers, choisissez Réseau, Gérer. Affichez les cibles de montage créé par Amazon EFS dans chaque zone de disponibilité de la région dans laquelle réside votre VPC. Pour chaque zone de disponibilité pour vos instances, vérifiez que la valeur de Security group (Groupe de sécurité) correspond au groupe de sécurité que vous avez créé dans Prérequis (p. 1283).
9. Choisissez Enregistrer.

Étape 2 : Monter le système de fichiers

Utilisez la procédure suivante pour lancer deux instances `t2.micro`. Notez que les instances T2 doivent être lancées dans un sous-réseau. Vous pouvez utiliser un VPC par défaut ou un VPC personnalisé.

> **Note**
>
> Il existe d'autres façons de monter le volume (par exemple, sur une instance déjà en cours d'exécution). Pour plus d'informations, consultez la section Montage des systèmes de fichiers du Amazon Elastic File System Guide de l'utilisateur.

Pour lancer deux instances et monter un système de fichiers EFS

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Choisissez Launch Instances.
3. Pour Step 1: Choose an Amazon Machine Image (AMI) [Étape 1 : Choisir une AMI (Amazon Machine Image)], sélectionnez une AMI Amazon Linux.
4. Pour Step 2: Choose an Instance Type [Étape 2 : Choisir un type d'instance), conservez le type d'instance par défaut, `t2.micro`, et choisissez Next: Configure Instance Details (Suivant : Configurer les détails de l'instance).
5. Pour Etape 3 : Configurer les détails de l'instance, procédez comme suit :

a. Pour Number of instances (Nombre d'instances), entrez 2.
b. [VPC par défaut] Si vous avez un VPC par défaut, il correspond à la valeur par défaut du champ Network. Conservez le VPC par défaut et la valeur par défaut du champ Subnet pour utiliser le sous-réseau par défaut dans la zone de disponibilité choisie par Amazon EC2 pour vos instances.

   [VPC personnalisé] Sélectionnez votre VPC dans le champ Network (Réseau) et un sous-réseau public dans Subnet (Sous-réseau).
c. [VPC personnalisé] Dans Auto-assign Public IP, choisissez Enable. Dans le cas contraire, vos instances n'obtiennent pas d'adresses IP publiques ni de noms DNS publics.
d. Pour File systems (Systèmes de fichiers), choisissez Add file system (Ajouter un système de fichiers). Assurez-vous que la valeur correspond à l'ID du système de fichiers que vous avez créé dans Étape 1 : Créer un système de fichiers EFS (p. 1284). Le chemin affiché en regard de l'ID du système de fichiers correspond au point de montage que l'instance utilisera, et vous pouvez le modifier. Sous Advanced details (Détails avancés), les données utilisateur sont générées automatiquement et incluent les commandes nécessaires pour monter le système de fichiers.
e. Passez à l'étape 6 de l'assistant.

6. Sur la page Configure Security Group (Configurer le groupe de sécurité), choisissez Select an existing security group (Sélectionner un groupe de sécurité existant), puis le groupe de sécurité que vous avez créé dans Prérequis (p. 1283). Ensuite, choisissez Vérifier et lancer.
7. Sur la page Review Instance Launch, sélectionnez Launch.
8. Dans la boîte de dialogue Select an existing key pair or create a new key pair, sélectionnez Choose an existing key pair, puis choisissez votre paire de clés. Cochez la case de confirmation, puis sélectionnez Launch Instances.
9. Dans le panneau de navigation, choisissez Instances pour voir l'état de vos instances. Leur état d'origine est `pending`. Lorsqu'elles passent à l'état `running`, vos instances sont prêtes à être utilisées.

Votre instance est désormais configurée pour monter le système de fichiers Amazon EFS au lancement et à chaque redémarrage.

Étape 3 : Tester le système de fichiers

Vous pouvez vous connecter à vos instances et vérifier que le système de fichiers est bien monté dans le répertoire que vous avez indiqué (par exemple, /mnt/efs).

Pour vérifier que le système de fichiers est bien monté

1. Connectez-vous à vos instances. Pour plus d'informations, consultez Connectez-vous à votre instance Linux (p. 598).
2. Dans la fenêtre du terminal de chaque instance, exécutez la commande df -T pour vérifier que le système de fichiers EFS est bien monté.

```
$ df -T
Filesystem     Type              1K-blocks    Used          Available Use% Mounted on
/dev/xvda1     ext4                8123812 1949800            6073764  25% /
devtmpfs       devtmpfs            4078468      56            4078412   1% /dev
tmpfs          tmpfs               4089312       0            4089312   0% /dev/shm
efs-dns        nfs4       9007199254740992       0   9007199254740992   0% /mnt/efs
```

Notez que le nom du système de fichiers, qui s'affiche comme *`efs-dns`* dans l'exemple de résultat, est au format suivant.

```
file-system-id.efs.aws-region.amazonaws.com:/
```

3. (Facultatif) Créez un fichier dans le système de fichiers à partir d'une instance, et vérifiez ensuite que vous pouvez consulter ce fichier à partir de l'autre instance.

   a. Dans la première instance, exécutez la commande suivante pour créer le fichier.

   ```
   $ sudo touch /mnt/efs/test-file.txt
   ```

   b. Dans la deuxième instance, exécutez la commande suivante pour afficher le fichier.

   ```
   $ ls /mnt/efs
   test-file.txt
   ```

Étape 4 : Nettoyer

Lorsque vous avez terminé ce didacticiel, vous pouvez résilier les instances et supprimer le système de fichiers.

Pour résilier les instances

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez les instances à résilier.
4. Choisissez État de l'instance, Résilier l'instance.
5. Choisissez Résilier lorsque vous êtes invité à confirmer.

Pour supprimer le système de fichiers

1. Ouvrez la console Amazon Elastic File System à l'adresse https://console.aws.amazon.com/efs/.
2. Sélectionnez le système de fichiers à supprimer.
3. Choisissez Actions, Delete file system.
4. Lorsque vous êtes invité à confirmer, entrez l'ID du système de fichiers et choisissez Delete file system (Supprimer le système de fichiers).

Limites de volume d'instance

Le nombre maximal de volumes que votre instance peut avoir dépend du système d'exploitation et du type d'instance. Lorsque vous réfléchissez au nombre de volumes à ajouter à votre instance, vous devriez déterminer si vous avez besoin d'une plus grande bande passante E/S ou d'une plus grande capacité de stockage.

Sommaire

- Limites de volume du système Nitro (p. 1286)
- Limites de volume spécifiques à Linux (p. 1287)
- Bande passante et capacité (p. 1287)

Limites de volume du système Nitro

Les instances reposant sur le système Nitro (p. 218) prennent en charge un nombre maximum d'attachements, qui sont partagés entre les interfaces réseau, les volumes EBS et les volumes de stockage

d'instance NVMe. Chaque instance possède au moins un attachement d'interface réseau. Les volumes de stockage d'instance NVMe sont attachés automatiquement. Pour plus d'informations, consultez Interfaces réseau Elastic (p. 840) et Volumes de stockage d'instance (p. 1266).

La plupart de ces instances prennent en charge un maximum de 28 attachements. Si, par exemple, vous ne possédez pas d'interfaces réseau supplémentaires attachées sur une instance EBS uniquement, vous pouvez attacher jusqu'à 27 volumes EBS à cette instance. Si une instance avec 2 volumes de stockage d'instance NVMe ne contient qu'une seule interface réseau supplémentaire, vous pouvez y attacher 24 volumes EBS.

Pour les autres instances, les limites suivantes s'appliquent :

- Les instances `inf1.xlarge` et `inf1.2xlarge`, prennent en charge jusqu'à 26 volumes EBS.
- Les instances `inf1.6xlarge` prennent en charge 23 volumes EBS maximum.
- Les instances `inf1.24xlarge` prennent en charge 11 volumes EBS maximum.
- La plupart des instances matériel nu prennent en charge un maximum de 31 volumes EBS.
- Les instances `u-6tb1.metal`, `u-9tb1.metal` et `u-12tb1.metal` prennent en charge un maximum de 19 volumes EBS s'ils sont lancés après le 12 mars 2020 et un maximum de 14 volumes EBS dans le cas contraire. Pour attacher plus de 14 volumes EBS à une instance lancée avant le 12 mars 2020, contactez votre équipe de compte pour mettre à niveau l'instance sans frais supplémentaires.
- Les instances `u-18tb1.metal` et `u-24tb1.metal` prennent en charge jusqu'à 19 volumes EBS.

Limites de volume spécifiques à Linux

Des échecs de démarrage peuvent se produire si vous attachez plus de 40 volumes. Ce nombre inclut le volume racine, ainsi que tous les volumes de stockage d'instance et les volumes EBS attachés. Si vous rencontrez des problèmes de démarrage sur une instance avec un grand nombre de volumes, arrêtez l'instance, détachez tous les volumes qui ne sont pas essentiels au processus de démarrage, puis rattachez les volumes une fois que l'instance est en cours d'exécution.

> **Important**
>
> Attacher plus de 40 volumes sur une instance Linux est pris en charge autant que possible et ce, sans garantie.

Bande passante et capacité

Pour des cas d'utilisation de bande passante régulière et prévisible, utilisez des instances de connectivité réseau optimisées EBS ou 10 gigabits, et des volumes SSD à usage général ou IOPS provisionnés SSD. Suivez les recommandations données dans Amazon EBS–Instances optimisées (p. 1213) pour associer les E/S par seconde que vous avez provisionnées pour vos volumes et la bande passante disponible depuis vos instances pour des performances maximales. Pour les configurations RAID, de nombreux administrateurs estiment que des grappes de plus de 8 volumes diminuent les retours de performances en raison d'une plus grande surcharge E/S. Testez la performance de votre application individuelle et ajustez-la si nécessaire.

Noms d'appareil pour les instances Linux

Lorsque vous associez un volume à votre instance, vous incluez un nom d'appareil pour le volume. Ce nom d'appareil est utilisé par Amazon EC2. Le pilote du périphérique de stockage en mode bloc de l'instance attribue le nom réel du volume au montage de celui-ci et le nom affecté peut être différent de celui recommandé par Amazon EC2.

Le nombre maximum de volumes que votre instance peut prendre en charge dépend du système d'exploitation. Pour plus d'informations, consultez Limites de volume d'instance (p. 1286).

Sommaire

- Noms d'appareil disponibles (p. 1288)
- Considérations sur les noms d'appareil (p. 1289)

Pour plus d'informations sur les noms d'appareil des instances Windows, consultez Noms d'appareil sur les instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Noms d'appareil disponibles

Deux types de virtualisation sont disponibles pour les instances Linux : la virtualisation paravirtuelle (PV) et la virtualisation HVM. Le type de virtualisation d'une instance est déterminé par l'AMI utilisée pour lancer cette instance. Tous les types d'instance prennent en charge les AMI HVM. Certains types d'instance de la génération précédente prennent en charge les AMI PV. Veillez à noter le type de virtualisation de votre AMI dans la mesure où les noms d'appareil recommandés et disponibles que vous utilisez dépendent du type de virtualisation de votre instance. Pour plus d'informations, consultez Types de virtualisation AMI Linux (p. 108).

Le tableau ci-après répertorie les noms d'appareils disponibles que vous pouvez spécifier dans un mappage de périphérique de stockage en mode bloc ou lorsque vous attachez un volume EBS.

Type de virtualisation	Disponible	Réservé pour la racine	Recommandé pour les volumes EBS	Volumes de stockage d'instance
Paravirtuel	/dev/sd[a-z] /dev/sd[a-z][1-15] /dev/hd[a-z] /dev/hd[a-z][1-15]	/dev/sda1	/dev/sd[f-p] /dev/sd[f-p][1-6]	/dev/sd[b-e]
HVM	/dev/sd[a-z] /dev/xvd[b-c][a-z]	Diffère selon l'AMI /dev/sda1 or /dev/xvda	/dev/sd[f-p] *	/dev/sd[b-e] /dev/sd[b-h] (h1.16xlarge) /dev/sd[b-y] (d2.8xlarge) /dev/sd[b-i] (i2.8xlarge) **

* Les noms d'appareil que vous spécifiez pour les volumes EBS NVMe dans un mappage de périphérique de stockage en mode bloc sont remplacés par les noms du périphérique NVMe (`/dev/nvme[0-26]n1`). Le pilote du périphérique de stockage en mode bloc peut attribuer les noms de périphériques NVMe dans un autre ordre que celui que vous avez spécifié pour les volumes dans le mappage de périphériques de stockage en mode bloc.

** Les volumes de stockage d'instance NVMe sont automatiquement énumérés et un nom d'appareil NVMe leur est automatiquement attribué.

Pour plus d'informations sur les volumes de stockage d'instance, consultez Stockage d'instance Amazon EC2 (p. 1264). Pour plus d'informations sur les volumes NVMe EBS (instances Nitro), notamment sur l'identification du périphérique EBS, consultez Amazon EBS et NVMe sur des instances Linux (p. 1209).

Considérations sur les noms d'appareil

Gardez les points suivants à l'esprit lorsque vous sélectionnez un nom d'appareil :

- Bien que vous puissiez relier vos volumes EBS à l'aide des noms d'appareil utilisés pour relier les volumes de stockage d'instances, nous vous recommandons fortement de ne pas le faire dans la mesure où les résultats peuvent être imprévisibles.
- Le nombre de volumes de stockage d'instance NVMe pour une instance dépend de la taille de cette dernière. Les volumes de stockage d'instance NVMe sont automatiquement énumérés et un nom de périphérique NVMe (`/dev/nvme[0-26]n1`) leur sont automatiquement attribués.
- En fonction du pilote du périphérique de stockage en mode bloc du noyau sélectionné, le périphérique peut être attaché avec un autre nom que celui spécifié. Par exemple, si vous spécifiez un nom de périphérique de `/dev/sdh`, votre appareil peut être renommé `/dev/xvdh` ou `/dev/hdh`. Dans la plupart des cas, la lettre finale reste la même. Dans certaines versions de Red Hat Enterprise Linux (et ses variantes, telles que CentOS), la lettre finale peut changer (`/dev/sda` peut devenir `/dev/xvde`). Dans ces cas, la lettre finale de chaque nom de périphérique est incrémentée le même nombre de fois. Par exemple, si `/dev/sdb` est renommé `/dev/xvdf`, alors `/dev/sdc` est renommé `/dev/xvdg`. Amazon Linux crée un lien symbolique pour le nom que vous avez spécifié pour le périphérique renommé. D'autres systèmes d'exploitation peuvent avoir un comportement différent.
- Les AMI HVM ne prennent pas en charge les chiffres à la fin des noms de périphériques, à l'exception de `/dev/sda1`, qui est réservé pour le périphérique racine, et de `/dev/sda2`. L'utilisation de `/dev/sda2` est possible, mais nous ne recommandons pas l'utilisation de ce mappage de périphérique avec les instances HVM.
- Lorsque vous utilisez des AMI PV, vous ne pouvez pas attacher de volumes qui partagent les mêmes lettres de périphérique, avec et sans chiffres à la fin. Par exemple, si vous attachez un premier volume en tant que `/dev/sdc` et un autre volume en tant que `/dev/sdc1`, seul `/dev/sdc` sera visible pour l'instance. Pour utiliser des chiffres à la fin des noms de périphériques, vous devez y avoir recours pour tous les noms de périphériques qui partagent les mêmes lettres de base (par exemple, `/dev/sdc1`, `/dev/sdc2`, `/dev/sdc3`).
- Certains noyaux personnalisés peuvent avoir des restrictions qui limitent l'utilisation à `/dev/sd[f-p]` ou `/dev/sd[f-p][1-6]`. Si vous rencontrez des difficultés en utilisant `/dev/sd[q-z]` ou `/dev/sd[q-z][1-6]`, essayez avec `/dev/sd[f-p]` ou `/dev/sd[f-p][1-6]`.

Mappage de périphérique de stockage en mode bloc

Chaque instance que vous lancez comporte un volume de périphérique racine associé, qui correspond à un volume Amazon EBS ou à un volume de stockage d'instance. Vous pouvez utiliser un mappage de périphérique de stockage en mode bloc pour spécifier des volumes EBS supplémentaires ou des volumes de stockage d'instance à attacher à une instance lorsqu'elle est lancée. Vous pouvez également attacher des volumes EBS supplémentaires à une instance en cours d'exécution. Consultez Attacher un volume Amazon EBS à une instance (p. 1109). Cependant, le seul moyen d'attacher des volumes de stockage d'instance à une instance est d'utiliser le mappage de périphérique de stockage en mode bloc pour les attacher lors du lancement de l'instance.

Pour plus d'informations sur les volumes du périphérique racine, consultez Modification du volume racine pour qu'il persiste (p. 25).

Table des matières

- Concepts de mappage de périphérique de stockage en mode bloc (p. 1290)
- Mappage de périphérique de stockage en mode bloc d'une AMI (p. 1293)
- Mappage de périphérique de stockage en mode bloc d'une instance (p. 1295)

Concepts de mappage de périphérique de stockage en mode bloc

Un périphérique de stockage en mode bloc est un dispositif de stockage qui déplace des données en séquence d'octets ou bits (blocs). Ces dispositifs prennent en charge l'accès aléatoire et utilisent généralement des E/S mis en mémoire tampon. Ce sont par exemple des disques durs, des lecteurs de CD-ROM et des lecteurs flash. Un périphérique de stockage en mode bloc peut être physiquement attaché à un ordinateur ou accessible à distance comme s'il était physiquement attaché à l'ordinateur.

Amazon EC2 prend en charge deux types de périphériques de stockage en mode bloc :

- Les volumes de stockage d'instance (périphériques virtuels dont le matériel sous-jacent est physiquement attaché à l'ordinateur hôte de l'instance)
- Les volumes EBS (périphériques de stockage à distance)

Un mappage de périphérique de stockage en mode bloc définit les périphériques de stockage en mode bloc (volumes de stockage d'instance et volumes EBS) qui doivent être attachés à l'instance. Vous pouvez spécifier un mappage de périphérique de stockage en mode bloc lors de la création d'une AMI, afin que le mappage soit utilisé par toutes les instances lancées à partir de l'AMI. Vous pouvez également spécifier un mappage de périphérique de stockage en mode bloc lorsque vous lancez une instance, afin que son mappage remplace celui spécifié dans l'AMI à partir de laquelle vous avez lancé l'instance. Notez que tous les volumes de stockage d'instance NVMe pris en charge par un type d'instance sont automatiquement énumérés et un nom de périphérique leur est automatiquement attribué au lancement de l'instance. Le fait de les ajouter dans votre mappage de périphérique de stockage en mode bloc n'a aucun effet.

Table des matières

- Entrées du mappage de périphérique de stockage en mode bloc (p. 1290)
- Mises en garde sur le stockage d'instance du mappage de périphérique de stockage en mode bloc (p. 1291)
- Exemple de mappage de périphérique de stockage en mode bloc (p. 1292)
- Mise à disposition d'appareils dans le système d'exploitation (p. 1292)

Entrées du mappage de périphérique de stockage en mode bloc

Lorsque vous créez un mappage de périphérique de stockage en mode bloc, vous spécifiez les informations suivantes pour chaque périphérique de stockage en mode bloc qui doit être attaché à l'instance :

- Le nom du périphérique utilisé dans Amazon EC2. Le pilote du périphérique de stockage en mode bloc de l'instance attribue le nom de volume réel lors du montage du volume. Le nom attribué peut être différent de celui recommandé par Amazon EC2. Pour plus d'informations, consultez Noms d'appareil pour les instances Linux (p. 1287).

Pour les volumes de stockage d'instance, vous spécifiez également les informations suivantes :

- Le nom du périphérique virtuel : `ephemeral[0-23]`. Notez que le nombre et la taille des volumes de stockage d'instance disponibles pour votre instance varient en fonction du type d'instance.

Pour les volumes de stockage d'instance NVMe, les informations suivantes s'appliquent également :

- Ces volumes sont automatiquement énumérés et un nom de périphérique leur est automatiquement attribué. Le fait de les ajouter dans votre mappage de périphérique de stockage en mode bloc n'a aucun effet.

Pour les volumes EBS, vous spécifiez également les informations suivantes :

- L'ID de l'instantané à utiliser pour créer le périphérique de stockage en mode bloc (snap-xxxxxxxx). Cette valeur est facultative si vous spécifiez une taille de volume.
- Taille du volume en Gio La taille spécifiée doit être supérieure ou égale à la taille de l'instantané spécifié.
- Suppression ou non du volume lors de l'arrêt de l'instance (`true` ou `false`). La valeur par défaut est `true` pour le volume du périphérique racine et `false` pour les volumes attachés. Lorsque vous créez une AMI, son mappage de périphérique de stockage en mode bloc hérite de ce paramètre de l'instance. Lorsque vous lancez une instance, elle hérite de ce paramètre de l'AMI.
- Le type de volume, qui peut être `gp2` pour SSD à usage général, `io1` ou `io2` pour IOPS provisionnés SSD, `st1` pour HDD à débit optimisé, `sc1` pour Cold HDD ou `standard` pour magnétique. La valeur par défaut est `gp2`.
- Le nombre d'opérations d'E/S par seconde (IOPS) prises en charge par le volume. (Non utilisé avec les volumes `gp2`, `st1`, `sc1` ou `standard`.)

Mises en garde sur le stockage d'instance du mappage de périphérique de stockage en mode bloc

Vous devez prendre en compte plusieurs mises en garde lorsque vous lancez des instances avec des AMIs comportant des volumes de stockage d'instance dans leurs mappages de périphérique de stockage en mode bloc.

- Certains types d'instance comprennent un plus grand nombre de volumes de stockage d'instance que d'autres et certains types d'instance ne contiennent aucun volume de stockage d'instance. Si votre type d'instance prend en charge un volume de stockage d'instance et que votre AMI comporte des mappages pour deux volumes de stockage d'instance, l'instance est lancée avec un volume de stockage d'instance.
- Les volumes de stockage d'instance peuvent uniquement être mappés au moment du lancement. Vous ne pouvez pas arrêter une instance sans volume de stockage d'instance (comme `t2.micro`), modifier le type de l'instance par un type prenant en charge les volumes de stockage d'instance, puis redémarrer l'instance avec des volumes de stockage d'instance. En revanche, vous pouvez créer une AMI à partir de l'instance et la lancer sur un type d'instance prenant en charge les volumes de stockage d'instance, et mapper ces volumes de stockage d'instance à l'instance.
- Si vous lancez une instance à laquelle sont mappés des volumes de stockage d'instance, puis arrêtez l'instance, en modifiez le type avec un nombre inférieur de volumes de stockage d'instance et redémarrez l'instance, les mappages des volumes de stockage d'instance du lancement initial apparaissent toujours dans les métadonnées de l'instance. Cependant, l'instance n'a accès qu'au nombre maximum de volumes de stockage d'instance pris en charge pour ce type d'instance.

  Note

  Lorsqu'une instance est arrêtée, toutes les données stockées sur les volumes de stockage d'instance sont perdues.
- En fonction de la capacité de stockage de l'instance au moment du lancement, les instances M3 peuvent ignorer les mappages de périphérique de stockage en mode bloc du stockage d'instance AMI au moment du lancement, sauf s'ils sont spécifiés au moment du lancement. Il est conseillé de spécifier les mappages de périphérique de stockage en mode bloc du stockage d'instance au moment du lancement, même si les volumes d'instance de stockage de l'AMI que vous lancez sont mappés dans l'AMI, afin de garantir la disponibilité des volumes de stockage d'instance au lancement de l'instance.

Exemple de mappage de périphérique de stockage en mode bloc

L'illustration suivante montre un exemple de mappage de périphérique de stockage en mode bloc pour une instance basée sur les volumes EBS. `/dev/sdb` est mappé à `ephemeral0` et deux volumes EBS sont mappés. L'un à `/dev/sdh` et l'autre à `/dev/sdj`. La figure illustre également le volume EBS qui est le volume du périphérique racine, `/dev/sda1`.

Notez que cet exemple de mappage de périphérique de stockage en mode bloc est utilisé dans les exemples de commandes et d'API de cette rubrique. Les exemples de commandes et d'API qui créent les mappages de périphérique de stockage en mode bloc sont disponibles dans les sections Spécification d'un mappage de périphérique de stockage en mode bloc pour une AMI (p. 1293) et Mise à jour du mappage de périphérique de stockage en mode bloc lors du lancement d'une instance (p. 1296).

Mise à disposition d'appareils dans le système d'exploitation

Les noms de périphériques tels que `/dev/sdh` et `xvdh` sont utilisés par Amazon EC2 pour décrire des périphériques de stockage en mode bloc. Le mappage de périphérique de stockage en mode bloc est utilisé par Amazon EC2 pour spécifier les périphériques de stockage en mode bloc qui doivent être attachés à une instance EC2. Lorsqu'un périphérique de stockage en mode bloc est attaché à une instance, il doit être monté par le système d'exploitation pour que vous puissiez accéder au dispositif de stockage. Lorsqu'un périphérique de stockage en mode bloc est détaché d'une instance, il doit être démonté par le système d'exploitation. Ainsi, vous ne pouvez plus accéder au dispositif de stockage.

Avec une instance Linux, les noms des appareils spécifiés dans le mappage de périphérique de stockage en mode bloc sont mappés à leurs périphériques de stockage en mode bloc correspondants au premier démarrage de l'instance. Le type d'instance détermine les volumes de stockage d'instance qui sont formatés et montés par défaut. Vous pouvez monter des volumes de stockage d'instance supplémentaires au moment du lancement, à condition de ne pas dépasser le nombre de volumes de stockage d'instance disponibles pour votre type d'instance. Pour plus d'informations, consultez Stockage d'instance Amazon EC2 (p. 1264). Le pilote du périphérique de stockage en mode bloc pour l'instance détermine les périphériques utilisés lorsque les volumes sont formatés et montés. Pour plus d'informations, consultez Attacher un volume Amazon EBS à une instance (p. 1109).

Mappage de périphérique de stockage en mode bloc d'une AMI

Chaque AMI comporte un mappage de périphérique de stockage en mode bloc qui spécifie les périphériques de stockage en mode bloc à attacher à une instance lancée à partir de l'AMI. Une AMI fournie par Amazon comprend uniquement un périphérique racine. Pour ajouter d'autres périphériques de stockage en mode bloc à une AMI, vous devez créer votre propre AMI.

Sommaire

- Spécification d'un mappage de périphérique de stockage en mode bloc pour une AMI (p. 1293)
- Affichage des volumes EBS dans le mappage de périphérique de stockage en mode bloc d'une AMI (p. 1295)

Spécification d'un mappage de périphérique de stockage en mode bloc pour une AMI

Lorsque vous créez une AMI, il existe deux façons de spécifier des volumes en plus du volume racine. Si vous avez déjà attaché des volumes à une instance en cours d'exécution avant de créer une AMI à partir de l'instance, le mappage de périphérique de stockage en mode bloc pour l'AMI comprend ces mêmes volumes. Pour les volumes EBS, les données existantes sont enregistrées dans un nouvel instantané. C'est ce nouvel instantané qui est spécifié dans le mappage de périphérique de stockage en mode bloc. Pour les volumes de stockage d'instance, les données ne sont pas conservées.

Pour une AMI basée sur des volumes EBS, vous pouvez ajouter des volumes EBS et des volumes de stockage d'instance à l'aide d'un mappage de périphérique de stockage en mode bloc. Pour une AMI basée sur le stockage d'instance, vous pouvez ajouter des volumes de stockage d'instance uniquement en modifiant les entrées du mappage de périphérique de stockage en mode bloc dans le fichier manifest des images lors de l'enregistrement de l'image.

> **Note**
>
> Pour les instances M3, vous devez spécifier les volumes de stockage d'instance dans le mappage de périphérique de stockage en mode bloc de l'instance lorsque cette dernière est lancée. Lorsque vous lancez une instance M3, les volumes de stockage d'instance spécifiés dans le mappage de périphérique de stockage en mode bloc de l'AMI peuvent être ignorés s'ils ne sont pas spécifiés dans le cadre du mappage de périphérique de stockage en mode bloc de l'instance.

Pour ajouter des volumes à une AMI à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, choisissez Instances.
3. Sélectionnez une instance, puis choisissez Actions, Image, Create Image.
4. Dans la boîte de dialogue Create Image, choisissez Add New Volume.

5. Sélectionnez un type de volume dans la liste Type et un nom de périphérique dans la liste Device. Pour un volume EBS, vous pouvez éventuellement spécifier un instantané, une taille de volume et un type de volume.
6. Choisissez Créer une image.

Pour ajouter des volumes à une AMI à l'aide de la ligne de commande

Utilisez la commande AWS CLI create-image pour spécifier un mappage de périphérique de stockage en mode bloc pour une AMI basée sur des volumes EBS. Utilisez la commande AWS CLI register-image afin de spécifier un mappage de périphérique de stockage en mode bloc pour une AMI basée sur le stockage d'instance.

Spécifiez le mappage de périphérique de stockage en mode bloc à l'aide du paramètre `--block-device-mappings`. Les arguments encodés en JSON peuvent être fournis soit directement depuis la ligne de commande soit par référence à un fichier :

```
--block-device-mappings [mapping, ...]
--block-device-mappings [file://mapping.json]
```

Pour ajouter un volume de stockage d'instance, utilisez le mappage suivant :

```
{
    "DeviceName": "/dev/sdf",
    "VirtualName": "ephemeral0"
}
```

Pour ajouter un volume `gp2` vide de 100 Gio, utilisez le mappage suivant :

```
{
    "DeviceName": "/dev/sdg",
    "Ebs": {
      "VolumeSize": 100
    }
}
```

Pour ajouter un volume EBS basé sur un instantané, utilisez le mappage suivant :

```
{
    "DeviceName": "/dev/sdh",
    "Ebs": {
      "SnapshotId": "snap-xxxxxxxx"
    }
}
```

Pour omettre un mappage pour un périphérique, utilisez le mappage suivant :

```
{
    "DeviceName": "/dev/sdj",
    "NoDevice": ""
}
```

Vous pouvez aussi utiliser le paramètre `-BlockDeviceMapping` avec les commandes suivantes (Outils AWS pour Windows PowerShell) :

- New-EC2Image
- Register-EC2Image

Affichage des volumes EBS dans le mappage de périphérique de stockage en mode bloc d'une AMI

Vous pouvez facilement énumérer les volumes EBS du mappage de périphérique de stockage en mode bloc pour une AMI.

Pour afficher les volumes EBS pour une AMI à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, sélectionnez AMI.
3. Choisissez EBS images dans la liste Filter pour obtenir la liste des AMI basées sur EBS.
4. Sélectionnez l'AMI souhaitée et consultez l'onglet Details. Au minimum, les informations suivantes sont disponibles pour le périphérique racine :

   - Type de périphérique racine (`ebs`)
   - Nom du périphérique racine (par exemple, `/dev/sda1`)
   - Block Devices (par exemple, `/dev/sda1=snap-1234567890abcdef0:8:true`)

   Si l'AMI a été créée avec des volumes EBS supplémentaires à l'aide d'un mappage de périphérique de stockage en mode bloc, le champ Block Devices affiche également le mappage pour ces volumes supplémentaires. Notez que cet écran n'affiche pas les volumes de stockage d'instance.

Pour afficher les volumes EBS d'une AMI à l'aide de la ligne de commande

Utilisez la commande describe-images (AWS CLI) ou la commande Get-EC2Image (Outils AWS pour Windows PowerShell) afin d'énumérer les volumes EBS dans le mappage de périphérique de stockage en mode bloc pour une AMI.

Mappage de périphérique de stockage en mode bloc d'une instance

Par défaut, une instance que vous lancez comprend tous les périphériques de stockage spécifiés dans le mappage de périphérique de stockage en mode bloc de l'AMI à partir de laquelle vous avez lancé l'instance. Vous pouvez spécifier les modifications apportées au mappage de périphérique de stockage en mode bloc d'une instance lorsque vous la lancez. Ces mises à jour remplacent le mappage de périphérique de stockage en mode bloc de l'AMI ou fusionnent avec.

Limites

- Pour le volume racine, vous pouvez uniquement modifier les données informations suivantes : taille du volume, type de volume et indicateur Delete on Termination.
- Lorsque vous modifiez un volume EBS, vous ne pouvez pas en diminuer la taille. Vous devez donc spécifier un instantané dont la taille est égale ou supérieure à celle de l'instantané spécifié dans le mappage de périphérique de stockage en mode bloc de l'AMI.

Table des matières

- Mise à jour du mappage de périphérique de stockage en mode bloc lors du lancement d'une instance (p. 1296)
- Mise à jour du mappage de périphérique de stockage en mode bloc d'une instance en cours d'exécution (p. 1297)

- Affichage des volumes EBS dans le mappage de périphérique de stockage en mode bloc d'une instance (p. 1298)
- Affichage du mappage de périphérique de stockage en mode bloc d'une instance pour les volumes de stockage d'instances (p. 1299)

Mise à jour du mappage de périphérique de stockage en mode bloc lors du lancement d'une instance

Vous pouvez ajouter des volumes EBS et des volumes de stockage d'instance à une instance lors de son lancement. Notez que la mise à jour du mappage de périphérique de stockage en mode bloc d'une instance n'entraîne pas de modification permanente du mappage de périphérique de stockage en mode bloc de l'AMI à partir de laquelle il a été lancé.

Pour ajouter des volumes à une instance à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Sur le tableau de bord, choisissez Lancer une instance.
3. Sur la page Choose an Amazon Machine Image (AMI), sélectionnez l'AMI à utiliser, puis choisissez Select.
4. Suivez l'Assistant pour compléter les pages Choisir un type d'instance et Configurer les détails de l'instance.
5. Sur la page Add Storage, vous pouvez modifier le volume racine, les volumes EBS et les volumes de stockage d'instance de la façon suivante :
   - Pour modifier la taille du volume racine, recherchez le volume Root dans la colonne Type, et modifiez le champ Size.
   - Pour supprimer un volume EBS spécifié par le mappage de périphérique de stockage en mode bloc de l'AMI utilisée pour lancer l'instance, recherchez le volume et cliquez sur l'icône Delete qui lui correspond.
   - Pour ajouter un volume EBS, choisissez Add New Volume, puis EBS dans la liste Type, et renseignez les champs (Device, Snapshot, etc).
   - Pour supprimer un volume de stockage d'instance spécifié par le mappage de périphérique de stockage en mode bloc de l'AMI utilisée pour lancer l'instance, recherchez le volume et choisissez l'icône Delete qui lui correspond.
   - Pour ajouter un volume de stockage d'instance, choisissez Add New Volume, sélectionnez Instance Store dans la liste Type, puis choisissez un nom de périphérique dans la liste Device.
6. Complétez les pages restantes de l'Assistant, puis sélectionnez Launch.

Pour ajouter des volumes à une instance à l'aide de l'AWS CLI

Utilisez la commande AWS CLI run-instances avec l'option `--block-device-mappings` pour spécifier un mappage de périphérique de bloc pour une instance au lancement.

Supposons par exemple qu'une AMI basée sur des volumes EBS spécifie le mappage de périphérique de stockage en mode bloc suivant :

- /dev/sdb=ephemeral0
- /dev/sdh=snap-1234567890abcdef0
- /dev/sdj=:100

Pour empêcher l'attachement de `/dev/sdj` à une instance lancée à partir de cette AMI, utilisez le mappage suivant.

```
{
    "DeviceName": "/dev/sdj",
    "NoDevice": ""
}
```

Pour augmenter la taille de `/dev/sdh` à 300 Gio, spécifiez le mappage suivant. Notez que vous ne devez pas spécifier l'ID d'instantané pour `/dev/sdh`, car le fait de spécifier le nom du périphérique suffit à identifier le volume.

```
{
    "DeviceName": "/dev/sdh",
    "Ebs": {
      "VolumeSize": 300
    }
}
```

Pour augmenter la taille du volume racine au lancement de l'instance, appelez d'abord describe-images avec l'ID de l'AMI pour vérifier le nom de l'appareil du volume racine. Par exemple, `"RootDeviceName": "/dev/xvda"`. Pour remplacer la taille du volume racine, spécifiez le nom de l'appareil racine utilisé par l'AMI et la nouvelle taille du volume.

```
{
    "DeviceName": "/dev/xvda",
    "Ebs": {
      "VolumeSize": 100
    }
}
```

Pour attacher un volume de stockage d'instance supplémentaire, `/dev/sdc`, spécifiez le mappage suivant. Si le type d'instance ne prend pas en charge plusieurs volumes de stockage d'instance, ce mappage n'a aucun effet. Si l'instance prend en charge les volumes de stockage d'instance NVMe, ils sont automatiquement énumérés et un nom d'appareil NVMe leur est attribué.

```
{
    "DeviceName": "/dev/sdc",
    "VirtualName": "ephemeral1"
}
```

Pour ajouter des volumes à une instance à l'aide de l'Outils AWS pour Windows PowerShell

Utilisez le paramètre `-BlockDeviceMapping` avec la commande New-EC2Instance (Outils AWS pour Windows PowerShell)

Mise à jour du mappage de périphérique de stockage en mode bloc d'une instance en cours d'exécution

Vous pouvez utiliser la commande modify-instance-attribute de l'AWS CLI pour mettre à jour le mappage de périphérique de stockage en mode bloc d'une instance en cours d'exécution. Vous n'avez pas besoin d'arrêter l'instance avant de modifier cet attribut.

```
aws ec2 modify-instance-attribute --instance-id i-1a2b3c4d --block-device-mappings file://
mapping.json
```

Par exemple, pour conserver le volume racine à la clôture de l'instance, spécifiez les informations suivantes dans le fichier `mapping.json`.

```
[
  {
    "DeviceName": "/dev/sda1",
    "Ebs": {
      "DeleteOnTermination": false
    }
  }
]
```

Vous pouvez aussi utiliser le paramètre `-BlockDeviceMapping` avec la commande Edit-EC2InstanceAttribute (Outils AWS pour Windows PowerShell).

Affichage des volumes EBS dans le mappage de périphérique de stockage en mode bloc d'une instance

Vous pouvez facilement énumérer les volumes EBS mappés à une instance.

> **Note**
>
> Pour les instances lancées avant la version du 31/10/2009 de l'API, AWS ne peut pas afficher le mappage de périphérique de stockage en mode bloc. Vous devez détacher et attacher à nouveau les volumes afin qu'AWS puisse afficher le mappage de périphérique de stockage en mode bloc.

Pour afficher les volumes EBS pour une instance à l'aide de la console

1. Ouvrez la console Amazon EC2.
2. Dans le panneau de navigation, choisissez Instances.
3. Dans la barre de recherche, tapez Type de périphérique racine, puis choisissez EBS. Une liste des instances basées sur des volumes EBS s'affiche.
4. Sélectionnez l'instance souhaitée, puis consultez les informations affichées dans l'onglet Description. Au minimum, les informations suivantes sont disponibles pour le périphérique racine :

   - Type de périphérique racine (`ebs`)
   - Périphérique racine (par exemple, `/dev/sda1`)
   - Périphériques de stockage en mode bloc (par exemple, `/dev/sda1`, `/dev/sdh` et `/dev/sdj`)

   Si l'instance a été lancée avec des volumes EBS supplémentaires à l'aide d'un mappage de périphérique de stockage en mode bloc, le champ Périphériques de stockage en mode bloc affiche ces volumes supplémentaires ainsi que le périphérique racine. Notez que cet écran n'affiche pas les volumes de stockage d'instance.

5. Pour afficher plus d'informations à propos d'un périphérique de stockage en mode bloc, choisissez l'entrée correspondante en regard de Périphériques de stockage en mode bloc. Les informations suivantes s'affichent pour le périphérique de stockage en mode bloc :

   - ID d'EBS (vol-xxxxxxxx)
   - Type de périphérique racine (`ebs`)
   - Heure d'attachement (aaaa-mmThh:mm:ss.ssTZD)

- Statut du périphérique de stockage en mode bloc (`attaching`, `attached`, `detaching`, `detached`)
- Supprimer à la résiliation (`Yes`, `No`)

Pour afficher les volumes EBS d'une instance à l'aide de la ligne de commande

Utilisez la commande describe-instances (AWS CLI) ou Get-EC2Instance (Outils AWS pour Windows PowerShell) pour énumérer les volumes EBS du mappage de périphérique de stockage en mode bloc pour une instance.

Affichage du mappage de périphérique de stockage en mode bloc d'une instance pour les volumes de stockage d'instances

Lorsque vous affichez le mappage de périphérique de stockage en mode bloc de votre instance, vous pouvez uniquement voir les volumes EBS, mais vous ne pouvez pas voir les volumes de stockage d'instance. Vous pouvez utiliser des métadonnées d'instance pour interroger les volumes de stockage d'instances non VNMe dans le mappage d'appareils en bloc. Les volumes de stockage d'instances NVMe ne sont pas inclus.

L'URI de base pour toutes les requêtes de métadonnées des instances est `http://169.254.169.254/latest/`. Pour plus d'informations, consultez Métadonnées d'instance et données utilisateur (p. 701).

Commencez par vous connecter à votre instance en cours d'exécution. Utilisez cette requête à partir de l'instance pour obtenir son mappage de périphérique de stockage en mode bloc.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/block-device-mapping/
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/block-device-mapping/
```

La réponse inclut les noms des périphériques de stockage en mode bloc de l'instance. Par exemple, la sortie pour une instance `m1.small` basée sur un stockage d'instance ressemble à cela :

```
ami
ephemeral0
root
swap
```

Le périphérique `ami` est le périphérique racine tel que le voit l'instance. Les volumes de stockage d'instance sont nommés `ephemeral[0-23]`. Le périphérique `swap` est utilisé pour le fichier d'échange. Si vous avez également mappé des volumes EBS, ils apparaissent en tant que `ebs1`, `ebs2`, etc.

Pour obtenir des détails relatifs à un périphérique de stockage en mode bloc individuel dans le mappage de périphérique de stockage en mode bloc, ajoutez son nom à la requête précédente, comme illustré ici.

IMDSv2

```
[ec2-user ~]$ TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-
ec2-metadata-token-ttl-seconds: 21600"` \
```

```
&& curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/latest/meta-
data/block-device-mapping/ephemeral0
```

IMDSv1

```
[ec2-user ~]$ curl http://169.254.169.254/latest/meta-data/block-device-mapping/
ephemeral0
```

Le type d'instance détermine le nombre de volumes de stockage d'instances disponibles pour l'instance. Si le nombre de volumes de stockage d'instances dans un mappage d'appareils en bloc dépasse le nombre de volumes de stockage d'instances disponibles pour une instance, les volumes supplémentaires sont ignorés. Pour afficher les volumes de stockage d'instances disponibles sur votre instance, exécutez la commande lsblk. Pour voir combien de volumes de stockage d'instances sont pris en charge par chaque type d'instance, veuillez consulter Volumes de stockage d'instance (p. 1266).

Ressources et balises

Amazon EC2 fournit différentes ressources que vous pouvez créer et utiliser. Certaines de ces ressources incluent des images, des instances, des volumes et des instantanés. Lorsque vous créez une ressource, nous lui affectons un ID de ressource unique.

Certaines ressources peuvent être balisées avec des valeurs que vous définissez pour mieux les organiser et les identifier.

Les rubriques suivantes décrivent les ressources et les balises, et expliquent comment les utiliser.

Sommaire

- Emplacements des ressources (p. 1301)
- ID de ressource (p. 1302)
- Liste et filtrage de vos ressources (p. 1303)
- Balisage de vos ressources Amazon EC2 (p. 1309)
- Quotas de service Amazon EC2 (p. 1322)
- Rapports d'utilisation d'Amazon EC2 (p. 1324)

Emplacements des ressources

Certaines ressources peuvent être utilisées dans toutes les régions (international) et certaines autres sont spécifiques à la région ou la zone de disponibilité dans lesquelles elles résident.

Ressource	Type	Description
Compte AWS	Solution internationale	Vous pouvez utiliser le même compte AWS dans toutes les régions.
Paires de clés	Mondial ou régional	Les paires de clés que vous créez à l'aide d'Amazon EC2 sont liées à la région où vous les avez créées. Vous pouvez créer votre propre paire de clés RSA et la télécharger dans la région dans laquelle vous voulez l'utiliser. Par conséquent, vous pouvez la mettre à disposition dans le monde entier en la téléchargeant dans chaque région. Pour plus d'informations, consultez Paires de clés Amazon EC2 et instances Linux (p. 1047).
Identifiants de ressource Amazon EC2	Régional	Chaque identificateur de ressource (tel qu'un ID d'AMI, d'instance, de volume EBS ou d'instantané) est lié à sa région et peut être utilisé uniquement dans la région où vous avez créé la ressource.
Noms de ressource fournis par l'utilisateur	Régional	Chaque nom de ressource (comme un nom de groupe de sécurité ou de paire de clés) est lié à sa région et peut être utilisé uniquement dans la région où vous avez créé la ressource. Même si vous pouvez créer des ressources avec le même nom dans plusieurs régions, ces ressources ne sont pas liées.
AMI	Régional	Une AMI est liée à la région dans laquelle ses fichiers sont situés au sein d'Amazon S3. Vous pouvez

Ressource	Type	Description
		copier une AMI d'une région à une autre. Pour plus d'informations, consultez Copie une AMI (p. 173).
Adresses IP Elastic	Régional	Une adresse IP Elastic est liée à une région et ne peut être associée qu'à une instance de la même région.
Groupes de sécurité	Régional	Un groupe de sécurité est lié à une région et ne peut être affecté qu'aux instances de la même région. Vous ne pouvez pas permettre à une instance de communiquer avec une instance se trouvant en dehors de sa région à l'aide de règles de groupe de sécurité. Le trafic provenant d'une instance située dans une autre région est considéré comme un trafic à bande passante de réseau étendu (WAN).
Instantanés EBS	Régional	Un instantané (snapshot) EBS est lié à sa région et peut être uniquement utilisé pour créer des volumes dans la même région. Vous pouvez copier un instantané d'une région à une autre. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).
Volumes EBS	Zone de disponibilité	Un volume Amazon EBS est lié à sa zone de disponibilité et peut être uniquement attaché à des instances de la même zone de disponibilité.
Instances	Zone de disponibilité	Une instance est liée à la zone de disponibilité dans laquelle vous l'avez lancée. Cependant, son ID d'instance est lié à la région.

ID de ressource

Lorsque des ressources sont créées, nous affectons à chacune d'entre elles un ID de ressource unique. Un ID de ressource est constitué d'un identificateur de ressource (par exemple, `snap` pour un instantané) suivi d'un tiret et d'une combinaison unique de lettres et de chiffres.

Vous pouvez utiliser des ID de ressource pour rechercher vos ressources sur la console Amazon EC2. Si vous utilisez un outil de ligne de commande ou l'API Amazon EC2 pour gérer Amazon EC2, des ID de ressource sont requis pour certaines commandes. Par exemple, si vous utilisez la commande stop-instances AWS CLI pour arrêter une instance, vous devez spécifier l'ID de l'instance dans la commande.

Longueur des ID de ressource

Avant janvier 2016, les ID affectés aux ressources nouvellement créées de certains types de ressource utilisaient 8 caractères après le tiret (par exemple, i-1a2b3c4d). De janvier 2016 à juin 2018, nous avons modifié les ID de ces types de ressource pour utiliser 17 caractères après le tiret (par exemple, `i-1234567890abcdef0`). Selon le moment où votre compte a été créé, vous pouvez disposer de ressources des types suivants avec des ID courts, bien que toutes les nouvelles ressources de ces types reçoivent les ID longs :

- bundle
- conversion-task
- customer-gateway
- dhcp-options
- elastic-ip-allocation

- elastic-ip-association
- export-task
- flow-log
- image
- import-task
- instance
- internet-gateway
- network-acl
- network-acl-association
- network-interface
- network-interface-attachment
- prefix-list
- route-table
- route-table-association
- security-group
- instantané
- sous-réseau
- subnet-cidr-block-association
- réservation
- volume
- vpc
- vpc-cidr-block-association
- vpc-endpoint
- vpc-peering-connection
- vpn-connection
- vpn-gateway

Liste et filtrage de vos ressources

Vous pouvez obtenir la liste de certains types de ressource à l'aide de la console Amazon EC2. Vous pouvez obtenir une liste de chaque type de ressource à l'aide de sa commande ou de son action d'API correspondante. Si vous avez plusieurs ressources, vous pouvez filtrer les résultats pour n'inclure que les ressources qui correspondent à certains critères.

Sommaire

- Création d'une liste et filtrage des ressources à l'aide de la console (p. 1303)
- Affichage et filtrage à l'aide de l'interface de commande et des API (p. 1307)

Création d'une liste et filtrage des ressources à l'aide de la console

Table des matières

- Affichage des ressources à l'aide de la console (p. 1304)
- Filtrage des ressources à l'aide de la console (p. 1304)

Affichage des ressources à l'aide de la console

Vous pouvez afficher les types de ressource Amazon EC2 les plus courants à l'aide de la console. Pour afficher des ressources supplémentaires, utilisez l'interface ligne de commande ou les actions d'API.

Pour afficher les ressources EC2 à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le volet de navigation, choisissez l'option qui correspond à la ressource. Par exemple, pour créer une liste de vos instances, choisissez Instances.
3. La page affiche toutes les ressources du type de ressource sélectionné.

Filtrage des ressources à l'aide de la console

La fonction de recherche diffère légèrement entre l'ancienne et la nouvelle console Amazon EC2.

Nouvelle console

La nouvelle console prend en charge deux types de filtrage.

- Le filtrage des API se produit côté serveur. Le filtrage est appliqué à l'appel d'API et réduit le nombre de ressources renvoyées par le serveur. Il permet un filtrage rapide sur des ensembles volumineux de ressources et peut réduire le temps et le coût du transfert de données entre le serveur et le navigateur.
- Le filtrage client se produit du côté du client. Il vous permet de filtrer les données déjà disponibles dans le navigateur (en d'autres termes, les données qui ont déjà été renvoyées par l'API). Le filtrage client fonctionne parfaitement en conjonction avec un filtre d'API de manière à réduire le filtrage à de plus petits ensembles de données dans le navigateur.

La nouvelle console Amazon EC2 prend en charge les types de recherche suivants :

Recherche par mot-clé

La recherche par mot clé est une recherche de texte libre qui vous permet de rechercher une valeur parmi tous les attributs de vos ressources, sans spécifier l'attribut à rechercher.

> **Note**
>
> Toutes les recherches par mots-clés utilisent le filtrage client.

Pour rechercher par mot-clé, entrez ou collez ce que vous recherchez dans la zone de recherche, puis choisissez Enter (Entrer). Par exemple, rechercher `123` établit la correspondance avec toutes les instances dont un ou plusieurs attributs contiennent 123 (par exemple, dans une adresse IP, un ID d'instance, un ID de VPC ou un ID d'AMI). Si votre recherche de texte libre renvoie des correspondances inattendues, appliquez des filtres supplémentaires.

Recherche par attribut

La recherche par attribut vous permet de rechercher un attribut spécifique parmi toutes vos ressources.

> **Note**
>
> Les recherches par attribut utilisent le filtrage d'API ou le filtrage client, en fonction de l'attribut sélectionné. Lors d'une recherche par attribut, les attributs sont regroupés en conséquence.

Par exemple, vous pouvez rechercher l'attribut État de l'instance pour toutes vos instances afin de renvoyer uniquement les instances dont l'état est `stopped`. Pour cela :

1. Dans le champ de recherche de l'écran Instances, commencez à saisir `Instance state`. Lorsque vous entrez des caractères, une liste d'attributs correspondants s'affiche.

2. Sélectionnez État de l'instance dans la liste. Une liste des valeurs possibles pour l'attribut sélectionné s'affiche.
3. Sélectionnez Stopped (Arrêté) dans la liste.

Vous pouvez utiliser les techniques suivantes pour améliorer ou affiner vos recherches.

Recherche inversée

Les recherches inverses vous permettent de rechercher des ressources qui ne correspondent pas à une valeur spécifiée. Les recherches inverses s'effectuent en préfixant le mot clé de recherche d'un point d'exclamation (!). Par exemple, pour répertorier toutes les instances qui ne sont pas affectées au groupe de sécurité nommé `launch-wizard-1`, effectuez une recherche via l'attribut Security group name (Nom du groupe de sécurité) et entrez le mot-clé `!launch-wizard-1`.

> **Note**
>
> La recherche inverse est prise en charge avec des recherches par mot-clé et des recherches par attribut uniquement sur des filtres client. Elle n'est pas prise en charge avec des recherches par attribut sur les filtres d'API.

Recherche partielle

Avec les recherches partielles, vous pouvez rechercher des valeurs de chaîne partielles. Pour effectuer une recherche partielle, entrez uniquement une partie du mot-clé que vous souhaitez rechercher. Par exemple, pour rechercher toutes les instances `t2.micro`, `t2.small` et `t2.medium`, effectuez une recherche par l'attribut Instance Type (Type d'instance) puis saisissez le mot-clé `t2`.

> **Note**
>
> La recherche partielle est prise en charge avec les recherches par mot-clé et les recherches par attribut sur les filtres client uniquement. Elle n'est pas prise en charge avec des recherches par attribut sur les filtres d'API.

Recherche d'expression régulière

Pour utiliser les recherches d'expression régulière, vous devez activer Use regular expression matching (Utiliser la correspondance d'expression régulière) dans les préférences.

Les expressions régulières sont utiles quand vous avez besoin de faire correspondre les valeurs d'un champ à un modèle spécifique. Par exemple, pour rechercher une valeur qui commence par `s`, recherchez `^s`. Pour rechercher une valeur qui se termine par `xyz`, recherchez `xyz$`. Pour rechercher une valeur commençant par un nombre suivi d'un ou de plusieurs caractères, recherchez `[0-9]+.*`. La recherche par expression régulière n'est pas sensible à la casse.

> **Note**
>
> La recherche par expression régulière est prise en charge avec les recherches par mot-clé et les recherches par attribut uniquement sur les filtres client. Elle n'est pas prise en charge avec des recherches par attribut sur les filtres d'API.

Recherche par caractère générique

Utilisez le caractère générique `*` pour faire correspondre zéro ou plusieurs caractères. Utilisez le caractère générique `?` pour faire correspondre zéro ou un caractère. Par exemple, si vous disposez d'un ensemble de données avec les valeurs suivantes : prod, prods et production ; « `prod*` » correspond à toutes les valeurs, tandis que « `prod?` » correspond uniquement à `prod` et `prods`. Pour utiliser les valeurs littérales, échappez-les avec une barre oblique inverse (\). Par exemple, « `prod\*` » correspondrait à prod*.

> **Note**
>
> La recherche par caractère générique est prise en charge avec les recherches par attribut uniquement sur les filtres d'API. Elle n'est pas prise en charge avec les recherches par mot-clé et les recherches par attribut uniquement sur les filtres client.

Combinaison de recherches

En général, plusieurs filtres avec le même attribut sont automatiquement joints avec `OR`. Par exemple, la recherche `Instance State : Running` et `Instance State : Stopped` renvoie toutes les instances en cours d'exécution OU arrêtées. Pour joindre la recherche avec `AND`, recherchez sur différents attributs. Par exemple, la recherche `Instance State : Running` et `Instance Type : c4.large` renvoie uniquement les instances de type `c4.large` ET à l'état arrêté.

Ancienne console

L'ancienne console Amazon EC2 prend en charge les types de recherche suivants :

Recherche par mot-clé

La recherche par mot-clé est une recherche de texte libre qui vous permet de rechercher une valeur parmi tous les attributs de vos ressources. Pour rechercher par mot-clé, entrez ou collez ce que vous recherchez dans la zone de recherche, puis choisissez Enter (Entrer). Par exemple, rechercher `123` établit la correspondance avec toutes les instances dont un ou plusieurs attributs contiennent 123 (par exemple, dans une adresse IP, un ID d'instance, un ID de VPC ou un ID d'AMI). Si votre recherche de texte libre renvoie des correspondances inattendues, appliquez des filtres supplémentaires.

Recherche par attribut

La recherche par attribut vous permet de rechercher un attribut spécifique parmi toutes vos ressources. Par exemple, vous pouvez rechercher l'attribut State (État) pour toutes vos instances de manière à renvoyer uniquement les instances dont l'état est `stopped`. Pour cela :

1. Dans le champ de recherche de l'écran Instances, commencez à saisir `Instance State`. Lorsque vous entrez des caractères, une liste d'attributs correspondants s'affiche.
2. Sélectionnez Instance State dans la liste. Une liste des valeurs possibles pour l'attribut sélectionné s'affiche.
3. Sélectionnez Stopped (Arrêté) dans la liste.

Vous pouvez utiliser les techniques suivantes pour améliorer ou affiner vos recherches.

Recherche inversée

Les recherches inverses vous permettent de rechercher des ressources qui ne correspondent pas à une valeur spécifiée. Les recherches inverses s'effectuent en préfixant le mot clé de recherche d'un point d'exclamation (!). Par exemple, pour répertorier toutes les instances qui ne sont pas terminées, effectuez une recherche par l'attribut Instance State (État de l'instance) et entrez le mot clé `!Terminated`.

Recherche partielle

Avec les recherches partielles, vous pouvez rechercher des valeurs de chaîne partielles. Pour effectuer une recherche partielle, entrez uniquement une partie du mot-clé que vous souhaitez rechercher. Par exemple, pour rechercher toutes les instances `t2.micro`, `t2.small` et `t2.medium`, effectuez une recherche par l'attribut Instance Type (Type d'instance) puis saisissez le mot-clé `t2`.

Recherche d'expression régulière

Les expressions régulières sont utiles quand vous avez besoin de faire correspondre les valeurs d'un champ à un modèle spécifique. Par exemple, pour rechercher toutes les instances dont la valeur d'attribut commence par `s`, recherchez `^s`. Ou pour rechercher toutes les instances qui ont une valeur d'attribut qui se termine par `xyz`, recherchez `xyz$`. La recherche par expression régulière n'est pas sensible à la casse.

Combinaison de recherches

En général, plusieurs filtres avec le même attribut sont automatiquement joints avec `OR`. Par exemple, la recherche `Instance State : Running` et `Instance State : Stopped` renvoie toutes les

instances en cours d'exécution OU arrêtées. Pour joindre la recherche avec `AND`, recherchez sur différents attributs. Par exemple, la recherche `Instance State : Running` et `Instance Type : c4.large` renvoie uniquement les instances de type `c4.large` ET à l'état arrêté.

Pour filtrer une liste de ressources

1. Dans le panneau de navigation, sélectionnez un type de ressource (par exemple, Instances).
2. Choisissez le champ de recherche.
3. Choisissez le filtre dans la liste.
4. Spécifiez une valeur de filtre.
5. Une fois que vous avez terminé, retirez le filtre.

Affichage et filtrage à l'aide de l'interface de commande et des API

Chaque type de ressource possède une commande d'interface de ligne de commande ou une action d'API correspondante que vous utilisez pour afficher les ressources de ce type. Les listes de ressources qui en résultent peuvent être longues, de sorte qu'il peut être plus rapide et plus utile de filtrer les résultats pour inclure uniquement les ressources qui répondent à des critères spécifiques.

Considérations relatives au filtrage

- Vous pouvez spécifier plusieurs filtres et plusieurs valeurs de filtre dans une seule requête.
- Vous pouvez aussi utiliser des caractères génériques avec les valeurs de filtre. Un astérisque (*) correspond à zéro ou plusieurs caractères, et un point d'interrogation (?) correspond à zéro ou un caractère.
- Les valeurs de filtre sont sensibles à la casse.
- Votre recherche peut inclure les valeurs littérales des caractères génériques ; vous devez simplement leur associer une séquence d'échappement avec une barre oblique inverse devant le caractère. Par exemple, la valeur `\*amazon\?\\` recherche la chaîne littérale `*amazon?\`.

Filtres pris en charge

Pour découvrir les filtres pris en charge pour chaque ressource Amazon EC2, consultez la documentation suivante :

- AWS CLI : commandes `describe` dans le AWS CLI Command ReferenceAmazon EC2.
- Outils pour Windows PowerShell : commandes `Get` dans le Outils AWS pour PowerShell Cmdlet ReferenceAmazon EC2.
- API de requête : actions d'API `Describe` dans le document Amazon EC2 API Reference.

Example Exemple : spécifier un filtre unique

Vous pouvez lister vos instances Amazon EC2 à l'aide de la commande describe-instances. Sans aucun filtre, la réponse contient les informations pour toutes vos ressources. Vous pouvez utiliser la commande suivante pour inclure uniquement les instances en cours d'exécution dans votre sortie.

```
aws ec2 describe-instances --filters Name=instance-state-name,Values=running
```

Pour répertorier uniquement les ID des instances en cours d'exécution, ajoutez le paramètre `--query` comme suit.

```
aws ec2 describe-instances --filters Name=instance-state-name,Values=running --query
 "Reservations[*].Instances[*].InstanceId" --output text
```

Voici un exemple de sortie.

```
i-0ef1f57f78d4775a4
i-0626d4edd54f1286d
i-04a636d18e83cfacb
```

Example Exemple : spécifier plusieurs filtres ou valeurs de filtre

Si vous spécifiez plusieurs filtres ou plusieurs valeurs de filtre, la ressource doit correspondre à tous les filtres pour pouvoir apparaître dans les résultats.

Vous pouvez utiliser la commande suivante pour répertorier toutes les instances dont le type est `m5.large` ou `m5d.large`.

```
aws ec2 describe-instances --filters Name=instance-type,Values=m5.large,m5d.large
```

Vous pouvez utiliser la commande suivante pour répertorier toutes les instances arrêtées dont le type est `t2.micro`.

```
aws ec2 describe-instances --filters Name=instance-state-name,Values=stopped Name=instance-
type,Values=t2.micro
```

Example Exemple : utiliser des caractères génériques dans une valeur de filtre

Si vous spécifiez `database` comme valeur de filtre pour le filtre `description` lors de la description des instantanés EBS via describe-snapshots, la commande renvoie uniquement les instantanés dont la description correspond à « database ».

```
aws ec2 describe-snapshots --filters Name=description,Values=database
```

Le caractère générique * correspond à zéro ou plusieurs caractères. Si vous spécifiez `*database*` comme valeur de filtre, la commande renvoie uniquement les instantanés dont la description inclut ce terme.

```
aws ec2 describe-snapshots --filters Name=description,Values=*database*
```

Le caractère générique ? correspond à 1 seul caractère. Si vous spécifiez `database?` comme valeur de filtre, la commande renvoie uniquement les instantanés dont la description correspond à « database » ou à ce terme, suivi d'un caractère.

```
aws ec2 describe-snapshots --filters Name=description,Values=database?
```

Si vous indiquez `database????`, la commande renvoie uniquement les instantanés dont la description correspond à « database », suivi d'un maximum de quatre caractères. Elle exclut les descriptions contenant le terme « database » suivi de cinq caractères ou plus.

```
aws ec2 describe-snapshots --filters Name=description,Values=database????
```

Example Exemple : filtre basé sur la date

Avec l'AWS CLI, vous pouvez utiliser JMesPath pour filtrer les résultats à l'aide d'expressions. Par exemple, la commande describe-snapshots suivante affiche les ID de tous les instantanés créés par votre

compte AWS (représenté par *123456789012*) avant la date spécifiée (représentée par *2020-03-31*). Si vous ne spécifiez pas le propriétaire, les résultats incluent tous les instantanés publics.

```
aws ec2 describe-snapshots --filters Name=owner-id,Values=123456789012 --query "Snapshots[?
(StartTime<=`2020-03-31`)].[SnapshotId]" --output text
```

La commande suivante affiche les ID de tous les instantanés créés dans la plage de dates spécifiée.

```
aws ec2 describe-snapshots --filters Name=owner-id,Values=123456789012 --query "Snapshots[?
(StartTime>=`2019-01-01`) && (StartTime<=`2019-12-31`)].[SnapshotId]" --output text
```

Filtre basé sur les balises

Pour obtenir des exemples de filtrage d'une liste de ressources en fonction de leurs balises, consultez Utilisation des balises à l'aide de la ligne de commande (p. 1318).

Balisage de vos ressources Amazon EC2

Pour vous aider à gérer vos instances, images et autres ressources Amazon EC2, vous pouvez affectez vos propres métadonnées sous la forme de balises. Les balises vous permettent de classer vos ressources AWS de différentes manières, par exemple, par objectif, par propriétaire ou par environnement. Cette approche est utile lorsque vous avez de nombreuses ressources de même type. Elle vous permet d'identifier rapidement une ressource spécifique en fonction des balises que vous lui avez attribuées. Cette rubrique décrit les balises et vous montre comment les créer.

> **Warning**
>
> Les clés de balise et leurs valeurs sont renvoyées par différents appels d'API. Le fait de refuser l'accès à `DescribeTags` ne refuse pas automatiquement l'accès aux balises renvoyées par d'autres API. Nous vous recommandons de ne pas inclure de données sensibles dans vos balises.

Sommaire

- Principes de base des balises (p. 1309)
- Balisage de vos ressources (p. 1310)
- Restrictions liées aux balises (p. 1313)
- Balisage de vos ressources pour facturation (p. 1314)
- Gestion des balises à l'aide de la console (p. 1314)
- Utilisation des balises à l'aide de la ligne de commande (p. 1318)
- Ajout de balises à une ressource à l'aide de CloudFormation (p. 1321)

Principes de base des balises

Une balise est une étiquette que vous affectez à une ressource AWS. Chaque balise est constituée d'une clé et d'une valeur facultative que vous définissez.

Les balises vous permettent de classer vos ressources AWS de différentes manières, par exemple, par objectif, par propriétaire ou par environnement. Par exemple, vous pouvez définir pour les instances de votre compte Amazon EC2 un ensemble de balises qui vous aide à suivre le propriétaire et le niveau de stack de chaque instance.

Le graphique suivant illustre le fonctionnement du balisage. Dans cet exemple, vous avez affecté deux balises à chacune de vos instances : une balise avec la clé `Owner` et une autre avec la clé `Stack`. Chaque balise possède également une valeur associée.

Nous vous recommandons de concevoir un ensemble de clés de balise répondant à vos besoins pour chaque type de ressource. L'utilisation d'un ensemble de clés de balise cohérent facilite la gestion de vos ressources. Vous pouvez rechercher et filtrer les ressources en fonction des balises que vous ajoutez. Pour plus d'informations sur la mise en œuvre d'une stratégie efficace de balisage des ressources, consultez le livre blanc AWS sur les bonnes pratiques en matière de balisage.

Les balises n'ont pas de signification sémantique pour Amazon EC2 et sont interprétées strictement comme des chaînes de caractères. De plus, les balises ne sont pas automatiquement affectées à vos ressources. Vous pouvez modifier les clés et valeurs de balise, et vous pouvez retirer des balises d'une ressource à tout moment. Vous pouvez définir la valeur d'une balise sur une chaîne vide, mais vous ne pouvez pas définir la valeur d'une balise sur null. Si vous ajoutez une balise ayant la même clé qu'une balise existante sur cette ressource, la nouvelle valeur remplace l'ancienne valeur. Si vous supprimez une ressource, les balises associées à celle-ci sont également supprimées.

Vous pouvez gérer les balises à l'aide de la AWS Management Console, de l'AWS CLI et de l'API Amazon EC2.

Si vous utilisez AWS Identity and Access Management (IAM), vous pouvez contrôler quels utilisateurs de votre compte AWS sont autorisés à créer, modifier ou supprimer des balises. Pour de plus amples informations, veuillez consulter Identity and Access Management pour Amazon EC2 (p. 976).

Balisage de vos ressources

Vous pouvez attribuer des balises à la plupart des ressources Amazon EC2 qui existent déjà dans votre compte. Le tableau (p. 1311) ci-dessous répertorie les ressources qui prennent en charge le balisage.

Si vous utilisez la console Amazon EC2, vous pouvez appliquer des balises aux ressources à l'aide de l'onglet Balises sur l'écran de ressource concerné, ou vous pouvez utiliser l'écran Balises. Certains écrans de ressource vous permettent de spécifier des balises pour une ressource lors de la création de cette

ressource ; par exemple, une balise avec une clé de `Name` et une valeur que vous indiquez. Dans la plupart des cas, la console applique les balises immédiatement après la création de la ressource (plutôt qu'au cours de la création de ressources). La console peut organiser des ressources en fonction de la balise `Name`, mais cette balise n'a pas de signification sémantique pour le service Amazon EC2.

Si vous utilisez l'API Amazon EC2, l'AWS CLI ou un kit SDK AWS, vous pouvez utiliser l'action d'API EC2 `CreateTags` pour appliquer des balises aux ressources existantes. En outre, certaines actions de création de ressources vous permettent de spécifier des balises pour une ressource lors de la création de cette dernière. Si les balises ne peuvent pas être appliquées au cours de la création de ressources, nous restaurons le processus de création de ressources. Cela permet de s'assurer que les ressources sont créées avec des balises ou qu'elles ne sont pas créées du tout, et qu'aucune ressource ne demeurent sans balise à tout moment. En attribuant des balises aux ressources au moment de la création, vous pouvez supprimer la nécessité d'exécuter des scriptes de balisage personnalisés après la création de ressources.

Le tableau suivant décrit les ressources Amazon EC2 qui peuvent être balisées, et les ressources qui peuvent être balisées lors de la création à l'aide de l'API Amazon EC2, de l'AWS CLI ou d'un kit SDK AWS.

Prise en charge du balisage pour les ressources Amazon EC2

Ressource	Prend en charge les balises	Prend en charge le balisage au moment de la création
AFI	Oui	Oui
AMI	Oui	Non
Tâche de bundle	Non	Non
Réservation de capacité	Oui	Oui
Passerelle transporteur	Oui	Oui
Point de terminaison VPN Client	Oui	Oui
Route VPN Client	Non	Non
Passerelle client	Oui	Oui
Hôte dédié	Oui	Oui
Réservation Hôte dédié	Oui	Oui
Option DHCP	Oui	Oui
Instantané EBS	Oui	Oui
Volume EBS	Oui	Oui
Flotte EC2	Oui	Oui
Passerelle Internet de sortie uniquement	Oui	Oui
Adresse IP Elastic	Oui	Non
Accélérateur Elastic Graphics	Oui	Non
Instance	Oui	Oui
Volume de stockage d'instance	S/O	S/O
Passerelle Internet	Oui	Oui

Ressource	Prend en charge les balises	Prend en charge le balisage au moment de la création
Pool d'adresses IP (BYOIP)	Oui	Oui
Paire de clés	Oui	Oui
Modèle de lancement	Oui	Oui
Version du modèle de lancement	Non	Non
Passerelle locale	Oui	Non
Table de routage de passerelle locale	Oui	Non
Interface virtuelle de passerelle locale	Oui	Non
Groupe d'interface virtuelle de passerelle locale	Oui	Non
Association de VPC de table de routage de passerelle locale	Oui	Oui
Association de groupe d'interface virtuelle de table de routage de passerelle locale	Oui	Non
Passerelle NAT	Oui	Oui
ACL réseau	Oui	Oui
Interface réseau	Oui	Oui
Groupe de placement	Oui	Oui
Listes de préfixes	Oui	Oui
Instance réservée	Oui	Non
Liste d'entités d'Instance réservée	Non	Non
Table de routage	Oui	Oui
Demande de parc d'instances Spot	Oui	Oui
Demande d'instance Spot	Oui	Oui
Groupe de sécurité	Oui	Oui
Sous-réseau	Oui	Oui
Filtre Traffic Mirror	Oui	Oui
Session Traffic Mirror	Oui	Oui
Cible Traffic Mirror	Oui	Oui
Passerelle de transit	Oui	Oui

Ressource	Prend en charge les balises	Prend en charge le balisage au moment de la création
Table de routage de passerelle de transit	Oui	Oui
Attachement de VPC de passerelle de transit	Oui	Oui
Passerelle réseau privé virtuel	Oui	Oui
VPC	Oui	Oui
Point de terminaison d'un VPC	Oui	Oui
Service de point de terminaison d'un VPC	Oui	Oui
Configuration de service de point de terminaison de VPC	Oui	Oui
Journal de flux VPC	Oui	Oui
Connexion d'appairage de VPC	Oui	Oui
Connexion VPN	Oui	Oui

Vous pouvez baliser les instances et volumes au moment de leur création en utilisant l'assistant Lancer des instances Amazon EC2 dans la console Amazon EC2. Vous pouvez baliser vos volumes EBS au moment de leur création en utilisant l'écran des volumes, ou les instantanés EBS dans l'écran d'instantanés. Vous pouvez également utiliser les API Amazon EC2 de création de ressources (par exemple, RunInstances) pour appliquer des balises lors de la création de votre ressource.

Vous pouvez appliquer des autorisations au niveau des ressources basées sur des balises dans vos stratégies IAM aux actions d'API Amazon EC2 qui prennent en charge le balisage à la création, afin de mettre en œuvre un contrôle détaillé des utilisateurs et des groupes qui peuvent baliser des ressources à leur création. Vos ressources sont correctement sécurisées depuis la création. Les balises sont appliquées immédiatement à vos ressources. Les autorisations de niveau ressource basées sur des balises sont donc effectives immédiatement. Vos ressources peuvent être suivies et signalées avec plus de précision. Vous pouvez appliquer l'utilisation du balisage sur les nouvelles ressources et contrôler que les clés et valeurs de balise sont définies sur vos ressources.

Vous pouvez également appliquer des autorisations au niveau des ressources pour les actions d'API Amazon EC2 `CreateTags` et `DeleteTags` dans vos stratégies IAM afin de contrôler les clés et valeurs de balise définies sur vos ressources existantes. Pour plus d'informations, consultez Exemple : Balisage de vos ressources (p. 1018).

Pour plus d'informations sur le balisage de vos ressources pour la facturation, consultez Utilisation des balises de répartition des coûts du AWS Billing and Cost Management Guide de l'utilisateur.

Restrictions liées aux balises

Les restrictions de base suivantes s'appliquent aux balises :

- Nombre maximal de balises par ressource – 50
- Pour chaque ressource, chaque clé de balise doit être unique, et chaque clé de balise peut avoir une seule valeur.
- Longueur de clé maximale – 128 caractères Unicode en UTF-8

- Longueur de valeur maximale – 256 caractères Unicode en UTF-8
- EC2 permet d'utiliser n'importe quel caractère dans ses balises ; d'autres services en revanche sont plus restrictifs. Les caractères autorisés pour les services sont les lettres, les chiffres et les espaces représentables en UTF-8, ainsi que les caractères suivants : + - = . _ : / @.
- Les clés et valeurs de balise sont sensibles à la casse.
- Le préfixe `aws:` est réservé à l'utilisation d'AWS. Lorsque la balise possède une clé de balise avec ce préfixe, vous ne pouvez pas modifier ou supprimer sa clé ou sa valeur. Les balises avec le préfixe `aws:` ne sont pas comptabilisées comme vos balises pour la limite de ressources.

Vous ne pouvez pas mettre fin à une ressource, ou l'arrêter ou la supprimer uniquement en fonction de ses balises ; vous devez spécifier l'identificateur de ressource. Par exemple, pour supprimer des instantanés (snapshot) que vous avez balisés avec une clé de balise appelée `DeleteMe`, vous devez utiliser l'action `DeleteSnapshots` avec les identificateurs de ressource des instantanés, tels que `snap-1234567890abcdef0`.

Vous pouvez baliser des ressources publiques ou partagées, mais les balises que vous affectez sont disponibles uniquement pour votre compte AWS, pas pour les autres comptes partageant la ressource.

Vous ne pouvez pas attribuer des balises à toutes les ressources. Pour plus d'informations, consultez Prise en charge du balisage pour les ressources Amazon EC2 (p. 1311).

Balisage de vos ressources pour facturation

Vous pouvez utiliser des balises pour organiser votre facture AWS afin de refléter votre propre structure de coût. Pour ce faire, inscrivez-vous pour obtenir votre facture de compte AWS avec les valeurs de clé de balise incluses. Pour plus d'informations sur la configuration d'un rapport de répartition des coûts avec des balises, consultez la section relative au rapport de répartition des coûts mensuel dans le AWS Billing and Cost Management Guide de l'utilisateur. Pour voir le coût de vos ressources combinées, vous pouvez organiser vos informations de facturation en fonction des ressources possédant les mêmes valeurs de clé de balise. Par exemple, vous pouvez baliser plusieurs ressources avec un nom d'application spécifique, puis organiser vos informations de facturation pour afficher le coût total de cette application dans plusieurs services. Pour plus d'informations, consultez Utilisation des balises de répartition des coûts dans le AWS Billing and Cost Management Guide de l'utilisateur.

> **Note**
>
> Si vous venez d'activer la création de rapports, les données du mois en cours peuvent être consultées après 24 heures.

Les balises de répartition des coûts peuvent indiquer quelles ressources contribuent aux coûts, mais la suppression ou la désactivation des ressources ne réduit pas toujours les coûts. Par exemple, des données d'instantané qui sont référencées par un autre instantané sont conservées, même si l'instantané qui contient les données d'origine est supprimé. Pour plus d'informations, consultez Volumes et instantanés Amazon Elastic Block Store dans le AWS Billing and Cost Management Guide de l'utilisateur.

> **Note**
>
> Les adresses IP Elastic étiquetées ne sont pas affichées dans votre rapport de répartition des coûts.

Gestion des balises à l'aide de la console

À l'aide de la console Amazon EC2, vous pouvez voir quelles balises sont utilisées dans toutes vos ressources Amazon EC2 au sein de la même région. Vous pouvez afficher les balises par ressource et par

type de ressource, et voir également combien d'éléments de chaque type de ressource sont associés à une balise spécifiée. Vous pouvez également utiliser la console Amazon EC2 afin d'appliquer ou supprimer des balises pour une ou plusieurs ressources à la fois.

Pour plus d'informations sur l'utilisation de filtres pour répertorier vos ressources, consultez Liste et filtrage de vos ressources (p. 1303).

A des fins de facilité d'utilisation et pour obtenir de meilleurs résultats, utilisez Tag Editor dans la AWS Management Console, qui offre une façon centrale et unifiée de créer et gérer vos balises. Pour plus d'informations, consultez Utilisation de Tag Editor dans le Mise en route avec AWS Management Console.

Tâches

- Affichage des balises (p. 1315)
- Ajout et suppression de balises sur une ressource individuelle (p. 1316)
- Ajout et suppression de balises pour un groupe de ressources (p. 1317)
- Ajout d'une balise lorsque vous lancez une instance (p. 1317)
- Filtrage d'une liste de ressources par balise (p. 1318)

Affichage des balises

Vous pouvez afficher des balises de deux manières différentes sur la console Amazon EC2 : afficher les balises pour une ressource individuelle ou pour toutes les ressources.

Affichage des balises pour des ressources individuelles

Lorsque vous sélectionnez une page spécifique d'une ressource sur la console Amazon EC2, celle-ci affiche une liste de ces ressources. Par exemple, si vous sélectionnez Instances dans le panneau de navigation, la console affiche une liste d'instances Amazon EC2. Lorsque vous sélectionnez une ressource dans l'une de ces listes (par exemple, un instance), si la ressource prend en charge les balises, vous pouvez afficher et gérer ses balises. Sur la plupart des pages de ressources, vous pouvez afficher les balises en sélectionnant l'onglet Balises.

Vous pouvez ajouter à la liste des ressources une colonne affichant toutes les valeurs pour les balises avec la même clé. Cette colonne vous permet de trier et filtrer la liste des ressources par balise. Vous pouvez ajouter à la liste des ressources une nouvelle colonne pour afficher vos balises de deux manières.

- Dans l'onglet Balises, sélectionnez Afficher la colonne. Une nouvelle colonne est alors ajoutée à la console.
- Choisissez l'icône en forme d'engrenage Afficher / Masquer les colonnes puis, dans la boîte de dialogue Afficher / Masquer les colonnes, sélectionnez la clé de balise sous Vos clés de balise.

Affichage des balises pour toutes les ressources

Vous pouvez afficher les balises de toutes les ressources en sélectionnant Balises à partir du panneau de navigation de la console Amazon EC2. L'image suivante montre le volet Balises qui répertorie toutes les balises utilisées par type de ressource.

Ajout et suppression de balises sur une ressource individuelle

Vous pouvez gérer les balises pour une ressource individuelle directement à partir de la page de la ressource.

Pour ajouter une balise à une ressource individuelle

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. À partir de la barre de navigation, sélectionnez la région répondant à vos besoins. Ce choix est important car certaines ressources Amazon EC2 peuvent être partagées entre des régions, contrairement à d'autres ressources. Pour de plus amples informations, veuillez consulter Emplacements des ressources (p. 1301).
3. Dans le panneau de navigation, sélectionnez un type de ressource (par exemple, Instances).
4. Sélectionnez la ressource à partir de la liste des ressources et choisissez l'onglet Balises.
5. Choisissez Gérer les balises, Ajouter une balise. Entrez la clé et la valeur de la balise. Lorsque vous avez terminé d'ajouter des balises, choisissez Enregistrer.

Pour supprimer une balise d'une ressource individuelle

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. À partir de la barre de navigation, sélectionnez la région répondant à vos besoins. Ce choix est important car certaines ressources Amazon EC2 peuvent être partagées entre des régions, contrairement à d'autres ressources. Pour de plus amples informations, veuillez consulter Emplacements des ressources (p. 1301).
3. Dans le panneau de navigation, sélectionnez un type de ressource (par exemple, Instances).
4. Sélectionnez la ressource à partir de la liste des ressources et choisissez l'onglet Balises.
5. Choisissez Manage tags (Gérer les balises). Pour chaque balise, choisissez Supprimer. Lorsque vous avez terminé de supprimer des balises, choisissez Enregistrer.

Ajout et suppression de balises pour un groupe de ressources

Pour ajouter une balise à un groupe de ressources

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. À partir de la barre de navigation, sélectionnez la région répondant à vos besoins. Ce choix est important car certaines ressources Amazon EC2 peuvent être partagées entre des régions, contrairement à d'autres ressources. Pour de plus amples informations, veuillez consulter Emplacements des ressources (p. 1301).
3. Dans le panneau de navigation, sélectionnez Tags.
4. En haut du volet de contenu, sélectionnez Gérer les balises.
5. Dans Filtre, sélectionnez le type de ressource (par exemple, les instances).
6. Dans la liste des ressources, activez la case à cocher en regard de chaque ressource.
7. Sous Ajouter une balise, entrez la clé et la valeur de la balise, puis choisissez Ajouter une balise.

   Note

   Si vous ajoutez une nouvelles balise ayant la même clé de balise qu'une balise existante, la nouvelle balise remplace l'ancienne balise.

Pour supprimer une balise d'un groupe de ressources

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. À partir de la barre de navigation, sélectionnez la région répondant à vos besoins. Ce choix est important car certaines ressources Amazon EC2 peuvent être partagées entre des régions, contrairement à d'autres ressources. Pour de plus amples informations, veuillez consulter Emplacements des ressources (p. 1301).
3. Dans le panneau de navigation, sélectionnez Balises, Gérer les balises.
4. Pour afficher les balises utilisées, sélectionnez l'icône en forme d'engrenage Afficher / Masquer les colonnes, puis, dans la boîte de dialogue Afficher / Masquer les colonnes, sélectionnez les clés de balise à afficher, puis Fermer.
5. Dans Filtre, sélectionnez le type de ressource (par exemple, les instances).
6. Dans la liste des ressources, activez la case à cocher en regard de chaque ressource.
7. Sous Supprimer la balise, entrez la clé de balise et choisissez Supprimer la balise.

Ajout d'une balise lorsque vous lancez une instance

Pour ajouter une balise à l'aide de l'assistant de lancement

1. À partir de la barre de navigation, sélectionnez la région souhaitée pour l'instance. Ce choix est important car certaines ressources Amazon EC2 peuvent être partagées entre des régions, contrairement à d'autres ressources. Sélectionnez la région qui répond à vos besoins. Pour de plus amples informations, veuillez consulter Emplacements des ressources (p. 1301).
2. Choisissez Launch Instances.
3. La page Sélection d'une Amazon Machine Image (AMI) affiche une liste de configurations de base appelées Amazon Machine Images (AMIs). Sélectionnez l'AMI à utiliser, puis Sélectionner. Pour de plus amples informations sur la sélection d'une AMI, veuillez consulter Rechercher une AMI.
4. Sur la page Configurer les détails de l'instance, configurez les paramètres d'instance requis, puis cliquez sélectionnez Next: Add Storage (Suivant : Ajouter le stockage).
5. Sur la page Ajouter le stockage, vous pouvez spécifier des volumes de stockage supplémentaires pour votre instance. Choisissez Next: Add Tags (Suivant : Ajouter des balises) une fois que vous avez terminé.

6. Sur la page Ajouter des balises, spécifiez des balises pour l'instance, les volumes ou les deux. Choisissez Ajouter une autre balise pour ajouter plusieurs balises à votre instance. Choisissez Suivant : Configurer le groupe de sécurité une fois que vous avez terminé.
7. Sur la page Configurer le groupe de sécurité, vous pouvez sélectionner un groupe de sécurité existant parmi ceux que vous possédez ou laisser l'assistant créer un groupe de sécurité pour vous. Choisissez Vérifier et lancer lorsque vous avez terminé.
8. Vérifiez vos paramètres. Lorsque vous êtes satisfait de vos sélections, sélectionnez Lancer. Sélectionnez une paire de clés existante ou créez-en une, cochez la case de confirmation, puis cliquez sur Lancer des instances.

Filtrage d'une liste de ressources par balise

Vous pouvez filtrer votre liste de ressources selon une ou plusieurs clés de balise et valeurs de balise.

Pour filtrer une liste de ressources par balise

1. Dans le panneau de navigation, sélectionnez un type de ressource (par exemple, Instances).
2. Choisissez le champ de recherche.
3. Choisissez la clé de balise dans la liste.
4. Choisissez la valeur de balise correspondante dans la liste.
5. Lorsque vous avez terminé, retirez le filtre.

Pour plus d'informations sur les filtres, consultez Liste et filtrage de vos ressources (p. 1303).

Utilisation des balises à l'aide de la ligne de commande

Vous pouvez ajouter des balises à de nombreuses ressources EC2 lors de leur création en utilisant le paramètre de spécifications de balise pour la commande créer. Vous pouvez afficher les balises d'une ressource à l'aide de la commande décrire de la ressource. Vous pouvez également ajouter, mettre à jour ou supprimer des balises pour vos ressources existantes à l'aide des commandes suivantes.

Tâche	AWS CLI	Outils AWS pour Windows PowerShell
Ajouter ou remplacer une ou plusieurs balises	create-tags	New-EC2Tag
Supprimer une ou plusieurs balises.	delete-tags	Remove-EC2Tag
Décrire une ou plusieurs balises.	describe-tags	Get-EC2Tag

Tâches

- Ajout de balises lors de la création de ressources (p. 1318)
- Ajout de balises à une ressource existante (p. 1319)
- Décrire les ressources balisées (p. 1320)

Ajout de balises lors de la création de ressources

Les exemples suivants montrent comment appliquer des balises lorsque vous créez des ressources.

La manière dont vous entrez des paramètres au format JSON sur la ligne de commande varie selon le système d'exploitation. Sous Linux, macOS ou Unix et Windows PowerShell, la structure de données JSON est placée entre guillemets simples ('). Omettez les guillemets simples lorsque vous utilisez les commandes depuis la ligne de commande Windows. Pour plus d'informations, veuillez consulter Spécification de valeurs de paramètre pour l'AWS Command Line Interface.

Example Exemple : Lancez une instance et appliquez des balises à l'instance et au volume

La commande run-instances lance une instance et applique une balise avec une clé **webserver** et une valeur de **production** à cette dernière. La commande applique également une balise avec une clé de **cost-center** et une valeur **cc123** à n'importe quel volume EBS qui est créé (dans ce cas, le volume racine).

```
aws ec2 run-instances \
    --image-id ami-abc12345 \
    --count 1 \
    --instance-type t2.micro \
    --key-name MyKeyPair \
    --subnet-id subnet-6e7f829e \
    --tag-specifications 'ResourceType=instance,Tags=[{Key=webserver,Value=production}]'
 'ResourceType=volume,Tags=[{Key=cost-center,Value=cc123}]'
```

Vous pouvez appliquer les mêmes clés et valeurs de balise aux instances et aux volumes pendant le lancement. La commande suivante lance une instance et applique une balise avec une clé de **cost-center** et une valeur de **cc123** à l'instance et à n'importe quel volume EBS qui est créé.

```
aws ec2 run-instances \
    --image-id ami-abc12345 \
    --count 1 \
    --instance-type t2.micro \
    --key-name MyKeyPair \
    --subnet-id subnet-6e7f829e \
    --tag-specifications 'ResourceType=instance,Tags=[{Key=cost-center,Value=cc123}]'
 'ResourceType=volume,Tags=[{Key=cost-center,Value=cc123}]'
```

Example Exemple : Créez un volume et appliquez une balise

La commande create-volume crée un volume et applique deux balises : **purpose=production**, et **cost-center=cc123**.

```
aws ec2 create-volume \
    --availability-zone us-east-1a \
    --volume-type gp2 \
    --size 80 \
    --tag-specifications 'ResourceType=volume,Tags=[{Key=purpose,Value=production},
{Key=cost-center,Value=cc123}]'
```

Ajout de balises à une ressource existante

Les exemples suivants montrent comment ajouter des balises à une ressource existante à l'aide de la commande create-tags.

Example Exemple : Ajout d'une balise à une ressource

La commande suivante ajoute la balise **Stack=production** à l'image spécifiée ou remplace une balise existante pour l'AMI où la clé de balise est **Stack**. Si la commande réussit, aucune sortie n'est renvoyée.

```
aws ec2 create-tags \
```

```
    --resources ami-78a54011 \
    --tags Key=Stack,Value=production
```

Example Exemple : Ajout de balises à plusieurs ressources

Cet exemple ajoute (ou remplace) deux balises pour une AMI et une instance. L'une des balises contient simplement une clé (**webserver**), sans valeur (nous avons défini une chaîne vide comme valeur). L'autre balise est constituée d'une clé (**stack**) et d'une valeur (**Production**). Si la commande réussit, aucune sortie n'est renvoyée.

```
aws ec2 create-tags \
    --resources ami-1a2b3c4d i-1234567890abcdef0 \
    --tags Key=webserver,Value=  Key=stack,Value=Production
```

Example Exemple : Ajout de balises avec des caractères spéciaux

Cet exemple ajoute la balise **[Group]=test** à une instance. Les crochets (**[** et **]**) sont des caractères spéciaux, qui doivent être échappés.

Si vous utilisez Linux ou OS X, pour échapper les caractères spéciaux, placez l'élément avec le caractère spécial entre des guillemets doubles (**"**), puis placez toute la structure de clé et de valeur entre des guillemets simples (**'**).

```
aws ec2 create-tags \
    --resources i-1234567890abcdef0 \
    --tags 'Key="[Group]",Value=test'
```

Si vous utilisez Windows, pour échapper les caractères spéciaux, placez l'élément qui a des caractères spéciaux entre des guillemets doubles ("), puis faites précéder chaque guillemet double d'une barre oblique inverse (****), comme suit :

```
aws ec2 create-tags ^
    --resources i-1234567890abcdef0 ^
    --tags Key=\"[Group]\",Value=test
```

Si vous utilisez Windows PowerShell, pour échapper les caractères spéciaux, placez la valeur qui a des caractères spéciaux entre des guillemets doubles (**"**), faites précéder chaque caractère de guillemets doubles d'une barre oblique inverse (****), puis placez toute la structure de clé et de valeur entre des guillemets simples (**'**), comme suit :

```
aws ec2 create-tags `
    --resources i-1234567890abcdef0 `
    --tags 'Key=\"[Group]\",Value=test'
```

Décrire les ressources balisées

Les exemples suivants montrent comment utiliser des filtres avec describe-instances pour afficher des instances avec des balises spécifiques. Toutes les commandes décrire EC2 utilisent cette syntaxe pour filtrer par balise sur un seul type de ressource. Vous pouvez également utiliser la commande describe-tags pour filtrer par balise sur les types de ressources EC2.

Example Exemple : Décrire les instances avec la clé de balise spécifiée

La commande suivante décrit les instances avec une balise **Stack**, quelle que soit la valeur de la balise.

```
aws ec2 describe-instances \
    --filters Name=tag-key,Values=Stack
```

Example Exemple : Décrire les instances avec la balise spécifiée

La commande suivante décrit les instances avec la balise **`Stack=production`**.

```
aws ec2 describe-instances \
    --filters Name=tag:Stack,Values=production
```

Example Exemple : Décrire les instances avec la valeur de balise spécifiée

La commande suivante décrit les instances à l'aide d'une balise avec la valeur **`production`**, quelle que soit la clé de balise.

```
aws ec2 describe-instances \
    --filters Name=tag-value,Values=production
```

Example Exemple : Décrire toutes les ressources EC2 avec la balise spécifiée

La commande suivante décrit toutes les ressources EC2 avec la balise **`Stack=Test`**.

```
aws ec2 describe-tags \
    --filters Name=key,Values=Stack Name=value,Values=Test
```

Ajout de balises à une ressource à l'aide de CloudFormation

Avec les types de ressource Amazon EC2, vous spécifiez des balises à l'aide d'une propriété `TagSpecifications` ou `Tags`.

Les exemples suivants ajoutent la balise **`Stack=Production`** à AWS::EC2::Instance en utilisant sa propriété `Tags`.

Example Exemple : Tags dans YAML

```
Tags:
  - Key: "Stack"
    Value: "Production"
```

Example Exemple : Tags dans JSON

```
"Tags": [
    {
        "Key": "Stack",
        "Value": "Production"
    }
]
```

Les exemples suivants ajoutent la balise **`Stack=Production`** à AWS::EC2::LaunchTemplate LaunchTemplateData en utilisant sa propriété `TagSpecifications`.

Example Exemple : TagSpecifications dans YAML

```
TagSpecifications:
  - ResourceType: "instance"
    Tags:
    - Key: "Stack"
      Value: "Production"
```

Example Exemple : TagSpecifications dans JSON

```
"TagSpecifications": [
    {
        "ResourceType": "instance",
        "Tags": [
            {
                "Key": "Stack",
                "Value": "Production"
            }
        ]
    }
]
```

Quotas de service Amazon EC2

Amazon EC2 fournit différentes ressources que vous pouvez utiliser. Ces ressources incluent des images, des instances, des volumes et des instantanés (snapshot). Lorsque vous créez votre compte AWS, nous définissons des quotas par défaut (également appelés limites) sur ces ressources en fonction des régions. Par exemple, il existe un nombre maximal d'instances que vous pouvez lancer dans une région. Par exemple, si vous lancez une instance dans la région USA Ouest (Oregon), la demande ne doit pas faire en sorte que votre utilisation dépasse votre nombre maximal d'instances dans cette région.

La console Amazon EC2 fournit des informations sur les limites pour les ressources gérées par les consoles Amazon EC2 et Amazon VPC. Vous pouvez demander une augmentation pour la plupart de ces limites. Utilisez les informations sur les limites que nous fournissons pour gérer votre infrastructure AWS. Prévoyez de demander les augmentations de limite avant le moment où vous en aurez besoin.

Pour de plus amples informations, veuillez consulter Points de terminaison et quotas Amazon EC2 dans le document Référence générale d'Amazon Web Services. Pour obtenir des informations sur les quotas Amazon EBS, veuillez consulter Quotas Amazon EBS (p. 1264).

Affichages de vos limites actuelles

Utilisez la page Limites de la console Amazon EC2 afin d'afficher les limites actuelles des ressources fournies par Amazon EC2 et Amazon VPC, région par région.

Pour afficher vos limites actuelles

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez une région.

3. Dans le panneau de navigation, cliquez sur Restrictions.
4. Recherchez la ressource dans la liste. Vous pouvez utiliser les champs de recherche pour filtrer la liste par nom de ressource ou groupe de ressource. La colonne Current limit (Limite actuelle) affiche la valeur maximale en cours pour la ressource pour votre compte.

Demande d'augmentation

Utilisez la page Limites de la console Amazon EC2 pour demander une augmentation de vos ressources Amazon EC2 ou Amazon VPC, par région.

Vous pouvez également demander une augmentation en utilisant Quotas de service. Pour de plus amples information, veuillez consulter Demande d'augmentation du quota dans le Guide de l'utilisateur Quotas de service.

Pour demander une augmentation à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans la barre de navigation, sélectionnez une région.
3. Dans le panneau de navigation, cliquez sur Restrictions.
4. Sélectionnez la ressource et choisissez Demander une augmentation de limite.
5. Remplissez les champs obligatoires du formulaire de demande d'augmentation de limite et choisissez Submit (Envoyer). Nous vous répondrons en utilisant la méthode de contact que vous avez spécifiée.

Limites relatives à des e-mails envoyés à l'aide du port 25

Amazon EC2 limite le trafic sur le port 25 de toutes les instances par défaut. Vous pouvez demander que cette restriction soit supprimée. Pour de plus amples informations, veuillez consulter Comment supprimer la restriction du port 25 à partir de mon instance EC2 ? dans le Centre de connaissances AWS.

Rapports d'utilisation d'Amazon EC2

AWS fournit un outil de rapport gratuit appelé AWS Cost Explorer qui vous permet d'analyser le coût et l'utilisation de vos instances EC2, ainsi que l'utilisation de vos Instances réservées. Vous pouvez afficher les données pour une période allant jusqu'aux 13 derniers mois et prévoir vos dépenses pour les trois prochains mois. Vous pouvez utiliser Cost Explorer pour afficher des schémas de vos dépenses en ressources AWS au fil du temps, identifier les domaines qui méritent d'être approfondis et connaître des tendances que vous pouvez utiliser pour comprendre vos coûts. Vous pouvez également spécifier des plages de temps pour les données et afficher des données temporelles par jour ou par mois.

Voici un exemple des questions auxquelles vous pouvez répondre en utilisant Cost Explorer :

- Quel est le montant de mes dépenses par type d'instance ?
- Combien d'heures d'instances ai-je utilisé par service ?
- Quelle est la répartition de l'utilisation des instances par zone de disponibilité ?
- Quelle est la répartition de l'utilisation des instances par compte AWS ?
- Est-ce que j'utilise mes Instances réservées de façon optimale ?
- Mes Instances réservées me permettent-elles d'économiser de l'argent ?

Pour de plus amples informations sur l'utilisation des rapports dans Cost Explorer, notamment l'enregistrement des rapports, veuillez consulter Analyse de vos coûts à l'aide de Cost Explorer.

Dépannage des instances EC2

La documentation suivante peut vous aider à résoudre les problèmes que vous pouvez rencontrer avec votre instance.

Table des matières

- Résolution des problèmes de lancement d'instance (p. 1325)
- Résolution de la connexion à votre instance (p. 1327)
- Résolution des problèmes d'arrêt de votre instance (p. 1335)
- Résolution des problèmes de mise hors service (d'arrêt) de votre instance (p. 1337)
- Résolution des instances avec des contrôles de statut échoués (p. 1338)
- Résolution d'un problème d'instance inaccessible (p. 1360)
- Démarrage à partir du mauvais volume (p. 1363)
- Utilisation d'EC2Rescue pour Linux (p. 1364)
- Envoi d'une interruption de diagnostic (utilisateurs avancés uniquement) (p. 1375)

Pour plus d'informations sur les instances Windows, consultez Résolution des problèmes liés aux instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Résolution des problèmes de lancement d'instance

Le problème suivant vous empêche de lancer une instance.

Problèmes de lancement

- Dépassement de la limite d'instance (p. 1325)
- Capacité d'instance insuffisante (p. 1326)
- Mise hors service immédiate de l'instance (p. 1326)

Dépassement de la limite d'instance

Description

Vous obtenez l'erreur `InstanceLimitExceeded` lorsque vous essayez de lancer une nouvelle instance ou de redémarrer une instance arrêtée.

Cause

Si vous obtenez une erreur `InstanceLimitExceeded` lorsque vous essayez de lancer une nouvelle instance ou de redémarrer une instance arrêté, vous avez atteint la limite du nombre d'instances que vous pouvez lancer dans une région. Lorsque vous créez votre compte AWS, nous définissons des limites sur nombre d'instances que vous pouvez exécuter en fonction des régions.

Solution

Vous pouvez demander une augmentation de la limite d'instance par région. Pour plus d'informations, consultez Quotas de service Amazon EC2 (p. 1322).

Capacité d'instance insuffisante

Description

Vous obtenez l'erreur `InsufficientInstanceCapacity` lorsque vous essayez de lancer une nouvelle instance ou de redémarrer une instance arrêtée.

Cause

Si vous obtenez une erreur `InsufficientInstanceCapacity` lorsque vous essayez de lancer une instance ou de redémarrer une instance arrêtée, AWS n'a actuellement pas assez de capacité à la demande disponible pour répondre à votre demande.

Solution

Pour résoudre ce problème, essayez ce qui suit :

- Attendez quelques minutes, puis renvoyez votre demande. La capacité peut changer fréquemment.
- Envoyez une nouvelle demande avec un nombre réduit d'instances. Par exemple, si vous faites une demande simple pour lancer 15 instances, essayez de faire 3 demandes pour 5 instances ou 15 demandes pour 1 instance à la place.
- Si vous lancez une instance, soumettez une nouvelle demande sans spécifier de zone de disponibilité.
- Si vous lancez une instance, envoyez une nouvelle demande en utilisant un type d'instance différent (que vous pouvez redimensionner à un stade ultérieur). Pour plus d'informations, consultez Modification du type d'instance (p. 313).
- Si vous lancez des instances dans un groupe de placement du cluster, vous pouvez recevoir une erreur de capacité insuffisante. Pour plus d'informations, consultez Règles et restrictions des groupes de placement (p. 927).
- Créez une Réservation de capacité à la demande, qui permet de réserver la capacité Amazon EC2 pour n'importe quelle durée. Pour plus d'informations, consultez Réservations de capacité à la demande (p. 504).
- Essayez d'acheter des instances réservées qui constituent une réservation de capacité à long terme. Pour plus d'informations, consultez Instances réservées Amazon EC2.

Mise hors service immédiate de l'instance

Description

Votre instance passe de l'état `pending` à l'état `terminated`.

Cause

Voici quelques raisons qui expliquent pourquoi une instance peut se terminer immédiatement :

- Vous avez dépassé vos limites de volumes EBS. Pour plus d'informations, consultez Limites de volume d'instance (p. 1286).
- Un instantané EBS est corrompu.
- Le volume EBS racine est chiffré et vous n'êtes pas autorisé à accéder à la clé CMK pour le déchiffrement.
- Un instantané spécifié dans le mappage de périphérique de stockage en mode bloc pour l'AMI est chiffré et vous ne disposez pas des autorisations nécessaires pour accéder à la clé CMK pour la déchiffrer, ou vous n'avez pas accès à la clé CMK pour chiffrer les volumes restaurés.

- L'AMI basée sur le stockage d'instance que vous avez utilisée pour lancer l'instance ne possède par une partie obligatoire (un fichier image.part.xx).

Pour de plus amples informations, veuillez récupérer le motif de résiliation à l'aide de l'une des méthodes suivantes.

Pour obtenir la cause de la résiliation à l'aide de la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez l'instance.
3. Dans le premier onglet, recherchez le motif en regard de State transition reason (Motif de transition de l'état).

Pour obtenir la cause de la résiliation à l'aide de l'AWS Command Line Interface

1. Utilisez la commande describe-instances et spécifiez l'ID de l'instance.

```
aws ec2 describe-instances --instance-id instance_id
```

2. Vérifiez la réponse JSON renvoyée par la commande et notez les valeurs de l'élément de réponse `StateReason`.

   Le bloc de code suivant présente un exemple d'élément de réponse `StateReason`.

```
"StateReason": {
  "Message": "Client.VolumeLimitExceeded: Volume limit exceeded",
  "Code": "Server.InternalError"
},
```

Pour obtenir la cause de la résiliation à l'aide de l'AWS CloudTrail

Pour de plus amples informations, veuillez consulter Affichage des événements avec l'historique des événements CloudTrail dans le AWS CloudTrail User Guide.

Solution

En fonction de la cause de la résiliation, exécutez l'une des actions suivantes :

- **`Client.VolumeLimitExceeded: Volume limit exceeded`** — Supprimez les volumes inutilisés. Vous pouvez envoyer une demande d'augmentation de votre limite de volumes.
- **`Client.InternalError: Client error on launch`** — Assurez-vous que vous disposez des autorisations requises pour accéder aux clés CMK utilisées pour déchiffrer et chiffrer des volumes. Pour de plus amples informations, veuillez consulter Utilisation de stratégies de clé dans AWS KMS dans le AWS Key Management Service Developer Guide.

Résolution de la connexion à votre instance

Les informations suivantes peuvent vous aider à résoudre les problèmes de connexion à votre instance. Pour plus d'informations sur les instances Windows, consultez Résolution des problèmes liés aux instances Windows dans le Amazon EC2 Guide de l'utilisateur pour les instances Windows.

Problèmes et erreurs de connexion

- Causes courantes des problèmes de connexion (p. 1328)
- Erreur de connexion à votre instance : connexion expirée (p. 1329)
- Erreur : unable to load key… Expecting: ANY PRIVATE KEY (p. 1331)
- Erreur : clé de l'utilisateur non reconnue par le serveur (p. 1331)
- Erreur : autorisation refusée ou connexion fermée par [instance] port 22 (p. 1332)
- Erreur : fichier de clé privée non protégé (p. 1334)
- Erreur : La clé privée doit commencer par « -----BEGIN RSA PRIVATE KEY----- » et se terminer par « -----END RSA PRIVATE KEY----- » (p. 1334)
- Erreur : le serveur a refusé notre clé ou Aucune méthode d'authentification prise en charge disponible (p. 1334)
- Impossible d'envoyer une commande ping à l'instance (p. 1335)
- Erreur : le serveur a fermé la connexion réseau de manière inopinée (p. 1335)

Causes courantes des problèmes de connexion

Nous vous recommandons de commencer le dépannage en vérifiant les causes courantes des problèmes de connexion à votre instance.

Vérifiez le nom d'utilisateur de votre instance

Vous pouvez vous connecter à votre instance à l'aide du nom d'utilisateur de votre compte d'utilisateur ou du nom d'utilisateur par défaut de l'AMI que vous avez utilisée pour lancer votre instance.

- Obtenez le nom d'utilisateur de votre compte d'utilisateur.

  Pour de plus amples informations sur la création d'un compte utilisateur, veuillez consulter Gestion des comptes d'utilisateur sur votre instance Amazon Linux (p. 660).
- Obtenir le nom d'utilisateur par défaut pour l'AMI que vous avez utilisée pour lancer votre instance
  - Pour Amazon Linux 2 ou l'AMI Amazon Linux, le nom d'utilisateur est `ec2-user`.
  - Pour une AMI CentOS, le nom d'utilisateur est `centos`.
  - Pour une AMI Debian, le nom d'utilisateur est `admin`.
  - Pour une AMI Fedora, le nom d'utilisateur est `ec2-user` ou `fedora`.
  - Pour une AMI RHEL, le nom d'utilisateur est `ec2-user` ou `root`.
  - Pour une AMI SUSE, le nom d'utilisateur est `ec2-user` ou `root`.
  - Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.
  - Si `ec2-user` et `root` ne fonctionnent pas, renseignez-vous auprès de votre fournisseur AMI.

Vérifiez que les règles de votre groupe de sécurité autorisent le trafic

Assurez-vous que vos règles de groupe de sécurité autorisent le trafic entrant à partir de votre adresse IPv4 publique sur le port approprié. Pour connaître les étapes à vérifier, veuillez consulter Erreur de connexion à votre instance : connexion expirée (p. 1329)

Vérifiez que votre instance est prête

Une fois l'instance lancée, il peut falloir quelques minutes pour qu'elle soit prête pour que vous puissiez vous y connecter. Vérifiez votre instance pour vous assurer qu'elle est en cours d'exécution et qu'elle a réussi ses vérifications d'état.

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, puis sélectionnez votre instance.
3. Vérifiez que votre instance est à l'état `running` et qu'elle a réussi les contrôles de statut.

Vérifiez les prérequis généraux pour la connexion à votre instance.

Pour plus d'informations, consultez Prérequis généraux pour se connecter à votre instance (p. 598).

Erreur de connexion à votre instance : connexion expirée

Si vous essayez de vous connecter à votre instance et vous obtenez un message d'erreur `Network error: Connection timed out` ou `Error connecting to [instance], reason: -> Connection timed out: connect`, essayez ce qui suit :

- Vérifiez les règles du groupe de sécurité. Vous avez besoin d'un groupe de sécurité qui permet le trafic entrant à partir de votre adresse IPv4 publique sur le même port.

  1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
  2. Dans le panneau de navigation, choisissez Instances, puis sélectionnez votre instance.
  3. Dans l'onglet Description en bas de la page de la console, en regard de Groupes de sécurité, sélectionnez afficher les règles pour afficher la liste des règles en vigueur pour l'instance sélectionnée.
  4. Pour les instances Linux : lorsque vous sélectionnez afficher les règles, une fenêtre contenant les ports sur lesquels le trafic est autorisé s'affiche. Vérifiez qu'il existe une règle qui autorise le trafic de votre ordinateur au port 22 (SSH).

     Pour les instances Windows : lorsque vous sélectionnez afficher les règles, une fenêtre contenant les ports sur lesquels le trafic est autorisé s'affiche. Vérifiez qu'il existe une règle qui autorise le trafic de votre ordinateur au port 3389 (RDP).

     Chaque fois que vous redémarrez une instance, une nouvelle adresse IP (ainsi qu'un nom d'hôte) lui est affectée. Si votre groupe de sécurité possède une règle qui permet le trafic entrant à partir d'une seule adresse IP, il se peut que cette adresse ne soit pas statique si votre ordinateur est sur un réseau d'entreprise ou si vous vous connectez via un fournisseur de services Internet (ISP). Au lieu de cela, spécifiez la plage d'adresses IP utilisées par les ordinateurs clients. Si votre groupe de sécurité ne possède pas de règle qui permet le trafic entrant comme ce qui est décrit dans l'étape précédente, ajoutez une règle à votre règle de sécurité. Pour plus d'informations, consultez Autorisation de l'accès réseau à vos instances (p. 1044).

     Pour plus d'informations sur les règles des groupes de sécurité, consultez la rubrique Règles des groupes de sécurité dans le Guide de l'utilisateur d'Amazon VPC.
- Vérifiez la table de routage pour le sous-réseau. Vous avez besoin d'un itinéraire qui envoie tout le trafic destiné à l'extérieur du VPC vers la passerelle Internet du VPC.

  1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
  2. Dans le panneau de navigation, choisissez Instances, puis sélectionnez votre instance.
  3. Dans l'onglet Description, écrivez les valeurs de ID de VPC et ID de sous-réseau.
  4. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
  5. Dans le panneau de navigation, choisissez Passerelles Internet. Vérifiez qu'il existe une passerelle Internet attachée à votre VPC. Sinon, choisissez Créer une passerelle Internet pour créer une passerelle Internet. Sélectionnez la passerelle Internet, puis choisissez Attacher au VPC et suivez les instructions pour l'attacher à votre VPC.
  6. Dans le panneau de navigation, sélectionnez Sous-réseaux, puis sélectionnez votre sous-réseau.
  7. Dans l'onglet Table de routage, vérifiez qu'il existe une route avec `0.0.0.0/0` comme destination et la passerelle Internet pour votre VPC comme cible. Si vous vous connectez à votre instance à l'aide

de son adresse IPv6, vérifiez qu'il existe une route pour tout le trafic IPv6 (`::/0`) qui pointe vers la passerelle Internet. Sinon, procédez comme suit :

a. Choisissez l'ID de la table de routage (rtb-xxxxxxxx) pour accéder à cette dernière.

b. Dans l'onglet Routes, choisissez Edit routes (Modifier les routes). Choisissez Add route (Ajouter une route) et utilisez `0.0.0.0/0` comme destination et la passerelle Internet comme cible. Pour IPv6, choisissez Add route (Ajouter une route) et utilisez `::/0` comme destination et la passerelle Internet comme cible.

c. Choisissez Save routes (Enregistrer les routes).

- Vérifiez la liste de contrôle d'accès (ACL) du réseau pour le sous-réseau. L'ACL réseau doit permettre le trafic entrant et sortant à partir de votre adresse IP locale sur le même port. L'ACL réseau par défaut permet tout le trafic entrant et sortant.

  1. Ouvrez la console Amazon VPC à l'adresse https://console.aws.amazon.com/vpc/.
  2. Dans le panneau de navigation, sélectionnez Sous-réseaux et sélectionnez votre sous-réseau.
  3. Dans l'onglet Description, recherchez ACL réseau et choisissez son ID (acl-xxxxxxxx).
  4. Sélectionnez l'ACL réseau. Pour Règles entrantes, vérifiez que les règles autorisent le trafic à partir de votre ordinateur. Sinon, supprimez ou modifiez la règle qui bloque le trafic à partir de votre ordinateur.
  5. Pour Règles sortantes, vérifiez que les règles autorisent le trafic vers votre ordinateur. Sinon, supprimez ou modifiez la règle qui bloque le trafic vers votre ordinateur.

- Si votre ordinateur est sur un réseau d'entreprise, demandez à votre administrateur de réseau si le pare-feu interne permet le trafic entrant et sortant à partir de votre ordinateur sur le port 22 (pour les instances Linux) ou le port 3389 (pour les instances Windows).

  Si vous avez un pare-feu sur votre ordinateur, vérifiez s'il permet le trafic entrant et sortant à partir de votre ordinateur sur le port 22 (pour les instances Linux) ou le port 3389 (pour les instances Windows).

- Vérifiez que votre instance possède une adresse IPv4 publique. Si non, vous pouvez associer une adresse IP Elastic à votre instance. Pour plus d'informations, consultez Adresses IP Elastic (p. 831).

- Vérifiez la charge de l'UC sur votre instance. Il se peut que le serveur soit surchargé. AWS fournit automatiquement des données, comme les métriques Amazon CloudWatch et le statut des instances, que vous pouvez utiliser pour voir quelle charge de l'UC se trouve sur votre instance et, si nécessaire, pour ajuster la gestion de vos charges. Pour plus d'informations, consultez Surveillance de vos instances avec CloudWatch (p. 761).

  - Si votre charge est variable, vous pouvez automatiquement effectuer des mises à l'échelle ascendantes et descendantes de vos instances en utilisant l'Auto Scaling et l'Elastic Load Balancing.
  - Si votre charge augmente régulièrement, vous pouvez passer à un type d'instance plus important. Pour plus d'informations, consultez Modification du type d'instance (p. 313).

Pour vous connecter à votre instance à l'aide d'une adresse IPv6, vérifiez les points suivants :

- Votre sous-réseau doit être associé à une table de routage ayant une route pour le trafic IPv6 (`::/0`) vers une passerelle Internet.
- Vos règles de groupe de sécurité doivent autoriser le trafic entrant à partir de votre adresse IPv6 locale sur le port approprié (22 pour Linux et 3389 pour Windows).
- Vos règles ACL réseau doivent autoriser le trafic IPv6 entrant et sortant.
- Si vous avez lancé votre instance à partir d'une AMI plus ancienne, elle n'est peut-être pas configurée pour DHCPv6 (les adresses IPv6 ne sont pas automatiquement reconnues sur l'interface réseau). Pour plus d'informations, consultez Configuration d'IPv6 sur vos instances dans le Amazon VPC Guide de l'utilisateur.
- Votre ordinateur local doit avoir une adresse IPv6 et doit être configuré pour utiliser IPv6.

Erreur : unable to load key… Expecting: ANY PRIVATE KEY

Si vous essayez de vous connecter à votre instance et obtenez le message d'erreur `unable to load key ... Expecting: ANY PRIVATE KEY`, le fichier dans lequel la clé privée est stockée est mal configuré. Si le fichier de clé privée se termine par `.pem`, il est peut-être toujours mal configuré. Une cause possible de configuration incorrecte d'un fichier de clé privée est l'absence d'un certificat.

Si le fichier de clé privée est mal configuré, suivez ces étapes pour corriger l'erreur.

1. Créez une nouvelle paire de clés. Pour plus d'informations, consultez Option 1 : Créer une paire de clés à l'aide de Amazon EC2 (p. 1048).
2. Ajoutez la nouvelle paire de clés à votre instance. Pour plus d'informations, consultez Connexion à votre instance Linux en cas de perte de votre clé privée (p. 1056).
3. Connectez-vous à votre instance à l'aide de la nouvelle paire de clés.

Erreur : clé de l'utilisateur non reconnue par le serveur

Si vous utilisez SSH pour vous connecter à votre instance

- Utilisez `ssh -vvv` pour obtenir des informations très détaillées sur le débogage en vous connectant :

```
ssh -vvv -i path/my-key-pair.pem my-instance-user-
name@ec2-203-0-113-25.compute-1.amazonaws.com
```

L'exemple de données de sortie suivant montre que vous pouvez voir si vous étiez en train de vous connecter à votre instance avec une clé qui n'était pas reconnue par le serveur.

```
open/ANT/myusername/.ssh/known_hosts).
debug2: bits set: 504/1024
debug1: ssh_rsa_verify: signature correct
debug2: kex_derive_keys
debug2: set_newkeys: mode 1
debug1: SSH2_MSG_NEWKEYS sent
debug1: expecting SSH2_MSG_NEWKEYS
debug2: set_newkeys: mode 0
debug1: SSH2_MSG_NEWKEYS received
debug1: Roaming not allowed by server
debug1: SSH2_MSG_SERVICE_REQUEST sent
debug2: service_accept: ssh-userauth
debug1: SSH2_MSG_SERVICE_ACCEPT received
debug2: key: boguspem.pem ((nil))
debug1: Authentications that can continue: publickey
debug3: start over, passed a different list publickey
debug3: preferred gssapi-keyex,gssapi-with-mic,publickey,keyboard-interactive,password
debug3: authmethod_lookup publickey
debug3: remaining preferred: keyboard-interactive,password
debug3: authmethod_is_enabled publickey
debug1: Next authentication method: publickey
debug1: Trying private key: boguspem.pem
debug1: read PEM private key done: type RSA
debug3: sign_and_send_pubkey: RSA 9c:4c:bc:0c:d0:5c:c7:92:6c:8e:9b:16:e4:43:d8:b2
debug2: we sent a publickey packet, wait for reply
debug1: Authentications that can continue: publickey
debug2: we did not send a packet, disable method
debug1: No more authentication methods to try.
```

```
Permission denied (publickey).
```

Si vous utilisez PuTTY pour vous connecter à votre instance

- Vérifiez que votre fichier de clé privée (.pem) a été converti au format reconnu par PuTTY (.ppk). Pour plus d'informations sur la conversion de votre clé privée, consultez Connexion à votre instance Linux à partir de Windows à l'aide de PuTTY (p. 614).

  Note

  Dans PuTTYgen, chargez votre fichier de clé privée et sélectionnez Enregistrer la clé privée plutôt que Générer.
- Vérifiez que vous vous connectez avec le nom utilisateur approprié pour votre AMI. Saisissez le nom d'utilisateur dans le champ Nom d'hôte de la fenêtre Configuration de PuTTY.
  - Pour Amazon Linux 2 ou l'AMI Amazon Linux, le nom d'utilisateur est `ec2-user`.
  - Pour une AMI CentOS, le nom d'utilisateur est `centos`.
  - Pour une AMI Debian, le nom d'utilisateur est `admin`.
  - Pour une AMI Fedora, le nom d'utilisateur est `ec2-user` ou `fedora`.
  - Pour une AMI RHEL, le nom d'utilisateur est `ec2-user` ou `root`.
  - Pour une AMI SUSE, le nom d'utilisateur est `ec2-user` ou `root`.
  - Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.
  - Si `ec2-user` et `root` ne fonctionnent pas, renseignez-vous auprès de votre fournisseur AMI.
- Vérifiez que vous avez une règle entrante de groupe de sécurité pour permettre le trafic entrant vers le port approprié. Pour plus d'informations, consultez Autorisation de l'accès réseau à vos instances (p. 1044).

Erreur : autorisation refusée ou connexion fermée par [instance] port 22

Si vous vous connectez à votre instance à l'aide de SSH et que vous obtenez l'une des erreurs suivantes,`Host key not found in [directory]`, `Permission denied (publickey)`, `Authentication failed, permission denied` ou `Connection closed by [instance] port 22`, vérifiez que vous vous connectez avec nom d'utilisateur approprié pour votre AMI et que vous avez indiqué le bon fichier de clé privée (`.pem`)) pour votre instance.

Les noms utilisateur appropriés sont comme suit :

- Pour Amazon Linux 2 ou l'AMI Amazon Linux, le nom d'utilisateur est `ec2-user`.
- Pour une AMI CentOS, le nom d'utilisateur est `centos`.
- Pour une AMI Debian, le nom d'utilisateur est `admin`.
- Pour une AMI Fedora, le nom d'utilisateur est `ec2-user` ou `fedora`.
- Pour une AMI RHEL, le nom d'utilisateur est `ec2-user` ou `root`.
- Pour une AMI SUSE, le nom d'utilisateur est `ec2-user` ou `root`.
- Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.
- Si `ec2-user` et `root` ne fonctionnent pas, renseignez-vous auprès de votre fournisseur AMI.

Pa exemple, pour utiliser un client SSH et vous connecter à une instance Amazon Linux, utilisez la commande suivante :

```
ssh -i /path/my-key-pair.pem my-instance-user-name@ec2-203-0-113-25.compute-1.amazonaws.com
```

Confirmez que vous utilisez le fichier de clé privée qui correspond à la paire de clés que vous avez sélectionnée lorsque vous avez lancé l'instance.

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Sélectionnez votre instance. Dans l'onglet Description, vérifiez la valeur de Nom de la paire de clés.
3. Si vous n'avez pas spécifié une paire de clés lorsque vous avez lancé l'instance, vous pouvez mettre fin à l'instance et lancer une nouvelle instance en vous assurant de spécifier une paire de clés. S'il s'agit d'une instance que vous avez utilisée, mais que vous n'avez plus le fichier `.pem` pour votre paire de clés, vous pouvez remplacer la paire de clés par une nouvelle. Pour plus d'informations, consultez Connexion à votre instance Linux en cas de perte de votre clé privée (p. 1056).

Si vous avez généré votre propre paire de clés, assurez-vous que votre générateur de clés est configuré pour créer des clés RSA. Les clés DSA ne sont pas acceptées.

Si vous obtenez une erreur `Permission denied (publickey)` et qu'aucune des réponses ci-dessus ne s'applique (par exemple, vous avez pu vous connecter précédemment), les autorisations sur le répertoire de base de votre instance a peut-être été modifiées. Les autorisations pour `/home/`*`my-instance-user-name`*`/.ssh/authorized_keys` doivent être limitées au propriétaire uniquement.

Pour vérifier les autorisations sur votre instance

1. Arrêtez votre instance et détachez le volume racine. Pour plus d'informations, consultez Arrêt et démarrage de votre instance (p. 625) et Détacher un volume Amazon EBS d'une instance Linux (p. 1127).
2. Lancez une instance temporaire dans la même zone de disponibilité que votre instance actuelle (utilisez une AMI similaire ou la même AMI que vous avez utilisée pour votre instance actuelle) et attachez le volume racine à l'instance temporaire. Pour plus d'informations, consultez Attacher un volume Amazon EBS à une instance (p. 1109).
3. Connectez-vous à l'instance temporaire, créez un point de montage et montez le volume que vous avez joint. Pour plus d'informations, consultez Rendre un volume Amazon EBS disponible à l'utilisation sur le Linux (p. 1113).
4. A partir de l'instance temporaire, vérifiez les autorisations du répertoire `/home/`*`my-instance-user-name`*`/` du volume attaché. Si nécessaire, modifiez les autorisations comme suit :

   ```
   [ec2-user ~]$ chmod 600 mount_point/home/my-instance-user-name/.ssh/authorized_keys
   ```

   ```
   [ec2-user ~]$ chmod 700 mount_point/home/my-instance-user-name/.ssh
   ```

   ```
   [ec2-user ~]$ chmod 700 mount_point/home/my-instance-user-name
   ```

5. Démontez le volume, détachez-le de l'instance temporaire et attachez-le de nouveau à l'instance originale. Assurez-vous que vous avez spécifié le bon nom de périphérique pour le volume racine, par exemple, `/dev/xvda`.
6. Démarrez votre instance. Si vous n'avez plus besoin de l'instance temporaire, vous pouvez la mettre en service.

Erreur : fichier de clé privée non protégé

Votre fichier de clé privée doit être protégé des opérations de lecture et d'écriture des autres utilisateurs. Si n'importe qui sauf vous peut lire ou écrire sur votre clé privée, alors SSH ignore votre clé et vous voyez le message d'avertissement suivant.

```
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
@         WARNING: UNPROTECTED PRIVATE KEY FILE!          @
@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@@
Permissions 0777 for '.ssh/my_private_key.pem' are too open.
It is required that your private key files are NOT accessible by others.
This private key will be ignored.
bad permissions: ignore key: .ssh/my_private_key.pem
Permission denied (publickey).
```

Si vous voyez un message similaire lorsque vous essayez de vous connecter à votre instance, examinez la première ligne du message d'erreur pour vérifier que vous utiliser la bonne clé publique pour votre instance. L'exemple ci-dessus utilise la clé privée `.ssh/my_private_key.pem` avec les autorisations sur les fichiers de `0777` ce qui permet à n'importe qui de lire ou écrire sur ce fichier. Ce niveau d'autorisation n'est pas sûr du tout, donc SSH ignore cette clé. Pour corriger cette erreur, exécutez la commande suivante en remplaçant le chemin de votre fichier de clé privée.

```
[ec2-user ~]$ chmod 0400 .ssh/my_private_key.pem
```

Erreur : La clé privée doit commencer par « -----BEGIN RSA PRIVATE KEY----- » et se terminer par « -----END RSA PRIVATE KEY----- »

Si vous utilisez un outil tiers, tel que ssh-keygen, pour créer une paire de clés RSA, il génère la clé privée au format de clé OpenSSH. Lorsque vous vous connectez à votre instance, si vous utilisez la clé privée au format OpenSSH pour déchiffrer le mot de passe, vous obtenez l'erreur `Private key must begin with "-----BEGIN RSA PRIVATE KEY-----" and end with "-----END RSA PRIVATE KEY-----"`.

Pour résoudre cette erreur, la clé privée doit être au format PEM. Utilisez la commande suivante pour créer la clé privée au format PEM :

```
ssh-keygen -m PEM
```

Erreur : le serveur a refusé notre clé ou Aucune méthode d'authentification prise en charge disponible

Si vous utilisez PuTTY pour vous connecter à votre instance et que vous obtenez l'une des erreurs suivantes, Erreur : Le serveur a refusé votre clé ou Erreur : Méthodes d'authentification disponibles non prises en charge, vérifiez que vous vous connectez avec le nom d'utilisateur approprié pour votre AMI. Entrez le nom d'utilisateur dans le champ Nom d'utilisateur de la fenêtre Configuration de PuTTY.

Les noms d'utilisateur appropriés sont comme suit :

- Pour Amazon Linux 2 ou l'AMI Amazon Linux, le nom d'utilisateur est `ec2-user`.

- Pour une AMI CentOS, le nom d'utilisateur est `centos`.
- Pour une AMI Debian, le nom d'utilisateur est `admin`.
- Pour une AMI Fedora, le nom d'utilisateur est `ec2-user` ou `fedora`.
- Pour une AMI RHEL, le nom d'utilisateur est `ec2-user` ou `root`.
- Pour une AMI SUSE, le nom d'utilisateur est `ec2-user` ou `root`.
- Pour une AMI Ubuntu, le nom utilisateur est `ubuntu`.
- Si `ec2-user` et `root` ne fonctionnent pas, renseignez-vous auprès de votre fournisseur AMI.

Vous devriez aussi vérifier que votre fichier de clé privée (.pem) a été correctement converti au format reconnu par PuTTY (.ppk). Pour plus d'informations sur la conversion de votre clé privée, consultez Connexion à votre instance Linux à partir de Windows à l'aide de PuTTY (p. 614).

Impossible d'envoyer une commande ping à l'instance

La commande `ping` est un type de trafic ICMP. Si vous ne pouvez pas pinger votre instance, assurez-vous que vos règles entrantes de groupe de sécurité autorisent le trafic ICMP pour le message `Echo Request` de toutes les sources, ou de l'ordinateur ou de l'instance à partir desquels vous émettez la commande.

Si vous ne pouvez pas fournir une commande `ping` à partir de votre instance, assurez-vous que vos règles sortantes de groupe de sécurité autorisent le trafic ICMP pour le message `Echo Request` vers toutes les destinations ou vers l'hôte que vous essayer de pinger.

Les commandes `Ping` peuvent également être bloquées par un pare-feu ou un délai d'attente en raison de latence réseau ou de problèmes matériels. Vous devez consulter votre réseau local ou votre administrateur système pour obtenir de l'aide sur la résolution des problèmes supplémentaires.

Erreur : le serveur a fermé la connexion réseau de manière inopinée

Si vous vous connectez à votre instance via PuTTY et que vous recevez le message d'erreur « Le serveur a fermé la connexion réseau de manière inopinée », vérifiez que vous avez activé le paramètre keepalive dans la page de Connexion de la Configuration PuTTY, afin d'éviter de vous faire déconnecter. Certains serveurs déconnectent les clients lorsqu'ils n'ont pas reçu de données dans une période de temps spécifiée. Réglez les secondes entre keepalives à 59 secondes.

Si vous éprouvez encore des difficultés après avoir activé les keepalives, essayez de désactiver l'algorithme de Nagle dans la page de Connexion de la Configuration PuTTY.

Résolution des problèmes d'arrêt de votre instance

Si vous avez arrêté votre instance basée sur Amazon EBS et que celle-ci semble « bloquée » à l'état `stopping`, il peut y avoir un problème avec l'ordinateur hôte sous-jacent.

L'utilisation d'une instance est gratuite tant que l'instance n'est pas à l'état `running`.

Forcez l'arrêt de l'instance à l'aide de la console ou de l'AWS CLI.

- Pour forcer l'arrêt de l'instance à l'aide de la console, sélectionnez l'instance bloquée et choisissez État de l'instance, Arrêter l'instance et Oui, forcer l'arrêt.

- Pour forcer l'arrêt de l'instance à l'aide de l'AWS CLI, utilisez la commande stop-instances et l'option `--force` de la façon suivante :

```
aws ec2 stop-instances --instance-ids i-0123ab456c789d01e --force
```

Si l'instance ne s'est pas arrêtée après 10 minutes, publiez une demande d'aide sur le Amazon EC2 forum. Pour contribuer à une résolution rapide du problème, incluez l'ID d'instance et décrivez les étapes que vous avez déjà effectuées. Sinon, si vous disposez d'un plan de support, créez une demande d'assistance technique dans le Centre de support.

Création d'une instance de remplacement

Pour essayer de résoudre le problème en attendant d'obtenir de l'aide de la part du Amazon EC2 forum ou du Centre de support, créez une instance de remplacement. Créez une AMI de l'instance bloquée et lancez une nouvelle instance à l'aide de cette AMI.

Pour créer une instance de remplacement à l'aide de la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances et choisissez l'instance bloquée.
3. Sélectionnez Actions, Image, Créer une image.
4. Dans la boîte de dialogue Créer une image, remplissez les champs suivants, puis choisissez Créer l'image :

   a. Attribuez un nom et une description pour l'AMI.
   b. Sélectionnez Pas de redémarrage.

   Pour plus d'informations, consultez Création d'une AMI Linux à partir d'une instance (p. 133).
5. Lancez une nouvelle instance à partir de l'AMI et vérifiez qu'elle fonctionne.
6. Sélectionnez l'instance bloquée et choisissez Actions, État de l'instance, Résilier. Si l'instance reste aussi bloquée lors de la mise à fin, Amazon EC2 la force automatiquement à se terminer en quelques heures.

Pour créer une instance de remplacement à l'aide de l'interface de ligne de commande

1. Créez une AMI à partir de l'instance bloquée, en utilisant la commande create-image (AWS CLI) et l'option `--no-reboot` de la façon suivante .

```
aws ec2 create-image --instance-id i-0123ab456c789d01e --name "AMI" --description "AMI
 for replacement instance" --no-reboot
```

2. Lancez une nouvelle instance à partir de l'AMI en utilisant la commande run-instances (AWS CLI) de la façon suivante :

```
aws ec2 run-instances --image-id ami-1a2b3c4d --count 1 --instance-type c3.large --key-
name MyKeyPair --security-groups MySecurityGroup
```

3. Vérifiez que la nouvelle instance fonctionne.
4. Mettez fin à l'instance bloquée en utilisant la commande terminate-instances (AWS CLI) de la façon suivante :

```
aws ec2 terminate-instances --instance-ids i-1234567890abcdef0
```

Si vous ne pouvez pas créer une AMI à partir de l'instance comme décrit dans les procédures précédentes, vous pouvez configurer une instance de remplacement de la façon suivante :

(Alternative) Pour créer une instance de remplacement à l'aide de la console

1. Sélectionnez l'instance et choisissez Description, Périphériques de stockage en mode bloc. Sélectionnez chaque volume et écrivez son volume ID. Assurez-vous de noter quel volume correspond au volume racine.
2. Dans le panneau de navigation, choisissez Volumes. Sélectionnez chaque volume pour l'instance et sélectionnez Actions, Créer un instantané.
3. Dans le panneau de navigation, choisissez Snapshots. Sélectionnez l'instantané que vous venez de créer et choisissez Actions, Créer un volume.
4. Lancez une instance avec le même système d'exploitation que l'instance bloqué. Notez l'ID du volume et le nom de périphérique de son volume racine.
5. Dans le panneau de navigation, sélectionnez Instances, choisissez l'instance que vous venez de lancer, sélectionnez État de l'instance, Arrêter l'instance.
6. Dans le panneau de navigation, sélectionnez Volumes, choisissez le volume racine de l'instance arrêtée, et sélectionnez Actions, Détacher un volume.
7. Sélectionnez le volume racine que vous avez créé à partir de l'instance bloquée, choisissez Actions, Attacher un volume et attachez-le à la nouvelle instance comme volume racine (en utilisant le nom de périphérique que vous avez écrit). Attachez n'importe quel volume non-racine supplémentaire à l'instance.
8. Dans le panneau de navigation, sélectionnez Instances et choisissez l'instance de remplacement. Choisissez État de l'instance, Démarrer l'instance. Vérifiez que l'instance fonctionne.
9. Sélectionnez l'instance bloquée, choisissez État de l'instance, Résilier l'instance. Si l'instance reste également bloquée lors de la mise hors service, Amazon EC2 force automatiquement sa mise hors service en quelques heures.

Résolution des problèmes de mise hors service (d'arrêt) de votre instance

Vous n'êtes pas facturé pour l'utilisation d'une instance tant que l'instance n'est pas à l'état `running`. En d'autres termes, lorsque vous mettez fin à une instance, l'instance ne vous est plus facturée dès que son état passe à `shutting-down`.

Mise à fin d'instance retardée

Si votre instance reste à l'état `shutting-down` pendant plus que quelques minutes, elle peut être retardée à cause des scripts d'arrêt exécutés par l'instance.

Un autre cause possible est un problème avec l'ordinateur hôte sous-jacent. Si votre instance reste à l'état `shutting-down` pendant plusieurs heures, Amazon EC2 la considère comme une instance bloquée et la résilie de force.

S'il semble que votre instance est bloquée pendant la mise à fin et que cela dure depuis plus de quelques heures, publiez une demande d'aide sur le Amazon EC2 forum. Pour aider à accélérer la résolution d'un problème, incluez l'ID d'instance et décrivez les étapes que vous avez déjà effectuées. Sinon, si vous disposez d'un plan de support, créez une demande d'assistance technique dans le Centre de support.

Instance terminée toujours affichée

Après avoir mis fin à une instance, elle reste visible pendant un court instant avant d'être supprimée. L'état indique `terminated`. Si l'entrée n'est pas supprimée après plusieurs heures, contactez le support.

Instances lancées ou terminées automatiquement

De manière générale, ces comportements signifient que vous avez utilisé Amazon EC2 Auto Scaling ou Flotte EC2 pour mettre automatiquement à l'échelle vos ressources informatiques en fonction des critères que vous avez définis.

- Vous mettez fin à une instance et une nouvelle instance se lance automatiquement.
- Vous lancez une instance et l'une de vos instances se termine automatiquement.
- Vous arrêtez une instance, elle se termine et une nouvelle instance se lance automatiquement.

Pour arrêter la mise à l'échelle automatique, veuillez consulter Amazon EC2 Auto Scaling Guide de l'utilisateur ou Lancement d'instances à l'aide d'un Flotte EC2 (p. 558).

Résolution des instances avec des contrôles de statut échoués

Les informations suivantes peuvent vous aider à résoudre les problèmes si votre instance échoue à un contrôle de statut. Commencez par déterminer si vos applications présentent des problèmes. Si vous constatez que l'instance n'exécute pas vos applications comme prévu, passez en revue les informations de contrôle de statut et les journaux système.

Sommaire

- Examen des informations de contrôle de statut (p. 1339)
- Récupération des journaux système (p. 1339)
- Dépannage des erreurs du journal du système pour les instances basées sur Linux (p. 1340)
- Mémoire insuffisante : processus d'arrêt (p. 1341)
- ERROR: mmu_update failed (la mise à jour de la gestion de la mémoire a échoué) (p. 1341)
- Erreur d'E/S (échec du périphérique de stockage en mode bloc) (p. 1342)
- I/O ERROR: neither local nor remote disk (le périphérique de stockage en mode bloc distribué ne fonctionne plus) (p. 1343)
- request_module: runaway loop modprobe (modprobe en boucle sur le noyau hérité sur des versions Linux plus anciennes) (p. 1344)
- « FATAL: kernel too old » et « fsck: No such file or directory while trying to open /dev » (décalage entre le noyau et l'AMI) (p. 1345)
- « FATAL: Could not load /lib/modules » ou « BusyBox » (modules noyau manquants) (p. 1346)
- ERROR Invalid kernel (noyau incompatible EC2) (p. 1347)
- fsck: No such file or directory while trying to open... (système de fichiers non trouvé) (p. 1348)
- General error mounting filesystems (Montage en échec) (p. 1350)
- VFS: Unable to mount root fs on unknown-block (le système de fichiers racine ne correspond pas) (p. 1351)
- Erreur : Unable to determine major/minor number of root device... (décalage du système de fichiers/périphérique racine) (p. 1352)
- XENBUS : Device with no driver... (p. 1353)
- ... days without being checked, check forced (Contrôle du système de fichiers nécessaire) (p. 1354)
- fsck died with exit status... (périphérique manquant) (p. 1355)
- Invite GRUB (grubdom>) (p. 1356)
- Bringing up interface eth0: Device eth0 has different MAC address than expected, ignoring. (Adresse MAC codée de manière irréversible) (p. 1358)

- Unable to load SELinux Policy. Machine is in enforcing mode. Halting now. (Erreur de configuration SELinux) (p. 1359)
- XENBUS: Timeout connecting to devices (délai d'attente Xenbus) (p. 1360)

Examen des informations de contrôle de statut

Pour enquêter sur les instances dégradées en utilisant la console Amazon EC2

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, choisissez Instances, puis sélectionnez votre instance.
3. Dans le volet des détails, sélectionnez Contrôles des statuts pour voir les résultats individuels pour tous les Contrôles de statut de système et Contrôles de statut des instances.

Si un contrôle de statut d'un système a échoué, vous pouvez essayer l'une des options suivantes :

- Créez une alarme de récupération d'instance. Pour plus d'informations, consultez Créer des alarmes qui arrêtent, finissent, redémarrent ou récupèrent une instance (p. 785).
- Si vous avez modifié le type d'instance pour définir une instance basée sur le système Nitro (p. 218), les contrôles de statut échouent si vous avez migré à partir d'une instance qui ne possède pas les pilotes ENA et NVMe requis. Pour plus d'informations, consultez Compatibilité pour le redimensionnement des instances (p. 313).
- Pour une instance qui utilise une AMI basée sur des volumes Amazon EBS, arrêtez et relancez l'instance.
- Pour une instance qui utilise une AMI basée sur le stockage d'instance, arrêtez l'instance et lancez-en une autre.
- Attendez qu'Amazon EC2 résolve le problème.
- Publiez votre problème sur le Amazon EC2 forum.
- Si votre instance est dans un groupe Auto Scaling, le service Amazon EC2 Auto Scaling lance automatiquement une instance de remplacement. Pour plus d'informations, consultez Vérification de l'état des instances Auto Scaling dans le Amazon EC2 Auto Scaling Guide de l'utilisateur.
- Récupérez le journal du système et recherchez les erreurs.

Récupération des journaux système

Si un contrôle de statut d'instance échoue, vous pouvez relancer l'instance et récupérer les journaux du système. Les journaux peuvent révéler une erreur que peut vous aider à résoudre le problème. Le redémarrage efface les informations inutiles des journaux.

Pour redémarrer une instance et récupérer le journal du système

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation, sélectionnez Instances, puis choisissez votre instance.
3. Choisissez État de l'instance, Redémarrer l'instance. Le redémarrage de votre instance peut prendre quelques minutes.
4. Vérifiez si le problème existe encore. Dans certains cas, le redémarrage peut résoudre le problème.
5. Lorsque l'instance possède l'état `running`, sélectionnez Actions, Paramètres de l'instance, Obtenir le journal système.
6. Consultez le journal qui apparaît à l'écran et utilisez la liste ci-dessous des déclarations d'erreurs connues du journal du système afin de résoudre votre problème.

7. Si votre expérience diffère de nos résultats de contrôle ou si vous rencontrez un problème avec votre instance que nos contrôles n'ont pas détecté, sélectionnez Envoyer les commentaires dans l'onglet Contrôles des statuts pour nous aider à améliorer nos tests de détection.
8. Si votre problème n'est pas résolu, vous ne pouvez pas publier votre problème sur le Amazon EC2 forum.

Dépannage des erreurs du journal du système pour les instances basées sur Linux

Pour les instances basées sur Linux qui ont échoué à un contrôle de statut d'instance, comme le contrôle d'accessibilité de l'instance, vérifiez que vous avez suivi les étapes ci-dessous pour récupérer le journal du système. La liste suivante contient certaines erreurs communes du journal du système et les actions suggérées que vous pouvez prendre pour résoudre le problème correspondant à chaque erreur.

Memory Errors

- Mémoire insuffisante : processus d'arrêt (p. 1341)
- ERROR: mmu_update failed (la mise à jour de la gestion de la mémoire a échoué) (p. 1341)

Device Errors

- Erreur d'E/S (échec du périphérique de stockage en mode bloc) (p. 1342)
- I/O ERROR: neither local nor remote disk (le périphérique de stockage en mode bloc distribué ne fonctionne plus) (p. 1343)

Kernel Errors

- request_module: runaway loop modprobe (modprobe en boucle sur le noyau hérité sur des versions Linux plus anciennes) (p. 1344)
- « FATAL: kernel too old » et « fsck: No such file or directory while trying to open /dev » (décalage entre le noyau et l'AMI) (p. 1345)
- « FATAL: Could not load /lib/modules » ou « BusyBox » (modules noyau manquants) (p. 1346)
- ERROR Invalid kernel (noyau incompatible EC2) (p. 1347)

File System Errors

- fsck: No such file or directory while trying to open... (système de fichiers non trouvé) (p. 1348)
- General error mounting filesystems (Montage en échec) (p. 1350)
- VFS: Unable to mount root fs on unknown-block (le système de fichiers racine ne correspond pas) (p. 1351)
- Erreur : Unable to determine major/minor number of root device... (décalage du système de fichiers/périphérique racine) (p. 1352)
- XENBUS : Device with no driver... (p. 1353)
- ... days without being checked, check forced (Contrôle du système de fichiers nécessaire) (p. 1354)
- fsck died with exit status... (périphérique manquant) (p. 1355)

Operating System Errors

- Invite GRUB (grubdom>) (p. 1356)

- Bringing up interface eth0: Device eth0 has different MAC address than expected, ignoring. (Adresse MAC codée de manière irréversible) (p. 1358)
- Unable to load SELinux Policy. Machine is in enforcing mode. Halting now. (Erreur de configuration SELinux) (p. 1359)
- XENBUS: Timeout connecting to devices (délai d'attente Xenbus) (p. 1360)

Mémoire insuffisante : processus d'arrêt

Une erreur de mémoire insuffisante est indiquée par une entrée dans le journal système similaire à celle indiquée ci-dessous :

```
[115879.769795] Out of memory: kill process 20273 (httpd) score 1285879
or a child
[115879.769795] Killed process 1917 (php-cgi) vsz:467184kB, anon-
rss:101196kB, file-rss:204kB
```

Cause potentielle

Mémoire épuisée

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Effectuez l'une des actions suivantes : • Arrêtez l'instance et modifiez l'instance pour utiliser un type d'instance différent, puis relancez l'instance. Par exemple, un type d'instance plus importante ou optimisée pour la mémoire. • Redémarrez l'instance pour la renvoyer vers un statut non-défaillant. Le problème se reproduira probablement à moins que vous ne changiez de type d'instance.
Basée sur le stockage d'instance	Effectuez l'une des actions suivantes : • Arrêtez l'instance et lancez une nouvelle instance en spécifiant un type d'instance différent. Par exemple, un type d'instance plus importante ou optimisée pour la mémoire. • Redémarrez l'instance pour la renvoyer vers un statut non-défaillant. Le problème se reproduira probablement à moins que vous ne changiez de type d'instance.

ERROR: mmu_update failed (la mise à jour de la gestion de la mémoire a échoué)

Les échecs de la mise à jour de la gestion de la mémoire sont indiqués par une entrée du journal du système qui est similaire à ce qui suit :

```
...
Press `ESC' to enter the menu... 0   [H[J  Booting 'Amazon Linux 2011.09
 (2.6.35.14-95.38.amzn1.i686)'


root (hd0)

 Filesystem type is ext2fs, using whole disk

kernel /boot/vmlinuz-2.6.35.14-95.38.amzn1.i686 root=LABEL=/ console=hvc0 LANG=

en_US.UTF-8 KEYTABLE=us

initrd /boot/initramfs-2.6.35.14-95.38.amzn1.i686.img

ERROR: mmu_update failed with rc=-22
```

Cause potentielle

Problème avec Amazon Linux

Action suggérée

Publiez votre problème sur Forums dédiés aux développeurs ou contactez AWS Support.

Erreur d'E/S (échec du périphérique de stockage en mode bloc)

Une erreur d'entrée/sortie est indiquée par une entrée du journal du système qui est similaire à l'exemple suivant :

```
[9943662.053217] end_request: I/O error, dev sde, sector 52428288
[9943664.191262] end_request: I/O error, dev sde, sector 52428168
[9943664.191285] Buffer I/O error on device md0, logical block 209713024
[9943664.191297] Buffer I/O error on device md0, logical block 209713025
[9943664.191304] Buffer I/O error on device md0, logical block 209713026
[9943664.191310] Buffer I/O error on device md0, logical block 209713027
[9943664.191317] Buffer I/O error on device md0, logical block 209713028
[9943664.191324] Buffer I/O error on device md0, logical block 209713029
[9943664.191332] Buffer I/O error on device md0, logical block 209713030
[9943664.191339] Buffer I/O error on device md0, logical block 209713031
[9943664.191581] end_request: I/O error, dev sde, sector 52428280
[9943664.191590] Buffer I/O error on device md0, logical block 209713136
[9943664.191597] Buffer I/O error on device md0, logical block 209713137
[9943664.191767] end_request: I/O error, dev sde, sector 52428288
[9943664.191970] end_request: I/O error, dev sde, sector 52428288
[9943664.192143] end_request: I/O error, dev sde, sector 52428288
[9943664.192949] end_request: I/O error, dev sde, sector 52428288
[9943664.193112] end_request: I/O error, dev sde, sector 52428288
[9943664.193266] end_request: I/O error, dev sde, sector 52428288
...
```

Causes potentielles

Type d'instance	Cause potentielle
Basée sur Amazon EBS	Un volume Amazon EBS en échec

Type d'instance	Cause potentielle
Basée sur le stockage d'instance	Un lecteur physique en échec

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance. 2. Dissociez le volume. 3. Essayez de récupérer le volume. Note Il est recommandé de faire souvent des instantanés de vos volumes Amazon EBS. Cela diminue considérablement le risque de perte de données suite à un échec. 4. Attachez de nouveau le volume à l'instance. 5. Démarrez l'instance.
Basée sur le stockage d'instance	Mettez fin à l'instance et lancez une nouvelle instance. Note Les données ne peuvent pas être récupérées. Récupérez-les grâce aux sauvegardes. Note Il est recommandé d'utiliser soit Amazon S3, soit Amazon EBS pour les sauvegardes. Les volumes de stockage d'instance sont directement reliés aux échecs d'un hôte et d'un disque uniques.

I/O ERROR: neither local nor remote disk (le périphérique de stockage en mode bloc distribué ne fonctionne plus)

Une erreur d'entrée/sortie sur le périphérique est indiquée par une entrée du journal du système qui est similaire à l'exemple suivant :

```
...
block drbd1: Local IO failed in request_timer_fn. Detaching...

Aborting journal on device drbd1-8.
```

```
block drbd1: IO ERROR: neither local nor remote disk

Buffer I/O error on device drbd1, logical block 557056

lost page write due to I/O error on drbd1

JBD2: I/O error detected when updating journal superblock for drbd1-8.
```

Causes potentielles

Type d'instance	Cause potentielle
Basée sur Amazon EBS	Un volume Amazon EBS en échec
Basée sur le stockage d'instance	Un lecteur physique en échec

Action suggérée

Mettez fin à l'instance et lancez une nouvelle instance.

Pour une instance basée sur Amazon EBS, vous pouvez récupérer des données à partir d'un instantané récent en créant une image à partir de celle-ci. Toutes les données ajoutées après l'instantané ne peuvent pas être récupérées.

request_module: runaway loop modprobe (modprobe en boucle sur le noyau hérité sur des versions Linux plus anciennes)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous. L'utilisation d'un noyau Linux instable ou ancien (par exemple, 2.6.16-xenU) peut entraîner une condition de boucle interminable au démarrage.

```
Linux version 2.6.16-xenU (builder@xenbat.amazonsa) (gcc version 4.0.1
20050727 (Red Hat 4.0.1-5)) #1 SMP Mon May 28 03:41:49 SAST 2007

BIOS-provided physical RAM map:

 Xen: 0000000000000000 - 0000000026700000 (usable)

0MB HIGHMEM available.
...

request_module: runaway loop modprobe binfmt-464c

request_module: runaway loop modprobe binfmt-464c

request_module: runaway loop modprobe binfmt-464c

request_module: runaway loop modprobe binfmt-464c

request_module: runaway loop modprobe binfmt-464c
```

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez un noyau plus récent, soit basé sur GRUB ou statique, avec l'une des options suivantes: Option 1 : Arrêtez l'instance et lancez une nouvelle instance en spécifiant les paramètres `-kernel` et `-ramdisk`. Option 2 : 1. Arrêtez l'instance. 2. Modifiez les attributs de noyau et de ramdisk pour utiliser un noyau plus récent. 3. Démarrez l'instance.
Basée sur le stockage d'instance	Arrêtez l'instance et lancez une nouvelle instance en spécifiant les paramètres `-kernel` et `-ramdisk`.

« FATAL: kernel too old » et « fsck: No such file or directory while trying to open /dev » (décalage entre le noyau et l'AMI)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
Linux version 2.6.16.33-xenU (root@dom0-0-50-45-1-a4-ee.z-2.aes0.internal)
(gcc version 4.1.1 20070105 (Red Hat 4.1.1-52)) #2 SMP Wed Aug 15 17:27:36 SAST 2007
...
FATAL: kernel too old
Kernel panic - not syncing: Attempted to kill init!
```

Causes potentielles

Noyau et identifiant incompatibles

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance. 2. Modifiez la configuration pour utiliser un noyau plus récent. 3. Démarrez l'instance.

Pour ce type d'instance	Faire ceci
Basée sur le stockage d'instance	Utilisez la procédure suivante. 1. Créez une AMI qui utilise un noyau plus récent. 2. Mettez fin à l'instance. 3. Démarrez une nouvelle instance à partir de l'AMI que vous avez créée.

« FATAL: Could not load /lib/modules » ou « BusyBox » (modules noyau manquants)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
[    0.370415] Freeing unused kernel memory: 1716k freed
Loading, please wait...
WARNING: Couldn't open directory /lib/modules/2.6.34-4-virtual: No such file or directory
FATAL: Could not open /lib/modules/2.6.34-4-virtual/modules.dep.temp for writing: No such
 file or directory
FATAL: Could not load /lib/modules/2.6.34-4-virtual/modules.dep: No such file or directory
Couldn't get a file descriptor referring to the console
Begin: Loading essential drivers... ...
FATAL: Could not load /lib/modules/2.6.34-4-virtual/modules.dep: No such file or directory
FATAL: Could not load /lib/modules/2.6.34-4-virtual/modules.dep: No such file or directory
Done.
Begin: Running /scripts/init-premount ...
Done.
Begin: Mounting root file system... ...
Begin: Running /scripts/local-top ...
Done.
Begin: Waiting for root file system... ...
Done.
Gave up waiting for root device.  Common problems:
 - Boot args (cat /proc/cmdline)
   - Check rootdelay= (did the system wait long enough?)
   - Check root= (did the system wait for the right device?)
 - Missing modules (cat /proc/modules; ls /dev)
FATAL: Could not load /lib/modules/2.6.34-4-virtual/modules.dep: No such file or directory
FATAL: Could not load /lib/modules/2.6.34-4-virtual/modules.dep: No such file or directory
ALERT! /dev/sda1 does not exist. Dropping to a shell!


BusyBox v1.13.3 (Ubuntu 1:1.13.3-1ubuntu5) built-in shell (ash)
Enter 'help' for a list of built-in commands.

(initramfs)
```

Causes potentielles

Une ou plusieurs conditions suivantes peuvent entraîner ce problème :

- Ramdisk manquant
- Modules corrects manquants pour le ramdisk
- Le volume racine Amazon EBS n'est pas attaché correctement en tant que `/dev/sda1`

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Sélectionnez ramdisk corrigé pour le volume Amazon EBS. 2. Arrêtez l'instance. 3. Détachez le volume et réparez-le. 4. Attachez le volume à l'instance. 5. Démarrez l'instance. 6. Modifiez l'AMI pour utiliser le ramdisk corrigé.
Basée sur le stockage d'instance	Utilisez la procédure suivante. 1. Arrêtez l'instance et lancez une nouvelle instance avec le bon ramdisk. 2. Créez une nouvelle AMI avec le bon ramdisk.

ERROR Invalid kernel (noyau incompatible EC2)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
...
root (hd0)

 Filesystem type is ext2fs, using whole disk

kernel /vmlinuz root=/dev/sda1 ro

initrd /initrd.img

ERROR Invalid kernel: elf_xen_note_check: ERROR: Will only load images
built for the generic loader or Linux images
xc_dom_parse_image returned -1

Error 9: Unknown boot failure

  Booting 'Fallback'

root (hd0)

 Filesystem type is ext2fs, using whole disk

kernel /vmlinuz.old root=/dev/sda1 ro

Error 15: File not found
```

Causes potentielles

Une ou deux des conditions suivantes peuvent entraîner ce problème :

- Le noyau fourni n'est pas pris en charge par GRUB

- Le noyau de rechange n'existe pas

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance. 2. Remplacez-le avec un noyau qui fonctionne. 3. Installez un noyau de rechange. 4. Modifiez l'AMI en corrigeant le noyau.
Basée sur le stockage d'instance	Utilisez la procédure suivante. 1. Arrêtez l'instance et lancez une nouvelle instance avec le bon noyau. 2. Créez une AMI avec le noyau correct. 3. (Facultatif) Demandez une assistance technique pour la récupération des données en utilisant AWS Support.

fsck: No such file or directory while trying to open... (système de fichiers non trouvé)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
  Welcome to Fedora
  Press 'I' to enter interactive startup.
Setting clock : Wed Oct 26 05:52:05 EDT 2011 [  OK  ]

Starting udev: [  OK  ]

Setting hostname localhost:  [  OK  ]

No devices found
Setting up Logical Volume Management: File descriptor 7 left open
  No volume groups found
[  OK  ]

Checking filesystems
Checking all file systems.
[/sbin/fsck.ext3 (1) -- /] fsck.ext3 -a /dev/sda1
/dev/sda1: clean, 82081/1310720 files, 2141116/2621440 blocks
[/sbin/fsck.ext3 (1) -- /mnt/dbbackups] fsck.ext3 -a /dev/sdh
fsck.ext3: No such file or directory while trying to open /dev/sdh

/dev/sdh:
The superblock could not be read or does not describe a correct ext2
filesystem.  If the device is valid and it really contains an ext2
filesystem (and not swap or ufs or something else), then the superblock
is corrupt, and you might try running e2fsck with an alternate superblock:
    e2fsck -b 8193 <device>
```

```
[FAILED]


*** An error occurred during the file system check.
*** Dropping you to a shell; the system will reboot
*** when you leave the shell.
Give root password for maintenance
(or type Control-D to continue):
```

Causes potentielles

- Un bogue existe dans les définitions du système de fichiers ramdisk /etc/fstab
- Définitions du système de fichiers mal configurées dans /etc/fstab
- Lecteur manquant/en échec

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance, détachez le volume racine, réparez/modifiez le volume dans le fichier /etc/fstab, attachez le volume à l'instance et lancez l'instance. 2. Corrigez le ramdisk pour inclure le fichier /etc/fstab modifié (le cas échéant). 3. Modifiez l'AMI pour utiliser un ramdisk plus récent. Le sixième champ de fstab définit les exigences de disponibilité du montage. Une valeur non nulle implique qu'un fsck sera effectué sur ce volume et doit réussir. L'utilisation de ce champ peut être problématique dans Amazon EC2, car un échec entraîne généralement une invite de la console interactive qui n'est actuellement pas disponible dans Amazon EC2. Faites attention avec cette fonction et lisez la page sur la commande man Linux en ce qui concerne fstab.
Basée sur le stockage d'instance	Utilisez la procédure suivante. 1. Mettez fin à l'instance et lancez une nouvelle instance. 2. Détachez tous les volumes Amazon EBS errants et l'instance redémarré. 3. (Facultatif) Demandez une assistance technique pour la récupération des données en utilisant AWS Support.

General error mounting filesystems (Montage en échec)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
Loading xenblk.ko module
xen-vbd: registered block device major 8

Loading ehci-hcd.ko module
Loading ohci-hcd.ko module
Loading uhci-hcd.ko module
USB Universal Host Controller Interface driver v3.0

Loading mbcache.ko module
Loading jbd.ko module
Loading ext3.ko module
Creating root device.
Mounting root filesystem.
kjournald starting.  Commit interval 5 seconds

EXT3-fs: mounted filesystem with ordered data mode.

Setting up other filesystems.
Setting up new root fs
no fstab.sys, mounting internal defaults
Switching to new root and running init.
unmounting old /dev
unmounting old /proc
unmounting old /sys
mountall:/proc: unable to mount: Device or resource busy
mountall:/proc/self/mountinfo: No such file or directory
mountall: root filesystem isn't mounted
init: mountall main process (221) terminated with status 1

General error mounting filesystems.
A maintenance shell will now be started.
CONTROL-D will terminate this shell and re-try.
Press enter for maintenance
(or type Control-D to continue):
```

Causes potentielles

Type d'instance	Cause potentielle
Basée sur Amazon EBS	• Volume Amazon EBS détaché ou en échec. • Système de fichiers corrompu. • Décalage de la combinaison de ramdisk et d'AMI (par exemple, ramdisk Debian avec une AMI SUSE).
Basée sur le stockage d'instance	• Un lecteur en échec. • Un système de fichiers corrompu. • Un décalage de la combinaison de ramdisk et d'AMI (par ex., ramdisk Debian avec une AMI SUSE).

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance. 2. Détachez le volume racine. 3. Attachez le volume racine à une instance connue en fonctionnement. 4. Exécutez le contrôle du système de fichiers (fsck –a /dev/...). 5. Corrigez toutes les erreurs. 6. Détachez le volume de l'instance connue en fonctionnement. 7. Attachez le volume à l'instance arrêtée. 8. Démarrez l'instance. 9. Revérifiez le statut de l'instance.
Basée sur le stockage d'instance	Essayez l'une des actions suivantes : • Démarrez une nouvelle instance. • (Facultatif) Demandez une assistance technique pour la récupération des données en utilisant AWS Support.

VFS: Unable to mount root fs on unknown-block (le système de fichiers racine ne correspond pas)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
Linux version 2.6.16-xenU (builder@xenbat.amazonsa) (gcc version 4.0.1
 20050727 (Red Hat 4.0.1-5)) #1 SMP Mon May 28 03:41:49 SAST 2007
...
Kernel command line:  root=/dev/sda1 ro 4
...
Registering block device major 8
...
Kernel panic - not syncing: VFS: Unable to mount root fs on unknown-block(8,1)
```

Causes potentielles

Type d'instance	Cause potentielle
Basée sur Amazon EBS	• Le périphérique n'est pas attaché correctement. • Le périphérique racine n'est pas attaché au bon point périphérique. • Le système de fichiers n'est pas au format attendu.

Type d'instance	Cause potentielle
	• Utilisez le noyau hérité (par exemple, 2.6.16-XenU). • Mise à jour récente du noyau sur votre instance (mise à jour défaillante ou bogue de mise à jour)
Basée sur le stockage d'instance	Échec du périphérique matériel.

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Effectuez l'une des actions suivantes : • Arrêtez puis redémarrez l'instance. • Modifiez le volume racine pour l'attacher au bon point périphérique, comme /dev/sda1 au lieu de /dev/sda. • Arrêtez et modifiez pour le noyau moderne. • Pour plus d'informations sur les bogues de mise à jour connus, consultez la documentation de votre distribution Linux. Modifiez ou réinstallez le noyau.
Basée sur le stockage d'instance	Arrêtez l'instance et lancez une nouvelle instance en utilisant un noyau moderne.

Erreur : Unable to determine major/minor number of root device... (décalage du système de fichiers/périphérique racine)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
...
XENBUS: Device with no driver: device/vif/0
XENBUS: Device with no driver: device/vbd/2048
drivers/rtc/hctosys.c: unable to open rtc device (rtc0)
Initializing network drop monitor service
Freeing unused kernel memory: 508k freed
:: Starting udevd...
done.
:: Running Hook [udev]
:: Triggering uevents...<30>udevd[65]: starting version 173
done.
Waiting 10 seconds for device /dev/xvda1 ...
Root device '/dev/xvda1' doesn't exist. Attempting to create it.
ERROR: Unable to determine major/minor number of root device '/dev/xvda1'.
You are being dropped to a recovery shell
    Type 'exit' to try and continue booting
sh: can't access tty; job control turned off
[ramfs /]#
```

Causes potentielles

- Pilote du périphérique de stockage en mode bloc virtuel manquant ou configuré de façon incorrecte
- Conflit de l'énumération du périphérique (sda versus xvda ou sda au lieu de sda1)
- Choix incorrect du noyau de l'instance

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance. 2. Dissociez le volume. 3. Corrigez le problème du mappage du périphérique. 4. Démarrez l'instance. 5. Modifiez l'AMI pour traiter les problèmes du mappage du périphérique.
Basée sur le stockage d'instance	Utilisez la procédure suivante. 1. Créez une nouvelle AMI avec la solution appropriée (mapper le périphérique de stockage en mode bloc correctement). 2. Arrêtez l'instance et lancez une nouvelle instance à partir de l'AMI que vous avez créée.

XENBUS : Device with no driver...

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
XENBUS: Device with no driver: device/vbd/2048
drivers/rtc/hctosys.c: unable to open rtc device (rtc0)
Initalizing network drop monitor service
Freeing unused kernel memory: 508k freed
:: Starting udevd...
done.
:: Running Hook [udev]
:: Triggering uevents...<30>udevd[65]: starting version 173
done.
Waiting 10 seconds for device /dev/xvda1 ...
Root device '/dev/xvda1' doesn't exist. Attempting to create it.
ERROR: Unable to determine major/minor number of root device '/dev/xvda1'.
You are being dropped to a recovery shell
    Type 'exit' to try and continue booting
sh: can't access tty; job control turned off
[ramfs /]#
```

Causes potentielles

- Pilote du périphérique de stockage en mode bloc virtuel manquant ou configuré de façon incorrecte

- Conflit de l'énumération du périphérique (sda versus xvda)
- Choix incorrect du noyau de l'instance

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance. 2. Dissociez le volume. 3. Corrigez le problème du mappage du périphérique. 4. Démarrez l'instance. 5. Modifiez l'AMI pour traiter les problèmes du mappage du périphérique.
Basée sur le stockage d'instance	Utilisez la procédure suivante. 1. Créez une AMI avec la solution appropriée (mapper le périphérique de stockage en mode bloc correctement). 2. Arrêtez l'instance et lancez une nouvelle instance en utilisant l'AMI que vous avez créée.

... days without being checked, check forced (Contrôle du système de fichiers nécessaire)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
...
Checking filesystems
Checking all file systems.
[/sbin/fsck.ext3 (1) -- /] fsck.ext3 -a /dev/sda1
/dev/sda1 has gone 361 days without being checked, check forced
```

Causes potentielles

La durée de contrôle du système de fichiers est dépassée ; un contrôle du système de fichiers est en train d'être forcé.

Actions suggérées

- Patientez jusqu'à ce que le contrôle du système de fichiers se termine. Un contrôle de système de fichiers peut prendre longtemps en fonction de la taille du système de fichiers racine.
- Modifiez vos systèmes de fichiers pour supprimer l'application du contrôle du système de fichiers (fsck) en utilisant tune2fs ou des outils appropriés pour votre système de fichiers.

fsck died with exit status... (périphérique manquant)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
Cleaning up ifupdown....
Loading kernel modules...done.
...
Activating lvm and md swap...done.
Checking file systems...fsck from util-linux-ng 2.16.2
/sbin/fsck.xfs: /dev/sdh does not exist
fsck died with exit status 8
[31mfailed (code 8).[39;49m
```

Causes potentielles

- Ramdisk à la recherche d'un lecteur manquant
- Contrôle de cohérence forcé du système de fichiers
- Lecteur en échec ou détaché

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Essayez une ou plusieurs des solutions suivants pour résoudre le problème : • Arrêtez l'instance, attachez le volume à une instance existante en cours d'exécution. • Exécutez manuellement des contrôles de cohérence. • Corrigez le ramdisk pour inclure les utilitaires pertinents. • Modifiez les paramètres de réglage du système de fichiers pour supprimer les exigences de cohérence (non recommandé).
Basée sur le stockage d'instance	Essayez une ou plusieurs des solutions suivants pour résoudre le problème : • Regrouper le ramdisk avec les bons outils. • Modifiez les paramètres de réglage du système de fichiers pour supprimer les exigences de cohérence (non recommandé). • Mettez fin à l'instance et lancez une nouvelle instance. • (Facultatif) Demandez une assistance technique pour la récupération des données en utilisant AWS Support.

Invite GRUB (grubdom>)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
    GNU GRUB  version 0.97  (629760K lower / 0K upper memory)

       [ Minimal BASH-like line editing is supported.   For

         the   first   word,  TAB  lists  possible  command

         completions.  Anywhere else TAB lists the possible

         completions of a device/filename. ]

grubdom>
```

Causes potentielles

Type d'instance	Causes potentielles
Basée sur Amazon EBS	• Fichier de configuration GRUB manquant. • Mauvaise image GRUB utilisée ; fichier de configuration GRUB attendu à un emplacement différent. • Système de fichiers non pris en charge utilisé pour stoker votre fichier de configuration GRUB (par exemple, en transformant votre système de fichiers racine en un type qui n'est pas pris en charge par une version plus récente de GRUB).
Basée sur le stockage d'instance	• Fichier de configuration GRUB manquant. • Mauvaise image GRUB utilisée ; fichier de configuration GRUB attendu à un emplacement différent. • Système de fichiers non pris en charge utilisé pour stoker votre fichier de configuration GRUB (par exemple, en transformant votre système de fichiers racine en un type qui n'est pas pris en charge par une version plus récente de GRUB).

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Option 1 : Modifiez l'AMI et relancez l'instance : 1. Modifiez l'AMI source pour créer un fichier de configuration GRUB à l'emplacement standard (/boot/grub/menu.lst). 2. Vérifiez que votre version de GRUB prend en charge le type de système de fichiers sous-jacent et mettez à niveau GRUB si nécessaire.

Pour ce type d'instance	Faire ceci
	3. Choisissez l'image GRUB appropriée, (hd0 - 1er lecteur ou hd00 – 1er lecteur, première partition). 4. Arrêtez l'instance et lancez-en une nouvelle en utilisant l'AMI que vous avez créée. Option 2 : Corrigez l'instance existante: 1. Arrêtez l'instance. 2. Détachez le système de fichiers racine. 3. Attachez le système de fichiers racine à une instance connue en fonctionnement. 4. Montez le système de fichiers. 5. Créez un fichier de configuration GRUB. 6. Vérifiez que votre version de GRUB prend en charge le type de système de fichiers sous-jacent et mettez à niveau GRUB si nécessaire. 7. Détachez le système de fichiers. 8. Attachez-le à l'instance originale. 9. Modifiez l'attribut noyau pour utiliser l'image GRUB appropriée (1er disque ou 1ère partition sur 1er disque). 10. Démarrez l'instance.
Basée sur le stockage d'instance	Option 1 : Modifiez l'AMI et relancez l'instance : 1. Créez la nouvelle AMI avec un fichier de configuration GRUB à l'emplacement standard (/boot/grub/menu.lst). 2. Choisissez l'image GRUB appropriée, (hd0 - 1er lecteur ou hd00 – 1er lecteur, première partition). 3. Vérifiez que votre version de GRUB prend en charge le type de système de fichiers sous-jacent et mettez à niveau GRUB si nécessaire. 4. Arrêtez l'instance et lancez une nouvelle instance en utilisant l'AMI que vous avez créée. Option 2 : Arrêtez l'instance et lancez une nouvelle instance en spécifiant le noyau correct. Note Pour récupérer les données de l'instance existante, contactez AWS Support.

Bringing up interface eth0: Device eth0 has different MAC address than expected, ignoring. (Adresse MAC codée de manière irréversible)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
...
Bringing up loopback interface:  [  OK  ]

Bringing up interface eth0:  Device eth0 has different MAC address than expected, ignoring.
[FAILED]

Starting auditd: [  OK  ]
```

Causes potentielles

Il s'agit d'une interface MAC codée en dur dans la configuration d'AMI

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Effectuez l'une des actions suivantes : • Modifiez l'AMI pour supprimer le codage en dur et relancez l'instance. • Modifiez l'instance pour supprimer l'adresse MAC codée en dur. OU Utilisez la procédure suivante. 1. Arrêtez l'instance. 2. Détachez le volume racine. 3. Attachez le volume à une autre instance et modifiez le volume pour supprimer l'adresse MAC codée en dur. 4. Attachez le volume à l'instance originale. 5. Démarrez l'instance.
Basée sur le stockage d'instance	Effectuez l'une des actions suivantes : • Modifiez l'instance pour supprimer l'adresse MAC codée en dur. • Mettez fin à l'instance et lancez une nouvelle instance.

Unable to load SELinux Policy. Machine is in enforcing mode. Halting now. (Erreur de configuration SELinux)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
audit(1313445102.626:2): enforcing=1 old_enforcing=0 auid=4294967295
Unable to load SELinux Policy. Machine is in enforcing mode. Halting now.
Kernel panic - not syncing: Attempted to kill init!
```

Causes potentielles

SELinux a été activé par erreur :

- Le noyau fourni n'est pas pris en charge par GRUB
- Le noyau de rechange n'existe pas

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Utilisez la procédure suivante. 1. Arrêtez l'instance en échec. 2. Détachez le volume racine de l'instance en échec. 3. Attachez le volume racine à une autre instance Linux en fonctionnement (appelée plus tard instance de récupération). 4. Connectez-vous à l'instance de récupération et montez le volume racine de l'instance en échec. 5. Désactivez SELinux sur le volume racine monté. Ce processus varie selon les distributions Linux. Pour plus d'informations, consultez la documentation spécifique à votre système d'exploitation. Note Sur certaines systèmes, vous désactivez SELinux en réglant `SELINUX=disabled` dans le fichier `/mount_point/etc/sysconfig/selinux` où *`mount_point`* est l'emplacement où vous avez monté le volume sur votre instance de récupération. 6. Démontez et détachez le volume racine à partir de l'instance de récupération et attachez-le de nouveau à l'instance originale. 7. Démarrez l'instance.

Pour ce type d'instance	Faire ceci
Basée sur le stockage d'instance	Utilisez la procédure suivante. 1. Mettez fin à l'instance et lancez une nouvelle instance. 2. (Facultatif) Demandez une assistance technique pour la récupération des données en utilisant AWS Support.

XENBUS: Timeout connecting to devices (délai d'attente Xenbus)

Cette condition est indiquée par une entrée du journal du système qui est similaire à celle indiquée ci-dessous.

```
Linux version 2.6.16-xenU (builder@xenbat.amazonsa) (gcc version 4.0.1
20050727 (Red Hat 4.0.1-5)) #1 SMP Mon May 28 03:41:49 SAST 2007
...
XENBUS: Timeout connecting to devices!
...
Kernel panic - not syncing: No init found.  Try passing init= option to kernel.
```

Causes potentielles

- Le périphérique de stockage en mode bloc n'est pas connecté à l'instance
- Cette instance utilise un ancien noyau de l'instance

Actions suggérées

Pour ce type d'instance	Faire ceci
Basée sur Amazon EBS	Effectuez l'une des actions suivantes : • Modifiez l'AMI et l'instance pour utiliser un noyau moderne et relancez l'instance. • Redémarrez l'instance.
Basée sur le stockage d'instance	Effectuez l'une des actions suivantes : • Mettez fin à l'instance. • Modifiez l'AMI pour utiliser un noyau moderne et lancez une nouvelle instance en utilisant cette AMI.

Résolution d'un problème d'instance inaccessible

Vous pouvez utiliser les méthodes suivantes pour résoudre un problème d'instance inaccessible.

Sommaire

- Redémarrage d'instance (p. 1361)
- Sortie de la console de l'instance (p. 1361)
- Création d'une capture d'écran d'une instance inaccessible (p. 1362)
- Récupération d'instance en cas de plantage de l'ordinateur hôte (p. 1363)

Redémarrage d'instance

La capacité de redémarrer des instances qui sont généralement inaccessibles est précieuse pour le dépannage et la gestion générale des instances.

Tout comme vous pouvez réinitialiser un ordinateur en appuyant sur le bouton approprié, vous pouvez réinitialiser les instances EC2 en utilisant la console, l'interface ligne de commande ou l'API d'Amazon EC2. Pour plus d'informations, consultez Redémarrer votre instance (p. 641).

> Warning
>
> Pour les instances Windows, cette opération effectue un redémarrage matériel qui peut entraîner une corruption des données.

Sortie de la console de l'instance

La sortie de la console est un outil de valeur pour le diagnostic des problèmes. Elle est particulièrement utile pour la résolution des problèmes liés au noyau et à la configuration des services qui pourraient mettre fin à une instance ou la rendre inaccessible avant que son programme fantôme SSH ne puisse être démarré.

Pour Linux/Unix, la sortie de la console de l'instance affiche la sortie de la console exacte que serait normalement affichée sur un écran physique attaché à un ordinateur. La sortie de la console renvoie des informations mises en mémoire tampon qui ont été publiées après un état de transition d'instance (démarrage, arrêt, redémarrage et résiliation). La sortie publiée n'est pas continuellement mise à jour, uniquement lorsqu'elle est probablement très bénéfique.

Pour les instances Windows, la sortie de la console de l'instance affiche les trois dernières erreurs du journal des événements du système.

Vous pouvez éventuellement extraire la dernière sortie de console série à tout moment au cours du cycle de vie de l'instance. Cette option est prise en charge uniquement sur Instances reposant sur le système Nitro (p. 218). Elle n'est pas prise en charge via la console Amazon EC2.

> Note
>
> Seuls les 64 Ko les plus récents de la sortie publiée sont stockés et disponibles pendant au moins 1 heure après la dernière publication.

Seul le propriétaire de l'instance peut accéder à la sortie de la console. Vous pouvez supprimer la sortie de la console de vos instances en utilisant la console ou la ligne de commande.

Pour obtenir la sortie de la console en utilisant la console

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation sur la gauche, choisissez Instances, puis sélectionnez l'instance.
3. Choisissez Actions, Paramètres de l'instance, Obtenir le journal système.

Pour obtenir la sortie de la console en utilisant la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- get-console-output (AWS CLI)
- Get-EC2ConsoleOutput (Outils AWS pour Windows PowerShell)

Pour plus d'informations sur les erreurs communes du journal du système, consultez le didacticiel Dépannage des erreurs du journal du système pour les instances basées sur Linux (p. 1340).

Création d'une capture d'écran d'une instance inaccessible

Si vous ne pouvez pas joindre votre instance via SSH ou RDP, vous pouvez créer une capture d'écran de l'instance et l'afficher sous forme d'image. Cette image permet de voir le statut de l'instance et de résoudre le problème plus rapidement. Vous pouvez générer des captures d'écran pendant que l'instance s'exécute ou après son blocage. Aucun coût de transfert de données n'est facturé pour cette capture d'écran. L'image est générée au format JPG et ne dépasse pas 100 Ko. Cette fonctionnalité n'est pas prise en charge lorsque l'instance utilise un pilote NVIDIA GRID, est sur des instances nue (instances de type `*.metal`) ou est basée sur des processeurs Graviton ou Graviton 2 basés sur ARM. Cette fonction est disponible dans les régions suivantes :

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

Pour accéder à la console de l'instance

1. Ouvrez la console Amazon EC2 à l'adresse https://console.aws.amazon.com/ec2/.
2. Dans le panneau de navigation de gauche, choisissez Instances.
3. Sélectionnez l'instance à capturer.
4. Choisissez Actions, Paramètres de l'instance.
5. Sélectionnez Obtenir la capture de l'instance.

Cliquez sur l'image avec le bouton droit de la souris pour la télécharger et l'enregistrer.

Pour créer un instantané via la ligne de commande

Vous pouvez utiliser l'une des commandes suivantes. Le contenu renvoyé est codé en base64. Pour obtenir plus d'informations sur les CLI (interface ligne de commande), consultez le didacticiel Accès à Amazon EC2 (p. 3).

- get-console-screenshot (AWS CLI)
- GetConsoleScreenshot (API de requête Amazon EC2)

Récupération d'instance en cas de plantage de l'ordinateur hôte

S'il existe un problème irrécupérable lié au matériel d'un ordinateur hôte sous-jacent, AWS peut planifier un évènement d'arrêt d'instance. Vous êtes averti d'un tel événement en avance par e-mail.

Pour récupérer une instance basée sur Amazon EBS en cours d'exécution sur un ordinateur hôte qui a planté

1. Sauvegardez les données importantes qui se trouvent sur les volumes de stockage d'instance sur Amazon EBS ou Amazon S3.
2. Arrêtez l'instance.
3. Démarrez l'instance.
4. Restaurez toutes les données importantes.

Pour plus d'informations, consultez Arrêt et démarrage de votre instance (p. 625).

Pour récupérer une instance basée sur le stockage d'instance et exécutée sur un ordinateur hôte qui a planté

1. Créez une AMI à partir de l'instance.
2. Chargez l'image vers Amazon S3.
3. Sauvegardez les données importantes sur Amazon EBS ou Amazon S3.
4. Mettez fin à l'instance.
5. Lancez une nouvelle instance depuis l'AMI.
6. Restaurez toutes les données importantes sur la nouvelle instance.

Pour plus d'informations, consultez Création d'une AMI Linux basée sur le stockage d'instance (p. 136).

Démarrage à partir du mauvais volume

Dans certaines situations, vous pouvez trouver qu'un volume autre que celui attaché à `/dev/xvda` ou `/dev/sda` est devenu le volume racine de votre instance. Cela peut arriver lorsque vous avez attaché le volume racine d'une autre instance, ou un volume créé à partir de l'instantané d'un volume racine, à une instance avec un volume racine existant.

Ceci est dû à la façon de fonctionner du ramdisk initial dans Linux. Il choisit le volume définit comme `/` dans le fichier `/etc/fstab`, et dans certaines distributions. Ceci est déterminé par l'étiquette attachée à la partition du volume. Plus spécifiquement, vous trouvez que le fichier `/etc/fstab` ressemble à ce qui suit :

```
LABEL=/ / ext4 defaults,noatime 1 1
tmpfs /dev/shm tmpfs defaults 0 0
devpts /dev/pts devpts gid=5,mode=620 0 0
sysfs /sys sysfs defaults 0 0
proc /proc proc defaults 0 0
```

Si vous vérifiez l'étiquette des deux volumes, vous verrez qu'ils contiennent tous les deux l'étiquette `/` :

```
[ec2-user ~]$ sudo e2label /dev/xvda1
/
[ec2-user ~]$ sudo e2label /dev/xvdf1
/
```

Dans cet exemple, `/dev/xvdf1` pourrait devenir le périphérique racine où votre instance démarre après l'exécution initiale de ramdisk, au lieu du volume `/dev/xvda1` à partir duquel vous aviez essayé de démarrer. Pour résoudre ce problème, utilisez la même commande e2label pour changer l'étiquette du volume attaché à partir duquel vous ne souhaitez pas démarrer.

Dans certains cas, spécifier un UUID dans `/etc/fstab` peut résoudre ce problème. Cependant, si les deux volumes proviennent du même instantané ou si le deuxième est créé à partir d'un instantané du volume principal, ils partagent un UUID.

```
[ec2-user ~]$ sudo blkid
/dev/xvda1: LABEL="/" UUID=73947a77-ddbe-4dc7-bd8f-3fe0bc840778 TYPE="ext4"
 PARTLABEL="Linux" PARTUUID=d55925ee-72c8-41e7-b514-7084e28f7334
/dev/xvdf1: LABEL="old/" UUID=73947a77-ddbe-4dc7-bd8f-3fe0bc840778 TYPE="ext4"
 PARTLABEL="Linux" PARTUUID=d55925ee-72c8-41e7-b514-7084e28f7334
```

Pour changer l'étiquette d'un volume ext4 attaché

1. Utilisez la commande e2label pour remplacer l'étiquette du volume par autre chose que `/`.

   ```
   [ec2-user ~]$ sudo e2label /dev/xvdf1 old/
   ```

2. Vérifiez que le volume possède la nouvelle étiquette.

   ```
   [ec2-user ~]$ sudo e2label /dev/xvdf1
   old/
   ```

Pour changer l'étiquette d'un volume xfs attaché

- Utilisez la commande xfs_admin pour remplacer l'étiquette du volume par autre chose que `/`.

  ```
  [ec2-user ~]$ sudo xfs_admin -L old/ /dev/xvdf1
  writing all SBs
  new label = "old/"
  ```

Après avoir modifié l'étiquette du volume comme indiqué, vous devriez pouvoir redémarrer l'instance et avoir le bon volume sélectionné par le ramdisk initial lorsque l'instance démarre.

> **Important**
>
> Si vous prévoyez de détacher le volume avec la nouvelle étiquette et de le renvoyer vers une autre instance pour l'utiliser comme volume racine, vous devez ré-exécuter la procédure ci-dessus et réattribuer à l'étiquette du volume sa valeur d'origine. Sinon, l'autre instance ne démarre pas, car le ramdisk ne peut pas trouver le volume avec l'étiquette `/`.

Utilisation d'EC2Rescue pour Linux

EC2Rescue pour Linux est un outil open source, facile d'utilisation, qui peut être exécuté sur une instance Amazon EC2 Linux pour diagnostiquer et résoudre des problèmes courants à l'aide de sa bibliothèque de

plus de 100 modules. Quelques cas d'utilisation généralisés pour EC2Rescue pour Linux incluent le recueil du journal système et des journaux du gestionnaire de package, la collecte des données d'utilisation des ressources, ainsi que le diagnostic/la résolution des paramètres de noyau problématiques connus et des problèmes OpenSSH courants.

Si vous utilisez une instance Windows, consultez EC2Rescue pour Windows Server.

Table des matières

- Installation d'EC2Rescue pour Linux (p. 1365)
- Utilisation d'EC2Rescue pour Linux (p. 1368)
- Développement des modules EC2Rescue (p. 1370)

Installation d'EC2Rescue pour Linux

L'outil EC2Rescue pour Linux peut être installé sur une instance Amazon EC2 Linux qui satisfait les prérequis suivants.

Prérequis

- Systèmes d'exploitation pris en charge :
  - Amazon Linux 2
  - Amazon Linux 2016.09+
  - SUSE Linux Enterprise Server 12+
  - RHEL 7+
  - Ubuntu 16.04+
- Configuration logicielle requise :
  - Python 2.7.9+ ou 3.2+

Si votre système dispose de la version Python requise, vous pouvez installer la build standard. Dans le cas contraire, vous pouvez installer la build de la solution groupée, qui inclut une copie minimale de Python.

Pour installer la build standard

1. Sur une instance Linux active, téléchargez l'outil EC2Rescue pour Linux :

```
curl -O https://s3.amazonaws.com/ec2rescuelinux/ec2rl.tgz
```

2. (Facultatif) Avant de poursuivre, vous pouvez vérifier la signature du fichier d'installation EC2Rescue pour Linux. Pour plus d'informations, consultez (Facultatif) Vérification de la signature de EC2Rescue pour Linux (p. 1366).
3. Téléchargez le fichier de hachage sha256 :

```
curl -O https://s3.amazonaws.com/ec2rescuelinux/ec2rl.tgz.sha256
```

4. Vérifiez l'intégrité du tarball :

```
sha256sum -c ec2rl.tgz.sha256
```

5. Déballez le tarball :

```
tar -xvf ec2rl.tgz
```

6. Vérifiez l'installation en énumérant le fichier d'aide :

```
cd ec2rl-<version_number>
./ec2rl help
```

Pour installer la build de la solution groupée

Pour obtenir un lien jusqu'à la page de téléchargement et la liste des restrictions, consultez EC2Rescue pour Linux sur github.

(Facultatif) Vérification de la signature de EC2Rescue pour Linux

La procédure suivante est celle recommandée pour vérifier la validité du package EC2Rescue pour Linux pour les systèmes d'exploitation Linux.

Lorsque vous téléchargez une application à partir d'Internet, nous vous recommandons d'authentifier l'identité de l'éditeur du logiciel et de vérifier que l'application n'a pas été modifiée ou corrompue depuis sa publication. Cela vous évitera d'installer une version de l'application contenant un virus ou tout autre code malveillant.

Si, après l'exécution de la procédure décrite dans cette rubrique, vous déterminez que le logiciel de EC2Rescue pour Linux a été modifié ou corrompu, n'exécutez pas le fichier d'installation. Contactez plutôt Amazon Web Services.

Les fichiers EC2Rescue pour Linux pour les systèmes d'exploitation Linux sont signés à l'aide de GnuPG, une mise en œuvre Open Source de la norme Pretty Good Privacy (OpenPGP) pour les signatures numériques sécurisées. GnuPG (également appelée GPG) assure l'authentification et le contrôle d'intégrité par le biais d'une signature numérique. AWS publie une clé publique et des signatures que vous pouvez utiliser pour vérifier le package EC2Rescue pour Linux téléchargé. Pour plus d'informations sur PGP et GnuPG (GPG), consultez http://www.gnupg.org.

La première étape consiste à établir une approbation avec l'éditeur du logiciel. Téléchargez la clé publique de l'éditeur du logiciel, vérifiez que le propriétaire de cette clé publique est bien celui qu'il prétend être, puis ajoutez la clé publique à votre porte-clés. Votre porte-clés est un ensemble de clés publiques connues. Après avoir établi l'authenticité de la clé publique, vous pouvez l'utiliser pour vérifier la signature de l'application.

Tâches

- Installation des outils GPG (p. 1366)
- Authentification et importation de la clé publique (p. 1367)
- Vérification de la signature du package (p. 1367)

Installation des outils GPG

Si votre système d'exploitation est Linux ou Unix, les outils GPG peuvent déjà être installés. Pour tester si les outils sont installés sur votre système, entrez gpg2 à partir d'une invite de commande. Si les outils GPG sont installés, une invite de commande GPG s'affiche. Si les outils GPG ne sont pas installés, vous voyez une erreur indiquant que la commande est introuvable. Vous pouvez installer le package GnuPG à partir d'un référentiel.

Pour installer les outils GPG sur un système Linux basé sur Debian

- Depuis un terminal, exécutez la commande suivante :

```
apt-get install gnupg2
```

Pour installer les outils GPG sur un système Linux basé sur Red Hat

- Depuis un terminal, exécutez la commande suivante :

```
yum install gnupg2
```

Authentification et importation de la clé publique

L'étape suivante consiste à authentifier la clé publique EC2Rescue pour Linux et à l'ajouter en tant que clé de confiance à votre porte-clés GPG.

Pour authentifier et importer la clé publique EC2Rescue pour Linux

1. À l'invite de commande, utilisez la commande suivante pour obtenir une copie de votre clé publique GPG :

```
curl -O https://s3.amazonaws.com/ec2rescuelinux/ec2rl.key
```

2. À l'invite de commande, dans le répertoire où vous avez enregistré `ec2rl.key`, exécutez la commande suivante pour importer la clé publique EC2Rescue pour Linux dans votre porte-clés :

```
gpg2 --import ec2rl.key
```

La commande renvoie un résultat semblable à ce qui suit :

```
gpg: /home/ec2-user/.gnupg/trustdb.gpg: trustdb created
gpg: key 2FAE2A1C: public key "ec2autodiag@amazon.com <EC2 Rescue for Linux>" imported
gpg: Total number processed: 1
gpg:               imported: 1  (RSA: 1)
```

Vérification de la signature du package

Après avoir installé les outils GPG, authentifié et importé la clé publique d'EC2Rescue pour Linux et vérifié que la clé publique d'EC2Rescue pour Linux est approuvée, vous êtes prêt à vérifier la signature du script d'installation d'EC2Rescue pour Linux.

Pour vérifier la signature du script d'installation EC2Rescue pour Linux

1. À l'invite de commande, exécutez la commande suivante pour télécharger le fichier signature du script d'installation :

```
curl -O https://s3.amazonaws.com/ec2rescuelinux/ec2rl.tgz.sig
```

2. Vérifiez la signature en exécutant la commande suivante à l'invite de commande dans le répertoire où vous avez enregistré `ec2rl.tgz.sig` et le fichier d'installation d'EC2Rescue pour Linux. Ces deux fichiers doivent être présents.

```
gpg2 --verify ./ec2rl.tgz.sig
```

Le résultat doit ressembler à ce qui suit :

```
gpg: Signature made Thu 12 Jul 2018 01:57:51 AM UTC using RSA key ID 6991ED45
gpg: Good signature from "ec2autodiag@amazon.com <EC2 Rescue for Linux>"
gpg: WARNING: This key is not certified with a trusted signature!
gpg:          There is no indication that the signature belongs to the owner.
Primary key fingerprint: E528 BCC9 0DBF 5AFA 0F6C  C36A F780 4843 2FAE 2A1C
     Subkey fingerprint: 966B 0D27 85E9 AEEC 1146  7A9D 8851 1153 6991 ED45
```

Si le résultat contient l'expression `Good signature from "ec2autodiag@amazon.com <EC2 Rescue for Linux>"`, cela signifie que la signature a été vérifiée et vous pouvez continuer à exécuter le script d'installation d'EC2Rescue pour Linux.

Si le résultat inclut l'expression `BAD signature`, vérifiez si vous avez effectué la procédure correctement. Si vous continuez à obtenir cette réponse, contactez Amazon Web Services et n'exécutez pas le fichier d'installation que vous avez précédemment téléchargé.

Voici les informations détaillées sur les avertissements qui peuvent s'afficher :

- WARNING: This key is not certified with a trusted signature! There is no indication that the signature belongs to the owner.Ce message fait référence à votre niveau de confiance personnel dans la conviction que vous possédez une clé publique authentique pour EC2Rescue pour Linux. Dans un monde idéal, vous visiteriez un bureau Amazon Web Services et recevriez la clé en personne. Cependant, vous la téléchargez le plus souvent à partir d'un site Web. Dans le cas présent, le site Web est un site Amazon Web Services.
- gpg2: no ultimately trusted keys found. Cela signifie que la clé spécifique n'est pas « approuvée en dernier lieu » par vous-même (ou par d'autres personnes de confiance).

Pour plus d'informations, consultez http://www.gnupg.org.

Utilisation d'EC2Rescue pour Linux

Les tâches suivantes sont des tâches courantes que vous pouvez effectuer pour commencer à utiliser cet outil.

Tâches

- Exécution d'EC2Rescue pour Linux (p. 1368)
- Téléchargement des résultats (p. 1369)
- Création de sauvegardes (p. 1369)
- Obtenir de l'aide (p. 1370)

Exécution d'EC2Rescue pour Linux

Vous pouvez exécuter EC2Rescue pour Linux, comme illustré dans les exemples suivants..

Example Exemple : Exécuter tous les modules

Pour exécuter tous les modules, exécutez EC2Rescue pour Linux sans options :

```
./ec2rl run
```

Certains modules nécessitent un accès racine. Si vous n'êtes pas un utilisateur racine, utilisez sudo pour exécuter ces modules comme suit :

```
sudo ./ec2rl run
```

Example Exemple : Exécuter un module spécifique

Pour exécuter uniquement des modules spécifiques, utilisez le paramètre --only-modules :

```
./ec2rl run --only-modules=module_name --arguments
```

Par exemple, cette commande exécute le module dig pour interroger le domaine `amazon.com` :

```
./ec2rl run --only-modules=dig --domain=amazon.com
```

Example Exemple : Afficher les résultats

Vous pouvez afficher les résultats dans `/var/tmp/ec2rl` :

```
cat /var/tmp/ec2rl/logfile_location
```

Par exemple, affichez le fichier journal pour le module dig :

```
cat /var/tmp/ec2rl/2017-05-11T15_39_21.893145/mod_out/run/dig.log
```

Téléchargement des résultats

Si AWS Support a demandé les résultats ou pour partager les résultats à partir d'un compartiment S3, téléchargez-les à l'aide de l'outil d'interface de ligne de commande EC2Rescue pour Linux. La sortie des commandes EC2Rescue pour Linux doit fournir les commandes que vous avez besoin d'utiliser.

Example Exemple : Charger les résultats dans AWS Support

```
./ec2rl upload --upload-directory=/var/tmp/ec2rl/2017-05-11T15_39_21.893145 --support-
url="URLProvidedByAWSSupport"
```

Example Exemple : Charger les résultats dans un compartiment S3

```
./ec2rl upload --upload-directory=/var/tmp/ec2rl/2017-05-11T15_39_21.893145 --presigned-
url="YourPresignedS3URL"
```

Pour plus d'informations sur la génération d'URL pré-signées pour Amazon S3, consultez Téléchargement d'objets utilisant des URL pré-signées.

Création de sauvegardes

Créez une sauvegarde pour votre instance, un ou plusieurs volumes, ou un ID d'appareil spécifique à l'aide des commandes suivantes.

Example Exemple : Sauvegarder une instance avec une image machine Amazon (IMA)

```
./ec2rl run --backup=ami
```

Example Exemple : Sauvegarder tous les volumes associés à l'instance

```
./ec2rl run --backup=allvolumes
```

Example Exemple : Sauvegarder un volume spécifique

```
./ec2rl run --backup=volumeID
```

Obtenir de l'aide

EC2Rescue pour Linux inclut un fichier d'aide qui vous fournit des informations et la syntaxe pour chaque commande disponible.

Example Exemple : Afficher l'aide générale

```
./ec2rl help
```

Example Exemple : Répertorier les modules disponibles

```
./ec2rl list
```

Example Exemple : Afficher l'aide pour un module spécifique

```
./ec2rl help module_name
```

Par exemple, utilisez la commande suivante pour afficher le fichier d'aide pour le module dig :

```
./ec2rl help dig
```

Développement des modules EC2Rescue

Les modules sont écrits en YAML, une norme de sérialisation des données. Le fichier YAML d'un module se compose d'un document unique, qui représente le module et ses attributs.

Ajout d'attributs de modules

Le tableau suivant répertorie les attributs de module disponibles.

Attribut	Description
name	Nom du module. La longueur du nom doit être inférieure ou égale à 18 caractères.
Version	Numéro de version du module.
title	Titre court et descriptif du module. La longueur de cette valeur doit être inférieure ou égale à 50 caractères.
helptext	Description étendue du module. La longueur de chaque ligne doit être inférieure ou égale à 75 caractères. Si le module consomme des arguments, obligatoires ou facultatifs, incluez-les dans la valeur helptext. Par exemple : `helptext: !!str |`

Attribut	Description

	`Collect output from ps for system analysis`
	`Consumes --times= for number of times to repeat`
	`Consumes --period= for time period between repetition`

placement	Étape dans laquelle le module doit être exécuté. Valeurs prises en charge : • prediagnostic • run • postdiagnostic
langage	Langage dans lequel le code du module est écrit. Valeurs prises en charge : • bash • python Note Le code Python doit être compatible avec Python 2.7.9+ et Python 3.2+.
remediation	Indique si le module prend en charge la correction. Les valeurs prises en charge sont `True` ou `False`. Le module utilise par défaut `False` si aucune valeur n'est indiquée. Il s'agit donc d'un attribut facultatif pour les modules qui ne prennent pas en charge la correction.
content	Intégralité du code de script.
contrainte	Nom de l'objet contenant les valeurs de contrainte.
domaine	Descripteur de la façon dont le module est regroupé ou classé. L'ensemble de modules inclus utilise les domaines suivants : • application • net • os • performances
class	Descripteur du type de tâche effectué par le module. L'ensemble de modules inclus utilise les classes suivantes : • collect (collecte la sortie des programmes) • diagnose (réussite/échec en fonction d'un ensemble de critères) • gather (copie les fichiers et écrit dans un fichier spécifique)

Attribut	Description
distro	Liste des distributions Linux que ce module prend en charge. L'ensemble de modules inclus utilise les distributions suivantes : • alami (Amazon Linux) • rhel • ubuntu • suse
obligatoire	Arguments obligatoires que le module consomme à partir des options de l'interface de ligne de commande.
facultatif	Arguments facultatifs que le module peut utiliser.
logiciel	Exécutables logiciels utilisés dans le module. Cet attribut a pour but de spécifier un logiciel qui n'est pas installé par défaut. La logique EC2Rescue pour Linux s'assure que ces programmes sont présents et exécutables avant l'exécution du module.
package	Package logiciel source pour un exécutable. Cet attribut a pour but de fournir des détails étendus sur le package avec le logiciel, y compris une URL pour le téléchargement ou pour obtenir de plus amples informations.
sudo	Indique si l'accès racine est obligatoire pour exécuter le module. Vous n'avez pas besoin d'implémenter des vérifications sudo dans le script du module. Si la valeur est true, la logique EC2Rescue pour Linux exécute uniquement le module lorsque l'utilisateur qui exécute le module possède un accès racine.
perfimpact	Indique si le module peut avoir un impact important sur les performances qui affecte l'environnement dans lequel il est exécuté. Si la valeur est true et que l'argument `--perfimpact=true` n'est pas présent, le module est ignoré.
parallelexclusive	Spécifie un programme qui requiert une exclusivité mutuelle. Par exemple, tous les modules qui spécifient « bpf » sont exécutés en série.

Ajout de variables d'environnement

Le tableau suivant répertorie les variables d'environnement disponibles.

Variable d'environnement	Description
`EC2RL_CALLPATH`	Chemin vers `ec2rl.py`. Ce chemin peut être utilisé pour localiser le répertoire lib et utiliser les modules Python fournis.
`EC2RL_WORKDIR`	Répertoire tmp principal pour l'outil de diagnostic. Valeur par défaut: `/var/tmp/ec2rl`.
`EC2RL_RUNDIR`	Répertoire dans lequel toutes les sorties sont stockées. Valeur par défaut: `/var/tmp/ec2rl/<date&timestamp>`.
`EC2RL_GATHEREDDIR`	Répertoire racine dans lequel placer les données collectées sur le module. Valeur par défaut:`/var/tmp/ec2rl/<date&timestamp>/mod_out/gathered/`.
`EC2RL_NET_DRIVER`	Pilote utilisé pour la première interface réseau non virtuelle, triée par ordre alphabétique, de l'instance. Exemples : • xen_netfront • ixgbevf • ena
`EC2RL_SUDO`	True si EC2Rescue pour Linux est en cours d'exécution en tant que racine ; sinon la valeur est false.
`EC2RL_VIRT_TYPE`	Type de virtualisation, tel que fourni par les métadonnées d'instance. Exemples : • default-hvm • default-paravirtual
`EC2RL_INTERFACES`	Liste énumérée des interfaces du système. La valeur est une chaîne contenant des noms, tels que `eth0`, `eth1`, etc. Elle est générée via `functions.bash` et est disponible uniquement pour les modules dont elle provient.

Utilisation de la syntaxe YAML

Tenez compte des points suivants lorsque vous créez vos fichiers YAML de module :

- Le triple trait d'union (`---`) indique le début explicite d'un document.
- La balise `!ec2rlcore.module.Module` indique à l'analyseur YAML le constructeur à appeler lors de la création de l'objet à partir du flux de données. Vous trouverez le constructeur dans le fichier `module.py`.

- La balise `!!str` indique à l'analyseur YAML de ne pas tenter de déterminer le type des données, et d'interpréter plutôt le contenu comme un littéral de chaîne.
- Le caractère pipe (`|`) indique à l'analyseur YAML que la valeur est une scalaire littérale. Dans ce cas, l'analyseur inclut tous les espaces. C'est important pour les modules car les caractères de mise en retrait et de saut de ligne sont conservés.
- La mise en retrait standard YAML correspond à deux espaces, comme illustré dans les deux exemples suivants. Veillez à conserver la mise en retrait standard (par exemple, quatre espaces pour Python) pour votre script, puis mettez en retrait l'intégralité du contenu à l'aide de deux espaces dans le fichier du module.

Exemples de modules

Example un (`mod.d/ps.yaml`) :

```
--- !ec2rlcore.module.Module
# Module document. Translates directly into an almost-complete Module object
name: !!str ps
path: !!str
version: !!str 1.0
title: !!str Collect output from ps for system analysis
helptext: !!str |
  Collect output from ps for system analysis
  Requires --times= for number of times to repeat
  Requires --period= for time period between repetition
placement: !!str run
package:
  - !!str
language: !!str bash
content: !!str |
  #!/bin/bash
  error_trap()
  {
      printf "%0.s=" {1..80}
      echo -e "\nERROR: "$BASH_COMMAND" exited with an error on line ${BASH_LINENO[0]}"
      exit 0
  }
  trap error_trap ERR

  # read-in shared function
  source functions.bash
  echo "I will collect ps output from this $EC2RL_DISTRO box for $times times every $period
 seconds."
  for i in $(seq 1 $times); do
      ps auxww
      sleep $period
  done
constraint:
  requires_ec2: !!str False
  domain: !!str performance
  class: !!str collect
  distro: !!str alami ubuntu rhel suse
  required: !!str period times
  optional: !!str
  software: !!str
  sudo: !!str False
  perfimpact: !!str False
  parallelexclusive: !!str
```

Envoi d'une interruption de diagnostic (utilisateurs avancés uniquement)

> **Warning**
>
> Les interruptions de diagnostic sont conçues pour être utilisées par des utilisateurs avancés. Une utilisation incorrecte peut avoir un impact négatif sur votre instance. L'envoi d'une interruption de diagnostic à une instance peut déclencher le blocage ou le redémarrage de l'instance, ce qui peut se traduire par une perte de données.

Vous pouvez envoyer une interruption de diagnostic à une instance Linux inaccessible ou qui ne répond pas afin de déclencher manuellement une panique de noyau.

Les systèmes d'exploitation Linux tombent généralement en panne et redémarrent en cas de panique de noyau. Le comportement spécifique du système d'exploitation dépend de sa configuration. Vous pouvez aussi utiliser une panique de noyau pour que le noyau système du système d'exploitation de l'instance effectue des tâches telles que la génération d'un fichier de vidage sur incident. Vous pouvez alors utiliser les informations du fichier de vidage sur incident pour effectuer l'analyse de la cause de la panne et le débogage de l'instance.

Les données de vidage sur incident sont générées localement par le système d'exploitation sur l'instance elle-même.

Avant d'envoyer une interruption de diagnostic à votre instance, nous vous recommandons de consulter la documentation pour votre système d'exploitation et d'effectuer les modifications de configuration nécessaires.

Sommaire

- Types d'instance pris en charge (p. 1375)
- Prérequis (p. 1375)
- Envoi d'une interruption de diagnostic (p. 1378)

Types d'instance pris en charge

L'interruption de diagnostic est prise en charge sur tous les types d'instance basés sur Nitro, sauf A1. Pour plus d'informations, consultez Instances reposant sur le système Nitro (p. 218).

Prérequis

Avant d'utiliser une interruption de diagnostic, vous devez configurer le système d'exploitation de votre instance. Cela permet de s'assurer qu'elle effectuera les actions dont vous avez besoin en cas de panique de noyau.

Pour configurer Amazon Linux 2 pour générer un vidage sur incident en cas de panique de noyau

1. Connectez-vous à votre instance.
2. Installez kexec et kdump.

```
[ec2-user ~]$ sudo yum install kexec-tools -y
```

3. Configurez le noyau afin qu'il réserve une quantité appropriée de mémoire pour le noyau secondaire. La quantité de mémoire à réserver dépend de la quantité de mémoire totale disponible de votre

instance. Ouvrez le fichier `/etc/default/grub` à l'aide de votre éditeur de texte préféré, localisez la ligne commençant par `GRUB_CMDLINE_LINUX_DEFAULT`, puis ajoutez le paramètre `crashkernel` au format suivant : `crashkernel=`*`memory_to_reserve`*. Par exemple, pour réserver `160MB`, modifiez le fichier `grub` comme suit :

```
GRUB_CMDLINE_LINUX_DEFAULT="crashkernel=160M console=tty0 console=ttyS0,115200n8
 net.ifnames=0 biosdevname=0 nvme_core.io_timeout=4294967295 rd.emergency=poweroff
 rd.shell=0"
GRUB_TIMEOUT=0
GRUB_DISABLE_RECOVERY="true"
```

4. Enregistrez les modifications, puis fermez le fichier `grub`.
5. Générez à nouveau le fichier de configuration GRUB2.

```
[ec2-user ~]$ sudo grub2-mkconfig -o /boot/grub2/grub.cfg
```

6. Sur les instances basées sur les processeurs Intel et AMD, la commande `send-diagnostic-interrupt` envoie une interruption non masquable (NMI) inconnue à l'instance. Vous devez configurer le noyau pour tomber en panne lorsqu'il reçoit l'interruption NMI inconnue. Ouvrez le fichier `/etc/sysctl.conf` à l'aide de l'éditeur de texte de votre choix et ajoutez ce qui suit.

```
kernel.unknown_nmi_panic=1
```

7. Redémarrez votre instance et reconnectez-la.
8. Vérifiez que le noyau a été démarré avec le paramètre `crashkernel` correct.

```
$ grep crashkernel /proc/cmdline
```

L'exemple de sortie suivant illustre une configuration réussie.

```
BOOT_IMAGE=/boot/vmlinuz-4.14.128-112.105.amzn2.x86_64 root=UUID=a1e1011e-
e38f-408e-878b-fed395b47ad6 ro crashkernel=160M console=tty0 console=ttyS0,115200n8
 net.ifnames=0 biosdevname=0 nvme_core.io_timeout=4294967295 rd.emergency=poweroff
 rd.shell=0
```

9. Vérifiez que le service kdump est en cours d'exécution.

```
[ec2-user ~]$ systemctl status kdump.service
```

L'exemple de sortie suivant présente le résultat lorsque le service kdump est en cours d'exécution.

```
kdump.service - Crash recovery kernel arming
   Loaded: loaded (/usr/lib/systemd/system/kdump.service; enabled; vendor preset:
 enabled)
   Active: active (exited) since Fri 2019-05-24 23:29:13 UTC; 22s ago
  Process: 2503 ExecStart=/usr/bin/kdumpctl start (code=exited, status=0/SUCCESS)
 Main PID: 2503 (code=exited, status=0/SUCCESS)
```

Note

Par défaut, le fichier de vidage sur incident est enregistré dans `/var/crash/`. Pour modifier cet emplacement, modifiez le fichier `/etc/kdump.conf` à l'aide de l'éditeur de texte de votre choix.

Pour configurer Amazon Linux pour générer un vidage sur incident en cas de panique de noyau

1. Connectez-vous à votre instance.

2. Installez kexec et kdump.

```
[ec2-user ~]$ sudo yum install kexec-tools -y
```

3. Configurez le noyau afin qu'il réserve une quantité appropriée de mémoire pour le noyau secondaire. La quantité de mémoire à réserver dépend de la quantité de mémoire totale disponible de votre instance.

```
$ sudo grubby --args="crashkernel=memory_to_reserve" --update-kernel=ALL
```

Par exemple, pour réserver `160MB` pour le noyau d'incident, utilisez la commande qui suit.

```
$ sudo grubby --args="crashkernel=160M" --update-kernel=ALL
```

4. Sur les instances basées sur les processeurs Intel et AMD, la commande `send-diagnostic-interrupt` envoie une interruption non masquable (NMI) inconnue à l'instance. Vous devez configurer le noyau pour tomber en panne lorsqu'il reçoit l'interruption NMI inconnue. Ouvrez le fichier `/etc/sysctl.conf` à l'aide de l'éditeur de texte de votre choix et ajoutez ce qui suit.

```
kernel.unknown_nmi_panic=1
```

5. Redémarrez votre instance et reconnectez-la.
6. Vérifiez que le noyau a été démarré avec le paramètre `crashkernel` correct.

```
$ grep crashkernel /proc/cmdline
```

L'exemple de sortie suivant illustre une configuration réussie.

```
root=LABEL=/ console=tty1 console=ttyS0 selinux=0 nvme_core.io_timeout=4294967295
 LANG=en_US.UTF-8 KEYTABLE=us crashkernel=160M
```

7. Vérifiez que le service kdump est en cours d'exécution.

```
[ec2-user ~]$ sudo service kdump status
```

Si le service est en cours d'exécution, la commande renvoie la réponse `Kdump is operational`.

Note

Par défaut, le fichier de vidage sur incident est enregistré dans `/var/crash/`. Pour modifier cet emplacement, modifiez le fichier `/etc/kdump.conf` à l'aide de l'éditeur de texte de votre choix.

Pour configurer SUSE Linux Enterprise, Ubuntu ou Red Hat Enterprise Linux

Consultez les sites web suivants :

- SUSE Linux Enterprise
- Ubuntu
- Red Hat Enterprise Linux (RHEL)

Note

Sur les instances basées sur les processeurs Intel et AMD, la commande `send-diagnostic-interrupt` envoie une interruption non masquable (NMI) inconnue à l'instance. Vous devez

configurer le noyau pour tomber en panne lorsqu'il reçoit l'interruption NMI inconnue. Ajoutez ce qui suit au fichier de configuration.

```
kernel.unknown_nmi_panic=1
```

Envoi d'une interruption de diagnostic

Une fois que vous avez effectué les modifications de configuration nécessaires, vous pouvez envoyer une interruption de diagnostic à votre instance à l'aide de l'AWS CLI ou de l'API Amazon EC2.

Pour envoyer une interruption de diagnostic à votre instance (AWS CLI)

Utilisez la commande send-diagnostic-interrupt et spécifiez l'ID de l'instance.

```
aws ec2 send-diagnostic-interrupt --instance-id i-1234567890abcdef0
```

Historique du document

Le tableau ci-dessous décrit les ajouts majeurs apportés à la documentation de Amazon EC2 depuis 2019. Nous mettons aussi la documentation à jour régulièrement pour prendre en compte les commentaires qui nous sont envoyés.

update-history-change	update-history-description	update-history-date
Prise en charge de la mise en veille prolongée pour les types d'instance I3, M5ad et R5ad	Désormais, vous pouvez mettre en veille prolongée vos instances nouvellement lancées s'exécutant sur les types d'instances I3, M5ad et R5ad.	October 21, 2020
Limitations de vCPU instance Spot	Les limitations instance Spot sont maintenant gérées en fonction du nombre de vCPU que votre Instances Spot en cours d'exécution utilise ou utilisera en attendant le traitement des demandes ouvertes.	October 1, 2020
Réservations de capacité dans Zones locales	Réservations de capacité peut maintenant être créé et utilisé dans Zones locales.	September 30, 2020
Amazon Data Lifecycle Manager	Les stratégies Amazon Data Lifecycle Manager peuvent être configurées avec quatre planifications au maximum.	September 17, 2020
Instances T4g (p. 1379)	Les nouvelles instances à usage général optimisées par les processeurs AWS Graviton2 sont basées sur des cœurs Arm Neoverse 64 bits et sont conçues en silicium par AWS pour offrir de meilleures performances à moindre coût.	September 14, 2020
Prise en charge de la mise en veille prolongée pour M5a et R5a	Désormais, vous pouvez mettre en veille prolongée vos instances nouvellement lancées s'exécutant sur les types d'instances M5a et 5Ra.	August 28, 2020
Volumes IOPS provisionnés SSD (io2) pour Amazon EBS	Les volumes IOPS provisionnés SSD (`io2`) sont conçus pour offrir une durabilité de 99,999 % avec un AFR ne dépassant pas 0,001 %.	August 24, 2020
Les métadonnées d'instance fournissent des informations sur l'emplacement et le placement	Nouveaux champs de métadonnées d'instance dans la catégorie `placement` : région, nom du groupe de placement,	August 24, 2020

	numéro de partition, ID d'hôte et ID de zone de disponibilité.	
Instances C5ad (p. 1379)	Nouvelles instances optimisées pour le calcul avec processeurs AMD EYPC de deuxième génération.	August 13, 2020
Zones de longueur d'onde	Une zone de longueur d'onde est une zone isolée située à l'emplacement du transporteur où l'infrastructure de longueur d'onde est déployée.	August 6, 2020
Groupes Réservation de capacité	Vous pouvez utiliser AWS Resource Groups pour créer des collections logiques de réservations de capacité, puis lancer des instances cibles dans ces groupes.	July 29, 2020
Instances C6gd, M6gd et R6gd (p. 1379)	Les nouvelles instances à usage général optimisées par les processeurs AWS Graviton2 sont basées sur des cœurs Arm Neoverse 64 bits et sont conçues en silicium par AWS pour offrir de meilleures performances à moindre coût.	July 27, 2020
Restauration d'instantané rapide	Vous pouvez activer la restauration d’instantané rapide pour les instantanés partagés avec vous.	July 21, 2020
Instances C6g et R6g (p. 1379)	Les nouvelles instances à usage général optimisées par les processeurs AWS Graviton2 sont basées sur des cœurs Arm Neoverse 64 bits et sont conçues en silicium par AWS pour offrir de meilleures performances à moindre coût.	June 10, 2020
Instances en métal nu pour G4 (p. 1379)	Nouvelles instances qui offrent à vos applications un accès direct aux ressources physiques du serveur hôte.	June 5, 2020
Instances C5a (p. 1379)	Nouvelles instances optimisées pour le calcul avec processeurs AMD EYPC de deuxième génération.	June 4, 2020
Apportez vos propres adresses IPv6	Vous pouvez fournir tout ou partie de votre plage d'adresses IPv6 depuis votre réseau sur site vers votre compte AWS.	May 21, 2020

Instances M6g (p. 1379)	Les nouvelles instances à usage général optimisées par les processeurs AWS Graviton2 sont basées sur des cœurs Arm Neoverse 64 bits et sont conçues en silicium par AWS pour offrir de meilleures performances à moindre coût.	May 11, 2020
Lancement des instances à l'aide d'un paramètre Systems Manager	Vous pouvez spécifier un paramètre AWS Systems Manager au lieu d'une AMI lorsque vous lancez une instance.	May 5, 2020
Personnaliser les notifications d'événements planifiés	Vous pouvez personnaliser les notifications d'événements planifiés pour inclure des balises dans la notification par e-mail.	May 4, 2020
Amazon Linux 2 Kernel Live Patching	Kernel Live Patching pour Amazon Linux 2 vous permet d'appliquer des correctifs de vulnérabilité de sécurité et de bogues critiques à un noyau Linux en cours d'exécution, sans redémarrer ni interrompre les applications en cours d'exécution.	April 28, 2020
Amazon EBS Multi-Attach	Vous pouvez désormais attacher un volume SSD d'IOPS provisionnés (io1) à un maximum de 16 instances basées sur Nitro se trouvant dans la même zone de disponibilité.	February 14, 2020
Arrêt et démarrage d'une instance Spot	Vous pouvez maintenant arrêter vos instances Spot sauvegardées par Amazon EBS et les démarrer à votre gré, au lieu de reposer sur le comportement d'arrêt sur interruption.	January 13, 2020
Balisage des ressources (p. 1379)	Vous pouvez baliser des passerelles Internet de sortie uniquement, des passerelles locales, des tables de routage de passerelle locale, des interfaces virtuelles de passerelle locale, des groupes d'interfaces virtuelles de passerelle locale, des associations de VPC de table de routage de passerelle locale et des associations de groupes d'interface virtuelle de table de routage de passerelle locale.	January 10, 2020

Connexion à votre instance à l'aide de Session Manager	Vous pouvez démarrer une session Session Manager avec une instance à partir de la console Amazon EC2.	December 18, 2019
Instances Inf1 (p. 1379)	Nouvelles instances dotées d'AWS Inferentia, une puce d'inférence de machine learning conçue pour offrir des performances élevées à faible coût.	December 3, 2019
Partage d'Hôte dédié	Vous pouvez désormais partager vos Hôtes dédiés entre des comptes AWS.	December 2, 2019
Hôtes dédiés et groupes de ressources hôte	Les Hôtes dédiés peuvent désormais être utilisés avec des groupes de ressources hôte.	December 2, 2019
Spécification de crédits par défaut au niveau du compte	Vous pouvez définir la spécification de crédits par défaut pour chaque famille d'instances à capacité extensible au niveau du compte pour chaque région AWS.	November 25, 2019
Découverte du type d'instance	Vous pouvez identifier un type d'instance qui répond à vos besoins.	November 22, 2019
Hôtes dédiés (p. 1379)	Vous pouvez désormais configurer un Hôte dédié pour prendre en charge plusieurs types d'instances au sein d'une famille d'instances.	November 21, 2019
Restauration d'instantané rapide Amazon EBS	Vous pouvez activer les restaurations d’instantané rapides sur un instantané EBS pour vous assurer que les volumes EBS créés à partir de l’instantané sont entièrement initialisés à la création et fournissent instantanément la totalité des performances allouées.	November 20, 2019
Service des métadonnées d'instance Version 2	Vous pouvez utiliser Service des métadonnées d'instance Version 2, qui est une méthode orientée session de demande de métadonnées d'instance.	November 19, 2019
Elastic Fabric Adapter (EFA) (p. 1379)	Les adaptateurs Elastic Fabric Adapter peuvent désormais être utilisés avec Intel MPI 2019 Update 6.	November 15, 2019

Achats d'Instances réservées mis en file d'attente	Vous pouvez mettre l'achat d'une Instance réservée en file d'attente jusqu'à trois ans en avance.	October 4, 2019
Instances G4 (p. 1379)	Nouvelles instances disposant de GPU NVIDIA Tesla.	September 19, 2019
Interruption de diagnostic	Vous pouvez envoyer une interruption de diagnostic à une instance inaccessible ou qui ne répond pas afin de déclencher une panique de noyau.	August 14, 2019
Stratégie d'allocation optimisée pour la capacité	À l'aide de Flotte EC2 ou Parc d'instances Spot, vous pouvez désormais lancer des Instances Spot depuis des pools d'instances Spot avec une capacité optimale pour le nombre d'instances que vous lancez.	August 12, 2019
Partage d'Réservation de capacité à la demande	Vous pouvez désormais partager une Réservations de capacité entre des comptes AWS.	July 29, 2019
Elastic Fabric Adapter (EFA) (p. 1379)	EFA prend désormais en charge MPI 3.1.4 et Intel MPI 2019 Update 4.	July 26, 2019
Balisage des ressources (p. 1379)	Vous pouvez baliser des modèles de lancement à la création.	July 24, 2019
EC2 Instance Connect	EC2 Instance Connect est une solution simple et sécurisée pour vous connecter à vos instances à l'aide de Secure Shell (SSH).	June 27, 2019
Récupération de l'hôte	Redémarre automatiquement vos instances sur un nouvel hôte en cas de panne matérielle soudaine sur un Hôte dédié.	June 5, 2019
Instantanés multi-volumes Amazon EBS	Vous pouvez prendre des instantanés en lien avec la panne, aux données coordonnées et à un instant donné exact, sur plusieurs volumes EBS attachés à une instance EC2.	May 29, 2019
Balisage des ressources (p. 1379)	Vous pouvez baliser les Réservations d'hôtes dédiés.	May 27, 2019

Chiffrement Amazon EBS par défaut	Une fois que vous avez activé le chiffrement par défaut dans une région, tous les nouveaux volumes EBS que vous créez dans cette région sont chiffrés à l'aide de la clé CMK par défaut pour le chiffrement EBS.	May 23, 2019
Balisage des ressources (p. 1379)	Vous pouvez baliser des points de terminaison de VPC, des services de points de terminaison et des configurations de services de points de terminaison.	May 13, 2019
Instances I3en (p. 1379)	Les nouvelles instances I3en peuvent utiliser jusqu'à 100 Gb/s de bande passante réseau.	May 8, 2019
Assistant de recréation de plateformes Windows vers Linux pour les bases de données Microsoft SQL Server	Déplacez les charges de travail Microsoft SQL Server d'un système d'exploitation Windows vers un système d'exploitation Linux.	May 8, 2019
Elastic Fabric Adapter (EFA)	Vous pouvez attacher un Elastic Fabric Adapter à vos instances pour accélérer les applications HPC (Calcul Haute Performance).	April 29, 2019
Instances T3a (p. 1379)	Nouvelles instances équipées de processeurs AMD EYPC.	April 24, 2019
Instances M5ad et R5ad (p. 1379)	Nouvelles instances équipées de processeurs AMD EYPC.	March 27, 2019
Balisage des ressources (p. 1379)	Vous pouvez attribuer des balises personnalisées à vos réservations Hôte dédié pour les catégoriser de différentes façons.	March 14, 2019
Instances matériel nu pour M5, M5d, R5, R5d et z1d (p. 1379)	Nouvelles instances qui offrent à vos applications un accès direct aux ressources physiques du serveur hôte.	February 13, 2019

Historique des années précédentes

Le tableau ci-dessous décrit les ajouts majeurs apportés à la documentation de Amazon EC2 en 2018 et avant.

Fonction	Version de l'API	Description	Date de parution
Groupes de placement par partition	15-11-2016	Les groupes de placement par partition répartissent les instances entre les partitions logiques, en	20 décembre 2018

Fonction	Version de l'API	Description	Date de parution
		s'assurant que les instances d'une partition ne partagent pas le matériel sous-jacent avec les instances d'autres partitions. Pour plus d'informations, consultez Groupes de placement par partition (p. 926).	
Instances p3dn.24xlarge	15-11-2016	Les nouvelles instances p3dn.xlarge fournissent 100 Gb/s de bande passante réseau.	7 décembre 2018
Mise en veille d'instances EC2 Linux	15-11-2016	Vous pouvez mettre en veille une instance Linux si cette dernière a été activée pour la mise en veille et répond aux exigences de la mise en veille. Pour plus d'informations, consultez Mettre votre instance Linux en veille prolongée (p. 628).	28 novembre 2018
Accélérateurs Amazon Elastic Inference	15-11-2016	Vous pouvez attacher un accélérateur Amazon EI à vos instances pour ajouter une accélération alimentée par GPU afin de réduire le coût d'inférence de deep learning. Pour plus d'informations, consultez Amazon Elastic Inference (p. 736).	28 novembre 2018
Instances équipées de 100 Gb/s de bande passante réseau.	15-11-2016	Les nouvelles instances C5n peuvent utiliser jusqu'à 100 Gb/s de bande passante réseau.	26 novembre 2018
Instances équipées de processeurs basés sur Arm	15-11-2016	Les nouvelles instances A1 permettent de réaliser des économies considérables et sont parfaitement adaptées aux charges de travail adaptatives et basées sur Arm.	26 novembre 2018
Parc d'instances recommandé par la console Spot	15-11-2016	La console Spot recommande un parc d'instances basé sur les bonnes pratiques Spot (diversification des instances) pour répondre aux spécifications matérielles minimales (vCPU, mémoire et stockage) de vos besoins applicatifs. Pour plus d'informations, consultez Création d'une demande de Parc d'instances Spot (p. 416).	20 novembre 2018
Nouveau type de demande Flotte EC2 : `instant`	15-11-2016	Flotte EC2 prend désormais en charge un nouveau type de demande, `instant`, que vous pouvez utiliser pour allouer de manière synchrone une capacité sur des types d'instance et des modèles d'achat. La demande `instant` renvoie les instances lancées dans la réponse d'API, sans aucune action supplémentaire. Cela vous permet de contrôler si et quand les instances sont lancées. Pour plus d'informations, consultez Types de demande Flotte EC2 (p. 562).	14 novembre 2018

Fonction	Version de l'API	Description	Date de parution
Instances équipées de processeurs AMD EYPC	15-11-2016	Les nouvelles instances à usage générale (M5a) et à mémoire optimisée (R5a) offrent des options à prix réduit pour microservices, des bases de données de petite et moyenne taille, des postes de travail virtuels, des environnements de développement et de test, des applications métier, etc.	6 novembre 2018
Informations d'économies Spot	15-11-2016	Vous pouvez afficher les économies réalisées à partir de l'utilisation de Instances Spot pour un Parc d'instances Spot unique ou pour tous les Instances Spot. Pour plus d'informations, consultez Économies réalisées grâce à l'achat d'Instances Spot (p. 388).	5 novembre 2018
Prise en charge de la console pour l'optimisation des options d'UC	15-11-2016	Lorsque vous lancez une instance, vous pouvez optimiser les options d'UC pour répondre à des besoins métier ou des charges de travail spécifiques à l'aide de la console Amazon EC2. Pour plus d'informations, consultez Optimisations des options d'UC (p. 673).	31 octobre 2018
Prise en charge de la console pour la création d'un modèle de lancement à partir d'une instance	15-11-2016	Vous pouvez créer un modèle de lancement en utilisant une instance comme base d'un nouveau modèle de lancement à l'aide de la console Amazon EC2. Pour plus d'informations, consultez Création d'un modèle de lancement (p. 539).	30 octobre 2018
Réservations de capacité à la demande	15-11-2016	Vous pouvez réserver de la capacité pour vos instances Amazon EC2 dans une zone de disponibilité spécifique pour la durée de votre choix. Cela vous permet de créer et de gérer des réservations de capacité indépendamment des remises de facturation offertes par les instances réservées (IR). Pour plus d'informations, consultez Réservations de capacité à la demande (p. 504).	25 octobre 2018
Fourniture de vos propres adresses IP (BYOIP)	15-11-2016	Vous pouvez fournir tout ou partie de votre plage d'adresses IPv4 publiques depuis votre réseau sur site vers votre compte AWS. Une fois que vous avez fourni la plage d'adresses à AWS, celle-ci s'affiche dans votre compte en tant que groupe d'adresses. Vous pouvez créer une adresse IP Elastic à partir de votre groupe d'adresses et l'utiliser avec vos ressources AWS. Pour plus d'informations, consultez Fourniture de vos propres adresses IP (BYOIP) dans Amazon EC2 (p. 825).	23 octobre 2018
Instances g3s.xlarge	15-11-2016	Étend la famille d'instances G3 à calcul accéléré avec le lancement des instances g3s.xlarge.	11 octobre 2018
Balisage de Hôte dédié à la création et prise en charge de la console	15-11-2016	Vous pouvez baliser vos Hôtes dédiés à la création et gérer vos balises Hôte dédié à l'aide de la console Amazon EC2. Pour plus d'informations, consultez Allocation d'Hôtes dédiés (p. 470).	08 octobre 2018

Fonction	Version de l'API	Description	Date de parution
Instances à mémoire élevée	15-11-2016	Ces instances sont conçues pour exécuter des grandes bases de données en mémoire. Elles offrent des performances bare metal avec accès direct au matériel hôte. Pour plus d'informations, consultez Instances de mémoire optimisée (p. 278).	27 septembre 2018
Instances f1.4xlarge	15-11-2016	Étend la famille d'instances F1 à calcul accéléré avec le lancement des instances f1.4xlarge.	25 septembre 2018
Prise en charge par la console du dimensionnement planifié pour Parc d'instances Spot	15-11-2016	Augmente ou réduit la capacité actuelle du parc en fonction de la date et de l'heure. Pour plus d'informations, consultez Mise à l'échelle d'un Parc d'instances Spot à l'aide de la mise à l'échelle planifiée (p. 445).	20 septembre 2018
Instances T3	15-11-2016	Les instances T3 représentent un type d'instance à usage général, extensible, de nouvelle génération, qui fournit des performances d'UC de base avec la possibilité d'étendre l'utilisation de l'UC à tout moment et aussi longtemps que nécessaire. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).	21 août 2018
Stratégies d'allocation pour les Flottes EC2	15-11-2016	Vous pouvez spécifier si l'affectation de capacité à la demande est traitée par prix (prix le plus bas en premier) ou par priorité (priorité la plus élevée en premier). Vous pouvez spécifier le nombre de groupes d'instances Spot auxquels allouer votre capacité Spot cible. Pour plus d'informations, consultez Stratégies d'allocation pour Instances Spot (p. 562).	26 juillet 2018
Stratégies d'allocation pour les Parcs d'instances Spot	15-11-2016	Vous pouvez spécifier si l'affectation de capacité à la demande est traitée par prix (prix le plus bas en premier) ou par priorité (priorité la plus élevée en premier). Vous pouvez spécifier le nombre de groupes d'instances Spot auxquels allouer votre capacité Spot cible. Pour plus d'informations, consultez Stratégie d'allocation pour les Instances Spot (p. 381).	26 juillet 2018
Instances R5 et R5d	15-11-2016	Les instances R5 et R5d conviennent parfaitement aux bases de données hautes performances, aux caches en mémoire distribués et aux analyses en mémoire. Les instances R5d sont fournies avec des volumes de stockage d'instance NVMe. Pour plus d'informations, consultez Instances de mémoire optimisée (p. 278).	25 juillet 2018

Fonction	Version de l'API	Description	Date de parution
Instances z1d	15-11-2016	Ces instances sont conçues pour les applications qui exigent des performances par cœur élevées avec une grande quantité de mémoire, comme les applications d'automatisation de la conception électronique (EDA) et les bases de données relationnelles. Ces instances sont fournies avec des volumes de stockage d'instance NVMe. Pour plus d'informations, consultez Instances de mémoire optimisée (p. 278).	25 juillet 2018
Automatisation du cycle de vie des instantanés	15-11-2016	Vous pouvez utiliser Amazon Data Lifecycle Manager pour automatiser la création et la suppression d'instantanés pour vos volumes EBS. Pour plus d'informations, consultez Automatisation du cycle de vie des instantanés Amazon EBS (p. 1169).	12 juillet 2018
Options d'UC liées aux modèles de lancement	15-11-2016	Lorsque vous créez un modèle de lancement à l'aide des outils de ligne de commande, vous pouvez optimiser les options d'UC afin de les adapter à des charges de travail spécifiques ou à vos besoins métier. Pour plus d'informations, consultez Création d'un modèle de lancement (p. 539).	11 juillet 2018
Balisage des Hôtes dédiés	15-11-2016	Vous pouvez baliser vos Hôtes dédiés. Pour plus d'informations, consultez Balisage des entités Hôte dédié (p. 481).	3 juillet 2018
Instances `i3.metal`	15-11-2016	Les instances `i3.metal` offrent à vos applications un accès direct aux ressources physiques du serveur hôte, telles que les processeurs et la mémoire. Pour plus d'informations, consultez Instances de stockage optimisé (p. 289).	17 mai 2018
Obtenir la dernière sortie de console	15-11-2016	Vous pouvez extraire la dernière sortie de la console pour certains types d'instances quand vous utilisez la commande get-console-output de l'AWS CLI.	9 mai 2018
Optimiser les options d'UC	15-11-2016	Lorsque vous lancez une instance, vous pouvez optimiser les options d'UC pour répondre à des besoins métier ou des charges de travail spécifiques. Pour plus d'informations, consultez Optimisations des options d'UC (p. 673).	8 mai 2018
Flotte EC2	15-11-2016	Vous pouvez utiliser le Flotte EC2 pour lancer un groupe d'instances entre différents types d'instance EC2 et zones de disponibilité, et entre les modèles d'achat instance à la demande, Instance réservée et instance Spot. Pour plus d'informations, consultez Lancement d'instances à l'aide d'un Flotte EC2 (p. 558).	2 mai 2018

Fonction	Version de l'API	Description	Date de parution
Instances à la demande dans des Parcs d'instances Spot	15-11-2016	Vous pouvez inclure une capacité à la demande dans votre demande de Parc d'instances Spot pour garantir que vous avez toujours la capacité d'instance. Pour plus d'informations, consultez Fonctionnement de Parc d'instances Spot (p. 380).	2 mai 2018
Baliser les instantanés EBS à la création	15-11-2016	Vous pouvez appliquer des balises aux instantanés au moment de la création. Pour plus d'informations, consultez Création d'instantanés Amazon EBS (p. 1132).	2 avril 2018
Modifier les groupes de placement	15-11-2016	Vous pouvez déplacer une instance à l'intérieur ou à l'extérieur d'un groupe de placement, ou modifier son groupe de placement. Pour plus d'informations, consultez Modification du groupe de placement d'une instance (p. 935).	1 mars 2018
ID de ressource plus longs	15-11-2016	Vous pouvez activer le format d'ID long pour d'autres types de ressource. Pour plus d'informations, consultez ID de ressource (p. 1302).	9 février 2018
Améliorations des performances réseau	15-11-2016	Les instances qui se trouvent en dehors d'un groupe de placement de cluster peuvent à présent profiter d'une bande passante plus élevée pour l'envoi ou la réception de trafic réseau entre d'autres instances ou Amazon S3. Pour plus d'informations, consultez Fonctions de mise en réseau et de stockage (p. 219).	24 janvier 2018
Balisage de vos adresses IP Elastic	15-11-2016	Vous pouvez baliser vos adresses IP Elastic. Pour plus d'informations, consultez Balisage d'une adresse IP Elastic (p. 834).	21 décembre 2017
Amazon Linux 2	15-11-2016	Amazon Linux 2 est une nouvelle version d'Amazon Linux. Ce service fournit une base haute performance stable et sécurisée pour vos applications. Pour plus d'informations, consultez Amazon Linux (p. 186).	13 décembre 2017
Amazon Time Sync Service	15-11-2016	Amazon Time Sync Service permet de garder une heure précise sur votre instance. Pour plus d'informations, consultez Réglage de l'heure pour votre instance Linux (p. 668).	29 novembre 2017
T2 illimité	15-11-2016	Les instances T2 illimité peuvent dépasser le niveau de base aussi longtemps que nécessaire. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).	29 novembre 2017
Modèles de lancement	15-11-2016	Un modèle de lancement peut contenir tout ou partie des paramètres permettant de lancer une instance. Il est donc inutile de les spécifier à chaque lancement d'une instance. Pour plus d'informations, consultez Lancement d'une instance à partir d'un modèle de lancement (p. 537).	29 novembre 2017

Fonction	Version de l'API	Description	Date de parution
Placement par répartition	15-11-2016	Les groupes de placement par répartition sont recommandés pour les applications ayant un petit nombre d'instances critiques, qui doivent être séparées les unes des autres. Pour plus d'informations, consultez Groupes de placement par répartition (p. 927).	29 novembre 2017
Instances H1	15-11-2016	Les instances H1 sont conçues pour des charges de travail Big Data hautes performances. Pour plus d'informations, consultez Instances de stockage optimisé (p. 289).	28 novembre 2017
Instances M5	15-11-2016	Les instances M5 appartiennent à la nouvelle génération d'instances de calcul à usage général. Elles offrent des ressources de calcul, de mémoire, de stockage et de réseau équilibrées.	28 novembre 2017
Mise en veille prolongée d'instances Spot	15-11-2016	Le service d'instances Spot peut mettre les instances Spot en veille prolongée en cas d'interruption. Pour plus d'informations, consultez Mise en veille prolongée des Instances Spot interrompues (p. 454).	28 novembre 2017
Suivi de la cible de Parc d'instances Spot	15-11-2016	Vous pouvez configurer des stratégies de dimensionnement Suivi de la cible pour votre Parc d'instances Spot. Pour plus d'informations, consultez Mise à l'échelle d'un Parc d'instances Spot en utilisant une stratégie Suivi de la cible (p. 441).	17 novembre 2017
Parc d'instances Spot s'intègre à Elastic Load Balancing.	15-11-2016	Vous pouvez attacher un ou plusieurs équilibreurs de charge à un Parc d'instances Spot.	10 novembre 2017
Instances X1e	15-11-2016	Les instances X1e conviennent parfaitement aux bases de données hautes performances, aux bases de données en mémoire et aux autres applications d'entreprise utilisant beaucoup de mémoire. Pour plus d'informations, consultez Instances de mémoire optimisée (p. 278).	28 novembre 2017
Instances C5	15-11-2016	Les instances C5 sont conçues pour des applications de calculs intensifs. Pour plus d'informations, consultez Instances de calcul optimisé (p. 270).	6 novembre 2017
Fusionner et diviser des Instances réservées convertibles	15-11-2016	Vous pouvez échanger (ou fusionner) deux Instances réservées convertibles ou plus pour obtenir une nouvelle Instance réservée convertible. Vous pouvez également utiliser le processus de modification pour diviser une Instance réservée convertible en plus petites réservations. Pour plus d'informations, consultez Échange d'Instances réservées convertibles (p. 365).	6 novembre 2017

Fonction	Version de l'API	Description	Date de parution
Instances P3	15-11-2016	Les instances P3 appartiennent à la nouvelle génération d'instances GPU optimisées pour le calcul. Pour plus d'informations, consultez Instances à calcul accéléré Linux (p. 296).	25 octobre 2017
Modification de la location de VPC	15-11-2016	Vous pouvez modifier l'attribut de location d'instance d'un VPC en remplaçant `dedicated` par `default`. Pour plus d'informations, consultez Modification de la location d'un VPC (p. 503).	16 octobre 2017
Facturation par seconde	15-11-2016	Amazon EC2 facture l'utilisation Linux par seconde, avec des frais minimum d'une minute.	2 octobre 2017
Arrêt sur une interruption	15-11-2016	Vous pouvez préciser si Amazon EC2 doit arrêter ou résilier les Instances Spot lorsqu'elles sont interrompues. Pour plus d'informations, consultez Comportements d'interruption (p. 453).	18 septembre 2017
Baliser des passerelles NAT	15-11-2016	Vous pouvez baliser votre passerelle NAT. Pour plus d'informations, consultez Balisage de vos ressources (p. 1310).	7 septembre 2017
Descriptions des règles des groupes de sécurité	15-11-2016	Vous pouvez ajouter des descriptions aux règles des groupes de sécurité. Pour plus d'informations, consultez Règles des groupes de sécurité (p. 1063).	31 août 2017
Récupération d'adresses IP Elastic	15-11-2016	Si vous avez libéré une adresse IP Elastic à utiliser dans un VPC, vous pouvez essayer de la récupérer. Pour plus d'informations, consultez Récupération d'une adresse IP Elastic (p. 838).	11 août 2017
Baliser les instances de Parc d'instances Spot	15-11-2016	Vous pouvez configurer votre Parc d'instances Spot pour baliser les instances qu'il lance.	24 juillet 2017
Instances G3	15-11-2016	Les instances G3 offrent une plate-forme économique à hautes performances pour les applications graphiques qui utilisent DirectX ou OpenGL. Les instances G3 fournissent également des fonctions de station de travail virtuelle NVIDIA GRID, qui prennent en charge 4 écrans avec des résolutions pouvant atteindre 4096x2160. Pour plus d'informations, consultez Instances à calcul accéléré Linux (p. 296).	13 juillet 2017
Instances F1	15-11-2016	Les instances F1 représentent la prochaine génération d'instances à calcul accéléré. Pour plus d'informations, consultez Instances à calcul accéléré Linux (p. 296).	19 avril 2017

Fonction	Version de l'API	Description	Date de parution
Baliser des ressources pendant la création	15-11-2016	Vous pouvez appliquer des balises à des instances et des volumes au moment de la création. Pour plus d'informations, consultez Balisage de vos ressources (p. 1310). De plus, vous pouvez utiliser des autorisations de niveau ressources basées sur des balises pour contrôler les balises appliquées. Pour plus d'informations, consultez Octroi de l'autorisation de baliser les ressources lors de la création (p. 985).	28 mars 2017
Instances I3	15-11-2016	Les instances I3 représentent la prochaine génération d'instances de stockage optimisé. Pour plus d'informations, consultez Instances de stockage optimisé (p. 289).	23 février 2017
Effectuez des modifications sur les volumes EBS attachés	15-11-2016	Avec la plupart des volumes EBS attachés à la plupart des instances EC2, vous pouvez modifier la taille de volume, le type et les E/S par seconde sans détacher le volume, ni arrêter l'instance. Pour plus d'informations, consultez Amazon EBS Elastic Volumes (p. 1182).	13 février 2017
Attachement d'un rôle IAM	15-11-2016	Vous pouvez attacher, détacher ou remplacer un rôle IAM pour une instance existante. Pour plus d'informations, consultez Rôles IAM pour Amazon EC2 (p. 1034).	9 février 2017
Instances Spot dédiées	15-11-2016	Vous pouvez exécuter les Instances Spot sur un matériel à client unique dans un Virtual Private Cloud (VPC). Pour plus d'informations, consultez Définition d'une location pour vos Instances Spot (p. 393).	19 janvier 2017
Prise en charge d'IPv6	15-11-2016	Vous pouvez associer un bloc d'adresse CIDR IPv6 au VPC et aux sous-réseaux, et attribuer des adresses IPv6 aux instances de votre VPC. Pour plus d'informations, consultez Adressage IP des instances Amazon EC2 (p. 808).	1er décembre 2016
Instances R4	15-09-2016	Les instances R4 représentent la prochaine génération d'instances à mémoire optimisée. Les instances R4 conviennent aux charges de travail qui exigent beaucoup de mémoire et sont sensibles à la latence, par exemple BI, l'exploration et l'analyse des données, les bases de données en mémoire, la mise en cache web en mémoire distribuée à grande échelle et le traitement performant en temps réel des données non structurées par les applications. Pour plus d'informations, consultez Instances de mémoire optimisée (p. 278)	30 novembre 2016

Fonction	Version de l'API	Description	Date de parution
Nouveaux types d'instance `t2.xlarge` et `t2.2xlarge`	15-09-2016	Les instances T2 sont conçues pour offrir des performance de base modérées et la possibilité d'atteindre des performances nettement supérieures si votre charge de travail l'exige. Elles sont destinées aux applications qui requièrent une réactivité et des performances élevées pendant des périodes limitées, ainsi qu'un coût faible. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).	30 novembre 2016
Instances P2	15-09-2016	Les instances P2 utilisent des GPU NVIDIA Tesla K80 et sont conçues pour le calcul GPU à usage général à l'aide des modèles de programmation CUDA ou OpenCL. Pour plus d'informations, consultez Instances à calcul accéléré Linux (p. 296).	29 septembre 2016
Instances `m4.16xlarge`	01-04-2016	Étend la famille M4 à usage général avec la mise en place d'instances `m4.16xlarge`, dotées de 64 vCPU et 256 Gio de RAM.	6 septembre 2016
Dimensionnement automatique pour un Parc d'instances Spot		Vous pouvez désormais configurer des stratégies de dimensionnement pour votre Parc d'instances Spot. Pour plus d'informations, consultez Dimensionnement automatique pour un Parc d'instances Spot (p. 439).	1 septembre 2016
Elastic Network Adapter (ENA)	01-04-2016	Vous pouvez désormais utiliser ENA pour une mise en réseau améliorée. Pour plus d'informations, consultez Prise en charge de la mise en réseau améliorée (p. 864).	28 juin 2016
Prise en charge améliorée de l'affichage et de la modification des ID longs	01-04-2016	Vous pouvez maintenant afficher et modifier les paramètres d'ID long des autres utilisateurs IAM, des rôles IAM ou de l'utilisateur racine. Pour plus d'informations, consultez ID de ressource (p. 1302).	23 juin 2016
Copie d'instantanés Amazon EBS chiffrés entre des comptes AWS	01-04-2016	Vous pouvez maintenant copier les instantanés EBS chiffrés entre des comptes AWS. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).	21 juin 2016
Création d'une capture d'écran d'une console d'instance	01-10-2015	Vous pouvez désormais obtenir des informations supplémentaires lors du débogage d'instances inaccessibles. Pour plus d'informations, consultez Création d'une capture d'écran d'une instance inaccessible (p. 1362).	24 mai 2016
Instances X1	01-10-2015	Instances à mémoire optimisée conçues pour exécuter les bases de données dans la mémoire, moteurs de traitement du Big Data et applications de calcul hautes performances (HPC). Pour plus d'informations, consultez Instances de mémoire optimisée (p. 278).	18 mai 2016

Fonction	Version de l'API	Description	Date de parution
Deux nouveaux types de volumes EBS	01-10-2015	Vous pouvez désormais créer des volumes HDD à débit optimisé (st1) et des volumes HDD à froid (sc1). Pour plus d'informations, consultez Types de volume Amazon EBS (p. 1089).	19 avril 2016
Ajout de nouvelles métriques NetworkPacketsIn et NetworkPacketsOut pour Amazon EC2		Ajout de nouvelles métriques NetworkPacketsIn et NetworkPacketsOut pour Amazon EC2 Pour plus d'informations, consultez Métriques des instances (p. 764).	23 mars 2016
Métriques CloudWatch pour un Parc d'instances Spot		Vous pouvez maintenant obtenir des métriques CloudWatch pour votre Parc d'instances Spot. Pour plus d'informations, consultez Métriques CloudWatch pour un Parc d'instances Spot (p. 437).	21 mars 2016
Instances planifiées	01-10-2015	Les instances réservées planifiées (instances planifiées) vous permettent d'acheter des réservations de capacité récurrentes sur une base quotidienne, hebdomadaire ou mensuelle, avec une date de début et une durée spécifiées. Pour plus d'informations, consultez Instances réservées planifiées (p. 369).	13 janvier 2016
ID de ressource plus longs	01-10-2015	Nous introduisons progressivement des ID plus longs pour certains types de ressource Amazon EC2 et Amazon EBS. Durant la période d'abonnement, vous pouvez activer le format d'ID plus long pour les types de ressources pris en charge. Pour plus d'informations, consultez ID de ressource (p. 1302).	13 janvier 2016
ClassicLink Support DNS	01-10-2015	Vous pouvez activer la prise en charge de DNS ClassicLink pour votre VPC afin que les noms d'hôte DNS qui sont adressés entre les instances EC2-Classic et les instances du VPC soient résolus en adresses IP privées et non en adresses IP publiques. Pour plus d'informations, consultez Activation de la prise en charge de DNS ClassicLink (p. 955).	11 janvier 2016
Nouveau `t2.nano` type d'instance	01-10-2015	Les instances T2 sont conçues pour offrir des performance de base modérées et la possibilité d'atteindre des performances nettement supérieures si votre charge de travail l'exige. Elles sont destinées aux applications qui requièrent une réactivité et des performances élevées pendant des périodes limitées, ainsi qu'un coût faible. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).	15 décembre 2015

Fonction	Version de l'API	Description	Date de parution
Hôtes dédiés	01-10-2015	Un hôte dédié Amazon EC2 est un serveur physique avec une capacité d'instance dédiée à votre utilisation. Pour plus d'informations, consultez Hôtes dédiés (p. 465).	23 novembre 2015
Durée des instance Spot	01-10-2015	Vous pouvez désormais spécifier une durée pour vos Instances Spot. Pour plus d'informations, consultez Définition d'une durée pour vos Instances Spot (p. 392).	6 octobre 2015
Demande de modification de Parc d'instances Spot	01-10-2015	Vous pouvez désormais modifier la capacité cible de votre demande de Parc d'instances Spot. Pour plus d'informations, consultez Modification d'une demande de Parc d'instances Spot (p. 427).	29 septembre 2015
Stratégie d'allocation diversifiée de Parc d'instances Spot	15-04-2015	Vous pouvez désormais allouer des Instances Spot dans plusieurs pools d'instances Spot à l'aide d'une seule demande de Parc d'instances Spot. Pour plus d'informations, consultez Stratégie d'allocation pour les Instances Spot (p. 381).	15 septembre 2015
Pondération d'instance de Parc d'instances Spot	15-04-2015	Vous pouvez désormais définir les unités de capacité par lesquelles chaque type d'instance contribue aux performances de votre application et ajuster en conséquence le montant que vous êtes prêt à payer pour des Instances Spot pour chaque pool d'instances Spot. Pour plus d'informations, consultez Pondération d'instance de Parc d'instances Spot (p. 384).	31 août 2015
Nouvelle action d'alarme de redémarrage et nouveau rôle IAM à utiliser avec les actions d'alarme		Ajout de la nouvelle action d'alarme de redémarrage et du nouveau rôle IAM à utiliser avec les actions d'alarme. Pour plus d'informations, consultez Créer des alarmes qui arrêtent, finissent, redémarrent ou récupèrent une instance (p. 785).	23 juillet 2015
Nouveau `t2.large` type d'instance		Les instances T2 sont conçues pour offrir des performance de base modérées et la possibilité d'atteindre des performances nettement supérieures si votre charge de travail l'exige. Elles sont destinées aux applications qui requièrent une réactivité et des performances élevées pendant des périodes limitées, ainsi qu'un coût faible. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).	16 juin 2015
Instances M4		La prochaine génération d'instances à visée générale qui fournissent un équilibre entre les ressources de calcul, de mémoire et de réseau. Les instances M4 sont alimentées par un processeur personnalisé Intel 2,4 GHz Intel® Xeon® E5 2676v3 (Haswell) avec AVX2.	11 juin 2015

Fonction	Version de l'API	Description	Date de parution
Parcs d'instances Spot	15-04-2015	Vous pouvez gérer un ensemble, ou parc, d'Instances Spot au lieu de gérer des demandes d'instance Spot distinctes. Pour plus d'informations, consultez Fonctionnement de Parc d'instances Spot (p. 380).	18 mai 2015
Migrer les adresses IP Elastic vers EC2-Classic	15-04-2015	Vous ne pouvez pas migrer une adresse IP Elastic que vous avez allouée pour être utilisée dans EC2-Classic afin qu'elle soit utilisée dans un VPC. Pour plus d'informations, consultez Migration d'une Adresse IP Elastic à partir de EC2-Classic (p. 946).	15 mai 2015
Importation de machines virtuelles avec plusieurs disques comme AMI	01-03-2015	Le processus VM Import prend désormais en charge l'importation de machines virtuelles avec plusieurs disques comme AMI. Pour plus d'informations, consultez Importation d'un ordinateur virtuel comme image à l'aide de VM Import/Export dans le VM Import/Export Guide de l'utilisateur.	23 avril 2015
Nouveau `g2.8xlarge` type d'instance		La nouvelle instance `g2.8xlarge` est supportée par quatre GPU NVIDIA hautes performances, ce qui la rend parfaitement adaptée aux charges de travail de calcul GPU, y compris le rendu à grande échelle, le transcodage, le Machine Learning et autres charges de travail côté client qui nécessitent une importante puissance de traitement en parallèle.	7 avril 2015
Instances D2		Prochaine génération d'instances à stockage dense Amazon EC2 optimisées pour les applications nécessitant un accès séquentiel à d'importants volumes de données sur un stockage d'instance en attachement direct. Les instances D2 sont conçues pour offrir le meilleur rapport prix/performance de la gamme de stockage dense. Alimentées par des processeurs 2,4 GHz Intel® Xeon® E5 2676v3 (Haswell), les instances D2 constituent une amélioration par rapport aux instances HS1 en fournissant une puissance de calcul supplémentaire, plus de mémoire et une mise en réseau améliorée. De plus, les instances D2 sont disponibles en quatre tailles d'instance, avec des options de stockage de 6 To, 12 To, 24 To et 48 To. Pour plus d'informations, consultez Instances de stockage optimisé (p. 289).	24 mars 2015

Fonction	Version de l'API	Description	Date de parution
Récupération automatique des instances EC2		Vous pouvez créer une alarme Amazon CloudWatch qui surveille une instance Amazon EC2 et récupère automatiquement l'instance si cette dernière est dégradée suite à une défaillance du matériel sous-jacent ou à un problème nécessitant une intervention d'AWS pour sa résolution. Une instance récupérée est identique à l'instance d'origine, y compris son ID d'instance, les adresses IP privées et toutes les métadonnées d'instance. Pour plus d'informations, consultez Récupération de votre instance (p. 652).	12 janvier 2015
Instances C4		Prochaine génération d'instances optimisées pour le calcul qui fournissent des performances d'UC très élevées à un prix économique. Les instances C4 sont basées sur des processeurs 2,9 GHz Intel® Xeon® E5-2666 v3 (Haswell) personnalisés. Grâce à Turbo boost, la vitesse de l'horloge du processeur des instances C4 instances peut atteindre des fréquences aussi élevées que 3,5 Ghz avec 1 ou 2 cœur(s) Turbo. Se développant sur les capacités des instances C3 optimisées pour le calcul, les instances C4 offrent aux clients les performances de processeur les plus élevées parmi les instances EC2. Ces instances sont idéalement adaptées aux applications web à trafic élevé, à la diffusion de publicités, au traitement par batch, au codage vidéo, à l'analyse distribuée, à la physique haute énergie, à l'analyse du génome et à la modélisation numérique en dynamique des fluides. Pour plus d'informations, consultez Instances de calcul optimisé (p. 270).	11 janvier 2015
ClassicLink	01-10-2014	ClassicLink vous permet de relier votre instance EC2-Classic à un VPC de votre compte. Vous pouvez associer des groupes de sécurité VPC à l'instance EC2-Classic en activant la communication entre votre instance EC2-Classic et des instances de votre VPC à l'aide d'adresses IP privées. Pour plus d'informations, consultez ClassicLink (p. 949).	7 janvier 2015

Fonction	Version de l'API	Description	Date de parution
Avis d'interruption d'instance Spot		Le meilleur moyen de vous protéger contre une interruption d'instance Spot est de faire en sorte que votre application soit tolérante aux pannes au niveau de son architecture. En outre, vous pouvez tirer parti des avis de résiliation d'instance Spot, qui vous préviennent deux minutes avant qu'Amazon EC2 n'interrompe ou ne résilie votre instance Spot. Pour plus d'informations, consultez Avis d'interruption d'une instance Spot (p. 457).	5 janvier 2015
Prise en charge de la pagination `DescribeVolumes`	01-09-2014	L'appel de l'API `DescribeVolumes` prend désormais en charge la pagination des résultats à l'aide des paramètres `MaxResults` et `NextToken`. Pour plus d'informations, consultez DescribeVolumes dans le manuel Amazon EC2 API Reference.	23 octobre 2014
Instances T2	15-06-2014	Les instances T2 sont conçues pour offrir des performance de base modérées et la possibilité d'atteindre des performances nettement supérieures si votre charge de travail l'exige. Elles sont destinées aux applications qui requièrent une réactivité et des performances élevées pendant des périodes limitées, ainsi qu'un coût faible. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).	30 juin 2014
Nouvelle page Limites de service EC2		Utilisez la page Limites de service EC2 de la console Amazon EC2 afin d'afficher les limites actuelles des ressources fournies par Amazon EC2 et Amazon VPC, région par région.	19 juin 2014
Volumes Amazon EBS SSD à usage général	01-05-2014	Les volumes SSD à usage général offrent un stockage économique idéal pour un large éventail de charges de travail. Ces volumes offrent des latences inférieures à 10 millisecondes, la capacité d'augmenter jusqu'à 3 000 IOPS pour une durée étendue et une performance de base de 3 IOPS/Gio. La taille des volumes SSD à usage général peut aller de 1 Gio à 1 Tio. Pour plus d'informations, consultez Volumes SSD à usage général (`gp2`) (p. 1091).	16 juin 2014

Fonction	Version de l'API	Description	Date de parution
Chiffrement Amazon EBS	01-05-2014	Chiffrement Amazon EBS offre un chiffrement transparent des instantanés et des volumes de données EBS sans que vous ayez à développer et à maintenir une infrastructure de gestion de clés sécurisée. EBS Encryption assure la sécurité des données au repos en chiffrant vos données à l'aide de clés gérées par Amazon. Le chiffrement est effectué sur les serveurs hébergeant des instances EC2, assurant le chiffrement des données lorsqu'elles se déplacent entre les instances EC2 et le stockage EBS. Pour plus d'informations, consultez Chiffrement Amazon EBS (p. 1194).	21 mai 2014
Instances R3	01-02-2014	Prochaine génération d'instances optimisées pour la mémoire avec le meilleur niveau de prix par Gio de RAM et des performances élevées. Ces instances conviennent parfaitement pour les bases de données NoSQL et relationnelles, les solutions d'analyse en mémoire, les applications de calcul scientifique et autres applications très gourmandes en mémoire qui peuvent tirer parti de la mémoire élevée par processeur virtuel, des hautes performances de calcul et des capacités de mise en réseau améliorée des instances R3. Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.	9 avril 2014
Nouvelle version Amazon Linux AMI		Amazon Linux AMI 2014.03 est disponible.	27 mars 2014
Amazon EC2 Usage Reports		Amazon EC2 Usage Reports est un ensemble de rapports qui affiche les données de coût et d'utilisation d'EC2. Pour plus d'informations, consultez Rapports d'utilisation d'Amazon EC2 (p. 1324).	28 janvier 2014
Instances M3 supplémentaires	15-10-2013	Les tailles `m3.medium` et `m3.large` des instances M3 sont désormais prises en charge. Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.	20 janvier 2014

Fonction	Version de l'API	Description	Date de parution
Instances I2	15-10-2013	Ces instances fournissent des IOPS très élevées et prennent en charge la commande TRIM sur les instances Linux pour de meilleures performances d'écriture SSD. Les instances I2 prennent aussi en charge la gestion de la mise en réseau qui offrent des latences inter-instance améliorées, une instabilité réseau moindre et des performances en termes de paquet par seconde (PPS) significativement plus élevées. Pour plus d'informations, consultez Instances de stockage optimisé (p. 289).	19 décembre 2013
Instances M3 mises à jour	15-10-2013	Les tailles `m3.xlarge` et `m3.2xlarge` d'instance M3 prennent désormais en charge le stockage d'instance avec les volumes SSD.	19 décembre 2013
Importation de machines virtuelles Linux	15-10-2013	Le processus VM Import prend désormais en charge l'importation d'instances Linux. Pour plus d'informations, consultez la VM Import/Export Guide de l'utilisateur.	16 décembre 2013
Autorisations au niveau des ressources pour RunInstances	15-10-2013	Vous pouvez désormais créer des politiques dans AWS Identity and Access Management pour contrôler les permissions au niveau des ressources pour l'action de l'API RunInstances d'Amazon EC2. Pour plus d'informations et obtenir des exemples de stratégie, consultez Identity and Access Management pour Amazon EC2 (p. 976).	20 novembre 2013
Instances C3	15-10-2013	Instances optimisées pour le calcul qui fournissent des performances d'UC très élevées à un prix économique. Les instances C3 prennent aussi en charge la gestion de la mise en réseau qui offrent des latences inter-instance améliorées, une instabilité réseau moindre et des performances en termes de paquet par seconde (PPS) significativement plus élevées. Ces instances sont idéalement adaptées aux applications web à trafic élevé, à la diffusion de publicités, au traitement par batch, au codage vidéo, à l'analyse distribuée, à la physique haute énergie, à l'analyse du génome et à la modélisation numérique en dynamique des fluides. Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.	14 novembre 2013
Lancement d'une instance depuis AWS Marketplace		Vous pouvez désormais lancer une instance à partir de AWS Marketplace à l'aide de l'Assistant de lancement Amazon EC2. Pour plus d'informations, consultez Lancement d'une instance AWS Marketplace (p. 557).	11 novembre 2013

Fonction	Version de l'API	Description	Date de parution
Instances G2	01-10-2013	Ces instances conviennent parfaitement pour les services de création vidéo, les visualisations 3D, la diffusion d'applications gourmandes en graphiques et autres charges de travail côté serveur nécessitant une importante puissance de traitement en parallèle. Pour plus d'informations, consultez Instances à calcul accéléré Linux (p. 296).	4 novembre 2013
Nouvel Assistant de lancement		L'Assistant de lancement EC2 a été refondu. Pour plus d'informations, consultez Lancement d'une instance à l'aide de l'assistant de lancement d'instance (p. 530).	10 octobre 2013
Modification des types d'instance des instances réservées d'Amazon EC2	01-10-2013	Vous pouvez désormais modifier le type d'instances des instances réservées Linux d'une même famille (par exemple, M1, M2, M3, C1). Pour plus d'informations, consultez Modification des Instances réservées (p. 357).	09 octobre 2013
Nouvelle version Amazon Linux AMI		Amazon Linux AMI 2013.09 est disponible.	30 septembre 2013
Modification des instances réservées Amazon EC2	15-08-2013	Vous pouvez désormais modifier les instances réservées d'une région. Pour plus d'informations, consultez Modification des Instances réservées (p. 357).	11 septembre 2013
Attribution d'une adresse IP publique	15-07-2013	Vous pouvez désormais attribuer une adresse IP publique quand vous lancez une instance dans VPC. Pour plus d'informations, consultez Attribution d'une adresse IPv4 publique lors du lancement d'une instance (p. 813).	20 août 2013
Attribution d'autorisations au niveau des ressources	15-06-2013	Amazon EC2 prend en charge les nouveaux Amazon Resource Names (ARN) et clés de condition. Pour plus d'informations, consultez Stratégies IAM pour Amazon EC2 (p. 979).	8 juillet 2013
Copies d'instantané incrémentielles	01-02-2013	Vous pouvez désormais effectuer des copies d'instantané incrémentielles. Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).	11 juin 2013
Nouvelle page Balises		Nouvelle page Balises dans la console Amazon EC2. Pour plus d'informations, consultez Balisage de vos ressources Amazon EC2 (p. 1309).	04 avril 2013
Nouvelle version Amazon Linux AMI		Amazon Linux AMI 2013.03 est disponible.	27 mars 2013

Fonction	Version de l'API	Description	Date de parution
Types d'instance optimisées EBS additionnelles	01-02-2013	Les types d'instance suivants peuvent désormais être lancés en tant qu'instances optimisées EBS : `c1.xlarge`, `m2.2xlarge`, `m3.xlarge` et `m3.2xlarge`. Pour plus d'informations, consultez Amazon EBS–Instances optimisées (p. 1213).	19 mars 2013
Copier une AMI d'une région vers une autre	01-02-2013	Vous pouvez copier une AMI d'une région vers une autre, ce qui vous permet de lancer des instances cohérentes dans plusieurs régions AWS rapidement et facilement. Pour plus d'informations, consultez Copie une AMI (p. 173).	11 mars 2013
Lancer une instance dans un VPC par défaut	01-02-2013	Votre compte AWS est capable de lancer des instances sur EC2-Classic ou sur un VPC, ou seulement sur VPC, région par région. Si vous pouvez uniquement lancer des instances dans un VPC, nous créerons un VPC par défaut pour vous. Lorsque vous lancez une instance, nous la lançons dans votre VPC par défaut, sauf si vous créez un VPC personnalisé et que vous le spécifiez au lancement de l'instance.	11 mars 2013
Type d'instance cluster (cr1.8xlarge) à mémoire élevée	01-12-2012	Possibilité d'avoir d'importantes quantités de mémoire couplées avec des performances UC et réseau élevées. Ces instances conviennent parfaitement pour l'analyse en mémoire, l'analyse des graphiques et les applications de calcul scientifique.	21 janvier 2013
Type d'instance de stockage (`hs1.8xlarge`) élevé	01-12-2012	Les instances à stockage élevé fournissent une densité de stockage très haute et d'excellentes performances en lecture/écriture par instance. Elles sont parfaitement adaptées à l'entreposage des données, à Hadoop/MapReduce et aux systèmes de fichiers en parallèle.	20 décembre 2012
Copie d'instantané EBS	01-12-2012	Vous pouvez utiliser les copies d'instantané pour créer des sauvegardes de données, des volumes Amazon EBS ou des Amazon Machine Images (AMI). Pour plus d'informations, consultez Copie d'un instantané Amazon EBS (p. 1138).	17 décembre 2012
Métriques EBS mises à jour et contrôles de statut pour les volumes IOPS provisionnés SSD	01-10-2012	Métriques EBS mises à jour pour inclure deux nouvelles métriques pour les volumes IOPS provisionnés SSD. Pour plus d'informations, consultez Métriques Amazon CloudWatch pour Amazon EBS (p. 1247). Ajout de nouveaux contrôles de statut pour les volumes IOPS provisionnés SSD. Pour plus d'informations, consultez Vérifications du statut du volume EBS (p. 1119).	20 novembre 2012

Fonction	Version de l'API	Description	Date de parution
Noyaux Linux		ID AKI mis à jour ; noyaux de distribution réorganisés ; section PVOps mise à jour.	13 novembre 2012
Instances M3	01-10-2012	Il existe de nouveaux types d'instance M3 Extra Large et M3 Double Extra Large. Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.	31 octobre 2012
Statut de demande d'instance Spot	01-10-2012	Le statut de demande d'instance Spot simplifie la détermination de l'état de vos demandes Spot.	14 octobre 2012
Nouvelle version Amazon Linux AMI		Amazon Linux AMI 2012.09 est disponible.	11 octobre 2012
Place de marché des instances réservées Amazon EC2	15-08-2012	Le Place de marché des instances réservées met en correspondance les vendeurs ayant des instances réservées Amazon EC2 dont ils n'ont plus besoin et les acheteurs en quête d'achat de capacité supplémentaire. Les instances réservées achetées et vendues via le Place de marché des instances réservées fonctionnent comme toute autre instance réservée, si ce n'est qu'il peut ne leur rester qu'une durée standard complète et qu'elles peuvent être vendues à différents prix.	11 septembre 2012
IOPS provisionnés SSD pour Amazon EBS	20-07-2012	Les volumes IOPS provisionnés SSD offrent des performances élevées et prévisibles pour les charges de travail gourmandes en E/S, telles que les applications de base de données, qui reposent sur des temps de réponse rapides et réguliers. Pour de plus amples informations, veuillez consulter Types de volume Amazon EBS (p. 1089).	31 juillet 2012
Instances d'E/S élevées pour Amazon EC2	15-06-2012	Les instances d'E/S élevées offrent des performances d'E/S disque très hautes et à faible latence à l'aide d'un stockage d'instance local basé sur SSD.	18 juillet 2012
Rôles IAM sur les instances Amazon EC2	01-06-2012	Les rôles IAM pour Amazon EC2 fournissent : • Clés d'accès AWS pour les applications s'exécutant sur des instances Amazon EC2. • Rotation automatique des clés d'accès AWS sur l'instance Amazon EC2. • Autorisations détaillées pour les applications s'exécutant sur les instances Amazon EC2 qui adressent des demandes à vos services AWS.	11 juin 2012

Fonction	Version de l'API	Description	Date de parution
Fonctions d'instance Spot qui facilitent le démarrage et la gestion d'une interruption potentielle.		Vous pouvez désormais gérer vos Instances Spot comme suit : • Spécifiez le montant que vous êtes prêt à payer pour des Instances Spot à l'aide de configurations de lancement Auto Scaling et configurez un calendrier pour indiquer ce montant pour des Instances Spot. Pour plus d'informations, consultez Lancement d'Instances Spot dans votre groupe Auto Scaling dans le Amazon EC2 Auto Scaling Guide de l'utilisateur. • Obtenez des notifications quand les instances sont lancées ou terminées. • Utilisez les modèles AWS CloudFormation pour lancer les Instances Spot dans une pile avec des ressources AWS.	7 juin 2012
Exportation d'instance EC2 et horodatage des contrôles de statut pour Amazon EC2	01-05-2012	Ajout de la prise en charge des horodatages sur le statut d'instance et le statut système pour indiquer la date et l'heure auxquelles un contrôle d'état a échoué.	25 mai 2012
Exportation d'instance EC2 et horodatages des contrôles de statut d'instance et de statut système pour Amazon VPC	01-05-2012	Ajout de la prise en charge de l'exportation d'instance EC2 vers Citrix Xen, Microsoft Hyper-V et VMware vSphere. Ajout de la prise en charge des horodatages dans les contrôles de statut d'instance et de statut système.	25 mai 2012
Instances cluster compute 8 Extra Large	01-04-2012	Ajout de la prise en charge des instances `cc2.8xlarge` dans un VPC.	26 avril 2012
AMI AWS Marketplace	01-04-2012	Ajout de la prise en charge des AMI AWS Marketplace. .	19 avril 2012
Nouvelle version Amazon Linux AMI		Amazon Linux AMI 2012.03 est disponible.	28 mars 2012
Nouvelle version AKI		La version AKI 1.03 et les AKI pour la région AWS GovCloud (US) sont disponibles.	28 mars 2012
Instances medium, prise en charge de 64 bits sur toutes les AMI et client SSH basé Java	15-12-2011	Ajout de la prise en charge d'un nouveau type d'instance et des informations 64 bits. Ajout de procédures pour l'utilisation du client SSH basé Java pour se connecter aux instances Linux.	7 mars 2012
Niveaux de tarification des instances réservées	15-12-2011	Ajout d'une nouvelle section expliquant comment tirer parti de la tarification des remises intégrée aux niveaux de tarification des instances réservées.	5 mars 2012

Fonction	Version de l'API	Description	Date de parution
Interfaces ENI (interface réseau Elastic) pour les instances EC2 dans Amazon Virtual Private Cloud	01-12-2011	Ajout d'une nouvelle section sur les interfaces ENI (interface réseau Elastic) pour les instances EC2 d'un VPC. Pour plus d'informations, consultez Interfaces réseau Elastic (p. 840).	21 décembre 2011
Nouveaux AKI et régions GRU		Ajout d'informations sur la publication de nouveaux AKI pour la région SA-East-1. Cette version remplace AKI version 1.01. AKI version 1.02 continuera d'être rétrocompatible.	14 décembre 2011
Nouveaux types d'offre pour les instances réservées Amazon EC2	01-11-2011	Vous avez le choix entre différentes offres d'instances réservées qui prennent en compte votre utilisation projetée de l'instance.	01 décembre 2011
Statut d'instance Amazon EC2	01-11-2011	Vous pouvez ajouter des détails supplémentaires sur le statut de vos instances, y compris les événements programmés planifiés par AWS susceptibles d'avoir un impact sur vos instances. Ces activités opérationnelles incluent les redémarrages d'instance requis pour appliquer les mises à jour logicielles ou les correctifs de sécurité, ou les exigences d'instance requises en cas de problèmes matériels. Pour plus d'informations, consultez Surveillance du statut de vos instances (p. 742).	16 novembre 2011
Type d'instance cluster compute Amazon EC2		Ajout de la prise en charge de cluster compute 8 Extra Large (cc2.8xlarge) dans Amazon EC2.	14 novembre 2011
Nouveaux AKI et régions PDX		Ajout d'informations sur la publication de nouveaux AKI pour la nouvelle région US-West 2.	8 novembre 2011
Instances Spot dans Amazon VPC	15-07-2011	Ajout d'informations sur la prise en charge des Instances Spot dans Amazon VPC. Avec cette mise à jour, les utilisateurs peuvent lancer des Instances Spot dans un Virtual Private Cloud (VPC). En lançant des Instances Spot dans un VPC, les utilisateurs d'Instances Spot peuvent profiter des avantages de Amazon VPC.	11 octobre 2011
Nouvelle version Amazon Linux AMI		Ajout d'informations sur la publication d'Amazon Linux AMI 2011.09. Cette mise à jour supprime la balise beta d'Amazon Linux AMI, prend en charge la possibilité de verrouiller les référentiels sur une version spécifiques et fournit une notification quand les mises à jour sont disponibles pour les packages installés, y compris les mises à jour de sécurité.	26 septembre 2011

Fonction	Version de l'API	Description	Date de parution
Processus VM Import simplifié pour les utilisateurs des outils d'interface ligne de commande	15-07-2011	Le processus VM Import est simplifié avec la fonctionnalité améliorée d'`ImportInstance` et d'`ImportVolume`, qui désormais effectuent le chargement des images dans Amazon EC2 après avoir créé la tâche d'importation. De plus, avec l'introduction de la commande `ResumeImport`, les utilisateurs peuvent redémarrer un chargement incomplet au point où la tâche s'est arrêtée.	15 septembre 2011
Prise en charge de l'importation au format de fichier VHD		VM Import peut désormais importer les fichiers image de machine virtuelle au format VHD. Le format de fichier VHD compatible avec les plateformes de virtualisation Citrix Xen et Microsoft Hyper-V. Avec cette version, VM Import prend désormais en charge les formats d'image RAW, VHD et VMDK (compatible VMware ESX). Pour plus d'informations, consultez la VM Import/Export Guide de l'utilisateur.	24 août 2011
Mise à jour d'Amazon EC2 VM Import Connector pour VMware vCenter		Ajout d'informations sur la version 1.1 version de l'appliance virtuelle Amazon EC2 VM Import Connector pour VMware vCenter (connecteur). Cette mise à jour inclut la prise en charge du proxy pour l'accès Internet, une meilleure gestion des erreurs, une précision accrue de la barre d'avancement des tâches et plusieurs correctifs de bogue.	27 juin 2011
Activation de Linux AMI pour exécuter les noyaux fournis par l'utilisateur		Ajout d'informations sur le changement de version d'AKI de 1.01 en 1.02. Cette version met à jour le PVGRUB pour traiter les défaillances de lancement associées aux instances Linux t1.micro. Pour plus d'informations, consultez Activation de vos propres noyaux Linux (p. 205).	20 juin 2011
Modifications de tarification des zones de disponibilité des Instances Spot	15-05-2011	Ajout d'informations sur la fonction de tarification des zones de disponibilité des Instances Spot. Dans cette version, nous avons ajouté les options de tarification des zones de disponibilité, comme parties intégrantes des informations retournées quand vous interrogez les demandes d'instance Spot et l'historique des prix Spot. Ces ajouts permettent de déterminer plus facilement le prix requis pour lancer une instance Spot dans une zone de disponibilité particulière.	26 mai 2011
AWS Identity and Access Management		Ajout d'informations sur AWS Identity and Access Management (IAM), qui permet aux utilisateurs de spécifier les actions Amazon EC2 qu'il est possible d'utiliser avec les ressources Amazon EC2 en général. Pour plus d'informations, consultez Identity and Access Management pour Amazon EC2 (p. 976).	26 avril 2011

Fonction	Version de l'API	Description	Date de parution
Activation de Linux AMI pour exécuter les noyaux fournis par l'utilisateur		Ajout d'informations sur l'activation d'une AMI Linux pour utiliser l'AKI (Amazon Kernel Image) PVGRUB afin d'exécuter un noyau fourni par l'utilisateur. Pour plus d'informations, consultez Activation de vos propres noyaux Linux (p. 205).	26 avril 2011
Instances dédiées		Lancées au sein de votre Amazon Virtual Private Cloud (Amazon VPC), les instances dédiées sont des instances physiquement isolées au niveau matériel hôte. Les instances dédiées vous laissent tirer profit des avantages d'Amazon VPC et du cloud AWS : mise en service Elastic à la demande et facturation limitée à ce que vous consommez, tout en isolant vos instances de calcul Amazon EC2 au niveau matériel. Pour plus d'informations, consultez Instances dédiées (p. 498).	27 mars 2011
Mises à jour des instances réservées dans AWS Management Console		Les mises à jour d'AWS Management Console simplifient pour les utilisateurs l'affichage de leurs instances réservées et l'achat d'instances réservées supplémentaires, y compris les instances réservées dédiées. Pour plus d'informations, consultez Instances réservées (p. 329).	27 mars 2011
Nouvel AMI de référence Amazon Linux		Le nouvel AMI de référence Amazon Linux remplace l'AMI de référence CentOS. Suppression d'informations sur l'AMI de référence CentOS, y compris la section nommée Correction de la dérive d'horloge pour les instances cluster de l'AMI CentOS 5.4.	15 mars 2011
Informations de métadonnées	01-01-2011	Ajout d'informations sur les métadonnées pour refléter les modifications de la version 2011-01-01. Pour plus d'informations, consultez Métadonnées d'instance et données utilisateur (p. 701) et Catégories de métadonnées d'instance (p. 721).	11 mars 2011
Amazon EC2 VM Import Connector pour VMware vCenter		Ajout d'informations sur l'appliance virtuelle Amazon EC2 VM Import Connector pour VMware vCenter (connecteur). Le connecteur est un plug-in pour VMware vCenter qui s'intègre à VMware vSphere Client et fournit une interface utilisateur graphique que vous pouvez utiliser pour importer vos machines virtuelles VMware sur Amazon EC2.	3 mars 2011
Forcer le détachement du volume		Vous pouvez désormais utiliser AWS Management Console pour forcer le détachement d'un volume Amazon EBS d'une instance. Pour plus d'informations, consultez Détacher un volume Amazon EBS d'une instance Linux (p. 1127).	23 février 2011

Fonction	Version de l'API	Description	Date de parution
Protection de la fin d'instance		Vous pouvez désormais utiliser AWS Management Console pour empêcher qu'une instance ne finisse. Pour plus d'informations, consultez Activation de la protection contre la résiliation (p. 648).	23 février 2011
Correction de la dérive d'horloge pour les instances cluster sur l'AMI CentOS 5.4		Ajout d'informations sur la façon de corriger la dérive d'horloge pour les instances s'exécutant sur l'AMI CentOS d'Amazon.	25 janvier 2011
VM Import	15-11-2010	Ajout d'informations sur VM Import, qui vous permet d'importer un volume ou une machine virtuelle dans Amazon EC2. Pour plus d'informations, consultez la VM Import/Export Guide de l'utilisateur.	15 décembre 2010
Surveillance basique pour les instances	31-08-2010	Ajout d'informations sur la surveillance basique pour les instances EC2.	12 décembre 2010
Filtres et balises	31-08-2010	Ajout d'informations sur les ressources d'affichage, de filtrage et de balisage. Pour plus d'informations, consultez Liste et filtrage de vos ressources (p. 1303) et Balisage de vos ressources Amazon EC2 (p. 1309).	19 septembre 2010
Lancement d'instance idempotente	31-08-2010	Ajout d'informations pour garantir l'idempotence lors de l'exécution des instances. Pour plus d'informations, consultez Garantie de l'idempotence dans le manuel Amazon EC2 API Reference.	19 septembre 2010
Micro-instances	15-06-2010	Amazon EC2 propose le type d'instance `t1.micro` pour certains types d'applications. Pour plus d'informations, consultez Instances à capacité extensible (p. 233).	8 septembre 2010
AWS Identity and Access Management pour Amazon EC2		Amazon EC2 s'intègre désormais à AWS Identity and Access Management (IAM). Pour plus d'informations, consultez Identity and Access Management pour Amazon EC2 (p. 976).	2 septembre 2010
Instances cluster	15-06-2010	Amazon EC2 propose les instances cluster compute pour les applications de Calcul Haute Performance (HPC). Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.	12 juillet 2010
Désignation de l'adresse IP Amazon VPC	15-06-2010	Les utilisateurs Amazon VPC peuvent désormais spécifier l'adresse IP pour attribuer une instance lancée dans un VPC.	12 juillet 2010
Surveillance Amazon CloudWatch pour les volumes Amazon EBS		La surveillance Amazon CloudWatch est désormais automatiquement disponible pour les volumes Amazon EBS. Pour plus d'informations, consultez Métriques Amazon CloudWatch pour Amazon EBS (p. 1247).	14 juin 2010

Fonction	Version de l'API	Description	Date de parution
Instances Extra Large à mémoire élevée	30-11-2009	Amazon EC2 prend désormais en charge un type d'instance Extra Large (m2.xlarge) à mémoire élevée. Pour de plus amples informations sur les spécifications matérielles pour chaque type d'instance Amazon EC2, veuillez consulter Types d'instance Amazon EC2.	22 février 2010