

Amazon Elastic Compute Cloud (Amazon EC2)

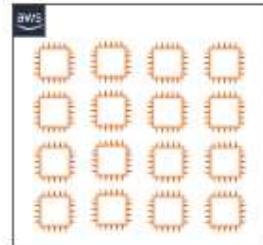


Photo de Taylor Vitit sur Unsplash

Serveurs sur site

Exemples d'utilisation des instances Amazon EC2

- ✓ Serveur d'application
- ✓ Serveur web
- ✓ Serveur de base de données
- ✓ Serveur de jeu
- ✓ Serveur de messagerie
- ✓ Serveur de médias
- ✓ Serveur de catalogue
- ✓ Serveur de fichiers
- ✓ Serveur de calcul
- ✓ Serveur proxy



Instances Amazon EC2



Photo de geralt sur Pixabay



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

9

L'exécution de serveurs sur site est coûteuse. Le matériel doit être acheté. Toutefois, cet approvisionnement s'appuie souvent sur des estimations plutôt que sur l'utilisation concrète des serveurs. Il faut créer des centres de données, les pourvoir en personnel et les maintenir : tout cela a un prix. Les entreprises doivent également prévoir en permanence une quantité suffisante de machines pour gérer les pics de trafic et les pics de charge de travail. Une fois les déploiements classiques sur site effectués, une partie de la capacité du serveur peut être inutilisée et inactive pendant un certain laps de temps au cours de l'exécution des serveurs, ce qui représente un gaspillage.

Amazon Elastic Compute Cloud (Amazon EC2) fournit des machines virtuelles sur lesquelles vous pouvez héberger les mêmes types d'applications que vous exécutez sur un serveur classique hébergé sur site. Ce service offre une capacité de calcul dans le cloud à la fois redimensionnable et sécurisée. Les instances EC2 prennent en charge diverses charges de travail. Parmi les utilisations courantes des instances EC2, citons (liste non exhaustive) :

- Serveurs d'application
- Serveurs web
- Serveurs de bases de données
- Serveurs de jeu
- Serveurs de messagerie
- Serveurs de médias
- Serveurs de catalogue
- Serveurs de fichiers
- Serveurs de calcul
- Serveurs proxy

Présentation d'Amazon EC2

• Amazon Elastic Compute Cloud (Amazon EC2)



- Fournit des machines virtuelles (appelées instances EC2) dans le cloud.
- Permet un *contrôle total* du système d'exploitation invité (Windows ou Linux) sur chaque instance.
- Vous pouvez lancer des instances, quelle que soit leur taille, dans une zone de disponibilité n'importe où dans le monde.
 - Lancez des instances à partir d'[Amazon Machine Images \(AMI\)](#).
 - Lancez en quelques clics ou à partir d'une ligne de code des instances qui seront prêtes en quelques minutes.
- Vous pouvez contrôler le trafic entrant et sortant des instances.



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

10

La dénomination EC2 dans Amazon EC2 est l'abréviation de Elastic Compute Cloud: **Elastic** fait référence à la possibilité offerte d'augmenter ou de diminuer facilement le nombre de serveurs utilisés pour prendre en charge une application automatiquement. Vous pouvez également augmenter ou diminuer la taille des serveurs existants. **Calcul** fait référence à la raison pour laquelle la plupart des utilisateurs exécutent des serveurs en premier lieu (hébergement des applications en cours d'exécution ou traitement des données). Ces actions nécessitent des ressources de calcul, y compris de la puissance de traitement(CPU) et de la mémoire(RAM). Et **Cloud** fait référence au fait que les instances EC2 que vous exécutez sont hébergées dans le cloud.

Amazon EC2 offre des machines virtuelles dans le cloud et donne un contrôle administratif total sur le système d'exploitation Windows ou Linux qui s'exécute sur l'instance. La plupart des systèmes d'exploitation sont pris en charge: Windows 2008, 2012, 2016 et 2019, Red Hat, SuSE, Ubuntu et Amazon Linux.

Un système d'exploitation qui s'exécute sur une machine virtuelle est souvent appelé système d'exploitation invité pour le distinguer du système d'exploitation hôte. Le système d'exploitation hôte est directement installé sur n'importe quel serveur hébergeant une ou plusieurs machines virtuelles.

Avec Amazon EC2, vous pouvez lancer un nombre illimité d'instances, quelle que soit leur taille, dans n'importe quelle zone de disponibilité, n'importe où dans le monde et ce en quelques minutes. Les instances sont lancées à partir d'Amazon Machine Image (AMI), qui sont en fait des modèles de machine virtuelle.

Lancement d'une instance AmazonEC2

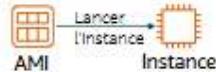
Lorsque vous lancerez une instance Amazon EC2 pour la première fois, vous utiliserez probablement l'assistant de lancement d'instance de la Console de gestion AWS.

L'assistant de lancement d'instance vous aide à lancer vos instances. Par exemple, si vous acceptez tous les paramètres par défaut, vous pouvez ignorer la plupart des étapes proposées par l'assistant et lancer une instance EC2 en quelques clics.

1. Sélection d'une AMI

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'Instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés



- Amazon Machine Image (AMI)

- Modèle utilisé pour créer une instance EC2 (**machine virtuelle, ou VM, qui est exécutée dans AWS Cloud**)
- Contient un système d'exploitation **Windows ou Linux**
- Certains **logiciels** sont souvent préinstallés

- Options d'AMI :

- Quick Start : *AMI Linux et Windows fournies par AWS*
- Mes AMI : *les AMI que vous avez créées*
- AWS Marketplace : *modèles préconfigurés de tiers*
- AMI de la communauté : *AMI partagées par des tiers ; leur utilisation est à vos risques*



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

12

Une Amazon Machine Image (AMI) fournit les informations nécessaires au lancement d'une instance EC2. Vous devez spécifier une AMI source lorsque vous lancez une instance. Vous pouvez utiliser différentes images AMI pour lancer différents types d'instances. Par exemple, vous pouvez choisir une AMI pour lancer une instance correspondant à un serveur Web et une autre AMI pour déployer une instance qui hébergera un serveur d'applications. Vous pouvez lancer plusieurs instances à partir d'une même image AMI.

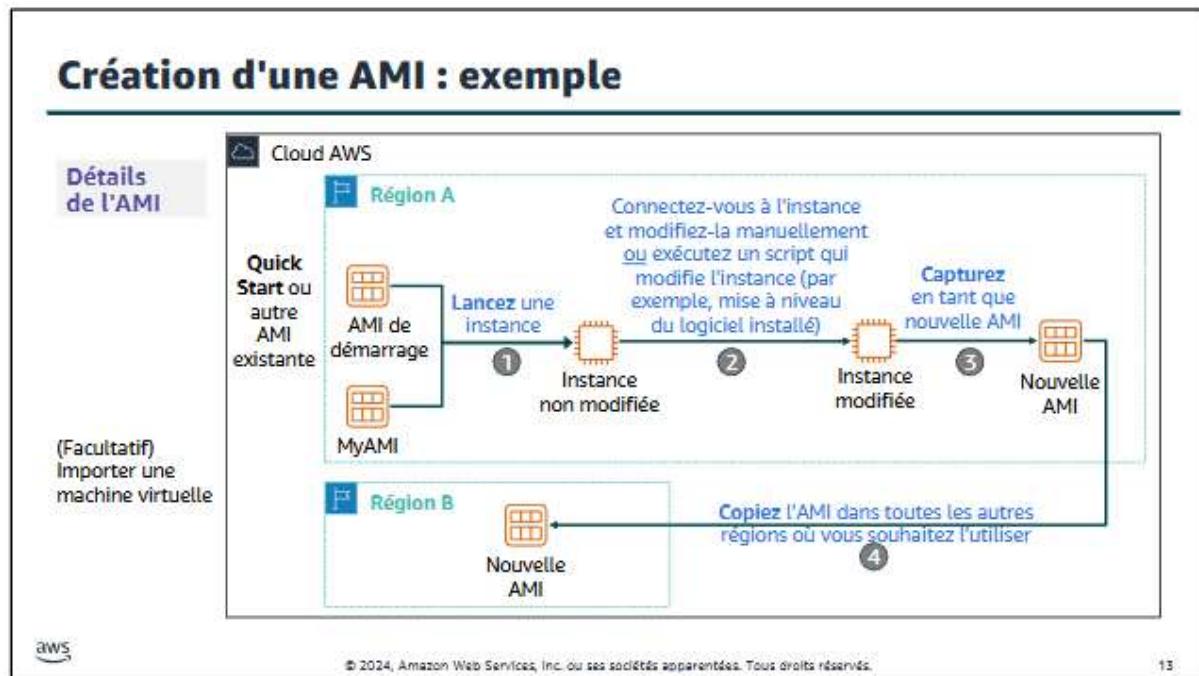
Une image AMI comprend les composants suivants :

- Un modèle pour le volume racine de l'instance. Un volume racine contient généralement un système d'exploitation et tout ce qui a été installé sur ce système d'exploitation (applications, bibliothèques, etc.). Le service Amazon EC2 copie le modèle sur le volume racine d'une nouvelle instance EC2, puis la démarre.
- Des autorisations de lancement qui gèrent les comptes AWS pouvant utiliser l'image AMI.
- Un mappage de périphérique de stockage en mode bloc qui indique le volume à attacher à l'instance (le cas échéant) lorsqu'elle est lancée.

Vous pouvez choisir plusieurs AMI :

- Quick Start : AWS propose plusieurs AMI prédéfinies pour lancer vos instances.
- Ces AMI incluent de nombreuses options Linux et Windows.
- Mes AMI : AMI que vous avez créé.
- AWS Marketplace : propose un catalogue numérique répertoriant des milliers de solutions logicielles. Ces AMI peuvent offrir des cas d'utilisation spécifiques pour aider à démarrer rapidement.
- AMI de la communauté : ces AMI sont créées par des tiers à travers le monde. **Elles ne sont pas vérifiées par AWS. Leurs utilisation est donc à vos risques.**

Création d'une AMI : exemple



Une AMI est créé à partir d'une instance EC2. Vous pouvez importer une machine virtuelle afin qu'elle devienne une instance EC2, puis enregistrer l'instance EC2 en tant qu'AMI. Ensuite lancer une instance EC2 à partir de cette AMI. Vous avez également la possibilité de commencer avec une AMI existante, telles que les AMI Quick Start fournies par AWS, pour créer une instance EC2.

Quelles que soient les options choisies (étape1), vous aurez ce que l'on appelle une instance non modifiée. À partir de cette instance, vous pouvez ensuite créer une instance principale, à savoir une machine virtuelle que vous avez configurée avec le système d'exploitation et les paramètres d'application spécifiques que vous souhaitez (étape2), puis la capturer en tant que nouvelle AMI (étape3).

Lorsque vous créez une AMI, AmazonEC2 arrête l'instance, génère un instantané de son volume racine, puis enregistre l'instantané en tant qu'AMI.

Une fois l'AMI enregistrée, elle peut être utilisée pour lancer de nouvelles instances dans la même région AWS. La nouvelle AMI peut maintenant être considérée comme une nouvelle AMI de démarrage. Vous souhaiterez peut-être également copier l'AMI dans d'autres régions (étape4), afin que les instances EC2 puissent également être lancées dans ces emplacements.

2. Sélection d'un type d'instance

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés

- + Votre cas d'utilisation
 - Comment l'instance EC2 que vous créez sera-t-elle utilisée ?
- + Le type d'instance que vous choisissez est associé aux éléments suivants :
 - Mémoire (RAM)
 - Puissance de traitement (CPU)
 - Espace disque et type de disque (stockage)
 - Performances réseau
- + Catégories des types d'Instances :
 - Polyvalent
 - Optimisée pour le calcul
 - Optimisée pour la mémoire
 - Optimisée pour le stockage
 - Calcul accéléré
- + Les types d'Instances offrent les options suivantes : *famille, génération et taille*



© 2006, Amazon Web Services, Inc. ou ses sociétés affiliées. Tous droits réservés.

14

Après avoir choisi l'AMI pour lancer l'instance, vous devez sélectionner un type d'instance.

Amazon EC2 offre une sélection de types d'instances optimisés pour différents cas d'utilisation. Les types d'instances se composent de plusieurs capacités liées au CPU, à la mémoire, au stockage et à la mise en réseau. Les différents types d'instances vous permettent de choisir une combinaison de ressources appropriée pour vos applications. Chaque type d'instance est disponible dans une ou plusieurs tailles d'instance, ce qui vous permet de mettre à l'échelle vos ressources en fonction des besoins de la charge de travail visée.

Les catégories des types d'instances sont les suivantes: polyvalentes, optimisées pour le calcul, optimisées pour la mémoire, optimisées pour le stockage et de calcul accéléré. Chaque catégorie de type d'instance offre un choix important de types d'instances.

Dénomination et taille des types d'instances EC2

Dénomination des types d'instances

- Exemple : **t3.large**

- T correspond au nom de la famille
- 3 correspond au numéro de génération
- Large correspond à la taille

Exemples de tailles d'instance

Nom de l'instance	vCPU	Mémoire (Go)	Stockage
t3.nano	2	0,5	EBS uniquement
t3.micro	2	1	EBS uniquement
t3.small	2	2	EBS uniquement
t3.medium	2	4	EBS uniquement
t3.large	2	8	EBS uniquement
t3.xlarge	4	16	EBS uniquement
t3.2xlarge	8	32	EBS uniquement



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

15

Lorsque vous regardez un type d'instance EC2, vous verrez que son nom comporte plusieurs parties. Prenons par exemple le type T.

T est le nom de famille, suivi d'un chiffre. Ici, ce chiffre est 3. Le chiffre est le numéro de génération de ce type. Par conséquent, une instance t3 est la troisième génération de la famille T.

En général, les types d'instances dont le numéro de génération est plus élevé sont plus puissants et offrent un meilleur rapport qualité-prix.

La partie suivante du nom correspond à la portion taille de l'instance. Lorsque vous comparez les tailles, il est important de noter la portion coefficient de la catégorie de taille.

Par exemple, une instance t3.2xlarge possède une vCPU et une mémoire deux fois plus importantes qu'une instance t3.xlarge. L'instance t3.xlarge dispose, à son tour, une vCPU et une mémoire deux fois plus importantes qu'une instance t3.large.

Il est également important de noter que la bande passante du réseau est également liée à la taille de l'instance EC2. Si vous exécutez des tâches qui consomment beaucoup de réseau, vous devrez peut-être augmenter les spécifications de l'instance pour répondre à vos besoins.

Sélection du type d'instance en fonction du cas d'utilisation

	Polyvalent	Optimisée pour le calcul	Optimisée pour la mémoire	Calcul accéléré	Optimisée pour le stockage
Types d'instances	a1, m4, m5, t2, t3	c4, c5	r4, r5, x1, z1	f1, g3, g4, p2, p3	d2, h1, i3
Cas d'utilisation	Large	Hautes performances	Bases de données en mémoire	Machine learning	Systèmes de fichiers distribués



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

16

Les types d'instances varient sur plusieurs points, notamment le type de CPU, le nombre de processeurs ou de coeurs, le type et la quantité de stockage, la quantité de mémoire et les performances réseau. Le graphique présente les différentes catégories d'instances, ainsi que les familles des types d'instances et des numéros de génération qui correspondent à chaque type de catégorie. Examinons plus en détail certains types d'instances:

Les instances T3 sont des instances à capacité extensible et polyvalentes qui servent de référence pour les performances en CPU, avec la possibilité de dépasser le seuil de base. Les cas d'utilisation de ce type d'instance comprennent les sites web et les applications web, les environnements de développement, les serveurs de développement, les référentiels de codes, les micro-services, les environnements de test et intermédiaires ainsi que les applications professionnelles.

Les instances C5 sont optimisées pour les applications de calcul intensives et offrent des performances élevées rentables pour un rapport prix/performances de calcul avantageux. Les cas d'utilisation incluent la modélisation scientifique, le traitement par lots, la diffusion de publicités, les jeux multi-joueurs pouvant être mis à l'échelle et l'encodage vidéo.

Les instances R5 sont optimisées pour les applications exigeantes en mémoire. Les cas d'utilisation comprennent des bases de données hautes performances, une exploitation et une analyse des données, des bases de données en mémoire, des caches en mémoire distribués à l'échelle web, des applications procédant au traitement en temps réel de données Big Data non structurées, des clusters Apache Hadoop ou Apache Spark et d'autres applications d'entreprise.

En plus de prendre en compte les besoins en matière de processeur, de RAM et de stockage de vos charges de travail, il est également important de prendre en considération les besoins en bande passante de son réseau. Chaque type d'instance fournit un niveau de performance réseau documenté. Par exemple, une instance a1.medium fournit jusqu'à 10Gbit/s, mais une instance p3dn.24xlarge peut aller jusqu'à 100Gbit/s. Choisissez un type d'instance qui répond à vos besoins.

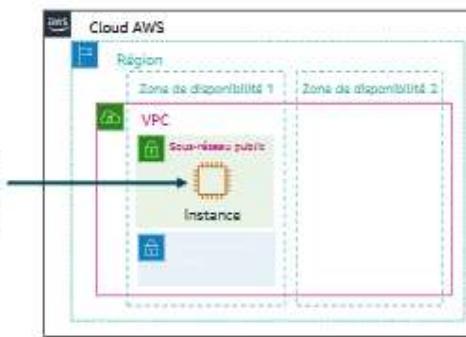
3. Définition des paramètres réseau

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés

- Où voulez-vous déployer l'instance ?
 - Identifiez le **VPC** et éventuellement le **sous-réseau**.
 - Voulez-vous que **l'adresse IP publique** soit attribuée automatiquement ?
 - Pour la rendre accessible sur Internet

Exemple : indiquez ici l'emplacement de déploiement de l'instance



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

18

Après avoir choisi une **AMI** et un **type d'instance**, vous devez indiquer l'emplacement réseau où doit être déployée l'instance EC2. Le choix de la région doit être fait avant de lancer l'assistant de lancement d'instance. **Vérifiez que vous êtes sur la page Région appropriée de la console Amazon EC2 avant de choisir Lancer une instance.**

Lorsque vous lancez une instance dans un VPC par défaut, AWS lui attribue par défaut une adresse IP publique. Lorsque vous lancez une instance dans un +VPC qui n'est pas défini par défaut, le sous-réseau possède un attribut qui détermine si les instances lancées dans ce sous-réseau reçoivent une adresse IP publique à partir du groupe d'adresses IPv4 publiques. Par défaut, AWS n'attribue pas d'adresse IP publique aux instances lancées dans un sous-réseau autre que celui par défaut.

Vous déterminez si votre instance recevra ou non une adresse IP publique soit en modifiant l'attribut d'adressage IP public de votre sous-réseau, soit en activant ou en désactivant la fonction d'adressage IP public lors du lancement (qui remplacera l'attribut d'adressage IP public du sous-réseau).

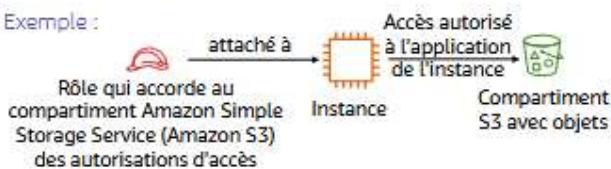
4. Attribution d'un rôle IAM (facultatif)

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés

- Les logiciels de l'instance EC2 doivent-ils communiquer avec d'autres services AWS ?
 - Si oui, attribuez un rôle IAM approprié
- Un rôle de Gestion des identités et des accès AWS (AWS IAM) attaché à une instance EC2 est conservé dans un profil d'instance.
- Vous n'êtes pas limité à l'attribution d'un rôle uniquement au lancement de l'instance.
- Vous pouvez également attacher un rôle à une instance qui existe déjà.

Exemple :



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

19

Il est courant d'utiliser des instances EC2 pour exécuter une application qui doit exécuter des appels API sécurisés vers d'autres services AWS. Pour prendre en charge ces cas d'utilisation, AWS vous permet d'attribuer un rôle de Gestion des identités et des accès AWS (AWS IAM) à une instance EC2. Sans cette fonction, vous pouvez décider de placer les autorisations AWS sur une instance EC2 pour pouvoir utiliser une application de cette instance. Toutefois, vous ne devriez jamais stocker les autorisations AWS sur une instance EC2. Cette approche n'est pas sûre.

Il est préférable d'attacher un rôle IAM à l'instance EC2. Le rôle IAM autorisera alors les demandes d'Application Program Interface (API) envoyées aux applications qui s'exécutent sur l'instance EC2.

Un profil d'instance est un conteneur pour un rôle IAM. Si vous utilisez la Console de gestion AWS pour créer un rôle pour Amazon EC2, la console crée automatiquement un profil d'instance et lui attribue le même nom que le rôle. Lorsque vous utiliserez ensuite la console Amazon EC2 pour lancer une instance avec un rôle IAM, vous pourrez associer ce rôle à une instance.

Dans la console, la liste qui s'affiche est une liste de noms de profils d'instance.

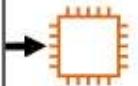
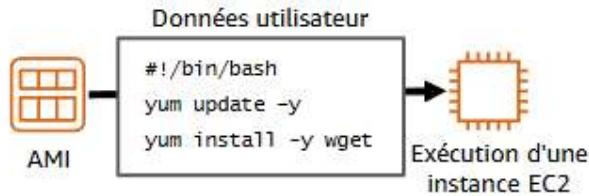
Dans l'exemple du haut, le rôle IAM est utilisé pour accorder des autorisations à une application qui s'exécute sur une instance EC2. L'application doit accéder à un compartiment Amazon S3. Vous pouvez attacher un rôle IAM lorsque vous lancez l'instance, mais vous pouvez également attacher un rôle à une instance EC2 déjà en cours d'exécution.

Lorsque vous définissez un rôle à utiliser par une instance EC2, vous définissez les comptes ou les services AWS qui peuvent adopter ce rôle. Si vous modifiez un rôle, la modification est propagée à toutes les instances aux quelles le rôle est attaché.

5. Script de données utilisateur (facultatif)

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés



Exécution d'une instance EC2

- Vous pouvez éventuellement spécifier un script de données utilisateur au lancement de l'instance
- Utilisez des scripts de données utilisateur pour personnaliser l'environnement d'exécution de votre instance
 - Le script s'exécute au démarrage initial de l'instance
- Peut être utilisé de manière stratégique
 - Par exemple, pour réduire le nombre d'AMI personnalisées créées et gérées



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

20

Lorsque vous créez des instances EC2, vous avez la possibilité de transmettre les données utilisateur à l'instance. Les données utilisateur peuvent automatiser l'exécution des installations et des configurations au lancement de l'instance. Par exemple, un script de données utilisateur peut appliquer des correctifs et mettre à jour le système d'exploitation de l'instance, récupérer et installer des clés de licence logicielle ou installer des logiciels supplémentaires.

Lorsque l'instance EC2 est créée, le script de données utilisateur s'exécute avec les priviléges du compte racine pendant les étapes finales du processus de démarrage. Sur les instances Linux, l'exécution est effectuée par le service cloud-init. Sur les instances Windows, l'exécution est effectuée par l'utilitaire EC2Config ou EC2Launch. Par défaut, les données utilisateur s'exécutent uniquement au démarrage initial de l'instance. Toutefois, si vous souhaitez que votre script de données utilisateur soit exécuté à chaque démarrage de l'instance, vous pouvez créer un script de données utilisateur MIME (Multipurpose Internet Mail Extensions) (ce processus n'est pas courant).

6. Configuration du stockage

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés

- Configurez le volume racine
 - Emplacement d'installation du système d'exploitation invité
- Attachez des volumes de stockage supplémentaires (facultatif)
 - L'AMI peut déjà inclure plusieurs volumes
- Pour chaque volume, indiquez :
 - La taille du disque (en Go)
 - Le type de volume
 - Différents types de disques SSD (Solid State Drive) et de disques durs (HDD) sont disponibles
 - Si le volume doit être supprimé lorsque l'instance est résiliée
 - Si vous souhaitez utiliser le chiffrement



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

21

Lorsque vous lancez une instance EC2, vous pouvez configurer les options de stockage. Par exemple, vous pouvez configurer la taille du volume racine sur lequel le système d'exploitation invité est installé. Vous pouvez également attacher des volumes de stockage supplémentaires lorsque vous lancez l'instance. Certaines images AMI sont également configurées pour lancer plusieurs volumes de stockage par défaut afin de fournir un espace de stockage distinct du volume racine.

Pour chaque volume de votre instance, vous pouvez définir la taille des disques, les types de volume et indiquer si le stockage est conservé lorsque l'instance est résiliée. Vous pouvez également choisir d'utiliser le chiffrement.

Options de stockage Amazon EC2

- **Amazon Elastic Block Store (Amazon EBS) :**
 - Volumes de stockage durables de type bloc
 - Si vous arrêtez l'instance et que vous la redémarrez, les données seront toujours présentes.
- **Stockage d'instance Amazon EC2 :**
 - Le stockage éphémère concerne les disques attachés à l'ordinateur hôte sur lequel l'instance EC2 est exécutée
 - Si l'instance s'arrête, les données stockées sont supprimées.
- **Autres options de stockage (ne concerne pas le volume racine) :**
 - Montez un système de fichiers **Amazon Elastic File System (Amazon EFS)**.
 - Connectez-vous à **Amazon Simple Storage Service (Amazon S3)**.



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

22

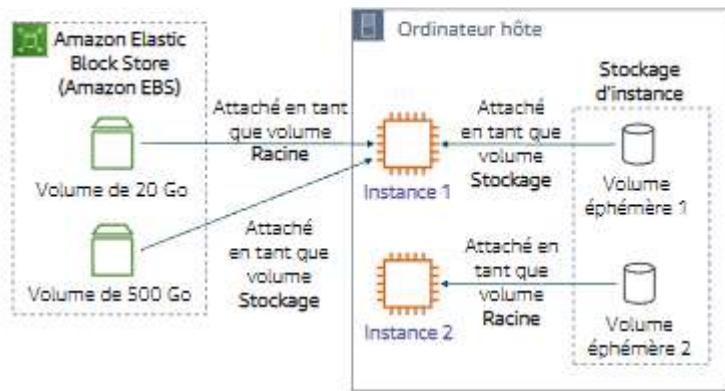
Amazon Elastic Block Store (Amazon EBS) est un service de stockage par bloc durable, facile à utiliser et à hautes performances, conçu pour être utilisé avec Amazon EC2 pour les charges de travail à débit élevé et à forte intensité de transactions. Avec Amazon EBS, vous pouvez choisir parmi quatre types de volumes différents afin de trouver un juste équilibre entre le prix et les performances. Vous pouvez modifier les types de volumes ou augmenter leur taille sans interrompre l'exécution de vos applications essentielles, ce qui vous permet de bénéficier d'un espace de stockage rentable dès que vous en avez besoin.

Le stockage d'instance Amazon EC2 fournit un espace de stockage par blocs éphémère (ou temporaire) pour votre instance. Ce stockage est localisé sur les disques physiquement associés à l'ordinateur hôte. Le stockage d'instance convient parfaitement lorsque vous devez stocker temporairement des informations qui changent fréquemment, telles que les tampons, les caches, les données temporaires et autres contenus temporaires. Vous pouvez également utiliser le stockage d'instance pour les données répliquées sur une flotte d'instances, comme un groupe de serveurs web à charge équilibrée. Si les instances sont arrêtées, en raison d'une erreur utilisateur ou d'une défaillance, les données du stockage d'instance seront supprimées.

Exemples d'options de stockage

- Caractéristiques de l'instance 1 :

- Inclut un type de volume racine Amazon EBS pour le système d'exploitation.
- Que se passe-t-il si l'instance est arrêtée puis redémarrée ?



- Caractéristiques de l'instance 2 :

- Possède un type de volume racine Stockage d'instance pour le système d'exploitation.
- Que se passe-t-il si l'instance s'arrête (en raison d'une erreur utilisateur ou d'une erreur système) ?



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

23

Si l'instance 1 était arrêtée puis redémarrée, le système d'exploitation survivrait et toutes les données stockées sur le volume Amazon EBS de 20Go ou de 500Go resteraient intactes. Cependant, toutes les données stockées sur le volume éphémère 1 seraient définitivement perdues.

L'exemple Instance2 montre que le volume racine se trouve sur un stockage d'instance (volume éphémère2). Une instance avec un volume racine de stockage d'instance ne peut pas être arrêtée par un appel d'API Amazon EC2. Elle ne peut qu'être résiliée. Cependant, elle pourrait être arrêtée depuis le système d'exploitation de l'instance (par exemple, en exécutant une commande shutdown) ou elle pourrait s'arrêter en raison d'un échec du système d'exploitation ou du disque, ce qui entraînerait la résiliation de l'instance. Si l'instance est résiliée, toutes les données stockées sur le volume éphémère 2 seront perdues, y compris le système d'exploitation. Vous ne pourrez pas redémarrer l'instance.

Par conséquent, ne comptez pas sur le stockage d'instance pour conserver des données précieuses sur le long terme. Utilisez plutôt un stockage de données plus durable, comme Amazon EBS, Amazon EFS ou Amazon S3.

7. Ajout de balises

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés

- Une **balise** est un étiquette que vous affectez à une ressource AWS.
 - Se compose d'une **clé** et d'une **valeur** facultative.
- Le **balisage** permet d'associer des **métadonnées** à une instance EC2.
- Avantages potentiels du **balisage** : filtrage, automatisation, répartition des coûts et contrôle d'accès.

Exemple :

Key	(128 characters maximum)	Value	(256 characters maximum)
Name		WebServer	
Add another tag (Up to 50 tags maximum)			



© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

24

Une balise est une étiquette que vous associez 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 les ressources AWS (telles que les instances EC2) de différentes manières. Par exemple, vous pouvez baliser les instances par objectif, propriétaire ou environnement. Le balisage permet d'associer des métadonnées à une instance EC2.

Les clés et valeurs des balises respectent la casse. Par exemple, une balise couramment utilisée pour les instances EC2 est une clé appelée Nom et une valeur qui décrit l'instance, par exemple Mon serveur web. La balise Nom est affichée par défaut sur la page Instances de la console AmazonEC2. Cependant, si vous créez une clé appelée nom (avec un n minuscule), elle n'apparaîtra pas dans la colonne Nom de la liste des instances (bien qu'elle apparaisse toujours dans le panneau des détails de l'instance dans l'onglet Balises).

8. Paramètres des groupes de sécurité

Options proposées par l'assistant de lancement d'instance :

1. AMI
2. Type d'instance
3. Paramètres réseau
4. Rôle IAM
5. Données utilisateur
6. Options de stockage
7. Étiquettes
8. Groupe de sécurité
9. Paire de clés

- Un groupe de sécurité est un **ensemble de règles de pare-feu** qui contrôlent le trafic entrant de l'instance.
 - Il existe *en dehors* du système d'exploitation invité de l'instance.
- Créez des **règles** qui indiquent la **source** et les **ports** que les communications réseau peuvent utiliser.
 - Indiquez le numéro de **port** et le **protocole**, par exemple TCP (Transmission Control Protocol), UDP (User Datagram Protocol) ou ICMP (Internet Control Message Protocol).
 - Indiquez la **source** (par exemple, une adresse IP ou un autre groupe de sécurité) autorisée à utiliser la règle.

Exemple de règle :

Type	Protocol	Port Range	Source
SSH	TCP	22	My IP 72.21.198.67/32

© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

25

Un groupe de sécurité fonctionne comme un pare-feu virtuel qui gère le trafic réseau d'une ou de plusieurs instances. Lorsque vous lancez une instance, vous pouvez définir un ou plusieurs groupes de sécurité. Si vous ne le faites pas, le groupe de sécurité par défaut est utilisé.

Vous pouvez ajouter des règles à chaque groupe de sécurité. Les règles autorisent le trafic sortant et entrant des instances associées. Vous pouvez modifier les règles d'un groupe de sécurité à tout moment. Les nouvelles règles sont appliquées automatiquement à toutes les instances associées au groupe de sécurité.

Lorsqu'AWS autorise le trafic vers une instance, toutes les règles de tous les groupes de sécurité associés à l'instance sont évaluées. Lorsque vous lancez une instance dans un cloud privé virtuel (VPC), vous devez créer un groupe de sécurité ou en utiliser un qui existe déjà dans ce VPC. Après avoir lancé une instance, vous pouvez modifier les groupes de sécurité associés.

Lorsque vous définissez une règle, vous pouvez indiquer la source autorisée de la communication réseau (règles entrantes) ou de destination (règles sortantes). La source peut être une adresse IP, une plage d'adresses IP, un autre groupe de sécurité, un point de terminaison d'un VPC de passerelle ou n'importe où (ce qui signifie que toutes les sources seront autorisées).

Par défaut, un groupe de sécurité inclut une règle de sortie qui autorise tout le trafic sortant. Vous pouvez retirer la règle et ajouter des règles sortantes qui n'autorisent qu'un trafic sortant spécifique. Si votre groupe de sécurité ne contient pas de règles sortantes, aucun trafic sortant provenant de votre instance n'est autorisé.

Dans l'exemple de règle plus haut, la règle autorise le trafic SSH (Secure Shell) sur le port 22 via le protocole TCP (Transmission Control Protocol)
Si la source de la demande est Mon IP.

L'adresse MyIP est calculée en déterminant l'adresse IP à partir de laquelle vous êtes actuellement connecté au Cloud AWS lorsque vous définissez la règle.

Les listes de contrôle d'accès réseau (listes ACL réseau) peuvent également être utilisées comme pare-feu pour protéger les sous-réseaux d'un VPC.

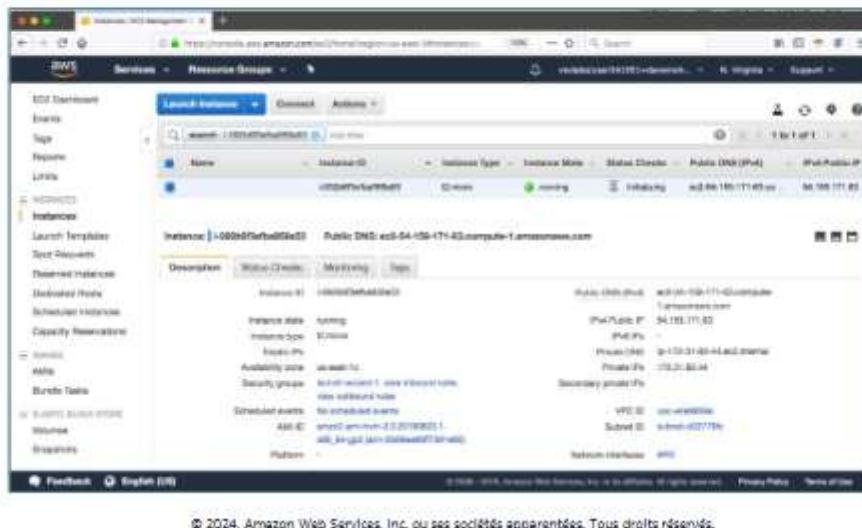
Une fois que vous avez défini toutes les configurations requises pour lancer une instance EC2 et après avoir personnalisé les paramètres de configuration facultatifs du Launch Wizard EC2, une fenêtre **Examiner le lancement de l'instance** est présentée. Si vous choisissez ensuite **Lancer**, une boîte de dialogue vous invite à sélectionner une paire de clés, de continuer sans paire de clés ou de créer une nouvelle paire de clés avant de choisir **Lancer des instances** et de créer l'instance EC2.

AmazonEC2 utilise le chiffrement à **clé publique** afin de chiffrer et de déchiffrer les informations de connexion. La technologie utilise une **clé publique** pour chiffrer une donnée, que le destinataire pourra ensuite déchiffrer à l'aide de la **clé privée**. Les clés publique et privée forment une paire de clés. Le chiffrement par clé publique vous permet d'accéder en toute sécurité à vos instances en utilisant une clé privée au lieu d'un mot de passe.

Pour vous connecter à une instance **Windows**, utilisez la clé privée pour obtenir le mot de passe administrateur, puis connectez-vous au bureau Windows de l'instance EC2 à l'aide du protocole RDP (Remote Desktop Protocol). Pour établir une connexion SSH entre une machine Windows et une instance Amazon EC2, vous pouvez utiliser l'outil PuTTY, qui requiert la même clé privée.

Avec les instances Linux, le contenu de la clé publique est placé sur l'instance au moment du démarrage. Une entrée est créée dans `~/.ssh/authorized_keys`. Pour vous connecter à votre instance Linux (par exemple, en utilisant SSH), vous devez fournir la clé privée lorsque vous établissez la connexion.

Vue de la console Amazon EC2 d'une instance EC2 en cours d'exécution



27

Après avoir choisi **Lancer les instances**, puis **Afficher les instances**, un écran qui ressemble à l'exemple s'affiche.

La plupart des paramètres que vous avez spécifiés lors du lancement se trouvent dans le panneau **Description**.

Les informations sur l'instance disponible incluent l'adresse IP et l'adresse DNS, le type d'instance, l'ID d'instance unique qui a été attribué à l'instance, l'ID d'AMI de l'AMI que vous avez utilisée pour lancer l'instance, l'ID de VPC, l'ID de sous-réseau, etc.

La plupart de ces détails fournissent des liens hypertexte que vous pouvez sélectionner pour en savoir plus sur les ressources qui sont pertinentes pour l'instance EC2 que vous avez lancée.

Vous pouvez également lancer des instances EC2 par programmation, soit à l'aide de l'Interface de la ligne de commande AWS (AWS CLI), soit à l'aide d'un des kits SDK AWS.

Autre option : lancement d'une instance EC2 avec l'interface de la ligne de commande AWS

- Les instances EC2 peuvent également être créées par programmation.



Interface de la ligne de commande AWS (AWS CLI)

- Cet exemple illustre la simplicité de la commande.
 - Cette commande suppose que la paire de clés et le groupe de sécurité existent déjà.
 - D'autres options peuvent être spécifiées. Pour plus d'informations, consultez le [document de référence sur les commandes AWS CLI](#).

Exemple de commande :

```
aws ec2 run-instances \
--image-id ami-1a2b3c4d \
--count 1 \
--instance-type c3.large \
--key-name MyKeyPair \
--security-groups MySecurityGroup \
--region us-east-1
```



© 2024, Amazon Web Services, Inc. ou ses sociétés affiliées. Tous droits réservés.

28

Dans l'exemple de l'interface de la ligne de commande AWS CLI, vous pouvez voir une commande unique qui spécifie les informations minimales nécessaires au lancement d'une instance. La commande comprend les données suivantes:

- aws : indique un appel de l'utilitaire de ligne de commande aws.
- ec2 : indique un appel de la commande de service ec2.
- run-instances: la sous-commande est appelée.

Le reste de la commande spécifie plusieurs paramètres, notamment:

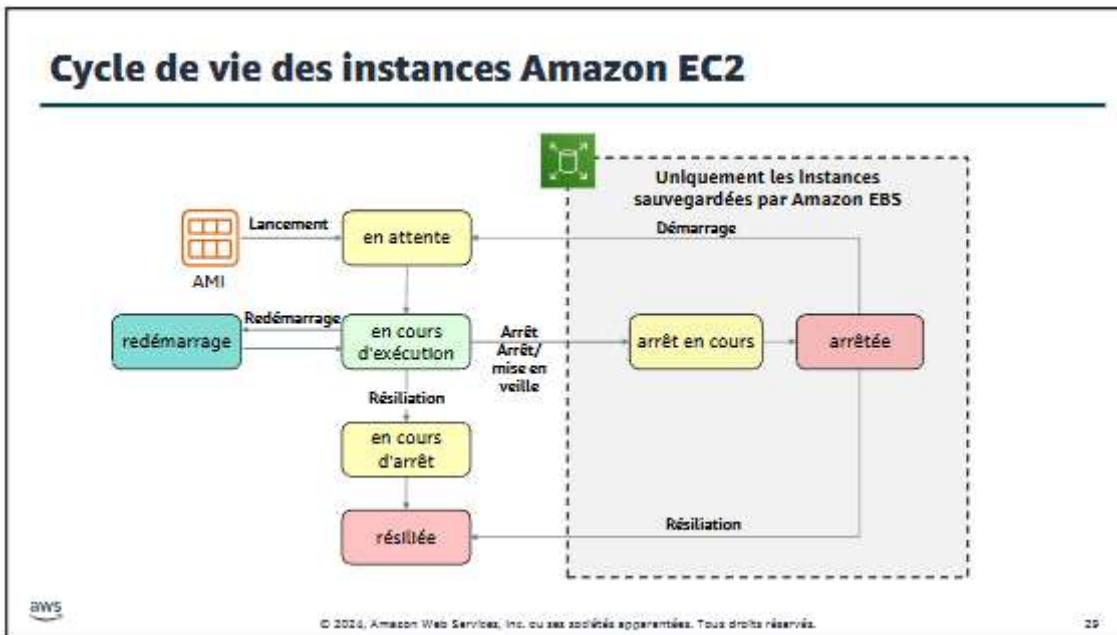
- image-id: ce paramètre est suivi d'un ID AMI. Toutes les AMI ont un ID unique.
- count: vous pouvez en spécifier plusieurs.
- instance-type: vous pouvez spécifier le type d'instance pour créer (par exemple) une instance c3.large
- key-name: dans cet exemple, supposons que MyKeyPair existe déjà.
- security-groups: dans cet exemple, supposons que MySecurityGroup existe déjà.
- region: les AMI existent dans une région AWS. Vous devez donc spécifier la région dans laquelle l'AWS CLI localisera l'AMI et lancera l'instanceEC2.

La commande doit créer avec succès une instance EC2 si :

- La syntaxe de la commande est correcte.
- Les ressources dont la commande a besoin existent déjà.
- Vous disposez des autorisations suffisantes pour l'exécuter

Si la commande est exécutée avec succès, l'API répond à la commande avec l'ID d'instance et d'autres données pertinentes que votre application utilisera dans les requêtes API suivantes.

Voici le cycle de vie d'une instance.



Les flèches indiquent les actions que vous pouvez effectuer et les cases indiquent l'état dans lequel l'instance entrera après cette action. Une instance peut présenter l'un des états suivants :

En attente : lorsqu'une instance est lancée pour la première fois à partir d'une AMI, ou lorsque vous démarrez une instance arrêtée, son état est en attente lorsque l'instance démarre et est déployée sur un ordinateur hôte. Le type d'instance spécifié au lancement détermine le matériel de l'ordinateur hôte utilisé pour votre instance.

En cours d'exécution : lorsque l'instance est entièrement démarrée et prête, elle passe de l'état en attente à l'état en cours d'exécution. Vous pouvez vous connecter via Internet à votre instance en cours d'exécution.

Redémarrage : AWS vous recommande de redémarrer une instance à l'aide de la console Amazon EC2, de l'AWSCLI ou des kits SDK AWS au lieu d'appeler un redémarrage depuis le système d'exploitation invité(OS). Une instance redémarrée reste sur le même hôte physique, conserve le même nom DNS public et la même adresse IP publique. Si elle possède des volumes de stockage d'instance, les données présentes sur ces volumes y sont conservées.

En cours d'arrêt : cet état est un état intermédiaire entre l'état En cours d'exécution et Résilié.

Résiliée : une instance résiliée demeure dans la console Amazon EC2 pendant un certain temps avant que la machine virtuelle ne soit supprimée. Vous ne pouvez ni vous connecter à une instance résiliée, ni la restaurer.

Arrêt en cours : les instances qui sont sauvegardées par Amazon EBS peuvent être arrêtées.

Leur état devient arrêt en cours avant de passer à l'état arrêté entièrement.

Arrêtée : une instance arrêtée n'encourt pas le même coût qu'une instance en cours d'exécution. Le démarrage d'une instance arrêtée la remet en état «En suspens», ce qui migre l'instance vers une nouvelle machine hôte.

Quand utiliser une adresse IP élastique ?

- Le redémarrage d'une instance *ne modifie aucune adresse IP ni aucun nom d'hôte DNS.*

- Lorsqu'une instance est **arrêtée puis redémarrée** :
 - L'adresse IPv4 *publique et le nom d'hôte DNS externe changent.*

 - L'adresse IPv4 *privée et le nom d'hôte DNS interne ne changent pas.*

- Si vous avez besoin d'une adresse IP publique persistante :
 - Associez une **adresse IP élastique** à l'instance.

- Caractéristiques des adresses IP élastiques :
 - Peuvent être associées aux instances de la région selon les besoins.

 - Restent allouées à votre compte jusqu'à ce que vous choisissiez de les libérer.



Adresse IP élastique

AWS © 2024, Amazon Web Services, Inc. ou ses sociétés affiliées. Tous droits réservés.

Une adresse IP publique est une adresse IP IPv4 qui est accessible depuis Internet. Chaque instance qui reçoit une adresse IP publique reçoit également un nom d'hôte DNS externe. Par exemple, si l'adresse IP publique attribuée à l'instance est 203.0.113.25, le nom d'hôte DNS externe pourrait être ec2-203-0-113-25.compute-1.amazonaws.com.

Si vous spécifiez qu'une adresse IP publique doit être attribuée à votre instance, elle est attribuée à partir du groupe AWS d'adresses IPv4 publiques.

L'adresse IP publique n'est pas associée à votre compte AWS.

Lorsqu'une adresse IP publique est dissociée de votre instance, elle est réintroduite dans le groupe d'adresses IPv4 publiques et vous ne pourrez pas spécifier que vous souhaitez la réutiliser.

AWS libérera l'adresse IP publique de votre instance lorsque cette dernière sera arrêtée ou résiliée. L'instance arrêtée recevra une nouvelle adresse IP publique lorsqu'elle sera redémarrée.

Si vous avez besoin d'une adresse IP publique persistante, pensez à associer une adresse IP élastique à l'instance. Pour associer une adresse IP élastique, vous devez d'abord allouer une nouvelle adresse IP élastique dans la région où l'instance existe. Une fois l'adresse IP élastique allouée, vous pouvez l'associer à une instance EC2.

Par défaut, tous les comptes AWS sont limités à cinq (5) adresses IP élastiques par région, car les adresses Internet publiques (IPv4) sont des ressources publiques relativement rares. Toutefois, cette limite est souple. Vous pouvez demander à ce qu'elle soit augmentée.

Amazon CloudWatch pour la surveillance

- Utilisez Amazon CloudWatch pour surveiller les instances EC2
 - Fournit des métriques en temps quasi réel
 - Vous donne accès à des graphiques dans l'onglet Surveillance de la console Amazon EC2
 - Conserve 15 mois de données historiques
- Surveillance basique
 - Activée par défaut, n'entraîne aucun coût supplémentaire
 - Données de métriques envoyées à CloudWatch toutes les 5 minutes
- Surveillance détaillée
 - Tarif mensuel fixe pour sept métriques présélectionnées
 - Données de métriques fournies toutes les minutes



The screenshot shows the CloudWatch Metrics dashboard for an EC2 instance named 'MyTestInstance'. It displays several line graphs for metrics like CPU Utilization, Network In, and Network Out. A callout box highlights the 'CloudWatch Metrics' section of the navigation bar at the top left.

© 2014, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

Vous pouvez surveiller vos instances à l'aide d'Amazon Cloud Watch, qui collecte et traite les données brutes d'Amazon EC2 en métriques lisibles en temps quasi réel. Ces statistiques sont enregistrées pendant une période de 15 mois. Vous pouvez ainsi accéder à des informations historiques et vous faire une idée plus précise des performances de votre application ou service web.

Par défaut, Amazon EC2 fournit une surveillance basique, qui envoie des données de métriques à Cloud Watch toutes les cinq minutes.

Pour envoyer les données de métriques correspondant à votre instance à Cloud Watch toutes les minutes, vous pouvez activer la surveillance détaillée sur l'instance.

Modèles de tarification Amazon EC2

Instances à la demande

- Facturation à l'heure
- Aucun engagement à long terme
- Donne droit à l'[offre gratuite AWS](#)

Hôtes dédiés

- Serveur physique doté d'une capacité d'instance EC2 entièrement dédiée à votre utilisation.

Instances dédiées

- Instances qui s'exécutent dans un VPC sur du matériel dédié à un seul client.

Instances réservées

- Paiement complet, paiement partiel ou aucun paiement à l'avance pour l'instance que vous réservez.
- Remise sur le tarif horaire de cette instance.
- Durée d'un an ou de trois ans.

Instances réservées planifiées

- Achetez une réservation de capacité qui est en permanence disponible selon un calendrier récurrent que vous définissez.
- Durée d'un an.

Instances Spot

- Les Instances sont exécutées tant qu'elles sont disponibles et que votre prix demandé est supérieur au prix de l'Instance Spot.
- Elles peuvent être interrompues par AWS à l'issue d'un délai de notification de 2 minutes.
- Les options d'interruption incluent la résiliation, l'arrêt ou la mise en veille prolongée.
- Les prix peuvent être nettement inférieurs à ceux des Instances à la demande.
- Choix judicieux si vous êtes flexible quant au moment où vos applications peuvent être exécutées.

La facturation à la seconde est disponible pour les instances à la demande, les instances réservées et les instances Spot sous Amazon Linux ou Ubuntu.

© 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés.

46

Amazon offre au choix différents modèles de tarification pour exécuter des instances EC2.

La facturation à la seconde n'est disponible que pour les instances à la demande, les instances réservées et les instances Spot sous Amazon Linux ou Ubuntu.

Les instances à la demande donnent droit à l'[offre gratuite d'AWS](#) (<https://aws.amazon.com/free/>). Elles présentent les frais initiaux les plus bas et la plus grande flexibilité. Elles n'impliquent aucun engagement initial ni contrat à long terme. Elles constituent un choix judicieux pour les applications avec des charges de travail à court terme, irrégulières ou imprévisibles.

Les hôtes dédiés sont des serveurs physiques dont des capacités d'instance vous sont dédiées. Ils vous permettent d'utiliser vos licences logicielles existantes par socket, par cœur ou par machine virtuelle, telles que Microsoft Windows ou Microsoft SQL Server.

Les instances dédiées sont des instances qui s'exécutent dans un cloud privé virtuel (VPC) sur du matériel dédié à un seul client. Elles sont physiquement isolées au niveau du matériel hôte des instances qui appartiennent à d'autres comptes AWS.

Les instances réservées vous permettent de réserver une capacité de calcul sur une durée d'un an ou de trois ans, avec des coûts horaires de fonctionnement réduits. Le prix d'utilisation, avec la remise, ne changera pas tant que vous possédez l'instance réservée. Si vous envisagez une utilisation constante et intensive, elles peuvent vous aider à réaliser d'importantes économies par rapport aux instances à la demande.

Les instances réservées planifiées vous permettent d'acheter des réservations de capacité sur une base quotidienne, hebdomadaire ou mensuelle, sur une durée spécifique, pendant un an.

Les instances Spot vous permettent d'enchérir sur des instances EC2 inutilisées, ce qui peut réduire vos coûts. Le prix horaire d'une instance Spot fluctue en fonction de l'offre et de la demande. Votre instance Spot s'exécute dès que votre offre dépasse le prix du marché actuel.

Chaque modèle de tarification Amazon EC2 offre un ensemble d'avantages différent.

Modèles de tarification Amazon EC2 : avantages



Instances à la demande	Instances Spot	Instances réservées	Hôtes dédiés
<ul style="list-style-type: none">• Faible coût et flexibilité	<ul style="list-style-type: none">• Charge de travail dynamique à grande échelle	<ul style="list-style-type: none">• La prévisibilité garantit que la capacité de calcul est disponible en cas de besoin	<ul style="list-style-type: none">• Réaliser des économies sur les coûts de licence• Satisfaire vos exigences en matière de conformité et de réglementation

aws © 2024, Amazon Web Services, Inc. ou ses sociétés apparentées. Tous droits réservés. 47

Les instances à la demande offrent la plus grande flexibilité, sans contrat à long terme et à des tarifs réduits.

Les instances Spot sont à grande échelle et à un prix fortement réduit.

Les instances réservées sont un choix judicieux si vos besoins en calcul sont prévisibles ou stables (par exemple, une instance dont vous savez qu'elle doit fonctionner en quasi-permanence pendant des mois ou des années).

Les hôtes dédiés constituent un choix judicieux lorsque vous avez des restrictions de licence pour le logiciel à exécuter sur Amazon EC2, ou lorsque les exigences en matière de réglementation et de conformité vous empêchent d'utiliser les autres options de déploiement.

Voyons quelques cas d'utilisation des différentes options de tarification.

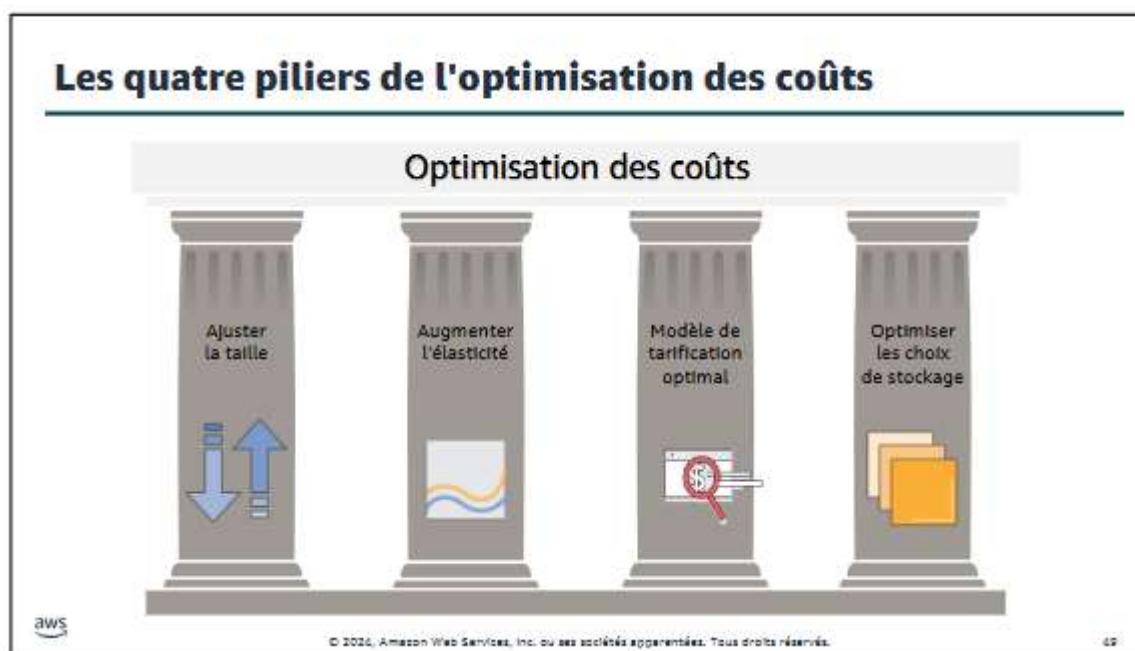
La tarification des **instances à la demande** convient parfaitement aux charges de travail irrégulières, ou si vous n'avez besoin de tester ou d'exécuter une application que sur une courte période (par exemple, pendant le développement ou les tests d'une application). Parfois, les charges de travail sont imprévisibles et les instances à la demande sont alors une option convenable.

Les instances Spot sont un choix judicieux si vos applications tolèrent les interruptions après notification dans un délai de 2 minutes. Par défaut, les instances sont résiliées, mais vous pouvez les configurer pour qu'elles s'arrêtent ou qu'elles soient mises en veille prolongée. Par défaut, les instances sont résiliées, mais vous pouvez les configurer pour qu'elles s'arrêtent ou qu'elles soient mises en veille prolongée. Les cas d'utilisation courants incluent les applications tolérantes aux pannes, telles que les serveurs web, les backends d'API et le traitement du Big Data. Les charges de travail qui enregistrent constamment des données dans un support de stockage persistant (comme Amazon S3) sont également de bons candidats.

Les instances réservées sont un choix judicieux lorsque vous disposez de charges de travail à long terme avec des modèles d'utilisation prévisibles, par exemple des serveurs que vous souhaitez exécuter de manière cohérente sur plusieurs mois.

Les hôtes dédiés sont un choix judicieux si vous utilisez des licences logicielles par socket, par cœur ou par machine virtuelle, ou lorsque vous devez exiger des exigences en matière de réglementation et de conformité spécifiques de l'entreprise.

Les quatre piliers de l'optimisation des coûts



Pour optimiser les coûts, vous devez prendre en compte quatre facteurs cohérents et puissants:

Ajuster la taille: choisissez le bon équilibre de types d'instance. Déterminez quand la taille des serveurs peut être réduite ou quand ils peuvent être arrêtés, tout en répondant à vos exigences de performances.

Augmenter l'élasticité: concevez vos déploiements de manière à réduire la quantité de capacité de serveur inactive en mettant en œuvre des déploiements élastiques, tels que des déploiements qui exploitent la mise à l'échelle automatique pour gérer les pics de charge.

Modèle de tarification optimal: identifiez les options de tarification disponibles. Analysez vos modèles d'utilisation afin d'exécuter les instances EC2 avec la bonne combinaison d'options tarifaires.

Optimiser les choix de stockage: analysez les besoins en stockage de vos déploiements. Réduisez la charge de stockage inutilisée lorsque cela est possible et choisissez des options de stockage plus économiques si elles répondent toujours à vos exigences en matière de performances de stockage.

Mesurer, surveiller et améliorer

- L'optimisation des coûts est un processus continu.



- Recommandations :

- Définissez et appliquez le **balisage de la répartition des coûts**.
- Définissez des métriques, fixez des objectifs et révisez-les régulièrement.
- Encouragez les équipes à **concevoir des architectures en fonction des coûts**.
- Attribuez la responsabilité de l'optimisation à un individu ou à une équipe.



Voici quelques points clés à retenir :

- Avec Amazon EC2, vous pouvez utiliser des machines virtuelles Windows et Linux dans le cloud.
- Vous lancez des instances EC2 à partir d'un modèle d'AMI dans un VPC de votre compte.
- Vous avez le choix entre de nombreux types d'instances. Chaque type d'instance offre différentes combinaisons de capacités de CPU, de RAM, de stockage et de mise en réseau.
- Vous pouvez configurer des groupes de sécurité pour contrôler l'accès aux instances (spécifiez les ports et la source autorisés).
- Les données utilisateur vous permettent de spécifier un script à exécuter la première fois qu'une instance se lance.
- Seules les instances qui sont soutenues par Amazon EBS peuvent être arrêtées.
- Vous pouvez utiliser Amazon CloudWatch pour capturer et examiner les métriques sur les instances EC2.