

Pilar de sostenibilidad



Pilar de sostenibilidad: AWS Well-Architected Framework

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

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

Resumen e introducción	i
Introducción	1
Sostenibilidad en la nube	3
El modelo de responsabilidad compartida	4
Sostenibilidad de la nube	5
Sostenibilidad en la nube	5
Sostenibilidad a través de la nube	5
Principios de diseño para la sostenibilidad en la nube	6
Proceso de mejora	8
Situación de ejemplo	9
Identifique los objetivos de mejora	9
Recursos	9
Evaluar las mejoras específicas	10
Métricas de proxy	10
Métricas empresariales	11
Indicadores clave de rendimiento	11
Estimar la mejora	12
Evaluar las mejoras	12
Priorizar y planificar las mejoras	13
Probar y validar las mejoras	14
Desplegar los cambios en producción	16
Medir los resultados y repetir los éxitos	16
Sostenibilidad como requisito no funcional	19
Prácticas recomendadas para la sostenibilidad en la nube	21
Selección de regiones	21
SUS01-BP01 Elegir la región basándose tanto en los requisitos empresariales como en los objetivos de sostenibilidad	21
Alineación con la demanda	23
SUS02-BP01 Escalar la infraestructura de la carga de trabajo dinámicamente	24
SUS02-BP02: Alineación de los SLA con los objetivos de sostenibilidad	28
SUS02-BP03: Detener la creación y el mantenimiento de los recursos no utilizados	29
SUS02-BP04 Optimizar la ubicación geográfica de las cargas de trabajo en función de sus requisitos de red	31

SUS02-BP05: Optimización de los recursos de los miembros del equipo para las actividades realizadas	34
SUS02-BP06 Implementar el almacenamiento en búfer o la limitación para aplanar la curva de demanda	35
Software y arquitectura	38
SUS03-BP01: Optimizar el software y la arquitectura para los trabajos asíncronos y programados	39
SUS03-BP02 Eliminar o refactorizar los componentes de cargas de trabajo con uso reducido o nulo	41
SUS03-BP03: Optimización de las áreas de código que consumen la mayor parte del tiempo o de los recursos	43
SUS03-BP04 Optimizar el impacto en los dispositivos y equipos	45
SUS03-BP05: Uso de los patrones de software y las arquitecturas que mejor respaldan los patrones de almacenamiento y el acceso a los datos	47
Administración de datos	50
SUS04-BP01: Implementación de una política de clasificación de datos	50
SUS04-BP02 Usar tecnologías que admiten patrones de almacenamiento y acceso a los datos	52
SUS04-BP03 Usar políticas para administrar el ciclo de vida de los conjuntos de datos	57
SUS04-BP04 Utilice la elasticidad y la automatización para ampliar el almacenamiento de bloques o el sistema de archivos	60
SUS04-BP05: Eliminación de datos innecesarios o redundantes	62
SUS04-BP06 Usar sistemas de archivos o almacenamiento compartidos para acceder a datos comunes	64
SUS04-BP07: Minimización del movimiento de datos entre redes	67
SUS04-BP08: Realización de copias de seguridad de los datos solo cuando sea difícil volver a crearlos	69
Hardware y servicios	71
SUS05-BP01 Usar la mínima cantidad de hardware para cumplir sus necesidades	72
SUS05-BP02: Uso de los tipos de instancia con el menor impacto	74
SUS05-BP03 Usar servicios administrados	78
SUS05-BP04 Optimizar el uso de aceleradores de computación basados en hardware	80
Proceso y cultura	82
SUS06-BP01 Adoptar métodos que permitan introducir mejoras en la sostenibilidad rápidamente	82
SUS06-BP02: Mantenimiento de una carga de trabajo actualizada	84

SUS06-BP03: Incremento de la utilización de los entornos de compilación	86
SUS06-BP04 Usar granjas de dispositivos administrados para pruebas	88
Conclusión	90
Colaboradores	91
Otra documentación	92
Revisiones del documento	93
Avisos	94
AWS Glossary	95

Pilar de sostenibilidad: AWS Well-Architected Framework

Fecha de publicación: 3 de octubre de 2023 ([Revisiones del documento](#))

Este documento técnico se centra en el pilar de sostenibilidad de Amazon Web Services (AWS) Well-Architected Framework. En él se incluyen principios de diseño, orientación operativa, prácticas recomendadas, posibles ventajas y desventajas, y planes de mejora que puede usar para lograr los objetivos de sostenibilidad de las cargas de trabajo de AWS.

Introducción

AWS Well-Architected Framework le ayuda a comprender las ventajas y desventajas de las decisiones que toma al crear cargas de trabajo en AWS. Mediante el uso del marco, podrá conocer las prácticas recomendadas de arquitectura para diseñar y operar cargas de trabajo en la Nube de AWS que sean seguras, fiables, eficaces, rentables y sostenibles. El marco ofrece una forma de medir sus arquitecturas de forma constante en función de las prácticas recomendadas de arquitectura y de identificar áreas de mejora. Si cuenta con cargas de trabajo con una buena arquitectura, aumentará la capacidad de respaldar los resultados empresariales.

El marco se basa en seis pilares:

- Excelencia operativa
- Seguridad
- Fiabilidad
- Eficiencia del rendimiento
- Optimización de costes
- Sostenibilidad

El presente documento se centra en el pilar de sostenibilidad y, dentro del ámbito de sostenibilidad, en la sostenibilidad medioambiental. Está destinado a aquellos que ocupan puestos en tecnología, como los directores de tecnología (CTO), arquitectos, desarrolladores y miembros del equipo de operaciones.

Después de leer este documento, comprenderá mejor las recomendaciones y estrategias actuales de AWS que puede utilizar cuando diseñe arquitecturas en la nube teniendo en cuenta la sostenibilidad.

Al adoptar las prácticas que se indican en este documento, podrá crear arquitecturas que aumentarán al máximo la eficiencia y reducirán el desperdicio.

Sostenibilidad en la nube

La sostenibilidad como disciplina aborda el impacto medioambiental, económico y social a largo plazo de sus actividades empresariales. La [Comisión Mundial sobre el Medio Ambiente y el Desarrollo de las Naciones Unidas](#) define el desarrollo sostenible como «aquel que permite satisfacer las necesidades del presente sin comprometer la habilidad de las futuras generaciones de satisfacer sus necesidades propias». Su organización o negocio puede tener repercusiones negativas en el medioambiente, como emisiones de carbono directas o indirectas, residuos no reciclables y daños a recursos compartidos, como el agua no contaminada.

Cuando se crean cargas de trabajo en la nube, la sostenibilidad consiste en comprender las repercusiones de los servicios que se usan, cuantificar el impacto durante todo el ciclo de vida de la carga de trabajo y aplicar tanto principios de diseño como prácticas recomendadas para reducir estas repercusiones. Este documento se centra en los impactos medioambientales, sobre todo en la eficiencia y el consumo energéticos, ya que son impulsores importantes que ayudan a los arquitectos a promover la adopción de medidas directas destinadas a reducir el uso de los recursos.

Al centrarse en los impactos medioambientales, debe entender cómo se contabilizan normalmente estos impactos y las repercusiones posteriores en la contabilidad de las emisiones de su organización. La [Protocolo para gases de efecto invernadero](#) organiza las emisiones de carbono en los siguientes ámbitos, junto con ejemplos de emisiones relevantes dentro de cada alcance para un proveedor de servicios en la nube como AWS:

- Alcance 1: todas las emisiones directas de las actividades de una organización o bajo su control. Por ejemplo, la quema de combustible por parte de los generadores de reserva de los centros de datos.
- Alcance 2: emisiones indirectas procedentes de la electricidad comprada y utilizada para alimentar los centros de datos y otras instalaciones. Por ejemplo, las emisiones de la generación de energía comercial.
- Alcance 3: todas las demás emisiones indirectas de las actividades de una organización procedentes de orígenes que no controla. Los ejemplos de AWS incluyen las emisiones relacionadas con la construcción de centros de datos y la fabricación y el transporte del hardware informático desplegado en los centros de datos.

Desde el punto de vista del cliente de AWS, las emisiones de sus cargas de trabajo que se ejecutan en AWS se contabilizan como emisiones indirectas y forman parte de sus emisiones de alcance

3. Cada carga de trabajo desplegada genera una fracción de las emisiones totales de AWS de cada uno de los alcances anteriores. La cantidad real varía según la carga de trabajo y depende de varios factores, como los servicios de AWS utilizados, la energía consumida por esos servicios, la intensidad de carbono de las redes eléctricas que sirven a los centros de datos de AWS donde se ejecutan y la adquisición de energía renovable por parte de AWS.

En este documento se describe, en primer lugar, un modelo de responsabilidad compartida para la sostenibilidad medioambiental y, después, ofrece las prácticas recomendadas de arquitectura para que pueda minimizar el impacto de sus cargas de trabajo mediante la reducción del total de recursos necesarios para su funcionamiento en los centros de datos de AWS.

El modelo de responsabilidad compartida

La sostenibilidad medioambiental es una responsabilidad compartida entre clientes y AWS.

- AWS es responsable de optimizar la sostenibilidad de la nube: proporcionar infraestructuras eficientes y compartidas, administrar el agua y abastecerse de energía renovable.
- Los clientes son responsables de la sostenibilidad in la nube: optimizar las cargas de trabajo y la utilización de los recursos y minimizar el total de recursos necesarios para desplegar sus cargas de trabajo.



Modelo de responsabilidad compartida

Sostenibilidad de la nube

Los proveedores de servicios en la nube tienen una menor huella de carbono y son más eficientes desde el punto de vista energético que las típicas alternativas locales porque invierten en tecnología eficiente de alimentación y refrigeración, operan con poblaciones de servidores eficientes desde el punto de vista energético y logran altas tasas de utilización de los servidores. Las cargas de trabajo en la nube reducen el impacto al aprovechar los recursos compartidos, como las redes, la energía, la refrigeración y las instalaciones físicas. Puede migrar sus cargas de trabajo en la nube a tecnologías más eficientes a medida que estén disponibles y utilizar los servicios basados en la nube para transformar sus cargas de trabajo y conseguir una mayor sostenibilidad.

Recursos

- [The Carbon Reduction Opportunity of Moving to Amazon Web Services \(La oportunidad de reducir las emisiones de carbono al migrar a Amazon Web Services\)](#)
- [AWS habilita soluciones de sostenibilidad](#)

Sostenibilidad en la nube

La sostenibilidad en la nube es un esfuerzo continuo centrado principalmente en la reducción de la energía y la eficiencia en todos los componentes de una carga de trabajo, con lo que se logra el máximo beneficio de los recursos provisionados y se minimizan los recursos totales necesarios. Este esfuerzo puede abarcar desde la selección inicial de un lenguaje de programación eficiente, la adopción de algoritmos modernos, el uso de técnicas eficientes de almacenamiento de datos, el despliegue de una infraestructura de computación eficiente y de tamaño correcto, y la minimización de los requisitos de un hardware de alta potencia para el usuario final.

Sostenibilidad a través de la nube

Además de minimizar el impacto de las cargas de trabajo que ha desplegado, puede utilizar la Nube de AWS para ejecutar cargas de trabajo diseñadas para respaldar sus desafíos de sostenibilidad más amplios. Algunos ejemplos de estos desafíos son la reducción de las emisiones de carbono, la disminución del consumo de energía, el reciclaje del agua o la reducción de los residuos en otras áreas de su empresa u organización.

Sostenibilidad a través de la nube es cuando se utiliza la tecnología de AWS para resolver un desafío de sostenibilidad más amplio. Por ejemplo, puede utilizar un servicio de machine learning como

[Amazon Monitron](#) para detectar comportamientos anómalos en la maquinaria industrial. Con estos datos de detección, puede llevar a cabo un mantenimiento preventivo para reducir el riesgo de incidentes medioambientales provocados por errores inesperados del equipo y garantizar que la maquinaria siga funcionando con la máxima eficacia.

Principios de diseño para la sostenibilidad en la nube

Aplique estos principios de diseño al diseñar sus cargas de trabajo en la nube para maximizar la sostenibilidad y minimizar el impacto.

- **Analice su impacto:** Mida la repercusión de su carga de trabajo en la nube y modele el impacto futuro de dicha carga. Incluya todas las fuentes de impacto, incluidas las repercusiones resultantes del uso de sus productos por parte de los clientes y de su eventual cierre y retirada. Revise los recursos y las emisiones que se requieren por unidad de trabajo para comparar el rendimiento productivo con el impacto total de sus cargas de trabajo en la nube. Use estos datos para establecer indicadores clave de rendimiento (KPI), evaluar formas de mejorar la productividad a la vez que se reduce el impacto y calcular la repercusión de los cambios propuestos a lo largo del tiempo.
- **Establezca objetivos de sostenibilidad:** Establezca objetivos de sostenibilidad a largo plazo, como la reducción de los recursos de computación y almacenamiento requeridos por transacción, para cada una de las cargas de trabajo en la nube. Modele el rendimiento de la inversión de las mejoras de sostenibilidad para las cargas de trabajo existentes y proporcione a los propietarios los recursos que necesitan para invertir en objetivos de sostenibilidad. Planifique el crecimiento y diseñe la arquitectura de sus cargas de trabajo para que ese crecimiento se traduzca en una reducción de la intensidad del impacto y elija una unidad de medida adecuada, por ejemplo, por usuario o por transacción. Los objetivos en general ayudan a respaldar el cumplimiento de los objetivos de sostenibilidad más amplios de su organización o negocio, a identificar las regresiones y a priorizar las áreas susceptibles de mejora.
- **Maximice el uso:** Aplique el tamaño adecuado a sus cargas de trabajo e implemente un diseño eficaz para garantizar una alta utilización y maximizar la eficiencia energética del hardware subyacente. Dos hosts que se ejecutan al 30 % son menos eficientes que uno solo que se ejecute al 60 %, debido al consumo energético base por host. Al mismo tiempo, elimine o minimice los recursos inactivos, el procesamiento y el almacenamiento para reducir la energía total necesaria para ejecutar su carga de trabajo.
- **Anticípese y adopte nuevas ofertas de hardware y software más eficaces:** Admita las mejoras que hagan sus socios y proveedores como ayuda para reducir el impacto de sus cargas de trabajo en

la nube. Supervise y evalúe de forma continua las nuevas ofertas de hardware y software más eficaces. Aporte flexibilidad al diseño para permitir una adopción rápida de tecnologías nuevas y eficaces.

- **Use servicios administrados:** El uso compartido de servicios en una amplia base de clientes ayuda a maximizar la utilización de los recursos, lo cual reduce la cantidad de infraestructura necesaria para admitir las cargas de trabajo en la nube. Por ejemplo, los clientes pueden compartir el impacto de los componentes de un centro de datos común, como la potencia y las redes, mediante la migración de las cargas de trabajo a la Nube de AWS y la adopción de servicios administrados, como AWS Fargate para contenedores sin servidor, en los que AWS opera a escala y es responsable de su funcionamiento eficiente. Use servicios administrados que le ayuden a minimizar su impacto, como pasar datos a los que no se accede con frecuencia a almacenamiento en frío de forma automática con configuraciones del ciclo de vida de Amazon S3 o Amazon EC2 Auto Scaling, a fin de ajustar la capacidad para satisfacer la demanda.
- **Reduzca el impacto ulterior de sus cargas de trabajo en la nube:** Reduzca la cantidad de energía o los recursos necesarios para usar sus servicios. Reduzca o elimine la necesidad de que los clientes tengan que actualizar sus dispositivos para usar sus servicios. Pruebe a usar granjas de dispositivos para comprender el impacto esperado y realice pruebas con los clientes para que entiendan el impacto real del uso de sus servicios.

Proceso de mejora

El proceso de mejora de la arquitectura incluye entender qué tiene y qué puede hacer para mejorar, para lo que debe seleccionar objetivos de mejora, probar las mejoras, adoptar las mejoras de éxito, cuantificar el éxito y compartir lo aprendido para poder replicarlo en otras partes y así, después, repetir el ciclo.

Los objetivos de sus mejoras pueden ser los siguientes:

- Eliminar los residuos, la baja utilización y los recursos inactivos o no utilizados
- Maximizar el valor de los recursos que consume

Note

Utilice todos los recursos que aprovisiona y realice el mismo trabajo con los mínimos recursos posibles.

En las etapas tempranas de la optimización, elimine primero las áreas desaprovechadas o con una baja utilización y, después, pase a optimizaciones más específicas que se ajusten a su carga de trabajo concreta.

Supervise los cambios en el consumo de los recursos a lo largo del tiempo. Identifique si la acumulación de cambios da como resultado un aumento ineficaz o significativo en el consumo de los recursos. Determine la necesidad de mejoras para abordar los cambios en el consumo e implementar las mejoras que establezca como prioritarias.

Los siguientes pasos se han diseñado como un proceso iterativo que evalúa, prioriza, prueba y despliega mejoras centradas en la sostenibilidad para las cargas de trabajo en la nube.

1. Identifique los objetivos de mejora: revise sus cargas de trabajo comparándolas con las prácticas recomendadas para la sostenibilidad que se indican en este documento e identifique los objetivos de mejora.
2. Evalúe las mejoras específicas: evalúe los cambios específicos para la mejora potencial, el coste previsto y el riesgo empresarial.
3. Priorice y planifique las mejoras: priorice los cambios que ofrezcan las mayores mejoras con el menor coste y riesgo, y establezca un plan de pruebas e implementación.

4. Pruebe y valide las mejoras: implemente los cambios en los entornos de prueba para validar su potencial de mejora.
5. Despliegue los cambios en producción: implemente los cambios en los entornos de producción.
6. Mida los resultados y repita los éxitos: busque oportunidades para replicar los éxitos en las cargas de trabajo y revierta los cambios con resultados inaceptables.

Situación de ejemplo

Más adelante en este documento se hace referencia a la siguiente situación de ejemplo para ilustrar cada paso del proceso de mejora.

Su empresa tiene una carga de trabajo que realiza complejas manipulaciones de imágenes en instancias Amazon EC2 y almacena los archivos modificados y originales para el acceso de los usuarios. Las actividades de procesamiento requieren un uso intensivo de la CPU y los archivos de salida son extremadamente grandes.

Identifique los objetivos de mejora

Conozca las prácticas recomendadas que pueden ayudarle a alcanzar sus objetivos de sostenibilidad. Puede encontrar descripciones detalladas de estas [prácticas recomendadas](#) y recomendaciones de mejora más adelante en este documento.

Revise sus cargas de trabajo y los recursos utilizados. Identificación de los puntos críticos, como los despliegues grandes y los recursos de uso frecuente. Evalúe estos puntos críticos para buscar oportunidades de mejorar la utilización eficaz de sus recursos y reducir el total de recursos necesarios para conseguir los resultados empresariales.

Revise su carga de trabajo según las prácticas recomendadas e identifique los candidatos de mejora.

Con la aplicación de este paso a [Situación de ejemplo](#), podrá identificar las siguientes prácticas recomendadas como objetivos probables de mejora:

- Uso de la mínima cantidad de hardware para cumplir sus necesidades
- Uso de tecnologías que admiten mejor sus patrones de almacenamiento y acceso a los datos

Recursos

- [Optimización de la infraestructura de AWS para la sostenibilidad, Parte I: computación](#)

- [Optimización de la infraestructura de AWS para la sostenibilidad, Parte II: almacenamiento](#)
- [Optimización de la infraestructura de AWS para la sostenibilidad, Parte III: redes](#)

Evaluar las mejoras específicas

Entienda los recursos provisionados por su carga de trabajo para completar una unidad de trabajo. Evalúe las posibles mejoras y calcule su impacto potencial, el coste de implementarlas y los riesgos asociados.

Para medir las mejoras a lo largo del tiempo, entienda primero qué ha provisionado en AWS y cómo se están consumiendo esos recursos.

Comience con un resumen de su uso de AWS y utilice los informes de coste y uso de AWS como ayuda para identificar los puntos críticos. Use este [código de ejemplo de AWS](#) para revisar y analizar su informe con la ayuda de Amazon Athena.

Métricas de proxy

Al evaluar los cambios específicos, también debe evaluar qué métricas cuantifican mejor el efecto de ese cambio en el recurso asociado. Estas métricas se llaman métricas de proxy. Seleccione las métricas de proxy que mejor reflejen el tipo de mejora que está evaluando y los recursos a los que se dirige la mejora. Estas métricas podrían evolucionar con el tiempo.

Los recursos provisionados para respaldar su carga de trabajo incluyen recursos de computación, almacenamiento y redes. Evalúe los recursos provisionados mediante sus métricas de proxy para ver cómo se consumen esos recursos.

Utilice sus métricas de proxy para medir los recursos provisionados a fin de alcanzar los resultados empresariales.

Recurso	Métricas de proxy de ejemplo	Objetivos de mejora
Computación	Minutos de vCPU	Maximizar la utilización de los recursos provisionados
Storage (Almacenamiento)	GB provisionados	Reducir el total provisionado
Red	GB transferidos o paquetes transferidos	Reducir el total transferido y la distancia transferida

Métricas empresariales

Seleccione métricas empresariales para cuantificar la consecución de los resultados empresariales. Sus métricas empresariales deben reflejar el valor proporcionado por su carga de trabajo, por ejemplo, el número de usuarios activos simultáneos, las llamadas a la API servidas o el número de transacciones realizadas. Estas métricas pueden evolucionar con el tiempo. Extreme las precauciones al evaluar las métricas empresariales basadas en las finanzas, ya que la incoherencia en el valor de las transacciones invalida las comparaciones.

Indicadores clave de rendimiento

Mediante la siguiente fórmula, divida los recursos aprovisionados entre los resultados empresariales alcanzados para determinar los recursos aprovisionados por unidad de trabajo.

$$\text{Recursos aprovisionados por unidad de trabajo} = \frac{\text{Métrica de proxy para recurso aprovisionado}}{\text{Métrica empresarial para resultados}}$$

Fórmula de KPI

Utilice sus recursos por unidad de trabajo como sus KPI. Establezca bases de referencia basadas en los recursos aprovisionados como base para las comparaciones.

Recurso	KPI de ejemplo	Objetivos de mejora
Computación	Minutos de vCPU por transacción	Maximizar la utilización de los recursos aprovisionados
Storage (Almacenamiento)	GB por transacción	Reducir el total aprovisionado
Red	GB transferidos por transacción o paquetes transferidos por transacción	Reducir el total transferido y la distancia transferida

Estimar la mejora

Estime la mejora tanto como la reducción cuantitativa de los recursos aprovisionados (según indican sus métricas de proxy) como el cambio porcentual con respecto a su base de referencia de recursos aprovisionados por unidad de trabajo.

Recurso	KPI de ejemplo	Objetivos de mejora
Computación	% de reducción de minutos de vCPU por transacción	Maximice el uso
Storage (Almacenamiento)	% de reducción de GB por transacción	Reducir el total aprovisionado
Red	% de reducción de GB transferidos por transacción o paquetes transferidos por transacción	Reducir el total transferido y la distancia transferida

Evaluar las mejoras

Evaluar las posibles mejoras en relación con el beneficio neto previsto. Evalúe el tiempo, el coste y el nivel de esfuerzo para implementar y mantener, así como los riesgos empresariales, como los impactos imprevistos.

Las mejoras dirigidas suelen representar compensaciones entre los tipos de recursos consumidos. Por ejemplo, para reducir el consumo de computación, puede almacenar un resultado, o, para limitar los datos transferidos, puede procesar los datos antes de enviar el resultado a un cliente. Estos [compensaciones](#) se explican con más detalle más adelante.

Incluya los requisitos no funcionales a la hora de evaluar los riesgos para su carga de trabajo, incluida la seguridad, la fiabilidad, la eficiencia del rendimiento, la optimización de costes y el impacto de las mejoras en su capacidad para utilizar su carga de trabajo.

Con la aplicación de este paso a [Situación de ejemplo](#), evalúa las mejoras objetivo con los siguientes resultados:

Práctica recomendada	Mejora objetivo	Potencial	actual	Riesgo
Uso de la mínima cantidad de hardware para cumplir sus necesidades	Implementación del escalado predictivo para reducir los periodos de baja utilización	Mediana	Bajo	Bajo
Uso de tecnologías que admiten mejor sus patrones de almacenamiento y acceso a los datos	Implementación de mecanismos de compresión más eficaces para reducir el almacenamiento total y el tiempo para conseguirlo	Alto	Bajo	Bajo

La implementación de la programación predictiva reduce las horas de vCPU consumidas por las instancias infrautilizadas o no utilizadas y proporciona beneficios moderados sobre los mecanismos de escalado existentes, con una reducción estimada del 11 % de los recursos consumidos. Los costes implicados son bajos e incluyen la configuración de los recursos de la nube y el funcionamiento del escalado predictivo para Amazon EC2 Auto Scaling. El riesgo es la restricción del rendimiento cuando el escalado se realiza de forma reactiva en respuesta a una demanda que supera las previsiones.

La implementación de una compresión más efectiva tendrá un impacto significativo con grandes reducciones en el tamaño de los archivos en todas sus imágenes originales y manipuladas, con una reducción estimada del 25 % en los requisitos de almacenamiento en la producción. La implementación del nuevo algoritmo es una sustitución de bajo esfuerzo con poco riesgo.

Priorizar y planificar las mejoras

Priorice las mejoras identificadas basándose en el mayor impacto previsto con los costes más bajos y el riesgo aceptable.

Decida en qué mejoras se va a centrar inicialmente e inclúyalas en su plan de desarrollo y planificación de recursos.

Con la aplicación de este paso a [Situación de ejemplo](#), prioriza las mejoras objetivo de la siguiente manera:

Prioridad	de los KPI	Potencial	actual	Riesgo
1	Implementación de mecanismos de compresión más eficaces	Alto	Bajo	Bajo
2	Implementación del escalado predictivo	Mediana	Bajo	Bajo

El alto potencial, el bajo coste y el riesgo de la actualización de la compresión de archivos lo convierten en un objetivo de gran valor para su empresa y en una prioridad sobre la implementación del escalado predictivo. Determina que implementar el escalado predictivo, con su impacto potencial medio, su bajo coste y su bajo riesgo, debería ser la mejora prioritaria una vez completada la compresión de archivos.

Asigna a un miembro del equipo para que implemente una mejor compresión de archivos y agrega el escalado predictivo a sus tareas pendientes.

Probar y validar las mejoras

Realice pequeñas pruebas con una inversión mínima para reducir el riesgo de un esfuerzo a gran escala.

Implemente una copia representativa de su carga de trabajo en su entorno de pruebas para limitar el coste y el riesgo de realizar las pruebas y la validación. Realice un conjunto predefinido de transacciones de prueba, mida los recursos provisionados y determine los recursos que se usan por unidad de trabajo para establecer una base de referencia de las pruebas.

Implemente su objetivo de mejora en el entorno de pruebas y repita la prueba con la misma metodología y en las mismas condiciones. Después, mida los recursos aprovisionados y los recursos utilizados por unidad de trabajo con su mejora aplicada.

Calcule el cambio porcentual desde su base de referencia de los recursos aprovisionados por unidad de trabajo y determine la reducción cuantitativa prevista de los recursos aprovisionados en su entorno de producción. Compare estos valores con los previstos. Determine si el resultado es un nivel de mejora aceptable. Evalúe si las compensaciones de los recursos adicionales consumidos convierten en inaceptable el beneficio neto de la mejora.

Determine si la mejora es un éxito y si deben invertirse recursos para implementar el cambio en la producción. Si el cambio se evalúa como infructuoso en este momento, reoriente sus recursos para probar y validar su próximo objetivo y continúe el ciclo de mejora.

% de reducción de los recursos aprovisionados por unidad de trabajo	Reducción cuantitativa de los recursos aprovisionados	Acción
Ha cumplido las expectativas	Ha cumplido las expectativas	Continuar con la mejora
No ha cumplido las expectativas	Ha cumplido las expectativas	Continuar con la mejora
Ha cumplido las expectativas	No ha cumplido las expectativas	Buscar alternativas de mejora
No ha cumplido las expectativas	No ha cumplido las expectativas	Buscar alternativas de mejora

Con la aplicación de este paso a [Situación de ejemplo](#), realiza pruebas para validar el éxito.

Después de realizar las pruebas con el algoritmo de compresión mejorado, el porcentaje de reducción de los recursos aprovisionados por unidad de trabajo (el almacenamiento necesario tanto para la imagen original como para la imagen modificada) ha cumplido las expectativas con una reducción media del 30 % en el almacenamiento aprovisionado y un aumento insignificante de la carga de computación.

Determina que los recursos de computación adicionales necesarios para aplicar el algoritmo de compresión mejorado a los archivos existentes en producción son insignificantes en comparación

con la reducción de almacenamiento conseguida. Ha confirmado el éxito con la reducción cuantitativa de los recursos necesarios (TB de almacenamiento) y la mejora está aprobada para desplegarla en producción.

Desplegar los cambios en producción

Implemente las mejoras probadas, validadas y aprobadas en producción. Implemente mediante despliegues limitados, confirme la funcionalidad de su carga de trabajo, pruebe la reducción real de los recursos provisionados y de los recursos consumidos por unidad de trabajo en el despliegue limitado y compruebe las consecuencias no deseadas del cambio. Continúe con los despliegues completos después de haberlos probado correctamente.

Revierta los cambios si no se superan las pruebas o si encuentra consecuencias no deseadas inaceptables de su cambio.

Con la aplicación de este paso a [Situación de ejemplo](#), realiza las acciones siguientes.

Implementa los cambios en producción mediante un despliegue limitado a través de una metodología de despliegue azul-verde. Las pruebas de funcionalidad con respecto a las instancias recién desplegadas son correctas. Obtiene una reducción media del 26 % en el almacenamiento provisionado para los archivos de imagen originales y manipulados. No se aprecia un aumento de la carga de computación al comprimir nuevos archivos.

Observa una disminución imprevista del tiempo transcurrido para comprimir archivos de imagen y lo atribuye al código muy optimizado del nuevo algoritmo de compresión.

Continúa con el despliegue completo de la nueva versión.

Medir los resultados y repetir los éxitos

Mida los resultados y repita los éxitos de las siguientes maneras:

- Mida la mejora inicial de los recursos provisionados por unidad de trabajo y la disminución cuantitativa de los recursos provisionados.
- Compare las estimaciones iniciales y los resultados de las pruebas con sus mediciones de producción. Identifique los factores que puedan haber contribuido a las diferencias y actualice sus metodologías de estimación y comprobación cuando resulte apropiado.
- Determine el éxito, y su grado, y comparta los resultados con las partes interesadas.

- Si ha tenido que revertir los cambios debido a pruebas erróneas o a consecuencias negativas no deseadas del cambio, identifique los factores que han contribuido a ello. Itere cuando sea viable o evalúe nuevos enfoques para lograr los objetivos del cambio.
- Tome lo que ha aprendido, establezca estándares y aplique las mejoras correctas a otros sistemas que puedan beneficiarse de forma similar. Capture y comparta su metodología, los artefactos relacionados y los beneficios netos, entre equipos y organizaciones para que otros puedan adoptar su estándar y repetir su éxito.
- Supervise los recursos aprovisionados por unidad de trabajo y haga un seguimiento de los cambios y del impacto total a lo largo del tiempo. Los cambios en la carga de trabajo, o en la forma en que sus clientes consumen su carga de trabajo, pueden tener impacto en la eficacia de la mejora. Vuelva a evaluar las oportunidades de mejora si observa una disminución significativa de la eficacia a corto plazo o una reducción acumulada de la eficacia a lo largo del tiempo.
- Cuantifique el beneficio neto de su mejora a lo largo del tiempo (incluidos los beneficios obtenidos por otros equipos que han aplicado su mejora, si están disponibles) para mostrar el retorno de la inversión de sus actividades de mejora.

Con la aplicación de este paso a [Situación de ejemplo](#), mide los resultados siguientes.

Su carga de trabajo presenta una mejora inicial del 23 % de reducción de los requisitos de almacenamiento tras desplegar y aplicar el nuevo algoritmo de compresión a los archivos de imagen existentes.

El valor medido coincide en gran medida con las estimaciones iniciales (25 %) y se determina que la diferencia significativa con respecto a las pruebas (30 %) es el resultado de que los archivos de imagen utilizados en las pruebas no son representativos de los archivos de imagen presentes en producción. Modifica el conjunto de imágenes de prueba para reflejar más adecuadamente las imágenes en producción.

La mejora se considera un éxito total. La reducción total del almacenamiento aprovisionado es un 2 % menos que el 25 % estimado, pero el 23 % sigue siendo una gran mejora en el impacto de la sostenibilidad y va acompañada de un ahorro de costes equivalente.

Las únicas consecuencias no deseadas del cambio son la beneficiosa reducción del tiempo transcurrido para realizar la compresión y una reducción equivalente de vCPU consumida. Estas mejoras se atribuyen a la alta optimización del código.

Establece un proyecto interno de código abierto en el que comparte su código, los artefactos asociados, la orientación sobre cómo implementar el cambio y los resultados de su implementación.

El proyecto interno de código abierto facilita que sus equipos adopten el código para todos sus casos de uso de almacenamiento de archivos persistentes. Sus equipos adoptan la mejora como un estándar. Los beneficios secundarios del proyecto interno de código abierto son que quienes adoptan la solución se benefician de las mejoras de la misma y que cualquiera puede contribuir a mejorar el proyecto.

Publica su éxito y comparte el proyecto de código abierto en toda la organización. Cada equipo que adopta la solución repite el beneficio con una inversión mínima y lo agrega al beneficio neto recibido de su inversión. Publica estos datos como una historia de éxito continuada.

Seguirá supervisando el impacto de la mejora a lo largo del tiempo y realizará los cambios necesarios en el proyecto interno de código abierto.

Sostenibilidad como requisito no funcional

Incluir la sostenibilidad en su lista de requisitos empresariales puede tener como consecuencia soluciones más rentables. Centrarse en obtener más valor de los recursos que utiliza (y utilizar menos de estos recursos) se traduce directamente en un ahorro económico en AWS, ya que solo tendrá que pagar por lo que utilice.

Al cumplir con los objetivos de sostenibilidad, es posible que no se reflejen compensaciones equivalentes en una o varias de las otras métricas tradicionales, como el tiempo de actividad, la disponibilidad o el tiempo de respuesta. Puede lograr excelentes resultados en la sostenibilidad sin que se vean afectados los niveles de servicio. Cuando sean necesarias compensaciones menores, las mejoras en sostenibilidad logradas por ellas pueden tener más importancia que la pérdida en la calidad de servicio.

Motive a los miembros del equipo para que sigan experimentando con mejoras en la sostenibilidad a medida que van desarrollando requisitos funcionales. Los equipos también deben insertar métricas de proxy al establecer objetivos para garantizar que evalúen la intensidad de los recursos al desarrollar las cargas de trabajo.

A continuación, se muestran ejemplos de compensaciones que pueden reducir los recursos de la nube que utiliza:

Ajustar la calidad del resultado: puede compensar una mejora en la calidad de resultados (QoR) con una reducción en la intensidad de la carga de trabajo a través de computación aproximada. La práctica de computación aproximada busca oportunidades para sacar partido de la carencia entre lo que los clientes necesitan y lo que su empresa produce. Por ejemplo, si almacena los datos en una estructura de datos establecida, puede colocar el operador `ORDER BY` en SQL para eliminar cualquier procesamiento innecesario, guardando los recursos mientras se sigue ofreciendo una respuesta aceptable.

Ajustar la respuesta rápida: un tiempo de respuesta más lento puede reducir la huella de carbono, ya que se reduce al mínimo el uso de recursos compartidos. El procesamiento de tareas ad hoc y efímeras puede provocar gastos mayores durante la puesta en marcha de una empresa. Agrupe y procese tareas en lotes en lugar de tener que pagar más cada vez que llegue una tarea. El procesamiento por lotes sacrifica el tiempo de respuesta para obtener una reducción en el uso de recursos compartidos que tienen que inicializar una instancia, descargar el código fuente y ejecutar el proceso.

Ajustar la disponibilidad: con AWS, puede añadir redundancia y lograr los objetivos de alta disponibilidad con tan solo hacer unos cuantos clics. Puede aumentar la redundancia a través de técnicas como, por ejemplo, la estabilidad estática mediante el aprovisionamiento de recursos inactivos, lo que siempre se traduce en un uso menor. Evalúe las necesidades del negocio al establecer objetivos. Las compensaciones relativamente menores con respecto a la disponibilidad pueden tener como consecuencia una gran mejora en la utilización. Por ejemplo, el patrón de arquitectura de estabilidad estática implica el aprovisionamiento de capacidad de conmutación por error inactiva para asumir inmediatamente la carga después de producirse un error en algún componente. Al establecer un requisito de disponibilidad menos estricto, ya no será necesario contar con una capacidad en línea inactiva y se contará con tiempo para automatizar y desplegar recursos de sustitución. Al incorporar la capacidad por error bajo demanda, se impulsará una mayor utilización en general sin que el negocio se vea afectado durante las operaciones normales, además de lograr una reducción de los costes.

Prácticas recomendadas para la sostenibilidad en la nube

Optimice la ubicación de la carga de trabajo y la arquitectura en función de la demanda, el software, los datos, el hardware y los procesos para aumentar la eficiencia energética. En cada una de estas áreas, pueden aplicarse prácticas recomendadas para reducir el impacto en la sostenibilidad de su carga de trabajo en la nube; para ello, se maximiza su utilización y se minimiza el desaprovechamiento y el total de recursos que se implementan y se conectan para admitir la carga de trabajo.

Temas

- [Selección de regiones](#)
- [Alineación con la demanda](#)
- [Software y arquitectura](#)
- [Administración de datos](#)
- [Hardware y servicios](#)
- [Proceso y cultura](#)

Selección de regiones

La elección de la región para su carga de trabajo afecta significativamente a sus KPI, incluidos el rendimiento, el coste y la huella de carbono. Para mejorar eficazmente estos KPI, debe elegir las regiones para sus cargas de trabajo basándose tanto en los requisitos empresariales como en los objetivos de sostenibilidad.

Prácticas recomendadas

- [SUS01-BP01 Elegir la región basándose tanto en los requisitos empresariales como en los objetivos de sostenibilidad](#)

SUS01-BP01 Elegir la región basándose tanto en los requisitos empresariales como en los objetivos de sostenibilidad

Elija una región para su carga de trabajo basándose tanto en los requisitos empresariales como en los objetivos de sostenibilidad para optimizar sus KPI, incluidos el rendimiento, el coste y la huella de carbono.

Patrones comunes de uso no recomendados:

- Selecciona la región de la carga de trabajo en función de la propia ubicación.
- Consolida todos los recursos de la carga de trabajo en una ubicación geográfica.

Beneficios de establecer esta práctica recomendada: la colocación de una carga de trabajo cerca de proyectos de energías renovables de Amazon o de regiones con baja intensidad de carbono publicada puede ayudar a reducir la huella de carbono de una carga de trabajo en la nube.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: medio

Guía para la implementación

La Nube de AWS es una red en constante expansión de regiones y puntos de presencia (POP), con una infraestructura de red global que los une. La elección de la región para su carga de trabajo afecta significativamente a sus KPI, incluidos el rendimiento, el coste y la huella de carbono. Para mejorar eficazmente estos KPI, debe elegir las regiones para su carga de trabajo basándose tanto en los requisitos empresariales como en los objetivos de sostenibilidad.

Pasos para la aplicación

- Siga estos pasos para evaluar y preseleccionar las posibles regiones para la carga de trabajo en función de los requisitos empresariales, incluido el cumplimiento, las características disponibles, el coste y la latencia:
 - Confirme que estas regiones cumplen con la normativa local vigente.
 - Utilice las [listas de servicios regionales de AWS](#) para comprobar si las regiones disponen de los servicios y las características que necesita para ejecutar su carga de trabajo.
 - Calcule el coste de la carga de trabajo en cada región mediante [AWS Pricing Calculator](#).
 - Pruebe la latencia de la red entre las ubicaciones de sus usuarios finales y cada Región de AWS.
- Elija regiones cerca de proyectos de energías renovables de Amazon y regiones en las que la intensidad de carbono recogida en la cuadrícula sea más baja que en otras ubicaciones (o regiones).
 - Determine las directrices de sostenibilidad pertinentes para realizar un seguimiento y comparar las emisiones de carbono de un año a otro basándose en el [protocolo de gases de efecto invernadero](#) (métodos basados en el mercado y en la ubicación).

- Elija la región en función del método que utilice para hacer un seguimiento de las emisiones de carbono. Para obtener más detalles sobre la elección de una región en función de sus directrices de sostenibilidad, consulte [How to select a Region for your workload based on sustainability goals](#) (Cómo seleccionar una región para su carga de trabajo en función de los objetivos de sostenibilidad).

Recursos

Documentos relacionados:

- [Descripción de las estimaciones de emisiones de carbono](#)
- [Amazon en todo el mundo](#)
- [Metodología de energía renovable](#)
- [Qué tener en cuenta al seleccionar una región para las cargas de trabajo](#)

Vídeos relacionados:

- [Architecting sustainably and reducing your AWS carbon footprint](#) (Arquitectura sostenible y reducción de la huella de carbono de AWS)

Alineación con la demanda

La forma en que los usuarios y las aplicaciones consumen sus cargas de trabajo y otros recursos puede ayudarle a identificar las mejoras necesarias para alcanzar sus objetivos de sostenibilidad. Escale la infraestructura para adaptarla continuamente a la demanda y compruebe que solo utiliza los recursos mínimos necesarios para prestar asistencia a sus usuarios. Alinee los niveles de servicio con las necesidades de los clientes. Posicione los recursos de forma que se limite el uso de red necesario para que los usuarios puedan consumirlos. Elimine los activos sin usar. Proporcione a los miembros de su equipo dispositivos que satisfagan sus necesidades con un impacto mínimo en la sostenibilidad.

Prácticas recomendadas

- [SUS02-BP01 Escalar la infraestructura de la carga de trabajo dinámicamente](#)
- [SUS02-BP02: Alineación de los SLA con los objetivos de sostenibilidad](#)
- [SUS02-BP03: Detener la creación y el mantenimiento de los recursos no utilizados](#)

- [SUS02-BP04 Optimizar la ubicación geográfica de las cargas de trabajo en función de sus requisitos de red](#)
- [SUS02-BP05: Optimización de los recursos de los miembros del equipo para las actividades realizadas](#)
- [SUS02-BP06 Implementar el almacenamiento en búfer o la limitación para aplanar la curva de demanda](#)

SUS02-BP01 Escalar la infraestructura de la carga de trabajo dinámicamente

Utilice la elasticidad de la nube y escale su infraestructura de forma dinámica para adaptar la oferta de recursos en la nube a la demanda y evitar un exceso de capacidad en su carga de trabajo.

Patrones comunes de uso no recomendados:

- No se escala la infraestructura con la carga de usuarios.
- La infraestructura se escala manualmente todo el tiempo.
- Deja la capacidad aumentada después de un evento de ajuste de escala en lugar de volver a desescalar verticalmente.

Ventajas de establecer esta práctica recomendada: la configuración y las pruebas de la elasticidad de la carga de trabajo contribuyen a ajustar de forma eficaz la oferta de recursos en la nube a la demanda y a evitar el exceso de capacidad aprovisionada. Puede aprovechar la elasticidad de la nube para escalar automáticamente la capacidad durante y después de los picos de demanda para asegurarse de que solo utiliza el número correcto de recursos necesarios para satisfacer los requisitos empresariales.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: medio

Guía para la implementación

La nube ofrece la flexibilidad de ampliar o reducir sus recursos de forma dinámica a través de diversos mecanismos para satisfacer los cambios en la demanda. La correspondencia óptima entre la oferta y la demanda ofrece el menor impacto medioambiental para una carga de trabajo.

La demanda puede ser fija o variable, lo que requiere métricas y automatización para garantizar que la administración no resulte difícil. Las aplicaciones pueden escalarse o desescalarsse verticalmente

mediante la modificación del tamaño de la instancia, escalarse o desescalarsse horizontalmente mediante la modificación del número de instancias, o una combinación de ambas.

Puede usar distintos enfoques para hacer que el suministro de recursos coincida con la demanda.

- Enfoque de seguimiento de objetivos: supervise su métrica de escalado y aumente o reduzca de forma automática la capacidad a medida que la necesite.
- Escalamiento predictivo: desescale horizontalmente para prever las tendencias diarias y semanales.
- Enfoque basado en la programación: configure su propia programación según los cambios de carga predecibles.
- Escalamiento de servicios: elija servicios (como los que son sin servidor) que se escalan de forma nativa por diseño o proporcione el escalamiento automático como una característica.

Identifique los períodos de uso reducido o inexistente y escale los recursos en consonancia para eliminar el exceso de capacidad y mejorar la eficiencia.

Pasos para la aplicación

- La elasticidad hace coincidir la oferta de los recursos que tiene con la demanda de esos recursos. Las instancias, los contenedores y las funciones proporcionan mecanismos de elasticidad, ya sea en combinación con el escalamiento automático o como características del servicio. AWS proporciona una serie de mecanismos de escalamiento automático para garantizar que las cargas de trabajo puedan desescalarsse verticalmente de forma rápida y sencilla durante los periodos con poca carga de usuarios. A continuación, se presentan algunos ejemplos de mecanismos de escalamiento automático:

Auto scaling mechanism	Where to use
Amazon EC2 Auto Scaling	Se usa para verificar que tiene el número correcto de instancias de Amazon EC2 disponibles para gestionar la carga de usuarios de su aplicación.
Application Auto Scaling	Se usa para escalar automáticamente los recursos de servicios de AWS individuales más allá de Amazon EC2, como funciones

Auto scaling mechanism	Where to use
Cluster Autoscaler de Kubernetes	<p>de Lambda o servicios de Amazon Elastic Container Service (Amazon ECS).</p> <p>Se usa para escalar automáticamente clústeres de Kubernetes en AWS.</p>

- A menudo se habla de escalamiento en relación con servicios de computación como instancias de Amazon EC2 o funciones AWS Lambda. Considere la configuración de servicios no computacionales como unidades de capacidad de lectura y escritura de [Amazon DynamoDB](#) o particiones de [Amazon Kinesis Data Streams](#) para ajustarse a la demanda.
- Verifique que las métricas para escalar o desescalar verticalmente se validan con respecto al tipo de carga de trabajo que se está desplegando. Si está desplegando una aplicación de transcodificación de vídeo, se espera una utilización del 100 % de la CPU y no debería ser su métrica principal. Puede usar una [métrica personalizada](#) (como la utilización de la memoria) para su política de escalamiento si es necesario. Para elegir las métricas adecuadas, tenga en cuenta las siguientes directrices para Amazon EC2:
 - La métrica debe ser una métrica de utilización válida y describir el grado de ocupación de una instancia.
 - El valor de la métrica debe aumentar o disminuir proporcionalmente al número de instancias del grupo de Auto Scaling.
- Utilice el [escalamiento dinámico](#) en lugar del [manual](#) para su grupo de Auto Scaling. También le recomendamos que use las [políticas de escalamiento de seguimiento de destino](#) en el escalamiento dinámico.
- Verifique que los despliegues de la carga de trabajo puedan manejar los eventos de escalamiento y desescalamiento horizontales. Cree escenarios de prueba para los eventos de escalamiento con el fin de verificar que la carga de trabajo se comporta del modo previsto y no afecta a la experiencia del usuario (como la pérdida de sesiones persistentes). Puede utilizar el [historial de actividades](#) para verificar una actividad de escalamiento correspondiente a un grupo de Auto Scaling.
- Evalúe los patrones predecibles de su carga de trabajo y escale de forma proactiva al anticiparse a los cambios previstos y planeados en la demanda. Con el escalamiento predictivo, puede eliminar la necesidad de aprovisionar capacidad en exceso. Para obtener más detalles, consulte [Escalamiento predictivo con Amazon EC2 Auto Scaling](#).

Recursos

Documentos relacionados:

- [Getting Started with Amazon EC2 Auto Scaling](#) (Introducción a Amazon EC2 Auto Scaling)
- [Predictive Scaling for EC2, Powered by Machine Learning \(Escalado predictivo para EC2, impulsado por el aprendizaje automático\)](#)
- [Analyze user behavior using Amazon OpenSearch Service, Amazon Data Firehose and Kibana](#) (Análisis del comportamiento del usuario con Amazon OpenSearch Service, Amazon Data Firehose y Kibana)
- [¿Qué es Amazon CloudWatch?](#)
- [Supervisión de la carga de bases de datos con Información sobre rendimiento en Amazon RDS](#)
- [Introducing Native Support for Predictive Scaling with Amazon EC2 Auto Scaling](#) (Introducción a la compatibilidad nativa para escalado predictivo con Amazon EC2 Auto Scaling)
- [Introducing Karpenter - An Open-Source, High-Performance Kubernetes Cluster Autoscaler](#) (Presentación de Karpenter: Cluster Autoscaler de Kubernetes de código abierto y alto rendimiento)
- [Deep Dive on Amazon ECS Cluster Auto Scaling](#) (Profundización en Auto Scaling de clúster de Amazon ECS)

Vídeos relacionados:

- [Build a cost-, energy-, and resource-efficient compute environment](#) (Crear un entorno de computación rentable, eficiente en términos de costes, energía y recursos)
- [Better, faster, cheaper compute: Cost-optimizing Amazon EC2](#) (Computación mejor, más rápida y más barata: Optimización de costes de Amazon EC2) (CMP202-R1)

Ejemplos relacionados:

- [Lab: Amazon EC2 Auto Scaling Group Examples](#) (Laboratorio: ejemplos de grupos de Amazon EC2 Auto Scaling)
- [Lab: Implement Autoscaling with Karpenter](#) (Laboratorio: Implementar escalado automático con Karpenter)

SUS02-BP02: Alineación de los SLA con los objetivos de sostenibilidad

Revise y optimice los acuerdos de nivel de servicio (SLA) de la carga de trabajo en función de sus objetivos de sostenibilidad a fin de minimizar los recursos necesarios para admitir la carga de trabajo sin dejar de satisfacer las necesidades empresariales.

Antipatrones usuales:

- Los SLA de carga de trabajo se desconocen o son ambiguos.
- Defina su SLA solo para la disponibilidad y el rendimiento.
- Utiliza el mismo patrón de diseño (como la arquitectura multi-AZ) para todas sus cargas de trabajo.

Beneficios de establecer esta práctica recomendada: la alineación de los SLA con los objetivos de sostenibilidad conlleva un uso óptimo de los recursos, al tiempo que se satisfacen las necesidades empresariales.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: bajo

Guía para la implementación

Los SLA definen el nivel de servicio que se espera de una carga de trabajo en la nube, como el tiempo de respuesta, la disponibilidad y la retención de datos. Influyen en la arquitectura, el uso de recursos y el impacto medioambiental de una carga de trabajo en la nube. Con una cadencia regular, revise los SLA y realice concesiones para reducir significativamente el uso de recursos a cambio de disminuciones aceptables en los niveles de servicio.

Pasos para la implementación

- Defina o rediseñe SLA que respalden sus objetivos de sostenibilidad y que, a la vez, cumplan sus requisitos empresariales sin superarlos.
- Haga concesiones para disminuir significativamente las repercusiones en la sostenibilidad a cambio de reducciones aceptables en los niveles de servicio.
 - Sostenibilidad y fiabilidad: las cargas de trabajo de alta disponibilidad tienden a consumir más recursos.
 - Sostenibilidad y rendimiento: el uso de más recursos para aumentar el rendimiento podría tener mayor impacto medioambiental.
 - Sostenibilidad y seguridad: las cargas de trabajo demasiado seguras podrían tener mayor impacto medioambiental.

- Use patrones de diseño, como [microservicios en AWS](#), que den prioridad a las funciones fundamentales y permitan unos niveles de servicio más bajos (como objetivos de tiempo de respuesta o de tiempo de recuperación) para las funciones que no sean esenciales.

Recursos

Documentos relacionados:

- [Acuerdos de nivel de servicios \(SLA\) de AWS](#)
- [Importance of Service Level Agreement for SaaS Providers](#) (Importancia de los acuerdos de nivel de servicio para los proveedores de SaaS)

Vídeos relacionados:

- [Delivering sustainable, high-performing architectures](#) (Entrega de arquitecturas sostenibles y de alto rendimiento)
- [Build a cost-, energy-, and resource-efficient compute environment \(Crear un entorno de computación rentable, eficiente en términos de costes, energía y recursos\)](#)

SUS02-BP03: Detener la creación y el mantenimiento de los recursos no utilizados

Retire los activos no utilizados de su carga de trabajo para reducir el número de recursos en la nube necesarios para atender su demanda y minimizar los residuos.

Patrones comunes de uso no recomendados:

- No analiza su aplicación en busca de activos redundantes o que ya no son necesarios.
- No elimina los activos que son redundantes o que ya no son necesarios.

Beneficios de establecer esta práctica recomendada: la eliminación de los activos no utilizados libera recursos y mejora la eficacia general de la carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: bajo

Guía para la implementación

Los activos no utilizados consumen recursos de la nube, como espacio de almacenamiento y potencia de computación. Con la identificación y eliminación de estos activos, podrá liberar estos recursos, lo que dará lugar a una arquitectura en la nube más eficiente. Realice análisis periódicos en los activos de aplicaciones (como los informes precompilados, los conjuntos de datos y las imágenes estáticas) y los patrones de acceso a los activos para identificar cualquier tipo de redundancia, infrautilización y los posibles objetivos de retirada. Elimine esos activos redundantes para reducir el despilfarro de recursos en su carga de trabajo.

Pasos para la implementación

- Utilice herramientas de supervisión para identificar los activos estáticos que ya no sean necesarios.
- Antes de eliminar un activo, evalúe el impacto de su eliminación en la arquitectura.
- Desarrolle un plan y elimine los activos que ya no sean necesarios.
- Consolide los recursos generados superpuestos para eliminar el procesamiento redundante.
- Actualice las aplicaciones para que dejen de producir y almacenar activos que no sean necesarios.
- Indique a terceros que administren en su nombre recursos que ya no son necesarios que dejen de producirlos y almacenarlos.
- Indique a terceros que consoliden los recursos redundantes producidos en su nombre.
- Revise periódicamente la carga de trabajo para identificar y eliminar los activos no utilizados.

Recursos

Documentos relacionados:

- [Optimización de la infraestructura de AWS para la sostenibilidad, parte II: almacenamiento](#)
- [¿Cómo puedo terminar los recursos activos que ya no necesito en mi Cuenta de AWS?](#)

Vídeos relacionados:

- [How do I check for and then remove active resources that I no longer need on my Cuenta de AWS? \(¿Cómo puedo comprobar y, a continuación, eliminar los recursos activos que ya no necesito en mi Cuenta de AWS?\)](#)

SUS02-BP04 Optimizar la ubicación geográfica de las cargas de trabajo en función de sus requisitos de red

Seleccione para su carga de trabajo una ubicación y unos servicios en la nube que acorten la distancia que debe recorrer el tráfico de red y reduzcan el total de recursos de red necesarios para admitir su carga de trabajo.

Patrones comunes de uso no recomendados:

- Se selecciona la región de la carga de trabajo en función de la propia ubicación.
- Consolida todos los recursos de la carga de trabajo en una ubicación geográfica.
- Todo el tráfico fluye a través de sus centros de datos existentes.

Beneficios de establecer esta práctica recomendada: Colocar una carga de trabajo cerca de sus usuarios permite obtener la menor latencia, al tiempo que disminuye el movimiento de datos a través de la red y reduce el impacto medioambiental.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Medio

Guía para la implementación

La infraestructura de Nube de AWS se crea en torno a opciones de ubicación como regiones, zonas de disponibilidad, grupos de ubicaciones y ubicaciones periféricas como [AWS Outposts](#) y [Zonas locales de AWS](#). Estas opciones de ubicación son las responsables de mantener la conectividad entre los componentes de las aplicaciones, los servicios en la nube, las redes periféricas y los centros de datos locales.

Analice los patrones de acceso a la red en su carga de trabajo para identificar cómo utilizar estas opciones de ubicación en la nube y reducir la distancia que debe recorrer el tráfico de red.

Pasos para la implementación

- Analice los patrones de acceso de la red en su carga de trabajo para identificar cómo utilizan los usuarios su aplicación.
 - Use herramientas de monitorización como [Amazon CloudWatch](#) y [AWS CloudTrail](#) para recopilar datos sobre las actividades de la red.
 - Analice los datos para identificar el patrón de acceso de la red.

- Seleccione las regiones para el despliegue de la carga de trabajo en función de los siguientes elementos clave:
 - Su objetivo de sostenibilidad: como se explica en la [Selección de regiones](#).
 - Dónde se encuentran sus datos: en el caso de las aplicaciones con gran cantidad de datos (como macrodatos y machine learning), el código de la aplicación debe ejecutarse lo más cerca posible de los datos.
 - Dónde se encuentran sus usuarios: para las aplicaciones orientadas al usuario, elija una región (o regiones) cercana a los usuarios de su carga de trabajo.
 - Otras restricciones: tenga en cuenta restricciones como la seguridad y el cumplimiento como se explica en [Qué tener en cuenta al seleccionar una región para las cargas de trabajo](#).
- Utilice almacenamiento en caché local o [Soluciones de almacenamiento en caché de AWS](#) para los recursos de uso frecuente con el fin de mejorar el rendimiento, reducir el movimiento de datos y disminuir el impacto medioambiental.

Servicio	Cuándo usar
Amazon CloudFront	Se usa para almacenar en caché el contenido estático como imágenes, scripts y vídeos, así como el contenido dinámico como respuestas de API y aplicaciones web.
Amazon ElastiCache	Se usa para almacenar en caché el contenido de las aplicaciones web.
DynamoDB Accelerator	Se usa para añadir aceleración en memoria a sus tablas de DynamoDB.

- Utilice servicios que puedan ayudarle a ejecutar el código más cerca de los usuarios de su carga de trabajo:

Servicio	Cuándo usar
Lambda@Edge	Se usa para las operaciones que utilizan muchos recursos de computación que se inician cuando los objetos no están en la memoria caché.

Servicio	Cuándo usar
Funciones de Amazon CloudFront	Se usan para casos de uso sencillos como las manipulaciones de solicitudes o respuestas HTTP(s) que pueden iniciarse mediante funciones de corta duración.
AWS IoT Greengrass	Se usa para ejecutar la computación local, la mensajería y el almacenamiento en caché de datos para los dispositivos conectados.

- Use la agrupación de conexiones para permitir reutilizar las conexiones y reducir la cantidad de recursos necesarios.
- Use los almacenes de datos distribuidos que no se basen en conexiones persistentes y en actualizaciones sincrónicas por coherencia para atender a las poblaciones regionales.
- Reemplace la capacidad de red estática preaprovisionada por capacidad dinámica compartida y comparta el impacto en la sostenibilidad de la capacidad de red con otros suscriptores.

Recursos

Documentos relacionados:

- [Optimización de la infraestructura de AWS para la sostenibilidad, parte III: redes](#)
- [Documentación de Amazon ElastiCache](#)
- [¿Qué es Amazon CloudFront?](#)
- [Características clave de Amazon CloudFront](#)

Vídeos relacionados:

- [Demystifying data transfer on AWS \(Desmitificación de la transferencia de datos en AWS\)](#)
- [Scaling network performance on next-gen Amazon EC2 instances \(Escalar el rendimiento de la red en instancias de EC2 de nueva generación\)](#)

Ejemplos relacionados:

- [AWS Networking Workshops \(Talleres de red de AWS\)](#)

- [Architecting for sustainability - Minimize data movement across networks \(Diseño de una arquitectura para la sostenibilidad: minimice el movimiento de datos entre las redes\)](#)

SUS02-BP05: Optimización de los recursos de los miembros del equipo para las actividades realizadas

Optimice los recursos proporcionados a los miembros del equipo para minimizar el impacto en la sostenibilidad medioambiental a la vez que se cubren sus necesidades.

Patrones comunes de uso no recomendados:

- Ignora el impacto de los dispositivos utilizados por los miembros de su equipo en la eficacia global de su aplicación en la nube.
- Administra y actualiza manualmente los recursos que utilizan los miembros del equipo.

Beneficios de establecer esta práctica recomendada: la optimización de los recursos de los miembros del equipo mejora la eficacia general de las aplicaciones basadas en la nube.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: bajo

Guía para la implementación

Analice los dispositivos que usan los miembros de su equipo para consumir sus servicios, el ciclo de vida que se espera que tengan y el impacto económico y en la sostenibilidad. Implemente estrategias para optimizar estos recursos. Por ejemplo, realice las operaciones complejas (como la representación y la compilación) en escritorios en una infraestructura escalable con un uso intensivo, en lugar de hacerlo en sistemas de usuarios únicos de gran potencia infrautilizados.

Pasos para la implementación

- Aprovechne las estaciones de trabajo y otros dispositivos para alinearlos con la forma en que se usan.
- Use escritorios virtuales y streaming de aplicaciones para limitar los requisitos de dispositivos y actualizaciones.
- Traslade a la nube las tareas con un uso intensivo del procesador o la memoria para usar su elasticidad.

- Evalúe el impacto de los procesos y los sistemas en el ciclo de vida de los dispositivos y seleccione aquellas soluciones que minimizan los requisitos para el reemplazo de dispositivos a la vez que satisfacen los requisitos empresariales.
- Implemente la administración remota de los dispositivos para reducir la necesidad de realizar viajes de negocios.
 - [Administrador de flotas de AWS Systems Manager](#) es una experiencia de interfaz de usuario (IU) unificada que le ayuda a administrar de forma remota sus nodos que se ejecutan en AWS o en un entorno local.

Recursos

Documentos relacionados:

- [¿Qué es Amazon WorkSpaces?](#)
- [Optimizador de costes para Amazon WorkSpaces](#)
- [Documentación de Amazon AppStream 2.0](#)
- [NICE DCV](#)

Vídeos relacionados:

- [Managing cost for Amazon WorkSpaces on AWS](#) (Administración de costes para Amazon WorkSpaces en AWS)

SUS02-BP06 Implementar el almacenamiento en búfer o la limitación para aplanar la curva de demanda

El almacenamiento en búfer y la limitación aplanan la curva de demanda y reducen la capacidad aprovisionada necesaria para su carga de trabajo.

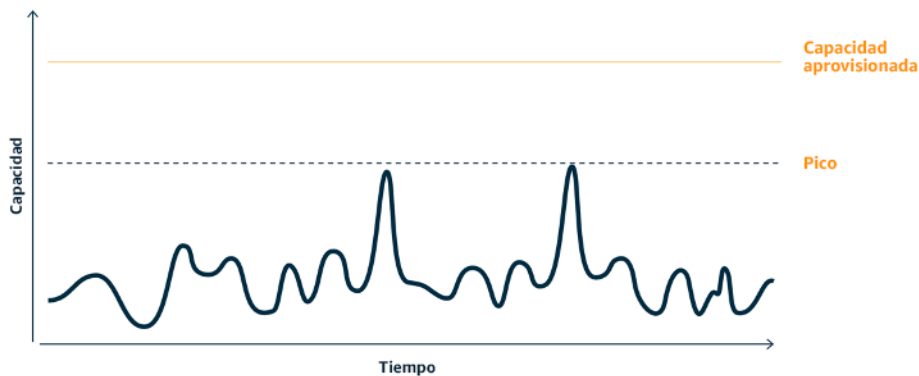
Patrones comunes de uso no recomendados:

- Procesa las solicitudes de los clientes inmediatamente mientras no es necesario.
- No analiza los requisitos de las solicitudes de los clientes.

Beneficios de establecer esta práctica recomendada: el aplanamiento de la curva de demanda reduce la capacidad aprovisionada necesaria para la carga de trabajo. La reducción de la capacidad aprovisionada implica un menor consumo de energía y un menor impacto medioambiental.

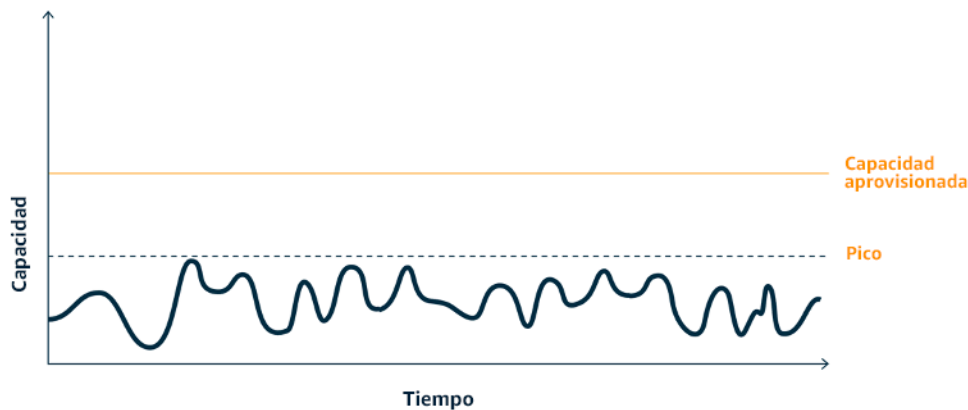
Nivel de riesgo expuesto si no se establece esta práctica recomendada: bajo

El aplanamiento de la curva de demanda de la carga de trabajo puede ayudarle a reducir la capacidad aprovisionada para una carga de trabajo y a reducir su impacto medioambiental. Supongamos una carga de trabajo con la curva de demanda que se muestra en la siguiente figura. Esta carga de trabajo tiene dos picos y, para gestionarlos, se aprovisiona la capacidad de recursos que muestra la línea naranja. Los recursos y la energía utilizados para esta carga de trabajo no están indicados por el área situada debajo de la curva de demanda, sino por el área situada debajo de la línea de capacidad aprovisionada, ya que esta capacidad se necesita para gestionar esos dos picos.



Curva de demanda con dos picos diferenciados que requieren una alta capacidad aprovisionada.

Puede utilizar el almacenamiento en búfer o la limitación para modificar la curva de demanda y suavizar los picos, lo que significa menos capacidad aprovisionada y menos energía consumida. Implemente limitaciones cuando sus clientes puedan realizar reintentos. Implemente el almacenamiento en búfer para almacenar la solicitud y aplazar el procesamiento para más adelante.



Efecto de la limitación sobre la curva de demanda y la capacidad aprovisionada.

Pasos para la implementación

- Analice las solicitudes de los clientes para determinar cómo responder a ellas. Entre las preguntas a tener en cuenta se incluyen las siguientes:
 - ¿Esta solicitud puede procesarse de forma asíncrona?
 - ¿El cliente tiene capacidad de reintentos?
- Si el cliente tiene capacidad de reintentos, puede implementar la limitación, que le indica al origen que si no puede atender la solicitud en el momento actual debe intentarlo más tarde.
 - Puede usar [Amazon API Gateway](#) para implementar la limitación.
- En el caso de los clientes que no pueden realizar reintentos, es necesario implementar un búfer para aplanar la curva de demanda. Un búfer aplaza el procesamiento de las solicitudes, por lo que permite a las aplicaciones que se ejecutan a diferentes ritmos comunicarse de forma efectiva. El enfoque basado en búfer utiliza una cola o una secuencia para aceptar mensajes de los productores. De este modo, los consumidores pueden leer y procesar los mensajes, lo que permite que dichos mensajes se ejecuten a la velocidad que cumpla con los requisitos empresariales de los consumidores.
 - [Amazon Simple Queue Service \(Amazon SQS\)](#) es un servicio administrado que proporciona colas que permiten que un solo consumidor lea mensajes individuales.
 - [Amazon Kinesis](#) ofrece una secuencia que permite que muchos consumidores lean los mismos mensajes.
- Analice la demanda general, la tasa de cambio y el tiempo de respuesta requerido para dimensionar correctamente la limitación o el búfer requeridos.

Recursos

Documentos relacionados:

- [Getting started with Amazon SQS](#) (Introducción a Amazon SQS)
- [Application integration Using Queues and Messages](#) (Integración de aplicaciones mediante colas y mensajes)

Vídeos relacionados:

- [Choosing the Right Messaging Service for Your Distributed App](#) (Elección del servicio de mensajería correcto para su aplicación distribuida)

Software y arquitectura

Implemente patrones que permitan suavizar la carga y mantener un uso elevado consistente de los recursos implementados para minimizar los recursos consumidos. Puede haber componentes que queden inactivos debido a la falta de uso relacionada con los cambios en el comportamiento de los usuarios a lo largo del tiempo. Revise los patrones y la arquitectura para consolidar los componentes infrutilizados a fin de incrementar el uso general. Retire los componentes que ya no son necesarios. Analice el rendimiento de los componentes de su carga de trabajo y optimice aquellos que consumen la mayor cantidad de recursos. Tenga en cuenta los dispositivos que usan los clientes para acceder a sus servicios e implemente patrones para minimizar la necesidad de realizar actualizaciones de los dispositivos.

Prácticas recomendadas

- [SUS03-BP01: Optimizar el software y la arquitectura para los trabajos asíncronos y programados](#)
- [SUS03-BP02 Eliminar o refactorizar los componentes de cargas de trabajo con uso reducido o nulo](#)
- [SUS03-BP03: Optimización de las áreas de código que consumen la mayor parte del tiempo o de los recursos](#)
- [SUS03-BP04 Optimizar el impacto en los dispositivos y equipos](#)
- [SUS03-BP05: Uso de los patrones de software y las arquitecturas que mejor respaldan los patrones de almacenamiento y el acceso a los datos](#)

SUS03-BP01: Optimizar el software y la arquitectura para los trabajos asíncronos y programados

Utilice patrones de software y arquitectura eficientes, como los basados en colas, para mantener una utilización elevada y coherente de los recursos desplegados.

Patrones comunes de uso no recomendados:

- Realiza un aprovisionamiento excesivo de los recursos de su carga de trabajo en la nube para hacer frente a picos imprevistos de la demanda.
- Su arquitectura no desacopla los emisores y los receptores de mensajes asíncronos mediante un componente de mensajería.

Beneficios de establecer esta práctica recomendada:

- Los patrones de software y arquitectura eficientes minimizan los recursos no utilizados en la carga de trabajo y mejoran la eficiencia global.
- Puede escalar el procesamiento independientemente de la recepción de mensajes asíncronos.
- Mediante un componente de mensajería, tendrá unos requisitos de disponibilidad más relajados que podrá cumplir con menos recursos.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Utilice patrones de arquitectura eficientes, como la [arquitectura basada en eventos](#), que dan como resultado una utilización uniforme de los componentes y minimizan el aprovisionamiento excesivo en la carga de trabajo. El uso de patrones de arquitectura eficientes minimiza los recursos inactivos por falta de uso debido a cambios en la demanda a lo largo del tiempo.

Comprenda los requisitos de los componentes de la carga de trabajo y adopte patrones de arquitectura que aumenten la utilización global de los recursos. Retire los componentes que ya no son necesarios.

Pasos para la aplicación

- Analice la demanda de su carga de trabajo para determinar cómo responder a ella.

- En el caso de solicitudes o trabajos que no requieran respuestas síncronas, utilice arquitecturas basadas en colas y empleados de escalamiento automático para maximizar la utilización. A continuación, encontrará algunos ejemplos de cuándo podría plantearse una arquitectura basada en colas:

Queuing mechanism	Description
Colas de trabajo de AWS Batch	Los trabajos de AWS Batch se envían a una cola de trabajos en la que permanecen hasta que pueden programarse para ejecutarse en un entorno de computación.
Instancias de spot de Amazon Simple Queue Service y Amazon EC2	Emparejamiento de Amazon SQS e instancias de spot para crear una arquitectura eficiente y tolerante a errores.

- En el caso de solicitudes o trabajos que puedan procesarse en cualquier momento, utilice mecanismos de programación para procesar los trabajos por lotes y obtener una mayor eficacia. A continuación, se presentan algunos ejemplos de mecanismos de programación en AWS:

Scheduling mechanism	Description
Programador de Amazon EventBridge	Una capacidad de Amazon EventBridge que le permite crear, ejecutar y administrar tareas programadas a escala.
Programación basada en tiempo de AWS Glue	Defina una programación basada en el tiempo para sus rastreadores y trabajos en AWS Glue.
Tareas programadas de Amazon Elastic Container Service (Amazon ECS)	Amazon ECS admite la creación de tareas programadas. Las tareas programadas utilizan reglas de Amazon EventBridge para ejecutar tareas según una programación o en respuesta a un evento de EventBridge.

Scheduling mechanism	Description
Instance Scheduler	Configure las programaciones de inicio y detención de sus instancias de Amazon EC2 y Amazon Relational Database Service.

- Si utiliza mecanismos de sondeo y webhooks en su arquitectura, reemplácelos por eventos. Utilice [arquitecturas basadas en eventos](#) para crear cargas de trabajo de elevada eficacia.
- Aproveche la [tecnología sin servidor en AWS](#) para eliminar la infraestructura aprovisionada en exceso.
- Dimensione correctamente los componentes individuales de su arquitectura para evitar recursos inactivos mientras se espera la entrada.

Recursos

Documentos relacionados:

- [¿Qué es Amazon Simple Queue Service?](#)
- [¿Qué es Amazon MQ?](#)
- [Escalamiento basado en Amazon SQS](#)
- [¿Qué es AWS Step Functions?](#)
- [¿Qué es AWS Lambda?](#)
- [Uso de AWS Lambda con Amazon SQS](#)
- [¿Qué es Amazon EventBridge?](#)

Vídeos relacionados:

- [Moving to event-driven architectures](#) (Migración a arquitecturas basadas en eventos)

SUS03-BP02 Eliminar o refactorizar los componentes de cargas de trabajo con uso reducido o nulo

Elimine los componentes que ya no se usan ni se necesitan y refactorice aquellos con un uso reducido para minimizar el desperdicio en su carga de trabajo.

Patrones comunes de uso no recomendados:

- No comprueba periódicamente el nivel de utilización de los componentes individuales de la carga de trabajo.
- No comprueba ni analiza recomendaciones de herramientas de dimensionamiento de AWS como [AWS Compute Optimizer](#).

Beneficios de establecer esta práctica recomendada: la eliminación de los activos no utilizados libera recursos y mejora la eficacia general de la carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Revise la carga de trabajo para identificar los componentes inactivos o no utilizados. Es un proceso de mejora iterativo que puede desencadenarse por cambios en la demanda o por el lanzamiento de un nuevo servicio en la nube. Por ejemplo, un descenso significativo del tiempo de ejecución de una función de [AWS Lambda](#) puede ser un indicador de la necesidad de reducir el tamaño de la memoria. Además, a medida que AWS lanza nuevos servicios y características, los servicios y la arquitectura óptimos para su carga de trabajo también pueden cambiar.

Supervise continuamente la actividad de la carga de trabajo y busque oportunidades para mejorar el nivel de uso de los componentes individuales. Con la eliminación de los componentes ociosos y la realización de actividades de redimensionamiento, cumplirá los requisitos de su empresa con el menor número de recursos en la nube.

Pasos para la implementación

- Supervise y capture las métricas de utilización de los componentes críticos de su carga de trabajo (como la utilización de la CPU, la utilización de la memoria o el rendimiento de la red en [métricas de Amazon CloudWatch](#)).
- Para cargas de trabajo estables, compruebe las herramientas de redimensionamiento de AWS como [AWS Compute Optimizer](#) a intervalos regulares para identificar los componentes ociosos, no utilizados o infrautilizados.
- En el caso de las cargas de trabajo efímeras, evalúe las métricas de utilización para identificar los componentes inactivos, no utilizados o infrautilizados.
- Retire los componentes y activos asociados (como las imágenes de Amazon ECR) que ya no sean necesarios.
- Refactorice o consolide los componentes infrautilizados con otros recursos para mejorar la eficiencia de uso. Por ejemplo, puede aprovisionar varias bases de datos pequeñas en una sola

instancia de base de datos de [Amazon RDS](#) en vez de ejecutar bases de datos en instancias individuales infrautilizadas.

- Entienda los [recursos aprovisionados por su carga de trabajo para completar una unidad de trabajo](#).

Recursos

Documentos relacionados:

- [AWS Trusted Advisor](#)
- [¿Qué es Amazon CloudWatch?](#)
- [Limpieza automatizada de imágenes no utilizadas en Amazon ECR](#)

Ejemplos relacionados:

- [Well-Architected Lab - Rightsizing with AWS Compute Optimizer](#) (Laboratorio de Well-Architected: redimensionamiento con AWS Compute Optimizer)
- [Well-Architected Lab - Optimize Hardware Patterns and Observe Sustainability KPIs](#) (Laboratorio de Well-Architected: optimizar los patrones de hardware y observar los KPI de sostenibilidad)

SUS03-BP03: Optimización de las áreas de código que consumen la mayor parte del tiempo o de los recursos

Optimice el código que se ejecuta en los distintos componentes de su arquitectura para minimizar el uso de los recursos y, a la vez, maximizar el rendimiento.

Patrones comunes de uso no recomendados:

- Ignora la optimización del código para el uso de recursos.
- Normalmente responde a los problemas de rendimiento con un aumento de los recursos.
- Su proceso de revisión y desarrollo del código no realiza un seguimiento de los cambios de rendimiento.

Beneficios de establecer esta práctica recomendada: El uso de código eficiente minimiza el uso de recursos y mejora el rendimiento.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Medio

Guía para la implementación

Es fundamental examinar cada área funcional, incluido el código de una aplicación con arquitectura de nube, para optimizar su uso de recursos y su rendimiento. Supervise continuamente el rendimiento de la carga de trabajo en los entornos de creación y producción e identifique oportunidades para mejorar los fragmentos de código que tienen un uso de recursos especialmente elevado. Adopte un proceso de revisión periódico para identificar errores o antipatrones en su código que utilicen los recursos de forma ineficiente. Use algoritmos sencillos y eficaces que produzcan los mismos resultados para su caso de uso.

Pasos para la implementación

- Durante el desarrollo de sus cargas de trabajo, adopte un proceso automatizado de revisión del código para mejorar la calidad e identificar errores y antipatrones.
 - [Automate code reviews with Amazon CodeGuru Reviewer \(Revisiones automáticas de código con Amazon CodeGuru Reviewer\)](#)
 - [Detecting concurrency bugs with Amazon CodeGuru \(Detección de errores de simultaneidad con Amazon CodeGuru\)](#)
 - [Raising code quality for Python applications using Amazon CodeGuru \(Mejora de la calidad del código para aplicaciones Python con Amazon CodeGuru\)](#)
- A medida que ejecute las cargas de trabajo, supervise los recursos para identificar los componentes con elevados requisitos de recursos por unidad de trabajo como objetivos de las revisiones de código.
- Para las revisiones de código, use un generador de perfiles de código para identificar las áreas de código que emplean más tiempo o recursos como objetivo de la optimización.
 - [Reducing your organization's carbon footprint with Amazon CodeGuru Profiler \(Reducción de la huella de carbono de su organización con Amazon CodeGuru Profiler\)](#)
 - [Understanding memory usage in your Java application with Amazon CodeGuru Profiler \(Descripción del uso de memoria en su aplicación Java con Amazon CodeGuru Profiler\)](#)
 - [Improving customer experience and reducing cost with Amazon CodeGuru Profiler \(Mejora de la experiencia del cliente y reducción de costes con Amazon CodeGuru Profiler\)](#)
- Use el sistema operativo y el lenguaje de programación más eficaces para la carga de trabajo. Para obtener más información sobre los lenguajes de programación energéticamente eficientes (incluido Rust), consulte [Sustainability with Rust \(Sostenibilidad con Rust\)](#).

- Reemplace los algoritmos que hacen un uso intensivo de la computación por versiones más sencillas y eficientes que produzcan el mismo resultado.
- Elimine el código innecesario, como la ordenación y el formato.

Recursos

Documentos relacionados:

- [¿Qué es Amazon CodeGuru Profiler?](#)
- [Instancias de FPGA](#)
- [SDK de AWS en Herramientas para crear en AWS](#)

Vídeos relacionados:

- [Improve Code Efficiency Using Amazon CodeGuru Profiler \(Mejora de la eficiencia del código con Amazon CodeGuru Profiler\)](#)
- [Automate Code Reviews and Application Performance Recommendations with Amazon CodeGuru \(Automatización de las revisiones de código y las recomendaciones de rendimiento de aplicaciones con Amazon CodeGuru\)](#)

SUS03-BP04 Optimizar el impacto en los dispositivos y equipos

Analice los dispositivos y los equipos empleados en la arquitectura y utilice estrategias para reducir su uso. Esto puede minimizar el impacto medioambiental global de su carga de trabajo en la nube.

Patrones comunes de uso no recomendados:

- Ignora el impacto medioambiental de los dispositivos que utilizan sus clientes.
- Administra y actualiza manualmente los recursos que utilizan los clientes.

Beneficios de establecer esta práctica recomendada: la implementación de patrones y características de software optimizados para el dispositivo del cliente puede reducir el impacto medioambiental general de la carga de trabajo en la nube.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

La implementación de patrones y características de software optimizados para el dispositivo del cliente puede reducir el impacto medioambiental general de la carga de trabajo en la nube.

- La implementación de nuevas características compatibles con versiones anteriores puede reducir el número de reemplazos de hardware.
- La optimización de una aplicación para que funcione de forma eficiente en los dispositivos puede contribuir a reducir su consumo de energía y a prolongar la duración de su batería (si funcionan con ella).
- La optimización de una aplicación para dispositivos también puede reducir la transferencia de datos a través de la red.

Comprenda los dispositivos y equipos utilizados en su arquitectura, su ciclo de vida previsto y el impacto de reemplazar esos componentes. Implemente patrones y características de software que puedan minimizar el consumo de energía del dispositivo, la necesidad de los clientes de reemplazarlo y también de actualizarlo manualmente.

Pasos para la implementación

- Realice un inventario de los dispositivos utilizados en su arquitectura. Los dispositivos pueden ser móviles, tabletas, dispositivos IoT, luces inteligentes o incluso dispositivos inteligentes en una fábrica.
- Optimice la aplicación que se ejecuta en los dispositivos:
 - Utilice estrategias como la ejecución de tareas en segundo plano para reducir su consumo de energía.
 - Tenga en cuenta la latencia y el ancho de banda de la red al crear cargas e implemente capacidades que ayuden al funcionamiento óptimo de las aplicaciones en enlaces de alta latencia y ancho de banda bajo.
 - Convierta las cargas útiles y los archivos a los formatos optimizados que requieren los dispositivos. Por ejemplo, puede utilizar [Amazon Elastic Transcoder](#) o [AWS Elemental MediaConvert](#) para convertir archivos multimedia digitales de gran tamaño y alta calidad a formatos que los usuarios puedan reproducir en dispositivos móviles, tabletas, navegadores web y televisores conectados.

- Realice las actividades con un uso intensivo de los recursos informáticos (como la representación de imágenes) en el lado del servidor o use el streaming de aplicaciones para mejorar la experiencia del usuario en los dispositivos más antiguos.
- Segmente y pague los resultados, sobre todo en las sesiones interactivas, para administrar las cargas y limitar los requisitos de almacenamiento local.
- Utilice el mecanismo automatizado vía inalámbrica (OTA) para desplegar actualizaciones en uno o varios dispositivos.
 - Puede utilizar una [canalización de CI/CD](#) para actualizar las aplicaciones móviles.
 - Puede utilizar [AWS IoT Device Management](#) para administrar a distancia los dispositivos conectados a escala.
- Para probar nuevas características y actualizaciones, utilice granjas de dispositivos administrados con conjuntos representativos de hardware e itere el desarrollo para maximizar los dispositivos admitidos. Para obtener más información, consulte [SUS06-BP04 Usar granjas de dispositivos administrados para pruebas](#).

Recursos

Documentos relacionados:

- [¿Qué es AWS Device Farm?](#)
- [Documentación de Amazon AppStream 2.0](#)
- [NICE DCV](#)
- [Tutorial de OTA para actualizar el firmware en dispositivos que ejecutan FreeRTOS](#)

Vídeos relacionados:

- [Introduction to AWS Device Farm](#)(Introducción a AWS Device Farm)

SUS03-BP05: Uso de los patrones de software y las arquitecturas que mejor respaldan los patrones de almacenamiento y el acceso a los datos

Analice cómo se usan los datos en la carga de trabajo, cómo los consumen los usuarios, cómo se transfieren y cómo se almacenan. Utilice patrones y arquitecturas de software que admitan mejor

el acceso a los datos y el almacenamiento para minimizar los recursos de computación, redes y almacenamiento necesarios para admitir la carga de trabajo.

Patrones comunes de uso no recomendados:

- Supone que todas las cargas de trabajo tienen patrones similares de almacenamiento y acceso a los datos.
- Solo utiliza un nivel de almacenamiento, asumiendo que todas las cargas de trabajo encajan en ese nivel.
- Supone que los patrones de acceso a los datos se mantendrán coherentes a lo largo del tiempo.
- Su arquitectura admite una posible ampliación de acceso a los datos, lo que provoca que los recursos permanezcan inactivos la mayor parte del tiempo.

Beneficios de establecer esta práctica recomendada: la selección y la optimización de su arquitectura en función de los patrones de acceso y almacenamiento de datos le ayudará a disminuir la complejidad del desarrollo y a aumentar la utilización general. Saber cuándo utilizar las tablas globales, las particiones de datos y el almacenamiento en caché le ayudará a disminuir la sobrecarga operativa y a escalar en función de sus necesidades de carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Utilice los patrones de software y arquitectura que mejor se adapten a las características de sus datos y a sus patrones de acceso. Por ejemplo, utilice una [arquitectura de datos moderna en AWS](#) que le permita usar servicios personalizados optimizados para sus casos de uso de análisis exclusivos. Estos patrones de arquitectura permiten un procesamiento de datos eficaz y reducen el uso de recursos.

Pasos para la implementación

- Analice las características de los datos y los patrones de acceso para identificar la configuración correcta de sus recursos en la nube. Entre las características clave que se deben tener en cuenta se incluyen:
 - Tipo de datos: estructurados, semiestructurados, no estructurados
 - Crecimiento de datos: limitados, no limitados
 - Durabilidad de datos: persistentes, efímeros, transitorios

- Patrones de acceso: lecturas o escrituras, frecuencia de actualización, con picos o constantes
- Utilice los patrones de arquitectura que mejor admitan los patrones de acceso y almacenamiento de datos.
- [Let's Architect! Modern data architectures](#) (Let's Architect! Arquitecturas de datos modernas)
- [Databases on AWS: The Right Tool for the Right Job](#) (Bases de datos de AWS: la herramienta adecuada para el trabajo adecuado.)
- Use tecnologías que funcionen de forma nativa con datos comprimidos.
- Utilice [servicios de análisis](#) personalizados para el procesamiento de datos en su arquitectura.
- Use el motor de base de datos que mejor admita su patrón de consulta dominante. Administre sus índices de base de datos para garantizar una ejecución eficaz de las consultas. Para más detalles, consulte [Bases de datos de AWS](#).
- Seleccione protocolos de red que reduzcan la cantidad de capacidad de red consumida en su arquitectura.

Recursos

Documentos relacionados:

- [Formatos de archivo de compatibilidad con la compresión de Athena](#)
- [Uso de COPY con formatos de datos de columnas con Amazon Redshift](#)
- [Conversión del formato de registro de entrada en Firehose](#)
- [Opciones de formato para las entradas y salidas de ETL en AWS Glue](#)
- [Mejora del rendimiento de las consultas en Amazon Athena con la conversión a formato de columnas](#)
- [Carga de archivos de datos comprimidos desde Amazon S3 con Amazon Redshift](#)
- [Supervisión de la carga de bases de datos con Información sobre rendimiento en Amazon Aurora](#)
- [Supervisión de la carga de bases de datos con Información sobre rendimiento en Amazon RDS](#)
- [Clase de almacenamiento de Amazon S3 Intelligent-Tiering](#)

Vídeos relacionados:

- [Building modern data architectures on AWS](#) (Creación de arquitecturas de datos modernas en AWS)

Administración de datos

Implemente prácticas de administración de datos para reducir el almacenamiento provisionado que se necesita para admitir la carga de trabajo y los recursos necesarios para su uso. Analice sus datos y use las configuraciones y tecnologías de almacenamiento que mejor admitan el valor empresarial de los datos y la forma en que se usan. Haga que el ciclo de vida de los datos incluya un almacenamiento más eficaz y de menor rendimiento cuando disminuyan los requisitos y elimine los datos que ya no se requieran.

Prácticas recomendadas

- [SUS04-BP01: Implementación de una política de clasificación de datos](#)
- [SUS04-BP02 Usar tecnologías que admiten patrones de almacenamiento y acceso a los datos](#)
- [SUS04-BP03 Usar políticas para administrar el ciclo de vida de los conjuntos de datos](#)
- [SUS04-BP04 Utilice la elasticidad y la automatización para ampliar el almacenamiento de bloques o el sistema de archivos](#)
- [SUS04-BP05: Eliminación de datos innecesarios o redundantes](#)
- [SUS04-BP06 Usar sistemas de archivos o almacenamiento compartidos para acceder a datos comunes](#)
- [SUS04-BP07: Minimización del movimiento de datos entre redes](#)
- [SUS04-BP08: Realización de copias de seguridad de los datos solo cuando sea difícil volver a crearlos](#)

SUS04-BP01: Implementación de una política de clasificación de datos

Clasifique los datos para comprender su criticidad para los resultados empresariales y elija el nivel de almacenamiento de bajo consumo adecuado para almacenar los datos.

Antipatrones usuales:

- No identifica activos de datos con características similares (como sensibilidad, criticidad empresarial o requisitos normativos) que se estén procesando o almacenando.
- No ha implementado un catálogo de datos para inventariar sus activos de datos.

Beneficios de establecer esta práctica recomendada: la implementación de una política de clasificación de datos le permite determinar el nivel de almacenamiento de bajo consumo para los datos.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

La clasificación de datos implica la identificación de los tipos de datos que se están procesando y almacenando en un sistema de información propiedad de una organización o controlado por ella. También implica tomar una determinación sobre la criticidad de los datos y la posible repercusión de su divulgación, pérdida o uso indebido.

Implemente la política de clasificación de datos mediante un trabajo en sentido inverso a partir del uso contextual de los datos y la creación de un esquema de categorización que tenga en cuenta el nivel de criticidad de un conjunto de datos determinado para las operaciones de una organización.

Pasos para la implementación

- Realice un inventario de los distintos tipos de datos que existen para su carga de trabajo.
 - Para obtener más detalles sobre las categorías de clasificación de datos, consulte el [documento técnico Data Classification](#) (Clasificación de datos).
- Determine la criticidad, la confidencialidad, la integridad y la disponibilidad de los datos en función del riesgo para la organización. Utilice estos requisitos para agrupar los datos en uno de los niveles de clasificación de datos que adopte.
 - Como ejemplo, consulte [Four simple steps to classify your data and secure your startup](#) (Cuatro sencillos pasos para clasificar sus datos y proteger su startup).
- Audite su entorno de forma periódica para detectar los datos no etiquetados y sin clasificar; a continuación, clasifique y etiquete los datos adecuadamente.
 - Como ejemplo, consulte [Catálogo de datos y rastreadores en AWS Glue](#).
- Establezca un catálogo de datos que proporcione capacidades de auditoría y gobernanza.
- Determine y documente los procedimientos de tratamiento de cada clase de datos.
- Utilice la automatización para auditar continuamente su entorno con el fin de identificar los datos sin etiquetar y sin clasificar, y clasifíquelos y etiquételos adecuadamente.

Recursos

Documentos relacionados:

- [Uso de Nube de AWS para respaldar la clasificación de datos](#)
- [Políticas de etiquetado de AWS Organizations](#)

Vídeos relacionados:

- [Enabling agility with data governance on AWS](#) (Facilitar la agilidad con la gobernanza de los datos en AWS)

SUS04-BP02 Usar tecnologías que admiten patrones de almacenamiento y acceso a los datos

Use las tecnologías de almacenamiento que mejor respalden la forma en que accede y guarda sus datos a fin de minimizar los recursos aprovisionados para admitir la carga de trabajo.

Patrones comunes de uso no recomendados:

- Supone que todas las cargas de trabajo tienen patrones similares de almacenamiento y acceso a los datos.
- Solo utiliza un nivel de almacenamiento, asumiendo que todas las cargas de trabajo encajan en ese nivel.
- Supone que los patrones de acceso a los datos se mantendrán coherentes a lo largo del tiempo.

Beneficios de establecer esta práctica recomendada: seleccionar y optimizar sus tecnologías de almacenamiento en función de los patrones de acceso y almacenamiento de datos le ayudará a reducir los recursos necesarios en la nube para satisfacer sus necesidades empresariales y a mejorar la eficacia general de la carga de trabajo en la nube.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Bajo

Guía para la implementación

Seleccione la solución de almacenamiento que mejor se adapte a sus patrones de acceso, o bien considere cambiar sus patrones de acceso de modo que se adapten a la solución de almacenamiento, a fin de maximizar la eficiencia del rendimiento.

- Evalúe las características de sus datos y su patrón de acceso para recopilar las características clave de sus necesidades de almacenamiento. Entre las características clave que se deben tener en cuenta se incluyen:
 - Tipo de datos: estructurados, semiestructurados y no estructurados
 - Crecimiento de los datos: delimitado, no delimitado
 - Durabilidad de los datos: persistentes, efímeros, transitorios
 - Patrones de acceso: lecturas o escrituras, frecuencia, con picos o constantes
- Migre los datos a la tecnología de almacenamiento adecuada que sea compatible con las características de sus datos y su patrón de acceso. A continuación, le presentamos algunos ejemplos de tecnologías de almacenamiento de AWS y sus principales características:

Tipo	Tecnología	Características clave
Clases de almacenamiento	Amazon S3	Un servicio de almacenamiento de objetos con escalabilidad ilimitada, alta disponibilidad y varias opciones de accesibilidad. La transferencia y el acceso a los objetos dentro y fuera de Amazon S3 puede utilizar un servicio como Aceleración de transferencia o bien Puntos de acceso para respaldar su ubicación, necesidades de seguridad y patrones de acceso.
Almacenamiento de archivos	Amazon S3 Glacier	Clase de almacenamiento de Amazon S3 desarrollada para el archivado de datos.
Sistema de archivos compartidos	Amazon Elastic File System (Amazon EFS)	Sistema de archivos montable al que pueden acceder varios tipos de soluciones de computaci

Tipo	Tecnología	Características clave
		<p>ón. Amazon EFS aumenta y reduce automáticamente el almacenamiento y su rendimiento se ha optimizado para ofrecer latencias bajas y constantes.</p>
Sistema de archivos compartidos	Amazon FSx	<p>Se basa en las últimas soluciones de computación de AWS para admitir cuatro sistemas de archivos de uso común: NetApp ONTAP, OpenZFS, Windows File Server y Lustre. En Amazon FSx, su latencia, rendimiento y E/S por segundo varían según el sistema de archivos y deben tenerse en cuenta a la hora de seleccionar el sistema de archivos adecuado para sus necesidades de carga de trabajo.</p>

Tipo	Tecnología	Características clave
Almacenamiento de bloques	Amazon Elastic Block Store (Amazon EBS)	Servicio de almacenamiento de bloques escalable y de alto rendimiento diseñado para Amazon Elastic Compute Cloud (Amazon EC2). Amazon EBS incluye almacenamiento respaldado por SSD para cargas de trabajo transaccionales y de IOPS intensivas, y almacenamiento respaldado por HDD para cargas de trabajo de rendimiento intensivo.
Base de datos relacional	Amazon Aurora , Amazon RDS , Amazon Redshift	Se han diseñado para respaldar las transacciones ACID (atomicidad, coherencia, aislamiento, durabilidad) y mantener la integridad referencial y una fuerte coherencia de datos. Muchas aplicaciones tradicionales, la planificación de recursos empresariales (ERP), la administración de las relaciones con los clientes (CRM) y los sistemas de comercio electrónico utilizan bases de datos relacionales para almacenar sus datos.

Tipo	Tecnología	Características clave
Base de datos de clave-valor	Amazon DynamoDB	Optimizada para patrones de acceso comunes, normalmente para almacenar y recuperar grandes volúmenes de datos. Las aplicaciones web con mucho tráfico, los sistemas de comercio electrónico y las aplicaciones de juegos son casos de uso típicos para las bases de datos de clave-valor.

- Para los sistemas de almacenamiento que tienen un tamaño fijo, como Amazon EBS o Amazon FSx, supervise el espacio de almacenamiento disponible y automatice la asignación de almacenamiento al alcanzar un umbral. Puede usar Amazon CloudWatch para recopilar y analizar diferentes métricas para [Amazon EBS](#) y [Amazon FSx](#).
- Las clases de almacenamiento de Amazon S3 pueden configurarse en el nivel de objeto y un único bucket puede contener objetos almacenados en todas las clases de almacenamiento.
- También puede utilizar las políticas de ciclo de vida de Amazon S3 para realizar transiciones automáticas de objetos entre clases de almacenamiento o eliminar datos sin necesidad de realizar cambios en la aplicación. En general, tiene que equilibrar la eficiencia de los recursos, la latencia de acceso y la fiabilidad cuando considere estos mecanismos de almacenamiento.

Recursos

Documentos relacionados:

- [Tipos de volumen de Amazon EBS](#)
- [Almacén de instancias de Amazon EC2](#)
- [Amazon S3 Intelligent-Tiering](#)
- [Características de E/S de Amazon EBS](#)
- [Uso de clases de almacenamiento de Amazon S3](#)
- [¿Qué es Amazon S3 Glacier?](#)

Vídeos relacionados:

- [Patrones de arquitectura para lagos de datos en AWS](#)
- [Deep dive on Amazon EBS \(STG303-R1\) \(Conocer en profundidad Amazon EBS \[STG303-R1\]\)](#)
- [Optimize your storage performance with Amazon S3 \(STG343\) \(Optimizar el rendimiento del almacenamiento con Amazon S3 \[STG343\]\)](#)
- [Building modern data architectures on AWS \(Creación de arquitecturas de datos modernas en AWS\)](#)

Ejemplos relacionados:

- [Amazon EFS CSI Driver \(Controlador CSI de Amazon EFS\)](#)
- [Amazon EBS CSI Driver \(Controlador CSI de Amazon EBS\)](#)
- [Amazon EFS Utilities \(Utilidades de Amazon EFS\)](#)
- [Amazon EBS Autoscale \(Escala automática de Amazon EBS\)](#)
- [Amazon S3 Examples \(Ejemplos de Amazon S3\)](#)

SUS04-BP03 Usar políticas para administrar el ciclo de vida de los conjuntos de datos

Administre el ciclo de vida de todos sus datos y aplique automáticamente la eliminación para minimizar el almacenamiento total necesario para su carga de trabajo.

Patrones comunes de uso no recomendados:

- Elimina los datos manualmente.
- No elimina ningún dato de su carga de trabajo.
- No traslada los datos a niveles de almacenamiento de mayor eficiencia energética en función de sus requisitos de retención y acceso.

Beneficios de establecer esta práctica recomendada: el uso de políticas de ciclo de vida de los datos garantiza un acceso y una conservación eficaces de los datos en una carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Los conjuntos de datos suelen tener distintos requisitos de conservación y acceso durante su ciclo de vida. Por ejemplo, su aplicación puede necesitar acceso frecuente a algunos conjuntos de datos durante un periodo de tiempo limitado. Después, se accede a esos conjuntos de datos con poca frecuencia.

Para administrar eficazmente los conjuntos de datos a lo largo de su ciclo de vida, configure las políticas de ciclo de vida, que son reglas que definen cómo administrar los conjuntos de datos.

Con las reglas de configuración del ciclo de vida, puede indicar al servicio de almacenamiento específico que realice la transición de un conjunto de datos a niveles de almacenamiento de mayor eficiencia energética, que lo archive o que lo elimine.

Pasos para la aplicación

- [Clasifique los conjuntos de datos de la carga de trabajo.](#)
- Defina procedimientos de gestión para cada clase de datos.
- Establezca políticas de ciclo de vida automatizadas para la aplicación de reglas de ciclo de vida. A continuación, se ofrecen algunos ejemplos de cómo configurar políticas automatizadas de ciclo de vida para distintos servicios de almacenamiento de AWS:

Storage service	How to set automated lifecycle policies
Amazon S3	<p>Puede usar Amazon S3 Lifecycle para administrar sus objetos a lo largo de su ciclo de vida. Si sus patrones de acceso son desconocidos, cambiantes o impredecibles, puede utilizar Amazon S3 Intelligent-Tiering, que supervisa los patrones de acceso y mueve automáticamente los objetos a los que no se ha accedido a niveles de acceso de menor coste. Puede aprovechar las métricas de Amazon S3 Storage Lens para identificar las oportunidades de optimización y las lagunas en la administración del ciclo de vida.</p>

Storage service	How to set automated lifecycle policies
Amazon Elastic Block Store	Puede utilizar Amazon Data Lifecycle Manager para automatizar la creación, retención y eliminación de instantáneas de Amazon EBS y AMI respaldadas por Amazon EBS.
Amazon Elastic File System	La administración del ciclo de vida de Amazon EFS se ocupa automáticamente el almacenamiento de archivos para sus sistemas de archivos.
Amazon Elastic Container Registry	Las políticas de ciclo de vida de Amazon ECR automatizan la limpieza de las imágenes de contenedor al hacer caducar las imágenes por antigüedad o cantidad.
AWS Elemental MediaStore	Puede utilizar una política de ciclo de vida de objetos que controle durante cuánto tiempo deben almacenarse los objetos en el contenedor de MediaStore.

- Elimine los volúmenes, las instantáneas y los datos no utilizados que estén fuera de su periodo de retención. Aproveche las características nativas del servicio, como el tiempo de vida de Amazon DynamoDB o la retención de registros de Amazon CloudWatch para su eliminación.
- Agregue y comprima datos cuando proceda en función de las reglas de ciclo de vida.

Recursos

Documentos relacionados:

- [Optimice sus reglas de Amazon S3 Lifecycle con el análisis de clases de almacenamiento de Amazon S3](#)
- [Evaluación de recursos con Reglas de AWS Config](#)

Vídeos relacionados:

- [Simplify Your Data Lifecycle and Optimize Storage Costs With Amazon S3 Lifecycle](#) (Simplifique su ciclo de vida de datos y optimice los costes de almacenamiento con Amazon S3 Lifecycle)
- [Reduce Your Storage Costs Using Amazon S3 Storage Lens](#) (Reduzca sus costes de almacenamiento con Amazon S3 Storage Lens)

SUS04-BP04 Utilice la elasticidad y la automatización para ampliar el almacenamiento de bloques o el sistema de archivos

Utilice la elasticidad y la automatización para ampliar el almacenamiento de bloques o el sistema de archivos a medida que crecen los datos para minimizar el almacenamiento total aprovisionado.

Patrones comunes de uso no recomendados:

- Adquiere un almacenamiento de bloques grande o un sistema de archivos de gran tamaño para necesidades futuras.
- Aprovisiona en exceso las operaciones de entrada y salida por segundo (IOPS) de su sistema de archivos.
- No supervisa el uso de sus volúmenes de datos.

Beneficios de establecer esta práctica recomendada: minimizar el aprovisionamiento excesivo del sistema de almacenamiento reduce los recursos inactivos y mejora la eficacia general de su carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Cree almacenamiento de bloques y sistemas de archivos con una asignación de tamaño, rendimiento y latencia adecuados para su carga de trabajo. Utilice la elasticidad y la automatización para ampliar el almacenamiento de bloques o el sistema de archivos a medida que crecen los datos sin tener que aprovisionar en exceso estos servicios de almacenamiento.

Pasos para la implementación

- En el caso del almacenamiento de tamaño fijo, como [Amazon EBS](#), asegúrese de supervisar la cantidad de almacenamiento utilizada en relación con el tamaño total del almacenamiento y cree una automatización, si es posible, para aumentar el tamaño del almacenamiento cuando se alcance un umbral.

- Use volúmenes elásticos y servicios administrados de datos en bloque para automatizar la asignación de almacenamiento adicional a medida que aumentan sus datos persistentes. Por ejemplo, puede utilizar [Volúmenes elásticos de Amazon EBS](#) para cambiar el tamaño de volumen, el tipo de volumen o ajustar el rendimiento de sus volúmenes de Amazon EBS.
- Elija la clase de almacenamiento, el modo de rendimiento y el modo de caudal adecuados para que su sistema de archivos responda a su necesidad empresarial, sin excederse.
 - [Rendimiento de Amazon EFS](#)
 - [Rendimiento de los volúmenes de Amazon EBS en instancias de Linux](#)
- Establezca niveles como objetivo de uso para los volúmenes de datos y ajuste el tamaño de los volúmenes que estén fuera de los intervalos esperados.
- Establezca el tamaño correcto de los volúmenes de solo lectura según los datos.
- Migre los datos a almacenes de objetos para evitar el aprovisionamiento del exceso de capacidad de los tamaños de volúmenes fijos en el almacenamiento en bloque.
- Revise periódicamente los volúmenes elásticos y los sistemas de archivos para terminar los volúmenes inactivos y reducir los recursos aprovisionados en exceso para ajustarlos al tamaño actual de los datos.

Recursos

Documentos relacionados:

- [Documentación de Amazon FSx](#)
- [What is Amazon Elastic File System? \(¿Qué es Amazon Elastic File System?\)](#)

Vídeos relacionados:

- [Deep Dive on Amazon EBS Elastic Volumes](#) (Profundización en los volúmenes elásticos de Amazon EBS)
- [Amazon EBS and Snapshot Optimization Strategies for Better Performance and Cost Savings](#) (Estrategias de optimización de Amazon EBS e instantáneas para mejorar el rendimiento y ahorrar costes)
- [Optimizing Amazon EFS for cost and performance, using best practices](#) (Optimización de Amazon EFS para costes y rendimiento, mediante prácticas recomendadas)

SUS04-BP05: Eliminación de datos innecesarios o redundantes

Elimine datos innecesarios o redundantes para minimizar los recursos de almacenamiento necesarios para guardar sus conjuntos de datos.

Patrones comunes de uso no recomendados:

- Duplica datos que se pueden obtener o recrear fácilmente.
- Realiza copia de seguridad de todos los datos sin tener en cuenta su criticidad.
- Solo elimina datos de forma irregular, en eventos operativos o no los elimina en absoluto.
- Almacena datos de forma redundante independientemente de la durabilidad del servicio de almacenamiento.
- Activa el control de versiones de Amazon S3 sin ninguna justificación empresarial.

Beneficios de establecer esta práctica recomendada: la eliminación de datos redundantes reduce el tamaño de almacenamiento necesario de la carga de trabajo y su impacto medioambiental.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

No almacene datos que no necesite. Automatice la eliminación de datos innecesarios. Use tecnologías que desduplicen los datos en el nivel de archivo y de bloque. Aproveche las características de replicación y redundancia de datos nativos de los servicios.

Pasos para la aplicación

- Evalúe si puede evitar almacenar datos mediante los conjuntos de datos existentes de disponibilidad pública en [AWS Data Exchange](#) y [Open Data on AWS](#) (Datos abiertos en AWS).
- Use mecanismos que puedan desduplicar los datos en el nivel de bloque y de objeto. A continuación, se ofrecen algunos ejemplos de cómo desduplicar datos en AWS:

Storage service	Deduplication mechanism
Amazon S3	Use AWS Lake Formation FindMatches para encontrar registros coincidentes en un conjunto de datos (incluidos los que no tienen

Storage service	Deduplication mechanism
	identificadores) con la nueva transformación de ML FindMatches.
Amazon FSx	Active la desduplicación de datos en Amazon FSx para Windows.
Instantáneas de Amazon Elastic Block Store	Las instantáneas son copias de seguridad progresivas, lo que significa que solo se guardan los bloques del dispositivo que han cambiado después de la instantánea más reciente.

- Analice el acceso de datos para identificar los datos innecesarios. Automatice las políticas de ciclo de vida. Aproveche las características nativas del servicio, como el [tiempo de vida de Amazon DynamoDB](#), [Amazon S3 Lifecycle](#) o la [retención de registros de Amazon CloudWatch](#) para su eliminación.
- Utilice las capacidades de virtualización de datos en AWS para mantener los datos en su origen y evitar la duplicación de datos.
 - [Cloud Native Data Virtualization on AWS](#) (Virtualización de datos nativos en la nube en AWS)
 - [Lab: Optimize Data Pattern Using Amazon Redshift Data Sharing](#) (Laboratorio: optimizar el patrón de datos mediante el uso compartido de datos de Amazon Redshift)
- Use una tecnología de copia de seguridad que pueda crear copias incrementales.
- Aproveche la durabilidad de [Amazon S3](#) y la [replicación de Amazon EBS](#) para conseguir sus objetivos de durabilidad en lugar de tecnologías autoadministradas (como una matriz redundante de discos independientes [RAID]).
- Centralice los datos de registro y de seguimiento, desduplicque las entradas de registro que sean idénticas y establezca mecanismos para ajustar los detalles cuando sea necesario.
- Rellene las memorias caché previamente solo cuando se justifique.
- Establezca la supervisión y automatización de la memoria caché para ajustar el tamaño de esta en consonancia.
- Quite los despliegues y los recursos desfasados de los almacenes de objetos y las memorias caché periféricas al introducir nuevas versiones de su carga de trabajo.

Recursos

Documentos relacionados:

- [Cambio de la retención de datos de registro en CloudWatch Logs](#)
- [Desduplicación de datos en Amazon FSx para Windows File Server](#)
- [Características de Amazon FSx para ONTAP, incluida la desduplicación de datos](#)
- [Invalidación de archivos en Amazon CloudFront](#)
- [Uso de AWS Backup para hacer copias de seguridad y restaurar sistemas de archivos de Amazon EFS](#)
- [¿Qué es Amazon CloudWatch Logs?](#)
- [Trabajar con copias de seguridad en Amazon RDS](#)

Vídeos relacionados:

- [Fuzzy Matching and Deduplicating Data with ML Transforms for AWS Lake Formation](#)
(Concordancia difusa y desduplicación de datos con transformaciones de ML para AWS Lake Formation)

Ejemplos relacionados:

- [¿Cómo analizo mis registros de acceso al servidor de Amazon S3 mediante Amazon Athena?](#)

SUS04-BP06 Usar sistemas de archivos o almacenamiento compartidos para acceder a datos comunes

Adopte sistemas de archivos o almacenamiento compartidos para evitar la duplicación de datos y posibilitar una infraestructura más eficiente para su carga de trabajo.

Patrones comunes de uso no recomendados:

- Aprovisiona almacenamiento para cada cliente.
- No desconecta el volumen de datos de los clientes inactivos.
- No proporciona acceso al almacenamiento a través de plataformas y sistemas.

Beneficios de establecer esta práctica recomendada: el uso de sistemas de archivos o almacenamiento compartidos permite compartir los datos con uno o varios consumidores sin tener que copiarlos. De este modo, se reducen los recursos de almacenamiento necesarios para la carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Si tiene varios usuarios o aplicaciones que acceden a los mismos conjuntos de datos, el uso de la tecnología de almacenamiento compartido es esencial para habilitar una infraestructura eficiente para su carga de trabajo. La tecnología de almacenamiento compartido proporciona una ubicación central para almacenar y administrar conjuntos de datos y evitar la duplicación de datos. También refuerza la coherencia de los datos entre los distintos sistemas. Además, la tecnología de almacenamiento compartido permite un uso más eficaz de la potencia de computación, ya que varios recursos de computación pueden acceder a los datos y procesarlos simultáneamente en paralelo.

Obtenga datos de estos servicios de almacenamiento compartido solo cuando los necesite y desconecte los volúmenes que no utilice para liberar recursos.

Pasos para la implementación

- Migre los datos al almacenamiento compartido cuando tengan varios consumidores. A continuación le mostramos algunos ejemplos de tecnología de almacenamiento compartido en AWS:

Storage option	When to use
Amazon EBS Multi-Attach	Amazon EBS Multi-Attach le permite adjuntar un único volumen SSD IOPS aprovisionadas (io1 o io2) a varias instancias que se encuentren en la misma zona de disponibilidad.
Amazon EFS	Consulte When to Choose Amazon EFS (Cuándo elegir Amazon EFS).

Storage option	When to use
Amazon FSx	Consulte Choosing an Amazon FSx File System (Elección de un sistema de archivos de Amazon FSx).
Amazon S3	Las aplicaciones que no requieren una estructura de sistema de archivos y están diseñadas para colaborar con el almacenamiento de objetos pueden utilizar Amazon S3 como una solución de almacenamiento de objetos escalable de forma masiva, duradera y de bajo coste.

- Copie datos en sistemas de archivos compartidos, o recupérellos de ellos, solo cuando sea necesario. Por ejemplo, puede crear un [sistema de archivos de Amazon FSx for Lustre respaldado por Amazon S3](#) y cargar solo el subconjunto de datos necesarios para procesar los trabajos en Amazon FSx.
- Elimine los datos según corresponda a sus patrones de uso, como se indica en [SUS04-BP03 Usar políticas para administrar el ciclo de vida de los conjuntos de datos](#).
- Desconecte los volúmenes de los clientes que no los estén usando de forma activa.

Recursos

Documentos relacionados:

- [Linking your file system to an Amazon S3 bucket](#) (Vinculación del sistema de archivos a un bucket de Amazon S3)
- [Using Amazon EFS for AWS Lambda in your serverless applications](#) (Uso de Amazon EFS para AWS Lambda en las aplicaciones sin servidor)
- [Amazon EFS Intelligent-Tiering Optimizes Costs for Workloads with Changing Access Patterns](#) (Amazon EFS Intelligent-Tiering optimiza los costes para las cargas de trabajo con patrones de acceso cambiantes)
- [Using Amazon FSx with your on-premises data repository](#) (Uso de Amazon FSx con su repositorio de datos local)

Vídeos relacionados:

- [Storage cost optimization with Amazon EFS](#) (Optimización de costes de almacenamiento con Amazon EFS)

SUS04-BP07: Minimización del movimiento de datos entre redes

Utilice sistemas de archivos o almacenamiento de objetos compartidos para acceder a los datos comunes y minimizar el total de recursos de redes necesarios para admitir el movimiento de datos para su carga de trabajo.

Patrones comunes de uso no recomendados:

- Almacena todos los datos en la misma Región de AWS independientemente de dónde se encuentren los usuarios de los datos.
- No optimiza el tamaño ni el formato de los datos antes de moverlos por la red.

Beneficios de establecer esta práctica recomendada: la optimización del movimiento de datos por la red reduce los recursos de redes totales necesarios para la carga de trabajo y disminuye su impacto medioambiental.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Medio

Guía para la implementación

El movimiento de datos por la organización requiere recursos de computación, red y almacenamiento. Utilice técnicas para minimizar el movimiento de datos y mejorar la eficacia general de su carga de trabajo.

Pasos para la implementación

- Considere la proximidad a los datos o a los usuarios como un factor de decisión cuando [seleccione una región para su carga de trabajo](#).
- Particione los servicios que se consumen regionalmente para que los datos específicos de una región se almacenen en la región en la que se consumen.
- Utilice formatos de archivo eficientes (como Parquet u ORC) y comprima los datos antes de moverlos por la red.

- No mueva los datos no utilizados. Algunos ejemplos que pueden ayudarle a evitar mover datos no utilizados:
 - Reduzca las respuestas de la API solo a los datos relevantes.
 - Agregue los datos cuando estén detallados (no se requiere información en el nivel de registro).
 - Consulte [Well-Architected Lab: Optimize Data Pattern Using Amazon Redshift Data Sharing \(Laboratorio de Well-Architected: optimizar el patrón de datos mediante el uso compartido de datos de Amazon Redshift\)](#).
 - Considere [el uso compartido de datos entre cuentas en AWS Lake Formation](#).
- Utilice servicios que puedan ayudarle a ejecutar el código más cerca de los usuarios de su carga de trabajo.

Servicio	Cuándo usar
Lambda@Edge	Se usa para las operaciones que utilizan muchos recursos de computación que se ejecutan cuando los objetos no están en la memoria caché.
Funciones de CloudFront	Se usan en casos de uso sencillos como las manipulaciones de solicitudes o respuestas HTTP(s) que pueden iniciarse mediante funciones de corta duración.
AWS IoT Greengrass	Ejecuta la computación local, la mensajería y el almacenamiento en caché de datos para los dispositivos conectados.

Recursos

Documentos relacionados:

- [Optimización de la infraestructura de AWS para la sostenibilidad, parte III: redes](#)
- [Infraestructura global de AWS](#)
- [Características clave de Amazon CloudFront, incluida la red perimetral global de CloudFront](#)
- [Compresión de solicitudes HTTP en Amazon OpenSearch Service](#)

- [Compresión de datos intermedia con Amazon EMR](#)
- [Carga de archivos de datos comprimidos desde Amazon S3 en Amazon Redshift](#)
- [Entrega archivos comprimidos con Amazon CloudFront](#)

Vídeos relacionados:

- [Demystifying data transfer on AWS \(Desmitificación de la transferencia de datos en AWS\)](#)

Ejemplos relacionados:

- [Architecting for sustainability - Minimize data movement across networks \(Diseño de una arquitectura para la sostenibilidad: minimice el movimiento de datos entre las redes\)](#)

SUS04-BP08: Realización de copias de seguridad de los datos solo cuando sea difícil volver a crearlos

Evite realizar copias de seguridad de datos que no tengan valor empresarial para minimizar los requisitos de recursos de almacenamiento para su carga de trabajo.

Patrones comunes de uso no recomendados:

- No dispone de una estrategia de copia de seguridad para los datos.
- Hace copias de seguridad de datos que pueden volver a crearse fácilmente.

Beneficios de establecer esta práctica recomendada: evitar las copias de seguridad de los datos que no son fundamentales reduce los recursos de almacenamiento necesarios para la carga de trabajo y disminuye su impacto medioambiental.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Evitar la copia de seguridad de datos innecesarios puede contribuir a reducir los costes y los recursos de almacenamiento utilizados por la carga de trabajo. Realice copias de seguridad únicamente de aquellos datos que tengan valor empresarial o que sean necesarios para satisfacer los requisitos de cumplimiento. Examine las políticas de copia de seguridad y excluya el almacenamiento efímero que no proporcione valor alguno en un escenario de recuperación.

Pasos para la implementación

- Implemente la política de clasificación de datos tal y como se indica en [SUS04-BP01: Implementación de una política de clasificación de datos](#).
- Utilice la criticidad de la clasificación de sus datos y diseñe la estrategia de copia de seguridad en función de su [objetivo de tiempo de recuperación \(RTO\) y de su objetivo de punto de recuperación \(RPO\)](#). Evite realizar copias de seguridad de datos no esenciales.
 - Excluya los datos que pueden volver a crearse fácilmente.
 - Excluya los datos efímeros de sus copias de seguridad.
 - Excluya las copias locales de los datos, a menos que el tiempo necesario para restaurar esos datos desde una ubicación común supere lo establecido en los acuerdos de nivel de servicio (SLA).
- Utilice una solución automatizada o un servicio administrado para realizar copias de seguridad de los datos fundamentales para la empresa.
 - [AWS Backup](#) es un servicio totalmente administrado que facilita la centralización y automatización de la protección de datos entre servicios de AWS, en la nube y en el entorno local. Para obtener orientación práctica sobre cómo crear copias de seguridad automatizadas con AWS Backup, consulte [Well-Architected Labs - Testing Backup and Restore of Data](#) (Laboratorios de Well-Architected: Pruebas de copia de seguridad y restauración de datos).
 - [Automate backups and optimize backup costs for Amazon EFS using AWS Backup](#) (Automatizar las copias de seguridad y optimizar los costes de copia de seguridad para Amazon EFS con AWS Backup).

Recursos

Prácticas recomendadas relacionadas:

- [REL09-BP01 Identificar todos los datos de los que se debe hacer una copia de seguridad y crearla o reproducir los datos a partir de los orígenes](#)
- [REL09-BP03 Realizar copias de seguridad de los datos automáticamente](#)
- [REL13-BP02 Usar estrategias de recuperación definidas para cumplir los objetivos de recuperación](#)

Documentos relacionados:

- [Uso de AWS Backup para hacer copias de seguridad y restaurar sistemas de archivos de Amazon EFS](#)
- [Instantáneas de Amazon EBS](#)
- [Trabajar con copias de seguridad en Amazon Relational Database Service](#)
- [Socio de APN: socios que pueden ayudar con la copia de seguridad](#)
- [AWS Marketplace: productos que pueden usarse para la copia de seguridad](#)
- [Copia de seguridad de Amazon EFS](#)
- [Copia de seguridad de Amazon FSx para Windows File Server](#)
- [Copia de seguridad y restauración para Amazon ElastiCache for Redis](#)

Vídeos relacionados:

- [AWS re:Invent 2021 - Backup, disaster recovery, and ransomware protection with AWS](#) (AWS re:Invent 2021: copia de seguridad, recuperación de desastres y protección contra ransomware con AWS)
- [AWS Backup Demo: Cross-Account and Cross-Region Backup](#) (Demostración de AWS Backup: copia de seguridad entre cuentas y entre regiones)
- [AWS re:Invent 2019: Deep dive on AWS Backup, ft. Rackspace \(STG341\)](#)

Ejemplos relacionados:

- [Well-Architected Lab - Testing Backup and Restore of Data](#) (Laboratorio de Well-Architected: probar la copia de seguridad y restauración de los datos)
- [Well-Architected Lab - Backup and Restore with Failback for Analytics Workload](#) (Laboratorio de Well-Architected: copia de seguridad y restauración con conmutación por recuperación para cargas de trabajo de análisis)
- [Well-Architected Lab - Disaster Recovery - Backup and Restore](#) (Laboratorio de Well-Architected: recuperación de desastres, copia de seguridad y restauración)

Hardware y servicios

Realice cambios en sus prácticas de administración de hardware como forma de reducir el impacto en la sostenibilidad de las cargas de trabajo. Minimice la cantidad de hardware necesario para

aprovisionar e implementar y seleccione el hardware y los servicios más eficaces para su carga de trabajo individual.

Prácticas recomendadas

- [SUS05-BP01 Usar la mínima cantidad de hardware para cumplir sus necesidades](#)
- [SUS05-BP02: Uso de los tipos de instancia con el menor impacto](#)
- [SUS05-BP03 Usar servicios administrados](#)
- [SUS05-BP04 Optimizar el uso de aceleradores de computación basados en hardware](#)

SUS05-BP01 Usar la mínima cantidad de hardware para cumplir sus necesidades

Utilice la cantidad mínima de hardware para su carga de trabajo a fin de satisfacer eficazmente sus necesidades empresariales.

Patrones comunes de uso no recomendados:

- No supervisa la utilización de los recursos.
- Tiene recursos con un bajo nivel de utilización en su arquitectura.
- No se revisa la utilización del hardware estático para determinar si debe redimensionarse.
- No establece objetivos de utilización de hardware para su infraestructura de computación en función de los KPI empresariales.

Beneficios de establecer esta práctica recomendada: el redimensionamiento de sus recursos en la nube contribuye a reducir el impacto medioambiental de una carga de trabajo, ahorrar dinero y mantener los niveles de rendimiento.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Seleccione de forma óptima el número total de hardware necesario para su carga de trabajo con el fin de mejorar su eficacia global. La Nube de AWS ofrece la flexibilidad de ampliar o reducir el número de recursos de forma dinámica a través de diversos mecanismos, como [AWS Auto Scaling](#), y satisfacer los cambios en la demanda. También proporciona [API y SDK](#) que permiten modificar los recursos con un mínimo esfuerzo. Use estas capacidades para hacer cambios frecuentes en las

implementaciones de su carga de trabajo. Además, utilice las directrices de dimensionamiento de las herramientas de AWS para usar eficazmente sus recursos en la nube y satisfacer sus necesidades empresariales.

Pasos para la implementación

- Elija el tipo de instancia que mejor se adapte a sus necesidades.
 - [How do I choose the appropriate Amazon EC2 instance type for my workload?](#) (¿Cómo elijo el tipo de instancia Amazon EC2 apropiado para mi carga de trabajo?)
 - [Attribute based instance type selection for Amazon EC2 Fleet](#) (Selección de tipo de instancia basada en atributos para Flota de Amazon EC2).
 - [Crear un grupo de Auto Scaling mediante la selección del tipo de instancia basada en atributos.](#)
- Escale mediante pequeños incrementos para las cargas de trabajo variables.
- Utilice varias opciones de compra de computación para equilibrar la flexibilidad de las instancias, la escalabilidad y el ahorro de costes.
 - Las [instancias bajo demanda](#) son las más adecuadas para las cargas de trabajo nuevas, con estado y con picos que no pueden ser flexibles en cuanto al tipo de instancia, la ubicación o el tiempo.
 - Las [instancias de spot](#) son una excelente forma de complementar las demás opciones para aplicaciones con tolerancia a errores y que son flexibles.
 - Use los [Savings Plans para computación](#) en cargas de trabajo en estado estable que permiten flexibilidad si cambian sus necesidades (como zona de disponibilidad, región, familias de instancias o tipos de instancia).
- Utilice la diversidad de instancias y zonas de disponibilidad para maximizar la disponibilidad de las aplicaciones y aprovechar el exceso de capacidad cuando sea posible.
- Use las recomendaciones de tamaño adecuado de las herramientas de AWS para adaptar su carga de trabajo.
 - [AWS Compute Optimizer](#)
 - [AWS Trusted Advisor](#)
- Negocie acuerdos de nivel de servicio (SLA) que permitan una reducción temporal de la capacidad mientras la automatización despliega recursos de reemplazo.

Recursos

Documentos relacionados:

- [Optimizing your AWS Infrastructure for Sustainability, Part I: Compute](#) (Optimización de la infraestructura de AWS para la sostenibilidad, parte I: computación)
- [Attribute based Instance Type Selection for Auto Scaling for Amazon EC2 Fleet](#) (Selección de tipo de instancia basada en atributos para Auto Scaling para Flota de Amazon EC2)
- [Documentación de AWS Compute Optimizer](#)
- [Operating Lambda: Performance optimization](#) (Operación de Lambda: optimización de rendimiento)
- [Documentación de Auto Scaling](#)

Vídeos relacionados:

- [Build a cost-, energy-, and resource-efficient compute environment \(Crear un entorno de computación rentable, eficiente en términos de costes, energía y recursos\)](#)

Ejemplos relacionados:

- [Well-Architected Labs - Rightsizing with AWS Compute Optimizer and Memory Utilization Enabled \(Level 200\)](#) (Laboratorios de Well-Architected: redimensionamiento con AWS Compute Optimizer y uso de memoria activado [nivel 200])

SUS05-BP02: Uso de los tipos de instancia con el menor impacto

Supervise y utilice continuamente nuevos tipos de instancias para aprovechar las mejoras de la eficiencia energética.

Patrones comunes de uso no recomendados:

- Solo utiliza una familia de instancias.
- Solo utiliza instancias x86.
- Especifica un tipo de instancia en su configuración de Amazon EC2 Auto Scaling.
- Utiliza instancias de AWS para fines para las que no fueron diseñadas (por ejemplo, utiliza instancias optimizadas para computación para una carga de trabajo de uso intensivo de memoria).
- No evalúa de forma regular nuevos tipos de instancia.
- No comprueba recomendaciones de herramientas de dimensionamiento de AWS como [AWS Compute Optimizer](#).

Beneficios de establecer esta práctica recomendada: Al utilizar instancias energéticamente eficientes y del tamaño adecuado, podrá reducir en gran medida el impacto medioambiental y el coste de su carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Medio

Guía para la implementación

El uso de instancias eficientes en la carga de trabajo en la nube es fundamental para el uso reducido de recursos y la rentabilidad. Supervise de forma continuada el lanzamiento de nuevos tipos de instancia y aproveche las mejoras de la eficiencia energética; se incluyen los tipos de instancia diseñados para admitir cargas de trabajo específicas, como el entrenamiento y la inferencia en machine learning y la transcodificación de vídeo.

Pasos para la implementación

- Conozca y explore los tipos de instancia que pueden reducir el impacto medioambiental de su carga de trabajo.
 - Suscríbase a [Novedades de AWS](#) para estar al día con las últimas tecnologías e instancias de AWS.
 - Conozca los diferentes tipos de instancias de AWS.
 - Conozca las instancias basadas en Graviton de AWS que ofrecen el mejor rendimiento por vatio de uso de energía en Amazon EC2 con [re:Invent 2020 - Conocer en profundidad las instancias de Amazon EC2 con procesador AWS Graviton2](#) y [Conocer en profundidad las instancias C7g de Amazon EC2 y AWS Graviton3](#).
- Planifique y realice la transición de su carga de trabajo a los tipos de instancia con el menor impacto.
 - Defina un proceso para evaluar nuevas funciones o instancias para su carga de trabajo. Aproveche la agilidad de la nube para probar rápidamente cómo los nuevos tipos de instancia pueden mejorar la sostenibilidad medioambiental de su carga de trabajo. Utilice las métricas proxy para medir cuántos recursos necesita para completar una unidad de trabajo.
 - Si es posible, modifique su carga de trabajo para que funcione con diversas cantidades de vCPU y de memoria para sacar el máximo partido de su elección de tipo de instancia.
 - Considere la posibilidad de cambiar su carga de trabajo a instancias basadas en Graviton para mejorar la eficiencia del rendimiento de su carga de trabajo.
 - [AWS Graviton Fast Start](#)

- [Considerations when transitioning workloads to AWS Graviton-based Amazon Elastic Compute Cloud instances \(Consideraciones al trasladar cargas de trabajo a instancias de Amazon EC2 basadas en AWS Graviton\)](#)
- [Graviton2 de AWS para ISV](#)
- Considere la selección de la opción de Graviton de AWS en el uso de [los servicios administrados de AWS](#).
- Migre su carga de trabajo a las regiones que ofrezcan las instancias con menor impacto en la sostenibilidad y que sigan cumpliendo sus requisitos empresariales.
- Para las cargas de trabajo de machine learning, utilice hardware personalizado específico para su carga de trabajo, como [AWS Trainium](#), [AWS Inferentia](#) y [Amazon EC2 DL1](#). Las instancias de AWS Inferentia, como las instancias Inf2, ofrecen hasta un 50 % más de rendimiento por vatio que las instancias de Amazon EC2 comparables.
- Utilice [Amazon SageMaker Inference Recommender](#) para el tamaño correcto del punto de conexión de inferencia de ML.
- Para cargas de trabajo con picos (cargas de trabajo con requisitos poco frecuentes de capacidad adicional), utilice [instancias de rendimiento ampliable](#).
- Para cargas de trabajo sin estado y tolerantes a errores, utilice [Instancias de spot de Amazon EC2](#) para incrementar el uso global de la nube y reducir el impacto en la sostenibilidad de los recursos no utilizados.
- Opere y optimice su instancia de carga de trabajo.
 - Para las cargas de trabajo efímeras, evalúe [las métricas de Amazon CloudWatch de instancias](#) como CPUUtilization para identificar si la instancia está inactiva o infrautilizada.
 - Para cargas de trabajo estables, compruebe las herramientas de redimensionamiento de AWS como [AWS Compute Optimizer](#) a intervalos regulares para identificar las oportunidades de optimizar y dimensionar las instancias.
 - [Well-Architected Lab - Rightsizing Recommendations \(Laboratorio de Well-Architected: recomendaciones de redimensionamiento\)](#)
 - [Well-Architected Lab - Rightsizing with Compute Optimizer \(Laboratorio de Well-Architected: redimensionamiento con Compute Optimizer\)](#)
 - [Well-Architected Lab - Optimize Hardware Patterns and Observe Sustainability KPIs \(Laboratorio de Well-Architected: optimizar los patrones de hardware y observar los KPI de sostenibilidad\)](#)

Recursos

Documentos relacionados:

- [Optimización de la infraestructura de AWS para la sostenibilidad, Parte I: computación](#)
- [AWS Graviton](#)
- [Amazon EC2 DL1](#)
- [Flotas de reservas de capacidad de Amazon EC2](#)
- [Flota de spot de Amazon EC2](#)
- [Funciones: configuración de funciones de Lambda](#)
- [Selección de tipo de instancia basada en atributos para la flota de Amazon EC2](#)
- [Building Sustainable, Efficient, and Cost-Optimized Applications on AWS \(Creación de aplicaciones sostenibles, eficientes y con optimización de costes en AWS\)](#)
- [How the Contino Sustainability Dashboard Helps Customers Optimize Their Carbon Footprint \(Cómo el panel de sostenibilidad de Contino ayuda a los clientes a optimizar su huella de carbono\)](#)

Vídeos relacionados:

- [Conocer en profundidad las instancias de Amazon EC2 con procesador AWS Graviton2](#)
- [Conocer en profundidad las instancias C7g de Amazon EC2 y AWS Graviton3](#)
- [Build a cost-, energy-, and resource-efficient compute environment \(Crear un entorno de computación rentable, eficiente en términos de costes, energía y recursos\)](#)

Ejemplos relacionados:

- [Solution: Guidance for Optimizing Deep Learning Workloads for Sustainability on AWS \(Solución: guía para optimizar las cargas de trabajo de aprendizaje profundo para la sostenibilidad en AWS\)](#)
- [Well-Architected Lab - Rightsizing Recommendations \(Laboratorio de Well-Architected: recomendaciones de redimensionamiento\)](#)
- [Well-Architected Lab - Rightsizing with Compute Optimizer \(Laboratorio de Well-Architected: redimensionamiento con Compute Optimizer\)](#)
- [Well-Architected Lab - Optimize Hardware Patterns and Observe Sustainability KPIs \(Laboratorio de Well-Architected: optimizar los patrones de hardware y observar los KPI de sostenibilidad\)](#)

- [Well-Architected Lab - Migrating Services to Graviton \(Laboratorio de Well-Architected: migración de servicios a Graviton\)](#)

SUS05-BP03 Usar servicios administrados

Utilice los servicios administrados para operar con mayor eficacia en la nube.

Patrones comunes de uso no recomendados:

- Utiliza instancias de Amazon EC2 con baja utilización para ejecutar sus aplicaciones.
- Su equipo interno solo administra la carga de trabajo, sin tiempo para centrarse en la innovación o las simplificaciones.
- Despliega y mantiene tecnologías para tareas que pueden ejecutarse con mayor eficacia en servicios administrados.

Beneficios de establecer esta práctica recomendada:

- El uso de servicios administrados traslada la responsabilidad a AWS, que dispone de información sobre millones de clientes que puede ayudar a impulsar nuevas innovaciones y eficiencias.
- El servicio administrado distribuye el impacto medioambiental del servicio entre muchos usuarios gracias a los planos de control de varios principios.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Mediana

Guía para la implementación

Los servicios administrados traspasan a AWS la responsabilidad del mantenimiento de un uso elevado y la optimización de la sostenibilidad del hardware desplegado. Los servicios administrados también eliminan la carga operativa y administrativa del mantenimiento de un servicio, lo que permite a su equipo tener más tiempo para centrarse en la innovación.

Revise su carga de trabajo para identificar los componentes que se pueden reemplazar por servicios administrados por AWS. Por ejemplo, [Amazon RDS](#), [Amazon Redshift](#) y [Amazon ElastiCache](#) proporcionan un servicio de base de datos administrada. [Amazon Athena](#), [Amazon EMR](#) y [Amazon OpenSearch Service](#) proporcionan un servicio de análisis administrados.

Pasos para la implementación

1. Realice un inventario de su carga de trabajo para servicios y componentes.
2. Evalúe e identifique los componentes que se pueden reemplazar por servicios administrados. A continuación, encontrará algunos ejemplos de cuándo podría plantearse el uso de un servicio administrado:

Task	What to use on AWS
Alojamiento de una base de datos	Use instancias de Amazon Relational Database Service (Amazon RDS) administradas en vez de mantener sus propias instancias de Amazon RDS en Amazon Elastic Compute Cloud (Amazon EC2) .
Alojamiento de una carga de trabajo de contenedores	Use AWS Fargate en vez de implementar su propia infraestructura de contenedores.
Alojamiento de aplicaciones web	Use AWS Amplify Hosting como servicio completamente administrado de CI/CD y de alojamiento para sitios web estáticos y aplicaciones web reproducidas en el servidor.

3. Identifique las dependencias y cree un plan de migraciones. Actualice los runbooks y las guías de estrategias según corresponda.
 - [AWS Application Discovery Service](#) recopila y presenta de modo automático la información detallada sobre el uso y las dependencias de aplicaciones para que pueda tomar decisiones más fundamentadas cuando planifique la migración.
4. Pruebe el servicio antes de migrar al servicio administrado.
5. Utilice el plan de migración para reemplazar los servicios autoadministrados por un servicio administrado.
6. Supervise continuamente el servicio una vez finalizada la migración para realizar los ajustes necesarios y optimizar el servicio.

Recursos

Documentos relacionados:

- [Productos de Nube de AWS](#)

- [Calculadora de coste total de propiedad \(TCO\) de AWS](#)
- [Amazon DocumentDB](#)
- [Amazon Elastic Kubernetes Service \(EKS\)](#)
- [Amazon Managed Streaming for Apache Kafka \(Amazon MSK\)](#)

Vídeos relacionados:

- [Cloud operations at scale with AWS Managed Services](#)(Operaciones en la nube a escala con AWS Managed Services)

SUS05-BP04 Optimizar el uso de aceleradores de computación basados en hardware

Optimice el uso de instancias de computación acelerada para reducir las demandas de infraestructura física de su carga de trabajo.

Patrones comunes de uso no recomendados:

- No supervisa el uso de GPU.
- Utiliza una instancia de uso general para la carga de trabajo, mientras que una instancia personalizada puede ofrecer mayor rendimiento, menor coste y mejor rendimiento por vatio.
- Utiliza aceleradores de computación basados en hardware para tareas en las que es más eficiente utilizar alternativas basadas en CPU.

Beneficios de establecer esta práctica recomendada: al optimizar el uso de los aceleradores basados en hardware, puede reducir las demandas de infraestructura física de su carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: Medio

Guía para la implementación

Si necesita una gran capacidad de procesamiento, puede beneficiarse del uso de instancias de computación acelerada, que proporcionan acceso a aceleradores de computación basados en hardware, como unidades de procesamiento gráfico (GPU) y matrices de puertas programables en campo (FPGA). Estos aceleradores de hardware realizan ciertas funciones, como el procesamiento gráfico o la concordancia de patrones de datos, de forma más eficiente que las alternativas basadas en CPU. Muchas cargas de trabajo aceleradas, como el renderizado, la transcodificación y el

machine learning, son muy variables en cuanto al uso de recursos. Ejecute este hardware solo durante el tiempo que sea necesario y retírelo mediante automatización cuando no se requiera para minimizar los recursos consumidos.

Pasos para la implementación

- Identifique qué [instancias de computación acelerada](#) pueden satisfacer sus necesidades.
- Para las cargas de trabajo de machine learning, utilice hardware personalizado específico para su carga de trabajo, como [AWS Trainium](#), [AWS Inferentia](#) y [Amazon EC2 DL1](#). Las instancias de AWS Inferentia, como las instancias Inf2, tienen hasta [un 50 % más de rendimiento por vatio en comparación con instancias de Amazon EC2 comparables](#).
- Recopile la métrica de uso de sus instancias de computación acelerada. Por ejemplo, puede usar un agente de CloudWatch para recopilar métricas como `utilization_gpu` y `utilization_memory` para sus GPU, como se muestra en [Collect NVIDIA GPU metrics with Amazon CloudWatch \(Recopilación de métricas de CPU de NVIDIA con Amazon CloudWatch\)](#).
- Optimice el código, el funcionamiento de la red y la configuración de los aceleradores de hardware para asegurarse de que se aprovecha al máximo el hardware subyacente.
 - [Optimizar la configuración de GPU](#)
 - [GPU Monitoring and Optimization in the Deep Learning AMI \(Supervisión y optimización de la GPU en la AMI de aprendizaje profundo\)](#)
 - [Optimizing I/O for GPU performance tuning of deep learning training in Amazon SageMaker \(Optimización de la E/S para el ajuste del rendimiento de la GPU en el entrenamiento del aprendizaje profundo en Amazon SageMaker\)](#)
- Utilice las bibliotecas de alto rendimiento y los controladores de GPU más recientes.
- Use la automatización para liberar instancias de GPU cuando no se estén usando.

Recursos

Documentos relacionados:

- [Computación acelerada](#)
- [Let's Architect! Architecting with custom chips and accelerators \(Arquitectura con chips y aceleradores personalizados\)](#)
- [How do I choose the appropriate Amazon EC2 instance type for my workload? \(¿Cómo elijo el tipo de instancia de EC2 apropiado para mi carga de trabajo?\)](#)

- [Instancias VT1 de Amazon EC2](#)
- [Choose the best AI accelerator and model compilation for computer vision inference with Amazon SageMaker \(Elija el mejor acelerador de IA y compilación de modelos para la inferencia de visión artificial con Amazon SageMaker\)](#)

Vídeos relacionados:

- [How to select Amazon EC2 GPU instances for deep learning \(Cómo seleccionar las instancias de GPU de Amazon EC2 para el aprendizaje profundo\)](#)
- [Deploying Cost-Effective Deep Learning Inference \(Despliegue rentable de la inferencia del aprendizaje profundo\)](#)

Proceso y cultura

Realice cambios en sus prácticas de desarrollo, prueba e implementación como forma de reducir el impacto en la sostenibilidad.

Prácticas recomendadas

- [SUS06-BP01 Adoptar métodos que permitan introducir mejoras en la sostenibilidad rápidamente](#)
- [SUS06-BP02: Mantenimiento de una carga de trabajo actualizada](#)
- [SUS06-BP03: Incremento de la utilización de los entornos de compilación](#)
- [SUS06-BP04 Usar granjas de dispositivos administrados para pruebas](#)

SUS06-BP01 Adoptar métodos que permitan introducir mejoras en la sostenibilidad rápidamente

Adopte métodos y procesos para validar las mejoras potenciales, minimizar los costes de las pruebas y ofrecer pequeñas mejoras.

Patrones comunes de uso no recomendados:

- La revisión de su solicitud de sostenibilidad es una tarea que se realiza solo una vez al comienzo de un proyecto.
- Su carga de trabajo se ha quedado obsoleta, ya que el proceso de lanzamiento es demasiado complejo para incorporar pequeños cambios para la eficiencia de los recursos.

- No dispone de mecanismos para mejorar su carga de trabajo para la sostenibilidad.

Beneficios de establecer esta práctica recomendada: si establece un proceso para incorporar mejoras de sostenibilidad y realizar un seguimiento de ellas, podrá adoptar continuamente nuevas características y capacidades, eliminar problemas y mejorar la eficacia de la carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: medio

Guía para la implementación

Pruebe y valide las mejoras de sostenibilidad potenciales antes de desplegarlas en producción. Tenga en cuenta el coste de las pruebas al calcular las posibles ventajas futuras de una mejora. Desarrolle métodos de prueba de bajo coste para ofrecer pequeñas mejoras.

Pasos para la implementación

- Añada requisitos para la mejora de la sostenibilidad a sus tareas pendientes de desarrollo.
- Utilice un [proceso de mejora](#) iterativo para identificar, evaluar, priorizar, probar y desplegar estas mejoras.
- Mejore y optimice continuamente sus procesos de desarrollo. Por ejemplo, [automatice su proceso de entrega de software mediante canalizaciones de integración y entrega continuas \(CI/CD\)](#) para probar y desplegar posibles mejoras con el fin de reducir el nivel de esfuerzo y limitar los errores provocados por los procesos manuales.
- Desarrolle y pruebe posibles mejoras con los componentes representativos mínimos viables para reducir el coste de las pruebas.
- Evalúe continuamente el impacto de las mejoras y realice los ajustes necesarios.

Recursos

Documentos relacionados:

- [AWS habilita soluciones de sostenibilidad](#)
- [Scalable agile development practices based on AWS CodeCommit](#)(Prácticas de desarrollo ágil escalables basadas en AWS CodeCommit)

Vídeos relacionados:

- [Delivering sustainable, high-performing architectures](#) (Entrega de arquitecturas sostenibles y de alto rendimiento)

Ejemplos relacionados:

- [Well-Architected Lab - Turning cost & usage reports into efficiency reports](#) (Laboratorio de Well-Architected: convertir informes sobre costes y uso en informes de eficiencia)

SUS06-BP02: Mantenimiento de una carga de trabajo actualizada

Mantenga actualizada su carga de trabajo para adoptar características eficaces, eliminar problemas y mejorar la eficacia general de su carga de trabajo.

Patrones comunes de uso no recomendados:

- Asume que su arquitectura actual es estática y no se actualizará con el tiempo.
- No dispone de sistemas ni de una cadencia regular para evaluar si los programas y paquetes actualizados son compatibles con su carga de trabajo.

Beneficios de establecer esta práctica recomendada: al establecer un proceso para mantener su carga de trabajo actualizada, puede adoptar nuevas funciones y capacidades, resolver problemas y mejorar la eficiencia de la carga de trabajo.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: bajo

Guía para la implementación

La actualización de sistemas operativos, tiempos de ejecución, middlewares, bibliotecas y aplicaciones puede mejorar la eficacia de la carga de trabajo y facilitar la adopción de tecnologías más eficientes. Un software actualizado también puede incluir características que midan el impacto de su carga de trabajo en la sostenibilidad de forma más precisa, ya que los proveedores ofrecen características para cumplir sus objetivos de sostenibilidad propios. Adopte una cadencia periódica para mantener su carga de trabajo al día de las últimas características y versiones.

Pasos para la implementación

- Defina un proceso y un calendario para evaluar nuevas funciones o instancias para su carga de trabajo. Aproveche la agilidad de la nube para probar rápidamente cómo las nuevas funciones pueden mejorar su carga de trabajo para:

- Reducir el impacto en la sostenibilidad.
- Lograr la eficacia operativa.
- Eliminar las barreras para una mejora planificada.
- Mejorar su capacidad a la hora de medir y administrar las repercusiones en la sostenibilidad.
- Inventariar el software y la arquitectura de su carga de trabajo e identificar los componentes que deben actualizarse.
- Puede usar [Inventario de AWS Systems Manager](#) para recopilar los metadatos del sistema operativo (SO), las aplicaciones y los metadatos de instancias de sus instancias de Amazon EC2 y saber rápidamente qué instancias están ejecutando el software y las configuraciones requeridas por su política de software así como las instancias que deben actualizarse.
- Entienda cómo actualizar los componentes de su carga de trabajo.

Workload component	How to update
Imágenes de máquina	Use EC2 Image Builder para administrar las actualizaciones de las imágenes de máquina de Amazon (AMI) para imágenes de servidor Linux o Windows.
Imágenes de contenedor	Use Amazon Elastic Container Registry (Amazon ECR) con su canalización existente para administrar las imágenes de Amazon Elastic Container Service (Amazon ECS) .
AWS Lambda	AWS Lambda incluye características de administración de versiones .

- Utilice la automatización del proceso de actualización para reducir el nivel de esfuerzo para desplegar nuevas funciones y limitar los errores causados por los procesos manuales.
- Puede utilizar [CI/CD](#) para actualizar automáticamente las AMI, las imágenes de contenedor y otros artefactos relacionados con su aplicación en la nube.
- Puede utilizar herramientas como [AWS Systems Manager Patch Manager](#) para automatizar el proceso de actualizaciones del sistema y programar la actividad con [Ventanas de mantenimiento de AWS Systems Manager](#).

Recursos

Documentos relacionados:

- [Centro de arquitectura de AWS](#)
- [Novedades de AWS](#)
- [Herramientas para desarrolladores de AWS](#)

Ejemplos relacionados:

- [Well-Architected Labs - Inventory and Patch Management](#) (Laboratorios de Well-Architected: administración de inventario y parches)
- [Laboratorio: AWS Systems Manager](#)

SUS06-BP03: Incremento de la utilización de los entornos de compilación

Aumente la utilización de recursos para desarrollar, probar y crear sus cargas de trabajo.

Patrones comunes de uso no recomendados:

- Aprovisiona o finaliza manualmente sus entornos de compilación.
- Mantiene sus entornos de compilación en funcionamiento independientemente de las actividades de prueba, compilación o lanzamiento (por ejemplo, ejecución de un entorno fuera del horario laboral de los miembros de su equipo de desarrollo).
- Aprovisiona en exceso los recursos para sus entornos de compilación.

Beneficios de establecer esta práctica recomendada: al aumentar la utilización de los entornos de compilación, puede mejorar la eficacia general de su carga de trabajo en la nube y, al mismo tiempo, asignar los recursos a los creadores para que desarrollen, prueben y creen de forma eficaz.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: bajo

Guía para la implementación

Use la automatización y la infraestructura como código para incorporar los entornos de compilación cuando sea necesario y retirarlos cuando no se utilicen. Un patrón común consiste en programar

períodos de disponibilidad que coincidan con las horas de trabajo de los miembros del equipo de desarrollo. Sus entornos de prueba deben parecerse mucho a la configuración de producción. Sin embargo, busque oportunidades para utilizar tipos de instancia con capacidad de ampliación, instancias de spot de Amazon EC2, servicios de base de datos de escalamiento automático, contenedores y tecnologías sin servidor para coordinar el desarrollo y la capacidad de prueba con el uso. Limite el volumen de datos para cumplir únicamente los requisitos de prueba. Si utiliza datos de producción en las pruebas, estudie las posibilidades de compartir los datos de producción y no trasladarlos.

Pasos para la implementación

- Utilice la infraestructura como código para aprovisionar sus entornos de compilación.
- Use la automatización para administrar el ciclo de vida de sus entornos de desarrollo y pruebas, y maximizar la eficiencia de sus recursos de compilación.
- Utilice estrategias para maximizar la utilización de los entornos de desarrollo y prueba.
 - Use el mínimo viable de entornos representativos para desarrollar y probar mejoras potenciales.
 - Utilice tecnologías sin servidor si es posible.
 - Use instancias bajo demanda para complementar sus dispositivos de desarrollador.
 - Use tipos de instancia con capacidad de ampliación, instancias de spot y otras tecnologías para alinear la capacidad de creación con el uso.
 - Adopte servicios nativos en la nube para obtener un acceso seguro al shell de instancias en lugar de implementar flotas de hosts bastión.
 - Escale automáticamente sus recursos de compilación en función de sus tareas de compilación.

Recursos

Documentos relacionados:

- [AWS Systems Manager Session Manager](#)
- [Instancias de rendimiento ampliable de Amazon EC2](#)
- [¿Qué es AWS CloudFormation?](#)
- [¿Qué es AWS CodeBuild?](#)
- [Instance Scheduler en AWS](#)

Vídeos relacionados:

- [Prácticas recomendadas de integración continua](#)

SUS06-BP04 Usar granjas de dispositivos administrados para pruebas

Utilice granjas de dispositivos administrados para probar eficazmente una nueva característica en un conjunto representativo de hardware.

Patrones comunes de uso no recomendados:

- Prueba y despliega manualmente su aplicación en dispositivos físicos individuales.
- No utiliza el servicio de pruebas de aplicaciones para probar e interactuar con sus aplicaciones (por ejemplo, Android, iOS y aplicaciones web) en dispositivos físicos reales.

Beneficios de establecer esta práctica recomendada: el uso de granjas de dispositivos administrados para probar aplicaciones con tecnología basada en la nube proporciona una serie de ventajas:

- Incluyen características más eficaces para probar la aplicación en una amplia gama de dispositivos.
- Eliminan la necesidad de una infraestructura interna para las pruebas.
- Ofrecen diversos tipos de dispositivos, incluido el hardware más antiguo y menos popular, lo que elimina la necesidad de actualizaciones innecesarias de los dispositivos.

Nivel de riesgo expuesto si no se establece esta práctica recomendada: bajo

Guía para la implementación

El uso de granjas de dispositivos administrados puede ayudarle a agilizar el proceso de prueba de nuevas características en un conjunto representativo de hardware. Las granjas de dispositivos administrados ofrecen diversos tipos de dispositivos, incluido el hardware más antiguo y menos popular, y evitan el impacto en la sostenibilidad para el cliente que tienen las actualizaciones innecesarias de dispositivos.

Pasos para la implementación

- Defina los requisitos y el plan de pruebas (como el tipo de prueba, los sistemas operativos y el calendario de pruebas).
 - Puede utilizar [Amazon CloudWatch RUM](#) para recopilar y analizar datos en el cliente y dar forma a su plan de pruebas.

- Seleccione la granja de dispositivos administrada que pueda admitir sus requisitos de pruebas. Por ejemplo, puede utilizar [AWS Device Farm](#) para probar y comprender el impacto de sus cambios en un conjunto representativo de hardware.
- Utilice la integración continua/despliegue continuo (CI/CD) para programar y ejecutar sus pruebas.
 - [Integrating AWS Device Farm with your CI/CD pipeline to run cross-browser Selenium tests](#) (Integración de AWS Device Farm con su canalización de CI/CD para ejecutar pruebas de Selenium en varios navegadores)
 - [Building and testing iOS and iPadOS apps with AWS DevOps and mobile services](#) (Compilar y probar aplicaciones de iOS y iPadOS con AWS DevOps y servicios móviles)
- Revise continuamente los resultados de sus pruebas y efectúe las mejoras necesarias.

Recursos

Documentos relacionados:

- [Lista de dispositivos de AWS Device Farm](#)
- [Visualización del panel de CloudWatch RUM](#)

Ejemplos relacionados:

- [Aplicación de muestra de AWS Device Farm para Android](#)
- [Aplicación de muestra de AWS Device Farm para iOS](#)
- [Pruebas web de Appium para AWS Device Farm](#)

Vídeos relacionados:

- [Optimize applications through end user insights with Amazon CloudWatch RUM](#) (Optimizar las aplicaciones mediante la información del usuario final con Amazon CloudWatch RUM)

Conclusión

Cada vez más organizaciones establecen objetivos de sostenibilidad en respuesta a los cambios en la normativa gubernamental, la ventaja competitiva y la demanda de los clientes, los empleados y los inversores. Los directores de tecnología, los arquitectos, los desarrolladores y los miembros del equipo de operaciones buscan formas de contribuir directamente a los objetivos de sostenibilidad de su organización. Gracias a estos principios de diseño y a las prácticas recomendadas respaldadas por los servicios de AWS, podrá tomar decisiones informadas que equilibren la seguridad, el coste, el rendimiento, la fiabilidad y la excelencia operativa con los resultados de sostenibilidad de sus cargas de trabajo de Nube de AWS. Cada acción que realice para disminuir el uso de recursos y aumentar la eficiencia en sus cargas de trabajo contribuye a reducir el impacto medioambiental y contribuye a los objetivos de sostenibilidad más amplios de su organización.

Colaboradores

Entre los colaboradores de este documento, están las siguientes personas:

- Sam Mokhtari, Senior Efficiency Lead Solutions Architect, Amazon Web Services
- Brendan Sisson, Principal Sustainability Solutions Architect, Amazon Web Services
- Margaret O'Toole, Sustainability Tech Leader, Amazon Web Services
- Steffen Grunwald, Principal Sustainability Solutions Architect, Amazon Web Services
- Ryan Eccles, Principal Engineer, Sustainability, Amazon
- Rodney Lester, Principal Architect, Amazon Web Services
- Adrian Cockcroft, VP Sustainability Architecture, Amazon Web Services
- Ian Meyers, Director of Technology, Solutions Architecture, Amazon Web Services

Otra documentación

Para obtener información adicional, consulte:

- [AWS Well-Architected](#)
- [Centro de arquitectura de AWS](#)
- [Sostenibilidad en la nube](#)
- [AWS habilita soluciones de sostenibilidad](#)
- [The Climate Pledge](#)
- [Objetivos de desarrollo sostenible de Naciones Unidas](#)
- [Protocolo para gases de efecto invernadero](#)

Revisiones del documento

Para recibir notificaciones sobre las actualizaciones de este documento técnico, suscríbase al canal RSS.

Cambio	Descripción	Fecha
Actualizaciones de los niveles de riesgo	Actualizaciones menores de los niveles de riesgo de las prácticas recomendadas.	October 3, 2023
Actualizaciones de la guía sobre las prácticas recomendadas	Actualizaciones de las prácticas recomendadas con nuevas guías en las siguientes áreas: Alineación con la demanda , Software y arquitectura , Datos y Hardware y servicios .	July 13, 2023
Actualizaciones del nuevo marco	Prácticas recomendadas actualizadas con guía prescriptiva y prácticas recomendadas añadidas.	April 10, 2023
Documento técnico actualizado	Prácticas recomendadas actualizadas con nueva guía de implementación.	December 15, 2022
Documento técnico actualizado	Se han ampliado las prácticas recomendadas y se han añadido planes de mejora.	October 20, 2022
Publicación inicial	Publicación de Pilar de sostenibilidad: AWS Well-Architected Framework.	December 2, 2021

Avisos

Los clientes son responsables de realizar sus propias evaluaciones de la información contenida en este documento. Este documento: (a) solo tiene fines informativos, (b) representa las prácticas y las ofertas de productos vigentes de AWS, que están sujetas a cambios sin previo aviso, y (c) no crea ningún compromiso ni garantía de AWS y sus empresas afiliadas, proveedores o concesionarios de licencias. Los productos o servicios de AWS se proporcionan “tal cual”, sin garantías, representaciones ni condiciones de ningún tipo, ya sean explícitas o implícitas. Las responsabilidades y obligaciones de AWS en relación con sus clientes se rigen por los acuerdos de AWS, y este documento no modifica ni forma parte de ningún acuerdo entre AWS y sus clientes.

© 2021 Amazon Web Services, Inc. o sus empresas afiliadas. Todos los derechos reservados.

AWS Glossary

For the latest AWS terminology, see the [AWS glossary](#) in the Glosario de AWS Reference.