



Prácticas recomendadas para las consultas de Amazon Redshift

AWS Orientación prescriptiva



AWS Orientación prescriptiva: Prácticas recomendadas para las consultas de Amazon Redshift

Copyright © 2026 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 registradas que no son propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, conectados o patrocinados por Amazon.

Table of Contents

Introducción	1
Descripción general	1
Destinatarios previstos	1
Objetivos	1
Componentes de arquitectura	2
Factores de rendimiento de las consultas	7
Propiedades de la tabla	7
Claves de clasificación	7
Compresión de datos	8
Distribución de datos	8
Información de tablas	8
Configuración del clúster	10
Tipo de nodo	10
Tamaño y cantidad de los nodos y sectores	10
Administración de la carga de trabajo	10
Aceleración de consultas cortas	11
Consulta SQL	11
Estructura de consulta	11
Compilación de código	12
Prácticas recomendadas para las tablas	13
Descripción del funcionamiento de las claves de clasificación	13
Consejos de ajuste de consultas	13
Evaluación de la eficacia de las claves de clasificación	14
Información sobre su tabla	15
Selección del estilo de distribución de tablas adecuado	15
Prácticas recomendadas para las consultas	17
Recomendación de evitar usar la instrucción <code>SELECT * FROM</code>	17
Identificación de los problemas con las consultas	17
Obtención de información resumida sobre la consulta	17
Recomendación de evitar las combinaciones cruzadas	18
Recomendación de evitar las funciones en los predicados de las consultas	18
Recomendación de evitar las conversiones de tipo innecesarias	19
Uso de expresiones CASE en las agregaciones complejas	19
Uso de subconsultas	20

Uso de predicados	20
Adición de predicados para filtrar tablas con combinaciones	21
Uso de los operadores menos costosos para los predicados	21
Uso de las claves de clasificación en las cláusulas GROUP BY	21
Recomendación de aprovechar las vistas materializadas	22
Precaución con las columnas de las cláusulas GROUP BY y ORDER BY	22
Prácticas recomendadas para Redshift Spectrum	23
Inserción de predicados en Redshift Spectrum	24
Consejos de ajuste de consultas para Redshift Spectrum	25
Recursos	26
Historial de documentos	27
Glosario	28
#	28
A	29
B	32
C	34
D	38
E	42
F	44
G	46
H	47
I	49
L	51
M	53
O	57
P	60
Q	63
R	63
S	66
T	70
U	72
V	73
W	73
Z	74
.....	lxxvi

Prácticas recomendadas para las consultas de Amazon Redshift

Ethan Stark, Amazon Web Services (AWS)

junio de 2024 ([historial de documentos](#))

Descripción general

En esta guía se proporcionan recomendaciones y prácticas recomendadas para optimizar el rendimiento de las consultas y tablas en [Amazon Redshift](#). Puede usar Amazon Redshift para consultar petabytes de datos estructurados y semiestructurados en su almacén de datos y su lago de datos mediante SQL estándar. En esta guía también se proporciona una descripción general de los componentes básicos de la arquitectura de un almacén de datos de Amazon Redshift. Este conocimiento, junto con la comprensión de los factores de rendimiento de las consultas (como las propiedades de las tablas, la configuración del clúster y la estructura de las consultas), puede ayudarlo a diseñar tablas y consultas eficientes y eficaces para su almacén de datos de Amazon Redshift.

Destinatarios previstos

Esta guía está prevista para ingenieros de datos, arquitectos de datos y analistas de datos que diseñan o utilizan tablas y consultas en Amazon Redshift.

Objetivos

Esta guía puede serle útil a usted y a su organización para lograr los objetivos siguientes:

- Diseñar tablas para realizar operaciones óptimas de almacenamiento y recuperación de datos
- Diseñar consultas para obtener un rendimiento y un ahorro de costos óptimos
- Optimizar el rendimiento de [Amazon Redshift Spectrum](#) para consultar datos directamente desde los archivos de [Amazon Simple Storage Service \(Amazon S3\)](#)

Componentes de arquitectura de un almacén de datos de Amazon Redshift

Le recomendamos que tenga un conocimiento básico de los principales componentes de la arquitectura de un almacén de datos de Amazon Redshift. Este conocimiento puede ayudarlo a comprender mejor cómo diseñar las consultas y las tablas para obtener un rendimiento óptimo.

Un almacén de datos en Amazon Redshift consta de los siguientes componentes de arquitectura principales:

- **Clústeres:** un clúster, compuesto por uno o más nodos de computación, es el componente de infraestructura principal de un almacén de datos de Amazon Redshift. Los nodos de computación son transparentes para las aplicaciones externas, pero la aplicación cliente solo interactúa directamente con el nodo principal. Los clústeres habituales tienen dos o más nodos de computación. Los nodos de computación se coordinan a través del nodo principal.
- **Nodo principal:** los nodos principales administran las comunicaciones de los programas cliente y todos los nodos de computación. Los nodos principales también preparan los planes para ejecutar una consulta cada vez que se envía una consulta a un clúster. Cuando los planes están listos, el nodo principal compila el código, lo distribuye a los nodos de computación y les asigna sectores de los datos a cada uno para procesar los resultados de la consulta.
- **Nodo de computación:** los nodos de computación ejecutan las consultas. El nodo principal compila un código para los elementos individuales del plan a fin de ejecutar la consulta y lo asigna a los nodos de computación individuales. Los nodos de computación ejecutan el código compilado y envían resultados intermedios de vuelta al nodo principal para su agregación final. Cada nodo de computación tiene su propia CPU dedicada, memoria y almacenamiento en disco integrado. A medida que la carga de trabajo crece, puede aumentar la capacidad de computación y almacenamiento de un clúster aumentando el número de nodos, actualizando el tipo de nodo o ambas.
- **Sector de nodo:** los nodos de computación están divididos en sectores. A cada sector de los nodos de computación se le asigna una parte de la memoria y del espacio en disco del nodo, donde se procesa una parte de la carga de trabajo asignada al nodo. A continuación, los sectores funcionan en paralelo para completar la operación. Los datos se distribuyen entre los sectores en función del [estilo de distribución](#) y la clave de distribución de una tabla en particular. Una distribución uniforme de los datos hace posible que Amazon Redshift asigne las cargas de trabajo de manera uniforme a los sectores y maximiza las ventajas del procesamiento en paralelo. El número de sectores

por nodo de computación se decide en función del tipo de nodo. Para obtener más información, consulte [Clústeres y nodos de Amazon Redshift](#) en la documentación de Amazon Redshift.

- **Procesamiento en paralelo masivo (MPP):** Amazon Redshift utiliza la arquitectura MPP para procesar datos de forma rápida, incluso las consultas complejas y con grandes cantidades de datos. Varios nodos de computación ejecutan el mismo código de consulta en partes de los datos para maximizar el procesamiento en paralelo.
- **Aplicación cliente:** Amazon Redshift se integra a distintas herramientas de carga de datos, extracción, transformación y carga (ETL), inteligencia empresarial (BI), generación de informes, minería de datos y análisis. Todas las aplicaciones cliente se comunican con el clúster solamente mediante el nodo principal.

En el siguiente diagrama se muestra cómo los componentes de la arquitectura de un almacén de datos de Amazon Redshift trabajan juntos para acelerar las consultas.



El ciclo de vida de la consulta consta de siete etapas:

1. Recepción y análisis de consultas:

- El nodo principal recibe la consulta y analiza el SQL.
- El analizador genera un árbol de consultas inicial que representa la estructura lógica de la consulta original.
- Amazon Redshift introduce este árbol de consultas en el optimizador de consultas.

2. Optimización de consultas:

- El optimizador evalúa la consulta y, de ser necesario, la reescribe para maximizar su eficacia.
- Este proceso de optimización puede implicar la creación de varias consultas relacionadas que reemplazan una misma consulta.

3. Generación de planes de consultas:

- El optimizador genera un plan de consultas (o varios planes, si es necesario) para su ejecución.
- El plan de consultas especifica las opciones de ejecución, como los tipos de combinaciones, el orden de las combinaciones, los métodos de agregación y los requisitos de distribución de datos.

4. Traducción del motor de ejecución:

- El motor de ejecución traduce el plan de consultas en pasos, segmentos y flujos separados:
 - Paso: representa una operación individual requerida durante la ejecución de la consulta. Los pasos se pueden combinar para permitir que los nodos de computación realicen consultas, combinaciones y otras operaciones en la base de datos.
 - Segmento: combina varios pasos que puede ejecutar un solo proceso. También es la unidad de compilación más pequeña que puede ejecutar un segmento de nodos de computación. (Un sector es la unidad de procesamiento en paralelo de Amazon Redshift).
 - Flujo: conjunto de segmentos distribuidos en los sectores de nodos de computación disponibles.
- El motor de ejecución genera un código compilado en función de estos pasos, segmentos y flujos. El código compilado se ejecuta más rápido que el código interpretado y consume menor capacidad de computación.
- El nodo principal transmite el código compilado a los nodos de computación.

5. Ejecución en paralelo:

- Este paso se produce una vez para cada flujo.

- Los sectores de los nodos de computación ejecutan los segmentos de la consulta en paralelo. 5

- Durante este proceso, Amazon Redshift optimiza la comunicación de red, el uso de la memoria y la administración de discos para transmitir los resultados intermedios de un paso del plan de consulta al siguiente.
- Esta optimización contribuye a una ejecución más rápida de las consultas.

6. Procesamiento de flujos:

- Este paso se produce una vez para cada flujo.
- El motor crea segmentos ejecutables para cada flujo para lograr un procesamiento en paralelo eficiente.

7. Clasificación y agregación finales:

- El nodo principal gestiona cualquier clasificación o agregación final que necesite la consulta.
- Cuando ha finalizado, el nodo principal devuelve los resultados al cliente.

Para obtener más información sobre los componentes de la arquitectura, consulte [Arquitectura del sistema de almacenamiento de datos](#) en la documentación de Amazon Redshift.

Factores de rendimiento de las consultas para Amazon Redshift

Hay una serie de factores que pueden afectar al rendimiento de las consultas. Los siguientes aspectos de sus datos, su clúster y de las operaciones de la base de datos influyen en el tiempo de procesamiento de sus consultas:

- [Propiedades de la tabla](#)
 - [Claves de clasificación](#) (Amazon Redshift Advisor)
 - [Compresión de datos](#) (automatizado)
 - [Distribución de datos](#) (automatizado)
 - [Información de tablas](#) (automatizado)
- [Configuración del clúster](#)
 - [Tipo de nodo](#)
 - [Tamaño y cantidad de los nodos y sectores](#)
 - [Administración de la carga de trabajo](#) (automatizado)
 - [Aceleración de consultas cortas](#) (automatizado)
- [Consulta SQL](#)
 - [Estructura de consulta](#)
 - [Compilación de código](#)

Propiedades de la tabla

Las tablas de Amazon Redshift son las unidades fundamentales para almacenar datos en Amazon Redshift y cada tabla tiene un conjunto de propiedades que determinan su comportamiento y accesibilidad. Algunas de estas propiedades son la clasificación, el estilo de distribución, la codificación por compresión y muchas otras. Entender estas propiedades es fundamental para optimizar el rendimiento, la seguridad y la rentabilidad de las tablas de Amazon Redshift.

Claves de clasificación

Amazon Redshift almacena los datos en disco en un determinado orden en función de la clave de clasificación de una tabla. El optimizador y el procesador de consultas utilizan la información

de ubicación de los datos en el nodo de computación para reducir la cantidad de bloques que se deben examinar. Esto mejora considerablemente la velocidad de las consultas al reducir la cantidad de datos que se tienen que procesar. Le recomendamos que utilice claves de clasificación para facilitar los filtros de la cláusula WHERE. Para obtener más información, consulte [Uso de las claves de clasificación](#) en la documentación de Amazon Redshift.

Compresión de datos

La compresión de datos reduce los requisitos de almacenamiento, lo que reduce el disco I/O y mejora el rendimiento de las consultas. Cuando se ejecuta una consulta, los datos comprimidos se leen en la memoria y luego se descomprimen cuando se ejecuta la consulta. Al cargar menos datos en la memoria, Amazon Redshift puede asignar más memoria al analizar los datos. Como el almacenamiento en columnas guarda los datos similares de forma secuencial, Amazon Redshift puede aplicar codificaciones de compresión adaptables que estén asociadas específicamente a los tipos de datos en columnas. La mejor forma de habilitar la compresión de datos en las columnas de las tablas es mediante la opción AUTO de Amazon Redshift de aplicar las codificaciones de compresión óptimas cuando se carga la tabla con los datos. Para obtener más información sobre el uso de la compresión automática de datos, consulte [Carga de tablas con compresión automática](#) en la documentación de Amazon Redshift.

Distribución de datos

Amazon Redshift almacena los datos en los nodos de computación en función del estilo de distribución de la tabla. Cuando ejecuta una consulta, el optimizador de consultas redistribuye los datos a los nodos informáticos según se necesite para realizar uniones y agregaciones. Elegir el estilo de distribución correcto para una tabla ayuda a reducir el impacto del paso de redistribución al localizar los datos en el lugar que deben estar, antes de que se realicen las combinaciones. Le recomendamos que utilice claves de distribución para facilitar las combinaciones más comunes. Para obtener más información, consulte [Uso de los estilos de distribución de datos](#) en la documentación de Amazon Redshift.

Información de tablas

Aunque, desde el primer momento, Amazon Redshift proporciona un rendimiento líder en el sector para la mayoría de las cargas de trabajo, mantener los clústeres de Amazon Redshift ejecutándose de forma correcta requiere tareas de mantenimiento. Al actualizar y eliminar datos, se crean filas inactivas que hay que eliminar, e incluso las tablas que solo se adjuntan se tienen que volver a ordenar si el orden de incorporación no es coherente con la clave de clasificación.

Vacuum

El proceso de succión en Amazon Redshift es esencial para el buen estado y el mantenimiento de su clúster de Amazon Redshift. También afecta al rendimiento de las consultas. Como las eliminaciones y las actualizaciones marcan los datos antiguos pero en realidad no los eliminan, debe utilizar la succión para recuperar el espacio en disco que ocupaban las filas de la tabla que las operaciones UPDATE y DELETE anteriores marcaron para su eliminación. Amazon Redshift puede ordenar y realizar una operación VACUUM DELETE de forma automática en las tablas en segundo plano.

Para limpiar las tablas tras una carga o una serie de actualizaciones incrementales, también puede ejecutar el comando VACUUM, ya sea en la base de datos completa o en tablas individuales. Si las tablas tienen claves de clasificación y las cargas de las tablas no se han optimizado para ordenarlas a medida que se insertan, debe utilizar la succión para volver a ordenar los datos (lo que puede ser crucial para el rendimiento). Para obtener más información, consulte [Limpieza de tablas](#) en la documentación de Amazon Redshift.

Análisis

La operación ANALYZE actualiza los metadatos estadísticos de las tablas de una base de datos de Amazon Redshift. Mantener las estadísticas actualizadas mejora el rendimiento de las consultas, ya que permite que el planificador de consultas elija los planes óptimos. Amazon Redshift monitorea constantemente la base de datos y, de forma automática, realiza operaciones de análisis en segundo plano. Para minimizar el impacto en el rendimiento del sistema, la operación ANALYZE se ejecuta automáticamente durante periodos en los que hay poca carga de trabajo. Si decide ejecutar explícitamente el comando ANALYZE, haga lo siguiente:

- Ejecute el comando ANALYZE antes de ejecutar cualquier consulta.
- Ejecute el comando ANALYZE en la base de datos de manera habitual al final de cada ciclo de carga o actualización regular.
- Ejecute el comando ANALYZE en las tablas nuevas que cree y en las tablas o columnas existentes que sufran cambios significativos.
- Considere ejecutar operaciones ANALYZE en programas diferentes para diferentes tipos de tablas y columnas, según su uso en consultas y su tendencia al cambio.
- Para ahorrar tiempo y recursos del clúster, al ejecutar el comando ANALYZE use la cláusula PREDICATE COLUMNS.

Configuración del clúster

Un clúster es una colección de nodos que realizan el almacenamiento y el procesamiento reales de los datos. Es fundamental configurar el clúster de Amazon Redshift correctamente si quiere lograr lo siguiente:

- Alta escalabilidad y simultaneidad
- Uso eficiente de Amazon Redshift
- Mejor rendimiento
- Menor costo

Tipo de nodo

Un clúster de Amazon Redshift puede utilizar uno de varios tipos de nodos (RA3 DC2, y DS2). Cada tipo de nodo ofrece diferentes tamaños y límites para ayudarlo a escalar su clúster de manera adecuada. El tamaño del nodo determina la capacidad de almacenamiento, la memoria, la CPU y el precio de cada nodo del clúster. La optimización del costo y el rendimiento comienza con la elección del tipo y tamaño de nodo correctos. Para obtener más información sobre los tipos de nodos, consulte [Información general de los clústeres de Amazon Redshift](#) en la documentación de Amazon Redshift.

Tamaño y cantidad de los nodos y sectores

Un nodo de computación está particionado en sectores. Si hay más nodos, hay más procesadores y sectores lo que permite procesar sus consultas con mayor rapidez al ejecutar porciones de una consulta de manera simultánea en todos los sectores. Sin embargo, cuantos más nodos, mayor será el gasto. Esto significa que debe encontrar el equilibrio adecuado entre costo y rendimiento para su sistema. Para obtener más información sobre la arquitectura de clústeres de Amazon Redshift, consulte [Arquitectura del sistema de almacenamiento de datos](#) en la documentación de Amazon Redshift.

Administración de la carga de trabajo

La administración de la carga de trabajo (WLM) de Amazon Redshift permite a los usuarios administrar las colas de cargas de trabajo con prioridades de manera flexible, por lo que las consultas cortas y de ejecución rápida no quedarán bloqueadas en las colas detrás de las consultas

de ejecución prolongada. La WLM automática utiliza algoritmos de machine learning (ML) para perfilar las consultas y colocarlas en la cola adecuada con los recursos adecuados, a la vez que administra la simultaneidad de las consultas y la asignación de memoria. Para obtener más información sobre la WLM, consulte [Implementación de la administración de la carga de trabajo](#) en la documentación de Amazon Redshift.

Aceleración de consultas cortas

La aceleración de consultas cortas (SQA) da prioridad a una serie de consultas que se ejecutan rápidamente frente a consultas que tardan más en ejecutarse. SQA ejecuta las consultas en un espacio dedicado, de forma que estas consultas no tienen que esperar en las colas detrás de otras consultas más largas. SQA solamente da prioridad a las consultas de corta ejecución y a las consultas que están en una cola definida por el usuario. Si usa SQA, las consultas cortas se ejecutan con mayor rapidez y tardará menos en ver los resultados. Si habilita SQA, podrá reducir o eliminar las colas de WLM dedicadas a las consultas cortas. Además, las consultas largas no necesitan competir por los espacios de una cola de WLM. Esto significa que puede configurar las colas de WLM para que utilicen menos espacios de consulta. Si se utiliza una simultaneidad más baja, el rendimiento de las consultas aumenta y el rendimiento de todo el sistema mejora con la mayoría de las cargas de trabajo. Para obtener más información sobre SQA, consulte [Uso de la aceleración de consultas cortas](#) en la documentación de Amazon Redshift.

Consulta SQL

Una consulta de base de datos es una solicitud de los datos de una base de datos. La solicitud debe proceder en un clúster de Amazon Redshift mediante SQL. Amazon Redshift admite herramientas de cliente SQL que se conectan a través de Java Database Connectivity (JDBC) y Open Database Connectivity (ODBC). Puede utilizar la mayoría de las herramientas de cliente SQL compatibles con los controladores JDBC u ODBC.

Estructura de consulta

La forma en la que está escrita la consulta afecta el rendimiento. Le recomendamos que escriba consultas de forma que se tenga que procesar y devolver la menor cantidad de datos que sea necesaria para satisfacer sus necesidades. Para obtener más información sobre cómo estructurar las consultas, consulte la sección [Prácticas recomendadas para diseñar consultas de Amazon Redshift](#) de esta guía.

Compilación de código

Amazon Redshift genera y compila código optimizado para cada plan de ejecución de consultas. El código compilado se ejecuta más rápido porque elimina los gastos generales de tener que recurrir a un intérprete. Para minimizar la latencia de las consultas nuevas y, al mismo tiempo, conservar las ventajas de rendimiento del código compilado, Amazon Redshift utiliza una técnica denominada composición. Composition genera una estructura ligera de lógica preexistente para procesar las nuevas consultas de forma inmediata y, al mismo tiempo, compila código altamente optimizado y específico para las consultas en segundo plano. Esto elimina la compilación de la ruta crítica de ejecución de las consultas. Esto significa que las consultas nuevas se inician más rápido y ofrecen un rendimiento coherente con las ejecuciones posteriores.

Amazon Redshift también utiliza un servicio de compilación sin servidor para escalar las compilaciones de consultas más allá de los recursos informáticos de un clúster de Amazon Redshift. Los segmentos de código compilados se almacenan en caché tanto localmente en el clúster como en una caché remota prácticamente ilimitada que persiste después de reiniciarse el clúster. Las ejecuciones posteriores de la misma consulta se realizan en menos tiempo, ya que se puede omitir la fase de compilación. Al utilizar un servicio de compilación escalable, Amazon Redshift compila el código en paralelo para ofrecer un rendimiento rápido y uniforme.

Prácticas recomendadas para el diseño de tablas de Amazon Redshift

En esta sección se brinda información general sobre las prácticas recomendadas para diseñar tablas de bases de datos. Le recomendamos que siga estas prácticas recomendadas para lograr un rendimiento y una eficacia óptimos en las consultas.

Descripción del funcionamiento de las claves de clasificación

Amazon Redshift almacena los datos en el disco en un determinado orden en función de la clave de ordenación. El optimizador de consultas de Amazon Redshift utiliza la ordenación cuando determina cuáles son los planes óptimos de consulta. Para usar las claves de clasificación de forma eficaz, le recomendamos que haga lo siguiente:

- Mantenga la tabla lo más ordenada posible.
- Utilice la clasificación VACUUM para restablecer un rendimiento óptimo.
- Evite comprimir la columna de claves de clasificación.
- Si la clave de clasificación está comprimida y la relación `sortkey1_skew` es significativamente alta, vuelva a crear la tabla sin habilitar la compresión de la clave de clasificación.
- Evite aplicar una función a las columnas de la clave de clasificación. Por ejemplo, en la siguiente consulta, la columna de claves de clasificación `trans_dt : TIMESTAMPTZ` no se usa si convierte su tipo a `DATE`:

```
select order_id, order_amt
from sales
where trans_dt::date = '2021-01-08'::date
```

- Realice las operaciones `INSERT` en el orden de las claves de clasificación.
- Utilice las claves de clasificación en la cláusula `GROUP BY` siempre que sea posible.

Consejos de ajuste de consultas

Le recomendamos que siga estos pasos para ajustar las consultas:

- Ordene siempre las claves de clasificación compuestas desde la cardinalidad más baja hasta la cardinalidad más alta para lograr una eficacia óptima.

- Si la clave principal de una clave de clasificación compuesta es relativamente única (es decir, tiene una cardinalidad alta), evite agregar más columnas a la clave de clasificación. Agregar más columnas tiene poco impacto en el rendimiento de las consultas, pero incrementa los costos de mantenimiento.

Evaluación de la eficacia de las claves de clasificación

Para optimizar las consultas, debe poder evaluar su eficacia. Le recomendamos que utilice la vista [SVL_QUERY_SUMMARY](#) para encontrar información general acerca de la ejecución de una consulta. En esta vista, puede utilizar el atributo `IS_RRSCAN` para determinar si un paso del plan EXPLAIN usa un análisis de rango restringido. También puede utilizar el atributo `rows_pre_filter` para determinar la selectividad de una clave de clasificación.

Además, puede usar una vista de administración de GitHub denominada [v_my_last_query_summary](#). En la vista se muestra información sobre la última consulta que se ejecutó.

La siguiente instrucción muestra cómo obtener información general acerca de la ejecución de una consulta.

```
select lpad(' ',stm+seg+step) || label as label,
       rows,
       bytes,
       is_diskbased,
       is_rrscan,
       rows_pre_filter
from svl_query_summary
where query = pg_last_query_id()
order by stm, seg, step;
```

La consulta anterior devuelve el siguiente ejemplo de salida.

label	<input type="checkbox"/> rows	bytes	is_diskbased	is_rrscan	rows_pre_filter
scan tbl=163860 name=orders	<input type="checkbox"/> 1500000	24000000	f	f	1500000
project	<input type="checkbox"/> 1500000	0	f	f	0
project	<input type="checkbox"/> 1500000	0	f	f	0
hash tbl=968	<input type="checkbox"/> 1500000	24000000	f	f	0
scan tbl=163852 name=lineitem	<input type="checkbox"/> 6001215	144029160	f	t	6001215
project	<input type="checkbox"/> 6001215	0	f	f	0
project	<input type="checkbox"/> 6001215	0	f	f	0
hjoin tbl=968	<input type="checkbox"/> 6001215	0	f	f	0
project	<input type="checkbox"/> 6001215	0	f	f	0
project	<input type="checkbox"/> 6001215	0	f	f	0

Información sobre su tabla

Es importante entender las propiedades fundamentales de la tabla. Para obtener más información acerca de la tabla, siga estos pasos:

- Use [PG_TABLE_DEF](#) para ver información acerca de las columnas de la tabla.
- Use [SVV_TABLE_INFO](#) para consultar información más exhaustiva acerca de una tabla, incluidos el sesgo de distribución de datos, el sesgo de distribución de claves, el tamaño de tabla y las estadísticas.

Selección del estilo de distribución de tablas adecuado

Cuando ejecuta una consulta, el optimizador de consultas redistribuye las filas a los nodos de computación según se necesite para realizar combinaciones y agregaciones. El objetivo al seleccionar un estilo de distribución de tablas es reducir el impacto del paso de redistribución al localizar los datos en el lugar que deben estar antes de que ejecute la consulta.

Recomendamos el siguiente enfoque para elegir el estilo de distribución de tablas adecuado:

- Para evitar la difusión y la redistribución en un plan de ejecución de consultas, coloque las filas dentro del mismo nodo. Por ejemplo, si selecciona DISTKEY, puede distribuir la tabla de hechos y la tabla unidimensional en sus columnas comunes. Seleccione la mayor dimensión según el tamaño del conjunto de datos filtrado. Solo se deben distribuir las filas que se usan en la combinación; por lo tanto, considere el tamaño del conjunto de datos después del filtrado, no el tamaño de la tabla.

- Asegúrese de que no haya asimetría en la columna en la que se cree la clave de distribución. De ser así, un nodo de computación podría realizar más tareas que otros. Si observa que hay asimetría, considere la posibilidad de cambiar la columna de claves de distribución. Se puede considerar que una columna es candidata a contener claves de distribución si sus valores están distribuidos uniformemente o tienen valores cardinales altos.
- Si la tabla utilizada en la condición de combinación es pequeña (menos de 1 GB), considere el estilo de distribución ALL.
- Puede comprimir la clave de distribución, pero debe evitar comprimir la columna de las claves de clasificación (especialmente la primera columna de las claves de clasificación).

Note

Si utiliza la optimización automática de tablas, no necesita elegir el estilo de distribución de la tabla. Para obtener más información, consulte [Uso de la optimización de tablas automática](#) en la documentación de Amazon Redshift. Para que Amazon Redshift elija el estilo de distribución adecuado, especifique AUTO en el estilo de distribución.

Prácticas recomendadas para diseñar consultas de Amazon Redshift

En esta sección se brinda información general sobre las prácticas recomendadas para diseñar consultas. Le recomendamos que siga las prácticas recomendadas de esta sección para lograr un rendimiento y una eficacia óptimos en las consultas.

Recomendación de evitar usar la instrucción SELECT * FROM

Le recomendamos que evite el uso de la instrucción SELECT * FROM. En su lugar, enumere siempre las columnas para su análisis. Esto reduce el tiempo de ejecución de las consultas y los costos de los análisis de las consultas de Amazon Redshift Spectrum.

Ejemplo de lo que se debe evitar

```
select *  
from sales;
```

Ejemplo de práctica recomendada

```
select sales_date, sales_amt  
from sales;
```

Identificación de los problemas con las consultas

Le recomendamos que consulte la vista [STL_ALERT_EVENT_LOG](#) para identificar y corregir posibles problemas con la consulta.

Obtención de información resumida sobre la consulta

Le recomendamos que use las vistas [SVL_QUERY_SUMMARY](#) y [SVL_QUERY_REPORT](#) para obtener información resumida sobre las consultas. Puede utilizar esta información para optimizar sus consultas.

Recomendación de evitar las combinaciones cruzadas

Le recomendamos que evite usar las combinaciones cruzadas a menos que sea absolutamente necesario. Si una condición de combinación, las combinaciones cruzadas son un producto cartesiano de dos tablas. Por lo general, las combinaciones cruzadas se ejecutan como combinaciones de bucles anidados (que son los tipos de combinación más lentos).

Ejemplo de lo que se debe evitar

```
select c.c_name,  
       n.n_name  
from tpch.customer c,  
     tpch.nation n;
```

Ejemplo de práctica recomendada

```
select c.c_name,  
       n.n_name  
from tpch.customer c,  
join tpch.nation n  
  on n.n_nationkey = c.c_nationkey;
```

Recomendación de evitar las funciones en los predicados de las consultas

Le recomendamos que evite el uso de las funciones en los predicados de las consultas. El uso de funciones en los predicados de las consultas puede afectar negativamente al rendimiento, ya que las funciones suelen agregar una sobrecarga de procesamiento adicional a cada fila y suelen ralentizar la ejecución general de la consulta.

Ejemplo de lo que se debe evitar

```
select sum(o_totalprice)  
from tpch.orders  
where datepart(year, o_orderdate) = 1992;
```

Ejemplo de práctica recomendada

```
select sum(o_totalprice)
```

```
from tpch.orders
where o_orderdate between '1992-01-01' and '1992-12-31';
```

Recomendación de evitar las conversiones de tipo innecesarias

Le recomendamos que evite usar las conversiones de tipo innecesarias en las consultas, ya que la conversión de tipos de datos requiere tiempo y recursos y ralentiza la ejecución de las consultas.

Ejemplo de lo que se debe evitar

```
select sum(o_totalprice)
from tpch.orders
where o_ordertime::date = '1992-01-01';
```

Ejemplo de práctica recomendada

```
select sum(o_totalprice)
from tpch.orders
where o_ordertime between '1992-01-01 00:00:00' and '1992-12-31 23:59:59';
```

Uso de expresiones CASE en las agregaciones complejas

Le recomendamos que utilice una [expresión CASE](#) para realizar agregaciones complejas en lugar de seleccionar de la misma tabla varias veces.

Ejemplo de lo que se debe evitar

```
select sum(sales_amt) as us_sales
from sales
where country = 'US';

select sum(sales_amt) as ca_sales
from sales
where country = 'CA';
```

Ejemplo de práctica recomendada

```
select sum(case when country = 'US' then sales_amt end) as us_sales,
```

```
sum(case when country = 'CA' then sales_amt end) as ca_sales
from sales;
```

Uso de subconsultas

Le recomendamos que utilice subconsultas en los casos donde una tabla en la consulta se utilice solamente para condiciones de predicado y la subconsulta devuelva una cantidad pequeña de filas (menos de 200).

Ejemplo de lo que se debe evitar

Si una subconsulta devuelve menos de 200 filas:

```
select sum(order_amt) as total_sales
from sales
where region_key IN
      (select region_key
       from regions
       where state = 'CA');
```

Ejemplo de práctica recomendada

Si una subconsulta devuelve 200 filas o más:

```
select sum(o.order_amt) as total_sales
from sales o
join regions r
  on r.region_key = o.region_key
  and r.state = 'CA';
```

Uso de predicados

Le recomendamos que utilice predicados para restringir el conjunto de datos tanto como sea posible. En SQL, los predicados se utilizan para filtrar y restringir los datos que se devuelven en una consulta. Al especificar las condiciones en un predicado, puede especificar qué filas deben incluirse en los resultados de la consulta en función de las condiciones especificadas. Esto le permite recuperar solo los datos que le interesan y mejora la eficacia y precisión de las consultas. Para obtener más información, consulte [Condiciones](#) en la documentación de Amazon Redshift.

Adición de predicados para filtrar tablas con combinaciones

Le recomendamos que agregue predicados para filtrar tablas que participen en combinaciones, aun cuando se apliquen los mismos filtros. El uso de predicados para filtrar tablas con combinaciones en SQL puede mejorar el rendimiento de las consultas, ya que reduce la cantidad de datos que se deben procesar y el tamaño del conjunto de resultados intermedios. Al especificar las condiciones de la operación de combinación en la cláusula WHERE, el motor de ejecución de consultas puede eliminar las filas que no coincidan con las condiciones antes de combinarlas. En consecuencia, se reduce el tamaño del conjunto de resultados y se incrementa la velocidad de las consultas.

Ejemplo de lo que se debe evitar

```
select p.product_name, sum(o.order_amt)
from sales o
join product p
  on r.product_key = o.product_key
where o.order_date > '2022-01-01';
```

Ejemplo de práctica recomendada

```
select p.product_name, sum(o.order_amt)
from sales o
join product p
  on p.product_key = o.product_key
  and p.added_date > '2022-01-01'
where o.order_date > '2022-01-01';
```

Uso de los operadores menos costosos para los predicados

En el predicado, utilice los operadores menos costosos que pueda. Los operadores de [condiciones de comparación](#) son preferibles a los operadores [LIKE](#). Los operadores LIKE siguen siendo preferibles a los operadores [SIMILAR TO](#) o [POSIX](#).

Uso de las claves de clasificación en las cláusulas GROUP BY

Utilice las claves de clasificación en la cláusula GROUP BY para que el planificador de consultas pueda utilizar la agrupación de manera más eficiente. Una consulta puede considerarse para la agregación en una fase cuando la lista GROUP BY tiene solamente columnas de clave de

clasificación, una de las cuales es también la clave de distribución. Las columnas con clave de clasificación en la lista GROUP BY deben incluir la primera clave de clasificación y, luego, las demás claves de clasificación que desee usar en el orden de la clave de clasificación.

Recomendación de aprovechar las vistas materializadas

Si es posible, reescriba la consulta sustituyendo el código complejo por una vista materializada, lo que mejorará considerablemente el rendimiento de la consulta. Para obtener más información, consulte [Creación de vistas materializadas en Amazon Redshift](#) en la documentación de Amazon Redshift.

Precaución con las columnas de las cláusulas GROUP BY y ORDER BY

Si utiliza las cláusulas GROUP BY y ORDER BY, asegúrese de colocar las columnas en el mismo orden en ambas cláusulas GROUP BY y ORDER BY. GROUP BY requiere implícitamente que los datos estén ordenados. Si la cláusula ORDER BY es diferente, los datos se deben ordenar dos veces.

Ejemplo de lo que se debe evitar

```
select a, b, c, sum(d)
from a_table
group by b, c, a
order by a, b, c
```

Ejemplo de práctica recomendada

```
select a, b, c, sum(d)
from a_table
group by a, b, c
order by a, b, c
```

Prácticas recomendadas para usar Amazon Redshift Spectrum

En esta sección se proporciona información general sobre las prácticas recomendadas para utilizar [Amazon Redshift Spectrum](#). Le recomendamos que siga estas prácticas recomendadas para lograr un rendimiento óptimo al utilizar Redshift Spectrum:

- Tenga en cuenta que los tipos de archivos tienen una influencia importante en el rendimiento de las consultas de Redshift Spectrum. Para mejorar el rendimiento, utilice archivos codificados en columnas como ORC o Parquet y utilice el formato CSV únicamente para las tablas de dimensiones muy pequeñas.
- Utilice las particiones basadas en prefijos para aprovechar la poda de particiones. Esto significa que debe usar filtros que estén vinculados mediante claves a las particiones del lago de datos.
- Redshift Spectrum se escala automáticamente para procesar solicitudes grandes, así que debería hacer todo lo posible en Redshift Spectrum (por ejemplo, insertar predicados).
- Preste atención a los archivos de partición en las columnas que se filtran frecuentemente. Si los datos se dividen en una o más columnas filtradas, Redshift Spectrum puede aprovechar la poda de particiones y omitir los análisis en busca de particiones y archivos innecesarios. Una práctica habitual es particionar datos según el tiempo.
- Puede comprobar la eficacia de sus particiones y la eficiencia de su consulta de Redshift Spectrum mediante la siguiente consulta.

```
Select query,
       segment,
       max(assigned_partitions) as total_partitions,
       max(qualified_partitions) as qualified_partitions
From svl_s3partition
Where query=pg_last_query_id()
Group by 1,2;
```

En la consulta anterior se muestra lo siguiente:

- `total_partitions` — El número de particiones reconocidas por AWS Glue Data Catalog
- `qualified_partitions`: número de prefijos en Amazon Simple Storage Service (Amazon S3) a los que se accede para la consulta de Redshift Spectrum

- También puede comprobar la tabla del sistema SVL_S3QUERY_SUMMARY para obtener información sobre la eficacia de sus particiones y la eficiencia de la consulta de Redshift Spectrum. Para ello, utilice la siguiente instrucción.

```
Select *
From svl_s3query_summary
Where query=pg_last_query_id();
```

En la consulta anterior se devuelve incluso más información, incluidos los valores `is_partitioned`, `s3_scanned_rows/bytes` y `s3_returned_rows/bytes`, además de archivos que indican la eficacia de la poda de particiones.

Inserción de predicados en Redshift Spectrum

El uso de la inserción de predicados evita consumir recursos del clúster de Amazon Redshift. Puede insertar muchas operaciones de SQL a la capa de Redshift Spectrum. Recomendamos que aproveche esta opción siempre que sea posible.

Tenga en cuenta lo siguiente:

- Dentro de la capa de Redshift Spectrum puede evaluar algunos tipos de operaciones de SQL completamente, incluidas las siguientes:
 - Cláusulas GROUP BY
 - Condiciones de comparación y búsqueda de coincidencias de patrones (por ejemplo, LIKE)
 - Funciones de agrupación (por ejemplo, COUNT, SUM, AVG, MIN y MAX)
 - `regex_replace`, `to_upper`, `date_trunc` y otras funciones
- Algunas operaciones, como DISTINCT y ORDER BY, no se pueden enviar a la capa de Redshift Spectrum. Realice ORDER BY solo en el nivel superior de la consulta si es posible, ya que la clasificación se realiza en el nodo principal.
- Examine el plan EXPLAIN de la consulta para comprobar si la inserción de predicados es efectiva. Para encontrar partes de Redshift Spectrum en un comando EXPLAIN, busque estos pasos:
 - S3 Seq Scan
 - S3 HashAggregate
 - S3 Query Scan
 - Escaneo marino PartitionInfo

- **Partition Loop**
- Utilice el menor número de columnas en la consulta. Redshift Spectrum puede eliminar las columnas para analizar si los datos están en formato Parquet u ORC.
- Haga un uso extensivo de las particiones para el procesamiento en paralelo y la eliminación de particiones, e intente que el tamaño de los archivos sea de al menos 64 MB si es posible.
- Establezca `TABLE PROPERTIES 'numRows' = 'nnn'` si usa `CREATE EXTERNAL TABLE` o `ALTER TABLE`. Amazon Redshift no analiza las tablas externas para generar las estadísticas de las tablas que el optimizador de consultas emplea a la hora de crear un plan de consulta. Si no se establecen estas estadísticas, Amazon Redshift supone que las tablas externas son las más grandes, mientras que las tablas locales son las más pequeñas.

Consejos de ajuste de consultas para Redshift Spectrum

Le recomendamos que tenga en cuenta lo siguiente para ajustar las consultas:

- El número de nodos de Redshift Spectrum que el clúster de Amazon Redshift puede utilizar para una consulta depende del número de sectores del clúster.
- Aumentar el tamaño el clúster puede beneficiar a los perfiles de computación locales, los perfiles de almacenamiento y las capacidades de consulta de la consulta del lago de datos de Amazon S3.
- El planificador de consultas de Amazon Redshift envía predicados y agrupaciones a la capa de consultas de Redshift Spectrum siempre que sea posible.
- Cuando se devuelven grandes cantidades de datos de Amazon S3, el procesamiento se ve limitado por los recursos del clúster.
- Como Redshift Spectrum se escala automáticamente para procesar solicitudes de gran tamaño, su rendimiento general mejora siempre que pueda transferir el procesamiento a la capa de Redshift Spectrum.

Recursos

- [Prácticas recomendadas de Amazon Redshift](#) (documentación de Amazon Redshift)
- [Prácticas recomendadas de Amazon Redshift para el diseño de consultas](#) (documentación de Amazon Redshift)
- [Ajuste del rendimiento de las consultas](#) (documentación de Amazon Redshift)
- [Plan de consultas](#) (documentación de Amazon Redshift)
- [Mejora del rendimiento de las consultas de Amazon Redshift Spectrum](#) (documentación de Amazon Redshift)
- [Descripción del ciclo de vida de las consultas en Amazon Redshift](#) (Recomendaciones de AWS)

Historial de documentos

En la siguiente tabla, se describen cambios significativos de esta guía. Si quiere recibir notificaciones de futuras actualizaciones, puede suscribirse a las [notificaciones RSS](#).

Cambio	Descripción	Fecha
Se eliminó AQUA	Eliminamos la información sobre Advanced Query Accelerator (AQUA).	14 de junio de 2024
Publicación inicial	—	3 de febrero de 2023

AWS Glosario de orientación prescriptiva

Los siguientes son términos de uso común en las estrategias, guías y patrones proporcionados por la Guía AWS prescriptiva. 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:

- **Refactor/re-architect** — Mueva una aplicación y modifique su arquitectura aprovechando al máximo las funciones nativas de 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 su base de datos Oracle local a la PostgreSQL-Compatible edición Amazon Aurora.
- **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: Migrar la base de datos Oracle en las instalaciones a Amazon Relational Database Service (Amazon RDS) para Oracle en la nube de 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 su sistema de gestión de relaciones con los clientes (CRM) a Salesforce.com.
- **Volver a alojar (migrar mediante lift-and-shift)**: traslade una aplicación a la nube sin hacer cambios para aprovechar las funcionalidades de la nube. Ejemplo: Migrar la 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. Los servidores se migran de una plataforma en las instalaciones a un servicio en la nube para la misma plataforma. Ejemplo: migrar una Microsoft Hyper-V aplicación a 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

A2A () Agent-to-Agent

Un protocolo completo para la colaboración entre agentes que facilita la delegación de tareas y la transferencia de estados.

ABAC

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

servicios abstractos

Consulte [servicios administrados](#).

ACID

Consulte [atomicidad, consistencia, aislamiento, 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 una [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 de origen gestiona las transacciones de las aplicaciones conectadas, mientras los datos se replican en la de destino. La base de datos de destino no acepta ninguna transacción durante la migración.

Agente

Un sistema de IA que puede razonar, planificar y tomar medidas de forma autónoma utilizando herramientas para alcanzar los objetivos.

Agent Ops

Prácticas operativas para crear, probar, implementar y ejecutar agentes de IA en producción a escala.

función de agregación

Función SQL que actúa en un grupo de filas y calcula un único valor de devolución para el grupo. Entre los ejemplos de funciones de agregación se incluyen SUM y MAX.

IA

Consulte [inteligencia artificial](#).

AIOps

Consulte [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

Enfoque de seguridad que permite usar de manera exclusiva 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. Para obtener más información, consulte [ABAC AWS en la](#) documentación 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

Una ubicación distinta dentro de una Región de AWS que está aislada de los fallos en otras zonas de disponibilidad y que proporciona una conectividad de red económica y de baja latencia a otras zonas de disponibilidad de la misma región.

AWS Marco de adopción de la nube (AWS CAF)

Un marco de directrices y mejores prácticas AWS para ayudar a las organizaciones a desarrollar un plan eficiente y eficaz para migrar con éxito a la nube. AWS CAF organiza la orientación en seis áreas de enfoque denominadas perspectivas: negocios, personas, gobierno, 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 proporciona orientación para el desarrollo, la formación y la comunicación de las personas a fin de preparar a 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 Marco de calificación de la carga de trabajo (AWS WQF)

Herramienta que evalúa las cargas de trabajo de migración de bases de datos, recomienda estrategias de migración y proporciona 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

bot malicioso

[Bot](#) destinado a causar interrupciones o daños a personas u organizaciones.

BCP

Consulte [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. Consulte también [endianidad](#).

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.

blue/green despliegue

Estrategia de implementación en la que se crean dos entornos separados, pero idénticos. La versión actual de la aplicación se ejecuta en un entorno (azul) y la nueva versión de la aplicación se ejecuta en el otro entorno (verde). Esta estrategia lo ayuda a hacer reversiones rápidas con un impacto mínimo.

bot

Aplicación de software que ejecuta tareas automatizadas a través de Internet y simula la actividad o interacción humana. Algunos bots son útiles o beneficiosos, como los rastreadores web que indexan la información de Internet. Otros bots, conocidos como bots maliciosos, tienen como objetivo causar interrupciones o daños a personas u organizaciones.

botnet

Redes de [bots](#) infectadas por [malware](#) y que están bajo el control de una sola parte, conocida como pastor de bots u operador de bots. Las botnets son el mecanismo más conocido para escalar los bots y su impacto.

branch

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](#) (GitHub documentación).

acceso de emergencia

En circunstancias excepcionales y mediante un proceso aprobado, es una forma rápida de que un usuario pueda acceder a un Cuenta de AWS sitio al que normalmente no tiene permisos de acceso. Para obtener más información, consulte el indicador de [implementación de procedimientos rompe-cristales](#) en la AWS Well-Architected guía.

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 [AWS Cloud Adoption Framework](#).

implementación canario

Lanzamiento lento e incremental de una versión para los usuarios finales. Cuando tenga mayor confianza en la nueva versión, la implementa y reemplaza la versión actual en su totalidad.

CCoE

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

CDC

Consulte [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

Introducción intencionada de 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 [integración continua y 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.

Desarrollador ciudadano

Un usuario empresarial que crea aplicaciones de IA utilizando plataformas sin code/low código sin conocimientos técnicos especializados.

cifrado del cliente

Cifrado de datos localmente, antes de que el objetivo los Servicio de AWS 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 de CCoE](#) en el blog de estrategia Nube de AWS empresarial.

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 relacionada con la tecnología de [computación de periferia](#).

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
- Re-invention — Optimizar los productos y servicios e innovar en la nube

Stephen Orban definió estas etapas en la entrada del blog [The Journey Toward Cloud-First & the Stages of Adoption del](#) blog Nube de AWS Enterprise Strategy. Para obtener información sobre su relación con la estrategia de AWS migración, consulte la [guía de preparación para la migración](#).

CMDB

Consulte [base de datos de administración de 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. Algunos repositorios en la nube comunes son GitHub o Bitbucket Cloud. Cada versión del código se denomina rama. En una estructura de microservicios, cada repositorio se encuentra dedicado a una única funcionalidad. Una sola CI/CD canalización 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 (CV)

Campo de la [IA](#) que utiliza el machine learning para analizar y extraer información de formatos visuales, como imágenes y videos digitales. Por ejemplo, Amazon SageMaker AI proporciona algoritmos de procesamiento de imágenes para CV.

deriva de configuración

En el caso de una carga de trabajo, un cambio en la configuración con respecto al estado esperado. Podría provocar que la carga de trabajo deje de cumplir las normas y, por lo general, es gradual e involuntaria.

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

Un conjunto de AWS Config reglas y medidas correctivas que puede reunir para personalizar sus controles de conformidad y seguridad. Puede implementar un paquete de conformidad como una entidad única en una región Cuenta de AWS y, o en una organización, mediante una plantilla YAML. Para obtener más información, consulta los [paquetes de conformidad](#) en la documentación. AWS Config

integración y entrega continuas (I) CI/CD

El proceso de automatización de las etapas de origen, creación, prueba, puesta en escena y producción del proceso de publicación del software. CI/CD se describe comúnmente como una canalización. CI/CD puede ayudarlo a automatizar los procesos, mejorar la productividad, mejorar la calidad del código y entregar más rápido. 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](#).

CV

Consulte [visión artificial](#).

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 los datos es un componente del pilar de seguridad del AWS Well-Architected Framework. Para obtener más información, consulte [Clasificación de datos](#).

deriva 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 deriva de 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.

mallado de datos

Marco de arquitectura que proporciona una propiedad de datos distribuida y descentralizada con una administración y una gobernanza centralizadas.

minimización de datos

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

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

Sistema de administración de datos que respalda la inteligencia empresarial, como los análisis. Los almacenes de datos suelen contener grandes cantidades de datos históricos y, por lo general, se utilizan para las consultas y los 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 [lenguaje de definición de bases 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.

defensa en profundidad

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. Al adoptar esta estrategia AWS, se añaden varios controles en diferentes capas de la AWS Organizations estructura para ayudar a proteger los recursos. Por ejemplo, un enfoque de defensa en profundidad podría combinar la autenticación multifactor, la segmentación de la red y el cifrado.

administrador delegado

En AWS Organizations, un servicio compatible puede registrar una cuenta de AWS miembro para administrar las cuentas de la organización y gestionar los permisos de 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 en 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 suelen utilizar para restringir consultas, filtrarlas y etiquetar los 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)

Estrategia y proceso que utiliza para minimizar el tiempo de inactividad y la pérdida de datos a causa de un [desastre](#). Para obtener más información, consulte [Recuperación de cargas de trabajo ante desastres en AWS: Recuperación en la nube](#) en el AWS Well-Architected marco.

DML

Consulte [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. Eric Evans introdujo este concepto en su libro *Domain-Driven Design: Tackling Complexity in the Heart of 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 ASP.NET Microsoft \(ASMX\) mediante contenedores y Amazon API Gateway](#).

DR

Consulte [recuperación ante desastres](#).

Detección de desviaciones

Seguimiento de las desviaciones con respecto a una configuración con línea de base. 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 [asignación de flujos de valor para el desarrollo](#).

E

EDA

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

EDI

Consulte [intercambio electrónico 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 la nube](#), la computación de periferia puede reducir la latencia de la comunicación y mejorar el tiempo de respuesta.

intercambio electrónico de datos (EDI)

Intercambio automatizado de documentos comerciales entre organizaciones. Para más información, consulte [¿Qué es el intercambio electrónico de datos?](#)

cifrado

Proceso de computación que transforma datos de texto plano, que son 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. Big-endian los sistemas almacenan primero el byte más significativo. Little-endian los sistemas almacenan primero el byte menos significativo.

punto de conexión

Consulte [punto de conexión de 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 final con AWS PrivateLink entidades principales Cuentas de AWS o AWS Identity and Access Management (de IAM) y conceder permisos a ellas. 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).

planificación de recursos empresariales (ERP)

Sistema que automatiza y administra los procesos empresariales clave (como la contabilidad, [MES](#) y la administración de proyectos) de una empresa.

cifrado de sobre

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

entorno

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 un CI/CD proceso, 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 AWS de seguridad de CAF incluyen la gestión 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](#).

ERP

Consulte [planificación de recursos empresariales](#).

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 hechos

Tabla central de un [esquema en estrella](#). Almacena datos cuantitativos sobre operaciones empresariales. Por lo general, 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.

Fail Fast

Filosofía que utiliza pruebas frecuentes e incrementales para reducir el ciclo de vida del desarrollo. Es una parte fundamental de los enfoques ágiles.

límite de aislamiento de errores

En el Nube de AWS, un límite, como una zona de disponibilidad Región 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 más información, consulte [AWS Fault Isolation Boundaries](#).

rama de característica

Consulte [rama](#).

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.

peticiones con pocos pasos

Proporcionar a un [LLM](#) una pequeña cantidad de ejemplos que demuestren la tarea y el resultado deseado antes de pedirle que lleve a cabo una tarea similar. Esta técnica es una aplicación del aprendizaje contextual, en el que los modelos aprenden a partir de ejemplos (tomas) integrados en las instrucciones. Few-shot Las indicaciones pueden ser eficaces para tareas que requieren un formato, un razonamiento o un conocimiento del dominio específicos. Consulte también [peticiones desde cero](#).

FGAC

Consulte [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 de cambio](#) 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.

FM

Consulte [modelo fundacional](#).

Modelo fundacional (FM)

Gran red neuronal de aprendizaje profundo que se entrenó con conjuntos de datos masivos de datos generalizados y no etiquetados. Los FM pueden hacer una amplia variedad de tareas generales, como comprender el lenguaje, generar texto e imágenes y conversar en lenguaje natural. Para más información, consulte [¿Qué son los modelos fundacionales?](#)

Puerta de enlace FM

Un intermediario centralizado que controla y normaliza el acceso a los modelos básicos. También se conoce como puerta de enlace LLM.

G

IA generativa

Subconjunto de modelos de [IA](#) que se entrenaron con grandes cantidades de datos y que pueden utilizar una simple petición de texto para crear contenido y artefactos nuevos, como imágenes, videos, texto y audio. Para más información, consulte [¿Qué es la IA generativa?](#)

bloqueo geográfico

Consulte [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 [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, mientras que el [flujo de trabajo basado en enlaces troncales](#) es el enfoque moderno preferido.

imagen dorada

Instantánea de un sistema o software que se usa como plantilla para implementar nuevas instancias de ese sistema o software. Por ejemplo, en la fabricación, una imagen dorada se puede utilizar para aprovisionar software en varios dispositivos y ayuda a mejorar la velocidad, la escalabilidad y la productividad de las operaciones de fabricación de dispositivos.

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 CSPM GuardDuty AWS Trusted Advisor, Amazon Inspector y AWS Lambda cheques personalizados.

barandas (AI)

Mecanismos de seguridad que filtran, validan y restringen las entradas y salidas de los [agentes](#) para ayudar a garantizar un comportamiento responsable y seguro de la IA.

H

HA

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.

datos de reserva

Parte de los datos históricos etiquetados que se ocultan de un conjunto de datos que se utiliza para entrenar un modelo de [machine learning](#). Puede utilizar los datos de reserva para evaluar el rendimiento del modelo mediante la comparación de las predicciones del modelo con los datos de reserva.

human-in-the-loop (HiTL)

Un patrón de flujo de trabajo en el que la ejecución de los [agentes](#) se detiene para su revisión y aprobación por parte de una persona en los puntos de decisión críticos.

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, una revisión suele 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

Consulte [infraestructura como código](#).

políticas basadas en identidades

Política asociada a uno o más directores de IAM que define sus permisos en el entorno. 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

Consulte [Internet de las cosas industrial](#).

infraestructura inmutable

Modelo que implementa una nueva infraestructura para las cargas de trabajo de producción en lugar de actualizar o modificar la infraestructura existente o aplicarle revisiones. Las infraestructuras inmutables son de manera intrínseca más coherentes, fiables y predecibles que las [infraestructuras mutables](#). Para obtener más información, consulte las mejores prácticas del [Framework para implementar con una infraestructura inmutable](#). AWS Well-Architected

VPC entrante (de entrada)

En una arquitectura de AWS cuentas múltiples, 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.

Industria 4.0

Un término que [Klaus Schwab](#) introdujo en 2016 para referirse a la modernización de los procesos de fabricación mediante avances en la conectividad, los datos en tiempo real, la automatización, el análisis y. AI/ML

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 AWS cuentas múltiples, una VPC centralizada que gestiona las inspecciones del tráfico de red entre las VPC (iguales o Regiones de AWS diferentes), Internet y las redes locales. 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 obtener más información, consulte Interpretabilidad del modelo [de aprendizaje automático](#) 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 [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 landing zone es un AWS entorno multicuenta bien diseñado, 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](#).

modelo de lenguaje de gran tamaño (LLM)

Modelo de [IA](#) de aprendizaje profundo que se entrenó previamente con una gran cantidad de datos. Un LLM puede llevar a cabo varias tareas, como responder preguntas, resumir documentos, traducir textos a otros idiomas y completar oraciones. Para más información, consulte [¿Qué es un LLM \(modelo de lenguaje de gran tamaño\)?](#)

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

Consulte [Las 7 R](#).

sistema little-endian

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

LLM

Consulte [modelo de lenguaje de gran tamaño](#).

entornos inferiores

Consulte [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

Consulte [rama](#).

malware

Software diseñado para comprometer la seguridad o la privacidad de la computadora. El malware podría interrumpir los sistemas informáticos, filtrar información confidencial u obtener acceso no autorizado. Algunos ejemplos de malware son los virus, los gusanos, el ransomware, los troyanos, el spyware y los registradores de pulsaciones de teclas.

Servicios administrados

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

sistema de ejecución de fabricación (MES)

Sistema de software para seguir, supervisar, documentar y controlar los procesos de producción que convierten las materias primas en productos acabados en la zona de producción.

MAP

Consulte [Programa de aceleración de la migración](#).

MCP

Consulte [Model Context Protocol](#).

Protocolo de contexto para modelos (MCP)

Un protocolo sin estado para la comunicación entre el [agente](#) y la [herramienta](#).

Servidor MCP

Un servicio que expone una o más [herramientas](#) a través del protocolo [Model Context](#).

mecanismo

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

cuenta de miembro

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

MES

Consulte [sistema de ejecución de fabricación](#).

Message Queuing Telemetry Transport (MQTT)

[Un protocolo de comunicación ligero de máquina a máquina \(M2M\), basado en el publish/subscribe patrón, para dispositivos de IoT con recursos limitados.](#)

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 [Integrar](#) microservicios mediante servicios sin servidor. 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)

Un AWS programa que proporciona soporte de consultoría, formación y servicios para ayudar a las organizaciones a crear una base operativa sólida para migrar a la nube y para 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

Cross-functional equipos que agilizan la migración de las cargas de trabajo mediante enfoques ágiles y automatizados. Los equipos de las fábricas de migración suelen estar compuestos por analistas y propietarios de operaciones, 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 la subred de destino, el grupo de seguridad y AWS la cuenta.

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: rehospede la migración a Amazon EC2 AWS con Application Migration Service.

Migration Portfolio Assessment (MPA)

Herramienta en línea que proporciona 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 iniciar sesión) está disponible de forma gratuita para todos los AWS consultores y consultores de los socios de APN.

Evaluación de la preparación para la migración (MRA)

Proceso que consiste en obtener información sobre el estado de preparación de una organización para la nube, identificar sus puntos fuertes y débiles y elaborar un plan de acción para cerrar las brechas identificadas mediante el 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 más información, consulte la entrada [Las 7 R](#) de este glosario y también [Mobilize your organization to accelerate large-scale migrations](#).

ML

Consulte [machine learning](#).

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 más información, consulte [Strategy for modernizing applications in the 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 más información, consulte [Evaluating modernization readiness for applications in the 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](#).

MPA

Consulte [Migration Portfolio Assessment](#).

MQTT

Consulte [Message Queuing Telemetry Transport](#).

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

Modelo que actualiza y modifica la infraestructura actual para las cargas de trabajo de producción. Para mejorar la coherencia, la confiabilidad y la previsibilidad, el AWS Well-Architected Marco recomienda el uso de una [infraestructura inmutable](#) como práctica recomendada.

O

OAC

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

OAI

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

OCM

Consulte [administración 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

Consulte [acuerdo de 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.

OPC-UA

Consulte [Open Process Communications: arquitectura unificada](#).

Comunicaciones de proceso abierto: arquitectura unificada () OPC-UA

Un protocolo de comunicación de máquina a máquina (M2M) para la automatización industrial. OPC-UA proporciona un estándar de interoperabilidad con esquemas de cifrado, autenticación y autorización de datos.

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)

Lista de comprobación de preguntas y prácticas recomendadas asociadas que son útiles para comprender, evaluar, prevenir o reducir el alcance de los incidentes y posibles errores. Para obtener más información, consulte [las revisiones de preparación operativa \(ORR\)](#) en el AWS Well-Architected marco.

tecnología operativa (TO)

Sistemas de hardware y software que funcionan con el entorno físico para controlar las operaciones, los equipos y la infraestructura industriales. En el sector de la fabricación, la integración de los sistemas de TO y tecnología de la información (TI) es un enfoque clave para las transformaciones de la [industria 4.0](#).

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

Un registro creado por y AWS CloudTrail que registra todos los eventos Cuentas de AWS de una organización 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 AWS migración, este marco se denomina aceleración de personal, debido a la velocidad de cambio que requieren 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 admite todos los buckets de S3 Regiones de AWS, el cifrado del lado del servidor con AWS KMS (SSE-KMS) y DELETE las solicitudes PUT y 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.

ORR

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

OT

Consulte [tecnología operativa](#).

VPC saliente (de salida)

En una arquitectura de AWS cuentas múltiples, 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.

PII

Consulte [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.

PLC

Consulte [controlador lógico programable](#).

PLM

Consulte [administración del ciclo de vida del producto](#).

policy

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

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.

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

Condición de consulta que devuelve `true` o `false`. En general, se encuentra en una cláusula `WHERE`.

inserción de predicados

Técnica de optimización de consultas en bases de datos que filtra los datos de la consulta antes de transferirlos. Esta técnica reduce la cantidad de datos de la base de datos relacional que se tienen que recuperar y procesar. Además, 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 AWS que puede realizar acciones y acceder a los recursos. Esta entidad suele ser un usuario raíz para un Cuenta de AWS 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

Enfoque de ingeniería de sistemas que tiene en cuenta la privacidad durante todo el proceso de desarrollo.

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

[Control de seguridad](#) que se diseñó para evitar la implementación de recursos que no cumplan con la normativa. Estos controles analizan los recursos antes de aprovisionarlos. Si el recurso no cumple con los requisitos del control, no se aprovisiona. 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.

administración del ciclo de vida del producto (PLM)

Administración de los datos y los procesos de un producto a lo largo de todo su ciclo de vida, desde el diseño, el desarrollo y el lanzamiento, pasando por el crecimiento y la madurez, hasta la reducción de su uso y su retirada.

entorno de producción

Consulte [entorno](#).

controlador lógico programable (PLC)

En el sector de la fabricación, computadora adaptable y altamente fiable que supervisa las máquinas y automatiza los procesos de fabricación.

encadenamiento de peticiones

Uso de la salida de una petición de [LLM](#) como entrada para la siguiente petición a fin de generar mejores respuestas. Esta técnica se utiliza para dividir una tarea compleja en tareas secundarias o para refinar o ampliar de forma iterativa una respuesta preliminar. Ayuda a mejorar la precisión y la relevancia de las respuestas de un modelo y permite obtener resultados más detallados y personalizados.

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.

publish/subscribe (pub/sub)

Patrón que permite establecer comunicaciones asíncronas entre microservicios para mejorar la escalabilidad y la capacidad de respuesta. Por ejemplo, en un [MES](#) basado en microservicios, un microservicio puede publicar mensajes de eventos en un canal al que se pueden suscribir otros microservicios. El sistema puede agregar nuevos microservicios sin cambiar el servicio de publicación.

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

Consulte [responsable, fiable, consultada e informada \(RACI\)](#).

RAG

Consulte [generación aumentada por recuperación](#).

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

Consulte [responsable, fiable, consultada e informada \(RACI\)](#).

RCAC

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

réplica de lectura

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

Consulte [Las 7 R](#).

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

Consulte [Las 7 R](#).

Region

Conjunto de AWS recursos en un área geográfica. Cada uno Región de AWS está aislado e independiente de los demás para proporcionar tolerancia a las fallas, estabilidad y resiliencia. Para más información, consulte [Specify which Regiones de AWS your account can use](#).

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

Consulte [Las 7 R](#).

versión

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

reubicar

Consulte [Las 7 R](#).

redefinir la plataforma

Consulte [Las 7 R](#).

recomprar

Consulte [Las 7 R](#).

resiliencia

Capacidad de una aplicación para resistir interrupciones o recuperarse de ellas. Al planificar la resiliencia en la Nube de AWS, la [alta disponibilidad](#) y la [recuperación ante desastres](#) son consideraciones comunes. Para más información, consulte [Resiliencia en la Nube de AWS](#).

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

Consulte [Las 7 R](#).

retirar

Consulte [Las 7 R](#).

Generación aumentada de recuperación (RAG)

Tecnología de [IA generativa](#) mediante la que un [LLM](#) hace referencia a un origen de datos autorizado que se encuentra fuera de sus orígenes de datos de entrenamiento antes de generar una respuesta. Por ejemplo, un modelo de RAG podría hacer una búsqueda semántica en la base de conocimientos o en los datos personalizados de una organización. Para más información, consulte [¿Qué es RAG \(generación aumentada por recuperación\)?](#)

rotación

Proceso mediante el que periódicamente se actualiza un [secreto](#) para que resulte más difícil que un atacante pueda acceder 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 [objetivo de punto de recuperación](#).

RTO

Consulte [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 función permite el inicio de sesión único (SSO) federado, de modo que los usuarios pueden iniciar sesión en la Consola de administración de AWS o llamar a las operaciones de la AWS API sin tener que crear un usuario en IAM para todos los miembros 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.

SCADA

Consulte [control de supervisión y adquisición de datos](#).

SCP

Consulte [política de control de servicio](#).

secreta

En AWS Secrets Manager, información confidencial o restringida, como una contraseña o credenciales de usuario, que se almacena de forma cifrada. Se compone del valor del secreto y de sus metadatos. El valor del secreto puede ser binario, una sola cadena o varias cadenas. Para más información, consulte [What's in a Secrets Manager secret?](#) en la documentación de Secrets Manager.

seguridad desde el diseño

Enfoque de ingeniería de sistemas que tiene en cuenta la seguridad durante todo el proceso de desarrollo.

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 de controles de seguridad principales: [preventivos](#), [de detección](#), [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

Acción predefinida y programada que está diseñada para responder automáticamente a un evento de seguridad o corregirlo. Estas automatizaciones sirven como controles de seguridad

[preventivos o adaptables](#) que le ayudan a implementar las mejores prácticas AWS de seguridad. La modificación de un grupo de seguridad de VPC, la aplicación de revisiones a una instancia de Amazon EC2 o la rotación de credenciales son algunos ejemplos de acciones de respuesta automatizadas.

cifrado del servidor

Cifrado de los datos en su destino, por parte de Servicio de AWS quien 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 [las políticas de control de servicios](#) en la AWS Organizations documentación.

punto de enlace de servicio

La URL del punto de entrada de 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)

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

modelo de responsabilidad compartida

Un modelo que describe la responsabilidad con AWS la que compartes la seguridad y el cumplimiento de 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](#).

Shadow AI

Aplicaciones de [IA](#) no autorizadas creadas o utilizadas fuera de los canales regulados dentro de una organización.

SIEM

Consulte [sistema de administración de eventos e información de seguridad](#).

único punto de error (SPOF)

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

SLA

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

SLI

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

SLO

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

modelo de dividir y sembrar

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 más información, consulte [Phased approach to modernizing applications in the Nube de AWS](#).

SPOF

Consulte [único punto de error](#).

esquema en estrella

Estructura organizativa de una base de datos que utiliza una tabla de hechos de gran tamaño para almacenar datos transaccionales o medidos y una o varias tablas dimensionales más pequeñas para almacenar los atributos de los datos. Esta estructura está diseñada para utilizarse 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 desmantelar 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 de cómo aplicar 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.

control de supervisión y adquisición de datos (SCADA)

En el sector de la fabricación, sistema que utiliza hardware y software para supervisar los activos físicos y las operaciones de producción.

cifrado simétrico

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

pruebas sintéticas

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

petición del sistema

Técnica para proporcionar contexto, instrucciones o pautas a un [LLM](#) para dirigir su comportamiento. Las peticiones del sistema ayudan a establecer el contexto y las reglas para las interacciones con los usuarios.

T

etiquetas

Key-value pares que actúan como metadatos para organizar sus AWS recursos. Las etiquetas pueden ayudar a administrar, identificar, organizar, buscar y filtrar recursos de . 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.

herramienta

Una función o API que un [agente](#) puede invocar para realizar operaciones en sistemas externos.

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 pasarela de tránsito](#) en la AWS Transit Gateway documentación.

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

Otorgar permisos a un servicio que especifique para realizar tareas en su organización 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 [AWS Organizations Utilización con otros AWS servicios](#) en la AWS Organizations documentación.

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.

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

Consulte [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 templados

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 hace un cálculo en un grupo de filas que se relacionan de alguna manera con el registro actual. Las funciones de ventana son útiles para las tareas de procesamiento, 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.

WORM

Consulte [escritura única y lectura múltiple](#).

WQF

Consulte [AWS Workload Qualification Framework](#).

escritura única y lectura múltiple (WORM)

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

Z

ataque de día cero

Ataque, normalmente de malware, que se aprovecha de 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.

peticiones desde cero

Proporcionar a un [LLM](#) instrucciones para llevar a cabo una tarea, pero sin ejemplos (pasos) que puedan ayudar a guiarlo. El LLM debe usar los conocimientos del entrenamiento previo para

llevar a cabo la tarea. La eficacia de la petición desde cero depende de la complejidad de la tarea y de la calidad de la petición. Consulte también [peticiones con pocos pasos](#).

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.