



Selección del hardware para su clúster de Amazon EMR

# AWS Guía prescriptiva



# AWS Guía prescriptiva: Selección del hardware para su clúster de Amazon EMR

Copyright © 2024 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Las marcas comerciales y la imagen comercial de Amazon no se pueden utilizar en relación con ningún producto o servicio que no sea de Amazon, de ninguna manera que pueda causar confusión entre los clientes y que menosprecie o desacredite a Amazon. Todas las demás marcas comerciales que no son propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, relacionados o patrocinados por Amazon.

---

# Table of Contents

Introducción .....	1
Información general .....	1
Configuración del clúster .....	2
Seleccionar e implementar .....	3
Pautas de selección .....	3
Selección de instancias .....	4
Calcular capacidad .....	5
Nodos principales .....	5
Nodos básicos y de tareas .....	5
Nodos básicos .....	5
Nodos de tarea .....	6
Elige almacenamiento .....	7
Sistema de archivos distribuidos Hadoop .....	7
Cuándo usar HDFS en Amazon EMR .....	7
Sistema de archivos EMR .....	7
Escalado automático .....	8
Scalado administrado de Amazon EMR .....	8
Escalado automático personalizado .....	8
Consejos de escalado .....	9
Prácticas recomendadas .....	10
Añadir instancias .....	10
Grupos de instancias .....	10
Spot Instances .....	11
Recursos .....	12
Historial de documentos .....	13
Glosario .....	14
# .....	14
A .....	15
B .....	18
C .....	19
D .....	22
E .....	27
F .....	29
G .....	30

---

H .....	31
I .....	32
L .....	34
M .....	35
O .....	39
P .....	41
Q .....	44
R .....	44
S .....	47
T .....	50
U .....	52
V .....	53
W .....	53
Z .....	54
.....	iv

# Selección del hardware para su clúster de Amazon EMR

Sayde Aguilar, Amiin Samatar y Diego Valencia, de Amazon Web Services (AWS)

agosto de 2023([historial de documentos](#))

Amazon EMR es una herramienta para el procesamiento de macrodatos. Utiliza software de código abierto, específicamente herramientas de Apache como Apache Spark y Apache Hudi. Además, ofrece varias opciones para configurar y utilizar un sistema de bajo costo, pay-as-you-go modelo.

En esta guía se explica cómo diseñar el clúster de Amazon EMR en función de esa elasticidad y se proporcionan las prácticas recomendadas a seguir a la hora de elegir el hardware.

## Información general

Amazon EMR se ha creado con Apache HadoopMapReduce, un marco para procesar grandes cantidades de datos. HadoopMapReduce procesa los datos en clústeres distribuidos al mismo tiempo mediante lógica paralela, lo que significa que cada proceso tiene su propio procesador. Amazon EMR usa un clúster Hadoop de servidores virtuales estructurado en Amazon Elastic Compute Cloud (Amazon EC2). Esto significa que todos los procesos paralelos se realizan en ordenadores independientes que se ejecutan en Amazon Web Services (AWS).

Un clúster de Hadoop es un tipo específico de clúster computacional que se utiliza para procesar grandes cantidades de datos no estructurados mediante entornos paralelos o distribuidos. Una característica clave de un clúster de Hadoop es que es altamente escalable y se puede configurar para aumentar la velocidad del procesamiento de datos. La escalabilidad se alcanza añadiendo o quitando nodos para aumentar o disminuir el rendimiento. En los clústeres de Hadoop, cada dato se copia entre los nodos del clúster, por lo que se pierden casi cero datos si se produce un error en un nodo.

En Amazon EMR, elasticidad hace referencia a la capacidad de redimensionamiento dinámico. Puede escalar automáticamente el clúster y realizar los cambios que necesite. No tiene que confiar en su diseño de hardware inicial.

En esta guía se explica cómo diseñar el clúster de Amazon EMR en función de esa elasticidad y se proporcionan las prácticas recomendadas a seguir a la hora de elegir el hardware.

# Configuración de los clústeres de Amazon EMR

Con Amazon EMR, configurar un clúster para procesar y analizar datos con marcos de macrodatos solo puede llevar unos minutos. Sin embargo, antes de implementar el clúster, debe preparar el almacenamiento, definir los límites de la operación y determinar la distribución más eficiente de la carga de trabajo.

Para su configuración de Amazon EMR, esta guía trata los siguientes temas:

- [Selección e implementación de un clúster de Amazon EMR](#)
- [Estimación de la capacidad del clúster de Amazon EMR](#)
- [Elegir el almacenamiento adecuado](#)
- [Elegir el enfoque de escalado automático](#)
- [Mejores prácticas de configuración](#)

# Selección e implementación de un clúster de Amazon EMR

Identifique y organice el tipo de nodos. Al definir el clúster de Amazon EMR, es importante entender su hardware. ¿Cómo funciona? ¿Cómo está compuesto? Las respuestas a estas preguntas incluyen tres partes:

- El tipo de nodos
- La función que desempeña cada nodo
- Los tipos de instancias de EC2 que son más eficientes para cada nodo

Inicialmente, el nodo principal se encarga de administrar los recursos generales. Ejecuta los componentes principales de la aplicación distribuida. Por ejemplo, ejecuta el Sistema de archivos distribuidos de Hadoop (HDFS) NameNode Servicio, realiza un seguimiento de los trabajos que se deben realizar en el clúster y monitorea el estado del sistema.

Además, Amazon EMR cuenta con nodos principales y nodos de tareas. Los nodos principales son gestionados por el nodo principal. Los nodos principales ejecutan nodos de tareas y se encargan de almacenar los datos en el HDFS del clúster. Los nodos de tareas se encargan de administrar las tareas que llegan al clúster. Un nodo de tareas no almacena datos. (Los nodos de tareas no son obligatorios).

Al configurar e implementar el clúster de Amazon EMR, es importante tener en cuenta la elección correcta de las instancias de EC2 que representarán los nodos del clúster. Hay varias formas de agregar instancias de EC2 a un clúster, en función de si utiliza la configuración de grupos de instancias o la configuración de flotas de instancias para el clúster. Para obtener más información acerca de los tipos de instancias admitidos, consulte [AWS documentación](#).

Las siguientes pautas se aplican a la mayoría [Clústeres de Amazon EMR](#). También puede revisar el [prácticas recomendadas de configuración de clústeres](#).

## Pautas para la selección de instancias

En general, las instancias preferidas para la implementación de Amazon EMR dependen del trabajo que esté ejecutando. Considere las siguientes preguntas:

- ¿Su trabajo consume mucha memoria?
- ¿Su trabajo consume mucha CPU?

- ¿Necesita grandes cantidades de almacenamiento?
- ¿Su trabajo requiere capacidad de GPU?

Estas preguntas le ayudarán a entender el tipo de instancias que necesita y las características reales que necesita. Determine cuántos trabajos quiere procesar al mismo tiempo y qué tan rápido necesita que se procesen los trabajos. Esto es importante porque el uso de Amazon EMR se cobra en incrementos por hora. Cuando activas un clúster, se te cobrará toda la hora.

Puede comprobar el costo de cada instancia que se ejecuta en diferentes AWS Regiones. Para comparar precios entre regiones, puede utilizar la [AWS Calculadora de precios](#) y cambia los valores en función de tu ubicación.

## Selección de instancias de EC2

Cuando haya respondido a las preguntas anteriores, es el momento de seleccionar las instancias en función de esos requisitos. Después de comprender las necesidades de su trabajo de procesamiento, determine el tipo de instancia en función de las características que necesite:

- Si necesitas instancias de uso general, elige M6g, T4g, o M5 instancias.
- Si necesitas instancias optimizadas para la computación, elige C6g o C5 instancias.
- Si necesitas instancias optimizadas para la memoria, elige R6g, X1, R5, o z1d instancias.
- Si debe optimizar el almacenamiento, elija I3 instancias, que proporcionan un alto rendimiento de E/S.
- Si necesita computación acelerada, como una GPU, elija P3, G4, o Inf1 instancias. Estos tipos de instancias ofrecen un alto rendimiento para el aprendizaje automático y la dinámica de fluidos, entre otros procesos.

Otra forma de entender los tipos de instancias y sus capacidades consiste en analizar las [memoria predeterminada](#) para cada tipo de instancia. Esta métrica le ayuda a ajustar y mejorar el rendimiento de su MapReduce trabajos. Para obtener más información, consulte [Parámetros de configuración del daemon de Hadoop](#).

Cuando sepa el tipo de instancias que necesita, puede planificar la capacidad de su clúster.



# Cálculo de la capacidad de un clúster de Amazon EMR

Si bien Amazon EMR es una plataforma de tamaño variable, es importante ajustar el tamaño de los clústeres. El dimensionamiento correcto evita que un clúster se vuelva lento si el tamaño es insuficiente o que los costos sean más altos si el clúster está sobredimensionado. Para anticipar estos problemas, puede calcular la cantidad y el tipo de nodos que se necesitarán para las cargas de trabajo.

## Nodos principales

Este tipo de nodo es el responsable de coordinar la distribución de los datos y los procesos. Como se mencionó anteriormente, los requisitos computacionales para los nodos principales son bajos. Puede utilizar un único nodo principal para administrar su clúster de Amazon EMR. Sin embargo, puede utilizar hasta tres nodos principales para no tener un único punto de falla. Si un nodo principal falla, Amazon EMR conmuta por error a uno de los otros dos nodos principales.

## Nodos básicos y de tareas

La diferencia entre los nodos básicos y los nodos de tareas es que los nodos de tareas no almacenan datos; solo proporcionan energía para realizar tareas computacionales paralelas.

Para calcular el número de nodos básicos y de tareas, debe conocer el tamaño de los datos y el uso aproximado de la memoria.

## Nodos básicos

Los nodos básicos son responsables de ejecutar las tareas para procesar los datos y también de almacenar datos en el Sistema de archivos distribuido de Hadoop (HDFS). Para calcular la capacidad de los nodos básicos, defina el número de nodos básicos y, a continuación, multiplique el número de nodos por el almacenamiento de Amazon Elastic Block Store (Amazon EBS) de cada nodo.

Por ejemplo, si define 10 nodos básicos para procesar 1 TiB de datos y tiene un tipo de instancia `m5.xlarge` con 64 GiB de almacenamiento en Amazon EBS, tiene una capacidad de  $10 \text{ nodos} \times 64 \text{ GiB}$  o 640 GiB. En función del [factor de replicación](#) HDFS de tres, el tamaño de los datos se replica tres veces en los nodos, por lo que 1 TiB de datos requiere una capacidad de 3 TiB. Como

este ejemplo tiene solo 640 GiB, debe aumentar la cantidad de nodos o cambiar el tipo de instancia hasta que tenga una capacidad de 3 TiB.

El tipo de instancia `m5.4xlarge` tiene 256 GiB de almacenamiento. Si se cambia a un tipo de instancia `m5.4xlarge` y se seleccionan 12 instancias, se obtiene suficiente capacidad.

```
12 instances × 256 GiB of storage = 3072 GiB = 3 TiB available
```

## Nodos de tarea

Los nodos de tareas solo ejecutan tareas. No almacenan datos. Para calcular el número de nodos de tareas, necesita un cálculo del uso de la memoria. Esta capacidad se puede dividir entre los nodos básicos y de tareas. Para calcular el número de nodos de tareas necesarios, puede restar del uso de la memoria aquella memoria proporcionada por los nodos básicos que calculó en el paso anterior.

Para tener un rango de memoria ampliado, se recomienda multiplicar la memoria necesaria por tres.

Suponga que tiene 28 procesos de 20 GiB cada uno.

```
3 × 28 processes × 20 GiB of memory = 1680 GiB of memory
```

Para este ejemplo, los nodos básicos tienen 64 GiB de memoria (instancias `m5.4xlarge`). Sus nodos básicos proporcionan  $64 \text{ GiB} \times 12 \text{ nodes} = 768 \text{ GiB of memory}$ , lo que no es suficiente en este ejemplo.

Para determinar lo que falta, reste la memoria del nodo básico de la memoria total requerida.

```
1680 GiB - 768 GiB core node memory = 912 GiB memory shortage.
```

Los nodos de tareas pueden proporcionar los 912 GiB de memoria restantes. Para este ejemplo, los nodos de tareas tienen 32 GiB de memoria (instancias `m5.2xlarge`). Para obtener el número de nodos de tareas necesarios, divida lo que falta de memoria entre la memoria del tipo de instancia.

```
912 GiB/32 GiB = 28.5 task nodes
```

No puede tener una fracción de un nodo de tarea, por lo que debe redondear hasta 29 nodos de tarea.

## Elegir el almacenamiento adecuado

Configure los [tipos de instancias y la capacidad](#) de su clúster. Los nodos principales y de tareas necesitan potencia de procesamiento y procesamiento, pero solo los nodos principales almacenan datos. Seleccione el tipo de almacenamiento más rentable para sus nodos principales.

Si utiliza Amazon EMR para procesar una gran cantidad de datos, tiene varias opciones para transferir datos desde Amazon S3. La mejor opción depende de tu carga de trabajo. En las siguientes secciones se proporcionan algunos puntos clave a tener en cuenta a la hora de decidir qué tipo de almacenamiento es el más adecuado para usted.

## Sistema de archivos distribuidos Hadoop

Hadoop Distributed File System (HDFS) es Hadoop Distributed File System (HDFS) es un sistema de archivos distribuido, escalable y portátil para Hadoop Distributed File System. Una ventaja de HDFS es el reconocimiento de datos entre los nodos de clúster de Hadoop que administran los clústeres y los nodos de clúster de Hadoop que administran los pasos individuales.

## Cuándo usar HDFS en Amazon EMR

Cuando tiene lecturas iterativas en el mismo conjunto de datos o cargas de trabajo con un uso intensivo de I/O de disco, puede utilizar HDFS para almacenar en caché los resultados intermedios y como almacenamiento activo para procesar los datos. El HDFS es efímero, lo que significa que se recupera cuando se terminan las instancias.

## Sistema de archivos EMR

El sistema de archivos EMR (EMRFS) es una implementación de HDFS que los clústeres de Amazon EMR suelen utilizar para leer y escribir archivos normales de Amazon EMR directamente en Amazon S3.

Puede usar EMRFS al leer el conjunto de datos una vez en cada ejecución. EMRFS desvincula el almacenamiento del procesamiento, por lo que no necesita aprovisionar nodos principales específicamente para almacenar datos y no necesita pagar por la replicación de datos en HDFS. Esto se traduce en costos más bajos y proporciona disponibilidad de los datos para varios clústeres. También tiene la ventaja de conservar los datos después de cerrar el clúster.

## Elegir el enfoque de escalado automático

La elasticidad es una de las principales ventajas de utilizar Amazon EMR. Hay dos opciones principales para escalar automáticamente los recursos:

- Escalado administrado
- Política de escalado personalizada

Con el escalado administrado o con una política de escalado automático personalizada, puede escalar y reducir los nodos para utilizar solo los recursos que necesita. La escalabilidad horizontal se utiliza para agregar más recursos cuando se necesita más capacidad. La escalabilidad se traduce en rentabilidad al eliminar los recursos que no se utilizan. Dentro del servicio Amazon EMR, CloudWatch las métricas de Amazon están habilitadas para monitorear sus recursos y así poder escalar su clúster. CloudWatch toma puntos de datos cada 5 minutos.

Existen diferentes consideraciones para cada uno de los enfoques de escalado automático.

## Escalado administrado de Amazon EMR

Utilice el [escalado administrado de EMR](#) si su carga de trabajo cumple los siguientes criterios:

- Se necesita una experiencia gestionada.
- Se utiliza Amazon EMR 5.330 o posterior.
- Necesita una frecuencia de evaluación de 1 minuto.
- La solución utiliza flotas de instancias para tener entre una y cinco opciones de instancia.
- Las aplicaciones se basan en Apache Spark, Apache Hive o Apache Hadoop YARN.

## Escalado automático personalizado

Utilice una [política de escalado automático personalizada](#) si su carga de trabajo cumple los siguientes criterios:

- Debe controlar la métrica para escalar.
- Se utiliza Amazon EMR 4.0+.
- No hay necesidad de una frecuencia de evaluación alta.

- No hay ningún requisito para controlar los períodos de enfriamiento entre cambios de tamaño consecutivos.
- Es importante controlar el número de instancias que se agregarán o eliminarán al escalar.
- La solución necesita acciones de escalado personalizadas. Por ejemplo, puede escalar más de un nodo en un período de 5 minutos. O puede que quieras ajustar el período de enfriamiento.
- No hay restricciones para usar diferentes tipos de instantáneas en un grupo de instancias.

## Consejos para añadir el escalado automático a su clúster

- Tenga en cuenta la cantidad de datos que procesará. Forecast pronósticos utilizando el caso con el mayor tamaño de datos.
- [Ajusta el tamaño](#) del clúster.
- Elija un [tipo de almacenamiento](#) que se adapte a sus necesidades.
- Comprenda las [métricas](#) de un clúster de Amazon EMR.
- Descubra cómo determinar la [métrica correcta](#) para escalar su clúster.
- Decida si utilizará instancias puntuales, grupos de instancias uniformes o flotas de instancias.
- En función de la información y las limitaciones, decida qué tipo de enfoque de escalado prefiere: el escalado administrado de Amazon EMR o una política de escalado automático personalizada.
- Configure el escalado administrado o la política personalizada.
- Si seleccionó una política de escalado automático personalizada, supervise las métricas de Amazon EMR para ajustar los umbrales de la política.

# Prácticas recomendadas de configuración de Amazon EMR

Al configurar su clúster de Amazon EMR, siga las siguientes prácticas recomendadas para agregar instancias, trabajar con grupos de instancias y usar instancias puntuales.

## Añadir instancias

Al configurar el clúster de EMR, es importante tener en cuenta la elección correcta de las instancias de EC2 que representarán los nodos del clúster. Recuerde que no puede cambiar el tipo de instancias, por ejemplo, cambiar las instancias puntuales a instancias bajo demanda, cuando el clúster está en ejecución. Para cambiar el nodo principal, debe cerrar el clúster y crear uno nuevo. Por eso debes elegir el tipo de instancia correcto para tener el menor tiempo de inactividad posible. Para obtener más información, consulte [Pautas y mejores prácticas de configuración de clústeres](#).

Hay varias formas de agregar instancias de EC2 a un clúster, en función de si utiliza la configuración de grupos de instancias o la configuración de flotas de instancias para el clúster:

- Agregar instancias de EC2 manualmente
- Agrega manualmente una tarea en el grupo de instancias para agregar automáticamente una instancia
- Configurar el escalado automático

## Grupos de instancias

Cuando añada instancias de EC2 a su configuración, considere la posibilidad de utilizar grupos de instancias. Si agrega instancias manualmente, puede agregar instancias del mismo tipo a los grupos de instancias principales y de tareas existentes. Además, puedes añadir un grupo de instancias de tareas, que puede usar un tipo de instancia diferente.

Por último, configure el escalado automático en Amazon EMR para un grupo de instancias. Las instancias se pueden añadir y eliminar automáticamente en función del valor de un AmazonCloudWatch métrica que especifique. De lo contrario, si usas flotas de instancias, agrega una sola flota de instancias de tareas. A continuación, cambie la capacidad objetivo para las instancias bajo demanda y las instancias puntuales para las flotas de instancias principales y de tareas existentes.

## Spot Instances

Utilice instancias puntuales en los nodos de tareas. Los nodos de tareas procesan datos, pero no contienen datos persistentes en el Sistema de archivos distribuidos de Hadoop (HDFS). Si los nodos de tareas se cierran porque el precio spot ha subido por encima del precio spot máximo, no se pierde ningún dato y el efecto en el clúster es mínimo.

Cuando lanza grupos de instancias de tareas como instancias puntuales, Amazon EMR aprovisiona tantos nodos de tareas como puede, utilizando su precio máximo de spot. Por ejemplo, puedes solicitar un grupo de instancias de tareas con seis nodos. Si solo hay cinco instancias puntuales disponibles al precio máximo de spot o por debajo de este, Amazon EMR lanza el grupo de instancias con cinco nodos. Amazon EMR añade el sexto nodo más adelante si es posible. Para obtener más información, consulte [Pautas y mejores prácticas de configuración de clústeres](#).

# Recursos

- [¿Qué es Amazon EMR?](#)
- [Características de Amazon EMR](#)
- [Consideraciones y prácticas recomendadas para Amazon EMR](#)
- [Tipos de instancias compatibles con Amazon EMR](#)
- [Pautas y prácticas recomendadas para la configuración de clústeres](#)
- [Comprenda los tipos de nodos: nodos principales, principales y de tareas](#)
- [Configuración de tareas](#)
- [Parámetros de configuración del daemon de Hadoop](#)
- [Utilice la consola de AWS gestión para configurar el escalado gestionado](#)
- [Supervise las métricas con CloudWatch](#)
- [Presentamos Amazon EMR Managed Scaling: cambie automáticamente el tamaño de los clústeres para reducir los costes](#)



## Historial de documentos

En la siguiente tabla se describen los cambios importantes de esta guía. Si quieres recibir notificaciones sobre futuras actualizaciones, puedes suscribirte a un [Fuente RSS](#).

Cambio	Descripción	Fecha
<a href="#">Aclaración</a>	Revisó <a href="#">ejemplos de estimación de capacidad</a> .	10 de agosto de 2023
=	Publicación inicial	29 de octubre de 2021

# Glosario de las Recomendaciones de AWS

Los siguientes son términos de uso común en las estrategias, guías y patrones que se ofrecen en las Recomendaciones de AWS. Para sugerir entradas, utilice el enlace [Enviar comentarios](#) al final del glosario.

## Números

### Las 7 R

Siete estrategias de migración comunes para trasladar aplicaciones a la nube. Estas estrategias se basan en las 5 R que Gartner identificó en 2011 y consisten en lo siguiente:

- **Refactorizar/rediseñar:** traslade una aplicación y modifique su arquitectura mediante el máximo aprovechamiento de las características nativas en la nube para mejorar la agilidad, el rendimiento y la escalabilidad. Por lo general, esto implica trasladar el sistema operativo y la base de datos. Ejemplo: Migre la base de datos de Oracle en las instalaciones a Amazon Aurora PostgreSQL-Compatible Edition.
- **Redefinir la plataforma (transportar y redefinir):** traslade una aplicación a la nube e introduzca algún nivel de optimización para aprovechar las capacidades de la nube. Ejemplo: Migre la base de datos Oracle en las instalaciones a Amazon Relational Database Service (Amazon RDS) para Oracle en la nube de AWS.
- **Recomprar (readquirir):** cambie a un producto diferente, lo cual se suele llevar a cabo al pasar de una licencia tradicional a un modelo SaaS. Ejemplo: Migre el sistema de administración de las relaciones con los clientes (CRM) a Salesforce.com.
- **Volver a alojar (migrar mediante lift-and-shift):** traslade una aplicación a la nube sin realizar cambios para aprovechar las capacidades de la nube. Ejemplo: Migre su base de datos de Oracle en las instalaciones a Oracle en una instancia de EC2 en la nube de AWS.
- **Reubicar (migrar el hipervisor mediante lift and shift):** traslade la infraestructura a la nube sin comprar equipo nuevo, reescribir aplicaciones o modificar las operaciones actuales. Este escenario de migración es específico de VMware Cloud en AWS, que permite la compatibilidad de máquinas virtuales (VM) y la portabilidad de cargas de trabajo entre el entorno en las instalaciones y de AWS. Puede utilizar las tecnologías de VMware Cloud Foundation desde los centros de datos en las instalaciones al migrar una infraestructura a VMware Cloud en AWS. Ejemplo: Reubicar el hipervisor que aloja la base de datos de Oracle a VMware Cloud en AWS.

- **Retener (revisitar):** conserve las aplicaciones en el entorno de origen. Estas pueden incluir las aplicaciones que requieren una refactorización importante, que desee posponer para más adelante, y las aplicaciones heredadas que desee retener, ya que no hay ninguna justificación empresarial para migrarlas.
- **Retirar:** retire o elimine las aplicaciones que ya no sean necesarias en un entorno de origen.

## A

### ABAC

Consulte control de [acceso basado en atributos](#).

servicios abstractos

Consulte [servicios gestionados](#).

### ACID

Consulte [atomicidad, consistencia, aislamiento y durabilidad](#).

migración activa-activa

Método de migración de bases de datos en el que las bases de datos de origen y destino se mantienen sincronizadas (mediante una herramienta de replicación bidireccional o mediante operaciones de escritura doble) y ambas bases de datos gestionan las transacciones de las aplicaciones conectadas durante la migración. Este método permite la migración en lotes pequeños y controlados, en lugar de requerir una transición única. Es más flexible, pero requiere más trabajo que la migración [activa-pasiva](#).

migración activa-pasiva

Método de migración de bases de datos en el que las bases de datos de origen y destino se mantienen sincronizadas, pero solo la base de datos de origen gestiona las transacciones de las aplicaciones conectadas, mientras los datos se replican en la base de datos de destino. La base de datos de destino no acepta ninguna transacción durante la migración.

función agregada

Función SQL que opera en un grupo de filas y calcula un único valor de retorno para el grupo. Entre los ejemplos de funciones agregadas se incluyen SUM y MAX.

### IA

Véase [inteligencia artificial](#).

## AIOps

Consulte las [operaciones de inteligencia artificial](#).

### anonimización

El proceso de eliminar permanentemente la información personal de un conjunto de datos. La anonimización puede ayudar a proteger la privacidad personal. Los datos anonimizados ya no se consideran datos personales.

### antipatronos

Una solución que se utiliza con frecuencia para un problema recurrente en el que la solución es contraproducente, ineficaz o menos eficaz que una alternativa.

### control de aplicaciones

Un enfoque de seguridad que permite el uso únicamente de aplicaciones aprobadas para ayudar a proteger un sistema contra el malware.

### cartera de aplicaciones

Recopilación de información detallada sobre cada aplicación que utiliza una organización, incluido el costo de creación y mantenimiento de la aplicación y su valor empresarial. Esta información es clave para [el proceso de detección y análisis de la cartera](#) y ayuda a identificar y priorizar las aplicaciones que se van a migrar, modernizar y optimizar.

### inteligencia artificial (IA)

El campo de la informática que se dedica al uso de tecnologías informáticas para realizar funciones cognitivas que suelen estar asociadas a los seres humanos, como el aprendizaje, la resolución de problemas y el reconocimiento de patrones. Para más información, consulte [¿Qué es la inteligencia artificial?](#)

### operaciones de inteligencia artificial (AIOps)

El proceso de utilizar técnicas de machine learning para resolver problemas operativos, reducir los incidentes operativos y la intervención humana, y mejorar la calidad del servicio. Para obtener más información sobre cómo se utiliza AIOps en la estrategia de migración de AWS, consulte la [Guía de integración de operaciones](#).

### cifrado asimétrico

Algoritmo de cifrado que utiliza un par de claves, una clave pública para el cifrado y una clave privada para el descifrado. Puede compartir la clave pública porque no se utiliza para el descifrado, pero el acceso a la clave privada debe estar sumamente restringido.

## atomicidad, consistencia, aislamiento, durabilidad (ACID)

Conjunto de propiedades de software que garantizan la validez de los datos y la fiabilidad operativa de una base de datos, incluso en caso de errores, cortes de energía u otros problemas.

## control de acceso basado en atributos (ABAC)

La práctica de crear permisos detallados basados en los atributos del usuario, como el departamento, el puesto de trabajo y el nombre del equipo. A fin de obtener más información, consulte [ABAC para AWS](#) en la documentación de AWS Identity and Access Management (IAM).

## origen de datos fidedigno

Ubicación en la que se almacena la versión principal de los datos, que se considera la fuente de información más fiable. Puede copiar los datos del origen de datos autorizado a otras ubicaciones con el fin de procesarlos o modificarlos, por ejemplo, anonimizarlos, redactarlos o seudonimizarlos.

## Zona de disponibilidad

Ubicación diferenciada de una Región de AWS que está aislada de los errores que se producen en otras zonas de disponibilidad y que brinda conectividad de red económica y de baja latencia a otras zonas de disponibilidad de la misma región.

## AWS Cloud Adoption Framework (AWS CAF)

Marco de directrices y prácticas recomendadas de AWS para ayudar a las empresas a desarrollar un plan eficiente y eficaz a fin de migrar con éxito a la nube. AWS CAF organiza la orientación en seis áreas de enfoque llamadas perspectivas: empresarial, humana, gobernanza, plataforma, seguridad y operaciones. Las perspectivas empresariales, humanas y de gobernanza se centran en las habilidades y los procesos empresariales; las perspectivas de plataforma, seguridad y operaciones se centran en las habilidades y los procesos técnicos. Por ejemplo, la perspectiva humana se dirige a las partes interesadas que se ocupan de los Recursos Humanos (RR. HH.), las funciones del personal y la administración de las personas. Desde esta perspectiva, AWS CAF brinda orientación para el desarrollo, la capacitación y la comunicación de las personas, con el fin de ayudar a preparar la organización para una adopción exitosa de la nube. Para obtener más información, consulte la [Página web de AWS CAF](#) y el [Documento técnico de AWS CAF](#).

## AWS Workload Qualification Framework (AWS WQF)

Herramienta que evalúa las cargas de trabajo de migración de bases de datos, recomienda estrategias de migración y brinda estimaciones de trabajo. AWS WQF se incluye con AWS

Schema Conversion Tool (AWS SCT). Analiza los esquemas de bases de datos y los objetos de código, el código de las aplicaciones, las dependencias y las características de rendimiento y proporciona informes de evaluación.

## B

### BCP

Consulte la [planificación de la continuidad del negocio](#).

### gráfico de comportamiento

Una vista unificada e interactiva del comportamiento de los recursos y de las interacciones a lo largo del tiempo. Puede utilizar un gráfico de comportamiento con Amazon Detective para examinar los intentos de inicio de sesión fallidos, las llamadas sospechosas a la API y acciones similares. Para obtener más información, consulte [Datos en un gráfico de comportamiento](#) en la documentación de Detective.

### sistema big-endian

Un sistema que almacena primero el byte más significativo. Véase también [endianness](#).

### clasificación binaria

Un proceso que predice un resultado binario (una de las dos clases posibles). Por ejemplo, es posible que su modelo de ML necesite predecir problemas como “¿Este correo electrónico es spam o no es spam?” o “¿Este producto es un libro o un automóvil?”.

### filtro de floración

Estructura de datos probabilística y eficiente en términos de memoria que se utiliza para comprobar si un elemento es miembro de un conjunto.

### rama

Área contenida de un repositorio de código. La primera rama que se crea en un repositorio es la rama principal. Puede crear una rama nueva a partir de una rama existente y, a continuación, desarrollar características o corregir errores en la rama nueva. Una rama que se genera para crear una característica se denomina comúnmente rama de característica. Cuando la característica se encuentra lista para su lanzamiento, se vuelve a combinar la rama de característica con la rama principal. Para obtener más información, consulte [Acerca de las sucursales \(documentación\)](#) GitHub .

## acceso con cristales rotos

En circunstancias excepcionales y mediante un proceso aprobado, un usuario puede acceder rápidamente a un sitio para el Cuenta de AWS que normalmente no tiene permisos de acceso. Para obtener más información, consulte el indicador [Implemente procedimientos de rotura de cristales en la guía Well-ArchitectedAWS](#).

## estrategia de implementación sobre infraestructura existente

La infraestructura existente en su entorno. Al adoptar una estrategia de implementación sobre infraestructura existente para una arquitectura de sistemas, se diseña la arquitectura en función de las limitaciones de los sistemas y la infraestructura actuales. Si está ampliando la infraestructura existente, puede combinar las estrategias de implementación sobre infraestructuras existentes y de [implementación desde cero](#).

## caché de búfer

El área de memoria donde se almacenan los datos a los que se accede con más frecuencia.

## capacidad empresarial

Lo que hace una empresa para generar valor (por ejemplo, ventas, servicio al cliente o marketing). Las arquitecturas de microservicios y las decisiones de desarrollo pueden estar impulsadas por las capacidades empresariales. Para obtener más información, consulte la sección [Organizado en torno a las capacidades empresariales](#) del documento técnico [Ejecutar microservicios en contenedores en AWS](#).

## planificación de la continuidad del negocio (BCP)

Plan que aborda el posible impacto de un evento disruptivo, como una migración a gran escala en las operaciones y permite a la empresa reanudar las operaciones rápidamente.

# C

## CAF

[Consulte el marco AWS de adopción de la nube.](#)

## CCoE

Consulte el [Centro de excelencia en la nube](#).

## CDC

Consulte la [captura de datos de cambios](#).

## captura de datos de cambio (CDC)

Proceso de seguimiento de los cambios en un origen de datos, como una tabla de base de datos, y registro de los metadatos relacionados con el cambio. Puede utilizar los CDC para diversos fines, como auditar o replicar los cambios en un sistema de destino para mantener la sincronización.

## ingeniería del caos

Introducir intencionalmente fallos o eventos disruptivos para poner a prueba la resiliencia de un sistema. Puedes usar [AWS Fault Injection Service\(AWS FIS\)](#) para realizar experimentos que estresen tus AWS cargas de trabajo y evalúen su respuesta.

## CI/CD

Consulte la [integración continua y la entrega continua](#).

## clasificación

Un proceso de categorización que permite generar predicciones. Los modelos de ML para problemas de clasificación predicen un valor discreto. Los valores discretos siempre son distintos entre sí. Por ejemplo, es posible que un modelo necesite evaluar si hay o no un automóvil en una imagen.

## cifrado del cliente

Cifrado de datos de forma local, antes de que el Servicio de AWS de destino los reciba.

## Centro de excelencia en la nube (CCoE)

Equipo multidisciplinario que impulsa los esfuerzos de adopción de la nube en toda la organización, incluido el desarrollo de las prácticas recomendadas en la nube, la movilización de recursos, el establecimiento de plazos de migración y la dirección de la organización durante las transformaciones a gran escala. Para obtener más información, consulte las [Publicaciones del CCoE](#) en el Blog de estrategia empresarial en la nube de AWS.

## computación en la nube

La tecnología en la nube que se utiliza normalmente para la administración de dispositivos de IoT y el almacenamiento de datos de forma remota. La computación en la nube suele estar conectada a la tecnología de [computación perimetral](#).



## modelo operativo en la nube

En una organización de TI, el modelo operativo que se utiliza para crear, madurar y optimizar uno o más entornos de nube. Para obtener más información, consulte [Creación de su modelo operativo de nube](#).

## etapas de adopción de la nube

Las siguientes son las cuatro fases por las que suelen pasar las empresas cuando migran a la nube de AWS:

- Proyecto: ejecución de algunos proyectos relacionados con la nube con fines de prueba de concepto y aprendizaje
- Fundamento: realización de inversiones fundamentales para escalar la adopción de la nube (p. ej., crear una zona de aterrizaje, definir un CCoE, establecer un modelo de operaciones)
- Migración: migración de aplicaciones individuales
- Reinención: optimización de productos y servicios e innovación en la nube

Stephen Orban definió estas etapas en la publicación del blog [The Journey Toward Cloud-First & the Stages of Adoption](#) (El camino hacia la nube como prioridad y las etapas de adopción) en el Blog de estrategia empresarial en la nube de AWS. Para obtener información sobre cómo se relacionan con la estrategia de migración de AWS, consulte la [Guía de preparación para la migración](#).

## CMDB

Consulte la [base de datos de gestión de la configuración](#).

## repositorio de código

Una ubicación donde el código fuente y otros activos, como documentación, muestras y scripts, se almacenan y actualizan mediante procesos de control de versiones. Los repositorios en la nube más comunes incluyen GitHub o AWS CodeCommit. Cada versión del código se denomina rama. En una estructura de microservicios, cada repositorio se encuentra dedicado a una única funcionalidad. Una sola canalización de CI/CD puede utilizar varios repositorios.

## caché en frío

Una caché de búfer que está vacía no está bien poblada o contiene datos obsoletos o irrelevantes. Esto afecta al rendimiento, ya que la instancia de la base de datos debe leer desde la memoria principal o el disco, lo que es más lento que leer desde la memoria caché del búfer.

## datos fríos

Datos a los que se accede con poca frecuencia y que suelen ser históricos. Al consultar este tipo de datos, normalmente se aceptan consultas lentas. Trasladar estos datos a niveles o clases de almacenamiento de menor rendimiento y menos costosos puede reducir los costos.

## visión artificial

Campo de IA utilizado por las máquinas para identificar personas, lugares y cosas en imágenes con una precisión igual o superior a la humana. Construido a menudo con modelos de aprendizaje profundo, automatiza la extracción, el análisis, la clasificación y la comprensión de información útil a partir de una sola imagen o una secuencia de imágenes.

## base de datos de administración de configuración (CMDB)

Repositorio que almacena y administra información sobre una base de datos y su entorno de TI, incluidos los componentes de hardware y software y sus configuraciones. Por lo general, los datos de una CMDB se utilizan en la etapa de detección y análisis de la cartera de productos durante la migración.

## paquete de conformidad

Una colección de acciones correctivas y reglas de AWS Config que puede reunir para personalizar sus controles de seguridad y conformidad. Puede implementar un paquete de conformidad como una sola entidad en una región y Cuenta de AWS, o en toda una organización, mediante una plantilla YAML. Para obtener más información, consulte [Paquetes de conformidad](#) en la documentación de AWS Config.

## integración y entrega continuas (CI/CD)

El proceso de automatización de las etapas de origen, compilación, prueba, presentación y producción del proceso de lanzamiento del software. La CI/CD se describe comúnmente como una canalización. La CI/CD puede ayudarlo a automatizar los procesos, mejorar la productividad, mejorar la calidad del código y entregar con mayor rapidez. Para obtener más información, consulte [Beneficios de la entrega continua](#). CD también puede significar implementación continua. Para obtener más información, consulte [Entrega continua frente a implementación continua](#).

# D

## datos en reposo

Datos que están estacionarios en la red, como los datos que se encuentran almacenados.

## clasificación de datos

Un proceso para identificar y clasificar los datos de su red en función de su importancia y sensibilidad. Es un componente fundamental de cualquier estrategia de administración de riesgos de ciberseguridad porque lo ayuda a determinar los controles de protección y retención adecuados para los datos. La clasificación de datos es un componente del pilar de seguridad del Marco de AWS Well-Architected. Para obtener más información, consulte [Clasificación de datos](#).

## desviación de datos

Una variación significativa entre los datos de producción y los datos que se utilizaron para entrenar un modelo de machine learning, o un cambio significativo en los datos de entrada a lo largo del tiempo. La desviación de los datos puede reducir la calidad, la precisión y la imparcialidad generales de las predicciones de los modelos de machine learning.

## datos en tránsito

Datos que se mueven de forma activa por la red, por ejemplo, entre los recursos de la red.

## minimización de datos

El principio de recopilar y procesar solo los datos estrictamente necesarios. Practicar la minimización de los datos en Nube de AWS puede reducir los riesgos de privacidad, los costos y la huella de carbono derivada de los análisis.

## perímetro de datos

Un conjunto de barreras preventivas en su AWS entorno que ayudan a garantizar que solo las identidades confiables accedan a los recursos confiables desde las redes esperadas. Para obtener más información, consulte [Crear un perímetro de datos sobre](#) AWS

## preprocesamiento de datos

Transformar los datos sin procesar en un formato que su modelo de ML pueda analizar fácilmente. El preprocesamiento de datos puede implicar eliminar determinadas columnas o filas y corregir los valores faltantes, incoherentes o duplicados.

## procedencia de los datos

El proceso de rastrear el origen y el historial de los datos a lo largo de su ciclo de vida, por ejemplo, la forma en que se generaron, transmitieron y almacenaron los datos.

## titular de los datos

Persona cuyos datos se recopilan y procesan.

## almacenamiento de datos

Un sistema de administración de datos que respalde la inteligencia empresarial, como el análisis. Los almacenes de datos suelen contener grandes cantidades de datos históricos y, por lo general, se utilizan para consultas y análisis.

## lenguaje de definición de datos (DDL)

Instrucciones o comandos para crear o modificar la estructura de tablas y objetos de una base de datos.

## lenguaje de manipulación de datos (DML)

Instrucciones o comandos para modificar (insertar, actualizar y eliminar) la información de una base de datos.

## DDL

Consulte el [lenguaje de definición de bases de datos](#) de datos.

## conjunto profundo

Combinar varios modelos de aprendizaje profundo para la predicción. Puede utilizar conjuntos profundos para obtener una predicción más precisa o para estimar la incertidumbre de las predicciones.

## aprendizaje profundo

Un subcampo del ML que utiliza múltiples capas de redes neuronales artificiales para identificar el mapeo entre los datos de entrada y las variables objetivo de interés.

## defense-in-depth

Un enfoque de seguridad de la información en el que se distribuyen cuidadosamente una serie de mecanismos y controles de seguridad en una red informática para proteger la confidencialidad, la integridad y la disponibilidad de la red y de los datos que contiene. Cuando se adopta esta estrategia en AWS, se suman varios controles en diferentes capas de la estructura de AWS Organizations para ayudar a proteger los recursos. Por ejemplo, un defense-in-depth enfoque podría combinar la autenticación multifactorial, la segmentación de la red y el cifrado.

## administrador delegado

En AWS Organizations, un servicio compatible puede registrar una cuenta de miembro de AWS a fin de administrar las cuentas de la organización y los permisos para ese servicio. Esta cuenta

se denomina administrador delegado para ese servicio. Para obtener más información y una lista de servicios compatibles, consulte [Servicios que funcionan con AWS Organizations](#) en la documentación de AWS Organizations.

## Implementación

El proceso de hacer que una aplicación, características nuevas o correcciones de código se encuentren disponibles en el entorno de destino. La implementación abarca implementar cambios en una base de código y, a continuación, crear y ejecutar esa base en los entornos de la aplicación.

### entorno de desarrollo

[Consulte entorno.](#)

### control de detección

Un control de seguridad que se ha diseñado para detectar, registrar y alertar después de que se produzca un evento. Estos controles son una segunda línea de defensa, ya que lo advierten sobre los eventos de seguridad que han eludido los controles preventivos establecidos. Para obtener más información, consulte [Controles de detección](#) en Implementación de controles de seguridad en AWS.

### asignación de flujos de valor para el desarrollo (DVSM)

Proceso que se utiliza para identificar y priorizar las restricciones que afectan negativamente a la velocidad y la calidad en el ciclo de vida del desarrollo de software. DVSM amplía el proceso de asignación del flujo de valor diseñado originalmente para las prácticas de fabricación ajustada. Se centra en los pasos y los equipos necesarios para crear y transferir valor a través del proceso de desarrollo de software.

### gemelo digital

Representación virtual de un sistema del mundo real, como un edificio, una fábrica, un equipo industrial o una línea de producción. Los gemelos digitales son compatibles con el mantenimiento predictivo, la supervisión remota y la optimización de la producción.

### tabla de dimensiones

En un [esquema en estrella](#), tabla más pequeña que contiene los atributos de datos sobre los datos cuantitativos de una tabla de hechos. Los atributos de la tabla de dimensiones suelen ser campos de texto o números discretos que se comportan como texto. Estos atributos se utilizan habitualmente para restringir consultas, filtrar y etiquetar conjuntos de resultados.

## desastre

Un evento que impide que una carga de trabajo o un sistema cumplan sus objetivos empresariales en su ubicación principal de implementación. Estos eventos pueden ser desastres naturales, fallos técnicos o el resultado de acciones humanas, como una configuración incorrecta involuntaria o un ataque de malware.

## recuperación de desastres (DR)

La estrategia y el proceso que se utilizan para minimizar el tiempo de inactividad y la pérdida de datos ocasionados por un [desastre](#). Para obtener más información, consulte [Recuperación de desastres de cargas de trabajo en AWS: Recuperación en la nube](#) en un marco Well-Architected AWS.

## DML

Consulte el [lenguaje de manipulación de bases](#) de datos.

## diseño basado en el dominio

Un enfoque para desarrollar un sistema de software complejo mediante la conexión de sus componentes a dominios en evolución, o a los objetivos empresariales principales, a los que sirve cada componente. Este concepto lo introdujo Eric Evans en su libro, *Diseño impulsado por el dominio: abordando la complejidad en el corazón del software* (Boston: Addison-Wesley Professional, 2003). Para obtener información sobre cómo utilizar el diseño basado en dominios con el patrón de higos estranguladores, consulte [Modernización gradual de los servicios web antiguos de Microsoft ASP.NET \(ASMX\) mediante contenedores y Amazon API Gateway](#).

## DR

Consulte [recuperación ante desastres](#).

## detección de deriva

Seguimiento de las desviaciones con respecto a una configuración de referencia. Por ejemplo, puedes usarlo AWS CloudFormation para [detectar desviaciones en los recursos del sistema](#) o puedes usarlo AWS Control Tower para [detectar cambios en tu landing zone](#) que puedan afectar al cumplimiento de los requisitos de gobierno.

## DVSM

Consulte [el mapeo del flujo de valor del desarrollo](#).

# E

## EDA

Consulte el [análisis exploratorio de datos](#).

### computación en la periferia

La tecnología que aumenta la potencia de cálculo de los dispositivos inteligentes en la periferia de una red de IoT. En comparación con [la computación en nube, la computación](#) perimetral puede reducir la latencia de la comunicación y mejorar el tiempo de respuesta.

### cifrado

Proceso informático que transforma datos de texto plano, legibles por humanos, en texto cifrado.

### clave de cifrado

Cadena criptográfica de bits aleatorios que se genera mediante un algoritmo de cifrado. Las claves pueden variar en longitud y cada una se ha diseñado para ser impredecible y única.

### endianidad

El orden en el que se almacenan los bytes en la memoria del ordenador. Los sistemas big-endianos almacenan primero el byte más significativo. Los sistemas Little-Endian almacenan primero el byte menos significativo.

### punto de conexión

[Consulte el punto final del servicio](#).

### servicio de punto de conexión

Servicio que puede alojar en una nube privada virtual (VPC) para compartir con otros usuarios. Puede crear un servicio de punto de conexión con AWS PrivateLink y conceder permisos a otras Cuentas de AWS o para entidades principales de AWS Identity and Access Management (IAM). Estas cuentas o entidades principales pueden conectarse a su servicio de punto de conexión de forma privada mediante la creación de puntos de conexión de VPC de interfaz. Para obtener más información, consulte [Creación de un servicio de punto de conexión](#) en la documentación de Amazon Virtual Private Cloud (Amazon VPC).

### cifrado de sobre

El proceso de cifrar una clave de cifrado con otra clave de cifrado. Para obtener más información, consulte [Cifrado de sobre](#) en la documentación de AWS Key Management Service (AWS KMS).

## environment

Una instancia de una aplicación en ejecución. Los siguientes son los tipos de entornos más comunes en la computación en la nube:

- entorno de desarrollo: instancia de una aplicación en ejecución que solo se encuentra disponible para el equipo principal responsable del mantenimiento de la aplicación. Los entornos de desarrollo se utilizan para probar los cambios antes de promocionarlos a los entornos superiores. Este tipo de entorno a veces se denomina entorno de prueba.
- entornos inferiores: todos los entornos de desarrollo de una aplicación, como los que se utilizan para las compilaciones y pruebas iniciales.
- entorno de producción: instancia de una aplicación en ejecución a la que pueden acceder los usuarios finales. En una canalización de CI/CD, el entorno de producción es el último entorno de implementación.
- entornos superiores: todos los entornos a los que pueden acceder usuarios que no sean del equipo de desarrollo principal. Esto puede incluir un entorno de producción, entornos de preproducción y entornos para las pruebas de aceptación por parte de los usuarios.

## epopeya

En las metodologías ágiles, son categorías funcionales que ayudan a organizar y priorizar el trabajo. Las epopeyas brindan una descripción detallada de los requisitos y las tareas de implementación. Por ejemplo, las epopeyas de seguridad de AWS CAF incluyen la administración de identidades y accesos, los controles de detección, la seguridad de la infraestructura, la protección de datos y la respuesta a incidentes. Para obtener más información sobre las epopeyas en la estrategia de migración de AWS, consulte la [Guía de implementación del programa](#).

## análisis de datos de tipo exploratorio (EDA)

El proceso de analizar un conjunto de datos para comprender sus características principales. Se recopilan o agregan datos y, a continuación, se realizan las investigaciones iniciales para encontrar patrones, detectar anomalías y comprobar las suposiciones. El EDA se realiza mediante el cálculo de estadísticas resumidas y la creación de visualizaciones de datos.



# F

## tabla de datos

La tabla central de un [esquema en forma de estrella](#). Almacena datos cuantitativos sobre las operaciones comerciales. Normalmente, una tabla de hechos contiene dos tipos de columnas: las que contienen medidas y las que contienen una clave externa para una tabla de dimensiones.

## fallan rápidamente

Una filosofía que utiliza pruebas frecuentes e incrementales para reducir el ciclo de vida del desarrollo. Es una parte fundamental de un enfoque ágil.

## límite de aislamiento de fallas

En elNube de AWS, un límite, como una zona de disponibilidadRegión de AWS, un plano de control o un plano de datos, que limita el efecto de una falla y ayuda a mejorar la resiliencia de las cargas de trabajo. Para obtener más información, consulte [Límites de AWS aislamiento](#) de errores.

## rama de característica

Consulte la [sucursal](#).

## características

Los datos de entrada que se utilizan para hacer una predicción. Por ejemplo, en un contexto de fabricación, las características pueden ser imágenes que se capturan periódicamente desde la línea de fabricación.

## importancia de las características

La importancia que tiene una característica para las predicciones de un modelo. Por lo general, esto se expresa como una puntuación numérica que se puede calcular mediante diversas técnicas, como las explicaciones aditivas de Shapley (SHAP) y los gradientes integrados. Para obtener más información, consulte [Interpretabilidad del modelo de aprendizaje automático con: AWS](#).

## transformación de funciones

Optimizar los datos para el proceso de ML, lo que incluye enriquecer los datos con fuentes adicionales, escalar los valores o extraer varios conjuntos de información de un solo campo de datos. Esto permite que el modelo de ML se beneficie de los datos. Por ejemplo, si divide la fecha del “27 de mayo de 2021 00:15:37” en “jueves”, “mayo”, “2021” y “15”, puede ayudar al algoritmo

de aprendizaje a aprender patrones matizados asociados a los diferentes componentes de los datos.

## FGAC

Consulte el control [de acceso detallado](#).

control de acceso preciso (FGAC)

El uso de varias condiciones que tienen por objetivo permitir o denegar una solicitud de acceso.  
migración relámpago

Método de migración de bases de datos que utiliza la replicación continua de datos mediante la [captura de datos modificados](#) para migrar los datos en el menor tiempo posible, en lugar de utilizar un enfoque gradual. El objetivo es reducir al mínimo el tiempo de inactividad.

## G

bloqueo geográfico

Consulta [las restricciones geográficas](#).

restricciones geográficas (bloqueo geográfico)

En Amazon CloudFront, una opción para impedir que los usuarios de países específicos accedan a las distribuciones de contenido. Puede utilizar una lista de permitidos o bloqueados para especificar los países aprobados y prohibidos. Para obtener más información, consulta [la sección Restringir la distribución geográfica del contenido](#) en la CloudFront documentación.

Flujo de trabajo de Gitflow

Un enfoque en el que los entornos inferiores y superiores utilizan diferentes ramas en un repositorio de código fuente. El flujo de trabajo de Gitflow se considera heredado, y el [flujo de trabajo basado en enlaces troncales](#) es el enfoque moderno preferido.

estrategia de implementación desde cero

La ausencia de infraestructura existente en un entorno nuevo. Al adoptar una estrategia de implementación desde cero para una arquitectura de sistemas, puede seleccionar todas las tecnologías nuevas sin que estas deban ser compatibles con una infraestructura existente, lo que también se conoce como [implementación sobre infraestructura existente](#). Si está ampliando la infraestructura existente, puede combinar las estrategias de implementación sobre infraestructuras existentes y de implementación desde cero.

## barrera de protección

Una regla de alto nivel que ayuda a regular los recursos, las políticas y la conformidad en todas las unidades organizativas (OU). Las barreras de protección preventivas aplican políticas para garantizar la alineación con los estándares de conformidad. Se implementan mediante políticas de control de servicios y límites de permisos de IAM. Las barreras de protección de detección detectan las vulneraciones de las políticas y los problemas de conformidad, y generan alertas para su corrección. Se implementan mediante Amazon AWS Config, AWS Security Hub, GuardDuty, AWS Trusted Advisor, Amazon Inspector y AWS Lambda cheques personalizados.

## H

### JA

Consulte [alta disponibilidad](#).

### migración heterogénea de bases de datos

Migración de la base de datos de origen a una base de datos de destino que utilice un motor de base de datos diferente (por ejemplo, de Oracle a Amazon Aurora). La migración heterogénea suele ser parte de un esfuerzo de rediseño de la arquitectura y convertir el esquema puede ser una tarea compleja. [AWS ofrece AWS SCT](#), lo cual ayuda con las conversiones de esquemas.

### alta disponibilidad (HA)

La capacidad de una carga de trabajo para funcionar de forma continua, sin intervención, en caso de desafíos o desastres. Los sistemas de alta disponibilidad están diseñados para realizar una conmutación por error automática, ofrecer un rendimiento de alta calidad de forma constante y gestionar diferentes cargas y fallos con un impacto mínimo en el rendimiento.

### modernización histórica

Un enfoque utilizado para modernizar y actualizar los sistemas de tecnología operativa (TO) a fin de satisfacer mejor las necesidades de la industria manufacturera. Un histórico es un tipo de base de datos que se utiliza para recopilar y almacenar datos de diversas fuentes en una fábrica.

### migración homogénea de bases de datos

Migración de la base de datos de origen a una base de datos de destino que comparte el mismo motor de base de datos (por ejemplo, Microsoft SQL Server a Amazon RDS para SQL Server). La migración homogénea suele formar parte de un esfuerzo para volver a alojar o redefinir la plataforma. Puede utilizar las utilidades de bases de datos nativas para migrar el esquema.

## datos recientes

Datos a los que se accede con frecuencia, como datos en tiempo real o datos traslacionales recientes. Por lo general, estos datos requieren un nivel o una clase de almacenamiento de alto rendimiento para proporcionar respuestas rápidas a las consultas.

## hotfix

Una solución urgente para un problema crítico en un entorno de producción. Debido a su urgencia, las revisiones suelen realizarse fuera del flujo de trabajo habitual de las DevOps versiones.

## periodo de hiperatención

Periodo, inmediatamente después de la transición, durante el cual un equipo de migración administra y monitorea las aplicaciones migradas en la nube para solucionar cualquier problema. Por lo general, este periodo dura de 1 a 4 días. Al final del periodo de hiperatención, el equipo de migración suele transferir la responsabilidad de las aplicaciones al equipo de operaciones en la nube.

## I

## laC

Vea [la infraestructura como código](#).

## políticas basadas en identidad

Una política asociada a una o más entidades principales de IAM que define sus permisos en el entorno de la Nube de AWS.

## aplicación inactiva

Aplicación que utiliza un promedio de CPU y memoria de entre 5 y 20 por ciento durante un periodo de 90 días. En un proyecto de migración, es habitual retirar estas aplicaciones o mantenerlas en las instalaciones.

## IIoT

Véase el [Internet industrial de las cosas](#).

## infraestructura inmutable

Un modelo que implementa una nueva infraestructura para las cargas de trabajo de producción en lugar de actualizar, parchear o modificar la infraestructura existente. [Las infraestructuras](#)

[inmutables son intrínsecamente más consistentes, fiables y predecibles que las infraestructuras mutables](#). Para obtener más información, consulte las prácticas recomendadas para [implementar con una infraestructura inmutable](#) en Well-Architected FrameworkAWS.

## VPC entrante (de entrada)

En una arquitectura de varias cuentas de AWS, una VPC que acepta, inspecciona y enruta las conexiones de red desde fuera de una aplicación. La [Arquitectura de referencia de seguridad de AWS](#) recomienda configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

## migración gradual

Estrategia de transición en la que se migra la aplicación en partes pequeñas en lugar de realizar una transición única y completa. Por ejemplo, puede trasladar inicialmente solo unos pocos microservicios o usuarios al nuevo sistema. Tras comprobar que todo funciona correctamente, puede trasladar microservicios o usuarios adicionales de forma gradual hasta que pueda retirar su sistema heredado. Esta estrategia reduce los riesgos asociados a las grandes migraciones.

## infraestructura

Todos los recursos y activos que se encuentran en el entorno de una aplicación.

## infraestructura como código (IaC)

Proceso de aprovisionamiento y administración de la infraestructura de una aplicación mediante un conjunto de archivos de configuración. La IaC se ha diseñado para ayudarlo a centralizar la administración de la infraestructura, estandarizar los recursos y escalar con rapidez a fin de que los entornos nuevos sean repetibles, fiables y consistentes.

## Internet de las cosas industrial (IIoT)

El uso de sensores y dispositivos conectados a Internet en los sectores industriales, como el productivo, el eléctrico, el automotriz, el sanitario, el de las ciencias de la vida y el de la agricultura. Para obtener más información, consulte [Creación de una estrategia de transformación digital del Internet de las cosas industrial \(IIoT\)](#).

## VPC de inspección

En una arquitectura de varias cuentas de AWS, una VPC centralizada que administra las inspecciones del tráfico de red entre VPC (en la misma o en diferentes Regiones de AWS), Internet y las redes en las instalaciones. La [Arquitectura de referencia de seguridad de AWS](#) recomienda configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

## Internet de las cosas (IoT)

Red de objetos físicos conectados con sensores o procesadores integrados que se comunican con otros dispositivos y sistemas a través de Internet o de una red de comunicación local. Para obtener más información, consulte [¿Qué es IoT?](#).

## interpretabilidad

Característica de un modelo de machine learning que describe el grado en que un ser humano puede entender cómo las predicciones del modelo dependen de sus entradas. Para más información, consulte [Interpretabilidad del modelo de machine learning con AWS](#).

## IoT

[Consulte Internet de las cosas.](#)

## biblioteca de información de TI (ITIL)

Conjunto de prácticas recomendadas para ofrecer servicios de TI y alinearlos con los requisitos empresariales. La ITIL proporciona la base para la ITSM.

## administración de servicios de TI (ITSM)

Actividades asociadas con el diseño, la implementación, la administración y el soporte de los servicios de TI para una organización. Para obtener información sobre la integración de las operaciones en la nube con las herramientas de ITSM, consulte la [Guía de integración de operaciones](#).

## ITIL

Consulte la [biblioteca de información de TI](#).

## ITSM

Consulte [Administración de servicios de TI](#).

## L

## control de acceso basado en etiquetas (LBAC)

Una implementación del control de acceso obligatorio (MAC) en la que a los usuarios y a los propios datos se les asigna explícitamente un valor de etiqueta de seguridad. La intersección entre la etiqueta de seguridad del usuario y la etiqueta de seguridad de los datos determina qué filas y columnas puede ver el usuario.

## zona de aterrizaje

Una zona de aterrizaje es un entorno de AWS correctamente diseñado, con varias cuentas, que es escalable y seguro. Este es un punto de partida desde el cual las empresas pueden lanzar e implementar rápidamente cargas de trabajo y aplicaciones con confianza en su entorno de seguridad e infraestructura. Para obtener más información sobre las zonas de aterrizaje, consulte [Configuración de un entorno de AWS seguro y escalable con varias cuentas](#).

## migración grande

Migración de 300 servidores o más.

## LBAC

Consulte control de [acceso basado en etiquetas](#).

## privilegio mínimo

La práctica recomendada de seguridad que consiste en conceder los permisos mínimos necesarios para realizar una tarea. Para obtener más información, consulte [Aplicar permisos de privilegio mínimo](#) en la documentación de IAM.

## migrar mediante lift-and-shift

Ver [7 Rs](#).

## sistema little-endian

Un sistema que almacena primero el byte menos significativo. Véase también [endianness](#).

## entornos inferiores

[Véase entorno](#).

# M

## machine learning (ML)

Un tipo de inteligencia artificial que utiliza algoritmos y técnicas para el reconocimiento y el aprendizaje de patrones. El ML analiza y aprende de los datos registrados, como los datos del Internet de las cosas (IoT), para generar un modelo estadístico basado en patrones. Para más información, consulte [Machine learning](#).

## rama principal

Ver [sucursal](#).

## servicios gestionados

Servicios de AWS para los que AWS opera la capa de infraestructura, el sistema operativo y las plataformas, y usted accede a los puntos finales para almacenar y recuperar datos. Amazon Simple Storage Service (Amazon S3) y Amazon DynamoDB son ejemplos de servicios gestionados. También se conocen como servicios abstractos.

## MAP

Consulte [Migration Acceleration Program](#).

## mecanismo

Un proceso completo en el que se crea una herramienta, se impulsa su adopción y, a continuación, se inspeccionan los resultados para realizar los ajustes necesarios. Un mecanismo es un ciclo que se refuerza y mejora a sí mismo a medida que funciona. Para obtener más información, consulte [Creación de mecanismos](#) en el AWS Well-Architected Framework.

## cuenta de miembro

Todas las Cuentas de AWS distintas de las cuentas de administración que forman parte de una organización en AWS Organizations. Una cuenta no puede pertenecer a más de una organización a la vez.

## microservicio

Un servicio pequeño e independiente que se comunica a través de API bien definidas y que, por lo general, es propiedad de equipos pequeños e independientes. Por ejemplo, un sistema de seguros puede incluir microservicios que se adapten a las capacidades empresariales, como las de ventas o marketing, o a subdominios, como las de compras, reclamaciones o análisis. Los beneficios de los microservicios incluyen la agilidad, la escalabilidad flexible, la facilidad de implementación, el código reutilizable y la resiliencia. Para obtener más información, consulte [Integración de microservicios mediante servicios sin servidor de AWS](#).

## arquitectura de microservicios

Un enfoque para crear una aplicación con componentes independientes que ejecutan cada proceso de la aplicación como un microservicio. Estos microservicios se comunican a través de una interfaz bien definida mediante API ligeras. Cada microservicio de esta arquitectura se puede actualizar, implementar y escalar para satisfacer la demanda de funciones específicas de una aplicación. Para obtener más información, consulte [Implementación de microservicios en AWS](#).



## Programa de aceleración de la migración (MAP)

Programa de AWS que brinda soporte de consultoría, capacitación y servicios para ayudar a las empresas a construir una base operativa sólida para migrar a la nube y ayudar a compensar el costo inicial de las migraciones. El MAP incluye una metodología de migración para ejecutar las migraciones antiguas de forma metódica y un conjunto de herramientas para automatizar y acelerar los escenarios de migración más comunes.

### migración a escala

Proceso de transferencia de la mayoría de la cartera de aplicaciones a la nube en oleadas, con más aplicaciones desplazadas a un ritmo más rápido en cada oleada. En esta fase, se utilizan las prácticas recomendadas y las lecciones aprendidas en las fases anteriores para implementar una fábrica de migración de equipos, herramientas y procesos con el fin de agilizar la migración de las cargas de trabajo mediante la automatización y la entrega ágil. Esta es la tercera fase de la [estrategia de migración de AWS](#).

### fábrica de migración

Equipos multifuncionales que agilizan la migración de las cargas de trabajo mediante enfoques automatizados y ágiles. Los equipos de Migration Factory suelen incluir operadores, analistas de negocio y propietarios, ingenieros de migración, desarrolladores y DevOps profesionales que trabajan a pasos agigantados. Entre el 20 y el 50 por ciento de la cartera de aplicaciones empresariales se compone de patrones repetidos que pueden optimizarse mediante un enfoque de fábrica. Para obtener más información, consulte la [discusión sobre las fábricas de migración](#) y la [Guía de fábricas de migración a la nube](#) en este contenido.

### metadatos de migración

Información sobre la aplicación y el servidor que se necesita para completar la migración. Cada patrón de migración requiere un conjunto diferente de metadatos de migración. Algunos ejemplos de metadatos de migración son las subredes de destino, los grupos de seguridad y las cuentas de AWS.

### patrón de migración

Tarea de migración repetible que detalla la estrategia de migración, el destino de la migración y la aplicación o el servicio de migración utilizados. Ejemplo: Volver a alojar la migración en Amazon EC2 con AWS Application Migration Service.

## Migration Portfolio Assessment (MPA)

Herramienta en línea que brinda información a fin de validar los argumentos comerciales necesarios para migrar a la nube de AWS. La MPA ofrece una evaluación detallada de la cartera (adecuación del tamaño de los servidores, precios, comparaciones del costo total de propiedad, análisis de los costos de migración), así como una planificación de la migración (análisis y recopilación de datos de aplicaciones, agrupación de aplicaciones, priorización de la migración y planificación de oleadas). La [herramienta MPA](#) (requiere inicio de sesión) está disponible de forma gratuita para todos los consultores de AWS y los consultores asociados de APN.

## Evaluación de la preparación para la migración (MRA)

Proceso que consiste en obtener información sobre el estado de preparación de la nube de una organización, identificar los puntos fuertes y débiles, y elaborar un plan de acción para cerrar las brechas identificadas, mediante AWS CAF. Para obtener más información, consulte la [Guía de preparación para la migración](#). La MRA es la primera fase de la [estrategia de migración de AWS](#).

## estrategia de migración

Enfoque utilizado para migrar una carga de trabajo a la nube de AWS. Para obtener más información, consulte la entrada de las [7 R](#) de este glosario y consulte [Movilice a su organización para acelerar las migraciones a gran escala](#).

## ML

[Consulte el aprendizaje automático.](#)

## MAPA

Consulte [la evaluación de la cartera de migración](#).

## modernización

Transformar una aplicación obsoleta (antigua o monolítica) y su infraestructura en un sistema ágil, elástico y de alta disponibilidad en la nube para reducir los gastos, aumentar la eficiencia y aprovechar las innovaciones. Para obtener más información, consulte [Estrategia para modernizar las aplicaciones en la Nube de AWS](#).

## evaluación de la preparación para la modernización

Evaluación que ayuda a determinar la preparación para la modernización de las aplicaciones de una organización; identifica los beneficios, los riesgos y las dependencias; y determina qué tan bien la organización puede soportar el estado futuro de esas aplicaciones. El resultado de la evaluación es un esquema de la arquitectura objetivo, una hoja de ruta que detalla las fases de desarrollo y los hitos del proceso de modernización y un plan de acción para abordar las

brechas identificadas. Para obtener más información, consulte [Evaluación de la preparación para la modernización de las aplicaciones en la nube de AWS](#).

#### aplicaciones monolíticas (monolitos)

Aplicaciones que se ejecutan como un único servicio con procesos estrechamente acoplados. Las aplicaciones monolíticas presentan varios inconvenientes. Si una característica de la aplicación experimenta un aumento en la demanda, se debe escalar toda la arquitectura. Agregar o mejorar las características de una aplicación monolítica también se vuelve más complejo a medida que crece la base de código. Para solucionar problemas con la aplicación, puede utilizar una arquitectura de microservicios. Para obtener más información, consulte [Descomposición de monolitos en microservicios](#).

#### clasificación multiclase

Un proceso que ayuda a generar predicciones para varias clases (predice uno de más de dos resultados). Por ejemplo, un modelo de ML podría preguntar “¿Este producto es un libro, un automóvil o un teléfono?” o “¿Qué categoría de productos es más interesante para este cliente?”.

#### infraestructura mutable

Un modelo que actualiza y modifica la infraestructura existente para las cargas de trabajo de producción. Para mejorar la coherencia, la fiabilidad y la previsibilidad, el AWS Well-Architected Framework recomienda el uso [de una infraestructura inmutable](#) como práctica recomendada.

## O

### OAC

[Consulte el control de acceso de origen.](#)

### OAI

Consulte la [identidad de acceso de origen](#).

### OCM

Consulte [gestión del cambio organizacional](#).

#### migración fuera de línea

Método de migración en el que la carga de trabajo de origen se elimina durante el proceso de migración. Este método implica un tiempo de inactividad prolongado y, por lo general, se utiliza para cargas de trabajo pequeñas y no críticas.

## OI

Consulte [integración de operaciones](#).

## OLA

Véase el [acuerdo a nivel operativo](#).

## migración en línea

Método de migración en el que la carga de trabajo de origen se copia al sistema de destino sin que se desconecte. Las aplicaciones que están conectadas a la carga de trabajo pueden seguir funcionando durante la migración. Este método implica un tiempo de inactividad nulo o mínimo y, por lo general, se utiliza para cargas de trabajo de producción críticas.

## acuerdo de nivel operativo (OLA)

Acuerdo que aclara lo que los grupos de TI operativos se comprometen a ofrecerse entre sí, para respaldar un acuerdo de nivel de servicio (SLA).

## revisión de la preparación operativa (ORR)

Una lista de preguntas y las mejores prácticas asociadas que le ayudan a comprender, evaluar, prevenir o reducir el alcance de los incidentes y posibles fallos. Para obtener más información, consulte [Operational Readiness Reviews \(ORR\)](#) en AWS Well-Architected Framework.

## integración de operaciones (OI)

Proceso de modernización de las operaciones en la nube, que implica la planificación de la preparación, la automatización y la integración. Para obtener más información, consulte la [Guía de integración de las operaciones](#).

## registro de seguimiento organizativo

Registro de seguimiento creado por AWS CloudTrail que registra todos los eventos para todas las Cuentas de AWS en una organización en AWS Organizations. Este registro de seguimiento se crea en cada Cuenta de AWS que forma parte de la organización y realiza un seguimiento de la actividad en cada cuenta. Para obtener más información, consulte [Crear un registro para una organización](#) en la CloudTrail documentación.

## administración del cambio organizacional (OCM)

Marco para administrar las transformaciones empresariales importantes y disruptivas desde la perspectiva de las personas, la cultura y el liderazgo. La OCM ayuda a las empresas a prepararse para nuevos sistemas y estrategias y a realizar la transición a ellos, al acelerar la adopción de cambios, abordar los problemas de transición e impulsar cambios culturales y organizacionales.

En la estrategia de migración de AWS, este marco se denomina aceleración de personas, debido a la velocidad de cambio requerida en los proyectos de adopción de la nube. Para obtener más información, consulte la [Guía de OCM](#).

#### control de acceso de origen (OAC)

En CloudFront, una opción mejorada para restringir el acceso y proteger el contenido del Amazon Simple Storage Service (Amazon S3). El OAC es compatible con todos los buckets de S3 en todas las Regiones de AWS, cifrado del servidor con AWS KMS (SSE-KMS), y solicitudes PUT y DELETE dinámicas al bucket de S3.

#### identidad de acceso de origen (OAI)

En CloudFront, una opción para restringir el acceso y proteger el contenido de Amazon S3. Cuando utiliza OAI, CloudFront crea un principal con el que Amazon S3 puede autenticarse. Los directores autenticados solo pueden acceder al contenido de un bucket de S3 a través de una distribución específica. CloudFront Consulte también el [OAC](#), que proporciona un control de acceso más detallado y mejorado.

## O

Consulte la [revisión de la preparación operativa](#).

#### VPC saliente (de salida)

En una arquitectura de varias cuentas de AWS, una VPC que gestiona las conexiones de red que se inician desde una aplicación. La [Arquitectura de referencia de seguridad de AWS](#) recomienda configurar su cuenta de red con VPC entrantes, salientes y de inspección para proteger la interfaz bidireccional entre su aplicación e Internet en general.

## P

#### límite de permisos

Una política de administración de IAM que se adjunta a las entidades principales de IAM para establecer los permisos máximos que puede tener el usuario o el rol. Para obtener más información, consulte [Límites de permisos](#) en la documentación de IAM.

#### información de identificación personal (PII)

Información que, vista directamente o combinada con otros datos relacionados, puede utilizarse para deducir de manera razonable la identidad de una persona. Algunos ejemplos de información de identificación personal son los nombres, las direcciones y la información de contacto.

## P11

Consulte la [información de identificación personal](#).

### manual de estrategias

Conjunto de pasos predefinidos que capturan el trabajo asociado a las migraciones, como la entrega de las funciones de operaciones principales en la nube. Un manual puede adoptar la forma de scripts, manuales de procedimientos automatizados o resúmenes de los procesos o pasos necesarios para operar un entorno modernizado.

### política

Objeto que puede definir los permisos (consulte la [política basada en la identidad](#)), especifique las condiciones de acceso (consulte la [política basada en los recursos](#)) o defina los permisos máximos para todas las cuentas de una organización AWS Organizations (consulte la política de control de [servicios](#)).

### persistencia políglota

Elegir de forma independiente la tecnología de almacenamiento de datos de un microservicio en función de los patrones de acceso a los datos y otros requisitos. Si sus microservicios tienen la misma tecnología de almacenamiento de datos, pueden enfrentarse a desafíos de implementación o experimentar un rendimiento deficiente. Los microservicios se implementan más fácilmente y logran un mejor rendimiento y escalabilidad si utilizan el almacén de datos que mejor se adapte a sus necesidades. Para obtener más información, consulte [Habilitación de la persistencia de datos en los microservicios](#).

### evaluación de cartera

Proceso de detección, análisis y priorización de la cartera de aplicaciones para planificar la migración. Para obtener más información, consulte la [Evaluación de la preparación para la migración](#).

### predicate

Una condición de consulta que devuelve `true` o `false`, por lo general, se encuentra en una cláusula `WHERE`.

### pulsar un predicado

Técnica de optimización de consultas de bases de datos que filtra los datos de la consulta antes de transferirlos. Esto reduce la cantidad de datos que se deben recuperar y procesar de la base de datos relacional y mejora el rendimiento de las consultas.

## control preventivo

Un control de seguridad diseñado para evitar que ocurra un evento. Estos controles son la primera línea de defensa para evitar el acceso no autorizado o los cambios no deseados en la red. Para obtener más información, consulte [Controles preventivos](#) en Implementación de controles de seguridad en AWS.

## entidad principal

Una entidad de AWS que puede realizar acciones y obtener acceso a los recursos. Esta entidad suele ser un usuario raíz de una Cuenta de AWS, un rol de IAM o un usuario. Para obtener más información, consulte Entidad principal en [Términos y conceptos de roles](#) en la documentación de IAM.

## Privacidad desde el diseño

Un enfoque de ingeniería de sistemas que tiene en cuenta la privacidad durante todo el proceso de ingeniería.

## zonas alojadas privadas

Contenedor que aloja información acerca de cómo desea que responda Amazon Route 53 a las consultas de DNS de un dominio y sus subdominios en una o varias VPC. Para obtener más información, consulte [Uso de zonas alojadas privadas](#) en la documentación de Route 53.

## control proactivo

Un [control de seguridad](#) diseñado para evitar el despliegue de recursos que no cumplan con las normas. Estos controles escanean los recursos antes de aprovisionarlos. Si el recurso no cumple con el control, significa que no está aprovisionado. Para obtener más información, consulte la [guía de referencia de controles](#) en la AWS Control Tower documentación y consulte [Controles proactivos](#) en Implementación de controles de seguridad en AWS.

## entorno de producción

Consulte [entorno](#).

## seudonimización

El proceso de reemplazar los identificadores personales de un conjunto de datos por valores de marcadores de posición. La seudonimización puede ayudar a proteger la privacidad personal. Los datos seudonimizados siguen considerándose datos personales.

## Q

### plan de consulta

Serie de pasos, como instrucciones, que se utilizan para acceder a los datos de un sistema de base de datos relacional SQL.

### regresión del plan de consulta

El optimizador de servicios de la base de datos elige un plan menos óptimo que antes de un cambio determinado en el entorno de la base de datos. Los cambios en estadísticas, restricciones, configuración del entorno, enlaces de parámetros de consultas y actualizaciones del motor de base de datos PostgreSQL pueden provocar una regresión del plan.

## R

### Matriz RACI

Véase [responsable, responsable, consultado, informado \(RACI\)](#).

### ransomware

Software malicioso que se ha diseñado para bloquear el acceso a un sistema informático o a los datos hasta que se efectúe un pago.

### Matriz RASCI

Véase [responsable, responsable, consultado, informado \(RACI\)](#).

### RCAC

Consulte control de [acceso por filas y columnas](#).

### read replica

Una copia de una base de datos que se utiliza con fines de solo lectura. Puede enrutar las consultas a la réplica de lectura para reducir la carga en la base de datos principal.

### rediseñar

Ver [7 Rs](#).



## objetivo de punto de recuperación (RPO)

La cantidad de tiempo máximo aceptable desde el último punto de recuperación de datos. Esto determina qué se considera una pérdida de datos aceptable entre el último punto de recuperación y la interrupción del servicio.

## objetivo de tiempo de recuperación (RTO)

La demora máxima aceptable entre la interrupción del servicio y el restablecimiento del servicio.

## refactorizar

Ver [7 Rs.](#)

## Región

Conjunto de recursos de AWS que se encuentran en un área geográfica. Cada Región de AWS está aislada y es independiente de las demás para ofrecer tolerancia a errores, estabilidad y resistencia. Para obtener más información, consulte [Administración de Regiones de AWS](#) en Referencia general de AWS.

## regresión

Una técnica de ML que predice un valor numérico. Por ejemplo, para resolver el problema de “¿A qué precio se venderá esta casa?”, un modelo de ML podría utilizar un modelo de regresión lineal para predecir el precio de venta de una vivienda en función de datos conocidos sobre ella (por ejemplo, los metros cuadrados).

## volver a alojar

Ver [7 Rs.](#)

## versión

En un proceso de implementación, el acto de promover cambios en un entorno de producción.

## trasladarse

Ver [7 Rs.](#)

## redefinir la plataforma

Ver [7 Rs.](#)

## recompra

Ver [7 Rs.](#)

## política basada en recursos

Una política asociada a un recurso, como un bucket de Amazon S3, un punto de conexión o una clave de cifrado. Este tipo de política especifica a qué entidades principales se les permite el acceso, las acciones compatibles y cualquier otra condición que deba cumplirse.

## matriz responsable, confiable, consultada e informada (RACI)

Una matriz que define las funciones y responsabilidades de todas las partes involucradas en las actividades de migración y las operaciones de la nube. El nombre de la matriz se deriva de los tipos de responsabilidad definidos en la matriz: responsable (R), contable (A), consultado (C) e informado (I). El tipo de soporte (S) es opcional. Si incluye el soporte, la matriz se denomina matriz RASCI y, si la excluye, se denomina matriz RACI.

## control receptivo

Un control de seguridad que se ha diseñado para corregir los eventos adversos o las desviaciones con respecto a su base de seguridad. Para obtener más información, consulte [Controles receptivos](#) en Implementación de controles de seguridad en AWS.

## retain

Ver [7 Rs](#).

## jubilarse

Ver [7 Rs](#).

## rotación

Proceso de actualizar periódicamente un [secreto](#) para dificultar el acceso de un atacante a las credenciales.

## control de acceso por filas y columnas (RCAC)

El uso de expresiones SQL básicas y flexibles que tienen reglas de acceso definidas. El RCAC consta de permisos de fila y máscaras de columnas.

## RPO

Consulte el [objetivo del punto de recuperación](#).

## RTO

Consulte el [objetivo de tiempo de recuperación](#).

## manual de procedimientos

Conjunto de procedimientos manuales o automatizados necesarios para realizar una tarea específica. Por lo general, se diseñan para agilizar las operaciones o los procedimientos repetitivos con altas tasas de error.

## S

### SAML 2.0

Un estándar abierto que utilizan muchos proveedores de identidad (IdPs). Esta característica permite el inicio de sesión único (SSO) federado a fin de que los usuarios puedan iniciar sesión en la AWS Management Console o llamar a la API de AWS sin necesidad de crear un usuario de IAM para cada persona de la organización. Para obtener más información sobre la federación basada en SAML 2.0, consulte [Acerca de la federación basada en SAML 2.0](#) en la documentación de IAM.

### SCP

Consulte la [política de control de servicios](#).

### secreta

Información confidencial o restringida, como una contraseña o credenciales de usuario, que almacene de forma cifrada. AWS Secrets Manager Se compone del valor secreto y sus metadatos. El valor secreto puede ser binario, una sola cadena o varias cadenas. Para obtener más información, consulte la documentación de [Secret](#) in the Secrets Manager.

### control de seguridad

Barrera de protección técnica o administrativa que impide, detecta o reduce la capacidad de un agente de amenazas para aprovechar una vulnerabilidad de seguridad. Existen cuatro tipos principales de controles de seguridad: [preventivos, de detección](#), con [capacidad](#) de [respuesta](#) y [proactivos](#).

### refuerzo de la seguridad

Proceso de reducir la superficie expuesta a ataques para hacerla más resistente a los ataques. Esto puede incluir acciones, como la eliminación de los recursos que ya no se necesitan, la implementación de prácticas recomendadas de seguridad consistente en conceder privilegios mínimos o la desactivación de características innecesarias en los archivos de configuración.

## sistema de información sobre seguridad y administración de eventos (SIEM)

Herramientas y servicios que combinan sistemas de administración de información sobre seguridad (SIM) y de administración de eventos de seguridad (SEM). Un sistema de SIEM recopila, monitorea y analiza los datos de servidores, redes, dispositivos y otras fuentes para detectar amenazas y brechas de seguridad y generar alertas.

## automatización de la respuesta de seguridad

Una acción predefinida y programada que está diseñada para responder automáticamente a un evento de seguridad o remediarlo. Estas automatizaciones sirven como controles de seguridad [detectables](#) o [adaptables](#) que le ayudan a implementar las mejores prácticas AWS de seguridad. Algunos ejemplos de acciones de respuesta automatizadas incluyen la modificación de un grupo de seguridad de VPC, la aplicación de parches a una instancia de Amazon EC2 o la rotación de credenciales.

## cifrado del servidor

Cifrado de los datos en su destino, por parte del Servicio de AWS que los recibe.

## política de control de servicio (SCP)

Una política que proporciona un control centralizado de los permisos de todas las cuentas de una organización en AWS Organizations. Las SCP definen barreras de protección o establecen límites a las acciones que un administrador puede delegar en los usuarios o roles. Puede utilizar las SCP como listas de permitidos o rechazados, para especificar qué servicios o acciones se encuentra permitidos o prohibidos. Para obtener más información, consulte [Políticas de control de servicio](#) en la documentación de AWS Organizations.

## punto de enlace de servicio

La URL del punto de entrada para un Servicio de AWS. Para conectarse mediante programación a un servicio de destino, puede utilizar un punto de conexión. Para obtener más información, consulte [Puntos de conexión de Servicio de AWS](#) en Referencia general de AWS.

## acuerdo de nivel de servicio (SLA)

Acuerdo que aclara lo que un equipo de TI se compromete a ofrecer a los clientes, como el tiempo de actividad y el rendimiento del servicio.

## indicador de nivel de servicio (SLI)

Medición de un aspecto del rendimiento de un servicio, como la tasa de errores, la disponibilidad o el rendimiento.

## objetivo de nivel de servicio (SLO)

[Una métrica objetivo que representa el estado de un servicio, medido mediante un indicador de nivel de servicio.](#)

## modelo de responsabilidad compartida

Modelo que describe la responsabilidad que comparte con AWS en cuanto a la conformidad y la seguridad en la nube. AWS es responsable de la seguridad de la nube, mientras que usted es responsable de la seguridad en la nube. Para obtener más información, consulte el [Modelo de responsabilidad compartida.](#)

## SIEM

Consulte [la información de seguridad y el sistema de gestión de eventos.](#)

## punto único de fallo (SPOF)

Una falla en un único componente crítico de una aplicación que puede interrumpir el sistema.

## SLA

Consulte el acuerdo [de nivel de servicio.](#)

## SLI

Consulte el indicador de [nivel de servicio.](#)

## ASÍ QUE

Consulte el objetivo de [nivel de servicio.](#)

## split-and-seed modelo

Un patrón para escalar y acelerar los proyectos de modernización. A medida que se definen las nuevas funciones y los lanzamientos de los productos, el equipo principal se divide para crear nuevos equipos de productos. Esto ayuda a ampliar las capacidades y los servicios de su organización, mejora la productividad de los desarrolladores y apoya la innovación rápida. Para obtener más información, consulte [Enfoque gradual para modernizar las aplicaciones en el.](#) Nube de AWS

## SPOF

Consulte el [punto único de falla.](#)

## esquema en forma de estrella

Estructura organizativa de una base de datos que utiliza una tabla de hechos grande para almacenar datos medidos o transaccionales y una o más tablas dimensionales más pequeñas para almacenar los atributos de los datos. Esta estructura está diseñada para usarse en un [almacén de datos](#) o con fines de inteligencia empresarial.

## patrón de higo estrangulador

Un enfoque para modernizar los sistemas monolíticos mediante la reescritura y el reemplazo gradual de las funciones del sistema hasta que se pueda dismantelar el sistema heredado. Este patrón utiliza la analogía de una higuera que crece hasta convertirse en un árbol estable y, finalmente, se apodera y reemplaza a su host. El patrón fue [presentado por Martin Fowler](#) como una forma de gestionar el riesgo al reescribir sistemas monolíticos. Para ver un ejemplo con la aplicación de este patrón, consulte [Modernización gradual de los servicios web antiguos de Microsoft ASP.NET \(ASMX\) mediante contenedores y Amazon API Gateway](#).

## subred

Un intervalo de direcciones IP en la VPC. Una subred debe residir en una sola zona de disponibilidad.

## cifrado simétrico

Un algoritmo de cifrado que utiliza la misma clave para cifrar y descifrar los datos.

## pruebas sintéticas

Probar un sistema de manera que simule las interacciones de los usuarios para detectar posibles problemas o monitorear el rendimiento. Puede usar [Amazon CloudWatch Synthetics](#) para crear estas pruebas.

# T

## etiquetas

Pares de clave y valor que funcionan como metadatos para organizar los recursos de AWS. Las etiquetas pueden ayudarle a administrar, identificar, organizar, buscar y filtrar recursos. Para obtener más información, consulte [Etiquetado de los recursos de AWS](#).

## variable de destino

El valor que intenta predecir en el ML supervisado. Esto también se conoce como variable de resultado. Por ejemplo, en un entorno de fabricación, la variable objetivo podría ser un defecto del producto.

## lista de tareas

Herramienta que se utiliza para hacer un seguimiento del progreso mediante un manual de procedimientos. La lista de tareas contiene una descripción general del manual de procedimientos y una lista de las tareas generales que deben completarse. Para cada tarea general, se incluye la cantidad estimada de tiempo necesario, el propietario y el progreso.

## entorno de prueba

Consulte [entorno](#).

## entrenamiento

Proporcionar datos de los que pueda aprender su modelo de ML. Los datos de entrenamiento deben contener la respuesta correcta. El algoritmo de aprendizaje encuentra patrones en los datos de entrenamiento que asignan los atributos de los datos de entrada al destino (la respuesta que desea predecir). Genera un modelo de ML que captura estos patrones. Luego, el modelo de ML se puede utilizar para obtener predicciones sobre datos nuevos para los que no se conoce el destino.

## puerta de enlace de tránsito

Centro de tránsito de red que puede utilizar para interconectar las VPC y las redes en las instalaciones. Para obtener más información, consulte [¿Qué es una puerta de enlace de tránsito?](#) en la documentación de AWS Transit Gateway.

## flujo de trabajo basado en enlaces troncales

Un enfoque en el que los desarrolladores crean y prueban características de forma local en una rama de característica y, a continuación, combinan esos cambios en la rama principal. Luego, la rama principal se adapta a los entornos de desarrollo, preproducción y producción, de forma secuencial.

## acceso de confianza

Concesión de permisos a un servicio que especifique para realizar tareas en su empresa en AWS Organizations y en sus cuentas en su nombre. El servicio de confianza crea un rol vinculado al servicio en cada cuenta, cuando ese rol es necesario, para realizar las tareas de administración

por usted. Para obtener más información, consulte [Uso de AWS Organizations con otros servicios de AWS](#) en la documentación de AWS Organizations.

## ajuste

Cambiar aspectos de su proceso de formación a fin de mejorar la precisión del modelo de ML. Por ejemplo, puede entrenar el modelo de ML al generar un conjunto de etiquetas, incorporar etiquetas y, luego, repetir estos pasos varias veces con diferentes ajustes para optimizar el modelo.

## equipo de dos pizzas

Un DevOps equipo pequeño al que puedes alimentar con dos pizzas. Un equipo formado por dos integrantes garantiza la mejor oportunidad posible de colaboración en el desarrollo de software.

# U

## incertidumbre

Un concepto que hace referencia a información imprecisa, incompleta o desconocida que puede socavar la fiabilidad de los modelos predictivos de ML. Hay dos tipos de incertidumbre: la incertidumbre epistémica se debe a datos limitados e incompletos, mientras que la incertidumbre aleatoria se debe al ruido y la aleatoriedad inherentes a los datos. Para más información, consulte la guía [Cuantificación de la incertidumbre en los sistemas de aprendizaje profundo](#).

## tareas indiferenciadas

También conocido como tareas arduas, es el trabajo que es necesario para crear y operar una aplicación, pero que no proporciona un valor directo al usuario final ni proporciona una ventaja competitiva. Algunos ejemplos de tareas indiferenciadas son la adquisición, el mantenimiento y la planificación de la capacidad.

## entornos superiores

Ver [entorno](#).



## V

### succión

Una operación de mantenimiento de bases de datos que implica limpiar después de las actualizaciones incrementales para recuperar espacio de almacenamiento y mejorar el rendimiento.

### control de versión

Procesos y herramientas que realizan un seguimiento de los cambios, como los cambios en el código fuente de un repositorio.

### Emparejamiento de VPC

Conexión entre dos VPC que permite enrutar el tráfico mediante direcciones IP privadas. Para obtener más información, consulte [¿Qué es una interconexión de VPC?](#) en la documentación de Amazon VPC.

### vulnerabilidad

Defecto de software o hardware que pone en peligro la seguridad del sistema.

## W

### caché caliente

Un búfer caché que contiene datos actuales y relevantes a los que se accede con frecuencia. La instancia de base de datos puede leer desde la caché del búfer, lo que es más rápido que leer desde la memoria principal o el disco.

### datos tibios

Datos a los que el acceso es infrecuente. Al consultar este tipo de datos, normalmente se aceptan consultas moderadamente lentas.

### función de ventana

Función SQL que realiza un cálculo en un grupo de filas que se relacionan de alguna manera con el registro actual. Las funciones de ventana son útiles para procesar tareas, como calcular una media móvil o acceder al valor de las filas en función de la posición relativa de la fila actual.

## carga de trabajo

Conjunto de recursos y código que ofrece valor comercial, como una aplicación orientada al cliente o un proceso de backend.

## flujo de trabajo

Grupos funcionales de un proyecto de migración que son responsables de un conjunto específico de tareas. Cada flujo de trabajo es independiente, pero respalda a los demás flujos de trabajo del proyecto. Por ejemplo, el flujo de trabajo de la cartera es responsable de priorizar las aplicaciones, planificar las oleadas y recopilar los metadatos de migración. El flujo de trabajo de la cartera entrega estos recursos al flujo de trabajo de migración, que luego migra los servidores y las aplicaciones.

## GUSANO

Mira, [escribe una vez, lee muchas](#).

## WQF

Consulte el [marco de calificación de cargas de trabajo de AWS](#).

## escribe una vez, lee muchas (WORM)

Un modelo de almacenamiento que escribe los datos una sola vez y evita que los datos se eliminen o modifiquen. Los usuarios autorizados pueden leer los datos tantas veces como sea necesario, pero no pueden cambiarlos. Esta infraestructura de almacenamiento de datos se considera [inmutable](#).

## Z

### ataque de día cero

Un ataque, normalmente de malware, que aprovecha una vulnerabilidad de [día cero](#).

### vulnerabilidad de día cero

Un defecto o una vulnerabilidad sin mitigación en un sistema de producción. Los agentes de amenazas pueden usar este tipo de vulnerabilidad para atacar el sistema. Los desarrolladores suelen darse cuenta de la vulnerabilidad a raíz del ataque.

### aplicación zombi

Aplicación que utiliza un promedio de CPU y memoria menor al 5 por ciento. En un proyecto de migración, es habitual retirar estas aplicaciones.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la versión original de inglés, prevalecerá la versión en inglés.