Outils de surveillance et d'alerte et meilleures pratiques pour Amazon RDS for MySQL et MariaDB

AWS Directives prescriptives

AWS Directives prescriptives: Outils de surveillance et d'alerte et meilleures pratiques pour Amazon RDS for MySQL et MariaDB

Table of Contents

Outils de surveillance et d'alerte et meilleures pratiques pour Amazon RDS for MySQL et MariaDB

Igor Obradovic, Amazon Web Services ()AWS

Mars 2025 (historique du document)

La surveillance des bases de données est le processus de mesure, de suivi et d'évaluation de la disponibilité, des performances et des fonctionnalités d'une base de données. Les solutions de surveillance et d'alerte aident les entreprises à garantir que leurs services de base de données, et donc leurs applications et charges de travail associées, sont sécurisés, performants, résilients et efficaces. Vous pouvez également collecter et analyser les journaux, les indicateurs, les événements et les traces de votre charge de travail afin de comprendre l'état de votre charge de travail et d'obtenir des informations sur les opérations au fil du temps. AWS

Vous pouvez surveiller vos ressources pour vous assurer qu'elles fonctionnent comme prévu, ainsi que pour détecter et résoudre les problèmes avant qu'ils n'affectent vos clients. Vous devez utiliser les métriques, les journaux, les événements et les traces que vous surveillez pour déclencher des alarmes lorsque les seuils sont dépassés.

Ce guide décrit les outils d'observabilité et de surveillance des bases de données ainsi que les meilleures pratiques pour les bases de données Amazon Relational Database Service (Amazon RDS). Le guide se concentre sur les bases de données MySQL et MariaDB, bien que la plupart des informations s'appliquent également aux autres moteurs de base de données Amazon RDS.

Ce guide s'adresse aux architectes de solutions, aux architectes de bases de données DBAs, aux DevOps ingénieurs seniors et aux autres membres de l'équipe qui participent à la conception, à la mise en œuvre et à la gestion de solutions de surveillance et d'observabilité pour leurs charges de travail de base de données exécutées dans le. AWS Cloud

Table des matières

Présentation

La surveillance et les alertes font partie des quatre piliers du AWS Well-Architected Framework.

- Le pilier de l'excellence opérationnelle prescrit que votre charge de travail doit être conçue pour inclure la télémétrie et la surveillance. AWS des services tels qu'Amazon Relational Database Service (Amazon RDS) fournissent les informations nécessaires pour comprendre l'état interne de votre charge de travail (par exemple, les métriques, les journaux, les événements et les traces). Lorsque vous exploitez vos bases de données Amazon RDS, vous devez comprendre l'état de santé de vos instances de bases de données, détecter les événements opérationnels et être en mesure de répondre aux événements planifiés et imprévus. AWS fournit des outils de surveillance qui vous aident à déterminer quand les résultats de l'organisation et de l'entreprise sont menacés, ou pourraient l'être, afin que vous puissiez prendre les mesures appropriées au bon moment.
- Le pilier de l'efficacité des performances prescrit que vous devez surveiller les performances de vos ressources, telles que les instances de base de données Amazon RDS, en collectant, en agrégeant et en traitant les indicateurs liés aux performances en temps réel. Vous pouvez identifier la dégradation des performances et corriger les facteurs (par exemple, des requêtes SQL non optimisées ou des paramètres de configuration inadéquats) à l'origine de cette dégradation. Vous pouvez déclencher des alarmes automatiquement lorsque les mesures dépassent les limites attendues. Nous vous recommandons d'utiliser les alarmes non seulement pour les notifications, mais également pour lancer des actions automatisées en réponse aux événements détectés. Vous pouvez évaluer les mesures que vous collectez par rapport à des seuils prédéfinis ou utiliser des algorithmes d'apprentissage automatique pour identifier les comportements anormaux. Par exemple, pour détecter une tendance à l'augmentation de l'utilisation du processeur, vous pouvez collecter et analyser la `cpuUtilization.total` métrique sur une période donnée. Le fait de signaler cette anomalie de manière proactive, avant que l'utilisation du processeur n'atteigne sa limite maximale, peut vous aider à résoudre le problème avant qu'il n'ait un impact sur vos clients.
- Le pilier de fiabilité définit la surveillance et les alertes comme essentielles pour garantir que vous répondez à vos exigences de disponibilité. Votre solution de surveillance doit être capable de détecter efficacement les défaillances. Lorsqu'il détecte des problèmes ou des défaillances, son objectif principal est d'alerter sur ces problèmes. La mise en œuvre de pratiques d'observabilité et de surveillance continues est essentielle pour garantir la résilience des architectures dans le cloud. Pour améliorer vos charges de travail, vous devez être en mesure de les mesurer et de comprendre leur état et leur état de santé. Les principes de conception pour la reprise automatique

en cas de panne, l'évolutivité horizontale et le provisionnement des capacités dépendent de la précision des services de surveillance et d'alerte.

- Le pilier de sécurité traite de la détection et de la prévention des modifications de configuration inattendues ou indésirables, ainsi que des comportements inattendus. Vous pouvez configurer vos instances de base de données Amazon RDS for MySQL et MariaDB avec le plug-in d'audit MariaDB pour enregistrer l'activité de la base de données, telle que les connexions des utilisateurs et les opérations spécifiques exécutées sur la base de données. Le plugin enregistre l'activité de la base de données dans un fichier journal, qui peut être intégré et importé dans les outils de surveillance et d'alerte. Le fichier journal est analysé en temps réel pour détecter tout comportement inattendu ou suspect dans votre base de données. Un tel comportement inattendu ou suspect peut indiquer que votre instance de base de données Amazon RDS a été compromise, ce qui indique des risques potentiels pour votre entreprise. Si l'outil de surveillance détecte un tel événement, il active une alarme pour déclencher une réponse à l'incident de sécurité, ce qui permet de lutter contre les activités suspectes et malveillantes.

Résultats commerciaux ciblés

La mise en œuvre des meilleures pratiques en matière de mécanismes de surveillance et d'alerte vous aide à garantir une infrastructure performante, résiliente, efficace, sécurisée et optimisée en termes de coûts pour vos applications et vos charges de travail. Vous pouvez utiliser des outils d'observabilité qui collectent, stockent et visualisent les métriques, les événements, les traces et les journaux en temps réel pour observer et analyser une vue d'ensemble de l'état et des performances de vos bases de données, et ainsi empêcher la dégradation ou l'interruption des services informatiques associés. Si une dégradation ou une interruption de service imprévue persiste, les outils de surveillance et d'alerte vous aident à détecter rapidement le problème, à l'escalader, à réagir, ainsi qu'à enquêter et à résoudre rapidement le problème. Une solution complète de surveillance et d'alerte pour les charges de travail de vos bases de données cloud vous aide à atteindre les résultats commerciaux suivants :

- Améliorez l'expérience client. Un service fiable améliore l'expérience de vos clients. Les bases de données sont souvent un élément clé des services numériques tels que les applications Web et mobiles, le streaming multimédia, les paiements business-to-business (B2B) APIs et les services d'intégration. Si vous pouvez surveiller et configurer des alertes sur vos bases de données pour détecter rapidement les problèmes, les étudier efficacement et y remédier le plus rapidement possible afin de minimiser les temps d'arrêt et autres perturbations, vous pouvez améliorer la disponibilité, la sécurité et les performances du service numérique pour vos clients.

- Renforcez la confiance des clients.De meilleures performances et une expérience utilisateur plus fluide vous aident à gagner la confiance de vos clients, ce qui peut se traduire par une augmentation du chiffre d'affaires sur votre plateforme. Par exemple, un fournisseur de services de traitement des paiements qui propose un service en ligne fiable peut s'attendre à une confiance et à une fidélité accrues de ses clients, ce qui se traduit par un plus grand nombre de clients et une meilleure fidélisation, une augmentation du nombre de transactions facturables et de nouveaux services innovants générant davantage de revenus.
- Évitez les pertes financières.Toute interruption imprévue de votre infrastructure de base de données peut avoir un impact sur les transactions commerciales que vos clients effectuent en utilisant votre application. Cela peut entraîner des pertes financières importantes dans certains cas. La violation des accords de niveau de service (SLAs) peut entraîner une perte de confiance des clients et, par conséquent, une perte de revenus. Cela peut également devenir une base légale pour des essais coûteux, dans le cadre desquels les clients peuvent exiger une indemnisation en fonction de votre responsabilité et de vos contrats de garantie. Selon une étude réalisée par Atlassian Corporation, une société de logiciels, le coût moyen d'une panne de service se situe entre 140 000 et 540 000 dollars de l'heure, selon le type et la taille de l'entreprise. Un environnement de base de données stable est essentiel pour éviter les interruptions prolongées et les pertes d'activité.
- Augmentez la valeur. Les mécanismes de surveillance et d'alerte peuvent vous aider à concevoir, développer et exploiter un service numérique hautement disponible, résilient, fiable, performant, rentable et sécurisé, mais ce n'est que le début. Vous voudrez que votre entreprise évolue et prospère au fil du temps, améliore les charges de travail cloud existantes et introduise de nouveaux services. Les nouveaux services apportent une valeur ajoutée à vos clients et augmentent les revenus de votre entreprise, ce qui a un effet moteur sur la croissance de votre entreprise.
- Améliorez la productivité des développeurs.Les développeurs productifs et efficaces, qui ne rencontrent pas de problèmes ni de blocages dans leurs tâches de développement, peuvent fournir des produits de haute qualité en moins de temps. Cependant, le génie logiciel et les opérations informatiques sont souvent confrontés à des défis complexes, et cette complexité augmente avec l'ampleur des charges de travail et de leurs architectures. Pour analyser les performances et la cohérence des applications distribuées, les développeurs ont besoin d'outils capables de fournir des mesures et des traces corrélées. Ils permettent d'identifier les artefacts de code et les composants d'infrastructure défectueux le plus rapidement possible et de déterminer les impacts sur les utilisateurs finaux. La bonne suite d'outils de surveillance et d'alerte peut aider les développeurs à coder et à tester de manière plus efficace et plus rapide.

- Améliorez l'efficacité et l'efficience opérationnelles.Lorsque vous gérez des charges de travail dans le cloud à grande échelle, même un faible pourcentage d'amélioration des performances peut se traduire par des économies de plusieurs millions de dollars. En surveillant vos bases de données et en analysant les métriques, les événements, les journaux et les traces, vous pouvez comprendre et prévoir vos futurs besoins en capacité, et tirer parti des économies de coûts disponibles dans le AWS Cloud. Comprendre les charges de travail et l'état de fonctionnement de votre Amazon RDS peut vous aider à réagir aux événements, à résoudre les problèmes et à planifier des améliorations.

Bonnes pratiques d'ordre général

Les meilleures pratiques suivantes vous aident à obtenir une visibilité suffisante sur l'état de votre charge de travail Amazon RDS et à prendre les mesures appropriées en réponse aux événements opérationnels et aux données de surveillance.

- Identifier KPIs. Identifiez les indicateurs de performance clés (KPIs) en fonction des résultats commerciaux souhaités. Évaluez KPIs pour déterminer le succès de la charge de travail. Par exemple, si votre activité principale est le commerce électronique, l'un des résultats commerciaux souhaités pourrait être que votre boutique en ligne soit disponible 24 heures sur 24, 7 jours sur 7 pour que vos clients puissent faire leurs achats. Pour atteindre ce résultat commercial, vous définissez le KPI de disponibilité pour la base de données Amazon RDS principale utilisée par votre application de boutique en ligne, et vous définissez le KPI de base à 99,99 % sur une base hebdomadaire. L'évaluation du KPI de disponibilité réel par rapport à la valeur de référence vous permet de déterminer si vous atteignez la disponibilité de base de données souhaitée de 99,99 % et si vous atteignez ainsi le résultat commercial d'un service 24 heures sur 24, 7 jours sur 7.
- Définissez les métriques de charge de travail. Définissez des métriques de charge de travail pour mesurer les quantités et les qualités de votre charge de travail Amazon RDS. Évaluez les métriques pour déterminer si la charge de travail atteint les résultats souhaités et pour comprendre l'état de la charge de travail. Par exemple, pour évaluer le KPI de disponibilité de votre instance de base de données Amazon RDS, vous devez mesurer des indicateurs tels que le temps de disponibilité et le temps d'arrêt de l'instance de base de données. Vous pouvez ensuite utiliser ces mesures pour calculer le KPI de disponibilité comme suit :

  ```
  availability = uptime / (uptime + downtime)
  ```

  Les métriques représentent des ensembles de points de données ordonnés dans le temps. Les métriques peuvent également inclure des dimensions, qui sont utiles pour la catégorisation et l'analyse.
- Collectez et analysez les métriques de charge de travail. Amazon RDS génère différents indicateurs et journaux, en fonction de votre configuration. Certains d'entre eux représentent des événements, des compteurs ou des statistiques d'instances de base de données tels que`db.Cache.innoDB_buffer_pool_hits`. D'autres indicateurs proviennent du système d'exploitation, par exemple`memory.Total`, qui mesure la quantité totale de mémoire de l'instance hôte Amazon Elastic Compute Cloud (Amazon EC2). L'outil de surveillance doit effectuer une

analyse régulière et proactive des mesures collectées afin d'identifier les tendances et de déterminer si des réponses appropriées sont nécessaires.

- Établissez des bases de référence pour les métriques de charge de travail Établissez des bases de référence pour les métriques afin de définir les valeurs attendues et d'identifier les bons ou les mauvais seuils. Par exemple, vous pouvez définir la base de référence jusqu'`ReadIOPS`à 1 000 dans le cadre des opérations de base de données normales. Vous pouvez ensuite utiliser cette base de référence à des fins de comparaison et pour identifier une surutilisation. Si vos nouveaux indicateurs indiquent régulièrement que les IOPS de lecture se situent entre 2 000 et 3 000, vous avez identifié un écart susceptible de déclencher une action d'investigation, d'intervention et d'amélioration.
- Alerte lorsque les résultats de la charge de travail sont menacés. Lorsque vous déterminez que les résultats commerciaux sont menacés, déclenchez une alerte. Vous pouvez alors soit résoudre les problèmes de manière proactive, avant qu'ils n'affectent vos clients, soit atténuer l'impact de l'incident en temps opportun.
- Identifiez les modèles d'activité attendus pour votre charge de travail. Sur la base de vos indicateurs de référence, établissez des modèles d'activité de la charge de travail afin d'identifier les comportements inattendus et de réagir en prenant les mesures appropriées si nécessaire. AWS fournit des outils de surveillance qui appliquent des algorithmes statistiques et d'apprentissage automatique pour analyser les métriques et détecter les anomalies.
- Alerte lorsque des anomalies de charge de travail sont détectées. Lorsque des anomalies sont détectées dans le fonctionnement des charges de travail Amazon RDS, déclenchez une alerte afin de pouvoir réagir en prenant les mesures appropriées si nécessaire.
- Révision, révision KPIs et indicateurs. Vérifiez que vos bases de données Amazon RDS répondent à vos exigences définies et identifiez les domaines susceptibles d'être améliorés pour atteindre vos objectifs commerciaux. Validez l'efficacité des mesures mesurées et évaluées KPIs, et révisez-les si nécessaire. Supposons, par exemple, que vous définissiez un KPI pour le nombre optimal de connexions simultanées à la base de données et que vous surveilliez les métriques relatives aux tentatives et aux échecs de connexion ainsi que les threads utilisateur créés et en cours d'exécution. Vous avez peut-être plus de connexions à la base de données que celles définies par votre référence de KPI. En analysant vos indicateurs actuels, vous pouvez détecter le résultat, mais il se peut que vous ne puissiez pas en déterminer la cause première. Si tel est le cas, vous devez revoir vos indicateurs et inclure des mesures de surveillance supplémentaires, telles que des compteurs pour les serrures de table. Les nouvelles mesures aideront à déterminer si l'augmentation du nombre de connexions à la base de données est due à des verrouillages de table inattendus.

Outils de surveillance

Nous vous recommandons d'utiliser des outils d'observabilité, de surveillance et d'alerte pour :

- Obtenez des informations sur les performances de votre environnement Amazon RDS
- Détectez les comportements inattendus et suspects
- Planifiez les capacités et prenez des décisions éclairées concernant l'allocation des instances Amazon RDS
- Analysez les métriques et les journaux pour prévoir les problèmes potentiels de manière proactive
- Générez des alertes lorsque les seuils sont dépassés afin de résoudre les problèmes avant que vos utilisateurs ne soient affectés

Vous avez le choix entre différentes options et solutions, notamment des outils et services d'observabilité et de surveillance natifs dans le cloud AWS fournis dans le cloud, des solutions logicielles open source gratuites et des solutions tierces commerciales pour la surveillance des instances de base de données Amazon RDS. Certains de ces outils sont décrits dans les sections qui suivent.

Pour déterminer quel outil répond le mieux à vos besoins, comparez les caractéristiques et les capacités de chaque outil aux exigences de votre organisation. Nous vous recommandons également d'évaluer les outils en termes de facilité de déploiement, de configuration et d'intégration, de mises à jour et de maintenance logicielles, de méthode de déploiement (par exemple, matérielle ou sans serveur), de licences, de prix et de tout autre facteur spécifique à votre organisation.

Sections

Outils inclus dans Amazon RDS

Amazon Relational Database Service (Amazon RDS) est un service de base de données géré dans le. AWS Cloud Amazon RDS étant un service géré, il vous libère de la plupart des tâches de gestion, telles que les sauvegardes de bases de données, les installations de systèmes d'exploitation (OS) et de logiciels de base de données, l'application de correctifs aux systèmes d'exploitation et aux logiciels, la configuration de haute disponibilité, le cycle de vie du matériel et les opérations des centres de données. AWS fournit également un ensemble complet d'outils qui vous permettent de créer une solution d'observabilité complète pour vos instances de base de données Amazon RDS.

Certains outils de surveillance sont inclus, préconfigurés et automatiquement activés dans le service Amazon RDS. Deux outils automatisés sont à votre disposition dès que vous démarrez votre nouvelle instance Amazon RDS :

- L'état de l'instance Amazon RDS fournit des informations sur l'état actuel de votre instance de base de données. Par exemple, les codes d'état incluent Disponible, Arrêté, Création, Sauvegarde et Échec. Vous pouvez utiliser la console Amazon RDS, le AWS Command Line Interface (AWS CLI) ou l'API Amazon RDS pour consulter le statut de l'instance. Pour plus d'informations, consultez la section Affichage du statut d'une instance de base de données Amazon RDS dans la documentation Amazon RDS.
- Les recommandations Amazon RDS fournissent des recommandations automatisées pour les instances de base de données, les répliques de lecture et les groupes de paramètres de base de données. Ces recommandations sont fournies en analysant l'utilisation des instances de base de données, les données de performance et la configuration, et sont fournies à titre indicatif. Par exemple, la recommandation relative à la version obsolète du moteur suggère que vos instances de base de données n'exécutent pas la dernière version du logiciel de base de données et que vous devez mettre à niveau votre instance de base de données pour bénéficier des derniers correctifs de sécurité et autres améliorations. Pour plus d'informations, consultez la section Affichage des recommandations Amazon RDS dans la documentation Amazon RDS.

CloudWatch espaces de noms

Amazon RDS s'intègre à Amazon CloudWatch, qui est un service de surveillance et d'alerte pour les ressources cloud et les applications qui s'y exécutent. AWS Amazon RDS collecte automatiquement les métriques, les fichiers journaux, les traces et les événements relatifs au fonctionnement, à

l'utilisation, aux performances et à l'état des instances de base de données, et les envoie à des CloudWatch fins de stockage, d'analyse et d'alerte à long terme.

Amazon RDS for MySQL et Amazon RDS for MariaDB publient automatiquement un ensemble de métriques par défaut à intervalles d'une minute, sans CloudWatch frais supplémentaires. Ces métriques sont collectées dans deux espaces de noms, qui sont des conteneurs pour les métriques :

- L'espace de noms AWS/RDS inclut des métriques au niveau de l'instance de base de données. Les exemples incluent `BinLogDiskUsage` (la quantité d'espace disque occupée par les journaux binaires), `CPUUtilization` (le pourcentage d'utilisation du processeur), `DatabaseConnections` (le nombre de connexions réseau client à l'instance de base de données), et bien d'autres encore.
- L'espace de noms AWS/Usage inclut des métriques d'utilisation au niveau du compte, qui sont utilisées pour déterminer si vous respectez vos quotas de service Amazon RDS. Les exemples incluent `DBInstances` (le nombre d'instances de base de données dans votre compte ou région AWS), `DBSubnetGroups` (le nombre de groupes de sous-réseaux de base de données dans votre AWS compte ou région) et `ManualSnapshots` (le nombre d'instantanés de base de données créés manuellement dans votre AWS compte ou région).

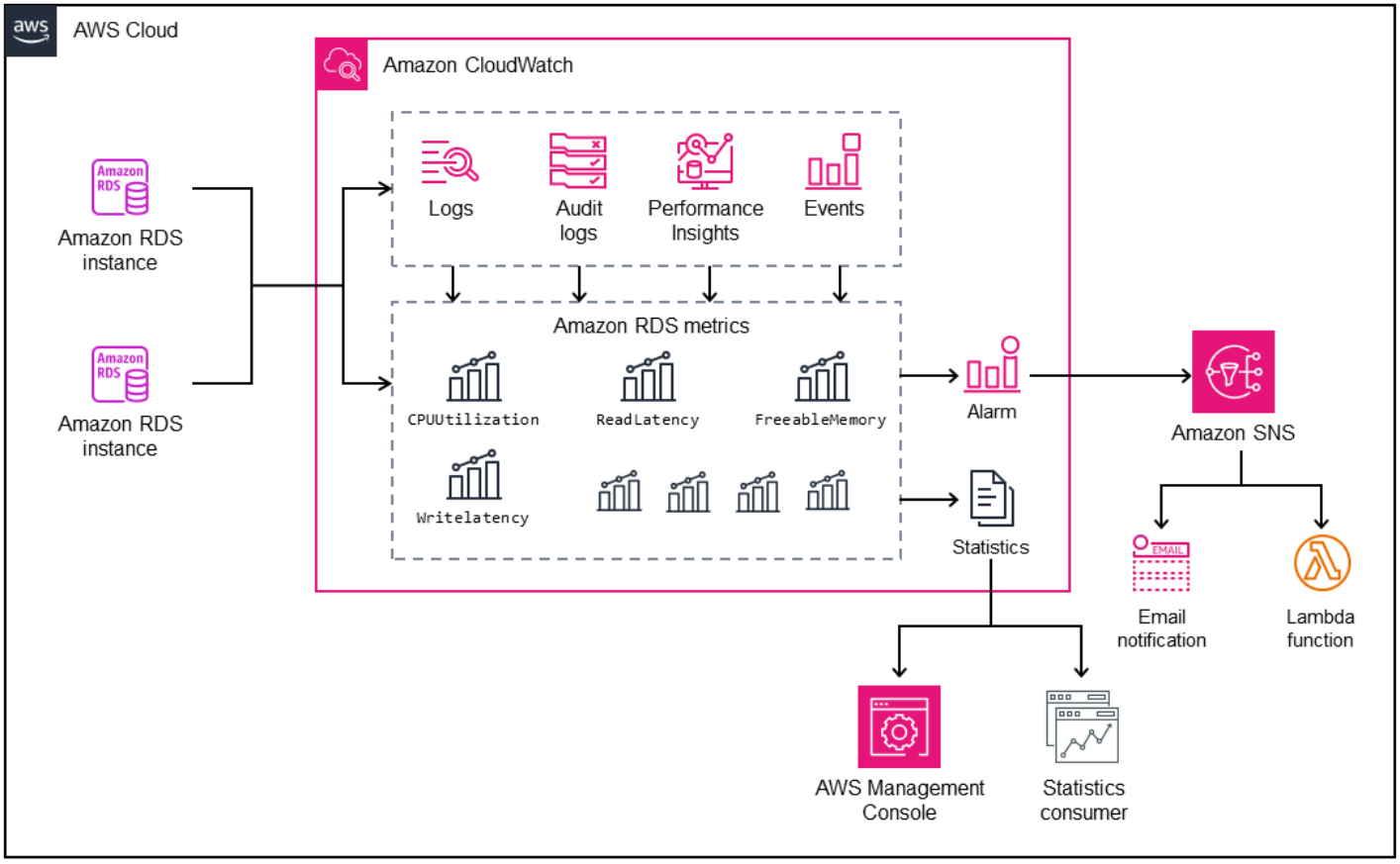
CloudWatch conserve les données métriques comme suit :

- 3 heures : les métriques personnalisées à haute résolution d'une période inférieure à 60 secondes sont conservées pendant 3 heures. Après 3 heures, les points de données sont agrégés en indicateurs de période d'une minute et conservés pendant 15 jours.
- 15 jours : Les points de données d'une durée de 60 secondes (1 minute) sont conservés pendant 15 jours. Après 15 jours, les points de données sont agrégés en métriques de 5 minutes et conservés pendant 63 jours.
- 63 jours : les points de données d'une durée de 300 secondes (5 minutes) sont conservés pendant 63 jours. Après 63 jours, les points de données sont agrégés en métriques sur une période d'une heure et conservés pendant 15 mois.
- 15 mois : les points de données d'une durée de 3 600 secondes (1 heure) sont disponibles pendant 15 mois (455 jours).

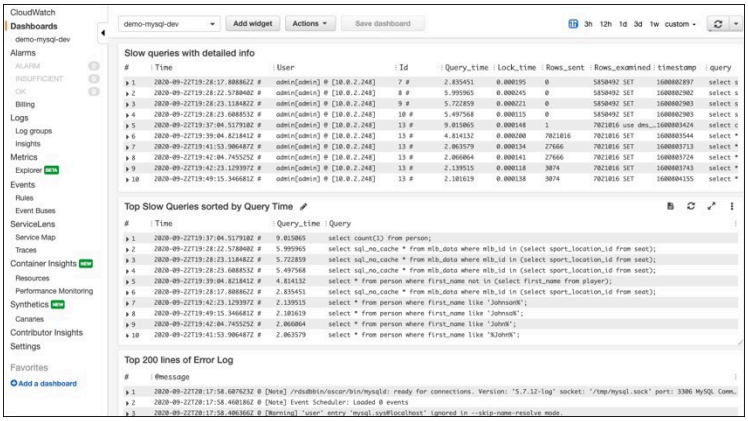
Pour plus d'informations, consultez la section Métriques dans la CloudWatch documentation.

CloudWatch alarmes et tableaux de bord

Vous pouvez utiliser les CloudWatch alarmes Amazon pour surveiller une métrique Amazon RDS spécifique sur une période donnée. Par exemple, vous pouvez surveiller`FreeStorageSpace`, puis effectuer une ou plusieurs actions si la valeur de la métrique dépasse le seuil que vous avez défini. Si vous définissez le seuil à 250 Mo et que l'espace de stockage disponible est de 200 Mo (moins que le seuil), l'alarme sera activée et pourra déclencher une action pour provisionner automatiquement du stockage supplémentaire pour l'instance de base de données Amazon RDS. L'alarme peut également envoyer un SMS de notification au DBA à l'aide d'Amazon Simple Notification Service (Amazon SNS). Le schéma suivant illustre ce processus.

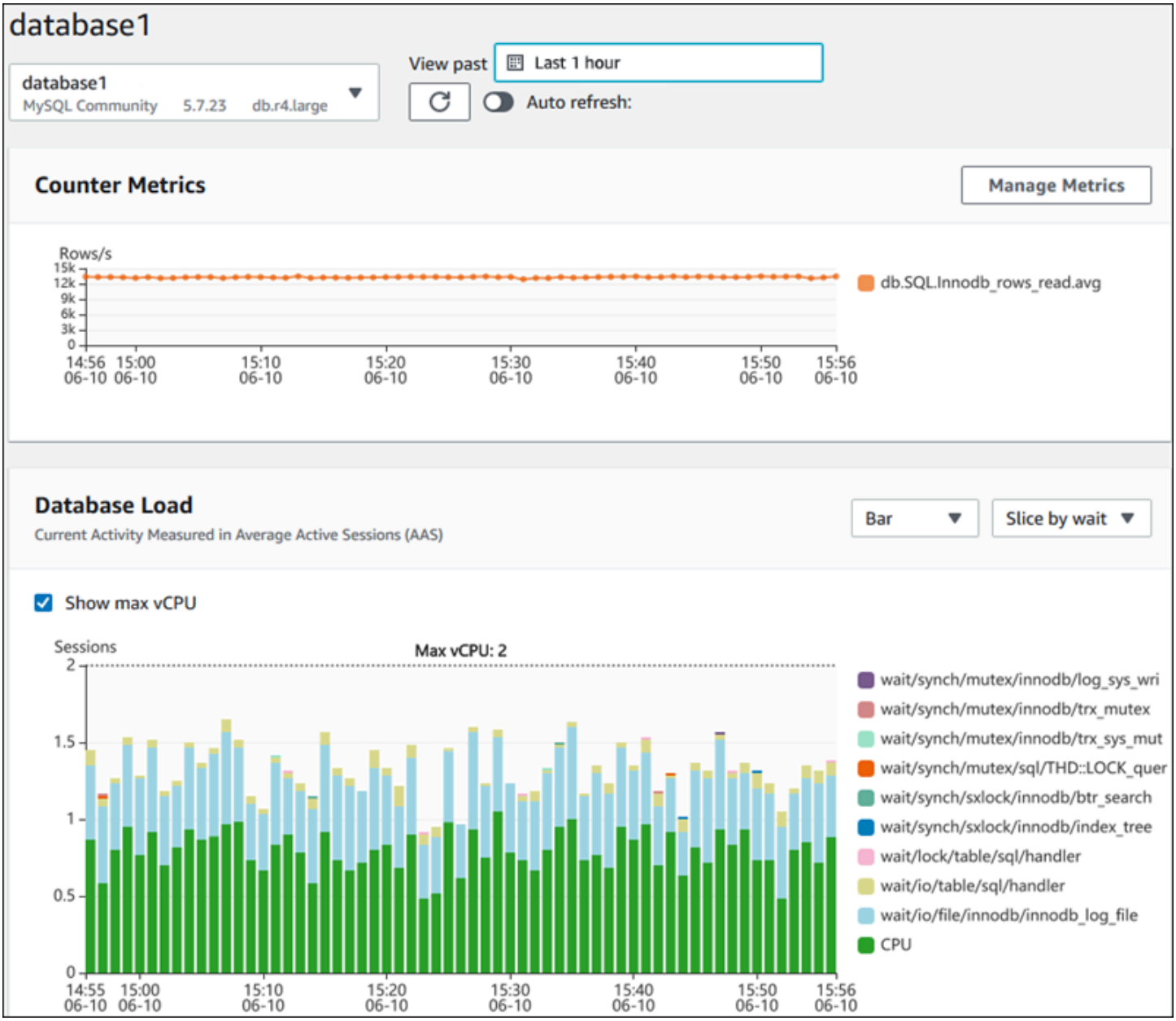


CloudWatch fournit également des tableaux de bord, que vous pouvez utiliser pour créer, personnaliser, interagir avec et enregistrer des vues personnalisées (graphiques) des métriques. Vous pouvez également utiliser CloudWatch Logs Insights pour créer un tableau de bord afin de surveiller le journal des requêtes lentes et le journal des erreurs, et pour recevoir des alertes si un schéma spécifique a été détecté dans ces journaux. L'écran suivant montre un exemple de CloudWatch tableau de bord.



Analyse des performances d’Amazon RDS

Amazon RDS Performance Insights est un outil de réglage et de surveillance des performances des bases de données qui étend les fonctionnalités de surveillance d'Amazon RDS. Il vous aide à analyser les performances de votre base de données en visualisant la charge de l'instance de base de données et en filtrant la charge par temps d'attente, instructions SQL, hôtes ou utilisateurs. L'outil combine plusieurs mesures dans un graphique interactif unique qui vous aide à identifier le type de goulot d'étranglement que peut rencontrer votre instance de base de données, comme les délais d'attente, la consommation élevée du processeur ou la latence des E/S, et à déterminer quelles instructions SQL sont à l'origine de ce goulot d'étranglement. L'écran suivant montre un exemple de visualisation.



Vous devez activer Performance Insights pendant le processus de création de l'instance de base de données afin de collecter des métriques pour les instances de base de données Amazon RDS de votre compte. Le niveau gratuit inclut sept jours d'historique des données de performance et un million de demandes d'API par mois. Vous pouvez éventuellement acheter des périodes de conservation plus longues. Pour des informations complètes sur les prix, consultez la section Performance Insights Pricing (Tarification de Performance Insights).

Pour plus d'informations sur la façon dont vous pouvez utiliser Performance Insights pour surveiller vos instances de base de données, consultez la section sur la surveillance des instances de base de données plus loin dans ce guide.

Performance Insights publie automatiquement les indicateurs sur CloudWatch. Outre l'utilisation de l'outil Performance Insights, vous pouvez tirer parti des fonctionnalités supplémentaires qu'il CloudWatch fournit. Vous pouvez examiner les métriques Performance Insights à l'aide de la CloudWatch console AWS CLI, de l'API ou de l' CloudWatch API. Vous pouvez également ajouter des CloudWatch alarmes, comme pour toute autre métrique. Par exemple, vous souhaiterez peut-être déclencher une notification par SMS DBAs ou prendre des mesures correctives si la `DBLoad` métrique dépasse le seuil que vous avez défini. Vous pouvez également ajouter les métriques Performance Insights à vos CloudWatch tableaux de bord existants.

Surveillance améliorée

Enhanced Monitoring est un outil qui capture les métriques en temps réel pour le système d'exploitation (OS) sur lequel s'exécute votre instance de base de données Amazon RDS. Ces mesures fournissent une granularité allant jusqu'à une seconde pour le processeur, la mémoire, les processus Amazon RDS et OS, le système de fichiers et les données d'E/S du disque, entre autres. Vous pouvez accéder à ces métriques et les analyser dans la console Amazon RDS. Comme pour Performance Insights, les métriques de surveillance améliorées sont fournies par Amazon RDS CloudWatch, où vous pouvez bénéficier de fonctionnalités supplémentaires telles que la conservation à long terme des métriques à des fins d'analyse, la création de filtres de métriques, l'affichage de graphiques sur le CloudWatch tableau de bord et la configuration d'alarmes. Par défaut, la surveillance améliorée est désactivée lorsque vous créez une nouvelle instance de base de données Amazon RDS. Vous pouvez activer cette fonctionnalité lorsque vous créez ou modifiez une instance de base de données. La tarification est basée sur la quantité de données transférées d'Amazon RDS vers CloudWatch Logs, ainsi que sur les taux de stockage. En fonction de la granularité et du nombre d'instances de base de données sur lesquelles la surveillance améliorée est activée, certaines parties des données de surveillance peuvent être incluses dans le niveau gratuit CloudWatch Logs. Pour obtenir des informations complètes sur les prix, consultez la section Amazon CloudWatch Pricing. Pour plus d'informations sur cet outil, consultez la documentation Amazon RDS et la FAQ sur la surveillance améliorée.

AWS Services supplémentaires

AWS fournit plusieurs services de support, qui s'intègrent également à Amazon RDS et CloudWatch qui améliorent encore l'observabilité de vos bases de données. Il s'agit notamment d'Amazon EventBridge, Amazon CloudWatch Logs et AWS CloudTrail.

- Amazon EventBridge est un bus d'événements sans serveur qui peut recevoir, filtrer, transformer, acheminer et diffuser des événements provenant de vos applications et AWS ressources, y compris de vos instances de base de données Amazon RDS. Un événement Amazon RDS indique un changement dans l'environnement Amazon RDS. Par exemple, lorsqu'une instance de base de données change son statut de Disponible à Arrêté, Amazon RDS génère l'événement`RDS-EVENT-0087 / The DB instance has been stopped`. Amazon RDS diffuse des événements dans le cadre d' CloudWatch Events et EventBridge en temps quasi réel. À l'aide de EventBridge and CloudWatch Events, vous pouvez définir des règles pour envoyer des alertes sur des événements Amazon RDS spécifiques présentant un intérêt et automatiser les actions à entreprendre lorsqu'un événement correspond à la règle. Diverses cibles sont disponibles en réponse à un événement, telles qu'une AWS Lambda fonction permettant d'effectuer une action corrective ou une rubrique Amazon SNS permettant d'envoyer un e-mail ou un SMS pour informer DBAs les DevOps ingénieurs de l'événement.
- Amazon CloudWatch Logs est un service qui centralise le stockage des fichiers journaux de toutes vos applications, systèmes et AWS services, y compris les instances de base de données Amazon RDS for MySQL et MariaDB et. AWS CloudTrail Si vous activez cette fonctionnalité pour vos instances de base de données, Amazon RDS publie automatiquement les journaux suivants dans CloudWatch Logs :
  - Journal des erreurs
  - Journal des requêtes lentes
  - Journal général
  - Journal d'audit

  Vous pouvez utiliser CloudWatch Logs Insights pour interroger et analyser les données du journal. Cette fonctionnalité inclut un langage de requête spécialement conçu pour vous aider à rechercher les événements du journal correspondant aux modèles que vous définissez. Par exemple, vous pouvez suivre la corruption des tables dans votre instance de base de données MySQL en surveillant le fichier journal des erreurs pour le modèle suivant : `"ERROR 1034 (HY000): Incorrect key file for table '*'; try to repair it OR Table * is marked as crashed"` Les données de journal filtrées peuvent être converties en CloudWatch métriques. Vous pouvez ensuite utiliser les métriques pour créer des tableaux de bord avec des graphiques ou des données tabulaires, ou définir une alarme si le seuil défini est dépassé. Cela est particulièrement utile lorsque vous utilisez le journal d'audit, car vous pouvez automatiquement surveiller, envoyer des alertes et prendre des mesures correctives si un comportement inattendu ou suspect est détecté. Vous pouvez accéder aux journaux de base de données et les gérer

à l'aide de la console de AWS gestion AWS CLI, de l'API Amazon RDS ou du AWS SDK pour CloudWatch les journaux.

- AWS CloudTrailenregistre et surveille en permanence l'activité des utilisateurs et des API dans votre Compte AWS. Il vous aide à effectuer l'audit, la surveillance de la sécurité et le dépannage opérationnel de vos instances de base de données Amazon RDS for MySQL ou MariaDB. CloudTrail est intégré à Amazon RDS. Toutes les actions peuvent être enregistrées et CloudTrail fournissent un enregistrement des actions entreprises par un utilisateur, un rôle ou un AWS service dans Amazon RDS. Par exemple, lorsqu'un utilisateur crée une nouvelle instance de base de données Amazon RDS, un événement est détecté et le journal inclut des informations sur l'action demandée (`"eventName": "CreateDBInstance"`), la date et l'heure de l'action (`"eventTime": "2022-07-30T22:14:06Z"`), les paramètres de la demande (`"requestParameters": {"dBInstanceIdentifier": "test-instance", "engine": "mysql", "dBInstanceClass": "db.m6g.large"}`), etc. Les événements enregistrés CloudTrail incluent à la fois les appels provenant de la console Amazon RDS et les appels provenant du code utilisant l'API Amazon RDS.

Outils de surveillance tiers

Dans certains scénarios, outre la suite complète d'outils d'observabilité et de surveillance natifs dans le cloud AWS fournie par Amazon RDS, vous souhaiterez peut-être utiliser des outils de surveillance d'autres fournisseurs de logiciels. Ces scénarios incluent les déploiements hybrides, dans lesquels un certain nombre de bases de données peuvent s'exécuter dans votre centre de données local et un autre ensemble de bases de données s'exécutant dans le. AWS Cloud Si vous avez déjà mis en place votre solution d'observabilité d'entreprise, vous souhaiterez peut-être continuer à utiliser vos outils existants et les étendre à vos AWS Cloud déploiements. Le défi lié à la mise en place d'une solution de surveillance tierce réside souvent dans les garanties imposées par Amazon RDS en tant que service géré dans le cloud. Par exemple, vous ne pouvez pas installer le logiciel agent sur le système d'exploitation hôte qui exécute l'instance de base de données, car l'accès à la machine hôte de base de données est refusé. Cependant, vous pouvez intégrer de nombreuses solutions de surveillance tierces à Amazon RDS en vous appuyant sur CloudWatch d'autres AWS Cloud services. Par exemple, les métriques, les journaux, les événements et les traces Amazon RDS peuvent être exportés puis importés dans l'outil de surveillance tiers pour une analyse, une visualisation et des alertes plus poussées. Certaines de ces solutions tierces incluent Prometheus, Grafana et Percona.

Prometheus et Grafana

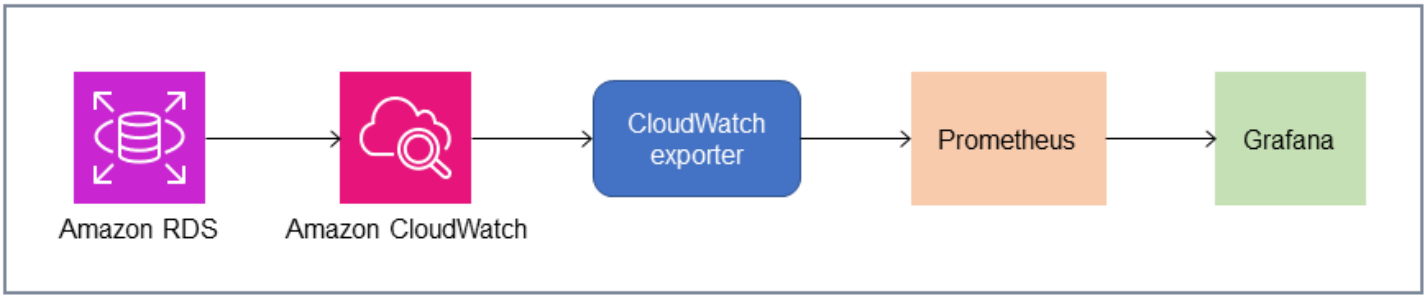
Prometheus est une solution de surveillance open source qui collecte des métriques à partir de cibles configurées à des intervalles donnés. Il s'agit d'une solution de surveillance polyvalente capable de surveiller n'importe quelle application ou service. Lorsque vous surveillez des instances de base de données Amazon RDS, CloudWatch collecte les métriques auprès d'Amazon RDS. Les métriques sont ensuite exportées vers le serveur Prometheus à l'aide d'un exportateur open source tel que l'exportateur YACE ou l'exportateur. CloudWatch

- L'exportateur YACE optimise les tâches d'exportation de données en récupérant plusieurs métriques en une seule demande adressée à l'API. CloudWatch Une fois les métriques stockées sur le serveur Prometheus, celui-ci évalue les expressions des règles et peut générer des alertes lorsque des conditions spécifiées sont respectées.
- CloudWatch L'exportateur est officiellement géré par Prometheus. Il récupère les CloudWatch métriques via l' CloudWatch API et les stocke sur le serveur Prometheus dans un format compatible avec Prometheus, en utilisant des requêtes d'API REST adressées au point de terminaison HTTP.

Lorsque vous choisissez un exportateur, que vous concevez votre modèle de déploiement et que vous configurez les instances d'exportateur, CloudWatchconsidérez les quotas de service et d'API and CloudWatch Logs, car l'exportation des CloudWatch métriques vers un serveur Prometheus est implémentée au-dessus de l'API. CloudWatch Par exemple, le déploiement de plusieurs instances d' CloudWatch Exporter dans une seule Compte AWS et même région pour surveiller des centaines d'instances de base de données Amazon RDS peut entraîner une erreur de régulation (ThrottlingException) et des erreurs de code 400. Pour surmonter ces limites, pensez à utiliser l'exportateur YACE, qui est optimisé pour collecter jusqu'à 500 mesures différentes en une seule demande. En outre, pour déployer un grand nombre d'instances de base de données Amazon RDS, vous devez envisager d'en utiliser plusieurs Comptes AWS, au lieu de centraliser la charge de travail en une seule Compte AWS et de limiter le nombre d'instances d'exportation dans chacune. Compte AWS

Les alertes sont générées par le serveur Prometheus et gérées par Alertmanager. Cet outil se charge de dédupliquer, de regrouper et d'acheminer les alertes vers le destinataire approprié, tel que les e-mails, les SMS ou Slack, ou de lancer une action de réponse automatique. Grafana, un autre outil open source, affiche des visualisations pour ces indicateurs. Grafana fournit des widgets de visualisation riches, tels que des graphiques avancés, des tableaux de bord dynamiques et des

fonctionnalités d'analyse telles que les requêtes ad hoc et l'analyse dynamique. Il peut également rechercher et analyser les journaux, et inclut des fonctionnalités d'alerte pour évaluer en permanence les métriques et les journaux, et envoyer des notifications lorsque les données correspondent aux règles d'alerte.



Percona

Percona Monitoring and Management (PMM) est une solution open source gratuite de surveillance, de gestion et d'observabilité de bases de données pour MySQL et MariaDB. PMM collecte des milliers de mesures de performance auprès des instances de base de données et de leurs hôtes. Il fournit une interface utilisateur Web pour visualiser les données dans des tableaux de bord et des fonctionnalités supplémentaires telles que des conseillers automatiques pour les évaluations de l'état de santé des bases de données. Vous pouvez utiliser PMM pour surveiller Amazon RDS. Cependant, le client PMM (agent) n'est pas installé sur les hôtes sous-jacents des instances de base de données Amazon RDS, car il n'a pas accès aux hôtes. Au lieu de cela, l'outil se connecte aux instances de base de données Amazon RDS, interroge les statistiques du serveur`INFORMATION_SCHEMA`, le schéma système et le schéma de performance, et utilise l' CloudWatch API pour acquérir des métriques, des journaux, des événements et des traces. PMM a besoin d'une clé d'accès utilisateur AWS Identity and Access Management (IAM) (rôle IAM) et découvre automatiquement les instances de base de données Amazon RDS disponibles pour la surveillance. L'outil PMM est profilé pour la surveillance des bases de données et collecte plus de métriques spécifiques à la base de données que Prometheus. Pour utiliser le tableau de bord PMM Query Analytics, vous devez configurer le schéma de performance comme source de requête, car l'agent Query Analytics n'est pas installé pour Amazon RDS et ne peut pas lire le journal des requêtes lentes. Au lieu de cela, il interroge directement les instances `performance_schema` de base de données MySQL et MariaDB pour obtenir des métriques. L'une des principales caractéristiques de PMM est sa capacité à alerter et à conseiller DBAs sur les problèmes identifiés par l'outil dans leurs bases de données. PMM propose des ensembles de contrôles capables de détecter les menaces de sécurité courantes, la dégradation des performances, la perte de données et la corruption de données.

Outre ces outils, il existe plusieurs solutions commerciales d'observabilité et de surveillance disponibles sur le marché qui peuvent s'intégrer à Amazon RDS. Les exemples incluent la surveillance des bases de données Datadog, la surveillanceDynatrace Amazon RDS et la surveillance des bases de données. AppDynamics

Surveillance des instances de base de données

Une instance de base de données est l'élément de base d'Amazon RDS. Il s'agit d'un environnement de base de données isolé qui s'exécute dans le cloud. Pour les bases de données MySQL et MariaDB, l'instance de base de données est le programme mysqld, également connu sous le nom de serveur MySQL, qui inclut plusieurs threads et composants tels que l'analyseur SQL, l'optimiseur de requêtes, le gestionnaire de threads/connexions, les variables système et d'état, et un ou plusieurs moteurs de stockage enfichables. Chaque moteur de stockage est conçu pour prendre en charge un cas d'utilisation spécifique. Le moteur de stockage par défaut et recommandé est InnoDB, un moteur de base de données relationnelle transactionnel à usage général conforme au modèle ACID (atomicité, cohérence, isolation et durabilité). InnoDB propose des structures en mémoire (pool de mémoire tampon, tampon de modification, index de hachage adaptatif, tampon de journal) ainsi que des structures sur disque (espaces de table, tables, index, journal d'annulation, journal de rétablissement, fichiers tampons à double écriture). Pour garantir que votre base de données adhère étroitement au modèle ACID, le moteur de stockage InnoDB met en œuvre de nombreuses fonctionnalités pour protéger vos données, notamment les transactions, le commit, le rollback, la reprise après incident, le verrouillage au niveau des lignes et le contrôle de la simultanéité multiversion (MVCC).

Tous ces composants internes d'une instance de base de données fonctionnent conjointement pour aider à maintenir la disponibilité, l'intégrité et la sécurité de vos données au niveau de performance attendu et satisfaisant. En fonction de votre charge de travail, chaque composant et fonctionnalité peut imposer des exigences en termes de ressources au niveau du processeur, de la mémoire, du réseau et des sous-systèmes de stockage. Lorsqu'une augmentation de la demande pour une ressource spécifique dépasse la capacité allouée ou les limites logicielles pour cette ressource (imposées par les paramètres de configuration ou par la conception du logiciel), l'instance de base de données peut subir une dégradation des performances ou une indisponibilité et une corruption complètes. Il est donc essentiel de mesurer et de surveiller ces composants internes, de les comparer aux valeurs de référence définies et de générer des alertes si les valeurs surveillées s'écartent des valeurs attendues.

Comme décrit précédemment, vous pouvez utiliser différents outils pour surveiller vos instances MySQL et MariaDB. Nous vous recommandons d'utiliser Amazon RDS Performance Insights et les CloudWatch outils de surveillance et d'alerte, car ces outils sont intégrés à Amazon RDS, collectent des métriques haute résolution, présentent les dernières informations de performance en temps quasi réel et génèrent des alarmes.

Quel que soit votre outil de surveillance préféré, nous vous recommandons d'activer le schéma de performance dans vos instances de base de données MySQL et MariaDB. Le schéma de performance est une fonctionnalité optionnelle permettant de surveiller le fonctionnement du serveur MySQL (l'instance de base de données) à un niveau bas. Il est conçu pour avoir un impact minimal sur les performances globales de la base de données. Vous pouvez gérer cette fonctionnalité à l'aide du `performance_schema` paramètre. Bien que ce paramètre soit facultatif, vous devez l'utiliser pour collecter des métriques haute résolution (une seconde) par SQL, des métriques de session active, des événements d'attente et d'autres informations de surveillance détaillées de bas niveau, collectées par Amazon RDS Performance Insights.

Sections

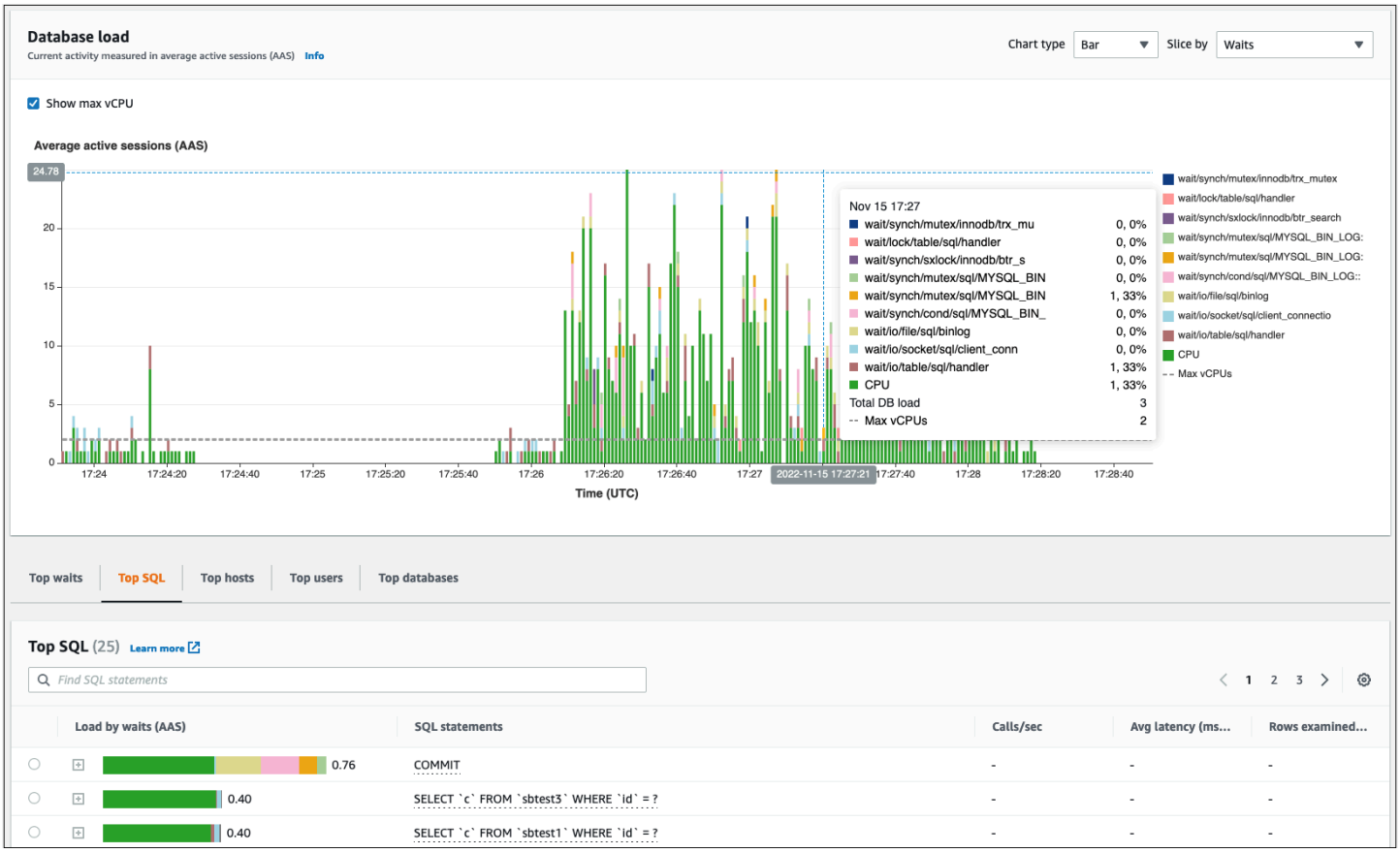
Mesures Performance Insights pour les instances de base de données

Performance Insights surveille différents types de mesures, comme indiqué dans les sections suivantes.

Charge de la base de données

Le chargement de la base de données (`DBLoad`) est un indicateur clé de Performance Insights qui mesure le niveau d'activité de votre base de données. Il est collecté chaque seconde et publié automatiquement sur Amazon CloudWatch. Il représente l'activité de l'instance de base de données dans le nombre moyen de sessions actives (AAS), c'est-à-dire le nombre de sessions exécutant simultanément des requêtes SQL. La `DBLoad` métrique est différente des autres métriques de séries chronologiques, car elle peut être interprétée à l'aide de l'une des cinq dimensions suivantes : temps d'attente, SQL, hôtes, utilisateurs et bases de données. Ces dimensions sont des sous-catégories de la `DBLoad` métrique. Vous pouvez les utiliser sous forme de tranches par catégories pour représenter les différentes caractéristiques de la charge de base de données. Pour une description détaillée de la façon dont nous calculons la charge de base de données, consultez la section Charge de base de données dans la documentation Amazon RDS.

L'illustration d'écran suivante montre l'outil Performance Insights.



Dimensions

- Les événements d'attente sont des conditions dans lesquelles une session de base de données attend la fin d'une ressource ou d'une autre opération afin de poursuivre son traitement. Si vous exécutez une instruction SQL telle que `SELECT * FROM big_table` et si cette table est beaucoup plus grande que le pool de tampons InnoDB alloué, votre session attendra probablement les événements d'`wait/io/file/innodb/innodb_data_file`attente, qui sont provoqués par des opérations d'E/S physiques sur le fichier de données. Les événements d'attente constituent une dimension importante pour la surveillance des bases de données, car ils indiquent d'éventuels problèmes de performance. Les événements d'attente indiquent les ressources et les opérations que les instructions SQL que vous exécutez au cours des sessions passent le plus de temps à attendre. Par exemple, l'`wait/synch/mutex/innodb/trx_sys_mutex`événement se produit lorsque l'activité de la base de données est élevée avec un grand nombre de transactions, et l'`wait/synch/mutex/innodb/buf_pool_mutex`événement se produit lorsqu'un thread a acquis un verrou sur le pool de mémoire tampon d'InnoDB pour accéder à une page en mémoire. Pour plus d'informations sur tous les événements d'attente MySQL et MariaDB, consultez les tableaux récapitulatifs des événements d'attente dans la documentation MySQL. Pour comprendre

comment interpréter les noms d'instruments, consultez les conventions de dénomination des instruments du schéma de performance dans la documentation MySQL.

- SQL indique les instructions SQL qui contribuent le plus à la charge totale de la base de données. Le tableau des principales dimensions, qui se trouve sous le graphique de charge de la base de données d'Amazon RDS Performance Insights, est interactif. Vous pouvez obtenir une liste détaillée des événements d'attente associés à l'instruction SQL en cliquant sur la barre dans la colonne Load by waits (AAS). Lorsque vous sélectionnez une instruction SQL dans la liste, Performance Insights affiche les événements d'attente associés dans le graphique de charge de la base de données et le texte de l'instruction SQL dans la section de texte SQL. Les statistiques SQL sont affichées sur le côté droit du tableau des principales dimensions.
- Les hôtes affichent les noms d'hôte des clients connectés. Cette dimension vous permet d'identifier les hôtes clients qui envoient la majeure partie de la charge à la base de données.
- Les utilisateurs regroupent la charge de base de données en fonction des utilisateurs connectés à la base de données.
- Les bases de données regroupent la charge de base de données selon le nom de la base de données à laquelle le client est connecté.

Métriques de compteur

Les contre-métriques sont des métriques cumulatives dont les valeurs ne peuvent être augmentées ou remises à zéro que lorsque l'instance de base de données redémarre. La valeur d'une contre-métrique ne peut pas être réduite à sa valeur précédente. Ces métriques représentent un compteur unique qui augmente de façon monotone.

- Les compteurs natifs sont des métriques définies par le moteur de base de données et non par Amazon RDS. Par exemple :
  - `SQL.Innodb_rows_inserted`représente le nombre de lignes insérées dans les tables InnoDB.
  - `SQL.Select_scan`représente le nombre de jointures ayant effectué une analyse complète de la première table.
  - `Cache.Innodb_buffer_pool_reads`représente le nombre de lectures logiques que le moteur InnoDB n'a pas pu récupérer depuis le pool de mémoire tampon et a dû lire directement depuis le disque.
  - `Cache.Innodb_buffer_pool_read_requests`représente le nombre de demandes de lecture logiques.

Pour les définitions de toutes les métriques natives, consultez la section Variables d'état du serveur dans la documentation MySQL.

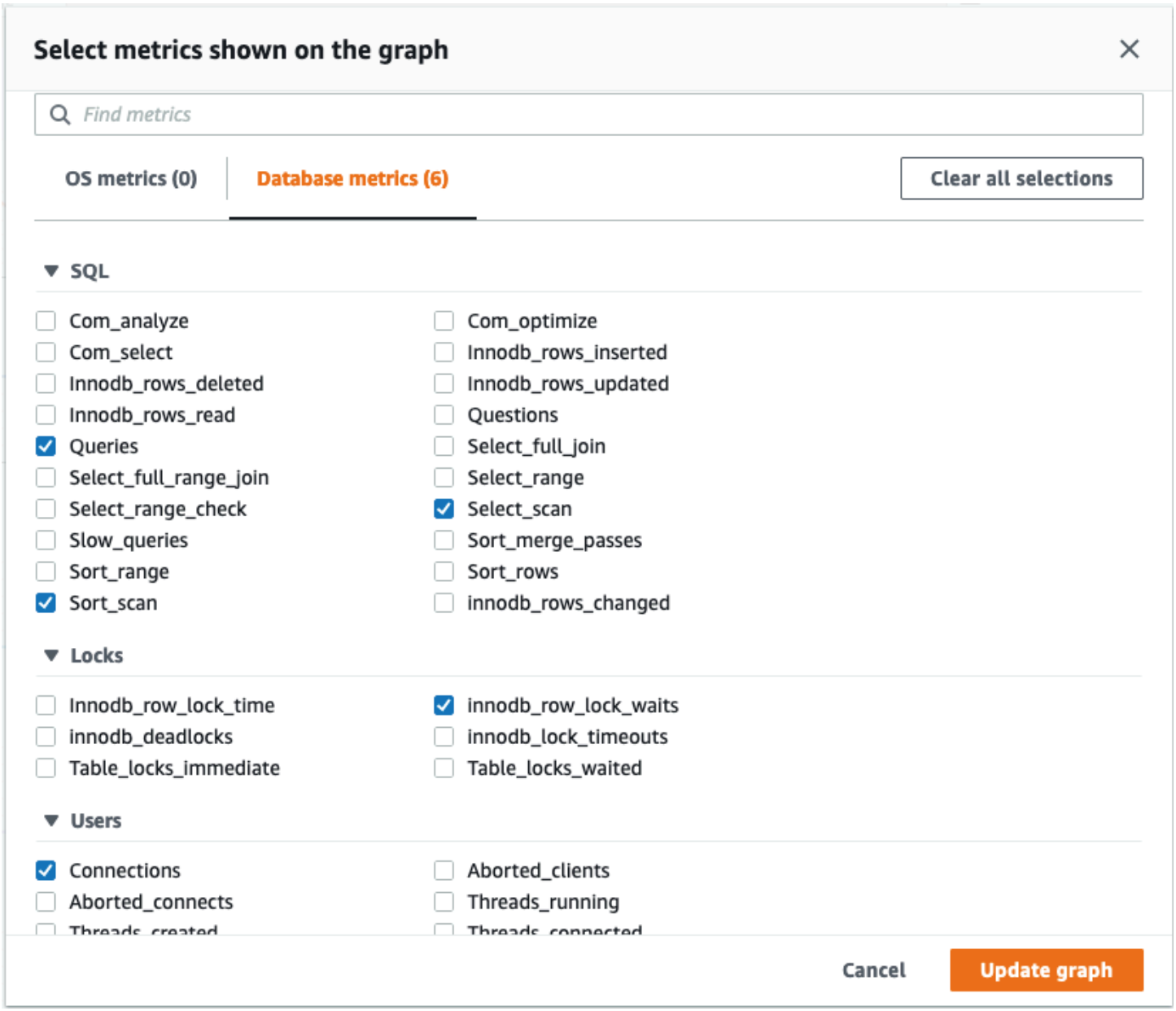
- Les compteurs non natifs sont définis par Amazon RDS. Vous pouvez obtenir ces mesures à l'aide d'une requête spécifique ou les dériver en utilisant au moins deux mesures natives dans les calculs. Les indicateurs de compteur non natifs peuvent représenter des latences, des ratios ou des taux de réussite. Par exemple :
  - `Cache.innoDB_buffer_pool_hits`représente le nombre d'opérations de lecture qu'InnoDB a pu récupérer depuis le pool de mémoire tampon sans utiliser le disque. Il est calculé à partir des métriques de compteur natives comme suit :

    ```
    db.Cache.Innodb_buffer_pool_read_requests - db.Cache.Innodb_buffer_pool_reads
    ```

  - `IO.innoDB_datafile_writes_to_disk`représente le nombre d'opérations d'écriture de fichiers de données InnoDB sur le disque. Il capture uniquement les opérations sur les fichiers de données, et non les opérations d'écriture en double écriture ou en journalisation à nouveau. Il est calculé comme suit :

    ```
    db.IO.Innodb_data_writes - db.IO.Innodb_log_writes - db.IO.Innodb_dblwr_writes
    ```

Vous pouvez visualiser les métriques des instances de base de données directement dans le tableau de bord Performance Insights. Choisissez Gérer les mesures, cliquez sur l'onglet Mesures de base de données, puis sélectionnez les mesures qui vous intéressent, comme indiqué dans l'illustration suivante.



Cliquez sur le bouton Mettre à jour le graphique pour afficher les mesures que vous avez sélectionnées, comme indiqué dans l'illustration suivante.

Statistiques SQL

Performance Insights rassemble des mesures relatives aux performances relatives aux requêtes SQL pour chaque seconde d'exécution d'une requête et pour chaque appel SQL. En général, Performance Insights collecte des statistiques SQL au niveau des instructions et du résumé. Toutefois, pour les instances de base de données MariaDB et MySQL, les statistiques sont collectées uniquement au niveau du condensé.

- Les statistiques Digest sont une métrique composite de toutes les requêtes qui ont le même modèle mais qui ont finalement des valeurs littérales différentes. Le condensé remplace des valeurs littérales spécifiques par une variable, par exemple :

```
SELECT department_id, department_name FROM departments WHERE location_id = ?
```

- Certaines métriques représentent les statistiques par seconde pour chaque instruction SQL digérée. Par exemple, `sql_tokenized.stats.count_star_per_sec` représente les appels par seconde (c'est-à-dire le nombre de fois par seconde où l'instruction SQL a été exécutée).

- Performance Insights inclut également des métriques qui fournissent des statistiques par appel pour une instruction SQL. Par exemple, `sql_tokenized.stats.sum_timer_wait_per_call` indique la latence moyenne de l'instruction SQL par appel, en millisecondes.

Les statistiques SQL sont disponibles dans le tableau de bord Performance Insights, dans l'onglet Top SQL du tableau des principales dimensions.

CloudWatch métriques pour les instances de base de données

Amazon contient CloudWatch également des métriques qu'Amazon RDS publie automatiquement. Les métriques qui résident dans l'`AWS/RDS`espace de noms sont des métriques au niveau de l'instance, qui font référence à l'instance Amazon RDS (c'est-à-dire à l'environnement de base de données isolé exécuté dans le cloud) plutôt qu'à l'instance de base de données au sens strict du processus mysqld. Par conséquent, la plupart de ces mesures par défaut entrent dans la catégorie des métriques du système d'exploitation, au sens strict du terme. Les exemples incluent :`CPUUtilization`,`WriteIOPS`,`SwapUsage`, et autres. Néanmoins, certaines métriques d'instance de base de données sont applicables à MariaDB et MySQL :

- `BinLogDiskUsage`— La quantité d'espace disque occupée par les journaux binaires.
- `DatabaseConnections`— Le nombre de connexions réseau client à l'instance de base de données.
- `ReplicaLag`— Durée pendant laquelle une instance de base de données de réplique en lecture est en retard par rapport à l'instance de base de données source.

Publication des indicateurs de Performance Insights sur CloudWatch

Amazon RDS Performance Insights surveille la plupart des métriques et dimensions des instances de base de données et les met à disposition via le tableau de bord Performance Insights de la console AWS de gestion. Ce tableau de bord convient parfaitement au dépannage des bases de données et à l'analyse des causes premières. Cependant, il n'est pas possible de créer des alarmes dans Performance Insights pour les mesures liées aux performances. Si vous souhaitez créer des alarmes basées sur les indicateurs de Performance Insights, ces indicateurs doivent être intégrés CloudWatch.

Performance Insights publie automatiquement les indicateurs sur CloudWatch. Vous pouvez interroger les mêmes données à partir de Performance Insights, mais l'intégration des métriques CloudWatch facilite l'ajout d' CloudWatch alarmes et l'ajout des métriques aux CloudWatch tableaux de bord existants. Les compteurs sont des indicateurs de performance du système d'exploitation et de la base de données, tels que `os.memory.free` oudb.Locks.Innodb_row_lock_time. La collecte des métriques du système d'exploitation dépend du paramètre de surveillance améliorée. Si la surveillance améliorée est désactivée, les métriques du système d'exploitation sont collectées une fois par minute. Si la surveillance améliorée est activée, les métriques du système d'exploitation sont collectées pour la période sélectionnée. Pour plus d'informations, consultez la section Activation et désactivation de la surveillance améliorée dans la documentation Amazon RDS.

Performance Insights vous permet d'exporter le tableau de bord des métriques préconfiguré ou personnalisé pour votre instance de base de données vers CloudWatch. Vous pouvez exporter le tableau de bord des métriques en tant que nouveau tableau de bord ou l'ajouter à un CloudWatch tableau de bord existant. L'exportation du tableau de bord des métriques Performance Insights vers le CloudWatch tableau de bord vous donne une vue unifiée et globale de l'état de votre système en fournissant une vue d'ensemble des métriques associées aux différentes ressources de votre système, telles que les EC2 instances, les ressources Amazon Elastic File System (Amazon EFS) et les ressources Elastic Load Balancing (ELB), ainsi que les métriques de votre instance de base de données.

Vous pouvez utiliser la fonction mathématique des CloudWatch `DB_PERF_INSIGHTS` métriques pour interroger et créer des alarmes et des graphiques basés sur les métriques Performance Insights provenant de CloudWatch. Pour créer une alarme sur une métrique Performance Insights, suivez les instructions de la CloudWatch documentation. Par exemple, si vous souhaitez déclencher une alarme lorsque le total des transactions actives dans votre instance de base de données atteint un

seuil spécifique, suivez les instructions de cette page, utilisez l'expression DB_PERF_INSIGHTS mathématique suivante, puis choisissez Appliquer :

```
DB_PERF_INSIGHTS('RDS', 'db-BQ2TPYY7HG2GDFC7APMB3BVB3M',
 'db.Transactions.active_transactions.avg')
```

où db-BQ2TPYY7HG2GDFC7APMB3BVB3M est l'ID de ressource de votre instance de base de données. Spécifiez la période (par exemple, 1 minute) et les conditions (par exemple, supérieure à 1 000). Pour finaliser la création de l'alarme, configurez les actions de l'alarme, ajoutez un nom et une description, puis prévisualisez et créez l'alarme.

Surveillance du système d'exploitation

Une instance de base de données dans Amazon RDS for MySQL ou MariaDB s'exécute sur le système d'exploitation Linux, qui utilise les ressources système sous-jacentes : processeur, mémoire, réseau et stockage.

```
MySQL [(none)]> SHOW variables LIKE 'version%';
+-------------------------+---------------------+
| Variable_name           | Value               |
+-------------------------+---------------------+
| version                 | 8.0.28              |
| version_comment         | Source distribution |
| version_compile_machine | aarch64             |
| version_compile_os      | Linux               |
| version_compile_zlib    | 1.2.11              |
+-------------------------+---------------------+
5 rows in set (0.00 sec)
```

Les performances globales de votre base de données et du système d'exploitation sous-jacent dépendent fortement de l'utilisation des ressources du système. Par exemple, le processeur est l'élément clé des performances de votre système, car il exécute les instructions du logiciel de base de données et gère les autres ressources du système. Si le processeur est surutilisé (c'est-à-dire si la charge nécessite plus de puissance processeur que celle allouée à votre instance de base de données), ce problème aura un impact sur les performances et la stabilité de votre base de données et, par conséquent, de votre application.

Le moteur de base de données alloue et libère de la mémoire de manière dynamique. Lorsque la mémoire RAM est insuffisante pour effectuer le travail en cours, le système écrit des pages de mémoire dans la mémoire d'échange, qui se trouve sur le disque. Le disque étant beaucoup plus lent que la mémoire, même s'il est basé sur la NVMe technologie SSD, une allocation excessive de mémoire entraîne une dégradation des performances. L'utilisation élevée de la mémoire entraîne une latence accrue des réponses de la base de données, car la taille d'un fichier de page augmente pour prendre en charge de la mémoire supplémentaire. Si l'allocation de mémoire est si élevée qu'elle épuise à la fois la RAM et les espaces de mémoire d'échange, le service de base de données peut devenir indisponible et les utilisateurs peuvent observer des erreurs telles que. `[ERROR] mysqld: Out of memory (Needed xyz bytes)`

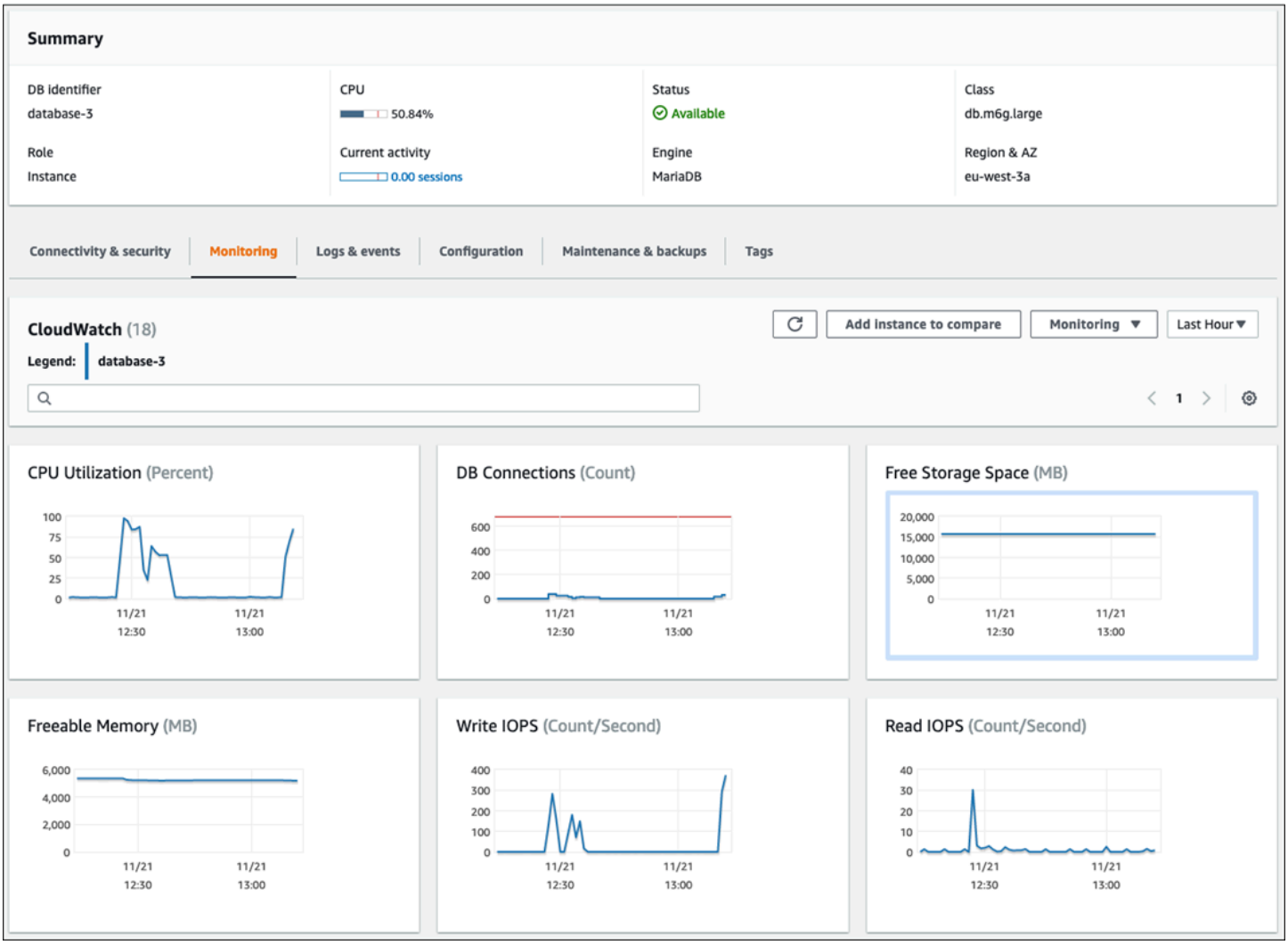
Les systèmes de gestion de bases de données MySQL et MariaDB utilisent le sous-système de stockage, qui comprend des disques qui stockent des structures sur disque telles que des tables, des index, des journaux binaires, des journaux de rétablissement, des journaux d'annulation et des fichiers tampons à double écriture. Par conséquent, la base de données, contrairement aux autres types de logiciels, doit effectuer une activité importante sur le disque. Pour un fonctionnement optimal de votre base de données, il est important de surveiller et de régler l'utilisation des E/S du disque ainsi que l'allocation de l'espace disque. Les performances de la base de données peuvent être affectées lorsque la base de données atteint les limites d'IOPS ou de débit maximal pris en charge par le disque. Par exemple, les accès aléatoires provoqués par une analyse d'index peuvent entraîner un grand nombre d'opérations d'E/S par seconde, ce qui peut éventuellement affecter les limites du stockage sous-jacent. Les scans complets des tables peuvent ne pas atteindre la limite d'IOPS, mais ils peuvent entraîner un débit élevé, mesuré en mégaoctets par seconde. Il est essentiel de surveiller et de générer des alertes concernant l'allocation d'espace disque, car des erreurs telles que celles-ci `OS error code 28: No space left on device` peuvent entraîner l'indisponibilité et la corruption de la base de données.

Amazon RDS fournit des métriques en temps réel pour le système d'exploitation sur lequel s'exécute votre instance de base de données. Amazon RDS publie automatiquement un ensemble de statistiques du système d'exploitation sur CloudWatch. Ces métriques sont disponibles pour affichage et analyse dans la console Amazon RDS et les CloudWatch tableaux de bord, et vous pouvez définir des alarmes sur les métriques sélectionnées dans. CloudWatch En voici quelques exemples :

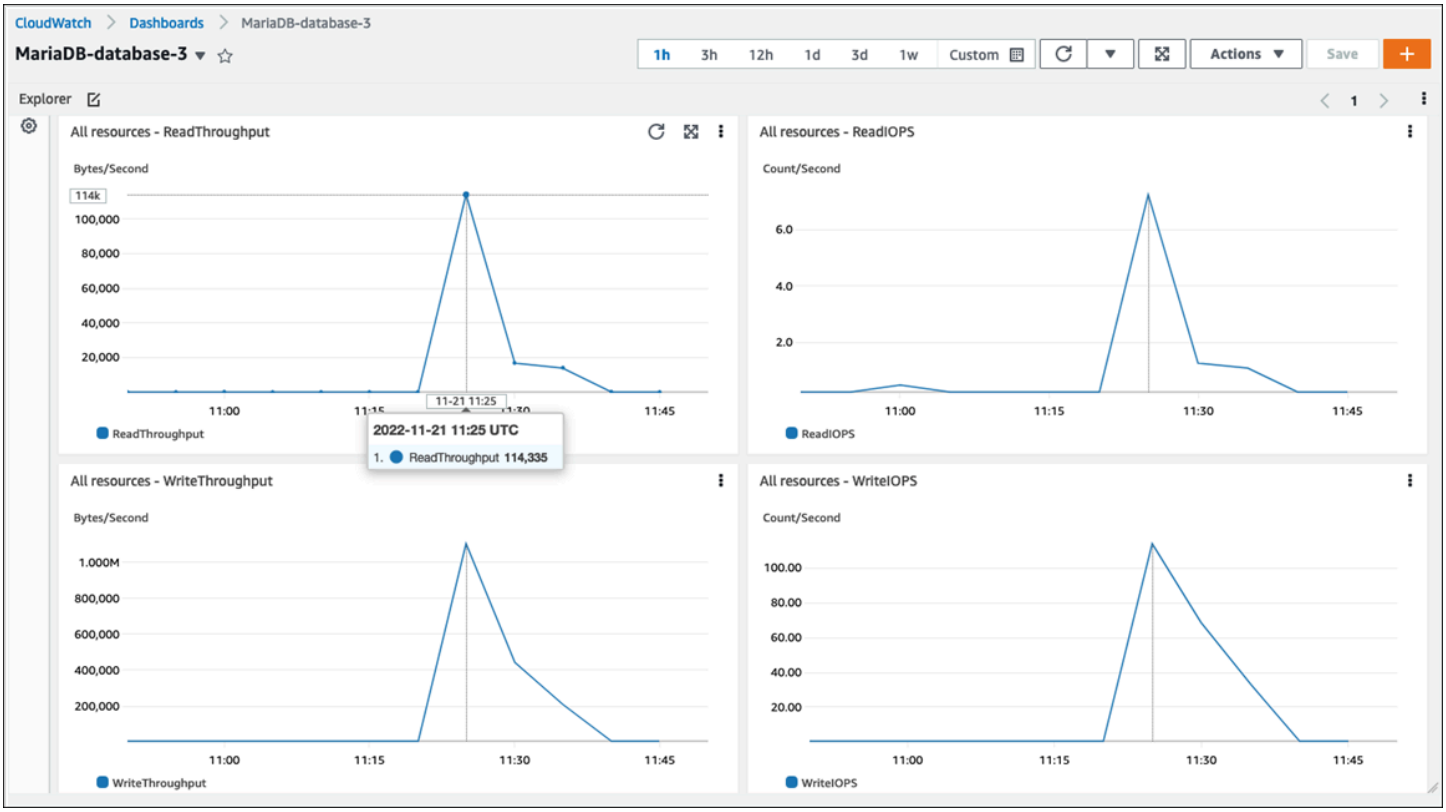
- `CPUUtilization`— Le pourcentage d'utilisation du processeur.
- `BinLogDiskUsage`— La quantité d'espace disque occupée par les journaux binaires.
- `FreeableMemory`— La quantité de mémoire à accès aléatoire disponible. Cela représente la valeur du `MemAvailable` champ de`/proc/meminfo`.
- `ReadIOPS`— Nombre moyen d'opérations d'E/S de lecture sur disque par seconde.
- `WriteThroughput`— Nombre moyen d'octets écrits sur le disque par seconde pour le stockage local.
- `NetworkTransmitThroughput`— Le trafic réseau sortant sur le nœud de base de données, qui combine à la fois le trafic de base de données et le trafic Amazon RDS utilisés pour la surveillance et la réplication.

Pour une référence complète de toutes les métriques publiées par Amazon RDS CloudWatch, consultez les CloudWatch métriques Amazon pour Amazon RDS dans la documentation Amazon RDS.

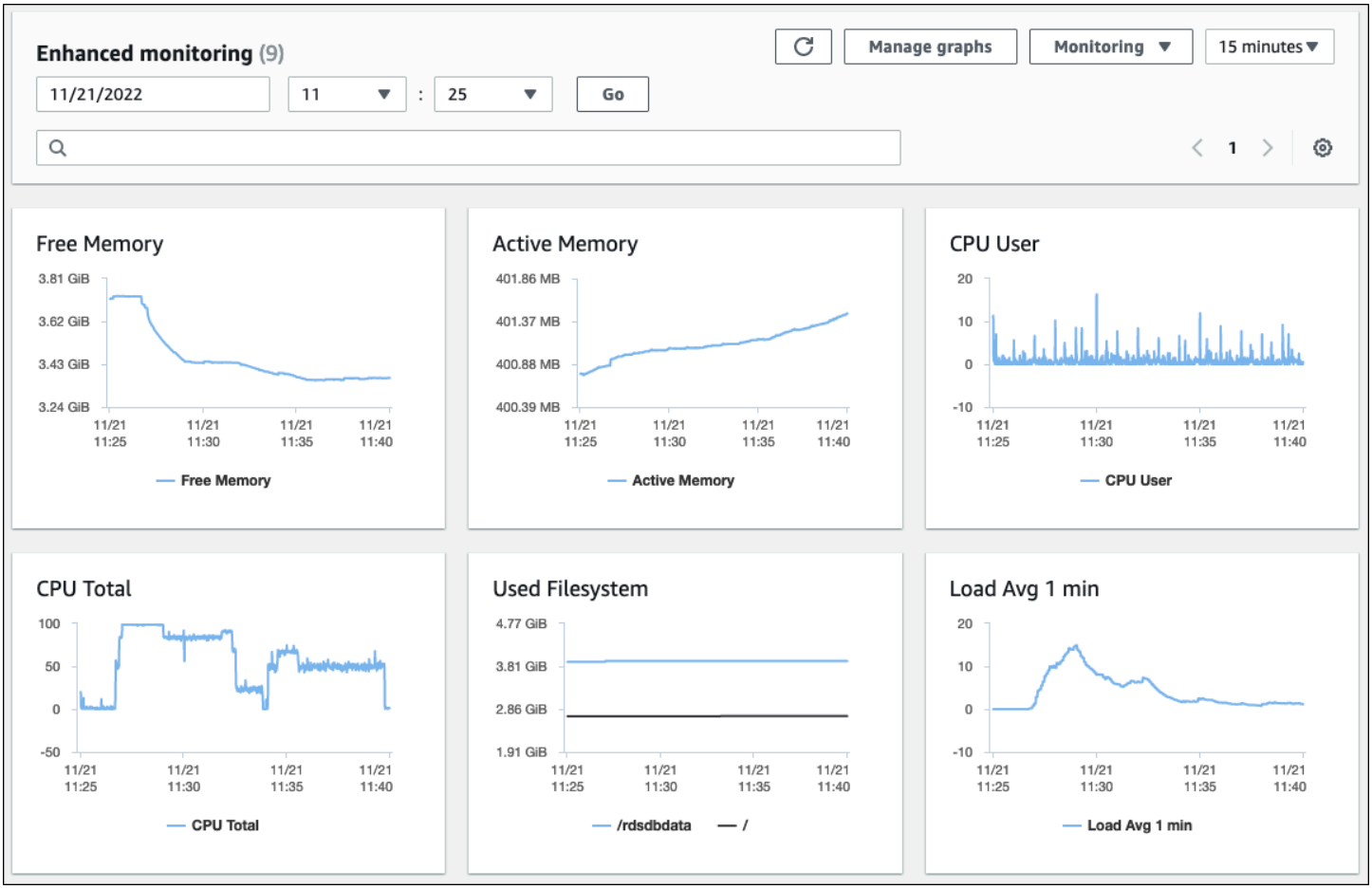
Le graphique suivant présente des exemples de CloudWatch métriques pour Amazon RDS affichées sur la console Amazon RDS.



Le graphique suivant montre des mesures similaires affichées dans le CloudWatch tableau de bord.



L'autre ensemble de métriques du système d'exploitation est collecté par Enhanced Monitoring pour Amazon RDS. Cet outil vous donne une meilleure visibilité sur l'état de vos instances de base de données Amazon RDS for MariaDB et Amazon RDS for MySQL, en fournissant des métriques système en temps réel et des informations sur les processus du système d'exploitation. Lorsque vous activez la surveillance améliorée sur votre instance de base de données et que vous définissez la granularité souhaitée, l'outil collecte les métriques du système d'exploitation et les informations de processus, que vous pouvez afficher et analyser sur la console Amazon RDS, comme indiqué dans l'écran suivant.



Certains des indicateurs clés fournis par Enhanced Monitoring sont les suivants :

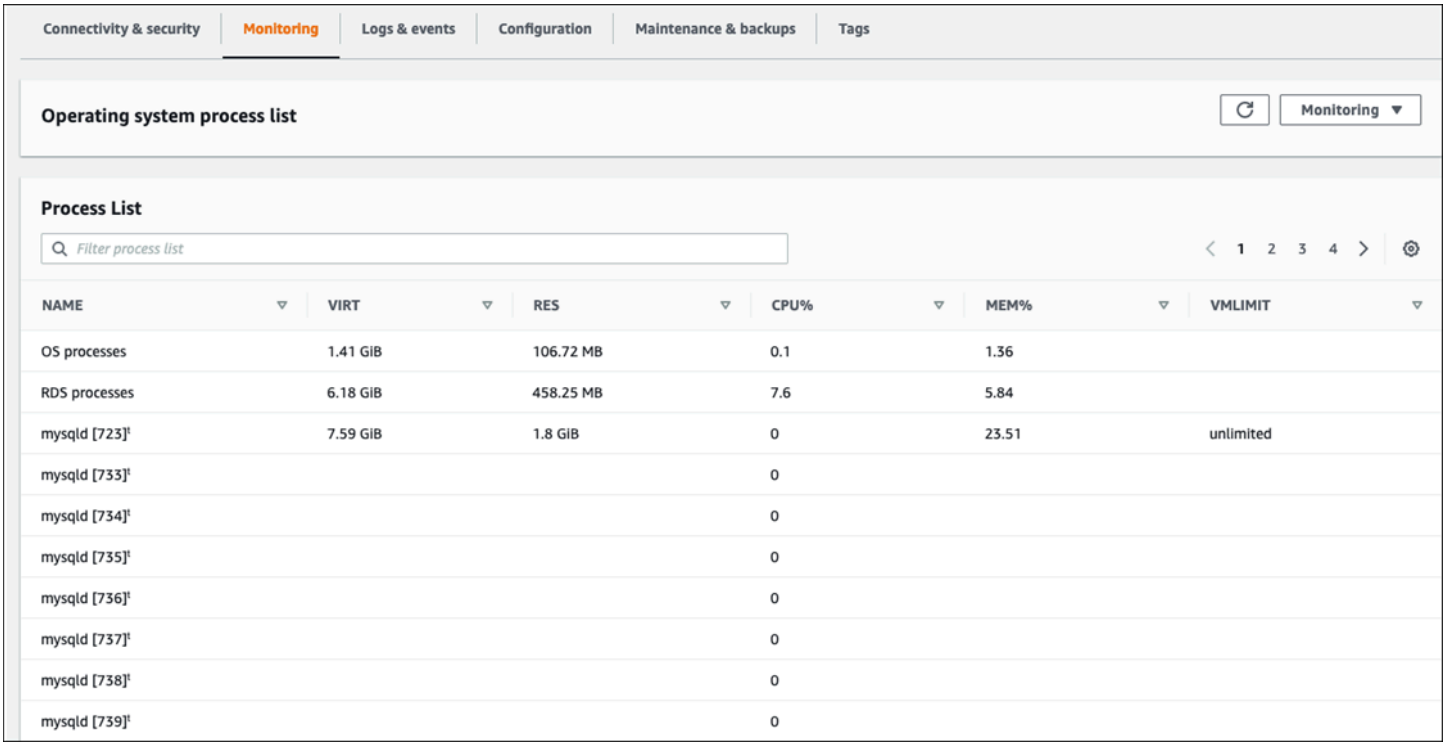
- `cpuUtilization.total`— Le pourcentage total du processeur utilisé.
- `cpuUtilization.user`— Pourcentage de CPU utilisé par les programmes utilisateur.
- `memory.active`— La quantité de mémoire attribuée, en kilo-octets.
- `memory.cached`— Quantité de mémoire utilisée pour la mise en cache des E/S basées sur le système de fichiers.
- `loadAverageMinute.one`— Le nombre de processus ayant demandé du temps processeur au cours de la dernière minute.

Pour une liste complète des métriques, consultez les métriques du système d'exploitation dans la section Enhanced Monitoring de la documentation Amazon RDS.

Sur la console Amazon RDS, la liste des processus du système d'exploitation fournit des informations détaillées sur chaque processus exécuté dans votre instance de base de données. La liste est organisée en trois sections :

- Processus du système d'exploitation – Cette section représente un résumé agrégé de tous les processus du noyau et du système. Ces processus ont généralement un impact minimal sur les performances de la base de données.
- Processus RDS — Cette section présente un résumé des AWS processus requis pour prendre en charge une instance de base de données Amazon RDS. Par exemple, il inclut l'agent de gestion Amazon RDS, les processus de surveillance et de diagnostic, ainsi que des processus similaires.
- Processus enfants RDS : cette section présente un résumé des processus Amazon RDS qui prennent en charge l'instance de base de données, dans ce cas, le `mysqld` processus et ses threads. Les `mysqld` fils apparaissent imbriqués sous le `mysqld` processus parent.

L'illustration d'écran suivante montre la liste des processus du système d'exploitation dans la console Amazon RDS.

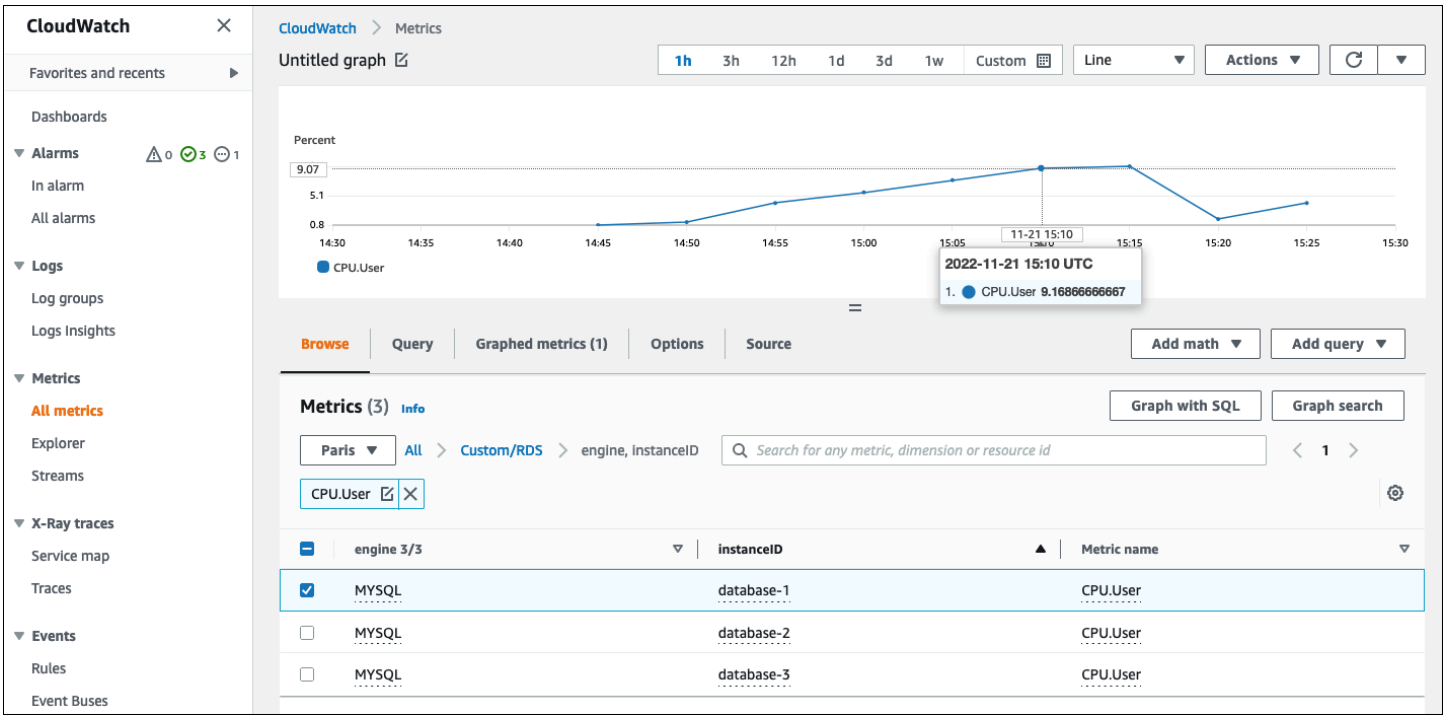


Amazon RDS fournit les statistiques issues de la surveillance améliorée dans votre compte CloudWatch Logs. Les données de surveillance affichées sur la console Amazon RDS sont extraites des CloudWatch journaux. Vous pouvez également récupérer les métriques d'une instance de base

de données sous forme de flux de journal à partir de CloudWatch Logs. Ces métriques sont stockées au format JSON. Vous pouvez utiliser la sortie JSON Enhanced Monitoring de CloudWatch Logs dans le système de surveillance de votre choix.

Pour afficher des graphiques sur le CloudWatch tableau de bord et créer des alarmes qui déclencheraient une action si une métrique franchissait le seuil défini, vous devez créer des filtres CloudWatch de métriques dans CloudWatch Logs. Pour obtenir des instructions détaillées, consultez l'article AWS Re:Post sur la façon de filtrer les CloudWatch journaux de surveillance améliorés afin de générer des métriques personnalisées automatisées pour Amazon RDS.

L'exemple suivant illustre la métrique personnalisée `CPU.User` dans l'espace de `Custom/RDS` noms. Cette métrique personnalisée est créée en filtrant la métrique `cpuUtilization.user` Enhanced Monitoring des CloudWatch journaux.



Lorsque la métrique est disponible dans le CloudWatch référentiel, vous pouvez l'afficher et l'analyser dans CloudWatch des tableaux de bord, appliquer d'autres opérations mathématiques et de requête, et définir une alarme pour surveiller cette métrique spécifique et générer des alertes si les valeurs observées ne sont pas conformes aux conditions d'alarme définies.

Événements, journaux et pistes d'audit

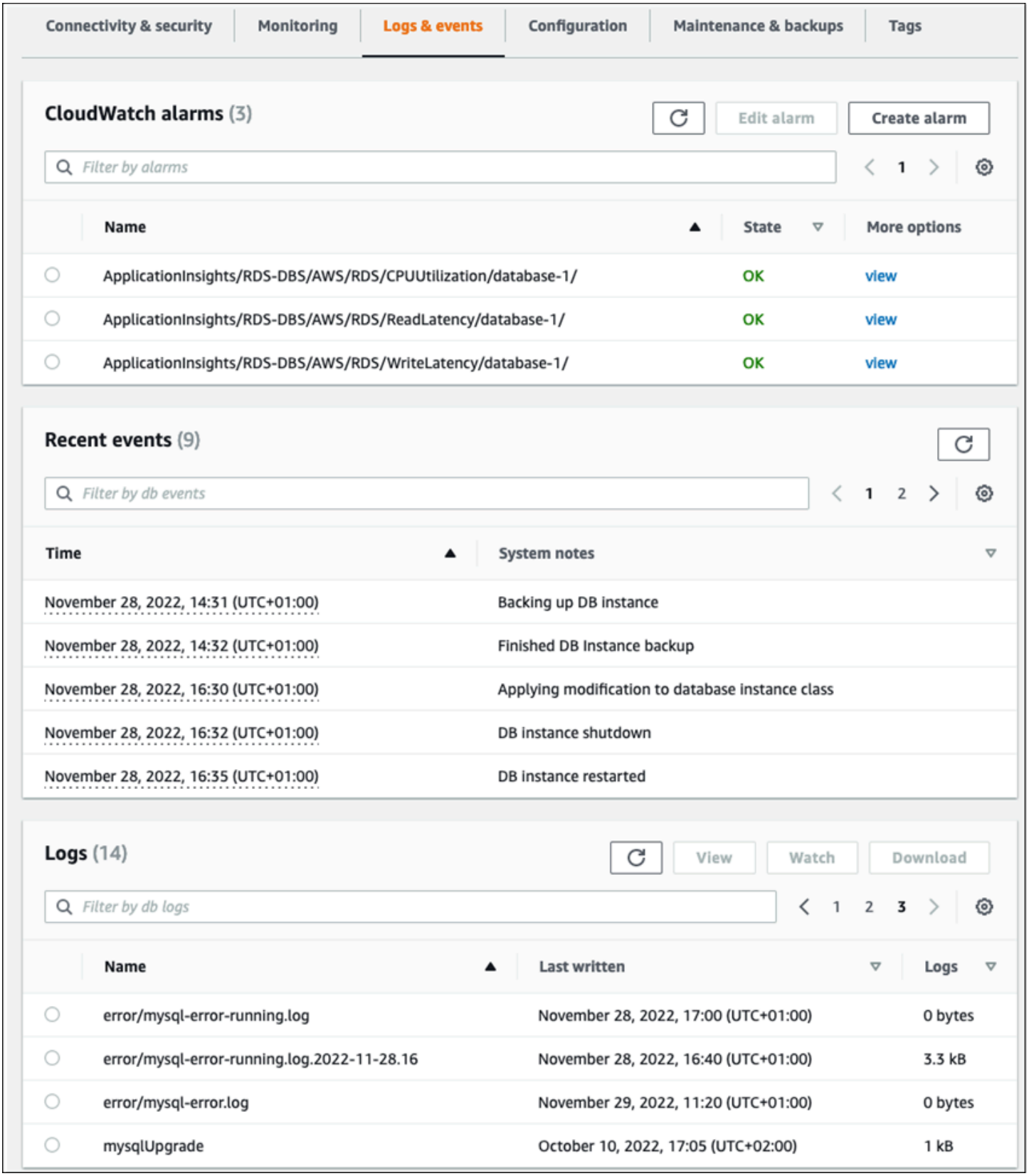
La surveillance des métriques des instances de base de données et des métriques du système d'exploitation, l'analyse des tendances, la comparaison des métriques avec les valeurs de référence et la génération d'alertes lorsque les valeurs dépassent les seuils définis sont autant de bonnes pratiques nécessaires pour vous aider à atteindre et à maintenir la fiabilité, la disponibilité, les performances et la sécurité de vos instances de base de données Amazon RDS. Cependant, une solution complète doit également surveiller les événements de base de données, les fichiers journaux et les pistes d'audit des bases de données MySQL et MariaDB.

Sections

- Événements Amazon RDS
- journaux de base de données
- Pistes d'audit

Événements Amazon RDS

Un événement Amazon RDS indique un changement dans l'environnement Amazon RDS. Par exemple, lorsque le statut de l'instance de base de données passe de Starting à Available, Amazon RDS génère l'événement`RDS-EVENT-0088 The DB instance has been started`. Amazon RDS diffuse des événements à Amazon EventBridge quasiment en temps réel. Vous pouvez accéder aux événements via la console Amazon RDS, la AWS CLI commande describe-events ou le fonctionnement de l'API Amazon RDS. DescribeEvents L'illustration d'écran suivante montre les événements et les journaux affichés sur la console Amazon RDS.


Connectivity & security
Monitoring
Logs & events
Configuration
Maintenance & backups
Tags
CloudWatch alarms (3)
Edit alarm
Create alarm
Filter by alarms
1
Name
State
More options
ApplicationInsights/RDS-DBS/AWS/RDS/CPUUtilization/database-1/
OK
view
ApplicationInsights/RDS-DBS/AWS/RDS/ReadLatency/database-1/
OK
view
ApplicationInsights/RDS-DBS/AWS/RDS/WriteLatency/database-1/
OK
view
Recent events (9)
Filter by db events
1
2
Time
System notes
November 28, 2022, 14:31 (UTC+01:00)
Backing up DB instance
November 28, 2022, 14:32 (UTC+01:00)
Finished DB Instance backup
November 28, 2022, 16:30 (UTC+01:00)
Applying modification to database instance class
November 28, 2022, 16:32 (UTC+01:00)
DB instance shutdown
November 28, 2022, 16:35 (UTC+01:00)
DB instance restarted
Logs (14)
View
Watch
Download
Filter by db logs
1
2
3
Name
Last written
Logs
error/mysql-error-running.log
November 28, 2022, 17:00 (UTC+01:00)
0 bytes
error/mysql-error-running.log.2022-11-28.16
November 28, 2022, 16:40 (UTC+01:00)
3.3 kB
error/mysql-error.log
November 29, 2022, 11:20 (UTC+01:00)
0 bytes
mysqlUpgrade
October 10, 2022, 17:05 (UTC+02:00)
1 kB

Amazon RDS émet différents types d'événements, notamment des événements d'instance de base de données, des événements de groupe de paramètres de base de données, des événements de groupe de sécurité de base de données, des événements de capture de base de données, des événements de proxy RDS et des événements de déploiement bleu/vert. Les informations incluent :

- Nom et type de source ; par exemple : `"SourceIdentifier": "database-1", "SourceType": "db-instance"`
- Date et heure de l'événement ; par exemple : `"Date": "2022-12-01T09:20:28.595000+00:00"`
- Message associé à l'événement, par exemple : `"Message": "Finished updating DB parameter group"`
- Catégorie d'événement ; par exemple : `"EventCategories": ["configuration change"]`

Pour une référence complète, consultez les catégories d'événements et les messages d'événements Amazon RDS dans la documentation Amazon RDS.

Nous vous recommandons de surveiller les événements Amazon RDS, car ils indiquent des changements de statut concernant la disponibilité des instances de base de données, des modifications de configuration, des modifications de l'état de lecture des répliques, des événements de sauvegarde et de restauration, des actions de basculement, des événements de défaillance, des modifications apportées aux groupes de sécurité et de nombreuses autres notifications. Par exemple, si vous avez configuré une instance de base de données de réplication en lecture afin d'améliorer les performances et la durabilité de votre base de données, nous vous recommandons de surveiller les événements Amazon RDS pour la catégorie d'événements de réplication en lecture associée aux instances de base de données. Cela est dû au fait que des événements tels que `RDS-EVENT-0057 Replication on the read replica was terminated` ceux indiquant que votre réplique en lecture ne se synchronise plus avec l'instance de base de données principale. Une notification à l'équipe responsable indiquant qu'un tel événement s'est produit pourrait aider à atténuer le problème en temps opportun. Amazon EventBridge et d'autres Services AWS entités AWS Lambda, telles qu'Amazon Simple Queue Service (Amazon SQS) et Amazon Simple Notification Service (Amazon SNS), peuvent vous aider à automatiser les réponses aux événements du système tels que les problèmes de disponibilité des bases de données ou les modifications des ressources.

Sur la console Amazon RDS, vous pouvez récupérer les événements des dernières 24 heures. Si vous utilisez l'API AWS CLI ou l'API Amazon RDS pour consulter les événements, vous pouvez

récupérer les événements des 14 derniers jours à l'aide de la commande describe-events comme suit.

```
$ aws rds describe-events --source-identifier database-1 --source-type db-instance
{
    "Events": [
        {
            "SourceIdentifier": "database-1",
            "SourceType": "db-instance",
            "Message": "CloudWatch Logs Export enabled for logs [audit, error, general,
 slowquery]",
            "EventCategories": [],
            "Date": "2022-12-01T09:20:28.595000+00:00",
            "SourceArn": "arn:aws:rds:eu-west-3:111122223333:db:database-1"
        },
        {
            "SourceIdentifier": "database-1",
            "SourceType": "db-instance",
            "Message": "Finished updating DB parameter group",
            "EventCategories": [
                "configuration change"
            ],
            "Date": "2022-12-01T09:22:40.413000+00:00",
            "SourceArn": "arn:aws:rds:eu-west-3:111122223333:db:database-1"
        }
    ]
}
```

Si vous souhaitez stocker des événements sur le long terme, soit jusqu'à la période d'expiration spécifiée, soit de façon permanente, vous pouvez utiliser les CloudWatch journaux pour enregistrer les informations relatives aux événements générés par Amazon RDS. Pour implémenter cette solution, vous pouvez utiliser une rubrique Amazon SNS pour recevoir des notifications d'événements Amazon RDS, puis appeler une fonction Lambda pour enregistrer l'événement dans Logs. CloudWatch

1. Créez une fonction Lambda qui sera appelée lors de l'événement et enregistrez les informations de l'événement dans Logs. CloudWatch CloudWatch Logs est intégré à Lambda et constitue un moyen pratique de consigner les informations relatives aux événements, en utilisant la fonction d'impression pour. `stdout`

2. Créez une rubrique SNS avec un abonnement à une fonction Lambda (définissez Protocol sur Lambda) et définissez le point de terminaison sur le nom de ressource Amazon (ARN) de la fonction Lambda que vous avez créée à l'étape précédente.
3. Configurez votre rubrique SNS pour recevoir des notifications d'événements Amazon RDS. Pour obtenir des instructions détaillées, consultez l'article AWS Re:Post sur la façon de faire en sorte que votre rubrique Amazon SNS reçoive des notifications Amazon RDS.
4. Sur la console Amazon RDS, créez un nouvel abonnement à un événement. Définissez Target sur l'ARN, puis sélectionnez la rubrique SNS que vous avez créée précédemment. Définissez le type de source et les catégories d'événements à inclure en fonction de vos besoins. Pour plus d'informations, consultez la section S'abonner aux notifications d'événements Amazon RDS dans la documentation Amazon RDS.

Journaux de base de données

Les bases de données MySQL et MariaDB génèrent des journaux auxquels vous pouvez accéder à des fins d'audit et de dépannage. Ces journaux sont les suivants :

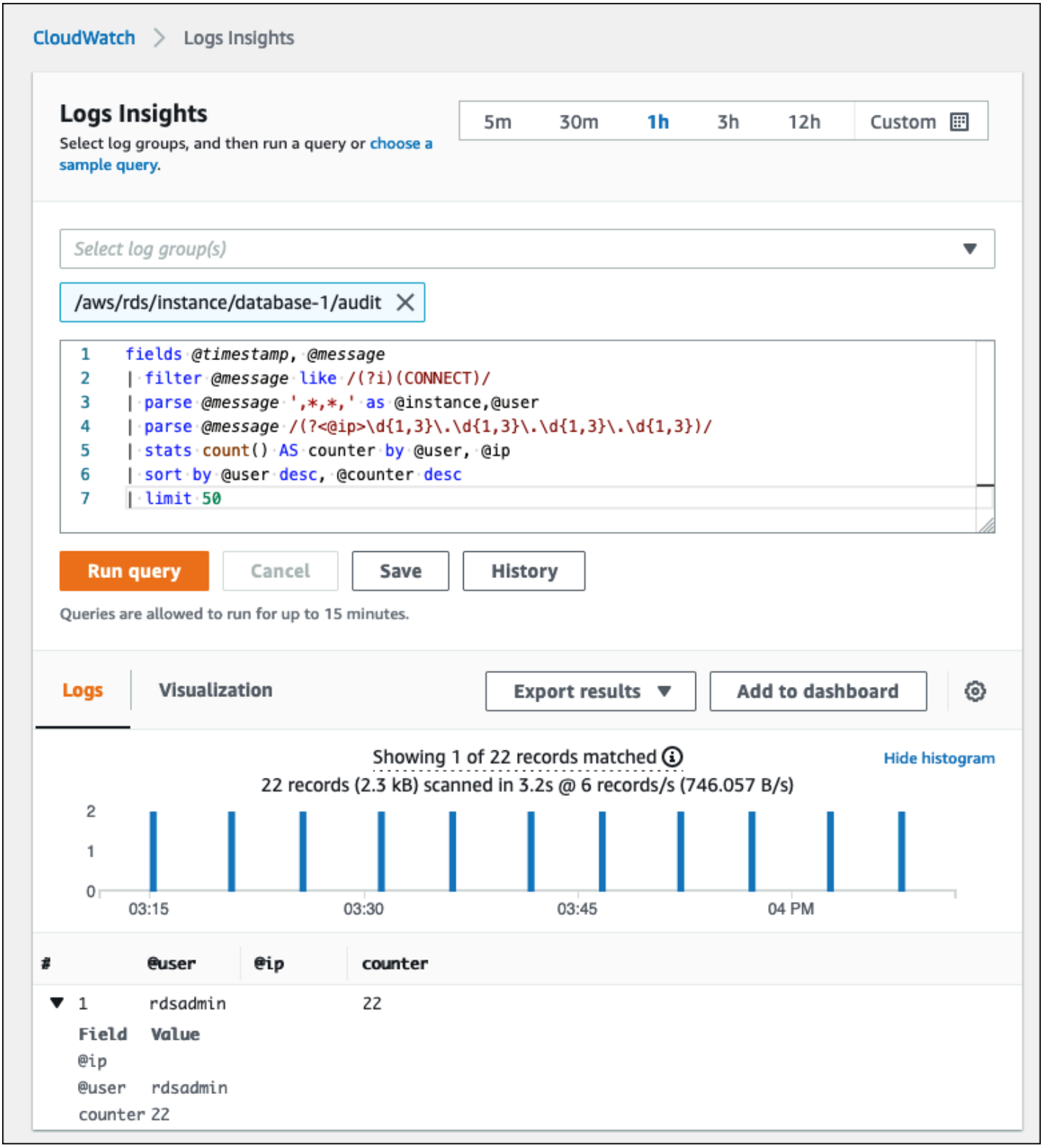
- Audit — La piste d'audit est un ensemble d'enregistrements qui enregistrent l'activité du serveur. Pour chaque session client, il enregistre qui s'est connecté au serveur (nom d'utilisateur et hôte), quelles requêtes ont été exécutées, quelles tables ont été consultées et quelles variables du serveur ont été modifiées.
- Erreur — Ce journal contient les heures de démarrage et d'arrêt du serveur (`mysqld`), ainsi que les messages de diagnostic tels que les erreurs, les avertissements et les notes qui apparaissent lors du démarrage et de l'arrêt du serveur, ainsi que pendant le fonctionnement du serveur.
- Général — Ce journal enregistre l'activité`mysqld`, y compris l'activité de connexion et de déconnexion pour chaque client, ainsi que les requêtes SQL reçues des clients. Le journal général des requêtes peut être très utile lorsque vous suspectez une erreur et que vous souhaitez savoir exactement à quoi le client a envoyé un message`mysqld`.
- Requête lente : ce journal fournit un enregistrement des requêtes SQL dont l'exécution a pris du temps.

Il est recommandé de publier les journaux de base de données d'Amazon RDS vers Amazon CloudWatch Logs. Avec CloudWatch Logs, vous pouvez effectuer une analyse en temps réel des données du journal, stocker les données dans un stockage hautement durable et gérer les données avec l'agent CloudWatch Logs. Vous pouvez accéder aux journaux de votre base de données et les

consulter depuis la console Amazon RDS. Vous pouvez également utiliser CloudWatch Logs Insights pour rechercher et analyser de manière interactive les données de vos CloudWatch journaux dans Logs. L'exemple suivant illustre une requête dans le journal d'audit qui vérifie combien de fois les `CONNECT` événements apparaissent dans le journal, qui s'est connecté et depuis quel client (adresse IP) ils se sont connectés. L'extrait du journal d'audit pourrait ressembler à ceci :

```
20221201 14:07:05,ip-10-22-1-51,rdsadmin,localhost,821,0,CONNECT,,,0,SOCKET
20221201 14:07:05,ip-10-22-1-51,rdsadmin,localhost,821,0,DISCONNECT,,,0,SOCKET
20221201 14:12:20,ip-10-22-1-51,rdsadmin,localhost,822,0,CONNECT,,,0,SOCKET
20221201 14:12:20,ip-10-22-1-51,rdsadmin,localhost,822,0,DISCONNECT,,,0,SOCKET
20221201 14:17:35,ip-10-22-1-51,rdsadmin,localhost,823,0,CONNECT,,,0,SOCKET
20221201 14:17:35,ip-10-22-1-51,rdsadmin,localhost,823,0,DISCONNECT,,,0,SOCKET
20221201 14:22:50,ip-10-22-1-51,rdsadmin,localhost,824,0,CONNECT,,,0,SOCKET
20221201 14:22:50,ip-10-22-1-51,rdsadmin,localhost,824,0,DISCONNECT,,,0,SOCKET
```

L'exemple de requête Log Insights montre qu'une personne est `rdsadmin` connectée à la base de données `localhost` toutes les 5 minutes, pour un total de 22 fois, comme indiqué dans l'illustration suivante. Ces résultats indiquent que l'activité provenait de processus internes d'Amazon RDS tels que le système de surveillance lui-même.



Les événements du journal incluent fréquemment des messages importants que vous souhaitez prendre en compte, tels que des avertissements ou des erreurs concernant les opérations associées

aux instances de base de données MySQL et MariaDB. Par exemple, si une opération échoue, une erreur peut survenir et être enregistrée dans le fichier journal des erreurs comme suit : `ERROR 1114 (HY000): The table zip_codes is full` Vous souhaiterez peut-être surveiller ces entrées pour comprendre l'évolution de vos erreurs. Vous pouvez créer des CloudWatch métriques personnalisées à partir des journaux Amazon RDS en utilisant des filtres pour activer la surveillance automatique des journaux des bases de données Amazon RDS afin de surveiller un journal spécifique pour détecter des modèles spécifiques et de générer une alarme en cas de violation du comportement attendu. Par exemple, créez un filtre métrique pour le groupe de journaux `/aws/rds/instance/database-1/error` qui surveillerait le journal des erreurs et rechercherait le modèle spécifique, tel que`ERROR`. Définissez le modèle de filtre sur `ERROR` et la valeur de la métrique sur`1`. Le filtre détectera chaque enregistrement de journal contenant le mot clé `ERROR` et augmentera le nombre de 1 pour chaque événement de journal contenant le mot « ERROR ». Après avoir créé le filtre, vous pouvez définir une alarme pour vous avertir en cas de détection d'erreurs dans le journal des erreurs MySQL ou MariaDB.

Pour en savoir plus sur la surveillance du journal des requêtes lentes et du journal des erreurs en créant un CloudWatch tableau de bord et en utilisant CloudWatch Logs Insights, consultez le billet de blog Création d'un tableau de CloudWatch bord Amazon pour surveiller Amazon RDS et Amazon Aurora MySQL.

Pistes d'audit

La piste d'audit (ou journal d'audit) fournit un enregistrement chronologique pertinent pour la sécurité des événements survenus dans votre environnement. Compte AWS Il inclut des événements pour Amazon RDS, qui fournissent des preuves documentaires de la séquence d'activités ayant affecté votre base de données ou votre environnement cloud. Dans Amazon RDS for MySQL ou MariaDB, l'utilisation de la piste d'audit implique :

- Surveillance du journal d'audit de l'instance de base de données
- Surveillance des appels d'API Amazon RDS AWS CloudTrail

Pour une instance de base de données Amazon RDS, les objectifs de l'audit sont généralement les suivants :

- Faciliter la responsabilisation dans les domaines suivants :
  - Modifications effectuées sur le paramètre ou la configuration de sécurité

  - Actions effectuées dans un schéma, une table ou une ligne de base de données, ou actions affectant un contenu spécifique
- Détection et investigation des intrusions
- Détection et investigation des activités suspectes
- Détection des problèmes d'autorisation ; par exemple, pour identifier les violations des droits d'accès par des utilisateurs réguliers ou privilégiés

La piste d'audit de base de données tente de répondre à ces questions typiques : qui a consulté ou modifié les données sensibles de votre base de données ? Quand est-ce que cela s'est produit ? D'où un utilisateur spécifique a-t-il accédé aux données ? Les utilisateurs privilégiés ont-ils abusé de leurs droits d'accès illimités ?

MySQL et MariaDB implémentent la fonctionnalité de journal d'audit des instances de base de données à l'aide du plugin d'audit MariaDB. Ce plugin enregistre les activités de la base de données, telles que la connexion des utilisateurs à la base de données et les requêtes exécutées sur la base de données. L'enregistrement de l'activité de la base de données est stocké dans un fichier journal. Pour accéder au journal d'audit, l'instance de base de données doit utiliser un groupe d'options personnalisé avec l'option `MARIADB_AUDIT_PLUGIN`. Pour plus d'informations, consultez la prise en charge du plug-in d'audit MariaDB pour MySQL dans la documentation Amazon RDS. Les enregistrements du journal d'audit sont stockés dans un format spécifique, tel que défini par le plugin. Vous trouverez plus de détails sur le format du journal d'audit dans la documentation du serveur MariaDB.

La piste AWS Cloud d'audit de votre AWS compte est fournie par le AWS CloudTrailservice. CloudTrail capture les appels d'API pour Amazon RDS sous forme d'événements. Toutes les actions Amazon RDS sont enregistrées. CloudTrail fournit un enregistrement des actions effectuées dans Amazon RDS par un utilisateur, un rôle ou un autre AWS service. Les événements incluent les actions effectuées dans la console de AWS gestion AWS CLI, AWS SDKs et APIs.

exemple

Dans un scénario d'audit classique, vous devrez peut-être combiner les AWS CloudTrail traces avec le journal d'audit de la base de données et la surveillance des événements Amazon RDS. Par exemple, vous pouvez avoir un scénario dans lequel les paramètres de base de données de votre instance de base de données Amazon RDS (par exemple,`database-1`) ont été modifiés et votre

tâche consiste à identifier qui a effectué la modification, ce qui a été modifié et quand le changement s'est produit.

Pour accomplir cette tâche, procédez comme suit :

1. Répertoriez les événements Amazon RDS survenus dans l'instance de base de données `database-1` et déterminez s'il existe un événement dans la catégorie `configuration change` contenant le message`Finished updating DB parameter group`.

```
$ aws rds describe-events --source-identifier database-1 --source-type db-instance
{
    "Events": [
        {
            "SourceIdentifier": "database-1",
            "SourceType": "db-instance",
            "Message": "Finished updating DB parameter group",
            "EventCategories": [
                "configuration change"
            ],
            "Date": "2022-12-01T09:22:40.413000+00:00",
            "SourceArn": "arn:aws:rds:eu-west-3:111122223333:db:database-1"
        }
    ]
}
```

2. Identifiez le groupe de paramètres de base de données utilisé par l'instance de base de données :

```
$ aws rds describe-db-instances --db-instance-identifier database-1 --query
 'DBInstances[*].[DBInstanceIdentifier,Engine,DBParameterGroups]'
[
    [
        "database-1",
        "mariadb",
        [
            {
                "DBParameterGroupName": "mariadb10-6-test",
                "ParameterApplyStatus": "pending-reboot"
            }
        ]
    ]
]
```

3. Utilisez le AWS CLI pour rechercher CloudTrail des événements dans la région où `database-1` est déployé, pendant la période autour de l'événement Amazon RDS découvert à l'étape 1, et où`EventName=ModifyDBParameterGroup`.

```
$ aws cloudtrail --region eu-west-3 lookup-events --lookup-attributes
 AttributeKey=EventName,AttributeValue=ModifyDBParameterGroup --start-time
 "2022-12-01, 09:00 AM" --end-time "2022-12-01, 09:30 AM"

{
    "eventVersion": "1.08",
    "userIdentity": {
        "accountId": "111122223333",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "sessionContext": {
            "sessionIssuer": {
                "type": "Role",
                "principalId": "AIDACKCEVSQ6C2EXAMPLE",
                "arn": "arn:aws:iam::111122223333:role/Role1",
                "accountId": "111122223333",
                "userName": "User1"
            }
        }
    },
    "eventTime": "2022-12-01T09:18:19Z",
    "eventSource": "rds.amazonaws.com",
    "eventName": "ModifyDBParameterGroup",
    "awsRegion": "eu-west-3",
    "sourceIPAddress": "AWS Internal",
    "userAgent": "AWS Internal",
    "requestParameters": {
        "parameters": [
            {
                "isModifiable": false,
                "applyMethod": "pending-reboot",
                "parameterName": "innodb_log_buffer_size",
                "parameterValue": "8388612"
            },
            {
                "isModifiable": false,
                "applyMethod": "pending-reboot",
                "parameterName": "innodb_write_io_threads",
                "parameterValue": "8"
            }
```

```
        ],
        "dBParameterGroupName": "mariadb10-6-test"
    },
    "responseElements": {
        "dBParameterGroupName": "mariadb10-6-test"
    },
    "requestID": "fdf19353-de72-4d3d-bf29-751f375b6378",
    "eventID": "0bba7484-0e46-4e71-93a8-bd01ca8386fe",
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "111122223333",
    "eventCategory": "Management",
    "sessionCredentialFromConsole": "true"
}
```


L' CloudTrail événement révèle que `User1` le rôle `Role1` du AWS compte 111122223333 a modifié le groupe `mariadb10-6-test` de paramètres de base de données utilisé par l'instance de base de données sur. `database-1` `2022-12-01 at 09:18:19` h Deux paramètres ont été modifiés et définis sur les valeurs suivantes :

- `innodb_log_buffer_size = 8388612`
- `innodb_write_io_threads = 8`

Fonctionnalités supplémentaires CloudTrail et CloudWatch journaux

Vous pouvez résoudre les incidents opérationnels et de sécurité survenus au cours des 90 derniers jours en consultant l'historique des événements sur la CloudTrail console. Pour prolonger la période de rétention et tirer parti des fonctionnalités de requête supplémentaires, vous pouvez utiliser AWS CloudTrail Lake. Avec AWS CloudTrail Lake, vous pouvez conserver les données d'événements dans un magasin de données d'événements pendant sept ans au maximum. En outre, le service prend en charge les requêtes SQL complexes qui offrent une vue plus approfondie et plus personnalisable des événements que les vues fournies par de simples recherches de valeurs-clés dans l'historique des événements.

Pour surveiller vos pistes d'audit, définir des alarmes et recevoir des notifications lorsqu'une activité spécifique se produit, vous devez configurer CloudTrail pour envoyer ses enregistrements de suivi à CloudWatch Logs. Une fois que les enregistrements de suivi sont stockés sous forme de CloudWatch journaux, vous pouvez définir des filtres métriques pour évaluer les événements du journal en

fonction des termes, des phrases ou des valeurs, et attribuer des métriques aux filtres métriques. En outre, vous pouvez créer des CloudWatch alarmes générées en fonction des seuils et des périodes que vous spécifiez. Par exemple, vous pouvez configurer des alarmes qui envoient des notifications aux équipes responsables, afin qu'elles puissent prendre les mesures appropriées. Vous pouvez également configurer CloudWatch pour exécuter automatiquement une action en réponse à une alarme.

Alerte

Les alertes sont l'une des sources d'informations les plus importantes en matière de sécurité, de disponibilité, de performance et de fiabilité de votre infrastructure informatique et de vos services informatiques. Ils notifient et informent vos équipes informatiques des menaces de sécurité continues, des pannes, des problèmes de performance ou des défaillances du système.

La bibliothèque d'infrastructure informatique (ITIL), en particulier les pratiques de gestion des services informatiques (ITSM), place les alertes automatisées au centre des meilleures pratiques de surveillance, de gestion des événements et de gestion des incidents.

L'alerte en cas d'incident se produit lorsque les outils de surveillance génèrent des alertes pour informer votre équipe et les outils automatisés (pour les éléments automatiquement exploitables) des modifications, des actions à haut risque ou des défaillances de l'environnement informatique. Les alertes informatiques constituent la première ligne de défense contre les pannes ou les modifications du système susceptibles de se transformer en incidents majeurs. En surveillant automatiquement les systèmes et en générant des alertes en cas de panne et de modifications risquées, les équipes informatiques peuvent minimiser les temps d'arrêt et réduire les coûts élevés qui en découlent.

En tant que meilleures pratiques, le AWS Well-Architected Framework prescrit que vous utilisiez la surveillance pour générer des notifications basées sur des alarmes, et que vous surveilliez et alertiez de manière proactive. Utilisez un service de surveillance tiers CloudWatch ou utilisez un service de surveillance tiers pour définir des alarmes indiquant lorsque les mesures dépassent les limites attendues.

L'objectif de la gestion des alertes est d'établir des procédures efficaces et standardisées pour gérer les événements et incidents liés à l'informatique par le biais de la journalisation, de la classification, de la définition et de la mise en œuvre des actions, de la clôture et des activités d'examen post-incident.

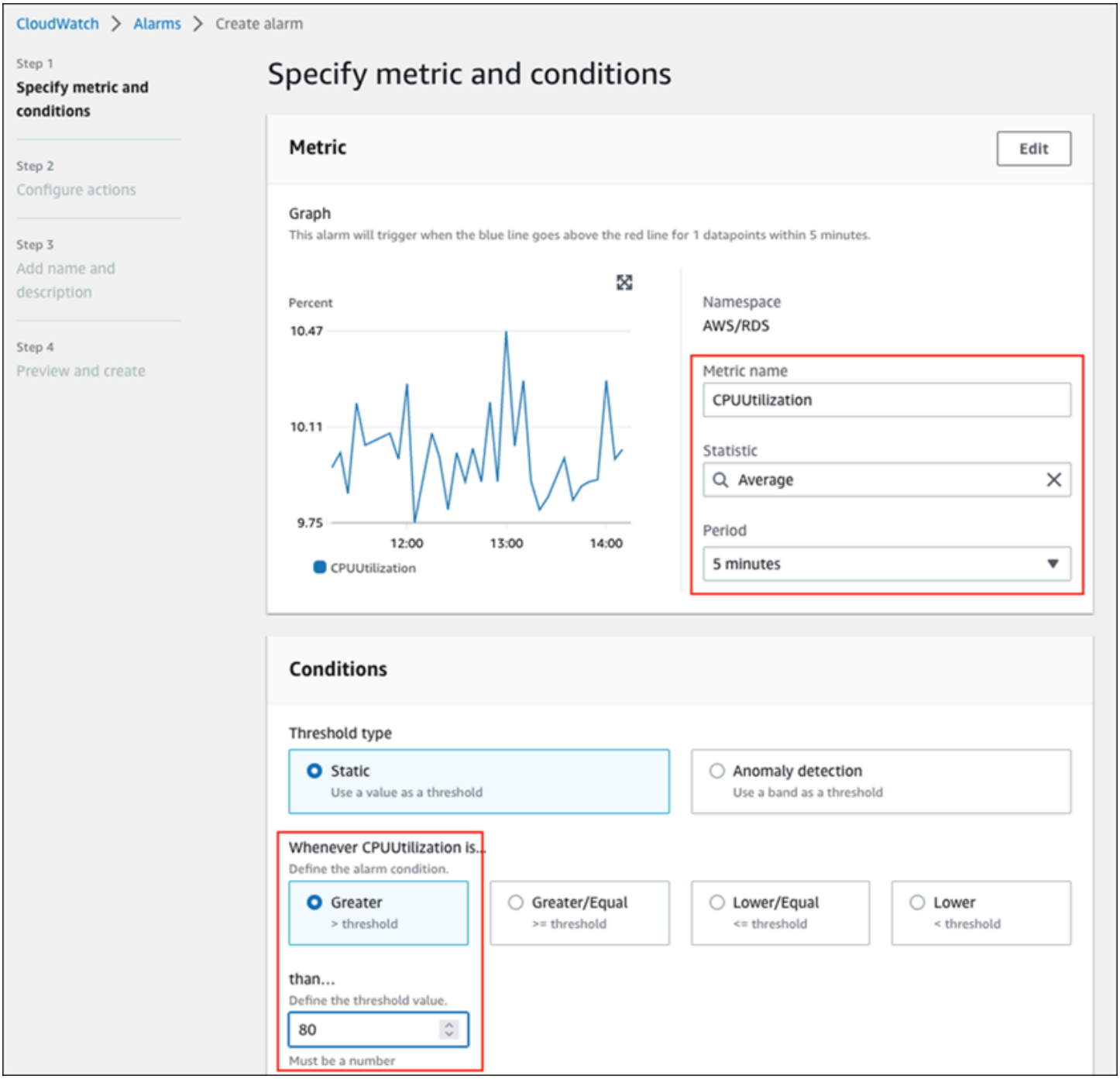
Sections

- CloudWatch alarmes
- EventBridge règles
- Spécification des actions, activation et désactivation des alarmes

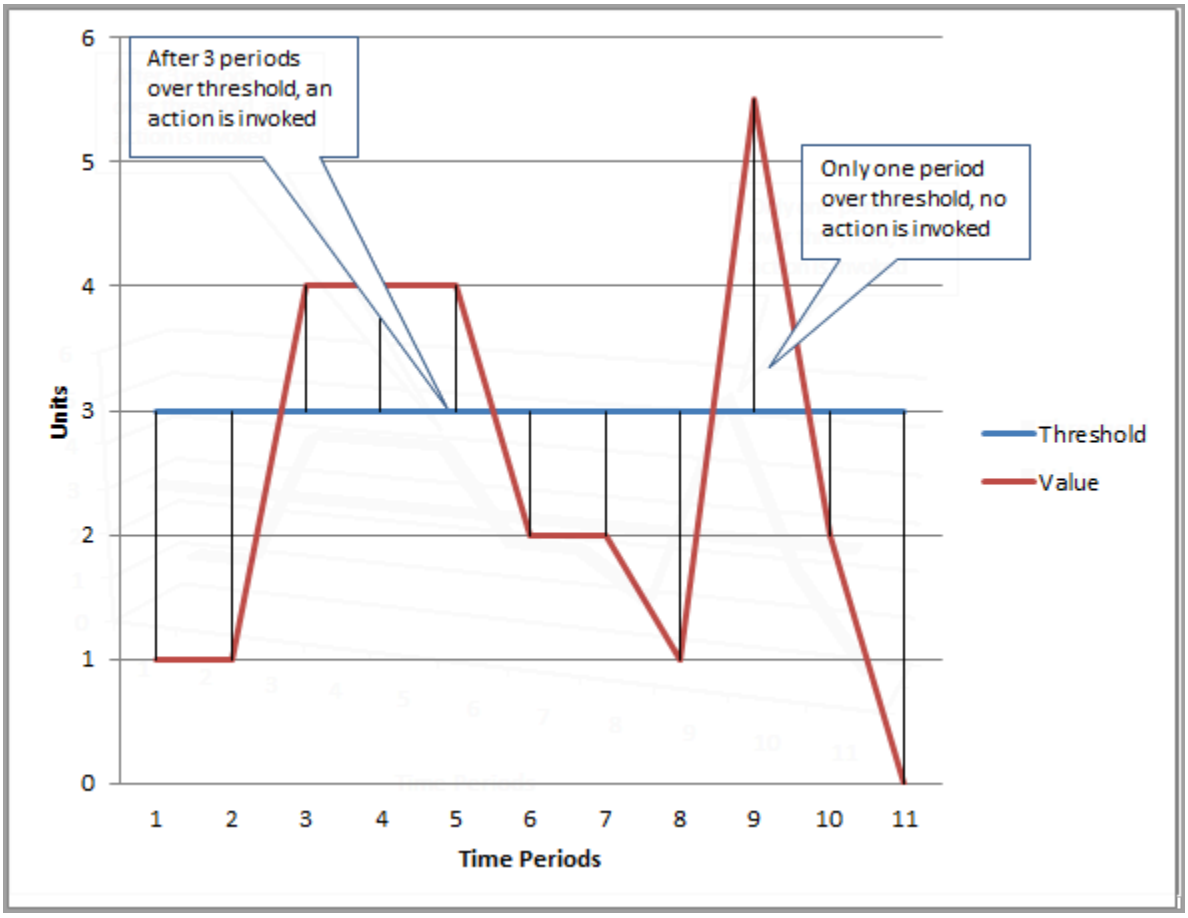
CloudWatch alarmes

Lorsque vous utilisez vos instances de base de données Amazon RDS, vous souhaitez surveiller et générer des alertes sur différents types de métriques, d'événements et de traces. Pour les bases de données MySQL et MariaDB, les principales sources d'informations sont les métriques des instances de base de données, les métriques du système d'exploitation, les événements, les journaux et les pistes d'audit. Nous vous recommandons d'utiliser des CloudWatch alarmes pour surveiller une seule métrique sur une période que vous spécifiez.

L'exemple suivant montre comment définir une alarme qui surveille la `CPUUtilization` métrique (pourcentage d'utilisation du processeur) sur toutes vos instances de base de données Amazon RDS. Vous configurez l'alarme pour qu'elle soit déclenchée si l'utilisation du processeur sur une instance de base de données est supérieure à 80 % pendant la période d'évaluation de 5 minutes.



Cela signifie que l'alarme passe à l'`ALARM`état si l'une de vos bases de données connaît une utilisation élevée du processeur (plus de 80 %) pendant 5 minutes ou plus. L'alarme reste active si `OK` le processeur atteint parfois un taux d'utilisation supérieur à 80 % pendant une courte période, puis retombe en dessous du seuil. Le graphique suivant illustre cette logique.



CloudWatch les alarmes prennent en charge les alarmes métriques et composites.

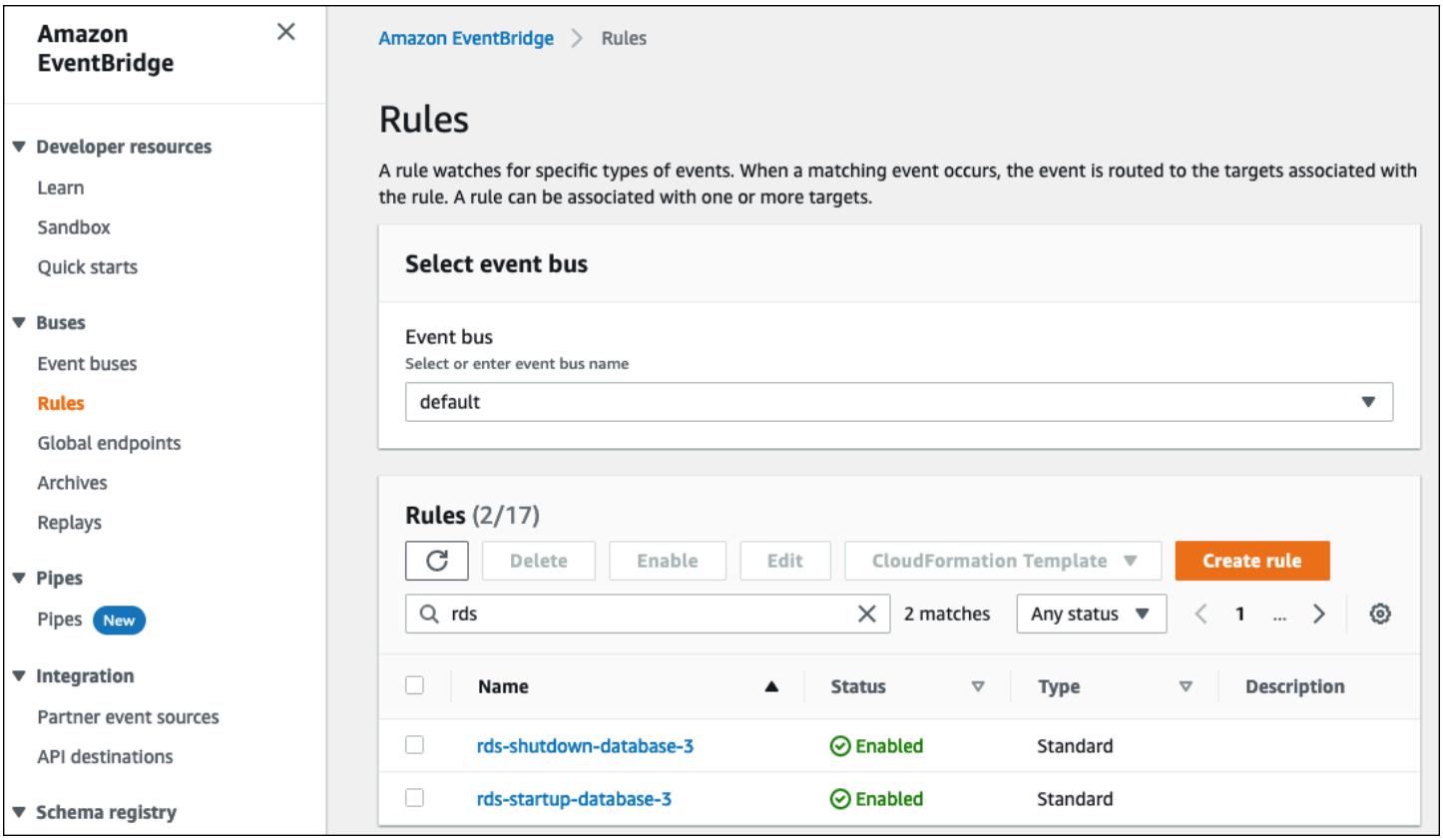
- Une alarme métrique surveille une seule CloudWatch métrique et peut exécuter des expressions mathématiques sur la métrique. Une alarme métrique peut envoyer des messages Amazon SNS, qui, à leur tour, peuvent effectuer une ou plusieurs actions en fonction de la valeur de la métrique par rapport à un seuil donné sur un certain nombre de périodes.
- Une alarme composite est basée sur une expression de règle, qui évalue l'état de plusieurs alarmes et passe à l'`ALARM`état uniquement si toutes les conditions de la règle sont remplies. Les alarmes composites sont généralement utilisées pour réduire le nombre d'alertes inutiles. Par exemple, vous pouvez avoir une alarme composite contenant plusieurs alarmes métriques configurées pour ne jamais effectuer d'actions. L'alarme composite enverrait une alerte lorsque toutes les alarmes métriques individuelles de l'alarme composite se trouvent déjà dans `ALARM`

CloudWatch les alarmes ne peuvent surveiller que CloudWatch les métriques. Si vous souhaitez créer une alarme en fonction des erreurs, des requêtes lentes ou des journaux généraux, vous devez créer des CloudWatch métriques à partir des journaux. Vous pouvez y parvenir, comme indiqué

précédemment dans les sections Surveillance du système d'exploitation et Événements, journaux et pistes d'audit, en utilisant des filtres pour créer des métriques à partir des événements du journal. De même, pour être alerté sur les métriques de surveillance améliorée, vous devez créer des filtres CloudWatch de métriques dans CloudWatch Logs.

EventBridge règles

Les événements Amazon RDS sont transmis à Amazon EventBridge, et vous pouvez utiliser des EventBridge règles pour réagir à ces événements. Par exemple, vous pouvez créer des EventBridge règles qui vous avertiront et effectueront une action si une instance de base de données spécifique s'arrête ou démarre, comme le montre l'écran suivant.

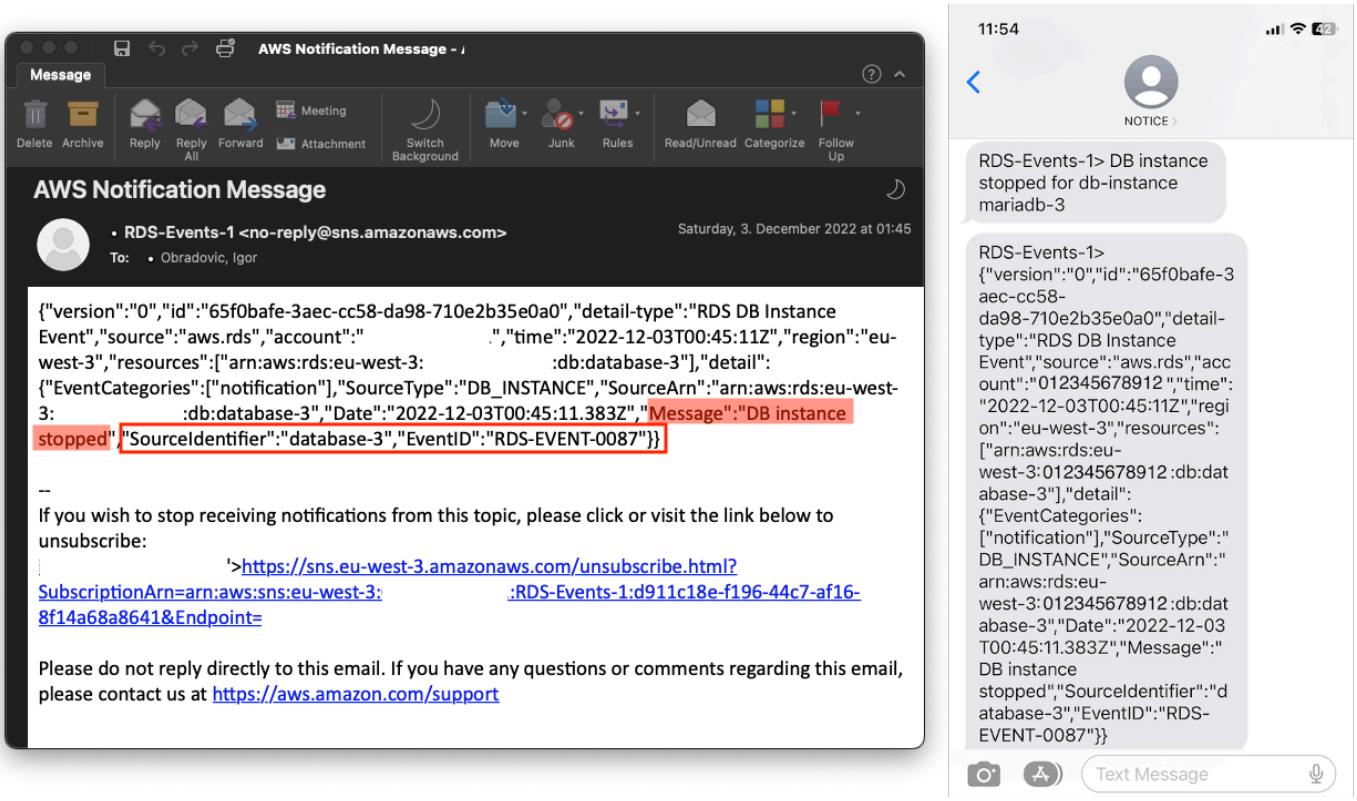


La règle qui détecte l'`The DB instance has been stopped`événement possède l'ID `RDS-EVENT-0087` d'événement Amazon RDS. Vous définissez donc la `Event Pattern` propriété de la règle comme suit :

```
{
  "source": ["aws.rds"],
  "detail-type": ["RDS DB Instance Event"],
```

```
  "detail": {
    "SourceArn": ["arn:aws:rds:eu-west-3:111122223333:db:database-3"],
    "EventID": ["RDS-EVENT-0087"]
  }
}
```

Cette règle surveille `database-3` uniquement l'instance de base de données et surveille l'`RDS-EVENT-0087`événement. Lorsqu'il EventBridge détecte l'événement, il l'envoie à une ressource ou à un point de terminaison, appelé cible. C'est ici que vous pouvez spécifier l'action que vous souhaitez effectuer si l'instance Amazon RDS s'arrête. Vous pouvez envoyer l'événement à de nombreuses cibles possibles, notamment un sujet SNS, une file d'attente Amazon Simple Queue Service (Amazon SQS), une AWS Lambda fonction, une AWS Systems Manager automatisation, une tâche, Amazon AWS Batch API Gateway, un plan de réponse dans Incident Manager, une fonctionnalité de AWS Systems Manager, etc. Par exemple, vous pouvez créer une rubrique SNS qui enverra un e-mail de notification et un SMS, et affecter cette rubrique SNS comme cible de la EventBridge règle. Si l'instance de base de données Amazon RDS `database-3` a été arrêtée, Amazon RDS envoie l'événement `RDS-EVENT-0087` à EventBridge l'endroit où il est détecté. EventBridgeappelle ensuite la cible, qui est le sujet SNS. La rubrique SNS est configurée pour envoyer un e-mail (comme indiqué dans l'illustration suivante) et un SMS.



Spécification des actions, activation et désactivation des alarmes

Vous pouvez utiliser une CloudWatch alarme pour spécifier les actions que l'alarme doit effectuer lorsqu'elle passe de l'état `OK`à `ALARM`, et à l'`INSUFFICIENT_DATA`état. CloudWatch intègre des rubriques SNS et plusieurs catégories d'actions supplémentaires qui ne sont pas applicables aux métriques Amazon RDS, telles que les actions Amazon Elastic Compute Cloud EC2 (Amazon) ou les actions de groupe Amazon EC2 Auto Scaling. EventBridge est généralement utilisé pour écrire des règles et définir des cibles qui prennent des mesures lorsque l'alarme est déclenchée pour les métriques Amazon RDS. CloudWatch envoie des événements à EventBridge chaque fois qu'une CloudWatch alarme change d'état. Vous pouvez utiliser ces événements de changement d'état d'alarme pour déclencher une cible d'événement EventBridge. Pour plus d'informations, consultez la section Événements d'alarme et EventBridge la CloudWatch documentation.

Vous devrez peut-être également gérer les alarmes, par exemple pour désactiver automatiquement une alarme lors de modifications de configuration planifiées ou de tests, puis réactiver l'alarme lorsque l'action planifiée est terminée. Par exemple, si vous avez une mise à niveau planifiée du logiciel de base de données qui nécessite une interruption de service et que des alarmes seront activées en cas d'indisponibilité de la base de données, vous pouvez désactiver et activer les alarmes à l'aide des actions DisableAlarmActionsde l'API et/ou des enable-alarm-actionscommandes disable-alarm-actionset du AWS CLI. EnableAlarmActions Vous pouvez également consulter l'historique de l'alarme sur la CloudWatch console ou en utilisant l'action de l'DescribeAlarmHistoryAPI ou la describe-alarm-historycommande du AWS CLI. CloudWatch conserve l'historique des alarmes pendant deux semaines. Sur la CloudWatch console, vous pouvez choisir le menu Favoris et récents dans le volet de navigation pour définir et accéder à vos alarmes préférées et aux alarmes les plus récemment visitées.

Prochaines étapes et ressources

Pour plus d'informations sur la migration de vos bases de données relationnelles vers le AWS Cloud, consultez la stratégie suivante sur le site Web des directives AWS prescriptives :

- Migration strategy for relational databases

Vous pouvez explorer les modèles de migration de bases de données dans AWS Prescriptive Guidance pour step-by-step obtenir des instructions concernant vos bases de données relationnelles spécifiques exécutées dans le AWS Cloud, y compris les tâches liées à la surveillance, à la migration et à la gestion des données.

Pour des ressources supplémentaires, consultez les rubriques suivantes :

- Guide de l'utilisateur d'Amazon Relational Database Service
- Guide de CloudWatch l'utilisateur Amazon
- Amazon RDS FAQs
- Performance Insights FAQs
- Fournissez les indicateurs de mesure Amazon RDS Performance Insights à un fournisseur de services tiers de surveillance des performances des applications à l'aide d'Amazon CloudWatch Metrics Stream (article de AWS blog)
- Création d'un CloudWatch tableau de bord Amazon pour surveiller Amazon RDS et Amazon Aurora MySQL (article de AWS blog)
- Optimisation d'Amazon RDS pour MySQL avec Performance Insights AWS (article de blog)

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
Informations mises à jour sur Performance Insights	Mise à jour de la section sur la publication des indicateurs Performance Insights afin CloudWatch d'y intégrer les informations les plus récentes.	11 mars 2025
Informations mises à jour sur les exportateurs	Mise à jour des informations sur les exportateurs et ajout de directives pour le choix d'un exportateur.	13 juin 2024
Publication initiale	—	30 juin 2023

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'un 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 ouBitbucket 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/CDpeut 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).

environment

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.

époppée

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 clics 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 troncs 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 replateforme. 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.

IIo 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' Comptes AWS 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 télémétrique 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.