Modélisation de données avec Amazon DynamoDB

AWS Conseils prescriptifs

AWS Conseils prescriptifs: Modélisation de données avec Amazon DynamoDB

Table of Contents

Modélisation de données avec Amazon DynamoDB

Processus, modèles et meilleures pratiques

Amazon Web Services (contributeurs)

Décembre 2023 (historique du document)

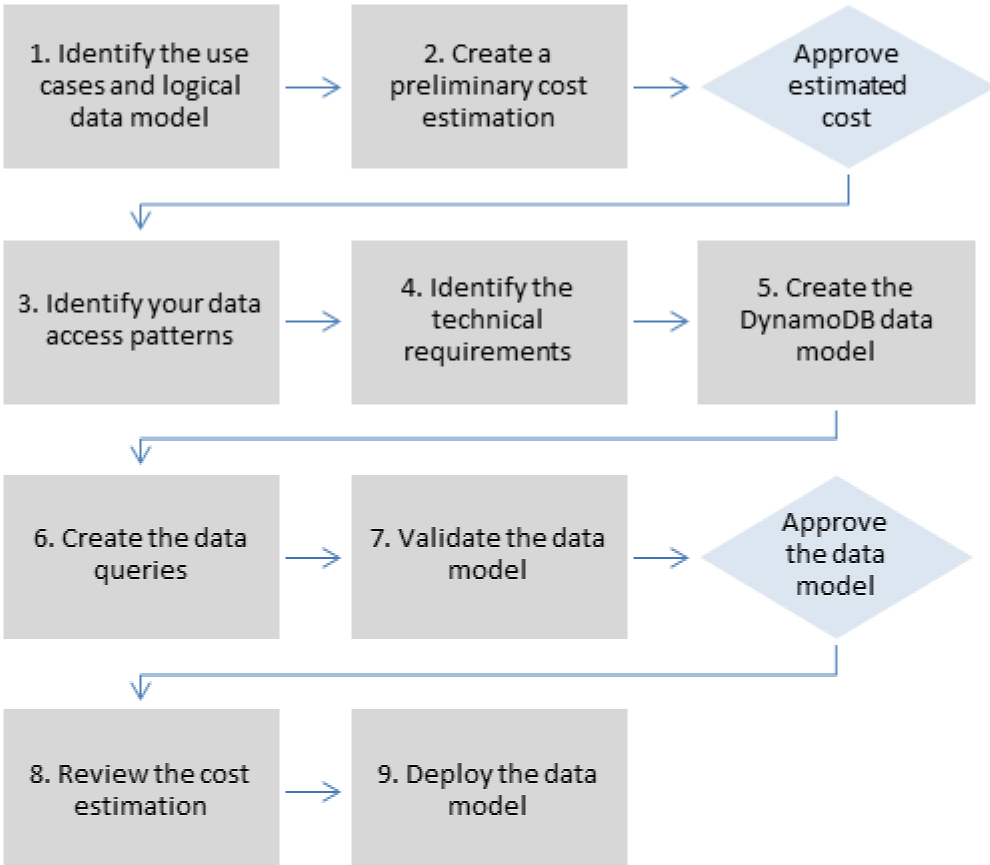
Les bases de données NoSQL fournissent des schémas flexibles pour créer des applications modernes. Ils sont largement reconnus pour leur facilité de développement, leurs fonctionnalités et leurs performances à grande échelle. Amazon DynamoDB fournit des performances rapides et prévisibles ainsi qu'une évolutivité sans faille pour les bases de données NoSQL dans le cloud Amazon Web Services ().AWS En tant que service de base de données entièrement géré, DynamoDB vous aide à vous décharger des charges administratives liées à l'exploitation et au dimensionnement d'une base de données distribuée. Vous n'avez pas à vous soucier de l'approvisionnement, de l'installation et de la configuration du matériel, de la réplication, de l'application de correctifs logiciels ou de la mise à l'échelle du cluster.

La conception de schéma NoSQL nécessite une approche différente de celle d'un système de gestion de base de données relationnelle (RDBMS) traditionnel. Le modèle de données RDBMS se concentre sur la structure des données et leurs relations avec les autres données. La modélisation des données NoSQL se concentre sur les modèles d'accès, c'est-à-dire sur la manière dont l'application va consommer les données. Elle stocke donc les données de manière à permettre des opérations de requête simples. Pour un SGBDR tel que Microsoft SQL Server ou IBM Db2, vous pouvez créer un modèle de données normalisé sans trop vous soucier des modèles d'accès. Vous pouvez étendre le modèle de données pour prendre en charge vos modèles et vos requêtes ultérieurement.

Ce guide présente un processus de modélisation des données pour l'utilisation de DynamoDB qui fournit les exigences fonctionnelles, les performances et les coûts effectifs. Ce guide est destiné aux ingénieurs de bases de données qui prévoient d'utiliser DynamoDB comme base de données opérationnelle pour leurs applications exécutées sur. AWS AWS Professional Services a utilisé le processus recommandé pour aider les entreprises à modéliser les données DynamoDB pour différents cas d'utilisation et charges de travail.

Flux de processus de modélisation des données

Nous recommandons le processus suivant lors de la modélisation de données à l'aide d'Amazon DynamoDB. Les étapes sont décrites en détail plus loin dans ce guide.



Matrice RACI

Certaines organisations utilisent une matrice d'attribution des responsabilités (également appelée matrice RACI) pour décrire les différents rôles impliqués dans un projet ou un processus métier spécifique. Ce guide propose une matrice RACI qui pourrait aider votre entreprise à identifier les bonnes personnes et les bonnes responsabilités pour le processus de modélisation des données DynamoDB. Pour chaque étape du processus, il répertorie les parties prenantes et leur implication :

- R — responsable de l'exécution de l'étape
- A — responsable de l'approbation et de la signature du travail

- C — consulté pour fournir des informations sur une tâche
- Je suis informé des progrès réalisés, mais je ne suis pas directement impliqué dans la tâche

En fonction de la structure de votre organisation et de votre équipe de projet, les rôles décrits dans la matrice RACI suivante peuvent être exécutés par la même partie prenante. Dans certaines situations, les parties prenantes sont à la fois responsables et responsables de mesures spécifiques. Par exemple, les ingénieurs de base de données peuvent être chargés à la fois de créer et d'approuver le modèle de données, car il s'agit de leur domaine de compétence.

Étape du processus	Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
1. Identifier les cas d'utilisation et le modèle de données logique	C	R/A	I	R		
2. Création d'une estimation préliminaire des coûts	C	A	I	R		
3. Identifiez vos modèles d'accès aux données	C	A	I	R		
4. Identifier les	C	C	A	R		

Étape du processus	Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
exigences techniques						
5. Création du modèle de données DynamoDB	I	I	I	R/A		
6. Création des requêtes de données	I	I	I	R/A	R	
7. Valider le modèle de données	A	R	I	C		
8. Passez en revue l'estimation des coûts	C	A	I	R		
9. Déployer le modèle de données DynamoDB	I	I	C	C		R/A

Étapes du processus de modélisation des données

Cette section détaille chaque étape du processus de modélisation des données recommandé pour Amazon DynamoDB.

Rubriques

Étape 1. Identifier les cas d'utilisation et le modèle de données logique

Objectifs

- Rassemblez les besoins de l'entreprise et les cas d'utilisation qui nécessitent une base de données NoSQL.
- Définissez le modèle de données logique à l'aide d'un diagramme entité-relation (ER).

Processus

- Les analystes commerciaux interrogent les utilisateurs professionnels pour identifier les cas d'utilisation et les résultats attendus.
- L'ingénieur de base de données crée le modèle de données conceptuel.
- L'ingénieur de base de données crée le modèle de données logique.

- L'ingénieur de base de données recueille des informations sur la taille des éléments, le volume de données et le débit de lecture et d'écriture attendu.

Outils et ressources

- Évaluation des exigences commerciales (voir modèle)
- Matrice des modèles d'accès (voir modèle)
- Votre outil préféré pour créer des diagrammes

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
C	R/A	I	R		

Outputs

- Cas d'utilisation documentés et exigences commerciales
- Modèle de données logique (diagramme ER)

Étape 2. Création d'une estimation préliminaire des coûts

Objectif

- Développez une estimation préliminaire des coûts pour DynamoDB.

Processus

- L'ingénieur de base de données crée l'analyse initiale des coûts à l'aide des informations disponibles et des exemples présentés sur la page de tarification de DynamoDB.
  - Créez une estimation des coûts pour la capacité à la demande (voir exemple).

- Créez une estimation des coûts pour la capacité allouée (voir exemple).
  - Pour le modèle de capacité provisionnée, obtenez le coût estimatif à l'aide du calculateur et appliquez un discount pour la capacité réservée.
- Comparez les coûts estimés des deux modèles de capacité.
- Créez une estimation pour tous les environnements (Dev, Prod, QA).
- L'analyste commercial examine et approuve ou rejette l'estimation préliminaire des coûts.

Outils et ressources

- AWS Calculateur de prix

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
C	A	I	R		

Outputs

- Estimation préliminaire des coûts

Étape 3. Identifiez vos modèles d'accès aux données

Les modèles d'accès ou de requêtes définissent la manière dont les utilisateurs et le système accèdent aux données pour répondre aux besoins de l'entreprise.

Objectif

- Documentez les modèles d'accès aux données.

Processus

- Un ingénieur de base de données et un analyste commercial interrogent les utilisateurs finaux pour déterminer comment les données seront interrogées à l'aide du modèle de matrice des modèles d'accès aux données.
  - Pour les nouvelles applications, ils passent en revue les témoignages des utilisateurs concernant les activités et les objectifs. Ils documentent les cas d'utilisation et analysent les modèles d'accès requis par ces cas d'utilisation.
  - Pour les applications existantes, ils analysent les journaux de requêtes pour découvrir comment les utilisateurs utilisent actuellement le système et pour identifier les principaux modèles d'accès.
- L'ingénieur de base de données identifie les propriétés suivantes des modèles d'accès :
  - Taille des données : connaître la quantité de données qui sera stockée et demandée en une seule fois permet de déterminer le moyen le plus efficace de partitionner les données (voir le billet de blog).
  - Forme des données : au lieu de remodéliser les données lors du traitement d'une requête (comme c'est le cas dans un SGBDR), une base de données NoSQL organise les données de manière à ce que leur forme dans la base de données corresponde aux requêtes. C'est un élément clé pour augmenter la vitesse et la scalabilité.
  - Vitesse des données : DynamoDB effectue une mise à l'échelle en augmentant le nombre de partitions physiques disponibles pour traiter les requêtes, et en répartissant de manière efficace les données sur ces partitions. Connaître à l'avance les pics de charge des requêtes peut aider à déterminer comment partitionner les données afin d'utiliser au mieux la capacité d'E/S.
- L'utilisateur professionnel hiérarchise les modèles d'accès ou de requête.
  - Les requêtes prioritaires sont généralement les requêtes les plus utilisées ou les plus pertinentes. Il est également important d'identifier les requêtes qui nécessitent une latence de réponse plus faible.

Outils et ressources

- Matrice des modèles d'accès (voir modèle)
- Choisir la bonne clé de partition DynamoDB (blog de base de données AWS)
- Conception NoSQL pour DynamoDB (documentation DynamoDB)

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
C	A	I	R		

Outputs

- Matrice des modèles d'accès aux données

exemple

Schéma d'accès	Priorité	Lire ou écrire	Description	Type (un seul élément, plusieurs éléments ou tous)	Attribut clé	Filtres	Ordre des résultats
Création d'un profil utilisateur	Élevé	Écrire	L'utilisateur crée un nouveau profil	Élément unique	Nom d'utilisateur	N/A	N/A
Mettre à jour le profil utilisateur	Moyen	Écrire	L'utilisateur met à jour son profil	Élément unique	Nom d'utilisateur	Nom d'utilisateur = utilisateur actuel	N/A

Étape 4 : Identifier les exigences techniques

Objectif

- Rassemblez les exigences techniques relatives à la base de données DynamoDB.

Processus

- Les analystes commerciaux interrogent l'utilisateur professionnel et DevOps l'équipe pour recueillir les exigences techniques à l'aide du questionnaire d'évaluation.

Outils et ressources

- Évaluation des exigences techniques (voir exemple de questionnaire)

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
C	C	A	R		

Outputs

- Document sur les exigences techniques

Étape 5. Création du modèle de données DynamoDB

Objectif

- Créez le modèle de données DynamoDB.

Processus

- L'ingénieur de base de données identifie le nombre de tables qui seront nécessaires pour chaque cas d'utilisation. Nous recommandons de conserver le moins de tables possible dans une application DynamoDB.
- Sur la base des modèles d'accès les plus courants, identifiez la clé primaire qui peut être de deux types : une clé primaire avec une clé de partition qui identifie les données, ou une clé primaire avec une clé de partition et une clé de tri. Une clé de tri est une clé secondaire permettant de regrouper et d'organiser les données afin de pouvoir les interroger efficacement au sein d'une partition. Vous pouvez utiliser des touches de tri pour définir des relations hiérarchiques dans vos données que vous pouvez interroger à n'importe quel niveau de la hiérarchie (voir le billet de blog).
  - Création de clés de partition
    - Définissez la clé de partition et évaluez sa distribution.
    - Identifiez le besoin de partitionnement des écritures pour répartir les charges de travail de manière uniforme.
  - Conception de clé de tri
    - Identifiez la clé de tri.
    - Identifiez le besoin d'une clé de tri composite.
    - Identifiez les besoins en matière de contrôle de version.
- Sur la base des modèles d'accès, identifiez les index secondaires répondant aux exigences des requêtes.
  - Identifiez le besoin d'index secondaires locaux (LSIs). Il s'agit d'index dotés de la même clé de partition que la table de base, mais d'une clé de tri différente.
    - Pour les tables avec LSIs, il existe une limite de taille de 10 Go par valeur de clé de partition. Une table LSIs peut stocker un nombre illimité d'éléments, à condition que la taille totale d'une valeur de clé de partition ne dépasse pas 10 Go.
  - Identifiez le besoin d'index secondaires globaux (GSIs). Il s'agit d'index dotés d'une clé de partition et d'une clé de tri qui peuvent être différentes de celles de la table de base (voir le billet de blog).
  - Définissez les projections de l'indice. Pensez à projeter moins d'attributs pour réduire la taille des éléments écrits sur l'index. Au cours de cette étape, vous devez déterminer si vous souhaitez utiliser les éléments suivants :
    - Indices clairsemés

  - Requêtes d'agrégation matérialisées
  - Surcharge GSI
  - Sharding GSI
  - Une réplique finalement cohérente utilisant le GSI
- L'ingénieur de base de données détermine si les données incluront des éléments volumineux. Dans ce cas, ils conçoivent la solution en utilisant la compression ou en stockant les données dans Amazon Simple Storage Service (Amazon S3).
- L'ingénieur de base de données détermine si des données de séries chronologiques seront nécessaires. Dans ce cas, ils utilisent le modèle de conception des séries chronologiques pour modéliser les données.
- L'ingénieur de base de données détermine si le modèle ER inclut many-to-many des relations. Dans ce cas, ils utilisent un modèle de conception de liste d'adjacence pour modéliser les données.

Outils et ressources

- NoSQL Workbench pour Amazon DynamoDB : fournit des fonctionnalités de modélisation, de visualisation des données, de développement et de test de requêtes pour vous aider à concevoir votre base de données DynamoDB
- Conception NoSQL pour DynamoDB (documentation DynamoDB)
- Choisir la bonne clé de partition DynamoDB (blog de base de données AWS)
- Bonnes pratiques d'utilisation des index secondaires dans DynamoDB (documentation DynamoDB)
- Comment concevoir des index secondaires globaux Amazon DynamoDB (blog de base de données AWS)

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
I	I	I	R/A		

Outputs

- Schéma de table DynamoDB qui répond à vos modèles d'accès et à vos exigences

exemple

La capture d'écran suivante montre NoSQL Workbench.

RetailDatabase | GSI: BundleProducts-VendorProducts-CustomerOrdersByOrderId | GSI: VendorOrdersByStatusDate-ProductInventory | GSI: ProductCatalog-CustomerOrdersByProduct

Primary Key		Attributes					
Partition Key: pk	Sort Key: sk						
P1	B1	GSI1-PK	GSI1-SK	name	desc		
		B1	P1	The Tiki Bundle	Everything you need for an island theme party.		
P4	B2	GSI1-PK	GSI1-SK	name	desc		
		B2	P4	Tiki Bar Set	Be the Mai Tai master with your very own Tiki Bar.		
P2	B1	name	desc	qty	GSI1-PK	GSI1-SK	location
		Tiki Torch	Bamboo tiki torch, 4 ft	6	B1	P2	W1-A9-S10-B52
	B2	name	desc	qty	GSI1-PK	GSI1-SK	location
		Tiki Torch	Bamboo tiki torch, 4 ft	2	B2	P2	W1-A9-S10-B52
	P2	name	desc	qty	location	reorderAt	GSI3-SK
		Tiki Torch	Bamboo tiki torch, 4 ft	656	W1-A9-S10-B52	100	/GardenOutdoor/OutdoorDecor/Lighting/LanternsT
	B1	name	desc	qty	GSI1-PK	GSI1-SK	location
		Tiki Statue - Pele	Tiki of the Hawaiian Fire Goddess Pele, 5 ft.	1	B1	P3	W1-A15-S6-B27

Étape 6. Création des requêtes de données

Objectif

- Créez les requêtes principales pour valider le modèle de données.

Processus

- L'ingénieur de base de données crée manuellement une table DynamoDB dans AWS la région ou sur son ordinateur (DynamoDB Local).
- L'ingénieur de base de données ajoute des exemples de données à la table DynamoDB.

- Un ingénieur de base de données crée des facettes à l'aide du NoSQL Workbench pour Amazon DynamoDB ou du AWS SDK pour Java ou Python afin de créer des exemples de requêtes (voir le billet de blog).

  Les facettes sont similaires à une vue de la table DynamoDB.
- L'ingénieur de base de données et le développeur cloud créent des exemples de requêtes en utilisant le AWS Command Line Interface (AWS CLI) ou le AWS SDK pour la langue préférée.

Outils et ressources

- Un AWS compte actif, pour accéder à la console DynamoDB
- DynamoDB Local (facultatif), si vous souhaitez créer la base de données sur votre ordinateur sans accéder au service Web DynamoDB
- NoSQL Workbench pour Amazon DynamoDB (téléchargement et documentation)
- AWS SDK dans le langage de votre choix (PythonJavaScript, PHP, .NET, Ruby, Java, Go, Node.js, C++ et SAP ABAP)

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
I	I	I	R/A	R	

Outputs

- Code pour interroger la table DynamoDB

Exemples

- Exemples DynamoDB utilisant AWS le SDK pour Java
- Exemples de Python
- JavaScriptexemples

Étape 7. Valider le modèle de données

Objectif

- Assurez-vous que le modèle de données répondra à vos exigences.

Processus

- L'ingénieur de base de données remplit la table DynamoDB avec des exemples de données.
- L'ingénieur de base de données exécute le code pour interroger la table DynamoDB.
- L'ingénieur de base de données collecte les résultats de la requête.
- L'ingénieur de base de données collecte les indicateurs de performance des requêtes.
- L'utilisateur professionnel vérifie que les résultats des requêtes répondent aux besoins de l'entreprise.
- Les analystes commerciaux valident les exigences techniques.

Outils et ressources

- Un AWS compte actif, pour accéder à la console DynamoDB
- DynamoDB Local (facultatif), si vous souhaitez créer la base de données sur votre ordinateur sans accéder au service Web DynamoDB
- AWS SDK dans la langue de votre choix

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
A	R	I	C		

Outputs

- Modèle de données approuvé

Étape 8. Passez en revue l'estimation des coûts

Objectifs

- Définissez le modèle de capacité et estimez les coûts DynamoDB pour affiner l'estimation des coûts à partir de l'étape 2.
- Obtenez l'approbation financière finale de l'analyste commercial et des parties prenantes.

Processus

- L'ingénieur de base de données identifie l'estimation du volume de données.
- L'ingénieur de base de données identifie les exigences en matière de transfert de données.
- L'ingénieur de base de données définit les unités de capacité de lecture et d'écriture requises.
- L'analyste commercial choisit entre les modèles de capacité à la demande et les modèles de capacité provisionnés.
- Un ingénieur de base de données identifie la nécessité d'une mise à l'échelle automatique DynamoDB.
- L'ingénieur de base de données saisit les paramètres dans l'outil Simple Monthly Calculator.
- L'ingénieur de base de données présente l'estimation finale du prix aux parties prenantes de l'entreprise.
- L'analyste commercial et les parties prenantes approuvent ou rejettent la solution.

Outils et ressources

- AWS Calculateur de prix

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
C	A	I	R		

Outputs

- Modèle de capacité
- Estimation des coûts révisée

Étape 9. Déployer le modèle de données

Objectif

- Déployez la table (ou les tables) DynamoDB sur. Région AWS

Processus

- DevOps L'architecte crée un AWS CloudFormation modèle ou un autre outil d'infrastructure sous forme de code (IaC) pour la table (ou les tables) DynamoDB. AWS CloudFormation fournit un moyen automatisé de provisionner et de configurer vos tables et les ressources associées.

Outils et ressources

- [AWS CloudFormation]

RACI

Utilisateur professionnel	Analyste commercial	Architecte de solutions	Ingénieur base de données	Développeur d'applications	DevOps ingénieur
I	I	C	C		R/A

Outputs

- AWS CloudFormation modèle

exemple

```
mySecondDDBTable:
   Type: AWS::DynamoDB::
   Table DependsOn: "myFirstDDBTable"
   Properties:
   AttributeDefinitions:
      -  AttributeName: "ArtistId"
         AttributeType: "S"
      -  AttributeName: "Concert"
         AttributeType: "S"
      -  AttributeName: "TicketSales"
         AttributeType: "S"
   KeySchema:
      -  AttributeName: "ArtistId"
         KeyType: "HASH"
      -  AttributeName: "Concert"
         KeyType: "RANGE"
   ProvisionedThroughput:
      ReadCapacityUnits:
         Ref: "ReadCapacityUnits"
      WriteCapacityUnits:
         Ref: "WriteCapacityUnits"
   GlobalSecondaryIndexes:
      -  IndexName: "myGSI"
         KeySchema:
```

```
          - AttributeName: "TicketSales"
            KeyType: "HASH"
        Projection:
            ProjectionType: "KEYS_ONLY"
        ProvisionedThroughput:
        ReadCapacityUnits:
            Ref: "ReadCapacityUnits"
        WriteCapacityUnits:
            Ref: "WriteCapacityUnits"
Tags:
    -  Key: mykey
       Value: myvalue
```

Modèles

Les modèles fournis dans cette section sont basés sur les données des joueurs de jeu modélisées avec Amazon DynamoDB sur le site Web. AWS

> **Note**
>
> Les tableaux de cette section utilisent MM comme abréviation pour million, et K comme abréviation pour mille.

Rubriques

Modèle d'évaluation des exigences commerciales

Fournissez une description du cas d'utilisation :

> **Description**
>
> Imaginez que vous créez un jeu multijoueur en ligne. Dans votre jeu, des groupes de 50 joueurs rejoignent une session pour jouer à un jeu, qui prend généralement environ 30 minutes. Pendant le jeu, vous devez mettre à jour le record d'un joueur spécifique pour indiquer le temps que le joueur a passé à jouer, ses statistiques ou s'il a gagné la partie. Les utilisateurs veulent voir les parties précédentes auxquelles ils ont joué, soit pour voir les gagnants des parties, soit pour regarder une rediffusion de l'action de chaque partie.

Fournissez des informations sur vos utilisateurs :

Utilisateur	Description	Numéro attendu
Joueur de jeu	Joueur de jeu en ligne.	1 MM

L'équipe de développement	Équipe interne qui utilisera les statistiques du jeu pour améliorer le expérience de jeu.	100

Fournissez des informations sur les sources de données et sur la manière dont les données seront ingérées :

Source	Description	Utilisateur
Jeu en ligne	Les joueurs créeront des profils et lanceront de nouvelles parties.	Joueur de jeu
Application de jeu	L'application de jeu collectera automatiquement des statistiq ues sur les jeux, telles que les heures de début et de fin, le nombre de joueurs, la position de chaque joueur et la carte du jeu.	

Fournissez des informations sur la manière dont les données seront consommées :

Consommateur	Description	Utilisateur
Jeu en ligne	Les joueurs pourront consulter leurs profils et consulter leurs statistiques de jeu.	Joueur de jeu
Analyses de données	L'équipe de développement du jeu extraira les statistiques du jeu pour analyser les données et améliorer l'expérience utilisateur. Les données seront	L'équipe de développement

	exportées depuis le magasin de données et importées dans Amazon S3 pour permettre l'analyse via une application Spark.	

Fournissez une liste des entités et indiquez comment elles sont identifiées :

Nom de l'entité	Description	Identifiant
Joueur de jeu	Stocke des informations telles que l'identification, l'adresse, les données démographiques, les intérêts de chaque utilisate ur (joueur).	Username
Instance de jeu	Fournit des informations sur chaque partie jouée, y compris le créateur, le début, la fin et la carte Yplayed.	Identifiant du jeu
Cartographie des utilisateurs du jeu	Représente les many-to-many relations entre les utilisateurs et les jeux.	ID de jeu et nom d'utilisateur

Créez un modèle ER pour les entités :

Fournissez des statistiques de haut niveau sur les entités :

Nom de l'entité	Nombre estimé d'enregistrements	Taille de l'enregis trement	Remarques
Joueur de jeu	1 MM	< 1 Ko	La plateforme de jeu compte environ 1 million d'utilisateurs.
Instance de jeu	6 MM (100 000 K/jour* 60 jours)	< 1 Ko	En moyenne, il y a 100 000 parties par jour. Nous devons enregistrer les 60 derniers jours.
Cartographie des utilisateurs du jeu	300 MM (6 jeux MM* 50 joueurs)	< 1 Ko	En moyenne, chaque jeu compte 50 joueurs sur lesquels nous devons stocker des informations.

Modèle d'évaluation des exigences techniques

Fournissez des informations sur les types d'ingestion de données :

Type d'ingestion de données	OU/N	Description	Frequency (Fréquence)
Accès aux applications	Y		
Passerelle API	Y		
Streaming de données	N		
Procédé Batch	N		
ETL	N		
Importation de données	N		
Séries chronologiques	N		

Fournissez des informations sur les types de consommation de données :

Type de consommation de données	OU/N	Description	Frequency (Fréquence)
Accès aux applications			
Passerelle API			
Exportation de données			
Analyses de données			

Agrégation de données

Génération de rapports

Rechercher

Streaming de données

ETL

Fournissez des estimations du volume de données :

Nom de l'entité	Nombre estimé d'enregistrements	Taille de l'enregistrement	Volume de données
Joueur de jeu	1 MM	< 1 Ko	~ 1 Go (1 MM* 1 KO)
Instance de jeu	6 MM (100K/jour * 60 jours)	< 1 Ko	~ 6 Go (6 MM* 1 KO)
Cartographie des utilisateurs du jeu	300 MM (6 jeux MM* 50 joueurs)	< 1 Ko	~ 300 Go (300 MM* 1 KO)

> **Note**
>
> La durée de conservation des données est de 60 jours. Après 60 jours, les données doivent être stockées dans Amazon S3 à des fins d'analyse, en utilisant DynamoDB Time to Live (TTL) pour déplacer automatiquement les données de DynamoDB vers Amazon S3.

Répondez aux questions suivantes concernant les modèles temporels :

- Dans quels délais l'application est-elle accessible à l'utilisateur (par exemple, 24 heures sur 24, 7 jours sur 7 ou de 9 h à 17 h en semaine) ?
- Y a-t-il un pic d'utilisation pendant la journée ? Combien d'heures ? Quel est le pourcentage d'utilisation des applications ?

Spécifiez les exigences en matière de débit d'écriture :

Nom de l'entité	Écrivres/jour	Heures/jour	écritures/seconde
Joueur de jeu	10 000 mises à jour	18	< 1
Instance de jeu	300 000	18	< 5
Cartographie des utilisateurs du jeu	1 800 000 000	18	~ 27,777

Remarques

Opérations d'écriture sur Game Player : 1 % des utilisateurs mettent à jour leur profil chaque jour. Nous prévoyons donc 10 000 mises à jour pour 1 000 000 d'utilisateurs.
Opérations d'écriture des instances de jeu : 100 000 jeux/jour. Pour chaque jeu, nous avons au moins 3 opérations d'écriture (à la création, au début et à la fin), soit un total de 300 000 opérations d'écriture.
Opérations d'écriture du mappage des utilisateurs du jeu : 100 000 jeux/jour pour chaque jeu à 50 joueurs. La durée moyenne d'une partie est de 30 minutes et la position du joueur est mise à jour toutes les 5 secondes. Nous estimons une moyenne de 360 mises à jour par joueur, donc le total est de 100 000 * 50 * 360 = 1 800 000 000 d'opérations d'écriture.

Spécifiez les exigences en matière de débit de lecture :

Nom de l'entité	Lectures par jour	Heures/jour	Lectures par seconde
Joueur de jeu	200 000	18	~ 3

Instance de jeu	5 000 000	18	~ 77
Cartographie des utilisateurs du jeu	1 800 000 000	18	~ 27,777

> **Remarques**
>
> Opérations de lecture du Game Player : 20 % des utilisateurs démarrent une partie, donc 1 MM* 0,2 = 200 000.
> Opérations de lecture des instances de jeu : 100 000 jeux/jour. Pour chaque jeu, nous avons au moins 1 opération de lecture par joueur et 50 joueurs par partie, soit un total de 5 000 000 d'opérations de lecture.
> Opérations de lecture du Game User Mapping : 100 000 parties par jour pour 50 joueurs. La durée moyenne d'une partie est de 30 minutes et la position du joueur est mise à jour toutes les 5 secondes. Nous estimons une moyenne de 360 mises à jour par joueur, et chaque mise à jour nécessite une opération de lecture. Le total est donc de 100 000 * 50 * 360 = 1 800 000 000 d'opérations de lecture.

Spécifiez les exigences de latence d'accès aux données :

Opération	99 percentiles	Latence maximale
Lecture	30 millisecondes	100 millisecondes
Write (Écrire)	10 millisecondes	50 millisecondes

Spécifiez les exigences en matière de disponibilité des données :

Exigence	OU/N	Métrique	Remarques
Haute disponibilité	Y	99,9 %	
RTO	Y	1 heure	Objectif en matière de temps de rétabliss ement

RPO	Y	1 heure	Objectif du point de rétablissement
Reprise après sinistre	N		
Réplication des données dans la région	N		
Réplication des données entre régions	N	Latence de 3 secondes	Lequel Régions AWS ?

Spécifiez les exigences de sécurité :

Exigence	OU/N	Remarques
Stockage de données sensibles	N	Informations de santé protégées (PHI), informati ons du secteur des cartes de paiement (PCI), informati ons personnelles identifiables (PII) ?
Chiffrement au repos	Y	
Chiffrement en transit	Y	
Chiffrement côté client	N	
Toute bibliothèque de chiffrement propriétaire ou tierce	N	
Enregistrement des accès aux données	N	
Audit de l'accès aux données	N	

Modèle de modèles d'accès

Collectez et documentez les informations sur les modèles d'accès pour le cas d'utilisation en utilisant les champs suivants :

Champ	Description
Schéma d'accès	Donnez un nom au modèle d'accès.
Description	Fournissez une description plus détaillée du modèle d'accès.
Priorité	Définissez une priorité pour le modèle d'accès (haute, moyenne ou faible). Cela définit les modèles d'accès les plus pertinents pour l'application.
Lire ou écrire	S'agit-il d'un modèle d'accès en lecture ou en écriture ?
Type	Le modèle accède-t-il à un seul élément, à plusieurs éléments ou à tous les éléments ?
Filtre	Le modèle d'accès nécessite-t-il un filtre ?
Tri	Le résultat nécessite-t-il un tri ?

Modèle

Schéma d'accès	Description	Priorité	Lire ou écrire	Type (article unique, multiple articles, ou tous)	Attribut clé	Filtres	Ordre des résultats

Création d'un profil utilisateur	L'utilisateur crée un nouveau profil.	Élevée	Write (Écrire)	Élément unique	Username	S/O	S/O
Mettre à jour le profil utilisateur	L'utilisateur met à jour son profil.	Moyenne	Write (Écrire)	Élément unique	Username	Nom d'utilisateur = utilisateur actuel	S/O
Obtenir le profil de l'utilisateur	L'utilisateur passe en revue son profil.	Élevée	Lecture	Élément unique	Username	Nom d'utilisateur = utilisateur actuel	S/O
Créez un jeu	L'utilisateur crée un nouveau jeu.	Élevée	Write (Écrire)	Élément unique	Identifiant du jeu	S/O	S/O

Trouvez des jeux ouverts	L'utilisateur recherche des jeux ouverts. Les résultats de recherche sont triés par horodatage de début dans l'ordre décroissant.	Élevée	Lecture	Objets multiples		GameStatus = ouvert	Horodatage de début descendant

Trouvez les jeux ouverts par carte	L'utilisateur recherche les parties ouvertes à l'aide d'une carte spécifique, triée par horodatage de début en ordre décroissant commande.	Moyenne	Lecture	Objets multiples		GameStatus = ouvert et carte = XYZ	Horodatage de début descendant
Voir le jeu	L'utilisateur passe en revue les détails d'un jeu.	Élevée	Lecture	Élément unique	Identifiant du jeu	S/O	S/O
Afficher les utilisateurs d'un jeu	L'utilisateur obtient la liste de tous les utilisateurs d'un jeu.	Moyenne	Lecture	Objets multiples		Identifiant du jeu = XYZ	S/O

Associer un utilisateur à un jeu	L'utilisateur rejoint une partie ouverte.	Élevée	Write (Écrire)	Élément unique	GameID et nom d'utilisateur	GameStatus = ouvert	S/O
Lancer un jeu	L'utilisateur démarre une nouvelle partie.	Élevée	Write (Écrire)	Élément unique	Identifiant du jeu	S/O	S/O
Mettre à jour le jeu pour l'utilisateur	Mettez à jour la position de l'utilisateur dans le jeu.	Moyenne	Write (Écrire)	Élément unique	GameID et nom d'utilisateur	S/O	S/O
Mettre à jour le jeu	Fin du jeu ; mise à jour des statistiques.	Moyenne	Write (Écrire)	Élément unique	Identifiant du jeu	S/O	S/O

Trouvez tous les jeux précédents pour un utilisateur	Répertoriez tous les jeux auxquels un utilisateur a joué, classés par date de début du jeu.	Faible	Lecture	Objets multiples	Nom d'utilisateur et GameID	Nom d'utilisateur = utilisateur actuel	Horodatage de début
Exporter des données pour l'analyse des données	L'équipe de développement exécutera un traitement par lots pour exporter les données vers Amazon S3.	Faible	Lecture	Tous	S/O	S/O	S/O

Bonnes pratiques

Envisagez d'appliquer les meilleures pratiques de conception DynamoDB suivantes :

- Conception de clés de partition : utilisez une clé de partition à cardinalité élevée pour répartir la charge de manière uniforme.
- Modèle de conception de liste d'adjacence : utilisez ce modèle de conception pour la gestion one-to-many et many-to-many les relations.
- Index clairsemé — Utilisez un index clairsemé pour vos index secondaires globaux (). GSIs Lorsque vous créez un GSI, vous spécifiez une clé de partition et éventuellement une clé de tri. Seuls les éléments de la table de base contenant une clé de partition de GSI correspondante apparaissent dans l'index fragmenté. Cela permet de rester GSIs plus petit.
- Surcharge d'index : utilisez le même GSI pour indexer différents types d'élément.
- Partitionnement d'écriture de GSI : partitionnez judicieusement afin de distribuer les données entre les partitions pour des requêtes plus efficaces et plus rapides.
- Objets de grande taille : stockez uniquement les métadonnées dans la table, enregistrez le blob dans Amazon S3 et conservez la référence dans DynamoDB. Divisez les éléments volumineux en plusieurs éléments et indexez efficacement à l'aide des clés de tri.

Pour en savoir plus sur les meilleures pratiques de conception, consultez la documentation Amazon DynamoDB.

Exemple de modélisation hiérarchique des données

Les sections suivantes utilisent un exemple de constructeur automobile pour montrer comment utiliser les étapes du processus de modélisation des données pour concevoir un système de gestion des composants à plusieurs niveaux dans DynamoDB.

Rubriques

Étape 1 : Identifier les cas d'utilisation et le modèle de données logique

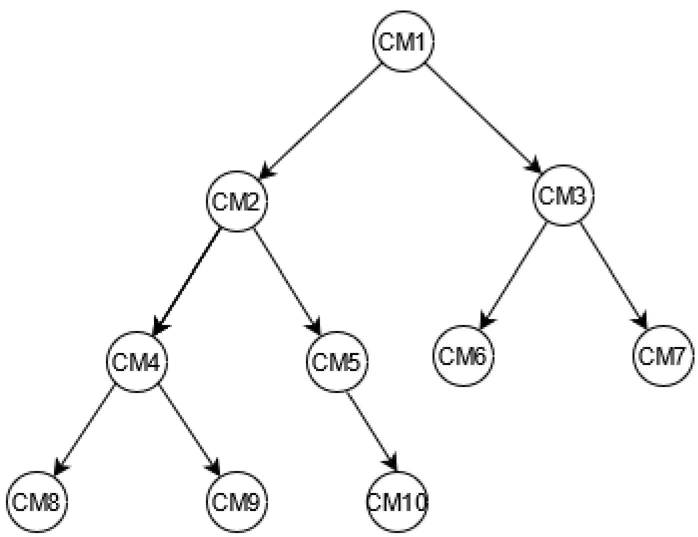
Une entreprise automobile souhaite créer un système de gestion des composants transactionnel pour stocker et rechercher toutes les pièces automobiles disponibles et pour établir des relations entre les différents composants et pièces. Par exemple, une voiture contient plusieurs batteries, chaque batterie contient plusieurs modules de haut niveau, chaque module contient plusieurs cellules et chaque cellule contient plusieurs composants de bas niveau.

Généralement, pour créer un modèle de relation hiérarchique, une base de données orientée graphe telle qu'Amazon Neptune constitue un meilleur choix. Dans certains cas, cependant, Amazon DynamoDB constitue une meilleure alternative pour la modélisation hiérarchique des données en raison de sa flexibilité, de sa sécurité, de ses performances et de son évolutivité.

Par exemple, vous pouvez créer un système dans lequel 80 à 90 % des requêtes sont transactionnelles, et dans lequel DynamoDB convient parfaitement. Dans cet exemple, les 10 à 20 % restants sont relationnels, une base de données de graphes telle que Neptune étant mieux adaptée.

Dans ce cas, l'inclusion d'une base de données supplémentaire dans l'architecture pour répondre à seulement 10 à 20 % des requêtes pourrait augmenter les coûts. Cela alourdit également la charge opérationnelle liée à la maintenance de plusieurs systèmes et à la synchronisation des données. Vous pouvez plutôt modéliser ces 10 à 20 % de requêtes relationnelles dans DynamoDB.

La création d'un schéma d'un exemple d'arborescence pour les composants automobiles peut vous aider à mapper la relation entre eux. Le schéma suivant montre un graphe de dépendance à quatre niveaux. CM1 est le composant de haut niveau de l'exemple de voiture lui-même. Il comporte deux sous-composants pour deux exemples de batteries, CM2 et CM3. Chaque batterie comporte deux sous-composants, à savoir les modules. CM2 possède des modules CM4 et CM5, et CM3 possède des modules CM6 et CM7. Chaque module comporte plusieurs sous-composants, à savoir les cellules. Le CM4 module comporte deux cellules, CM8 et CM9. CM5 comporte une cellule, CM1 0. CM6 et CM7 je n'ai pas encore de cellules associées.



Ce guide utilisera cette arborescence et les identifiants de ses composants comme référence. Un composant supérieur sera appelé parent, tandis qu'un sous-composant sera appelé enfant. Par exemple, le composant supérieur CM1 est le parent de CM2 et CM3. CM2 est le parent de CM4 et CM5. Il représente graphiquement les relations parent-enfant.

Dans l'arborescence, vous pouvez voir le graphique de dépendance complet d'un composant. Par exemple, CM8 dépend de CM4, qui dépend de CM2, qui dépend de CM1. L'arborescence définit le graphique de dépendance complet comme chemin. Un chemin décrit deux choses :

- Le graphique de dépendance
- La position dans l'arborescence

Remplissage des modèles pour répondre aux besoins de l'entreprise :

Fournissez des informations sur vos utilisateurs :

Utilisateur	Description
Employé	Employé interne de l'entreprise automobile qui a besoin d'informations sur les voitures et leurs composants

Fournissez des informations sur les sources de données et sur la manière dont les données seront ingérées :

Source	Description	Utilisateur
Système de gestion	Système qui stockera toutes les données relatives aux pièces automobiles disponibl es et à leurs relations avec les autres composants et pièces.	Employé

Fournissez des informations sur la manière dont les données seront consommées :

Consommateur	Description	Utilisateur
Système de gestion	Récupérez tous les composant s enfants immédiats pour un ID de composant parent.	Employé
Système de gestion	Récupérez une liste récursive de tous les composant s enfants pour un ID de composant.	Employé
Système de gestion	Consultez les ancêtres d'un composant.	Employé

Étape 2 : Création d'une estimation préliminaire des coûts

Il est important de calculer une estimation du coût pour tous les environnements de votre application afin de vérifier si la solution est financièrement viable. Une bonne pratique consiste à effectuer une estimation de haut niveau et à obtenir l'approbation de l'analyste commercial avant de procéder au développement et au déploiement.

- L'ingénieur de base de données crée l'analyse initiale des coûts à l'aide des informations disponibles et des exemples présentés sur la page de tarification de DynamoDB.
  - Créez une estimation des coûts pour la capacité à la demande (voir exemple).
  - Créez une estimation des coûts pour la capacité allouée (voir exemple).
    - Pour le modèle de capacité provisionnée, obtenez le coût estimé à l'aide du calculateur et appliquez la réduction pour la capacité réservée.
  - Comparez les coûts estimés des deux modèles de capacité.
  - Créez une estimation pour tous les environnements (Dev, Prod, QA).
- L'analyste commercial examine et approuve ou rejette l'estimation préliminaire des coûts.

À l'aide de ces valeurs de référence, vous pouvez créer un prix estimatif à soumettre pour approbation. Pour créer le budget, vous pouvez utiliser la page de tarification DynamoDB et le calculateur de prix AWS.

Étape 3 : Identifiez vos modèles d'accès aux données

Cet exemple de cas d'utilisation utilise les modèles d'accès suivants pour gérer les relations entre les différents composants de la voiture.

Schéma d'accès	Priorité	Lire ou écrire	Description	Type	Filtres	Ordre des résultats
Enfant immédiat	Élevé	Lecture	Récupérez tous les composant s enfants immédiats pour un	Plusieurs	`Component ID`	N/A

			ID de composant parent.			
Tous les composants pour enfants	Élevé	Lecture	Récupérez une liste récursive de tous les composants enfants pour un ID de composant.	Plusieurs	`Component ID`	N/A
Ancêtres	Élevé	Lecture	Récupérez les ancêtres d'un composant.	Plusieurs	`Component ID`	N/A

Étape 4 : Identifier les exigences techniques

Cet exemple ne comporte aucune exigence technique spécifique, qui n'entre pas dans le cadre de cet exemple. Dans des cas réels, il est recommandé de terminer cette étape et de valider que toutes les exigences techniques sont satisfaites avant de procéder au développement et au déploiement. Vous pouvez utiliser l'exemple de questionnaire pour effectuer cette étape dans votre analyse de rentabilisation. En outre, nous vous recommandons de valider les quotas du service DynamoDB pour vous assurer qu'il n'existe aucune limite stricte dans la solution que vous avez conçue.

Étape 5 : Création d'un modèle de données DynamoDB

Définissez les clés de partition pour votre table de base et vos index secondaires globaux (GSIs) :

- Conformément aux meilleures pratiques de conception des clés, `ComponentId` utilisez-la comme clé de partition pour la table de base dans cet exemple. Parce qu'il est unique, il `ComponentId` peut offrir une granularité. DynamoDB utilise la valeur de hachage de votre clé de partition pour déterminer la partition dans laquelle les données sont stockées physiquement. L'ID de composant unique génère une valeur de hachage différente, ce qui peut faciliter la distribution des données dans la table. Vous pouvez interroger la table de base à l'aide d'une clé de `ComponentId` partition.

- Pour trouver les enfants immédiats d'un composant, créez un GSI où se `ParentId` trouve la clé de partition et `ComponentId` la clé de tri. Vous pouvez interroger ce GSI en l'utilisant `ParentId` comme clé de partition.
- Pour trouver tous les enfants récursifs d'un composant, créez un GSI où `GraphId` est la clé de partition et `Path` est la clé de tri. Vous pouvez interroger ce GSI en utilisant `GraphId` comme clé de partition et l'opérateur `BEGINS_WITH(Path, "$path")` sur la clé de tri.

	Clé de partition	Clé de tri	Attributs de mappage
Table de base	`ComponentId`		`ParentId`, `GraphId`, `Path`
GSI1	`ParentId`	`ComponentId`	
GSI2	`GraphId`	`Path`	`ComponentId`

Stockage des composants dans la table

L'étape suivante consiste à créer chaque composant sur la table de base DynamoDB. Après avoir inséré tous les composants de l'arbre d'exemple, vous obtenez le tableau de base suivant.

ComponentId	ParentId	GraphId	Chemin
CM1		CM1#1	CM1
CM2	CM1	CM1#1	CM1|CM2
CM3	CM1	CM1#1	CM1|CM3

CM4	CM2	CM1#1	CM1|CM2|CM4
CM5	CM2	CM1#1	CM1|CM2|CM5
CM6	CM3	CM1#1	CM1|CM3|CM6
CM7	CM3	CM1#1	CM1|CM3|CM7
CM8	CM4	CM1#1	CM1|CM2|CM4|CM8
CM9	CM4	CM1#1	CM1|CM2|CM4|CM9
CM10	CM5	CM1#1	CM1| CM2 | CM5 | CM1 0

L' GSI1 indice

Pour vérifier tous les enfants immédiats d'un composant, vous créez un index utilisé `ParentId` comme clé de partition et `ComponentId` comme clé de tri. Le tableau croisé dynamique suivant représente l' GSI1 indice. Vous pouvez utiliser cet index pour récupérer tous les composants enfants immédiats à l'aide d'un ID de composant parent. Par exemple, vous pouvez savoir combien de batteries sont disponibles dans une voiture (CM1) ou quelles cellules sont disponibles dans un module (CM4).

ParentId	ComponentId
CM1	CM2

	CM3
CM2	CM4
	CM5
CM3	CM6
	CM7
CM4	CM8
	CM9
CM5	CM10

L' GSI2 indice

Le tableau croisé dynamique suivant représente l' GSI2 indice. Il est configuré à l'aide de `GraphId` en tant que clé de partition et `Path` comme clé de tri. À l'aide de `GraphI d` et de l'`begins_with`opération sur la touche de tri (`Path`), vous pouvez trouver la lignée complète d'un composant dans une arborescence.

GraphId	Chemin	ComponentId
CM1#1	CM1	CM1
	CM1|CM2	CM2
	CM1|CM3	CM3
	CM1|CM2|CM4	CM4
	CM1|CM2|CM5	CM5
	CM1|CM2|CM4|CM8	CM8
	CM1|CM2|CM4|CM9	CM9
	CM1| CM2 | CM5 | CM1 0	CM10

CM1|CM3|CM6	CM6
CM1|CM3|CM7	CM7

Étape 6 : Création de requêtes de données

Après avoir défini vos modèles d'accès et conçu votre modèle de données, vous pouvez interroger des données hiérarchiques dans la base de données DynamoDB. Afin de réduire les coûts et de garantir les performances, les exemples suivants utilisent uniquement l'opération de requête sans`Scan`.

- Trouvez les ancêtres d'un composant.

  Pour trouver les ancêtres (parent, grand-parent, arrière-grand-parent, etc.) du CM8 composant, interrogez la table de base à l'aide de. `ComponentId = "CM8"` La requête va renvoyer l'enregistrement suivant.

  Pour réduire la taille des données de résultat, vous pouvez utiliser une expression de projection pour renvoyer uniquement l'attribut `Path`.

ComponentId	ParentId	GraphId	Chemin
CM8	CM4	CM1#1	CM1|CM2|CM4|CM8

Chemin
CM1|CM2|CM4|CM8

  Maintenant, divisez le chemin à l'aide du tube (« | ») et prenez les premiers composants N-1 pour obtenir les ancêtres.

  Résultat de la requête : Les ancêtres de CM8 sont CM1, CM2, CM4.

- Trouvez les enfants immédiats d'un composant.

Pour obtenir tous les composants secondaires immédiats ou d'un niveau en aval du CM2 composant, effectuez une requête à l' GSI1 aide `ParentId = "CM2"` de. La requête va renvoyer l'enregistrement suivant.

ParentId	ComponentId
CM2	CM4
	CM5

- Trouvez tous les composants enfants en aval à l'aide d'un composant de niveau supérieur.

  Pour obtenir tous les composants enfants ou en aval pour le composant de niveau supérieur CM1, interrogez GSI2 en utilisant `GraphId = "CM1#1"` et `begins_with("Path", "CM1|")` et utilisez une expression de projection avec`ComponentId`. Elle renverra tous les composants liés à cette arborescence.

  Cet exemple comporte une seule arborescence, avec CM1 comme composant supérieur. En réalité, vous pourriez avoir des millions de composants de niveau supérieur dans la même table.

GraphId	ComponentId
	CM2
CM1#1	CM3
	CM4
	CM5
	CM8
	CM9
	CM10
	CM6
	CM7

- Trouvez tous les composants enfants en aval à l'aide d'un composant de niveau intermédiaire.

  Pour obtenir tous les composants enfants ou descendants de manière récursive par composant CM2, vous avez deux options. Vous pouvez effectuer une recherche récursive niveau par niveau, ou vous pouvez interroger l' GSI2index.

  - Interrogez GSI1, niveau par niveau, de manière récursive, jusqu'à atteindre le dernier niveau des composants enfants.

    1. Requête GSI1 à l'aide de`ParentId = "CM2"`. Cela va renvoyer l'enregistrement suivant.

ParentId	ComponentId
CM2	CM4
	CM5

    2. Encore une fois, interrogez GSI1 en utilisant`ParentId = "CM4"`. Cela va renvoyer l'enregistrement suivant.

ParentId	ComponentId
CM4	CM8
	CM9

    3. Encore une fois, interrogez GSI1 en utilisant`ParentId = "CM5"`. Cela va renvoyer l'enregistrement suivant.

       Continuez la boucle : interrogez pour chaque `ComponentId` jusqu'à ce que vous atteigniez le dernier niveau. Lorsqu'une requête utilisant `ParentId = "<ComponentId>"` ne renvoie aucun résultat, le résultat précédent provenait du dernier niveau de l'arborescence.

ParentId	ComponentId
CM5	CM10

    4. Fusionnez tous les résultats.

résultat= [CM4, CM5] + [CM8, CM9] + [CM10]

= [CM4, CM5 CM8, CM9, CM1 0]

- Query GSI2, qui stocke une arborescence hiérarchique pour un composant de niveau supérieur (une voiture ou CM1).

  1. Tout d'abord, recherchez le composant de premier niveau ou l'ancêtre supérieur et `Path` de CM2. Pour cela, interrogez la table de base en utilisant `ComponentId = "CM2"` pour trouver le chemin de ce composant dans l'arborescence hiérarchique. Sélectionnez les attributs `GraphId` et `Pat` h. La requête va renvoyer l'enregistrement suivant.

GraphId	Chemin
CM1#1	CM1|CM2

  2. Interrogez GSI2 en utilisant`GraphId = "CM1#1" AND BEGINS_WITH("Path", "CM1|CM2|")`. La requête va renvoyer les résultats suivants.

GraphId	Chemin	ComponentId
CM1#1	CM1|CM2|CM4	CM4
	CM1|CM2|CM5	CM5
	CM1|CM2|CM4|CM8	CM8
	CM1|CM2|CM4|CM9	CM9
	CM1| CM2 | CM5 | CM1 0	CM10

  3. Sélectionnez l'`ComponentId`attribut pour lequel vous souhaitez renvoyer tous les composants enfants CM2.

Étape 7 : Valider le modèle de données

Au cours de cette étape, l'utilisateur professionnel valide les résultats de la requête et vérifie s'ils répondent aux besoins de l'entreprise. Vous pouvez utiliser le tableau suivant pour vérifier les modèles d'accès par rapport aux exigences de l'utilisateur.

Question	Table de base//GSI	Interrogation
En tant qu'utilisateur, je souhaite récupérer tous les composants enfants immédiats pour un ID de composant parent.	GSI1	`ParentId = "<ComponentId>"` (Trouvez les enfants immédiats d'un composant.)
En tant qu'utilisateur, je souhaite récupérer une liste récursive de tous les composants enfants pour un ID de composant.	GSI1 ou GSI2	GSI1: `ParentId = "<ComponentId>"` or GSI2: `GraphId = "<TopLevelComponentId>#N" AND BEGINS_WITH("Path", "<PATH_OF_Component>")` (Trouvez tous les composants enfants de niveau inférieur à l'aide d'un composant de niveau supérieur. Trouvez tous les composants enfants de niveau inférieur à l'aide d'un composant de niveau intermédiaire.)
En tant qu'utilisateur, je souhaite voir les ancêtres d'un composant.	Table de base	`ComponentId = "<ComponentId>"` , puis sélectionnez l'attribut Path. (Trouvez les ancêtres d'un composant.)

Vous pouvez également implémenter un script (test) dans n'importe quel langage de programmation pour interroger directement DynamoDB et comparer les résultats aux résultats attendus.

Étape 8 : Revoir l'estimation des coûts

Passez en revue et affinez à nouveau l'estimation des coûts. En outre, il est recommandé de le valider auprès des parties prenantes de l'entreprise et d'obtenir l'approbation nécessaire pour passer à l'étape suivante.

Objectifs

- Définissez le modèle de capacité et estimez les coûts DynamoDB pour affiner l'estimation des coûts à partir de l'étape 2.
- Obtenez l'approbation financière finale de l'analyste commercial et des parties prenantes.

Processus

- L'ingénieur de base de données identifie l'estimation du volume de données.
- L'ingénieur de base de données identifie les exigences en matière de transfert de données.
- L'ingénieur de base de données définit les unités de capacité de lecture et d'écriture requises.
- L'analyste commercial choisit entre les modèles de capacité à la demande et les modèles de capacité provisionnés.
- Un ingénieur de base de données identifie la nécessité d'une mise à l'échelle automatique DynamoDB.
- L'ingénieur de base de données saisit les paramètres dans le Calculateur de tarification AWS.
- L'ingénieur de base de données présente l'estimation finale du prix aux parties prenantes de l'entreprise.
- L'analyste commercial et les parties prenantes approuvent ou rejettent la solution.

Étape 9 : Déployer le modèle de données

Pour cet exemple spécifique, le déploiement du modèle a été effectué à l'aide de NoSQL Workbench, une application destinée au développement et à l'exploitation de bases de données modernes. À l'aide de cet outil, vous avez la possibilité de créer un modèle de données, de télécharger des données et de les déployer directement sur votre Compte AWS. Si vous souhaitez implémenter cet exemple, vous pouvez utiliser le AWS CloudFormation modèle suivant, qui a été généré par NoSQL Workbench.

```
AWSTemplateFormatVersion: 2010-09-09
Resources:
  Components:
    Type: 'AWS::DynamoDB::Table'
    Properties:
      KeySchema:
        - AttributeName: ComponentId
          KeyType: HASH
      AttributeDefinitions:
        - AttributeName: ComponentId
          AttributeType: S
        - AttributeName: ParentId
          AttributeType: S
        - AttributeName: GraphId
          AttributeType: S
        - AttributeName: Path
          AttributeType: S
      GlobalSecondaryIndexes:
        - IndexName: GS1
          KeySchema:
            - AttributeName: ParentId
              KeyType: HASH
            - AttributeName: ComponentId
              KeyType: RANGE
          Projection:
            ProjectionType: KEYS_ONLY
        - IndexName: GSI2
          KeySchema:
            - AttributeName: GraphId
              KeyType: HASH
            - AttributeName: Path
              KeyType: RANGE
          Projection:
            ProjectionType: INCLUDE
            NonKeyAttributes:
              - ComponentId
      BillingMode: PAY_PER_REQUEST
      TableName: Components
```

Ressources supplémentaires

Informations supplémentaires sur DynamoDB

- Tarification DynamoDB
- Documentation DynamoDB
- Conception NoSQL pour DynamoDB
- Sharding d'écriture
- Index secondaires locaux () LSIs
- Index secondaires globaux () GSIs
- Surcharge GSIs
- Sharding GSI
- Utilisation GSIs pour créer une réplique finalement cohérente
- Indices clairsemés
- Requêtes d'agrégation matérialisées
- Modèle de conception de séries chronologiques
- Modèle de conception de liste de contiguïté
- Modèles de capacité à la demande et provisionnés
- Mise à l'échelle automatique DynamoDB
- Durée de vie de DynamoDB (TTL)
- Modélisation des données des joueurs avec DynamoDB (laboratoire)

AWS services

- AWS CloudFormation
- Amazon S3

Outils

- Calculateur de tarification AWS
- NoSQL Workbench pour DynamoDB
- DynamoDB Local

- DynamoDB et AWS SDKs

Bonnes pratiques

- Bonnes pratiques de conception et d'architecture avec DynamoDB (documentation DynamoDB)
- Bonnes pratiques d'utilisation des index secondaires (documentation DynamoDB)
- Bonnes pratiques pour le stockage d'éléments et d'attributs volumineux (documentation DynamoDB)
- Choisir la bonne cléAWS de partition DynamoDB (blog de base de données)
- Comment concevoir des index DBglobal secondaires Amazon Dynamo (blog de AWS base de données)
- Quelles sont les facettes de NoSQL Workbench pour Amazon DynamoDB (site Web de taille moyenne)

AWS ressources générales

- AWS Site Web d'orientation prescriptive
- AWS documentation
- AWS référence générale

Collaborateurs

Parmi les contributeurs à ce guide figurent :

- Camilo Gonzalez, architecte de données senior, AWS
- Moinul Al-Mamun, architecte senior du Big Data, AWS
- Santiago Segura, consultant en services professionnels, AWS
- Satheish Kumar Chandraprakasam, architecte d'applications cloud, AWS

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
Ajout d'une section sur les meilleures pratiques et d'un exemple de modélisation hiérarchique des données.	Nous avons ajouté un résumé des meilleures pratiques DynamoDB et step-by-step un exemple de conception et de validation d'un modèle hiérarchique.	5 décembre 2023
Publication initiale	—	26 octobre 2020

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 un Microsoft Hyper-V application à 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'indexation 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

Consultez la section 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

Consultez la base de données de gestion des configurations.

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 or Bitbucket Cloud. Chaque version du code est appelée branche. Dans une structure de microservice, chaque référentiel est consacré à une seule fonctionnalité. Un seul pipeline CI/CD peut utiliser plusieurs référentiels.

cache passif

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

données gelées

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

vision par ordinateur (CV)

Domaine de l'IA qui utilise l'apprentissage automatique pour analyser et extraire des informations à partir de formats visuels tels que des images numériques et des vidéos. Par exemple, AWS Panorama propose des appareils qui ajoutent des CV aux réseaux de caméras locaux, et 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

Consultez la section 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.

épopée

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

ERP

Voir Planification des ressources d'entreprise.

analyse exploratoire des données (EDA)

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

F

tableau des faits

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

échouer rapidement

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

limite d'isolation des défauts

Dans le AWS Cloud, une limite telle qu'une zone de disponibilité Région AWS, un plan de contrôle ou un plan de données qui limite l'effet d'une panne et contribue à améliorer la résilience des charges de travail. Pour plus d'informations, consultez la section Limites d'isolation des AWS pannes.

branche de fonctionnalités

Voir la succursale.

fonctionnalités

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

importance des fonctionnalités

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

transformation de fonctionnalité

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

invitation en quelques coups

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

FGAC

Découvrez le contrôle d'accès détaillé.

contrôle d'accès détaillé (FGAC)

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

migration instantanée (flash-cut)

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

FM

Voir le modèle de fondation.

modèle de fondation (FM)

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

G

IA générative

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

blocage géographique

Voir les restrictions géographiques.

restrictions géographiques (blocage géographique)

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

Flux de travail Gitflow

Approche dans laquelle les environnements inférieurs et supérieurs utilisent différentes branches dans un référentiel de code source. Le flux de travail Gitflow est considéré comme existant, et le flux de travail basé sur les 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 de l'historien

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

Consultez la section 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 la 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 AWS Well-Architected Framework.

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 : migration réhébergée 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

Consultez la section 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, à évaluer, à prévenir ou à 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 en matière de 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 de type « 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.