



Appliquer le framework AWS Well-Architected pour Amazon Timestream pour InfluxDB

AWS Conseils prescriptifs



AWS Conseils prescriptifs: Appliquer le framework AWS Well-Architected pour Amazon Timestream pour InfluxDB

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

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

Table of Contents

Introduction	1
Public visé	2
Objectifs	2
Excellence opérationnelle	3
Automatisez le déploiement en utilisant une approche IaC	3
Effectuez des modifications fréquentes, mineures et réversibles	4
Anticipez les défaillances	4
Tirez les leçons de toutes les défaillances opérationnelles	5
Utilisez les fonctionnalités de journalisation pour surveiller les activités non autorisées ou anormales	5
Sécurité	7
Mettre en œuvre la sécurité des données	7
Sécurisez vos réseaux	8
Mettre en œuvre l'authentification et l'autorisation	8
Fiabilité	10
Architecture de charge de travail, y compris les quotas de service et les modèles de déploiement	10
Modèles de déploiement	10
Gérez et dimensionnez Timestream pour InfluxDB	11
Efficacité des performances	13
Modélisation des données d'afflux et optimisation des requêtes	13
Optimisez les écritures	15
Optimisation des coûts	17
Comprendre les exigences et les coûts de votre cas d'utilisation	17
Sélection des ressources en tenant compte des coûts	17
Évolutivité pour répondre aux besoins de l'entreprise sans trop dépenser	18
Stockage et transfert de données adaptés	19
Durabilité	20
Région AWS sélection	20
Baser la consommation des ressources sur les modèles de comportement des utilisateurs	21
Optimisez le développement logiciel et les modèles d'architecture	21
Ressources	23
Références	23
Billets de blogs	23

Historique du document	24
Glossaire	25
#	25
A	26
B	29
C	31
D	34
E	39
F	41
G	43
H	44
I	46
L	48
M	50
O	54
P	57
Q	60
R	60
S	63
T	67
U	69
V	69
W	70
Z	71
.....	lxxii

Appliquer le framework AWS Well-Architected pour Amazon Timestream pour InfluxDB

Balwanth Bobilli, Amazon Web Services

Janvier 2025 ([historique du document](#))

Vous pouvez créer time-series-based des solutions sur Amazon Web Services (AWS) à l'aide d'[Amazon Timestream](#). Amazon Timestream propose des moteurs de base de données de séries chronologiques entièrement gérés et spécialement conçus pour les charges de travail, qu'il s'agisse de requêtes à faible latence ou d'ingestion de données à grande échelle. Avec [Amazon Timestream for InfluxDB](#), vous pouvez exécuter des bases de données InfluxDB open source pour des applications chronologiques, telles que les alertes en temps réel et AWS la fiabilité de l'infrastructure de surveillance, avec des temps de réponse en millisecondes. Timestream pour InfluxDB offre une disponibilité allant jusqu'à 99,9 %.

Ce guide fournit des conseils prescriptifs pour appliquer les principes du [AWS Well-Architected Framework](#) lorsque vous planifiez votre déploiement Timestream pour InfluxDB. Le AWS Well-Architected Framework vous aide à créer des infrastructures sécurisées, performantes, résilientes et efficaces pour une variété d'applications et de charges de travail. Il fournit également une approche cohérente vous permettant d'évaluer les architectures et de mettre en œuvre des conceptions évolutives.

Le AWS Well-Architected Framework repose sur les six piliers suivants :

- Excellence opérationnelle
- Sécurité
- Fiabilité
- Efficacité des performances
- Optimisation des coûts
- Durabilité

Ce guide fournit des informations sur les piliers de conception du AWS Well-Architected Framework. Envisagez d'appliquer ces meilleures pratiques lorsque vous déployez Amazon Timestream pour InfluxDB sur AWS.

 Note

Certains Services AWS ne sont pas disponibles dans toutes les Régions AWS. Pour connaître la disponibilité des régions, consultez la page [Points de terminaison et quotas du service](#) dans la AWS documentation et choisissez le lien du service.

Public visé

Ce guide est destiné aux ingénieurs de données, aux architectes de solutions et aux analystes de données qui conçoivent et mettent en œuvre des solutions pour les séries chronologiques de données sur AWS.

Objectifs

Ce guide peut vous aider, vous et votre organisation, à effectuer les tâches suivantes :

- Choisissez parmi les options de déploiement prises en charge, effectuez des écritures optimisées et implémentez un accès précis.
- Suivez les modèles de conception AWS Well-Architected qui contribuent à améliorer la résilience et la sécurité.
- Concevez vos requêtes pour des performances optimales.
- Découvrez comment être efficace sur le plan opérationnel lors de la gestion de votre instance Timestream for InfluxDB en production.

Pilier d'excellence opérationnelle

Le pilier de [l'excellence opérationnelle](#) du AWS Well-Architected Framework se concentre sur le fonctionnement et le suivi des systèmes, ainsi que sur l'amélioration continue des processus et des procédures afin de créer de la valeur commerciale. Le pilier de l'excellence opérationnelle inclut la capacité à soutenir le développement et à gérer efficacement les charges de travail, ainsi qu'à mieux comprendre leur fonctionnement.

Vous pouvez réduire la complexité opérationnelle grâce à des charges de travail autoréparables, qui détectent et corrigent la plupart des problèmes sans intervention humaine. Pour atteindre cet objectif, suivez les meilleures pratiques décrites dans cette section. Utilisez CloudWatch les métriques [Amazon](#) pour Amazon Timestream pour InfluxDB, le point de terminaison des métriques natives d'InfluxDB, et les mécanismes permettant de réagir lorsque votre APIs charge de travail s'écarte du comportement attendu.

Cette discussion sur le pilier de l'excellence opérationnelle met l'accent sur les domaines clés suivants :

- Infrastructure en tant que code (IaC)
- Gestion des modifications
- Stratégies de résilience
- Gestion des incidents
- Enregistrement et surveillance à des fins d'audit

Automatisez le déploiement en utilisant une approche IaC

Les meilleures pratiques pour automatiser le déploiement sur Timestream pour InfluxDB à l'aide d'IaC sont les suivantes :

- Appliquez IaC pour déployer Timestream pour InfluxDB chaque fois que possible. Pour une configuration cohérente de l'environnement [AWS Cloud Development Kit \(AWS CDK\)](#), utilisez un [AWS CloudFormation](#) modèle ou [HashiCorp Terraform](#) pour créer toutes les ressources requises pour votre instance.
- Automatisez les procédures opérationnelles Timestream pour InfluxDB, telles que le redimensionnement des instances.

- Utilisez des balises pour ajouter des métadonnées à votre flux temporel pour les ressources InfluxDB et suivez l'utilisation en fonction des balises. Pour plus d'informations, consultez la section [Marquage d'Amazon Timestream](#) pour InfluxDB.

Effectuez des modifications fréquentes, mineures et réversibles

Les recommandations suivantes mettent l'accent sur de petites modifications réversibles afin de minimiser la complexité et de réduire le risque d'interruption de la charge de travail :

- Stockez les modèles et les scripts IaC dans un service de contrôle de source, tel GitHub ou GitLab. Ne stockez pas les AWS informations d'identification dans le contrôle de source.
- Exigez que les déploiements IaC utilisent un service d'intégration et de livraison continues (CI/CD), tel que [AWS CodeDeploy](#) ou [AWS CodeBuild](#). Ces services compilent, testent et déploient du code dans un environnement hors production qui contient une instance InfluxDB éphémère avant d'affecter votre instance InfluxDB de production.
- Testez les requêtes d'infrastructure et d'application dans un environnement inférieur avant de les déployer en production. Cela minimise le risque d'interruption et permet de garantir qu'ils fonctionnent bien avec votre charge de travail et votre évolutivité.

Anticipez les défaillances

Une infrastructure autoréparante illustre l'excellence opérationnelle en anticipant les défaillances et en tentant de résoudre les problèmes sans intervention. Les recommandations suivantes vous aident à atteindre cette maturité avec Timestream pour InfluxDB :

- Utilisez des métriques pour surveiller l'utilisation de la mémoire, du processeur et du stockage. Vous pouvez configurer CloudWatch pour être informé quand les habitudes d'utilisation changent ou quand vous arrivez au bout de votre capacité de déploiement. De cette façon, vous pouvez maintenir les performances et la disponibilité du système.
- Augmentez la taille de votre instance de base de données lorsque vous approchez de la limite de ressources. Vous devrez disposer de capacités de mémoire et de stockage supplémentaires pour vous adapter aux hausses imprévues des besoins de vos applications.
- Si la charge de travail de votre base de données exige plus d'I/O que vous avez provisionné, la récupération suite à un basculement ou un échec de la base de données est lente. Pour

augmenter la capacité d'E/S d'une instance de base de données, migrez vers une autre instance de base de données dotée d'une capacité d'E/S supérieure.

- Si votre application cliente met en cache les données DNS de vos instances de base de données, définissez une valeur time-to-live (TTL) inférieure à 30 secondes. L'adresse IP sous-jacente d'une instance de base de données peut changer après un basculement. La mise en cache des données DNS pendant une période prolongée peut entraîner des échecs de connexion. Il se peut que votre application essaie de se connecter à une adresse IP qui n'est plus en service.
- Si votre application doit survivre à une Région AWS panne complète, envisagez de configurer la réplication ou d'écrire dans une autre région dans le cadre de vos plans de reprise après sinistre (DR). Comprenez les limites lors de la configuration de la réplication. Pour plus d'informations sur la réplication, consultez la documentation [InfluxDB](#).

Tirez les leçons de toutes les défaillances opérationnelles

Une infrastructure d'autoréparation est un effort à long terme que vous développez par itérations lorsque de rares problèmes surviennent ou que les réponses ne sont pas aussi efficaces que vous le souhaitez. Pour vous concentrer sur la mise en place d'une infrastructure autoréparante, adoptez les pratiques suivantes :

- Favorisez l'amélioration en tirant les leçons de tous les échecs.
- Partagez ce que vous avez appris au sein des équipes et de l'organisation. Si plusieurs équipes d'une organisation utilisent Timestream pour InfluxDB, créez un salon de discussion ou un groupe d'utilisateurs commun pour partager les leçons apprises et les meilleures pratiques.

Utilisez les fonctionnalités de journalisation pour surveiller les activités non autorisées ou anormales

Pour observer des modèles de performance et d'activité anormaux, prenez en compte les pratiques suivantes :

- Activez [la livraison des journaux](#) pour stocker les journaux InfluxDB dans [Amazon Simple Storage Service \(Amazon S3\)](#). Les journaux InfluxDB enregistrent des informations qui peuvent aider à vérifier les points suivants :
 - [Événements de l'API du plan de données](#)
 - Temps de réponse

- Détails du compactage
- Toute erreur critique ou tout avertissement rencontré par le système

Consultez les journaux pour détecter tout accès non autorisé ou toute anomalie. Dans l'ensemble, la journalisation fournit des informations de diagnostic pour le dépannage.

- Timestream for InfluxDB prend en charge la journalisation des actions du plan de contrôle en utilisant AWS CloudTrail. Pour plus d'informations, consultez [Logging Timestream pour les appels d'API InfluxDB](#) avec AWS CloudTrail
- Vous pouvez surveiller **CPUUtilizationMemoryUtilization**, et les **DiskUtilization** métriques depuis Timestream/InfluxDB > < Namespace > in. CloudWatch

Pour plus d'informations, consultez la documentation [Timestream for InfluxDB](#).

Pilier de sécurité

Le [pilier de sécurité](#) vous aide à comprendre comment appliquer le modèle de responsabilité partagée lors de l'utilisation de Timestream pour InfluxDB.

Le pilier de sécurité comprend les principaux domaines d'intérêt suivants :

- Sécurité des données
- Sécurité du réseau
- Authentification et autorisation

Les sections suivantes vous montrent comment configurer Timestream pour InfluxDB afin d'atteindre vos objectifs de sécurité et de conformité. Vous apprendrez également à en utiliser d'autres Services AWS qui vous aident à surveiller et à sécuriser votre Timestream pour les ressources InfluxDB.

Mettre en œuvre la sécurité des données

Les fuites de données et les violations mettent vos clients en danger et peuvent avoir un impact négatif important sur votre entreprise. Le [modèle de responsabilité AWS partagée s'applique](#) à la protection des données dans Timestream pour InfluxDB. Les pratiques suivantes permettent de protéger les données de vos clients contre toute exposition accidentelle ou malveillante :

- N'incluez pas d'informations confidentielles dans les noms d'instances, les balises, les groupes de paramètres, les rôles AWS Identity and Access Management (IAM) et les autres métadonnées. Ces données peuvent apparaître dans les journaux de facturation ou de diagnostic.
- Utilisez le chiffrement pour renforcer la protection des données de vos applications déployées dans le cloud. Les instances chiffrées fournissent une couche supplémentaire de protection des données en protégeant vos données contre tout accès non autorisé au stockage sous-jacent. Le chiffrement Timestream pour InfluxDB est activé par défaut.
- Utilisez le paramétrage dans la APIs mesure du possible.

Pour plus d'informations sur la protection et le chiffrement des données, consultez la documentation [Timestream for InfluxDB](#).

Sécurisez vos réseaux

Vous pouvez créer une instance Timestream pour InfluxDB uniquement dans un cloud privé virtuel (VPC) sur AWS. Pour renforcer la sécurité de votre instance, procédez comme suit :

- Limitez l'accès public au point de terminaison de l'instance en effectuant l'une des opérations suivantes :
 - Choisissez Non accessible au public sur la page Créer une base de données InfluxDB (sélectionnée par défaut).
 - Définissez la [PubliclyAccessible](#) propriété sur false in AWS CloudFormation (AWS::Timestream::InfluxDBInstance).

Lorsque vous choisissez Non accessible au public ou que vous définissez sur false, [PubliclyAccessible](#) les points de terminaison de l'instance ne sont accessibles qu'au sein du VPC. Les points de terminaison sont généralement accessibles depuis une instance [Amazon Elastic Compute Cloud \(Amazon EC2\)](#) exécutée dans le même VPC que l'instance Timestream for InfluxDB.

- Utilisez des groupes de sécurité pour sécuriser davantage l'accès de votre réseau à Timestream for InfluxDB au sein du VPC.
- Utilisez le protocole SSL/TLS pour communiquer avec les ressources. AWS Timestream pour InfluxDB nécessite le protocole TLS 1.2, et nous recommandons le protocole TLS 1.3.
- Connectez-vous en toute sécurité à votre instance en suivant les instructions de la section [Création et connexion à une instance Timestream pour InfluxDB](#).
- [Établissez une connexion privée entre votre VPC et Amazon Timestream pour les points de terminaison de l'API du plan de contrôle InfluxDB en créant un point de terminaison VPC d'interface.](#)

Pour plus d'informations sur la sécurité, consultez la documentation [Timestream for InfluxDB](#).

Mettre en œuvre l'authentification et l'autorisation

Utilisez les informations d'identification IAM pour contrôler les actions de gestion sur Timestream pour les instances InfluxDB. Lorsque vous vous connectez à Timestream pour InfluxDB à l'aide des informations d'identification IAM, votre rôle IAM doit disposer de politiques IAM qui accordent les autorisations requises pour effectuer des opérations de gestion de Timestream pour InfluxDB.

Assurez-vous de suivre le principe du moindre privilège, en n'accordant que les autorisations requises pour effectuer une tâche. Pour plus d'informations, consultez [Identity and Access Management pour Amazon Timestream pour InfluxDB](#).

Vous pouvez utiliser [Amazon Timestream pour les actions InfluxDB dans la politique IAM](#) afin de contrôler qui peut gérer votre instance Timestream pour InfluxDB.

[Pour fournir un contrôle d'accès précis à vos données stockées dans votre instance Amazon Timestream pour InfluxDB, offrez aux utilisateurs des jetons d'API InfluxDB.](#)

Pour interagir avec d'autres utilisateurs Services AWS, Amazon Timestream pour InfluxDB utilise des rôles liés au service IAM. Un rôle lié à un service est un type unique de rôle IAM directement lié à Timestream pour InfluxDB. Les rôles liés à un service sont prédéfinis par Timestream pour InfluxDB, et ils incluent toutes les autorisations dont le service a besoin pour appeler d'autres personnes en votre nom. Services AWS Pour plus d'informations, consultez [Utilisation de rôles liés à un service pour Amazon Timestream pour InfluxDB](#).

Pilier de fiabilité

Le [pilier de fiabilité](#) englobe la capacité d'une charge de travail à exécuter correctement et de manière cohérente la fonction prévue lorsqu'elle est censée le faire. Cela inclut la capacité d'exploiter et de tester la charge de travail tout au long de son cycle de vie.

La configuration d'une charge de travail fiable commence par des décisions de conception initiales pour le logiciel et l'infrastructure. Vos choix d'architecture ont un impact sur le comportement des charges de travail sur les différents piliers Well-Architected. Pour atteindre la fiabilité, vous devez suivre des modèles spécifiques.

Le pilier de fiabilité se concentre sur les domaines clés suivants :

- Architecture de charge de travail, y compris les quotas de service et les modèles de déploiement
- Gestion et dimensionnement des instances InfluxDB

Architecture de charge de travail, y compris les quotas de service et les modèles de déploiement

Chacun Compte AWS dispose de quotas pour les ressources qu'il propose Région AWS. Par exemple, chaque région dispose d'un [quota de diffusion temporelle pour les instances InfluxDB](#), quelle que soit la taille de l'instance. Une fois que vous avez atteint le nombre maximum d'instances dans une région, les appels supplémentaires pour créer des instances échouent, à une exception près. Un volume de stockage d'instance Timestream pour InfluxDB peut atteindre une taille maximale de 16 tébioctets (T) dans tous les cas pris en charge. TiBs Régions AWS

Modèles de déploiement

Pour [une haute disponibilité](#) et une prise en charge du basculement pour Timestream pour les instances InfluxDB, vous pouvez utiliser des déploiements multi-AZ avec une seule instance de base de données de secours. Ce type de déploiement est appelé déploiement d'instance de base de données multi-AZ. Amazon Timestream pour InfluxDB utilise la technologie de basculement Amazon. Dans le cadre d'un déploiement d'instance de base de données multi-AZ, Amazon Timestream approvisionne et gère automatiquement une réplique de secours synchrone dans une autre zone de disponibilité. Pour assurer la redondance des données, l'instance de base de données principale est répliquée de manière synchrone entre les zones de disponibilité vers la réplique de secours.

L'exécution d'une instance de base de données à haute disponibilité peut garantir la disponibilité en cas de défaillance de l'instance de base de données ou d'interruption de la zone de disponibilité. Si une panne imprévue de votre instance de base de données résulte d'un défaut d'infrastructure, Amazon Timestream for InfluxDB passe automatiquement à la réplique de secours. La durée du basculement dépend de l'activité de la base de données et d'autres conditions au moment où l'instance de base de données primaire est devenue indisponible.

Les durées de basculement oscillent généralement entre 60 et 120 secondes. Toutefois, des transactions importantes comportant des données à cardinalité élevée ou un long processus de restauration nécessitant un préchauffage peuvent augmenter le temps de basculement. Une fois le basculement terminé, un délai supplémentaire peut être nécessaire avant que la console Timestream reflète la nouvelle zone de disponibilité.

Si votre application doit rester disponible pendant une Région AWS panne complète, envisagez de configurer la réplication ou d'écrire dans une autre région dans le cadre de vos plans de reprise après sinistre (DR). Toutefois, avant de configurer la réplication, assurez-vous d'en connaître les limites. Pour plus d'informations, consultez la documentation [InfluxDB](#).

Amazon Timestream for InfluxDB effectue régulièrement des sauvegardes internes et les conserve pendant 24 heures pour garantir la disponibilité et la durabilité. Les instantanés sont pris lors des suppressions et conservés pendant 30 jours pour permettre les restaurations. Pour y accéder ou les utiliser, créez un dossier sur [AWS Support](#).

Gérez et dimensionnez Timestream pour InfluxDB

Timestream for InfluxDB prend en charge les classes d'instance idéales pour exécuter des charges de travail gourmandes en mémoire dans des bases de données InfluxDB open source. Les différentes [classes d'instances db.influx ont des](#) limites en termes de vCPUs, de mémoire, de stockage et de bande passante réseau. Pour choisir la classe d'instance qui répond aux exigences de latence d'écriture et de requête de votre application, observez les `DiskUtilization` indicateurs Amazon CloudWatch `CPUUtilization` et `MemoryUtilization`, lors des tests. Vous pouvez [redimensionner vos instances à la hausse ou à la baisse en](#) fonction de vos exigences en matière de charge de travail. Timestream for InfluxDB fournit plusieurs niveaux de stockage préconfigurés avec des IOPS et un débit optimaux requis pour différents types de charges de travail. Choisissez ce qui convient le mieux à votre charge de travail en fonction de vos besoins.

Si vos besoins de dimensionnement changent à des moments prévisibles, vous pouvez utiliser une [AWS Lambda fonction](#) ou un planificateur personnalisé et exécuter une API ou un SDK pour augmenter ou diminuer le dimensionnement avec un certain temps tampon.

Vous gérez votre configuration InfluxDB dans Timestream pour InfluxDB en utilisant des paramètres dans un groupe de paramètres. Les groupes de paramètres agissent comme un conteneur pour les options de configuration InfluxDB qui sont appliquées à une ou plusieurs instances de base de données. Lorsque vous modifiez des paramètres dans des groupes de paramètres, comprenez la différence entre les paramètres statiques et dynamiques, et apprenez comment et quand ils sont appliqués. Pour voir la configuration actuellement appliquée, utilisez l'action [GetDbParameterGroup](#)API.

Pilier d'efficacité des performances

Le [pilier de l'efficacité des performances](#) du AWS Well-Architected Framework se concentre sur la manière d'optimiser les performances lors de l'ingestion ou de l'interrogation de données. L'optimisation des performances est un processus progressif et continu comprenant les éléments suivants :

- Confirmation des exigences commerciales
- Mesurer les performances de la charge de travail
- Identification des composants sous-performants
- Ajustement des composants pour répondre aux besoins de votre entreprise

Le pilier relatif à l'efficacité des performances fournit des directives qui peuvent vous aider à choisir un modèle de données performant. Le pilier de l'efficacité des performances inclut les meilleures pratiques d'optimisation des requêtes et des écritures.

Le pilier de l'efficacité en matière de performance se concentre sur les domaines clés suivants :

- Modélisation des données d'afflux et optimisation des requêtes
- Optimisation de l'écriture

Modélisation des données d'afflux et optimisation des requêtes

La conception d'un schéma efficace est cruciale pour optimiser les performances et les capacités de requête des données de séries chronologiques dans InfluxDB. Commencez par choisir les bons tags et les bons champs. InfluxDB indexe les balises, de sorte que le moteur de requête n'a pas besoin de scanner chaque enregistrement d'une mesure pour localiser la valeur d'une balise. Cela signifie que l'interrogation des balises est plus efficace que l'interrogation des champs. Pour compacter et stocker les données, le moteur de stockage regroupe les valeurs de champ par clé de série, puis classe ces valeurs de champ par heure. Une clé de série est définie par une mesure, une clé et une valeur de balise, et une clé de champ. Pour plus d'informations sur la conception des données, consultez la documentation [InfluxDB](#).

Le moteur de stockage utilise un format de données TSM (Time-Structured Merge Tree). Pour plus d'informations sur le format de données TSM, consultez la documentation [InfluxDB](#).

Imaginez que vous collectez des données (`timestamp`, `host_id`, `region`, `cpu`, `memory`, `network_in_bytes`, `network_out_bytes`, `disk_io`) dans le cadre d'un cas d' DevOps utilisation. Les balises, y compris l'horodatage de l'enregistrement, fournissent un contexte qui permet d'identifier qui, quoi, quand et où se trouve un enregistrement. Les balises sont utilisées pour organiser et classer les données, ainsi que pour filtrer les données dans le cadre d'une requête.

Les `region` balises `host_id` et sont idéales pour organiser et classer les cas DevOps d'utilisation. Ces colonnes permettent de filtrer les données pour un hôte particulier ou d'exécuter une analyse en fonction de la colonne de région.

Les mesures fournissent la base pour effectuer des calculs mathématiques (tels que le calcul des totaux, des moyennes et des différences de taux de variation) et une analyse quantitative de vos données. Par conséquent, `cpu`, `memory`, `network_in_bytes`, `network_out_bytes`, et `disk_io` capturez les indicateurs importants liés à ceux DevOps qui évoluent au fil du temps. Vous pouvez utiliser ces métriques pour effectuer diverses analyses, telles que le calcul du processeur et de la mémoire sur différents hôtes. Vous pouvez utiliser ces valeurs métriques pour prendre des décisions basées sur les données afin d'éviter les interruptions de production et de planifier l'infrastructure.

La cardinalité est la combinaison de valeurs de balises uniques. Essayez de maintenir la cardinalité aussi faible que possible. Si votre application nécessite un identifiant unique pour chaque point de données, utilisez des valeurs de champ plutôt que des valeurs de balise. Cela se traduira par une latence des requêtes nettement meilleure. Une bonne conception du schéma peut empêcher une cardinalité élevée des séries, ce qui améliore les performances des requêtes. Si vous remarquez un ralentissement de la lecture et de l'écriture des données ou si vous souhaitez savoir comment la cardinalité affecte les performances, consultez la documentation [Timestream for](#) InfluxDB.

Si votre application émet des objets JSON, convertissez-les en colonnes individuelles (balises ou champs) et chargez les colonnes dans InfluxDB. InfluxDB est conçu pour les données de séries chronologiques. L'organisation de vos données avec des colonnes individuelles est donc une bonne pratique pour tirer pleinement parti des fonctionnalités du service.

Une seule instance OSS InfluxDB v2.7 prend en charge environ 20 compartiments InfluxDB écrits ou interrogés activement dans toutes les organisations. Plus de 20 compartiments peuvent avoir une incidence négative sur les performances. Certaines options de configuration d'InfluxDB sont limitées, et vous pouvez configurer certaines options en fonction de votre cas d'utilisation. Validez la configuration en fonction de la charge de travail de l'application pendant la phase de test. Les rétentions de données étant configurées au niveau du compartiment, les données présentant des

exigences de conservation différentes doivent être stockées dans différents compartiments. Pour plus d'informations sur les options de configuration, consultez la documentation [Timestream for InfluxDB](#).

Stockez les données dans des valeurs de balise ou des valeurs de champ, et non dans des clés de balise, des clés de champ ou des mesures. Si vous concevez votre schéma pour stocker les données sous forme de balises et de valeurs de champs, vos requêtes seront plus faciles à écrire et plus efficaces. Pour en savoir plus sur les meilleures pratiques en matière de modélisation des données, consultez la section [Conception axée sur les performances](#).

Utilisez les [tâches InfluxDB](#) pour pré-agrèger les données, charger les données dans différentes mesures ou compartiments, et générer des données pour les tableaux de bord et les visualisations à partir de celles-ci.

InfluxDB OSS expose un `/metrics` [point de terminaison](#) qui renvoie des métriques de performances, de ressources et d'utilisation formatées au format d'exposition en texte brut Prometheus. Utilisez les modèles InfluxDB pour configurer la [surveillance et les alertes](#) afin de détecter de manière proactive les problèmes, tels que la latence élevée des requêtes, la dégradation du débit d'écriture ou les pics d'utilisation des ressources.

Timestream for InfluxDB fournit un stockage inclus dans Influx IO. La sélection de la taille d'IOPS appropriée peut accélérer considérablement l'exécution des requêtes. Cela est particulièrement utile pour les requêtes qui doivent analyser de grandes quantités de données ou traiter un large éventail de demandes. Dans certains cas, il peut être nécessaire de combiner le dimensionnement de l'instance et l'amélioration des IOPS pour obtenir les améliorations de performances souhaitées.

Nous recommandons de faire correspondre les environnements de développement et de production (classe d'instance, type de stockage, configurations). Testez les modifications dans l'environnement inférieur pour chaque version avant de passer à la production. Sur les volumes de stockage inclus dans Influx IO, Timestream for InfluxDB fournit trois niveaux de stockage préconfigurés avec des IOPS optimaux (3 000, 12 000, 16 000) et un débit requis pour différents types de charges de travail. La plupart des cas d'utilisation nécessitent moins de 3 000 IOPS. Choisissez 12 000 ou 16 000 uniquement si les tests de performances indiquent un besoin d'IOPS élevées. Pour plus d'informations, consultez la [section Configuration](#) de la documentation Timestream for InfluxDB.

Optimisez les écritures

Pour optimiser les écritures dans InfluxDB, nous recommandons d'écrire les données par lots de 5 000 lignes de protocole de ligne par requête afin de minimiser la surcharge réseau. Pour

de meilleures performances, triez les balises par clé dans l'ordre lexicographique avant d'écrire des points de données. L'utilisation de la précision temporelle la plus grossière possible pour les horodatages, au lieu des nanosecondes, peut également améliorer les performances. L'activation de la compression gzip est un autre moyen d'accélérer les écritures et de réduire la bande passante du réseau. Dans la configuration du plugin de `influxdb_v2` sortie de votre `telegraf.conf` fichier, définissez l'`content_encoding` option `surgzip`. La mise en œuvre de ces optimisations peut améliorer de manière significative les performances et l'efficacité de l'écriture de données dans InfluxDB. Pour plus d'informations sur les meilleures pratiques d'écriture d'InfluxDB, voir [Optimiser les écritures](#) dans InfluxDB.

Les performances d'écriture d'InfluxDB sont souvent étroitement liées aux IOPS disponibles. Lors de l'écriture de données, InfluxDB doit effectuer un nombre important d'opérations d'E/S pour stocker les données. Lorsque vous augmentez les IOPS, InfluxDB peut traiter plus d'écritures par seconde.

Pilier d'optimisation des coûts

Le [pilier d'optimisation des coûts](#) du AWS Well-Architected Framework vise à éviter les coûts inutiles et à créer des architectures de manière optimisée en termes de coûts. Les recommandations suivantes peuvent vous aider à respecter les principes de conception d'optimisation des coûts et les meilleures pratiques architecturales pour Amazon Timestream for InfluxDB.

Le pilier de l'optimisation des coûts se concentre sur les domaines clés suivants :

- Comprendre les exigences et les coûts de votre cas d'utilisation
- Sélection des ressources en tenant compte des coûts
- Évolutivité pour répondre aux besoins de l'entreprise sans trop dépenser
- Stockage et transfert de données adaptés

Comprendre les exigences et les coûts de votre cas d'utilisation

Nous recommandons de ne pas utiliser Timestream pour InfluxDB dans les cas d'utilisation suivants :

- Si votre modèle de données contient des données relationnelles, Timestream pour InfluxDB n'est pas la bonne solution.
- Si vous ne pouvez pas utiliser de filtres temporels dans vos requêtes, Influx scannera toutes les séries, ce qui est inefficace.

Sélection des ressources en tenant compte des coûts

Les coûts des [instances InfluxDB](#) sont basés sur un taux horaire pour les heures d'exécution de votre instance. Les instances représentent, en moyenne, 85 % du coût total d'exploitation d'une base de données AWS, de sorte qu'un dimensionnement correct peut avoir des implications financières importantes. Le meilleur moyen de dimensionner correctement les instances est de tester les performances des applications :

- Les CPUUtilization et sont-ils MemoryUtilization constamment élevés ou bas ?
- Quel est l'équilibre entre prix et performance ?

Les coûts des instances évoluent de manière linéaire. Le coût horaire de l'`db.influx.2xlargeinstance` est le double de celui de l'`db.influx.xlargeinstance`, même si l'allocation de ressources y est également deux fois plus élevée. Le coût horaire de l'`db.influx.16xlargeinstance` est 16 fois supérieur à celui de l'`db.influx.xlargeinstance`.

Estimez le nombre d'écritures et de lectures correspondant à votre charge de travail pour une période donnée (seconde, minute, heure ou jour). Timestream pour les instances InfluxDB prend en charge 50 000 à plus de 500 000 écritures par seconde et 10 à 100 requêtes par seconde (QPS) selon le type d'instance. Par exemple, prend `db.influx.2xlarge` généralement en charge jusqu'à 150 000 écritures par seconde et environ 25 QPS. Avec un modèle de données et des requêtes efficaces, il peut dépasser ces performances. Si vos besoins varient en fonction de l'heure de la journée, de la semaine ou du mois, vous pouvez planifier la mise à l'échelle à la hausse ou à la baisse en procédant comme suit :

- [Créez et planifiez une AWS Lambda fonction](#).
- Utilisez un planificateur personnalisé et exécutez une API ou un SDK pour [augmenter ou diminuer la taille avec un certain temps](#) tampon.

Évolutivité pour répondre aux besoins de l'entreprise sans trop dépenser

Pour les expérimentations d'entrée de gamme avec Timestream pour InfluxDB, vous pouvez utiliser `db.influx.medium` `db.influx.large`. Ces instances sont suffisamment grandes pour que vous puissiez acquérir de l'expérience avec Timestream pour InfluxDB avant d'investir dans des instances plus importantes.

Les `db.influx.large` instances `db.influx.medium` et conviennent aux environnements de développement à faible coût. Cependant, ils ont une mémoire RAM plus petite (8 GiB et 16 GiB), moins de v (CPUs 1 vCPU et 2 VCPUs) et des performances réseau allant jusqu'à 10 Go seulement. Toutes les charges de travail ne sont pas adaptées à ces classes d'instances. Surveillez `CPUUtilization` `MemoryUtilization` et augmentez ou diminuez selon les besoins. Il existe souvent un ratio cohérent entre la mémoire et le processeur virtuel. La classe d'instance `db.influx` a un `memory-to-vCPU` ratio similaire à celui de la classe d'instance Amazon EC2 `r7g`. Nous vous recommandons vivement d'effectuer des tests de end-to-end performance ou de charge avant de passer à la production.

La modélisation efficace des données, l'écriture par lots et les requêtes optimisées nécessitent moins de mémoire et d'utilisation des ressources informatiques. Lorsque moins de ressources sont nécessaires, vous pouvez éventuellement utiliser des instances plus petites.

Stockage et transfert de données adaptés

Pour le stockage des données, appliquez les meilleures pratiques suivantes :

- Stockez uniquement les données de séries chronologiques dans Timestream pour InfluxDB.
- Définissez une rétention appropriée sur le bucket InfluxDB afin que les données antérieures à la rétention soient supprimées et que les partitions soient automatiquement compactées périodiquement. Pour plus d'informations, consultez la documentation [InfluxDB](#).
- Optimisez l'utilisation du disque pour les futures écritures.
- Supprimez tous les compartiments InfluxDB qui ne sont pas nécessaires pour vos charges de travail. InfluxDB prend en charge les suppressions. Vous pouvez effectuer des nettoyages planifiés si cela correspond à votre cas d'utilisation.

Pour le transfert de données, nous vous recommandons de déployer votre application de la même manière Région AWS que votre instance de base de données Timestream for InfluxDB afin d'éviter une surcharge réseau entre régions. Il peut également y avoir des frais de transfert de données. Pour plus d'informations sur le transfert de données, consultez la [page de tarification](#).

Pilier de durabilité

Le [pilier du développement durable](#) vise à minimiser les impacts environnementaux liés à l'exécution de charges de travail dans le cloud. Le pilier du développement durable contient les principaux domaines d'intérêt suivants :

- Comprendre votre impact
- Objectifs de durabilité
- Maximiser l'utilisation pour minimiser les ressources
- Anticiper et adopter de nouvelles offres matérielles et logicielles plus efficaces
- Utilisation de services gérés
- Réduire l'impact en aval

Ce guide vise à comprendre votre impact. Pour plus d'informations sur les autres principes de conception durable, consultez le [AWS Well-Architected Framework](#).

Vos choix et vos exigences ont un impact sur l'environnement. Pour augmenter la durabilité de votre charge de travail, procédez comme suit :

- Choisissez ceux Régions AWS qui ont une faible intensité en carbone.
- Dimensionnez vos ressources en fonction des besoins réels en termes de charge de travail au lieu de maximiser le temps de disponibilité et la durabilité.
- Optimisez votre modèle de données et optimisez l'utilisation des ressources informatiques.

Les sections suivantes présentent les pratiques que vous pouvez adopter pour réduire l'impact environnemental de la conception de votre charge de travail et des opérations en cours.

Région AWS sélection

Certains Régions AWS se trouvent à proximité de projets d'énergie renouvelable d'Amazon ou sont situés dans des zones où l'intensité en carbone publiée du réseau est inférieure à celle des autres réseaux. Évaluez les régions en fonction de vos [objectifs de durabilité](#) et de vos exigences en matière de charge de travail. Ensuite, croisez votre liste de régions viables avec les régions où [Timestream for InfluxDB](#) est disponible.

Baser la consommation des ressources sur les modèles de comportement des utilisateurs

L'ajustement de votre consommation en fonction du trafic et du comportement de vos utilisateurs permet de AWS minimiser l'impact des services sur l'environnement. Lors de la conception de votre solution, tenez compte des meilleures pratiques suivantes :

- Surveillez CloudWatch les indicateurs Amazon tels que `CPUUtilization` et `MemoryUtilization` pour déterminer quand votre demande est la plus élevée et la plus faible. Assurez-vous que les ressources de votre instance sont correctement dimensionnées pendant ces périodes.
- Envisagez d'aligner vos accords de niveau de service sur les objectifs de durabilité en plus des objectifs de continuité des activités. L'assouplissement des exigences telles que la reprise après sinistre multirégionale, la haute disponibilité ou la conservation des sauvegardes à long terme peut réduire la quantité de ressources nécessaires pour atteindre ces objectifs. Les environnements hors production et les charges de travail non critiques permettent de réduire les exigences.

Optimisez le développement logiciel et les modèles d'architecture

Pour éviter le gaspillage, optimisez votre modèle de données et vos requêtes. Partagez les ressources de calcul afin d'utiliser toutes les ressources disponibles dans l'instance Timestream for InfluxDB. Nous vous recommandons de mettre en œuvre les meilleures pratiques suivantes :

- Encouragez les équipes de développeurs à partager le stack Timestream for InfluxDB pour une meilleure utilisation des ressources dans la mesure du possible.
- Mettez en œuvre des modèles qui optimisent l'utilisation des ressources et minimisent les temps d'inactivité. Les exemples de modèles incluent l'utilisation de threads parallèles pour charger des données et le regroupement d'enregistrements dans le cadre d'une transaction plus importante.
- Optimisez vos requêtes et votre modèle de données InfluxDB afin de minimiser les ressources nécessaires au calcul des résultats.
- Utilisez les [tâches InfluxDB](#) pour pré-agrégier les données et réduire le temps d'analyse des mêmes données brutes par différents utilisateurs à des fins de visualisation ou de tableaux de bord.
- Maintenez votre Timestream pour les environnements InfluxDB à jour. Les dernières versions de Timestream pour InfluxDB prennent en charge les dernières EC2 instances, telles que Graviton, qui sont plus efficaces. Les dernières versions de la base de données incluent également des

améliorations de l'optimisation des requêtes et des corrections de bogues qui réduisent la quantité de ressources nécessaires au calcul de vos requêtes.

Ressources

Références

- [AWS Well-Architected](#)
- [AWS Documentation du framework Well-Architected](#)
- [Documentation Amazon Timestream pour InfluxDB](#)
- [Documentation d'InfluxDB OSS v2](#)

Billets de blogs

- [Utilisez le script de migration AWS InfluxDB pour migrer vos données InfluxDB OSS 2.x vers Amazon Timestream pour InfluxDB](#)
- [Exécutez et gérez des bases de données InfluxDB open source avec Amazon Timestream](#)

Historique du document

Le tableau suivant décrit les modifications importantes apportées à ce guide. Pour être averti des mises à jour à venir, abonnez-vous à un [fil RSS](#).

Modification	Description	Date
Publication initiale	—	28 janvier 2025

AWS Glossaire des directives prescriptives

Les termes suivants sont couramment utilisés dans les stratégies, les guides et les modèles fournis par les directives AWS prescriptives. Pour suggérer des entrées, veuillez utiliser le lien [Faire un commentaire](#) à la fin du glossaire.

Nombres

7 R

Sept politiques de migration courantes pour transférer des applications vers le cloud. Ces politiques s'appuient sur les 5 R identifiés par Gartner en 2011 et sont les suivantes :

- **Refactorisation/réarchitecture** : transférez une application et modifiez son architecture en tirant pleinement parti des fonctionnalités natives cloud pour améliorer l'agilité, les performances et la capacité de mise à l'échelle. Cela implique généralement le transfert du système d'exploitation et de la base de données. Exemple : migrez votre base de données Oracle sur site vers l'édition compatible avec Amazon Aurora PostgreSQL.
- **Replateformer (déplacer et remodeler)** : transférez une application vers le cloud et introduisez un certain niveau d'optimisation pour tirer parti des fonctionnalités du cloud. Exemple : migrez votre base de données Oracle sur site vers Amazon Relational Database Service (Amazon RDS) pour Oracle dans le AWS Cloud
- **Racheter (rachat)** : optez pour un autre produit, généralement en passant d'une licence traditionnelle à un modèle SaaS. Exemple : migrez votre système de gestion de la relation client (CRM) vers Salesforce.com.
- **Réhéberger (lift and shift)** : transférez une application vers le cloud sans apporter de modifications pour tirer parti des fonctionnalités du cloud. Exemple : migrez votre base de données Oracle locale vers Oracle sur une EC2 instance du AWS Cloud.
- **Relocaliser (lift and shift au niveau de l'hyperviseur)** : transférez l'infrastructure vers le cloud sans acheter de nouveau matériel, réécrire des applications ou modifier vos opérations existantes. Vous migrez des serveurs d'une plateforme sur site vers un service cloud pour la même plateforme. Exemple : migrer une Microsoft Hyper-V application vers AWS.
- **Retenir** : conservez les applications dans votre environnement source. Il peut s'agir d'applications nécessitant une refactorisation majeure, que vous souhaitez retarder, et d'applications existantes que vous souhaitez retenir, car rien ne justifie leur migration sur le plan commercial.

- Retirer : mettez hors service ou supprimez les applications dont vous n'avez plus besoin dans votre environnement source.

A

ABAC

Voir contrôle [d'accès basé sur les attributs](#).

services abstraits

Consultez la section [Services gérés](#).

ACIDE

Voir [atomicité, consistance, isolation, durabilité](#).

migration active-active

Méthode de migration de base de données dans laquelle la synchronisation des bases de données source et cible est maintenue (à l'aide d'un outil de réplication bidirectionnelle ou d'opérations d'écriture double), tandis que les deux bases de données gèrent les transactions provenant de la connexion d'applications pendant la migration. Cette méthode prend en charge la migration par petits lots contrôlés au lieu d'exiger un basculement ponctuel. Elle est plus flexible mais demande plus de travail qu'une migration [active-passive](#).

migration active-passive

Méthode de migration de base de données dans laquelle la synchronisation des bases de données source et cible est maintenue, mais seule la base de données source gère les transactions provenant de la connexion d'applications pendant que les données sont répliquées vers la base de données cible. La base de données cible n'accepte aucune transaction pendant la migration.

fonction d'agrégation

Fonction SQL qui agit sur un groupe de lignes et calcule une valeur de retour unique pour le groupe. Des exemples de fonctions d'agrégation incluent SUM et MAX.

AI

Voir [intelligence artificielle](#).

AIOps

Voir les [opérations d'intelligence artificielle](#).

anonymisation

Processus de suppression définitive d'informations personnelles dans un ensemble de données. L'anonymisation peut contribuer à protéger la vie privée. Les données anonymisées ne sont plus considérées comme des données personnelles.

anti-motif

Solution fréquemment utilisée pour un problème récurrent lorsque la solution est contre-productive, inefficace ou moins efficace qu'une alternative.

contrôle des applications

Une approche de sécurité qui permet d'utiliser uniquement des applications approuvées afin de protéger un système contre les logiciels malveillants.

portefeuille d'applications

Ensemble d'informations détaillées sur chaque application utilisée par une organisation, y compris le coût de génération et de maintenance de l'application, ainsi que sa valeur métier. Ces informations sont essentielles pour [le processus de découverte et d'analyse du portefeuille](#) et permettent d'identifier et de prioriser les applications à migrer, à moderniser et à optimiser.

intelligence artificielle (IA)

Domaine de l'informatique consacré à l'utilisation des technologies de calcul pour exécuter des fonctions cognitives généralement associées aux humains, telles que l'apprentissage, la résolution de problèmes et la reconnaissance de modèles. Pour plus d'informations, veuillez consulter [Qu'est-ce que l'intelligence artificielle ?](#)

opérations d'intelligence artificielle (AIOps)

Processus consistant à utiliser des techniques de machine learning pour résoudre les problèmes opérationnels, réduire les incidents opérationnels et les interventions humaines, mais aussi améliorer la qualité du service. Pour plus d'informations sur son AIOps utilisation dans la stratégie de AWS migration, consultez le [guide d'intégration des opérations](#).

chiffrement asymétrique

Algorithme de chiffrement qui utilise une paire de clés, une clé publique pour le chiffrement et une clé privée pour le déchiffrement. Vous pouvez partager la clé publique, car elle n'est pas utilisée pour le déchiffrement, mais l'accès à la clé privée doit être très restreint.

atomicité, cohérence, isolement, durabilité (ACID)

Ensemble de propriétés logicielles garantissant la validité des données et la fiabilité opérationnelle d'une base de données, même en cas d'erreur, de panne de courant ou d'autres problèmes.

contrôle d'accès par attributs (ABAC)

Pratique qui consiste à créer des autorisations détaillées en fonction des attributs de l'utilisateur, tels que le service, le poste et le nom de l'équipe. Pour plus d'informations, consultez [ABAC pour AWS](#) dans la documentation AWS Identity and Access Management (IAM).

source de données faisant autorité

Emplacement où vous stockez la version principale des données, considérée comme la source d'information la plus fiable. Vous pouvez copier les données de la source de données officielle vers d'autres emplacements à des fins de traitement ou de modification des données, par exemple en les anonymisant, en les expurgant ou en les pseudonymisant.

Zone de disponibilité

Un emplacement distinct au sein d'une Région AWS réseau isolé des défaillances dans d'autres zones de disponibilité et fournissant une connectivité réseau peu coûteuse et à faible latence aux autres zones de disponibilité de la même région.

AWS Cadre d'adoption du cloud (AWS CAF)

Un cadre de directives et de meilleures pratiques visant AWS à aider les entreprises à élaborer un plan efficace pour réussir leur migration vers le cloud. AWS La CAF organise ses conseils en six domaines prioritaires appelés perspectives : les affaires, les personnes, la gouvernance, les plateformes, la sécurité et les opérations. Les perspectives d'entreprise, de personnes et de gouvernance mettent l'accent sur les compétences et les processus métier, tandis que les perspectives relatives à la plateforme, à la sécurité et aux opérations se concentrent sur les compétences et les processus techniques. Par exemple, la perspective liée aux personnes cible les parties prenantes qui s'occupent des ressources humaines (RH), des fonctions de dotation en personnel et de la gestion des personnes. Dans cette perspective, la AWS CAF fournit des conseils pour le développement du personnel, la formation et les communications afin de préparer

l'organisation à une adoption réussie du cloud. Pour plus d'informations, veuillez consulter le [site Web AWS CAF](#) et le [livre blanc AWS CAF](#).

AWS Cadre de qualification de la charge de travail (AWS WQF)

Outil qui évalue les charges de travail liées à la migration des bases de données, recommande des stratégies de migration et fournit des estimations de travail. AWS Le WQF est inclus avec AWS Schema Conversion Tool (AWS SCT). Il analyse les schémas de base de données et les objets de code, le code d'application, les dépendances et les caractéristiques de performance, et fournit des rapports d'évaluation.

B

mauvais bot

Un [bot](#) destiné à perturber ou à nuire à des individus ou à des organisations.

BCP

Consultez la section [Planification de la continuité des activités](#).

graphique de comportement

Vue unifiée et interactive des comportements des ressources et des interactions au fil du temps. Vous pouvez utiliser un graphique de comportement avec Amazon Detective pour examiner les tentatives de connexion infructueuses, les appels d'API suspects et les actions similaires. Pour plus d'informations, veuillez consulter [Data in a behavior graph](#) dans la documentation Detective.

système de poids fort

Système qui stocke d'abord l'octet le plus significatif. Voir aussi [endianité](#).

classification binaire

Processus qui prédit un résultat binaire (l'une des deux classes possibles). Par exemple, votre modèle de machine learning peut avoir besoin de prévoir des problèmes tels que « Cet e-mail est-il du spam ou non ? » ou « Ce produit est-il un livre ou une voiture ? ».

filtre de Bloom

Structure de données probabiliste et efficace en termes de mémoire qui est utilisée pour tester si un élément fait partie d'un ensemble.

déploiement bleu/vert

Stratégie de déploiement dans laquelle vous créez deux environnements distincts mais identiques. Vous exécutez la version actuelle de l'application dans un environnement (bleu) et la nouvelle version de l'application dans l'autre environnement (vert). Cette stratégie vous permet de revenir rapidement en arrière avec un impact minimal.

bot

Application logicielle qui exécute des tâches automatisées sur Internet et simule l'activité ou l'interaction humaine. Certains robots sont utiles ou bénéfiques, comme les robots d'exploration Web qui indexent des informations sur Internet. D'autres robots, appelés « bots malveillants », sont destinés à perturber ou à nuire à des individus ou à des organisations.

botnet

Réseaux de [robots](#) infectés par des [logiciels malveillants](#) et contrôlés par une seule entité, connue sous le nom d'herder ou d'opérateur de bots. Les botnets sont le mécanisme le plus connu pour faire évoluer les bots et leur impact.

branche

Zone contenue d'un référentiel de code. La première branche créée dans un référentiel est la branche principale. Vous pouvez créer une branche à partir d'une branche existante, puis développer des fonctionnalités ou corriger des bogues dans la nouvelle branche. Une branche que vous créez pour générer une fonctionnalité est communément appelée branche de fonctionnalités. Lorsque la fonctionnalité est prête à être publiée, vous fusionnez à nouveau la branche de fonctionnalités dans la branche principale. Pour plus d'informations, consultez [À propos des branches](#) (GitHub documentation).

accès par brise-vitre

Dans des circonstances exceptionnelles et par le biais d'un processus approuvé, c'est un moyen rapide pour un utilisateur d'accéder à un accès auquel Compte AWS il n'est généralement pas autorisé. Pour plus d'informations, consultez l'indicateur [Implementation break-glass procedures](#) dans le guide Well-Architected AWS .

stratégie existante (brownfield)

L'infrastructure existante de votre environnement. Lorsque vous adoptez une stratégie existante pour une architecture système, vous concevez l'architecture en fonction des contraintes des systèmes et de l'infrastructure actuels. Si vous étendez l'infrastructure existante, vous pouvez combiner des politiques brownfield (existantes) et [greenfield](#) (inédites).

cache de tampon

Zone de mémoire dans laquelle sont stockées les données les plus fréquemment consultées.

capacité métier

Ce que fait une entreprise pour générer de la valeur (par exemple, les ventes, le service client ou le marketing). Les architectures de microservices et les décisions de développement peuvent être dictées par les capacités métier. Pour plus d'informations, veuillez consulter la section [Organisation en fonction des capacités métier](#) du livre blanc [Exécution de microservices conteneurisés sur AWS](#).

planification de la continuité des activités (BCP)

Plan qui tient compte de l'impact potentiel d'un événement perturbateur, tel qu'une migration à grande échelle, sur les opérations, et qui permet à une entreprise de reprendre ses activités rapidement.

C

CAF

Voir le [cadre d'adoption du AWS cloud](#).

déploiement de Canary

Diffusion lente et progressive d'une version pour les utilisateurs finaux. Lorsque vous êtes sûr, vous déployez la nouvelle version et remplacez la version actuelle dans son intégralité.

CCo E

Voir [le Centre d'excellence du cloud](#).

CDC

Voir [capture des données de modification](#).

capture des données de modification (CDC)

Processus de suivi des modifications apportées à une source de données, telle qu'une table de base de données, et d'enregistrement des métadonnées relatives à ces modifications. Vous pouvez utiliser la CDC à diverses fins, telles que l'audit ou la réplication des modifications dans un système cible afin de maintenir la synchronisation.

ingénierie du chaos

Introduire intentionnellement des défaillances ou des événements perturbateurs pour tester la résilience d'un système. Vous pouvez utiliser [AWS Fault Injection Service \(AWS FIS\)](#) pour effectuer des expériences qui stressent vos AWS charges de travail et évaluer leur réponse.

CI/CD

Découvrez [l'intégration continue et la livraison continue](#).

classification

Processus de catégorisation qui permet de générer des prédictions. Les modèles de ML pour les problèmes de classification prédisent une valeur discrète. Les valeurs discrètes se distinguent toujours les unes des autres. Par exemple, un modèle peut avoir besoin d'évaluer la présence ou non d'une voiture sur une image.

chiffrement côté client

Chiffrement des données localement, avant que la cible ne les Service AWS reçoive.

Centre d'excellence du cloud (CCoE)

Une équipe multidisciplinaire qui dirige les efforts d'adoption du cloud au sein d'une organisation, notamment en développant les bonnes pratiques en matière de cloud, en mobilisant des ressources, en établissant des délais de migration et en guidant l'organisation dans le cadre de transformations à grande échelle. Pour plus d'informations, consultez les [CCoarticles électroniques](#) du blog sur la stratégie AWS Cloud d'entreprise.

cloud computing

Technologie cloud généralement utilisée pour le stockage de données à distance et la gestion des appareils IoT. Le cloud computing est généralement associé à la technologie [informatique de pointe](#).

modèle d'exploitation du cloud

Dans une organisation informatique, modèle d'exploitation utilisé pour créer, faire évoluer et optimiser un ou plusieurs environnements cloud. Pour plus d'informations, consultez la section [Création de votre modèle d'exploitation cloud](#).

étapes d'adoption du cloud

Les quatre phases que les entreprises traversent généralement lorsqu'elles migrent vers AWS Cloud :

- **Projet** : exécution de quelques projets liés au cloud à des fins de preuve de concept et d'apprentissage
- **Base** : réaliser des investissements fondamentaux pour accélérer votre adoption du cloud (par exemple, créer une zone de landing zone, définir un CCo E, établir un modèle opérationnel)
- **Migration** : migration d'applications individuelles
- **Réinvention** : optimisation des produits et services et innovation dans le cloud

Ces étapes ont été définies par Stephen Orban dans le billet de blog [The Journey Toward Cloud-First & the Stages of Adoption](#) publié sur le blog AWS Cloud Enterprise Strategy. Pour plus d'informations sur leur lien avec la stratégie de AWS migration, consultez le [guide de préparation à la migration](#).

CMDB

Voir base de [données de gestion de configuration](#).

référentiel de code

Emplacement où le code source et d'autres ressources, comme la documentation, les exemples et les scripts, sont stockés et mis à jour par le biais de processus de contrôle de version. Les référentiels cloud courants incluent GitHub ou Bitbucket Cloud. Chaque version du code est appelée branche. Dans une structure de microservice, chaque référentiel est consacré à une seule fonctionnalité. Un seul pipeline CI/CD peut utiliser plusieurs référentiels.

cache passif

Cache tampon vide, mal rempli ou contenant des données obsolètes ou non pertinentes. Cela affecte les performances, car l'instance de base de données doit lire à partir de la mémoire principale ou du disque, ce qui est plus lent que la lecture à partir du cache tampon.

données gelées

Données rarement consultées et généralement historiques. Lorsque vous interrogez ce type de données, les requêtes lentes sont généralement acceptables. Le transfert de ces données vers des niveaux ou classes de stockage moins performants et moins coûteux peut réduire les coûts.

vision par ordinateur (CV)

Domaine de l'[IA](#) qui utilise l'apprentissage automatique pour analyser et extraire des informations à partir de formats visuels tels que des images numériques et des vidéos. Par exemple, Amazon SageMaker AI fournit des algorithmes de traitement d'image pour les CV.

dérive de configuration

Pour une charge de travail, une modification de configuration par rapport à l'état attendu. Cela peut entraîner une non-conformité de la charge de travail, et cela est généralement progressif et involontaire.

base de données de gestion des configurations (CMDB)

Référentiel qui stocke et gère les informations relatives à une base de données et à son environnement informatique, y compris les composants matériels et logiciels ainsi que leurs configurations. Vous utilisez généralement les données d'une CMDB lors de la phase de découverte et d'analyse du portefeuille de la migration.

pack de conformité

Ensemble de AWS Config règles et d'actions correctives que vous pouvez assembler pour personnaliser vos contrôles de conformité et de sécurité. Vous pouvez déployer un pack de conformité en tant qu'entité unique dans une région Compte AWS et, ou au sein d'une organisation, à l'aide d'un modèle YAML. Pour plus d'informations, consultez la section [Packs de conformité](#) dans la AWS Config documentation.

intégration continue et livraison continue (CI/CD)

Processus d'automatisation des étapes de source, de construction, de test, de préparation et de production du processus de publication du logiciel. CI/CD is commonly described as a pipeline. CI/CD peut vous aider à automatiser les processus, à améliorer la productivité, à améliorer la qualité du code et à accélérer les livraisons. Pour plus d'informations, veuillez consulter [Avantages de la livraison continue](#). CD peut également signifier déploiement continu. Pour plus d'informations, veuillez consulter [Livraison continue et déploiement continu](#).

CV

Voir [vision par ordinateur](#).

D

données au repos

Données stationnaires dans votre réseau, telles que les données stockées.

classification des données

Processus permettant d'identifier et de catégoriser les données de votre réseau en fonction de leur sévérité et de leur sensibilité. Il s'agit d'un élément essentiel de toute stratégie de gestion des risques de cybersécurité, car il vous aide à déterminer les contrôles de protection et de conservation appropriés pour les données. La classification des données est une composante du pilier de sécurité du AWS Well-Architected Framework. Pour plus d'informations, veuillez consulter [Classification des données](#).

dérive des données

Une variation significative entre les données de production et les données utilisées pour entraîner un modèle ML, ou une modification significative des données d'entrée au fil du temps. La dérive des données peut réduire la qualité, la précision et l'équité globales des prédictions des modèles ML.

données en transit

Données qui circulent activement sur votre réseau, par exemple entre les ressources du réseau.

maillage de données

Un cadre architectural qui fournit une propriété des données distribuée et décentralisée avec une gestion et une gouvernance centralisées.

minimisation des données

Le principe de collecte et de traitement des seules données strictement nécessaires. La pratique de la minimisation des données AWS Cloud peut réduire les risques liés à la confidentialité, les coûts et l'empreinte carbone de vos analyses.

périmètre de données

Ensemble de garde-fous préventifs dans votre AWS environnement qui permettent de garantir que seules les identités fiables accèdent aux ressources fiables des réseaux attendus. Pour plus d'informations, voir [Création d'un périmètre de données sur AWS](#).

prétraitement des données

Pour transformer les données brutes en un format facile à analyser par votre modèle de ML. Le prétraitement des données peut impliquer la suppression de certaines colonnes ou lignes et le traitement des valeurs manquantes, incohérentes ou en double.

provenance des données

Le processus de suivi de l'origine et de l'historique des données tout au long de leur cycle de vie, par exemple la manière dont les données ont été générées, transmises et stockées.

sujet des données

Personne dont les données sont collectées et traitées.

entrepôt des données

Un système de gestion des données qui prend en charge les informations commerciales, telles que les analyses. Les entrepôts de données contiennent généralement de grandes quantités de données historiques et sont généralement utilisés pour les requêtes et les analyses.

langage de définition de base de données (DDL)

Instructions ou commandes permettant de créer ou de modifier la structure des tables et des objets dans une base de données.

langage de manipulation de base de données (DML)

Instructions ou commandes permettant de modifier (insérer, mettre à jour et supprimer) des informations dans une base de données.

DDL

Voir [langage de définition de base](#) de données.

ensemble profond

Sert à combiner plusieurs modèles de deep learning à des fins de prédiction. Vous pouvez utiliser des ensembles profonds pour obtenir une prévision plus précise ou pour estimer l'incertitude des prédictions.

deep learning

Un sous-champ de ML qui utilise plusieurs couches de réseaux neuronaux artificiels pour identifier le mappage entre les données d'entrée et les variables cibles d'intérêt.

defense-in-depth

Approche de la sécurité de l'information dans laquelle une série de mécanismes et de contrôles de sécurité sont judicieusement répartis sur l'ensemble d'un réseau informatique afin de protéger la confidentialité, l'intégrité et la disponibilité du réseau et des données qu'il contient. Lorsque vous adoptez cette stratégie AWS, vous ajoutez plusieurs contrôles à différentes couches de

la AWS Organizations structure afin de sécuriser les ressources. Par exemple, une defense-in-depth approche peut combiner l'authentification multifactorielle, la segmentation du réseau et le chiffrement.

administrateur délégué

Dans AWS Organizations, un service compatible peut enregistrer un compte AWS membre pour administrer les comptes de l'organisation et gérer les autorisations pour ce service. Ce compte est appelé administrateur délégué pour ce service. Pour plus d'informations et une liste des services compatibles, veuillez consulter la rubrique [Services qui fonctionnent avec AWS Organizations](#) dans la documentation AWS Organizations .

déploiement

Processus de mise à disposition d'une application, de nouvelles fonctionnalités ou de corrections de code dans l'environnement cible. Le déploiement implique la mise en œuvre de modifications dans une base de code, puis la génération et l'exécution de cette base de code dans les environnements de l'application.

environnement de développement

Voir [environnement](#).

contrôle de détection

Contrôle de sécurité conçu pour détecter, journaliser et alerter après la survenue d'un événement. Ces contrôles constituent une deuxième ligne de défense et vous alertent en cas d'événements de sécurité qui ont contourné les contrôles préventifs en place. Pour plus d'informations, veuillez consulter la rubrique [Contrôles de détection](#) dans Implementing security controls on AWS.

cartographie de la chaîne de valeur du développement (DVSM)

Processus utilisé pour identifier et hiérarchiser les contraintes qui nuisent à la rapidité et à la qualité du cycle de vie du développement logiciel. DVSM étend le processus de cartographie de la chaîne de valeur initialement conçu pour les pratiques de production allégée. Il met l'accent sur les étapes et les équipes nécessaires pour créer et transférer de la valeur tout au long du processus de développement logiciel.

jumeau numérique

Représentation virtuelle d'un système réel, tel qu'un bâtiment, une usine, un équipement industriel ou une ligne de production. Les jumeaux numériques prennent en charge la maintenance prédictive, la surveillance à distance et l'optimisation de la production.

tableau des dimensions

Dans un [schéma en étoile](#), table plus petite contenant les attributs de données relatifs aux données quantitatives d'une table de faits. Les attributs des tables de dimensions sont généralement des champs de texte ou des nombres discrets qui se comportent comme du texte. Ces attributs sont couramment utilisés pour la contrainte des requêtes, le filtrage et l'étiquetage des ensembles de résultats.

catastrophe

Un événement qui empêche une charge de travail ou un système d'atteindre ses objectifs commerciaux sur son site de déploiement principal. Ces événements peuvent être des catastrophes naturelles, des défaillances techniques ou le résultat d'actions humaines, telles qu'une mauvaise configuration involontaire ou une attaque de logiciel malveillant.

reprise après sinistre (DR)

La stratégie et le processus que vous utilisez pour minimiser les temps d'arrêt et les pertes de données causés par un [sinistre](#). Pour plus d'informations, consultez [Disaster Recovery of Workloads on AWS : Recovery in the Cloud in the AWS Well-Architected Framework](#).

DML

Voir [langage de manipulation de base](#) de données.

conception axée sur le domaine

Approche visant à développer un système logiciel complexe en connectant ses composants à des domaines évolutifs, ou objectifs métier essentiels, que sert chaque composant. Ce concept a été introduit par Eric Evans dans son ouvrage Domain-Driven Design: Tackling Complexity in the Heart of Software (Boston : Addison-Wesley Professional, 2003). Pour plus d'informations sur l'utilisation du design piloté par domaine avec le modèle de figuier étrangleur, veuillez consulter [Modernizing legacy Microsoft ASP.NET \(ASMX\) web services incrementally by using containers and Amazon API Gateway](#).

DR

Voir [reprise après sinistre](#).

détection de dérive

Suivi des écarts par rapport à une configuration de référence. Par exemple, vous pouvez l'utiliser AWS CloudFormation pour [détecter la dérive des ressources du système](#) ou AWS Control Tower

pour [détecter les modifications de votre zone d'atterrissage](#) susceptibles d'affecter le respect des exigences de gouvernance.

DVSM

Voir la [cartographie de la chaîne de valeur du développement](#).

E

EDA

Voir [analyse exploratoire des données](#).

EDI

Voir échange [de données informatisé](#).

informatique de périphérie

Technologie qui augmente la puissance de calcul des appareils intelligents en périphérie d'un réseau IoT. Comparé au [cloud computing, l'informatique](#) de pointe peut réduire la latence des communications et améliorer le temps de réponse.

échange de données informatisé (EDI)

L'échange automatique de documents commerciaux entre les organisations. Pour plus d'informations, voir [Qu'est-ce que l'échange de données informatisé ?](#)

chiffrement

Processus informatique qui transforme des données en texte clair, lisibles par l'homme, en texte chiffré.

clé de chiffrement

Chaîne cryptographique de bits aléatoires générée par un algorithme cryptographique. La longueur des clés peut varier, et chaque clé est conçue pour être imprévisible et unique.

endianisme

Ordre selon lequel les octets sont stockés dans la mémoire de l'ordinateur. Les systèmes de poids fort stockent d'abord l'octet le plus significatif. Les systèmes de poids faible stockent d'abord l'octet le moins significatif.

point de terminaison

Voir [point de terminaison de service](#).

service de point de terminaison

Service que vous pouvez héberger sur un cloud privé virtuel (VPC) pour le partager avec d'autres utilisateurs. Vous pouvez créer un service de point de terminaison avec AWS PrivateLink et accorder des autorisations à d'autres Comptes AWS ou à AWS Identity and Access Management (IAM) principaux. Ces comptes ou principaux peuvent se connecter à votre service de point de terminaison de manière privée en créant des points de terminaison d'un VPC d'interface. Pour plus d'informations, veuillez consulter [Création d'un service de point de terminaison](#) dans la documentation Amazon Virtual Private Cloud (Amazon VPC).

planification des ressources d'entreprise (ERP)

Système qui automatise et gère les principaux processus métier (tels que la comptabilité, le [MES](#) et la gestion de projet) pour une entreprise.

chiffrement d'enveloppe

Processus de chiffrement d'une clé de chiffrement à l'aide d'une autre clé de chiffrement. Pour plus d'informations, consultez la section [Chiffrement des enveloppes](#) dans la documentation AWS Key Management Service (AWS KMS).

environnement

Instance d'une application en cours d'exécution. Les types d'environnement les plus courants dans le cloud computing sont les suivants :

- Environnement de développement : instance d'une application en cours d'exécution à laquelle seule l'équipe principale chargée de la maintenance de l'application peut accéder. Les environnements de développement sont utilisés pour tester les modifications avant de les promouvoir dans les environnements supérieurs. Ce type d'environnement est parfois appelé environnement de test.
- Environnements inférieurs : tous les environnements de développement d'une application, tels que ceux utilisés pour les générations et les tests initiaux.
- Environnement de production : instance d'une application en cours d'exécution à laquelle les utilisateurs finaux peuvent accéder. Dans un pipeline CI/CD, l'environnement de production est le dernier environnement de déploiement.
- Environnements supérieurs : tous les environnements accessibles aux utilisateurs autres que l'équipe de développement principale. Ils peuvent inclure un environnement de production, des

environnements de préproduction et des environnements pour les tests d'acceptation par les utilisateurs.

épopée

Dans les méthodologies agiles, catégories fonctionnelles qui aident à organiser et à prioriser votre travail. Les épopées fournissent une description détaillée des exigences et des tâches d'implémentation. Par exemple, les points forts de la AWS CAF en matière de sécurité incluent la gestion des identités et des accès, les contrôles de détection, la sécurité des infrastructures, la protection des données et la réponse aux incidents. Pour plus d'informations sur les épopées dans la stratégie de migration AWS , veuillez consulter le [guide d'implémentation du programme](#).

ERP

Voir [Planification des ressources d'entreprise](#).

analyse exploratoire des données (EDA)

Processus d'analyse d'un jeu de données pour comprendre ses principales caractéristiques. Vous collectez ou agrégez des données, puis vous effectuez des enquêtes initiales pour trouver des modèles, détecter des anomalies et vérifier les hypothèses. L'EDA est réalisée en calculant des statistiques récapitulatives et en créant des visualisations de données.

F

tableau des faits

La table centrale dans un [schéma en étoile](#). Il stocke des données quantitatives sur les opérations commerciales. Généralement, une table de faits contient deux types de colonnes : celles qui contiennent des mesures et celles qui contiennent une clé étrangère pour une table de dimensions.

échouer rapidement

Une philosophie qui utilise des tests fréquents et progressifs pour réduire le cycle de vie du développement. C'est un élément essentiel d'une approche agile.

limite d'isolation des défauts

Dans le AWS Cloud, une limite telle qu'une zone de disponibilité Région AWS, un plan de contrôle ou un plan de données qui limite l'effet d'une panne et contribue à améliorer la résilience des

charges de travail. Pour plus d'informations, consultez la section [Limites d'isolation des AWS pannes](#).

branche de fonctionnalités

Voir [succursale](#).

fonctionnalités

Les données d'entrée que vous utilisez pour faire une prédiction. Par exemple, dans un contexte de fabrication, les fonctionnalités peuvent être des images capturées périodiquement à partir de la ligne de fabrication.

importance des fonctionnalités

Le niveau d'importance d'une fonctionnalité pour les prédictions d'un modèle. Il s'exprime généralement sous la forme d'un score numérique qui peut être calculé à l'aide de différentes techniques, telles que la méthode Shapley Additive Explanations (SHAP) et les gradients intégrés. Pour plus d'informations, voir [Interprétabilité du modèle d'apprentissage automatique avec AWS](#).

transformation de fonctionnalité

Optimiser les données pour le processus de ML, notamment en enrichissant les données avec des sources supplémentaires, en mettant à l'échelle les valeurs ou en extrayant plusieurs ensembles d'informations à partir d'un seul champ de données. Cela permet au modèle de ML de tirer parti des données. Par exemple, si vous décomposez la date « 2021-05-27 00:15:37 » en « 2021 », « mai », « jeudi » et « 15 », vous pouvez aider l'algorithme d'apprentissage à apprendre des modèles nuancés associés à différents composants de données.

invitation en quelques coups

Fournir à un [LLM](#) un petit nombre d'exemples illustrant la tâche et le résultat souhaité avant de lui demander d'effectuer une tâche similaire. Cette technique est une application de l'apprentissage contextuel, dans le cadre de laquelle les modèles apprennent à partir d'exemples (prises de vue) intégrés dans des instructions. Les instructions en quelques étapes peuvent être efficaces pour les tâches qui nécessitent un formatage, un raisonnement ou des connaissances de domaine spécifiques. Voir également l'[invite Zero-Shot](#).

FGAC

Découvrez le [contrôle d'accès détaillé](#).

contrôle d'accès détaillé (FGAC)

Utilisation de plusieurs conditions pour autoriser ou refuser une demande d'accès.

migration instantanée (flash-cut)

Méthode de migration de base de données qui utilise la réplication continue des données par [le biais de la capture des données de modification](#) afin de migrer les données dans les plus brefs délais, au lieu d'utiliser une approche progressive. L'objectif est de réduire au maximum les temps d'arrêt.

FM

Voir le [modèle de fondation](#).

modèle de fondation (FM)

Un vaste réseau neuronal d'apprentissage profond qui s'est entraîné sur d'énormes ensembles de données généralisées et non étiquetées. FMs sont capables d'effectuer une grande variété de tâches générales, telles que comprendre le langage, générer du texte et des images et converser en langage naturel. Pour plus d'informations, voir [Que sont les modèles de base ?](#)

G

IA générative

Sous-ensemble de modèles d'[IA](#) qui ont été entraînés sur de grandes quantités de données et qui peuvent utiliser une simple invite textuelle pour créer de nouveaux contenus et artefacts, tels que des images, des vidéos, du texte et du son. Pour plus d'informations, consultez [Qu'est-ce que l'IA générative](#).

blocage géographique

Voir les [restrictions géographiques](#).

restrictions géographiques (blocage géographique)

Sur Amazon CloudFront, option permettant d'empêcher les utilisateurs de certains pays d'accéder aux distributions de contenu. Vous pouvez utiliser une liste d'autorisation ou une liste de blocage pour spécifier les pays approuvés et interdits. Pour plus d'informations, consultez [la section Restreindre la distribution géographique de votre contenu](#) dans la CloudFront documentation.

Flux de travail Gitflow

Approche dans laquelle les environnements inférieurs et supérieurs utilisent différentes branches dans un référentiel de code source. Le flux de travail Gitflow est considéré comme existant, et le [flux de travail basé sur les tronc](#) est l'approche moderne préférée.

image dorée

Un instantané d'un système ou d'un logiciel utilisé comme modèle pour déployer de nouvelles instances de ce système ou logiciel. Par exemple, dans le secteur de la fabrication, une image dorée peut être utilisée pour fournir des logiciels sur plusieurs appareils et contribue à améliorer la vitesse, l'évolutivité et la productivité des opérations de fabrication des appareils.

stratégie inédite

L'absence d'infrastructures existantes dans un nouvel environnement. Lorsque vous adoptez une stratégie inédite pour une architecture système, vous pouvez sélectionner toutes les nouvelles technologies sans restriction de compatibilité avec l'infrastructure existante, également appelée [brownfield](#). Si vous étendez l'infrastructure existante, vous pouvez combiner des politiques brownfield (existantes) et greenfield (inédites).

barrière de protection

Règle de haut niveau qui permet de régir les ressources, les politiques et la conformité au sein des unités organisationnelles (OUs). Les barrières de protection préventives appliquent des politiques pour garantir l'alignement sur les normes de conformité. Elles sont mises en œuvre à l'aide de politiques de contrôle des services et de limites des autorisations IAM. Les barrières de protection de détection détectent les violations des politiques et les problèmes de conformité, et génèrent des alertes pour y remédier. Ils sont implémentés à l'aide d'Amazon AWS Config AWS Security Hub GuardDuty AWS Trusted Advisor, d'Amazon Inspector et de AWS Lambda contrôles personnalisés.

H

HA

Découvrez [la haute disponibilité](#).

migration de base de données hétérogène

Migration de votre base de données source vers une base de données cible qui utilise un moteur de base de données différent (par exemple, Oracle vers Amazon Aurora). La migration hétérogène fait généralement partie d'un effort de réarchitecture, et la conversion du schéma peut s'avérer une tâche complexe. [AWS propose AWS SCT](#) qui facilite les conversions de schémas.

haute disponibilité (HA)

Capacité d'une charge de travail à fonctionner en continu, sans intervention, en cas de difficultés ou de catastrophes. Les systèmes HA sont conçus pour basculer automatiquement, fournir constamment des performances de haute qualité et gérer différentes charges et défaillances avec un impact minimal sur les performances.

modernisation des historiens

Approche utilisée pour moderniser et mettre à niveau les systèmes de technologie opérationnelle (OT) afin de mieux répondre aux besoins de l'industrie manufacturière. Un historien est un type de base de données utilisé pour collecter et stocker des données provenant de diverses sources dans une usine.

données de rétention

Partie de données historiques étiquetées qui n'est pas divulguée dans un ensemble de données utilisé pour entraîner un modèle d'[apprentissage automatique](#). Vous pouvez utiliser les données de blocage pour évaluer les performances du modèle en comparant les prévisions du modèle aux données de blocage.

migration de base de données homogène

Migration de votre base de données source vers une base de données cible qui partage le même moteur de base de données (par exemple, Microsoft SQL Server vers Amazon RDS for SQL Server). La migration homogène s'inscrit généralement dans le cadre d'un effort de réhébergement ou de replatforme. Vous pouvez utiliser les utilitaires de base de données natifs pour migrer le schéma.

données chaudes

Données fréquemment consultées, telles que les données en temps réel ou les données translationnelles récentes. Ces données nécessitent généralement un niveau ou une classe de stockage à hautes performances pour fournir des réponses rapides aux requêtes.

correctif

Solution d'urgence à un problème critique dans un environnement de production. En raison de son urgence, un correctif est généralement créé en dehors du flux de travail de DevOps publication habituel.

période de soins intensifs

Immédiatement après le basculement, période pendant laquelle une équipe de migration gère et surveille les applications migrées dans le cloud afin de résoudre les problèmes éventuels. En règle générale, cette période dure de 1 à 4 jours. À la fin de la période de soins intensifs, l'équipe de migration transfère généralement la responsabilité des applications à l'équipe des opérations cloud.

I

IaC

Considérez [l'infrastructure comme un code](#).

politique basée sur l'identité

Politique attachée à un ou plusieurs principaux IAM qui définit leurs autorisations au sein de l'AWS Cloud environnement.

application inactive

Application dont l'utilisation moyenne du processeur et de la mémoire se situe entre 5 et 20 % sur une période de 90 jours. Dans un projet de migration, il est courant de retirer ces applications ou de les retenir sur site.

Ilo T

Voir [Internet industriel des objets](#).

infrastructure immuable

Modèle qui déploie une nouvelle infrastructure pour les charges de travail de production au lieu de mettre à jour, d'appliquer des correctifs ou de modifier l'infrastructure existante. Les infrastructures immuables sont intrinsèquement plus cohérentes, fiables et prévisibles que les infrastructures [mutables](#). Pour plus d'informations, consultez les meilleures pratiques de [déploiement à l'aide d'une infrastructure immuable](#) dans le AWS Well-Architected Framework.

VPC entrant (d'entrée)

Dans une architecture AWS multi-comptes, un VPC qui accepte, inspecte et achemine les connexions réseau depuis l'extérieur d'une application. L'[architecture AWS de référence de sécurité](#) recommande de configurer votre compte réseau avec les fonctions entrantes, sortantes

et d'inspection VPCs afin de protéger l'interface bidirectionnelle entre votre application et l'Internet en général.

migration incrémentielle

Stratégie de basculement dans le cadre de laquelle vous migrez votre application par petites parties au lieu d'effectuer un basculement complet unique. Par exemple, il se peut que vous ne transfériez que quelques microservices ou utilisateurs vers le nouveau système dans un premier temps. Après avoir vérifié que tout fonctionne correctement, vous pouvez transférer progressivement des microservices ou des utilisateurs supplémentaires jusqu'à ce que vous puissiez mettre hors service votre système hérité. Cette stratégie réduit les risques associés aux migrations de grande ampleur.

Industry 4.0

Terme introduit par [Klaus Schwab](#) en 2016 pour désigner la modernisation des processus de fabrication grâce aux avancées en matière de connectivité, de données en temps réel, d'automatisation, d'analyse et d'IA/ML.

infrastructure

Ensemble des ressources et des actifs contenus dans l'environnement d'une application.

infrastructure en tant que code (IaC)

Processus de mise en service et de gestion de l'infrastructure d'une application via un ensemble de fichiers de configuration. IaC est conçue pour vous aider à centraliser la gestion de l'infrastructure, à normaliser les ressources et à mettre à l'échelle rapidement afin que les nouveaux environnements soient reproductibles, fiables et cohérents.

Internet industriel des objets (IIoT)

L'utilisation de capteurs et d'appareils connectés à Internet dans les secteurs industriels tels que la fabrication, l'énergie, l'automobile, les soins de santé, les sciences de la vie et l'agriculture.

Pour plus d'informations, voir [Élaboration d'une stratégie de transformation numérique de l'Internet des objets \(IIoT\) industriel](#).

VPC d'inspection

Dans une architecture AWS multi-comptes, un VPC centralisé qui gère les inspections du trafic réseau VPCs entre (identique ou Régions AWS différent), Internet et les réseaux locaux. L'[architecture AWS de référence de sécurité](#) recommande de configurer votre compte réseau

avec les fonctions entrantes, sortantes et d'inspection VPCs afin de protéger l'interface bidirectionnelle entre votre application et l'Internet en général.

Internet des objets (IoT)

Réseau d'objets physiques connectés dotés de capteurs ou de processeurs intégrés qui communiquent avec d'autres appareils et systèmes via Internet ou via un réseau de communication local. Pour plus d'informations, veuillez consulter la section [Qu'est-ce que l'IoT ?](#).

interprétabilité

Caractéristique d'un modèle de machine learning qui décrit dans quelle mesure un être humain peut comprendre comment les prédictions du modèle dépendent de ses entrées. Pour plus d'informations, voir [Interprétabilité du modèle d'apprentissage automatique avec AWS](#).

IoT

Voir [Internet des objets](#).

Bibliothèque d'informations informatiques (ITIL)

Ensemble de bonnes pratiques pour proposer des services informatiques et les aligner sur les exigences métier. L'ITIL constitue la base de l'ITSM.

gestion des services informatiques (ITSM)

Activités associées à la conception, à la mise en œuvre, à la gestion et à la prise en charge de services informatiques d'une organisation. Pour plus d'informations sur l'intégration des opérations cloud aux outils ITSM, veuillez consulter le [guide d'intégration des opérations](#).

ITIL

Consultez la [bibliothèque d'informations informatiques](#).

ITSM

Voir [Gestion des services informatiques](#).

L

contrôle d'accès basé sur des étiquettes (LBAC)

Une implémentation du contrôle d'accès obligatoire (MAC) dans laquelle une valeur d'étiquette de sécurité est explicitement attribuée aux utilisateurs et aux données elles-mêmes. L'intersection

entre l'étiquette de sécurité utilisateur et l'étiquette de sécurité des données détermine les lignes et les colonnes visibles par l'utilisateur.

zone de destination

Une zone d'atterrissage est un AWS environnement multi-comptes bien conçu, évolutif et sécurisé. Il s'agit d'un point de départ à partir duquel vos entreprises peuvent rapidement lancer et déployer des charges de travail et des applications en toute confiance dans leur environnement de sécurité et d'infrastructure. Pour plus d'informations sur les zones de destination, veuillez consulter [Setting up a secure and scalable multi-account AWS environment](#).

grand modèle de langage (LLM)

Un modèle d'[intelligence artificielle basé](#) sur le deep learning qui est préentraîné sur une grande quantité de données. Un LLM peut effectuer plusieurs tâches, telles que répondre à des questions, résumer des documents, traduire du texte dans d'autres langues et compléter des phrases. Pour plus d'informations, voir [Que sont LLMs](#).

migration de grande envergure

Migration de 300 serveurs ou plus.

LBAC

Voir contrôle d'[accès basé sur des étiquettes](#).

principe de moindre privilège

Bonne pratique de sécurité qui consiste à accorder les autorisations minimales nécessaires à l'exécution d'une tâche. Pour plus d'informations, veuillez consulter la rubrique [Accorder les autorisations de moindre privilège](#) dans la documentation IAM.

lift and shift

Voir [7 Rs](#).

système de poids faible

Système qui stocke d'abord l'octet le moins significatif. Voir aussi [endianité](#).

LLM

Voir le [grand modèle de langage](#).

environnements inférieurs

Voir [environnement](#).

M

machine learning (ML)

Type d'intelligence artificielle qui utilise des algorithmes et des techniques pour la reconnaissance et l'apprentissage de modèles. Le ML analyse et apprend à partir de données enregistrées, telles que les données de l'Internet des objets (IoT), pour générer un modèle statistique basé sur des modèles. Pour plus d'informations, veuillez consulter [Machine Learning](#).

branche principale

Voir [succursale](#).

malware

Logiciel conçu pour compromettre la sécurité ou la confidentialité de l'ordinateur. Les logiciels malveillants peuvent perturber les systèmes informatiques, divulguer des informations sensibles ou obtenir un accès non autorisé. Parmi les malwares, on peut citer les virus, les vers, les rançongiciels, les chevaux de Troie, les logiciels espions et les enregistreurs de frappe.

services gérés

Services AWS pour lequel AWS fonctionnent la couche d'infrastructure, le système d'exploitation et les plateformes, et vous accédez aux points de terminaison pour stocker et récupérer des données. Amazon Simple Storage Service (Amazon S3) et Amazon DynamoDB sont des exemples de services gérés. Ils sont également connus sous le nom de services abstraits.

système d'exécution de la fabrication (MES)

Un système logiciel pour le suivi, la surveillance, la documentation et le contrôle des processus de production qui convertissent les matières premières en produits finis dans l'atelier.

MAP

Voir [Migration Acceleration Program](#).

mécanisme

Processus complet au cours duquel vous créez un outil, favorisez son adoption, puis inspectez les résultats afin de procéder aux ajustements nécessaires. Un mécanisme est un cycle qui se renforce et s'améliore au fur et à mesure de son fonctionnement. Pour plus d'informations, voir [Création de mécanismes](#) dans le cadre AWS Well-Architected.

compte membre

Tous, à l'exception du compte de gestion, qui font partie d'une organisation dans AWS Organizations. Un compte ne peut être membre que d'une seule organisation à la fois.

MAILLES

Voir le [système d'exécution de la fabrication](#).

Transport télémetrique en file d'attente de messages (MQTT)

[Protocole de communication léger machine-to-machine \(M2M\), basé sur le modèle de publication/d'abonnement, pour les appareils IoT aux ressources limitées.](#)

microservice

Un petit service indépendant qui communique via un réseau bien défini APIs et qui est généralement détenu par de petites équipes autonomes. Par exemple, un système d'assurance peut inclure des microservices qui mappent à des capacités métier, telles que les ventes ou le marketing, ou à des sous-domaines, tels que les achats, les réclamations ou l'analytique. Les avantages des microservices incluent l'agilité, la flexibilité de la mise à l'échelle, la facilité de déploiement, la réutilisation du code et la résilience. Pour plus d'informations, consultez la section [Intégration de microservices à l'aide de services AWS sans serveur](#).

architecture de microservices

Approche de création d'une application avec des composants indépendants qui exécutent chaque processus d'application en tant que microservice. Ces microservices communiquent via une interface bien définie en utilisant Lightweight. APIs Chaque microservice de cette architecture peut être mis à jour, déployé et mis à l'échelle pour répondre à la demande de fonctions spécifiques d'une application. Pour plus d'informations, consultez la section [Implémentation de microservices sur AWS](#).

Programme d'accélération des migrations (MAP)

Un AWS programme qui fournit un support de conseil, des formations et des services pour aider les entreprises à établir une base opérationnelle solide pour passer au cloud, et pour aider à compenser le coût initial des migrations. MAP inclut une méthodologie de migration pour exécuter les migrations héritées de manière méthodique, ainsi qu'un ensemble d'outils pour automatiser et accélérer les scénarios de migration courants.

migration à grande échelle

Processus consistant à transférer la majeure partie du portefeuille d'applications vers le cloud par vagues, un plus grand nombre d'applications étant déplacées plus rapidement à chaque vague. Cette phase utilise les bonnes pratiques et les enseignements tirés des phases précédentes pour implémenter une usine de migration d'équipes, d'outils et de processus en vue de rationaliser la migration des charges de travail grâce à l'automatisation et à la livraison agile. Il s'agit de la troisième phase de la [stratégie de migration AWS](#).

usine de migration

Équipes interfonctionnelles qui rationalisent la migration des charges de travail grâce à des approches automatisées et agiles. Les équipes de Migration Factory comprennent généralement des responsables des opérations, des analystes commerciaux et des propriétaires, des ingénieurs de migration, des développeurs et DevOps des professionnels travaillant dans le cadre de sprints. Entre 20 et 50 % du portefeuille d'applications d'entreprise est constitué de modèles répétés qui peuvent être optimisés par une approche d'usine. Pour plus d'informations, veuillez consulter la rubrique [discussion of migration factories](#) et le [guide Cloud Migration Factory](#) dans cet ensemble de contenus.

métadonnées de migration

Informations relatives à l'application et au serveur nécessaires pour finaliser la migration. Chaque modèle de migration nécessite un ensemble de métadonnées de migration différent. Les exemples de métadonnées de migration incluent le sous-réseau cible, le groupe de sécurité et le AWS compte.

modèle de migration

Tâche de migration reproductible qui détaille la stratégie de migration, la destination de la migration et l'application ou le service de migration utilisé. Exemple : réorganisez la migration vers Amazon EC2 avec le service de migration AWS d'applications.

Évaluation du portefeuille de migration (MPA)

Outil en ligne qui fournit des informations pour valider l'analyse de rentabilisation en faveur de la migration vers le. AWS Cloud La MPA propose une évaluation détaillée du portefeuille (dimensionnement approprié des serveurs, tarification, comparaison du coût total de possession, analyse des coûts de migration), ainsi que la planification de la migration (analyse et collecte des données d'applications, regroupement des applications, priorisation des migrations et planification des vagues). L'[outil MPA](#) (connexion requise) est disponible gratuitement pour tous les AWS consultants et consultants APN Partner.

Évaluation de la préparation à la migration (MRA)

Processus qui consiste à obtenir des informations sur l'état de préparation d'une organisation au cloud, à identifier les forces et les faiblesses et à élaborer un plan d'action pour combler les lacunes identifiées, à l'aide du AWS CAF. Pour plus d'informations, veuillez consulter le [guide de préparation à la migration](#). La MRA est la première phase de la [stratégie de migration AWS](#).

stratégie de migration

L'approche utilisée pour migrer une charge de travail vers le AWS Cloud. Pour plus d'informations, reportez-vous aux [7 R](#) de ce glossaire et à [Mobiliser votre organisation pour accélérer les migrations à grande échelle](#).

ML

Voir [apprentissage automatique](#).

modernisation

Transformation d'une application obsolète (héritée ou monolithique) et de son infrastructure en un système agile, élastique et hautement disponible dans le cloud afin de réduire les coûts, de gagner en efficacité et de tirer parti des innovations. Pour plus d'informations, consultez [la section Stratégie de modernisation des applications dans le AWS Cloud](#).

évaluation de la préparation à la modernisation

Évaluation qui permet de déterminer si les applications d'une organisation sont prêtes à être modernisées, d'identifier les avantages, les risques et les dépendances, et qui détermine dans quelle mesure l'organisation peut prendre en charge l'état futur de ces applications. Le résultat de l'évaluation est un plan de l'architecture cible, une feuille de route détaillant les phases de développement et les étapes du processus de modernisation, ainsi qu'un plan d'action pour combler les lacunes identifiées. Pour plus d'informations, consultez la section [Évaluation de l'état de préparation à la modernisation des applications dans le AWS Cloud](#).

applications monolithiques (monolithes)

Applications qui s'exécutent en tant que service unique avec des processus étroitement couplés. Les applications monolithiques ont plusieurs inconvénients. Si une fonctionnalité de l'application connaît un pic de demande, l'architecture entière doit être mise à l'échelle. L'ajout ou l'amélioration des fonctionnalités d'une application monolithique devient également plus complexe lorsque la base de code s'élargit. Pour résoudre ces problèmes, vous pouvez utiliser une architecture de microservices. Pour plus d'informations, veuillez consulter [Decomposing monoliths into microservices](#).

MPA

Voir [Évaluation du portefeuille de migration](#).

MQTT

Voir [Message Queuing Telemetry](#) Transport.

classification multi-classes

Processus qui permet de générer des prédictions pour plusieurs classes (prédiction d'un résultat parmi plus de deux). Par exemple, un modèle de ML peut demander « Ce produit est-il un livre, une voiture ou un téléphone ? » ou « Quelle catégorie de produits intéresse le plus ce client ? ».

infrastructure mutable

Modèle qui met à jour et modifie l'infrastructure existante pour les charges de travail de production. Pour améliorer la cohérence, la fiabilité et la prévisibilité, le AWS Well-Architected Framework recommande l'utilisation [d'une infrastructure immuable comme](#) meilleure pratique.

O

OAC

Voir [Contrôle d'accès à l'origine](#).

OAI

Voir [l'identité d'accès à l'origine](#).

OCM

Voir [gestion du changement organisationnel](#).

migration hors ligne

Méthode de migration dans laquelle la charge de travail source est supprimée au cours du processus de migration. Cette méthode implique un temps d'arrêt prolongé et est généralement utilisée pour de petites charges de travail non critiques.

OI

Voir [Intégration des opérations](#).

OLA

Voir l'accord [au niveau opérationnel](#).

migration en ligne

Méthode de migration dans laquelle la charge de travail source est copiée sur le système cible sans être mise hors ligne. Les applications connectées à la charge de travail peuvent continuer à fonctionner pendant la migration. Cette méthode implique un temps d'arrêt nul ou minimal et est généralement utilisée pour les charges de travail de production critiques.

OPC-UA

Voir [Open Process Communications - Architecture unifiée](#).

Communications par processus ouvert - Architecture unifiée (OPC-UA)

Un protocole de communication machine-to-machine (M2M) pour l'automatisation industrielle. L'OPC-UA fournit une norme d'interopérabilité avec des schémas de cryptage, d'authentification et d'autorisation des données.

accord au niveau opérationnel (OLA)

Accord qui précise ce que les groupes informatiques fonctionnels s'engagent à fournir les uns aux autres, afin de prendre en charge un contrat de niveau de service (SLA).

examen de l'état de préparation opérationnelle (ORR)

Une liste de questions et de bonnes pratiques associées qui vous aident à comprendre, évaluer, prévenir ou réduire l'ampleur des incidents et des défaillances possibles. Pour plus d'informations, voir [Operational Readiness Reviews \(ORR\)](#) dans le AWS Well-Architected Framework.

technologie opérationnelle (OT)

Systèmes matériels et logiciels qui fonctionnent avec l'environnement physique pour contrôler les opérations, les équipements et les infrastructures industriels. Dans le secteur manufacturier, l'intégration des systèmes OT et des technologies de l'information (IT) est au cœur des transformations de [l'industrie 4.0](#).

intégration des opérations (OI)

Processus de modernisation des opérations dans le cloud, qui implique la planification de la préparation, l'automatisation et l'intégration. Pour en savoir plus, veuillez consulter le [guide d'intégration des opérations](#).

journal de suivi d'organisation

Un parcours créé par AWS CloudTrail qui enregistre tous les événements pour tous les membres Comptes AWS d'une organisation dans AWS Organizations. Ce journal de suivi est créé dans

chaque Compte AWS qui fait partie de l'organisation et suit l'activité de chaque compte. Pour plus d'informations, consultez [la section Création d'un suivi pour une organisation](#) dans la CloudTrail documentation.

gestion du changement organisationnel (OCM)

Cadre pour gérer les transformations métier majeures et perturbatrices du point de vue des personnes, de la culture et du leadership. L'OCM aide les organisations à se préparer et à effectuer la transition vers de nouveaux systèmes et de nouvelles politiques en accélérant l'adoption des changements, en abordant les problèmes de transition et en favorisant des changements culturels et organisationnels. Dans la stratégie de AWS migration, ce cadre est appelé accélération du personnel, en raison de la rapidité du changement requise dans les projets d'adoption du cloud. Pour plus d'informations, veuillez consulter le [guide OCM](#).

contrôle d'accès d'origine (OAC)

Dans CloudFront, une option améliorée pour restreindre l'accès afin de sécuriser votre contenu Amazon Simple Storage Service (Amazon S3). L'OAC prend en charge tous les compartiments S3 dans leur ensemble Régions AWS, le chiffrement côté serveur avec AWS KMS (SSE-KMS) et les requêtes dynamiques PUT adressées au compartiment S3. DELETE

identité d'accès d'origine (OAI)

Dans CloudFront, une option permettant de restreindre l'accès afin de sécuriser votre contenu Amazon S3. Lorsque vous utilisez OAI, il CloudFront crée un principal auprès duquel Amazon S3 peut s'authentifier. Les principaux authentifiés ne peuvent accéder au contenu d'un compartiment S3 que par le biais d'une distribution spécifique CloudFront . Voir également [OAC](#), qui fournit un contrôle d'accès plus précis et amélioré.

ORR

Voir l'[examen de l'état de préparation opérationnelle](#).

DE

Voir [technologie opérationnelle](#).

VPC sortant (de sortie)

Dans une architecture AWS multi-comptes, un VPC qui gère les connexions réseau initiées depuis une application. L'[architecture AWS de référence de sécurité](#) recommande de configurer votre compte réseau avec les fonctions entrantes, sortantes et d'inspection VPCs afin de protéger l'interface bidirectionnelle entre votre application et l'Internet en général.

P

limite des autorisations

Politique de gestion IAM attachée aux principaux IAM pour définir les autorisations maximales que peut avoir l'utilisateur ou le rôle. Pour plus d'informations, veuillez consulter la rubrique [Limites des autorisations](#) dans la documentation IAM.

informations personnelles identifiables (PII)

Informations qui, lorsqu'elles sont consultées directement ou associées à d'autres données connexes, peuvent être utilisées pour déduire raisonnablement l'identité d'une personne. Les exemples d'informations personnelles incluent les noms, les adresses et les informations de contact.

PII

Voir les [informations personnelles identifiables](#).

manuel stratégique

Ensemble d'étapes prédéfinies qui capturent le travail associé aux migrations, comme la fourniture de fonctions d'opérations de base dans le cloud. Un manuel stratégique peut revêtir la forme de scripts, de runbooks automatisés ou d'un résumé des processus ou des étapes nécessaires au fonctionnement de votre environnement modernisé.

PLC

Voir [contrôleur logique programmable](#).

PLM

Consultez la section [Gestion du cycle de vie des](#) produits.

politique

Objet capable de définir les autorisations (voir la [politique basée sur l'identité](#)), de spécifier les conditions d'accès (voir la [politique basée sur les ressources](#)) ou de définir les autorisations maximales pour tous les comptes d'une organisation dans AWS Organizations (voir la politique de contrôle des [services](#)).

persistance polyglotte

Choix indépendant de la technologie de stockage de données d'un microservice en fonction des modèles d'accès aux données et d'autres exigences. Si vos microservices utilisent la même

technologie de stockage de données, ils peuvent rencontrer des difficultés d'implémentation ou présenter des performances médiocres. Les microservices sont plus faciles à mettre en œuvre, atteignent de meilleures performances, ainsi qu'une meilleure capacité de mise à l'échelle s'ils utilisent l'entrepôt de données le mieux adapté à leurs besoins. Pour plus d'informations, veuillez consulter [Enabling data persistence in microservices](#).

évaluation du portefeuille

Processus de découverte, d'analyse et de priorisation du portefeuille d'applications afin de planifier la migration. Pour plus d'informations, veuillez consulter [Evaluating migration readiness](#).

predicate

Une condition de requête qui renvoie `true` ou `false`, généralement située dans une `WHERE` clause.

prédicat pushdown

Technique d'optimisation des requêtes de base de données qui filtre les données de la requête avant le transfert. Cela réduit la quantité de données qui doivent être extraites et traitées à partir de la base de données relationnelle et améliore les performances des requêtes.

contrôle préventif

Contrôle de sécurité conçu pour empêcher qu'un événement ne se produise. Ces contrôles constituent une première ligne de défense pour empêcher tout accès non autorisé ou toute modification indésirable de votre réseau. Pour plus d'informations, veuillez consulter [Preventative controls](#) dans *Implementing security controls on AWS*.

principal

Entité capable d'effectuer AWS des actions et d'accéder à des ressources. Cette entité est généralement un utilisateur root pour un Compte AWS rôle IAM ou un utilisateur. Pour plus d'informations, veuillez consulter la rubrique Principal dans [Termes et concepts relatifs aux rôles](#), dans la documentation IAM.

confidentialité dès la conception

Une approche d'ingénierie système qui prend en compte la confidentialité tout au long du processus de développement.

zones hébergées privées

Conteneur contenant des informations sur la manière dont vous souhaitez qu'Amazon Route 53 réponde aux requêtes DNS pour un domaine et ses sous-domaines au sein d'un ou de plusieurs

VPCs domaines. Pour plus d'informations, veuillez consulter [Working with private hosted zones](#) dans la documentation Route 53.

contrôle proactif

[Contrôle de sécurité](#) conçu pour empêcher le déploiement de ressources non conformes. Ces contrôles analysent les ressources avant qu'elles ne soient provisionnées. Si la ressource n'est pas conforme au contrôle, elle n'est pas provisionnée. Pour plus d'informations, consultez le [guide de référence sur les contrôles](#) dans la AWS Control Tower documentation et consultez la section [Contrôles proactifs dans Implémentation](#) des contrôles de sécurité sur AWS.

gestion du cycle de vie des produits (PLM)

Gestion des données et des processus d'un produit tout au long de son cycle de vie, depuis la conception, le développement et le lancement, en passant par la croissance et la maturité, jusqu'au déclin et au retrait.

environnement de production

Voir [environnement](#).

contrôleur logique programmable (PLC)

Dans le secteur manufacturier, un ordinateur hautement fiable et adaptable qui surveille les machines et automatise les processus de fabrication.

chaînage rapide

Utiliser le résultat d'une invite [LLM](#) comme entrée pour l'invite suivante afin de générer de meilleures réponses. Cette technique est utilisée pour décomposer une tâche complexe en sous-tâches ou pour affiner ou développer de manière itérative une réponse préliminaire. Cela permet d'améliorer la précision et la pertinence des réponses d'un modèle et permet d'obtenir des résultats plus précis et personnalisés.

pseudonymisation

Processus de remplacement des identifiants personnels dans un ensemble de données par des valeurs fictives. La pseudonymisation peut contribuer à protéger la vie privée. Les données pseudonymisées sont toujours considérées comme des données personnelles.

publish/subscribe (pub/sub)

Modèle qui permet des communications asynchrones entre les microservices afin d'améliorer l'évolutivité et la réactivité. Par exemple, dans un [MES](#) basé sur des microservices, un microservice peut publier des messages d'événements sur un canal auquel d'autres microservices

peuvent s'abonner. Le système peut ajouter de nouveaux microservices sans modifier le service de publication.

Q

plan de requête

Série d'étapes, telles que des instructions, utilisées pour accéder aux données d'un système de base de données relationnelle SQL.

régression du plan de requêtes

Le cas où un optimiseur de service de base de données choisit un plan moins optimal qu'avant une modification donnée de l'environnement de base de données. Cela peut être dû à des changements en termes de statistiques, de contraintes, de paramètres d'environnement, de liaisons de paramètres de requêtes et de mises à jour du moteur de base de données.

R

Matrice RACI

Voir [responsable, responsable, consulté, informé \(RACI\)](#).

CHIFFON

Voir [Retrieval Augmented Generation](#).

rançongiciel

Logiciel malveillant conçu pour bloquer l'accès à un système informatique ou à des données jusqu'à ce qu'un paiement soit effectué.

Matrice RASCI

Voir [responsable, responsable, consulté, informé \(RACI\)](#).

RCAC

Voir [contrôle d'accès aux lignes et aux colonnes](#).

réplica en lecture

Copie d'une base de données utilisée en lecture seule. Vous pouvez acheminer les requêtes vers le réplica de lecture pour réduire la charge sur votre base de données principale.

réarchitecte

Voir [7 Rs](#).

objectif de point de récupération (RPO)

Durée maximale acceptable depuis le dernier point de récupération des données. Il détermine ce qui est considéré comme étant une perte de données acceptable entre le dernier point de reprise et l'interruption du service.

objectif de temps de récupération (RTO)

Le délai maximum acceptable entre l'interruption du service et le rétablissement du service.

refactoriser

Voir [7 Rs](#).

Région

Un ensemble de AWS ressources dans une zone géographique. Chacun Région AWS est isolé et indépendant des autres pour garantir tolérance aux pannes, stabilité et résilience. Pour plus d'informations, voir [Spécifier ce que Régions AWS votre compte peut utiliser](#).

régression

Technique de ML qui prédit une valeur numérique. Par exemple, pour résoudre le problème « Quel sera le prix de vente de cette maison ? », un modèle de ML pourrait utiliser un modèle de régression linéaire pour prédire le prix de vente d'une maison sur la base de faits connus à son sujet (par exemple, la superficie en mètres carrés).

réhéberger

Voir [7 Rs](#).

version

Dans un processus de déploiement, action visant à promouvoir les modifications apportées à un environnement de production.

déplacer

Voir [7 Rs](#).

replateforme

Voir [7 Rs](#).

rachat

Voir [7 Rs](#).

résilience

La capacité d'une application à résister aux perturbations ou à s'en remettre. [La haute disponibilité et la reprise après sinistre](#) sont des considérations courantes lors de la planification de la résilience dans le AWS Cloud. Pour plus d'informations, consultez la section [AWS Cloud Résilience](#).

politique basée sur les ressources

Politique attachée à une ressource, comme un compartiment Amazon S3, un point de terminaison ou une clé de chiffrement. Ce type de politique précise les principaux auxquels l'accès est autorisé, les actions prises en charge et toutes les autres conditions qui doivent être remplies.

matrice responsable, redevable, consulté et informé (RACI)

Une matrice qui définit les rôles et les responsabilités de toutes les parties impliquées dans les activités de migration et les opérations cloud. Le nom de la matrice est dérivé des types de responsabilité définis dans la matrice : responsable (R), responsable (A), consulté (C) et informé (I). Le type de support (S) est facultatif. Si vous incluez le support, la matrice est appelée matrice RASCI, et si vous l'excluez, elle est appelée matrice RACI.

contrôle réactif

Contrôle de sécurité conçu pour permettre de remédier aux événements indésirables ou aux écarts par rapport à votre référence de sécurité. Pour plus d'informations, veuillez consulter la rubrique [Responsive controls](#) dans Implementing security controls on AWS.

retain

Voir [7 Rs](#).

se retirer

Voir [7 Rs](#).

Génération augmentée de récupération (RAG)

Technologie d'[IA générative](#) dans laquelle un [LLM](#) fait référence à une source de données faisant autorité qui se trouve en dehors de ses sources de données de formation avant de générer une

réponse. Par exemple, un modèle RAG peut effectuer une recherche sémantique dans la base de connaissances ou dans les données personnalisées d'une organisation. Pour plus d'informations, voir [Qu'est-ce que RAG ?](#)

rotation

Processus de mise à jour périodique d'un [secret](#) pour empêcher un attaquant d'accéder aux informations d'identification.

contrôle d'accès aux lignes et aux colonnes (RCAC)

Utilisation d'expressions SQL simples et flexibles dotées de règles d'accès définies. Le RCAC comprend des autorisations de ligne et des masques de colonnes.

RPO

Voir l'[objectif du point de récupération](#).

RTO

Voir l'[objectif relatif au temps de rétablissement](#).

runbook

Ensemble de procédures manuelles ou automatisées nécessaires à l'exécution d'une tâche spécifique. Elles visent généralement à rationaliser les opérations ou les procédures répétitives présentant des taux d'erreur élevés.

S

SAML 2.0

Un standard ouvert utilisé par de nombreux fournisseurs d'identité (IdPs). Cette fonctionnalité permet l'authentification unique fédérée (SSO), afin que les utilisateurs puissent se connecter AWS Management Console ou appeler les opérations de l' AWS API sans que vous ayez à créer un utilisateur dans IAM pour tous les membres de votre organisation. Pour plus d'informations sur la fédération SAML 2.0, veuillez consulter [À propos de la fédération SAML 2.0](#) dans la documentation IAM.

SCADA

Voir [Contrôle de supervision et acquisition de données](#).

SCP

Voir la [politique de contrôle des services](#).

secret

Dans AWS Secrets Manager des informations confidentielles ou restreintes, telles qu'un mot de passe ou des informations d'identification utilisateur, que vous stockez sous forme cryptée. Il comprend la valeur secrète et ses métadonnées. La valeur secrète peut être binaire, une chaîne unique ou plusieurs chaînes. Pour plus d'informations, voir [Que contient le secret d'un Secrets Manager ?](#) dans la documentation de Secrets Manager.

sécurité dès la conception

Une approche d'ingénierie système qui prend en compte la sécurité tout au long du processus de développement.

contrôle de sécurité

Barrière de protection technique ou administrative qui empêche, détecte ou réduit la capacité d'un assaillant d'exploiter une vulnérabilité de sécurité. Il existe quatre principaux types de contrôles de sécurité : [préventifs](#), [détectifs](#), [réactifs](#) et [proactifs](#).

renforcement de la sécurité

Processus qui consiste à réduire la surface d'attaque pour la rendre plus résistante aux attaques. Cela peut inclure des actions telles que la suppression de ressources qui ne sont plus requises, la mise en œuvre des bonnes pratiques de sécurité consistant à accorder le moindre privilège ou la désactivation de fonctionnalités inutiles dans les fichiers de configuration.

système de gestion des informations et des événements de sécurité (SIEM)

Outils et services qui associent les systèmes de gestion des informations de sécurité (SIM) et de gestion des événements de sécurité (SEM). Un système SIEM collecte, surveille et analyse les données provenant de serveurs, de réseaux, d'appareils et d'autres sources afin de détecter les menaces et les failles de sécurité, mais aussi de générer des alertes.

automatisation des réponses de sécurité

Action prédéfinie et programmée conçue pour répondre automatiquement à un événement de sécurité ou y remédier. Ces automatisations servent de contrôles de sécurité [détectifs](#) ou [réactifs](#) qui vous aident à mettre en œuvre les meilleures pratiques AWS de sécurité. Parmi les actions de réponse automatique, citons la modification d'un groupe de sécurité VPC, l'application de correctifs à une EC2 instance Amazon ou la rotation des informations d'identification.

chiffrement côté serveur

Chiffrement des données à destination, par celui Service AWS qui les reçoit.

Politique de contrôle des services (SCP)

Politique qui fournit un contrôle centralisé des autorisations pour tous les comptes d'une organisation dans AWS Organizations. SCPs définissez des garde-fous ou des limites aux actions qu'un administrateur peut déléguer à des utilisateurs ou à des rôles. Vous pouvez les utiliser SCPs comme listes d'autorisation ou de refus pour spécifier les services ou les actions autorisés ou interdits. Pour plus d'informations, consultez la section [Politiques de contrôle des services](#) dans la AWS Organizations documentation.

point de terminaison du service

URL du point d'entrée pour un Service AWS. Pour vous connecter par programmation au service cible, vous pouvez utiliser un point de terminaison. Pour plus d'informations, veuillez consulter la rubrique [Service AWS endpoints](#) dans Références générales AWS.

contrat de niveau de service (SLA)

Accord qui précise ce qu'une équipe informatique promet de fournir à ses clients, comme le temps de disponibilité et les performances des services.

indicateur de niveau de service (SLI)

Mesure d'un aspect des performances d'un service, tel que son taux d'erreur, sa disponibilité ou son débit.

objectif de niveau de service (SLO)

Mesure cible qui représente l'état d'un service, tel que mesuré par un indicateur de [niveau de service](#).

modèle de responsabilité partagée

Un modèle décrivant la responsabilité que vous partagez en matière AWS de sécurité et de conformité dans le cloud. AWS est responsable de la sécurité du cloud, alors que vous êtes responsable de la sécurité dans le cloud. Pour de plus amples informations, veuillez consulter [Modèle de responsabilité partagée](#).

SIEM

Consultez les [informations de sécurité et le système de gestion des événements](#).

point de défaillance unique (SPOF)

Défaillance d'un seul composant critique d'une application susceptible de perturber le système.

SLA

Voir le contrat [de niveau de service](#).

SLI

Voir l'indicateur de [niveau de service](#).

SLO

Voir l'objectif de [niveau de service](#).

split-and-seed modèle

Modèle permettant de mettre à l'échelle et d'accélérer les projets de modernisation. Au fur et à mesure que les nouvelles fonctionnalités et les nouvelles versions de produits sont définies, l'équipe principale se divise pour créer des équipes de produit. Cela permet de mettre à l'échelle les capacités et les services de votre organisation, d'améliorer la productivité des développeurs et de favoriser une innovation rapide. Pour plus d'informations, consultez la section [Approche progressive de la modernisation des applications dans](#) le AWS Cloud

SPOF

Voir [point de défaillance unique](#).

schéma en étoile

Structure organisationnelle de base de données qui utilise une grande table de faits pour stocker les données transactionnelles ou mesurées et utilise une ou plusieurs tables dimensionnelles plus petites pour stocker les attributs des données. Cette structure est conçue pour être utilisée dans un [entrepôt de données](#) ou à des fins de business intelligence.

modèle de figuier étrangleur

Approche de modernisation des systèmes monolithiques en réécrivant et en remplaçant progressivement les fonctionnalités du système jusqu'à ce que le système hérité puisse être mis hors service. Ce modèle utilise l'analogie d'un figuier de vigne qui se développe dans un arbre existant et qui finit par supplanter son hôte. Le schéma a été [présenté par Martin Fowler](#) comme un moyen de gérer les risques lors de la réécriture de systèmes monolithiques. Pour obtenir un

exemple d'application de ce modèle, veuillez consulter [Modernizing legacy Microsoft ASP.NET \(ASMX\) web services incrementally by using containers and Amazon API Gateway](#).

sous-réseau

Plage d'adresses IP dans votre VPC. Un sous-réseau doit se trouver dans une seule zone de disponibilité.

contrôle de supervision et acquisition de données (SCADA)

Dans le secteur manufacturier, un système qui utilise du matériel et des logiciels pour surveiller les actifs physiques et les opérations de production.

chiffrement symétrique

Algorithme de chiffrement qui utilise la même clé pour chiffrer et déchiffrer les données.

tests synthétiques

Tester un système de manière à simuler les interactions des utilisateurs afin de détecter les problèmes potentiels ou de surveiller les performances. Vous pouvez utiliser [Amazon CloudWatch Synthetics](#) pour créer ces tests.

invite du système

Technique permettant de fournir un contexte, des instructions ou des directives à un [LLM](#) afin d'orienter son comportement. Les instructions du système aident à définir le contexte et à établir des règles pour les interactions avec les utilisateurs.

T

balises

Des paires clé-valeur qui agissent comme des métadonnées pour organiser vos AWS ressources. Les balises peuvent vous aider à gérer, identifier, organiser, rechercher et filtrer des ressources. Pour plus d'informations, veuillez consulter la rubrique [Balisage de vos AWS ressources](#).

variable cible

La valeur que vous essayez de prédire dans le cadre du ML supervisé. Elle est également qualifiée de variable de résultat. Par exemple, dans un environnement de fabrication, la variable cible peut être un défaut du produit.

liste de tâches

Outil utilisé pour suivre les progrès dans un runbook. Liste de tâches qui contient une vue d'ensemble du runbook et une liste des tâches générales à effectuer. Pour chaque tâche générale, elle inclut le temps estimé nécessaire, le propriétaire et l'avancement.

environnement de test

Voir [environnement](#).

entraînement

Pour fournir des données à partir desquelles votre modèle de ML peut apprendre. Les données d'entraînement doivent contenir la bonne réponse. L'algorithme d'apprentissage identifie des modèles dans les données d'entraînement, qui mettent en correspondance les attributs des données d'entrée avec la cible (la réponse que vous souhaitez prédire). Il fournit un modèle de ML qui capture ces modèles. Vous pouvez alors utiliser le modèle de ML pour obtenir des prédictions sur de nouvelles données pour lesquelles vous ne connaissez pas la cible.

passerelle de transit

Un hub de transit réseau que vous pouvez utiliser pour interconnecter vos réseaux VPCs et ceux sur site. Pour plus d'informations, voir [Qu'est-ce qu'une passerelle de transit](#) dans la AWS Transit Gateway documentation.

flux de travail basé sur jonction

Approche selon laquelle les développeurs génèrent et testent des fonctionnalités localement dans une branche de fonctionnalités, puis fusionnent ces modifications dans la branche principale. La branche principale est ensuite intégrée aux environnements de développement, de préproduction et de production, de manière séquentielle.

accès sécurisé

Accorder des autorisations à un service que vous spécifiez pour effectuer des tâches au sein de votre organisation AWS Organizations et dans ses comptes en votre nom. Le service de confiance crée un rôle lié au service dans chaque compte, lorsque ce rôle est nécessaire, pour effectuer des tâches de gestion à votre place. Pour plus d'informations, consultez la section [Utilisation AWS Organizations avec d'autres AWS services](#) dans la AWS Organizations documentation.

réglage

Pour modifier certains aspects de votre processus d'entraînement afin d'améliorer la précision du modèle de ML. Par exemple, vous pouvez entraîner le modèle de ML en générant un ensemble d'étiquetage, en ajoutant des étiquettes, puis en répétant ces étapes plusieurs fois avec différents paramètres pour optimiser le modèle.

équipe de deux pizzas

Une petite DevOps équipe que vous pouvez nourrir avec deux pizzas. Une équipe de deux pizzas garantit les meilleures opportunités de collaboration possible dans le développement de logiciels.

U

incertitude

Un concept qui fait référence à des informations imprécises, incomplètes ou inconnues susceptibles de compromettre la fiabilité des modèles de ML prédictifs. Il existe deux types d'incertitude : l'incertitude épistémique est causée par des données limitées et incomplètes, alors que l'incertitude aléatoire est causée par le bruit et le caractère aléatoire inhérents aux données. Pour plus d'informations, veuillez consulter le guide [Quantifying uncertainty in deep learning systems](#).

tâches indifférenciées

Également connu sous le nom de « levage de charges lourdes », ce travail est nécessaire pour créer et exploiter une application, mais qui n'apporte pas de valeur directe à l'utilisateur final ni d'avantage concurrentiel. Les exemples de tâches indifférenciées incluent l'approvisionnement, la maintenance et la planification des capacités.

environnements supérieurs

Voir [environnement](#).

V

mise à vide

Opération de maintenance de base de données qui implique un nettoyage après des mises à jour incrémentielles afin de récupérer de l'espace de stockage et d'améliorer les performances.

contrôle de version

Processus et outils permettant de suivre les modifications, telles que les modifications apportées au code source dans un référentiel.

Appairage de VPC

Une connexion entre deux VPCs qui vous permet d'acheminer le trafic en utilisant des adresses IP privées. Pour plus d'informations, veuillez consulter la rubrique [Qu'est-ce que l'appairage de VPC ?](#) dans la documentation Amazon VPC.

vulnérabilités

Défaut logiciel ou matériel qui compromet la sécurité du système.

W

cache actif

Cache tampon qui contient les données actuelles et pertinentes fréquemment consultées. L'instance de base de données peut lire à partir du cache tampon, ce qui est plus rapide que la lecture à partir de la mémoire principale ou du disque.

données chaudes

Données rarement consultées. Lorsque vous interrogez ce type de données, des requêtes modérément lentes sont généralement acceptables.

fonction de fenêtre

Fonction SQL qui effectue un calcul sur un groupe de lignes liées d'une manière ou d'une autre à l'enregistrement en cours. Les fonctions de fenêtre sont utiles pour traiter des tâches, telles que le calcul d'une moyenne mobile ou l'accès à la valeur des lignes en fonction de la position relative de la ligne en cours.

charge de travail

Ensemble de ressources et de code qui fournit une valeur métier, par exemple une application destinée au client ou un processus de backend.

flux de travail

Groupes fonctionnels d'un projet de migration chargés d'un ensemble de tâches spécifique. Chaque flux de travail est indépendant, mais prend en charge les autres flux de travail du projet.

Par exemple, le flux de travail du portefeuille est chargé de prioriser les applications, de planifier les vagues et de collecter les métadonnées de migration. Le flux de travail du portefeuille fournit ces actifs au flux de travail de migration, qui migre ensuite les serveurs et les applications.

VER

Voir [écrire une fois, lire plusieurs](#).

WQF

Voir le [cadre AWS de qualification de la charge](#) de travail.

écrire une fois, lire plusieurs (WORM)

Modèle de stockage qui écrit les données une seule fois et empêche leur suppression ou leur modification. Les utilisateurs autorisés peuvent lire les données autant de fois que nécessaire, mais ils ne peuvent pas les modifier. Cette infrastructure de stockage de données est considérée comme [immuable](#).

Z

exploit Zero-Day

Une attaque, généralement un logiciel malveillant, qui tire parti d'une [vulnérabilité de type « jour zéro »](#).

vulnérabilité « jour zéro »

Une faille ou une vulnérabilité non atténuée dans un système de production. Les acteurs malveillants peuvent utiliser ce type de vulnérabilité pour attaquer le système. Les développeurs prennent souvent conscience de la vulnérabilité à la suite de l'attaque.

invite Zero-Shot

Fournir à un [LLM](#) des instructions pour effectuer une tâche, mais aucun exemple (plans) pouvant aider à la guider. Le LLM doit utiliser ses connaissances pré-entraînées pour gérer la tâche. L'efficacité de l'invite zéro dépend de la complexité de la tâche et de la qualité de l'invite. Voir également les instructions [en quelques clics](#).

application zombie

Application dont l'utilisation moyenne du processeur et de la mémoire est inférieure à 5 %. Dans un projet de migration, il est courant de retirer ces applications.

Les traductions sont fournies par des outils de traduction automatique. En cas de conflit entre le contenu d'une traduction et celui de la version originale en anglais, la version anglaise prévaudra.