



Optimice los costos de las cargas de trabajo de Microsoft en AWS

AWS Guía prescriptiva



AWS Guía prescriptiva: Optimice los costos de las cargas de trabajo de Microsoft en AWS

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

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

Table of Contents

Introducción	1
Descripción general de	1
Público	1
Cómo usar esta guía	2
Resultados empresariales específicos	3
Proceso de optimización de costos	4
Principales recomendaciones para optimizar los costos	7
Descripción general	7
Recomendaciones principales	7
Optimización y evaluación de licencias de AWS	9
Descripción general	9
Opciones de evaluación	10
Evaluación completa	10
Delimitación del alcance de las cargas de trabajo	11
Recopilación de datos	11
Análisis de datos	12
Planificación de los pasos siguientes	14
Impacto de la evaluación	16
Pasos a seguir a continuación	16
Recursos adicionales	16
Windows en Amazon EC2	17
Automatización de la programación de parada e inicio	18
Descripción general de	18
Casos prácticos	18
Escenario de optimización de costos	19
Recomendaciones de optimización de costos	21
Recursos adicionales	34
Dimensionamiento correcto de las cargas de trabajo de Windows	35
Descripción general de	35
Escenario de optimización de costos	35
Recomendaciones de optimización de costos	36
Recomendaciones	46
Recursos adicionales	46
Selección del tipo de instancia adecuado para las cargas de trabajo de Windows	47

Descripción general de	47
Recomendaciones de optimización de costos	48
Sigüientes pasos	57
Recursos adicionales	59
Uso de licencias propias para las cargas de trabajo de Windows y SQL Server	59
Descripción general de	59
Hosts dedicados de Amazon EC2	60
AWS opciones de licencia	65
Uso de licencias de Windows Server	65
Escenarios de optimización de costos	67
Recomendaciones de optimización de costos	73
Recursos adicionales	73
Optimización del gasto de Windows en Amazon EC2	74
Descripción general de	74
Descripción de Savings Plans	75
Escenarios de optimización de costos	81
Recomendaciones de optimización de costos	85
Recursos adicionales	87
Supervise los costes mediante herramientas AWS	87
Descripción general de	87
Recomendaciones de optimización de costos	88
Recursos adicionales	91
SQL Server	93
Elección de una solución de alta disponibilidad y recuperación ante desastres	94
Descripción general de	94
Grupos de disponibilidad Always On de SQL Server	95
Instancias de clústeres de conmutación por error Always On de SQL Server	97
SIOS DataKeeper	99
Grupos de disponibilidad AlwaysOn	101
Grupos de disponibilidad distribuida	102
Envío de registros	103
AWS Database Migration Service	105
AWS Elastic Disaster Recovery	106
Comparación de los costos	108
Recomendaciones de optimización de costos	112
Recursos adicionales	113

Explicación de las licencias de SQL Server	114
Descripción general de	114
AWS opciones de licenciamiento	114
El impacto en los costos de usar sus propias licencias	115
Optimización de las licencias	115
Recomendaciones de optimización de costos	116
Recursos adicionales	46
Selección de la instancia de EC2 adecuada para las cargas de trabajo de SQL Server	122
Descripción general de	123
Comparación de los costos	123
Escenario de optimización de costos	125
Recomendaciones de optimización de costos	126
Recursos adicionales	130
Unificación de instancias	130
Descripción general de	131
Escenario de optimización de costos	131
Recomendaciones de optimización de costos	133
Recursos adicionales	134
Comparación de las ediciones de SQL Server	134
Descripción general de	134
Impacto del costo	135
Recomendaciones de optimización de costos	137
Recursos adicionales	144
Evaluación de la edición Developer de SQL Server	144
Descripción general de	144
Impacto del costo	145
Recursos adicionales	46
Evaluación de SQL Server en Linux	148
Descripción general de	148
Impacto del costo	149
Recomendaciones de optimización de costos	151
Recursos adicionales	152
Optimización de las estrategias de copia de seguridad de SQL Server	152
Descripción general de	152
Copia de seguridad de servidores con instantáneas compatibles con VSS	153
Realice copias de seguridad de SQL Server mediante AWS Backup	155

Copia de seguridad a nivel de la base de datos	157
Recomendaciones de optimización de costos	167
Recursos adicionales	170
Modernización de las bases de datos de SQL Server	171
Descripción general de	171
Ofertas de bases de datos	171
Comparación entre Amazon RDS y Aurora	172
Recomendaciones de optimización de costos	174
Recursos adicionales	179
Optimización del almacenamiento para SQL Server	180
Descripción general de	180
Tipos de almacenamiento en SSD, rendimiento y costo para Amazon EBS	181
Optimización general de los costos de SSD para Amazon EBS	182
Recursos adicionales	184
Optimización de las licencias de SQL Server con Compute Optimizer	185
Descripción general de	185
Recomendaciones de optimización de costos	185
Configuración de Compute Optimizer	186
Recursos adicionales	188
Optimización del tamaño de SQL Server con Compute Optimizer	188
Descripción general de	188
Configuración de Compute Optimizer	189
Recursos adicionales	190
Revise Trusted Advisor las recomendaciones para las cargas de trabajo de SQL Server	190
Descripción general de	190
Recomendaciones de optimización de costos	190
Configuración Trusted Advisor	191
Recursos adicionales	192
Contenedores	193
Transferencia de aplicaciones de Windows a contenedores	194
Descripción general de	194
Beneficios en términos de costos	194
Recomendaciones de optimización de costos	196
Sigüientes pasos	200
Recursos adicionales	200
Optimice los costos de AWS Fargate las tareas en Amazon ECS	201

Descripción general de	201
Beneficios en términos de costos	201
Recomendaciones de optimización de costos	201
Sigüientes pasos	208
Recursos adicionales	208
Mejora la visibilidad de sus costos en Amazon EKS.	208
Descripción general de	208
Beneficios en términos de costos	209
Recomendaciones de optimización de costos	209
Sigüientes pasos	213
Recursos adicionales	213
Redefinición de la plataforma de las aplicaciones de Windows con App2Container	214
Descripción general de	214
Beneficios en términos de costos	215
Recomendaciones de optimización de costos	216
Sigüientes pasos	216
Recursos adicionales	216
Almacenamiento	217
Amazon EBS	217
Migración de volúmenes de Amazon EBS de gp2 a gp3	218
Edición de instantáneas de Amazon EBS	222
Eliminación de volúmenes de Amazon EBS no conectados	225
Amazon FSx	229
Elección del almacenamiento de archivos SMB adecuado	230
Habilite la deduplicación de datos en Amazon FSx	235
Conozca la fragmentación de datos en Windows FSx File Server	237
Comprenda el uso del volumen del disco duro en Amazon FSx	242
Uso de una única zona de disponibilidad	245
AWS Storage Gateway	248
Puerta de enlace de archivo de Amazon S3	248
Amazon FSx File Gateway	248
Impacto del costo	249
Recomendaciones de optimización de costos	251
Recursos adicionales	254
Active Directory	255
Active Directory autoadministrado en Amazon EC2	255

Descripción general de	255
Impacto del costo	255
Recomendaciones de optimización de costos	256
Recursos adicionales	261
AWS Managed Microsoft AD	261
Descripción general de	261
Impacto del costo	261
Recomendaciones de optimización de costos	261
Recursos adicionales	263
Conector de AD	263
Descripción general de	263
Impacto del costo	264
Recomendaciones de optimización de costos	264
Recursos adicionales	265
.NET	266
Refactorización para usar las versiones más modernas de .NET y migración a Linux	267
Descripción general de	267
Impacto del costo	267
Recomendaciones de optimización de costos	268
Consideraciones y recursos adicionales	269
Almacenamiento de aplicaciones .NET en contenedores	270
Descripción general de	270
Impacto del costo	270
Recomendaciones de optimización de costos	272
Recursos adicionales	275
Uso de instancias y contenedores Graviton	275
Descripción general de	275
Impacto del costo	275
Recomendaciones de optimización de costos	277
Recursos adicionales	278
Escalado dinámico para aplicaciones estáticas de .NET Framework	279
Descripción general de	279
Impacto del costo	284
Recomendaciones de optimización de costos	285
Recursos adicionales	287
Uso del almacenamiento en caché para reducir la demanda de la base de datos	287

Descripción general de	287
Impacto del costo	287
Recomendaciones de optimización de costos	288
Recursos adicionales	295
Uso de .NET sin servidor	295
Descripción general de	295
Impacto del costo	296
Recomendaciones de optimización de costos	297
Recursos adicionales	301
Uso de bases de datos personalizadas	301
Descripción general de	301
Impacto del costo	305
Recomendaciones de optimización de costos	308
Recursos adicionales	309
Pasos a seguir a continuación	310
Historial de documentos	311
Glosario	312
#	312
A	313
B	316
C	318
D	321
E	326
F	328
G	330
H	331
I	332
L	335
M	336
O	340
P	343
Q	346
R	346
S	349
T	353
U	355

V	356
W	356
Z	357
.....	ccclix

Optimice los costos de las cargas de trabajo de Microsoft en AWS

Bill Pfeiffer, Chase Lindeman y Kevin Sookhan, Amazon Web Services (AWS)

Octubre de 2025 ([historial del documento](#))

Descripción general de

Esta guía proporciona recomendaciones, prácticas recomendadas y estrategias para ayudar a optimizar los costes de las cargas de trabajo de Microsoft en AWS. La guía también incluye conocimientos básicos sobre AWS , técnicas de optimización de costos y arquitecturas de referencia para que pueda crear y automatizar cargas de trabajo rentables y de alto rendimiento que cumplan sus objetivos empresariales. En conjunto, esta guía se denomina Optimización de los costos de Microsoft en AWS (MACO). La guía MACO fue desarrollada por expertos del sector y se basa en escenarios del mundo real.

En esta guía se tratan las siguientes cargas de trabajo de Microsoft:

- Windows en Amazon Elastic Compute Cloud (Amazon EC2)
- SQL Server
- Contenedores
- Almacenamiento
- Active Directory
- .NET

Público

Esta guía está destinada a arquitectos, ingenieros, administradores, directores CTOs, responsables de la toma de decisiones técnicas y AWS socios. Es útil, pero no necesario, tener experiencia previa y conocimientos básicos sobre AWS facturación, tecnologías de Microsoft y administración de AWS sistemas.

Cómo usar esta guía

Puede utilizar esta guía para planificar e implementar el proceso MACO en su transición a la nube. Le recomendamos que lea esta guía de principio a fin para obtener una comprensión completa de las opciones y enfoques para optimizar los costos de sus cargas de trabajo de Microsoft. AWS Consulte las secciones de las cargas de trabajo que le interesen en función de las necesidades de su organización:

- [Windows en Amazon EC2](#)
- [SQL Server](#)
- [Contenedores](#)
- [Almacenamiento](#)
- [Active Directory](#)
- [.NET](#)

Important

Las muestras de código de esta guía se proporcionan únicamente con fines ilustrativos. Se recomienda probar todo el código en un entorno de desarrollo antes de usarlo en un entorno de producción. Antes de implementar cualquier código, le recomendamos que lo pruebe en lotes pequeños y, a continuación, revise los cambios en los costos resultantes con [AWS Cost Explorer](#). Esto puede ser útil para solucionar casos límite y otras incidencias que pueden convertirse en problemas más adelante.

Important

Los ejemplos de precios de esta guía se basan en los precios vigentes en el momento de la publicación. Los precios están sujetos a cambios. Además, sus costos pueden variar en función de su Región de AWS Servicio de AWS cuotas y otros factores relacionados con su entorno de nube.

Resultados empresariales específicos

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

- Aprenda a usar Optimización y evaluación de licencias de AWS (OLA de AWS) para evaluar y optimizar sus entornos en las instalaciones y en la nube actuales, en función del uso real de los recursos, las licencias de terceros y las dependencias entre aplicaciones.
- Desarrolle una propuesta empresarial para optimizar los costos con Calculadora para la modernización de AWS para cargas de trabajo de Microsoft.
- Optimice los costos de sus cargas de trabajo específicas de Microsoft, como las cargas de trabajo de Windows en Amazon Elastic Compute Cloud (Amazon EC2), SQL Server, contenedores, almacenamiento, Active Directory y .NET.

Proceso de optimización de costos

El alcance, el calendario y la ruta específica de su migración a la nube dependen de los objetivos empresariales, los requisitos técnicos y otros factores. En esta sección se ofrece un ejemplo de un proceso de migración a la nube que se centra en la [administración financiera en la nube con AWS](#) y sigue las recomendaciones y prácticas recomendadas de MACO. Puede usar este ejemplo para aprender a diseñar un proceso de migración a la nube para las cargas de trabajo de Microsoft.

Las siguientes tareas generales ilustran la estrategia que una organización podría adoptar para implementar las recomendaciones y prácticas recomendadas de MACO:

- Establezca una estrategia de etiquetado y active etiquetas de asignación de costos definidas por el usuario. Para obtener más información, consulte el documento técnico de AWS [Prácticas recomendadas para el etiquetado de los recursos de AWS](#).
- Defina los presupuestos en función de las aplicaciones, los equipos o los departamentos. Para obtener más información, consulte [Managing your costs with AWS Budgets](#) en la Guía del usuario de Administración de facturación y costos de AWS.
- Lleve a cabo una evaluación de Optimización y evaluación de licencias de AWS (OLA de AWS) para comenzar a ahorrar ya mismo. Para obtener más información, consulte [Optimización y evaluación de licencias de AWS](#) en la documentación de AWS.
- Traiga su propia licencia (BYOL) para cargas de trabajo de Windows y SQL Server mediante Amazon Elastic Compute Cloud Dedicated Hosts. Para más información, consulte la sección [Uso de licencias propias para las cargas de trabajo de Windows y SQL Server](#) de esta guía.
- Optimice sus licencias de SQL Server en AWS. Para más información, consulte la sección [Explicación de las licencias de SQL Server](#) de esta guía.
- Seleccione el tipo de instancia adecuado para las cargas de trabajo de Windows. Para más información, consulte la sección [Selección del tipo de instancia adecuado para las cargas de trabajo de Windows](#) de esta guía.
- Seleccione el tipo de instancia adecuado para las cargas de trabajo de SQL. Para más información, consulte la sección [Selección de la instancia de EC2 adecuada para las cargas de trabajo de SQL Server](#) de esta guía.
- Migre Amazon Elastic Block Store (Amazon EBS) de gp2 a gp3. Para más información, consulte la sección [Migración de volúmenes de Amazon EBS de gp2 a gp3](#) de esta guía.

- Controle las cargas de trabajo con Programador de instancias de EC2 en AWS. Para más información, consulte la sección [Automatización de la programación de parada e inicio](#) de esta guía.
- Elimine los costos de SQL Server derivados de las cargas de trabajo no relacionadas con la producción con SQL Server Developer Edition. Para más información, consulte la sección [Evaluación de la edición Developer de SQL Server](#) de esta guía.
- Utilice una única zona de disponibilidad para Amazon FSx para Windows File Server para cargas de trabajo de desarrollo y pruebas. Para más información, consulte la sección [Uso de una única zona de disponibilidad](#) de esta guía.
- Dimensione correctamente sus cargas de trabajo de Windows con AWS Compute Optimizer. Para más información, consulte la sección [Dimensionamiento correcto de las cargas de trabajo de Windows](#) de esta guía.
- Optimice el gasto en Windows en Amazon EC2 con Savings Plans. Para más información, consulte la sección [Optimización del gasto de Windows en Amazon EC2](#) de esta guía.
- Active la deduplicación de datos en FSx para Windows File Server Para más información, consulte la sección [Habilite la deduplicación de datos en Amazon FSx](#) de esta guía.
- Utilice las particiones de datos para sistemas de archivos en FSx para Windows File Server. Para más información, consulte la sección [Conozca la fragmentación de datos en Windows FSx File Server](#) de esta guía.
- Optimice sus estrategias de copia de seguridad de SQL Server. Para más información, consulte la sección [Optimización de las estrategias de copia de seguridad de SQL Server](#) de esta guía.
- Configure las aplicaciones estáticas de .NET Framework para que admitan el escalado dinámico. Para obtener más información, consulte la sección [Escalado dinámico para aplicaciones estáticas de .NET Framework](#) de esta guía.
- Utilice microservicios .NET sin servidor. Para más información, consulte la sección [Uso de .NET sin servidor](#) de esta guía.
- Almacene sus aplicaciones de Windows en contenedores. Para más información, consulte la sección [Almacenamiento de aplicaciones .NET en contenedores](#) de esta guía.
- Use [AWS Compute Optimizer](#) para dimensionar correctamente los contenedores de Windows que estén en marcha en AWS Fargate para Amazon Elastic Container Service (Amazon ECS). Para más información, consulte la sección [Activación de Compute Optimizer](#) de esta guía.
- Refactorice para usar las versiones más modernas de .NET y migre a Linux. Para más información, consulte la sección [Refactorización para usar las versiones más modernas de .NET y migración a Linux](#) de esta guía.

- Aproveche las instancias y los contenedores Graviton. Para más información, consulte la sección [Uso de instancias y contenedores Graviton](#) de esta guía.
- Modernice las bases de datos de SQL Server. Para más información, consulte la sección [Modernización de las bases de datos de SQL Server](#) de esta guía.
- Diseñe la infraestructura de Active Directory. Para más información, consulte la sección [Active Directory](#) de esta guía.

Para obtener más información sobre la experiencia de un cliente centrada en la administración financiera en la nube con AWS, consulte el documento técnico de AWS [Cloud Financial Management capability](#).

Principales recomendaciones para optimizar los costos

Descripción general

La optimización de costos es uno de los pilares de Marco de [AWS Well-Architected](#) y desempeña un papel fundamental en sus planes de migración a la nube. Encontrará recomendaciones para optimizar costos en esta guía, pero en esta sección se describen las recomendaciones que tienen un mayor impacto. Puede implementar estas recomendaciones rápidamente y tendrán un impacto significativo en su organización. Si aplica estas recomendaciones, podrá sentar las bases de todo el proceso de optimización de costos.

Recomendaciones principales

En la siguiente tabla, se enumeran las principales recomendaciones en cuanto a optimizaciones de costos de mayor impacto. La columna Dificultad de implementación clasifica cada optimización en función de una escala que va desde lo que es más fácil de implementar (1) hasta lo que es más difícil de implementar (5). La columna Ahorros estimados muestra una estimación basada en porcentajes de cuánto puede ahorrar su organización por cada optimización recomendada.

Optimizaciones	Dificultad de implementación	Ahorro estimado
Dimensionamiento correcto de las cargas de trabajo de Windows	3	25%
Uso de licencias propias para las cargas de trabajo de Windows y SQL Server	3	30%
Evaluación de la edición Developer de SQL Server	2	20%
Explicación de las licencias de SQL Server	2	Hasta un 50 %

Optimizaciones	Dificultad de implementación	Ahorro estimado
Automatización de la programación de parada e inicio	3	Hasta un 40 %
Selección del tipo de instancia adecuado para las cargas de trabajo de Windows	1	10-30 %
Refactorización para usar las versiones más modernas de .NET y migración a Linux	5	10-20 %
Optimización del gasto de Windows en Amazon EC2	3	Hasta un 20-40 %
Migración de volúmenes de Amazon EBS de gp2 a gp3	4	Hasta un 20 %

Important

Los ahorros estimados de la tabla anterior se aplican a cada ámbito técnico individual, no al gasto total en AWS de una cuenta. Por ejemplo, puede implementar Programador de instancias en diversos tipos y tamaños de entornos, lo que puede alterar los posibles ahorros. Las estimaciones se aplican específicamente a los costos de las instancias de Amazon EC2 y no implican ningún ahorro global para otros Servicios de AWS. Estas estimaciones se proporcionan como un indicador, no como una garantía.

Los expertos en MACO están disponibles para profundizar sobre las optimizaciones de costos. Para programar una reunión en la que analizar en profundidad su caso de uso, contacte con su equipo de cuentas o envíe un correo electrónico a optimize-microsoft@amazon.com.

Optimización y evaluación de licencias de AWS

Descripción general

Con [Optimización y evaluación de licencias de AWS \(OLA de AWS\)](#), puede evaluar y optimizar sus entornos en las instalaciones y en la nube actuales, en función del uso real de los recursos, las licencias de terceros y las dependencias entre aplicaciones. Puede usar OLA de AWS para ayudar a su organización a desarrollar una estrategia de migración y licencias que permita ahorrar costos al migrar a AWS o evaluar las cargas de trabajo de Microsoft existentes en AWS. OLA de AWS también puede ser de ayuda para lograr lo siguiente:

- Comprender las implementaciones, el rendimiento de las aplicaciones y los contratos existentes.
- Dimensionar correctamente los recursos.
- Desarrollar una hoja de ruta para la Nube de AWS.
- Reducir o eliminar los costos gracias al uso de las inversiones existentes y pagar solo por lo que utilice.

Le recomendamos que haga de OLA de AWS el primer paso en su [proceso de optimización de costos](#). Puede usar la AWS Partner Network para completar una evaluación de OLA de AWS. Esta le servirá para recopilar datos de evaluación y obtener recomendaciones para optimizar los costos asociados a las licencias e instancias.

En el diagrama siguiente se proporciona una descripción general del proceso de evaluación.



Opciones de evaluación

Puede elegir entre dos opciones de OLA de AWS para sus cargas de trabajo de Microsoft en AWS:

- **Versión básica:** en este caso de uso, todas sus cargas de trabajo están en VMware. Puede proporcionar a AWS un resultado de [RVTools](#). Luego, AWS puede ofrecer un tiempo de entrega de 1 a 5 días. Esta estrategia utiliza información puntual extraída directamente de VMware vCenter para desarrollar recomendaciones de tamaño y ofrecer opciones de precios bajo demanda..
- **Versión completa:** en este caso de uso, tiene un entorno mixto que se pone en marcha en diferentes proveedores de servicios en la nube, servidores físicos y servidores virtuales. AWS utiliza agentes del sistema operativo para recopilar datos de uso de 14 a 30 días. Esto le permite a AWS tomar decisiones fundamentadas sobre el tamaño de las instancias en función de los patrones de uso de las aplicaciones. AWS utiliza varias herramientas de terceros, como Cloudamize, para completar el análisis. AWS usa la AWS Partner Network para poder llevar a cabo la evaluación final del costo total de propiedad (TCO) con múltiples opciones de precios que tienen en cuenta los modelos de precios y las diferentes arquitecturas.

Evaluación completa

La evaluación completa de OLA de AWS comienza con una llamada telefónica de una hora. Durante esta llamada, AWS le ofrece su ayuda para dar con la infraestructura de AWS más óptima para respaldar su migración, elegir un método de recopilación de datos y establecer un cronograma para su finalización. La implementación de herramientas de detección en su organización depende del método de recopilación de datos, del tamaño de la organización y de las herramientas que utilice la organización para administrar su flota de servidores. La recopilación de datos de uso suele tardar dos semanas.

El proceso completo de evaluación de OLA de AWS dura entre 30 y 45 días y consta de las siguientes fases:

- Delimitación del alcance de las cargas de trabajo
- Recopilación de datos
- Análisis de datos
- Planificación de los pasos siguientes

Delimitación del alcance de las cargas de trabajo

En primer lugar, AWS trabaja con usted y su equipo para determinar el alcance de la evaluación. Por lo general, esta se divide por tipo de entorno (por ejemplo, de no producción y de producción). El alcance incluye la ubicación de las cargas de trabajo. Pueden ser cargas de trabajo que va a migrar a AWS, cargas de trabajo que ya están en marcha en AWS (por ejemplo, OLA de AWS para Amazon EC2) o cargas de trabajo que están en marcha en otros proveedores de servicios en la nube.

Recopilación de datos

A continuación, AWS implementa herramientas para poder detectar recursos y recopilar datos de rendimiento de sus servidores. Estas herramientas ofrecen cuatro opciones de implementación:

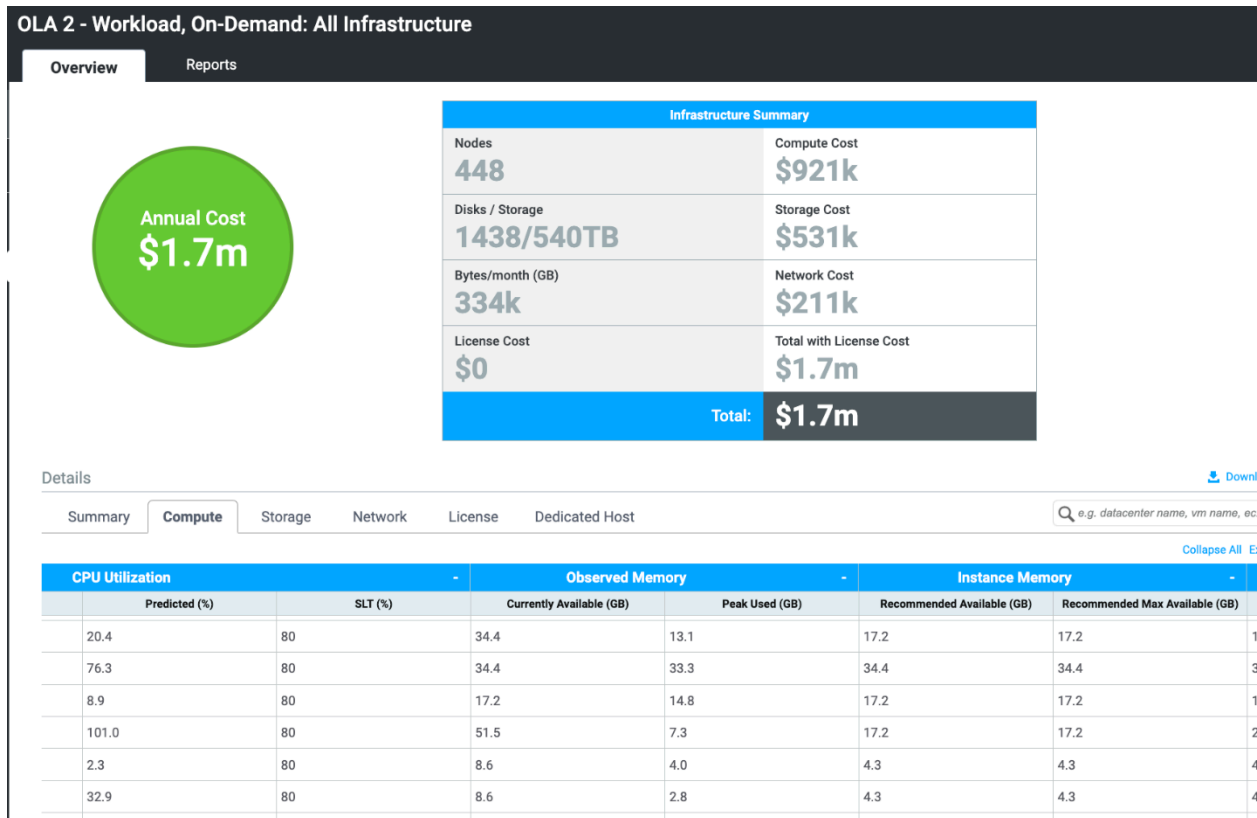
- Herramientas que pueden consultar el hipervisor (solo requieren credenciales de VMware vCenter o Hyper-V).
- Agentes que se pueden implementar en máquinas físicas o virtuales.
- Detección sin agente mediante SSH, Windows Remote Management (WinRM) o Windows Management Instrumentation (WMI) según el entorno y el sistema operativo.
- Recopilación y análisis de datos de archivos sin formato.

Para la implementación de sus herramientas, puede combinar cada opción y unificar los resultados. Es fundamental asegurarse de que la opción que elija no suponga una carga excesiva para sus recursos de TI. AWS se esfuerza por hacer que el proceso de evaluación sea lo más completo posible. Más allá de una breve llamada telefónica para ayudar con la configuración, el equipo de OLA de AWS y los arquitectos de soluciones especializados de Microsoft prepararán el análisis del costo total de propiedad (TCO) y las recomendaciones para su revisión.

La recopilación de datos suele tardar de dos a tres semanas cuando se analizan la utilización de la CPU, la utilización de la RAM, el rendimiento del almacenamiento, las IOPS y el rendimiento de la red. Lo ideal es que esta recopilación se lleve a cabo durante las horas de mayor actividad del mes laboral (por ejemplo, durante la elaboración de los informes financieros al final del mes). AWS quiere registrar los picos de uso porque de este modo se obtienen buenos ejemplos estadísticos de cuál debería ser el tamaño correcto para la instancia de AWS y, al mismo tiempo, se garantiza que el rendimiento supere lo que está disponible en las instalaciones. AWS combina las métricas de uso con las heurísticas de rendimiento de varias generaciones de procesadores para determinar exactamente la cantidad de CPU y RAM que requiere una carga de trabajo determinada. Estos

objetivos suelen ser inferiores a los que se asignan en las instalaciones. Esto no solo reduce el costo de la computación en función del tamaño de la instancia, sino que también optimiza los costos de las licencias.

La siguiente vista del panel muestra un ejemplo de los costos derivados de la infraestructura que se pueden registrar en una evaluación.

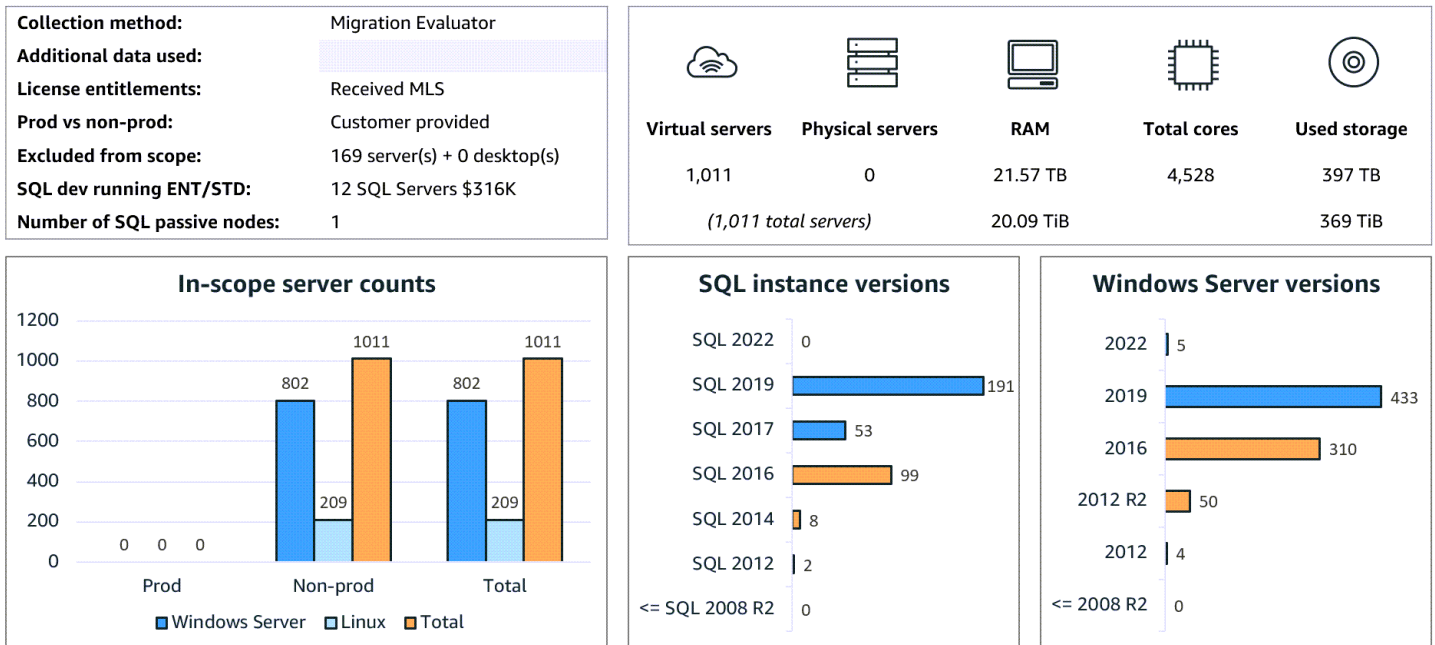


Análisis de datos

AWS ofrece una presentación resumida una vez finalizada la recopilación de datos. AWS revisa los datos, resume los resultados y, a continuación, hace recomendaciones para el uso en las instalaciones y la migración a la nube. Se pueden reducir los costos asociados a la computación y a las licencias analizando las oportunidades de unificación, las ganancias por elasticidad (la posibilidad de desactivar o ajustar las cargas de trabajo según la temporada) y las oportunidades de elección del SKU correcto (por ejemplo, aunque se utilice SQL Server Enterprise, los recursos necesarios y el uso de características indican que SQL Server Standard es suficiente). En el caso de productos como SQL Server, cuya licencia depende del núcleo, suele tener sentido desde el punto de vista económico colocar las cargas de trabajo en una instancia de computación más cara. Es decir, si el perfil de la CPU y la proporción de RAM por vCPU generan un efecto neto que reduce la cantidad de

núcleos con licencia tanto en los casos de uso de licencia incluida como en los de traiga su propia licencia (BYOL)

A continuación se muestra un ejemplo de análisis basado en los datos recopilados por la evaluación.



Los escenarios de optimización habituales incluyen la identificación tanto de las oportunidades de optimización de los recursos de AWS como de los ahorros en licencias de terceros.

Ejemplos de oportunidades de optimización de los recursos de AWS:

- Evitar el sobreaprovisionamiento para los picos de uso.
- Evitar asignar demasiados recursos o no usarlos lo suficiente.
- Dimensionar las instancias de forma adecuada y migrar a las generaciones más recientes de instancias de EC2.
- Ahorrar en los costos derivados de las operaciones migrando a bases de datos administradas.

Ejemplos de ahorros en licencias de terceros:

- Reducir los núcleos necesarios para poner en marcha la misma carga de trabajo.
- Eliminar la edición SQL Server Enterprise y los paquetes de complementos innecesarios.
- Eliminar los servidores zombis y sustituir el hardware obsoleto.
- Usar BYOL y las opciones con licencia incluida para reducir futuros contratos comerciales.

- Modernizar con soluciones de código abierto y nativas en la nube.

Planificación de los pasos siguientes

Por último, AWS utiliza los datos de rendimiento recopilados para estimar el tamaño y el costo específicos de la carga de trabajo. AWS también puede analizar en conjunto su entorno específico y proporcionar un análisis cuantitativo. Esto puede ser de ayuda para determinar si la mejor opción es una actualización en las instalaciones o una migración a AWS. Para crear un modelo empresarial económico basado en la nube, utilice el resumen del análisis del TCO (como se muestra en el siguiente ejemplo) que se proporciona al final de una evaluación de OLA de AWS.

	Option 1: Amazon EC2 shared	Option 1a: Amazon EC2 shared + power management	Option 2: Amazon EC2 mixed	Option 2a: Amazon EC2 mixed + power management
<i>Option details: compute</i>	100% Reserved Instances (RIs)	RIs + on-demand power management	100% RIs	RIs + on-demand power management
<i>Option details: Microsoft licenses</i>	WS LI and SQL BYOL	WS LI and SQL BYOL	WS BYOL or LI+SQL BYOL	WS BYOL or LI+SQL BYOL
Compute costs¹				
Year 1 compute cost	\$414,546	\$482,623	\$504,019	\$513,941
Year 1 vendor license included cost	\$392,858	\$244,415	\$9,804	\$4,783
	\$807,404	\$727,038	\$513,823	\$518,724
<i>Total compute savings in year 1, compared to Option 1</i>	—	10% (\$80,366)	36% (\$293,581)	36% (\$288,680)
Storage and networking costs²				
Annual estimated storage cost	\$336,494	\$336,494	\$336,494	\$336,494
Annual estimated networking cost	\$41,455	\$41,455	\$41,455	\$41,455
	\$377,949	\$377,949	\$377,949	\$377,949
Microsoft license costs**				
WS/CIS annual Software Assurance (SA) + current SPLA/Subs cost	\$0	\$0	\$0	\$0
WS/CIS license + SA + SPLA/Subs true-up cost	\$0	\$0	\$0	\$0
SQL annual SA + current SPLA/Subs cost	\$0	\$0	\$0	\$0
SQL license SA + current SPLA/Subs true-up cost	\$0	\$0	\$0	\$0
	\$0	\$0	\$0	\$0
Total estimated costs	\$1,185,353	\$1,104,987	\$891,772	\$896,673
<i>Annual TCO savings in year 1, compared to Option 1</i>	—	7% (\$80,366)	25% (\$293,581)	24% (\$288,680)

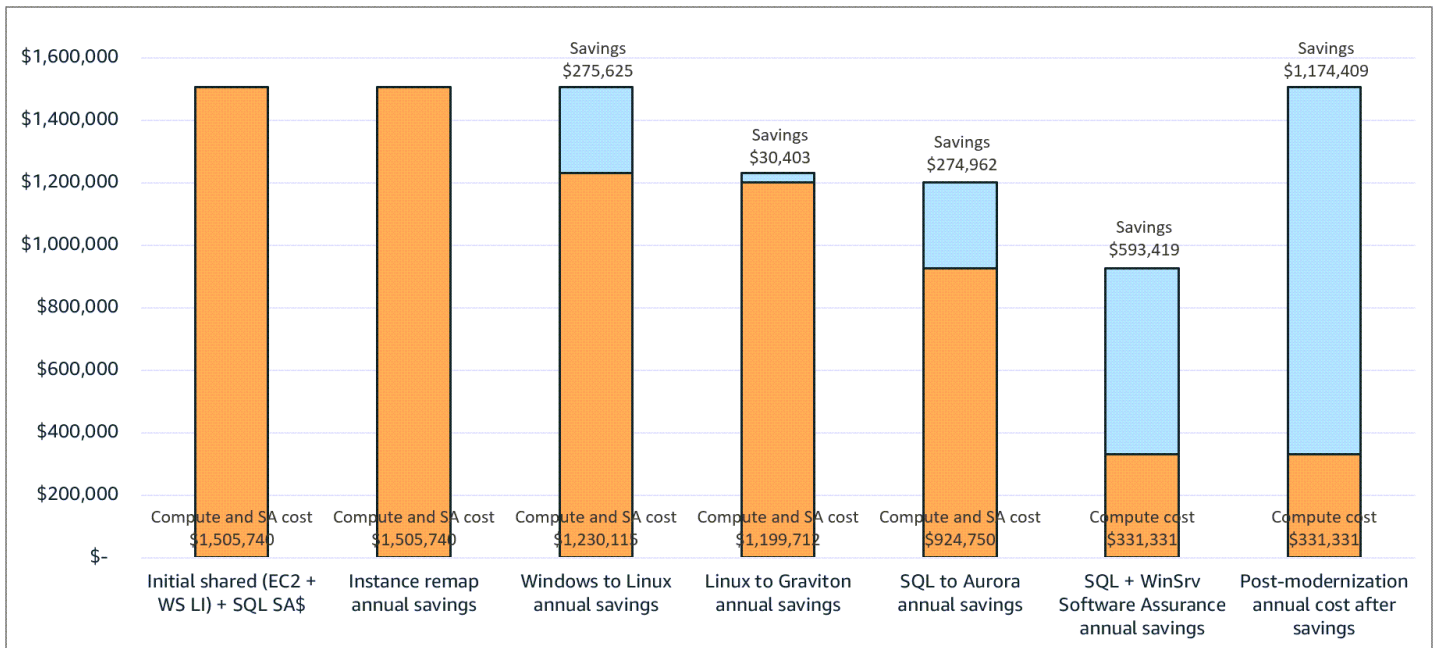
¹ Pricing model used: 3-year, no upfront RI

² Software Assurance and true-up costs provided by Microsoft

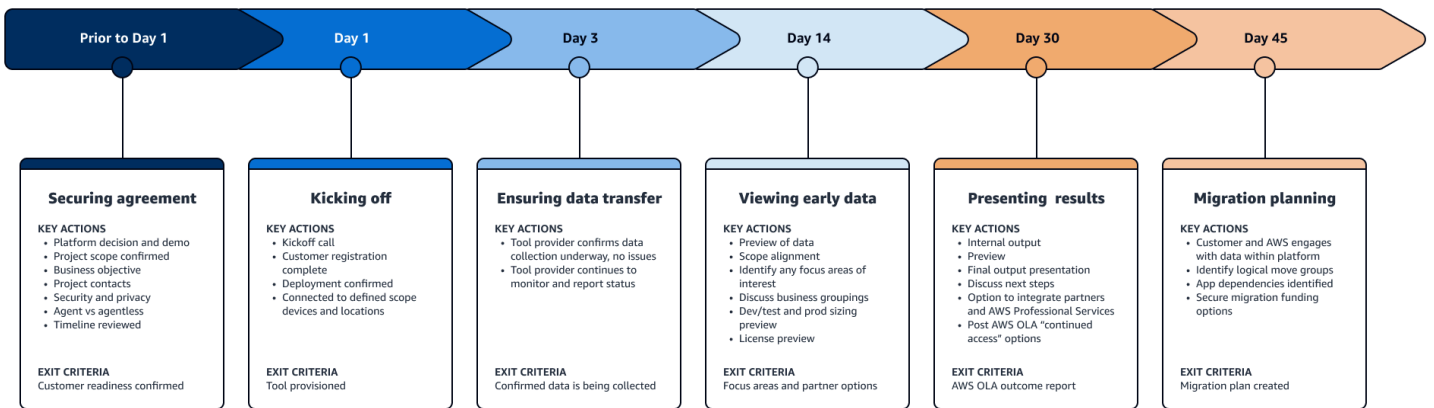
Una evaluación de OLA de AWS también proporciona información sobre el impacto que la modernización puede tener en sus cargas de trabajo actuales mediante sugerencias como las siguientes:

- Pasar a un sistema operativo Linux.
- Agregar compatibilidad con aplicaciones para procesadores ARM (AWS Graviton).
- Mover las cargas de trabajo de SQL Server a Amazon Aurora.
- Eliminar Software Assurance migrando las cargas de trabajo de Windows y SQL Server a tecnologías de código abierto.

El siguiente diagrama muestra la rentabilidad que se puede lograr mediante técnicas de modernización, como pasar de Windows a Linux o de SQL Server a Aurora.



El proceso completo de evaluación de OLA de AWS tarda aproximadamente 45 días. En el siguiente diagrama se muestra un cronograma de ejemplo.



Si tiene un entorno exclusivamente de VMware y puede proporcionar resultados de RVTools, puede reducir este plazo a una semana laboral. Además, AWS puede analizar un archivo sin formato que incluya datos de activos y uso, como el promedio de la CPU, el pico de la CPU, el promedio de la RAM y el pico de la RAM.

Impacto de la evaluación

Los clientes, por lo general, experimentan una disminución de costos de entre el 20 % y el 30 % tras llevar a cabo el dimensionamiento correcto de los recursos. Un buen dimensionamiento asigna la carga de trabajo de origen con las instancias de AWS con el tamaño más adecuado en función de los datos de uso. Estos ajustes de tamaño no solo reducen el costo mensual del entorno de AWS, sino que suelen resultar en ahorros en otros departamentos de la organización. Por ejemplo, un ahorro de entre el 20 % y el 30 % en licencias de Windows o SQL Server puede reducir el monto de la próxima compensación a Microsoft o liberar licencias para otras aplicaciones empresariales. La unificación y el dimensionamiento adecuado de las cargas de trabajo de SQL Server suelen ser las áreas en las que se obtienen los beneficios financieros más importantes.

AWS puede ser de ayuda a la hora clasificar los sistemas en buckets de modernización. Algunos sistemas son antiguos y no son viables desde el punto de vista financiero, mientras que otros pueden modernizarse y convertirse en contenedores o aplicaciones sin servidor, lo que supone los ahorros más significativos. La conversación con el equipo de AWS abarca desde temas generalizados sobre las posibilidades de la nube hasta conversaciones más específicas sobre cómo y por qué se deben modernizar cargas de trabajo específicas. AWS también le permite explorar posibles oportunidades de innovación.

Pasos a seguir a continuación

Si está empezando su proceso de optimización de costos para las cargas de trabajo de Microsoft que se ponen en marcha en entornos en las instalaciones o en AWS, contacte con el equipo de cuentas de AWS y solicite una evaluación de OLA de AWS. Los miembros del equipo de AWS pueden responder a sus preguntas y brindarle su ayuda para decidir si una evaluación de OLA de AWS es, en última instancia, la mejor opción para usted y su organización. Como alternativa, puede [solicitar una evaluación de OLA de AWS por Internet](#).

Recursos adicionales

- [Optimización y evaluación de licencias de AWS](#) (documentación de AWS)
- [AWS re:Invent 2022 - How to save costs and optimize Microsoft workloads on AWS \(ENT205\)](#) (YouTube)

Windows en Amazon EC2

[Amazon Elastic Compute Cloud \(Amazon EC2\)](#) es una plataforma de computación en la nube altamente flexible y escalable que es ideal para poner en marcha sus cargas de trabajo de Windows. Puede usar Amazon EC2 para implementar, administrar y escalar sus cargas de trabajo de Windows Server en la infraestructura segura, fiable, de alta disponibilidad y adaptable de la Nube de AWS. Tenga en cuenta las siguientes ventajas clave de poner en marcha cargas de trabajo de Windows en Amazon EC2:

- **Escalabilidad:** Amazon EC2 le permite escalar fácilmente sus cargas de trabajo de Windows para adaptarlas a los requisitos cambiantes. Puede crear nuevas instancias de EC2 con rapidez para hacer frente al aumento de la demanda y, con la misma facilidad, cancelar las instancias cuando ya no las necesite. Solo pagará por los recursos que utilice.
- **Flexibilidad:** Windows en Amazon EC2 admite una amplia gama de tipos de instancias que están diseñadas para satisfacer diversos requisitos de carga de trabajo, desde instancias de uso general hasta instancias optimizadas para memoria o procesamiento. Esta flexibilidad garantiza que pueda elegir el mejor tipo de instancia para sus aplicaciones específicas basadas en Windows, lo que maximiza el rendimiento y minimiza los costos.
- **Seguridad:** AWS proporciona varios niveles de seguridad para sus cargas de trabajo de Windows, incluidos firewalls de red, cifrado de datos y controles de acceso seguro. Esto significa que puede confiar en que sus aplicaciones y datos están protegidos y, al mismo tiempo, tener el control total de sus ajustes y configuraciones de seguridad.
- **Rentabilidad:** el modelo de pay-as-you-go precios le permite pagar solo por los recursos que utilice, lo que elimina la necesidad de realizar inversiones iniciales en hardware y software. Este modelo también le permite optimizar sus costos, reducir los gastos en capital y aumentar la eficiencia operativa. Es un modelo de precios ideal para empresas de todos los tamaños.

Esta sección de la guía abarca los siguientes temas:

- [Automatización de la programación de parada e inicio](#)
- [Dimensionamiento correcto de las cargas de trabajo de Windows](#)
- [Selección del tipo de instancia adecuado para las cargas de trabajo de Windows](#)
- [Uso de licencias propias para las cargas de trabajo de Windows y SQL Server](#)
- [Optimización del gasto de Windows en Amazon EC2](#)
- [Supervise los costes mediante herramientas AWS](#)

Automatización de la programación de parada e inicio

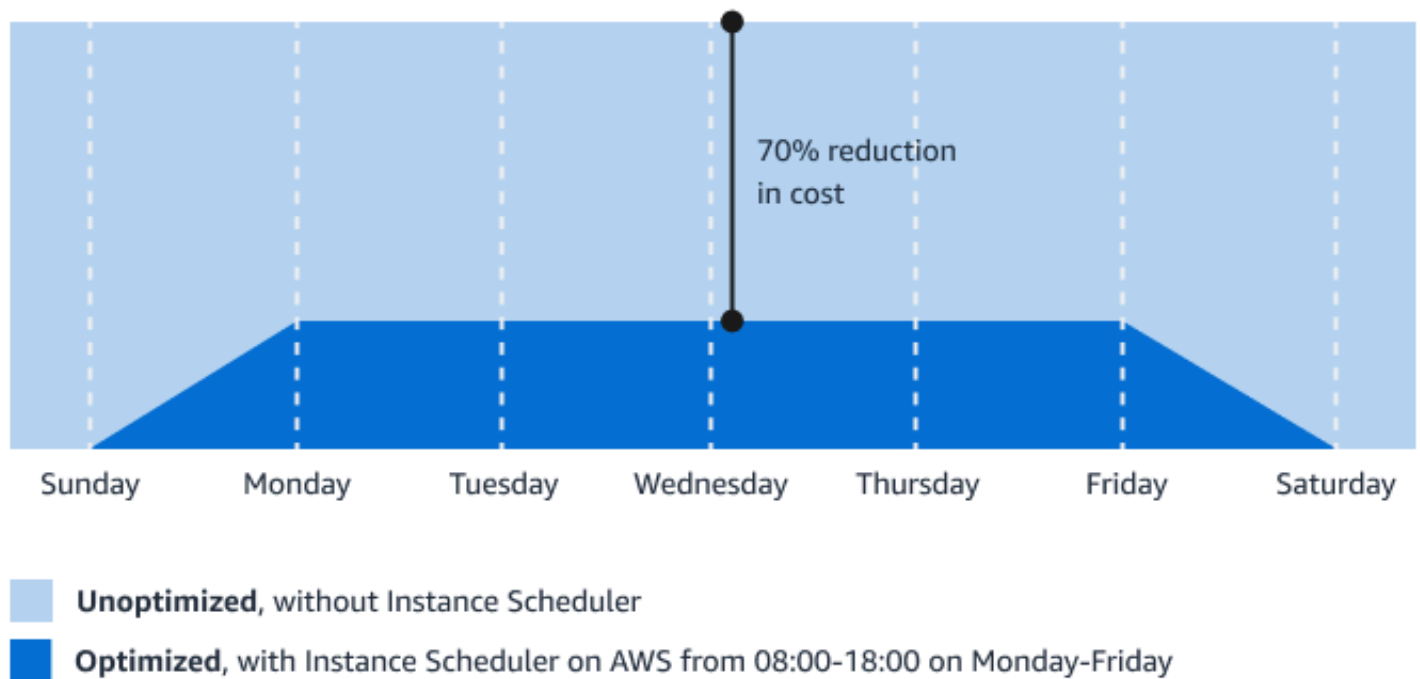
Descripción general de

[Programador de instancias en AWS](#) puede ser útil para reducir los costos operativos al automatizar el inicio y la detención de las instancias de [Amazon EC2](#) y [Amazon Relational Database Service \(Amazon RDS\)](#). Si deja todas sus instancias en funcionamiento a pleno rendimiento de forma continua, podría terminar pagando por los recursos que no se utilizan. El programador de instancias activado AWS le permite desactivar las instancias en momentos en que no se necesitan, como fuera del horario laboral, los fines de semana u otros períodos en los que el uso es bajo. Esto puede suponer un importante ahorro en costos a largo plazo.

El programador de instancias activado AWS también ofrece la programación de instancias entre cuentas, el etiquetado automatizado y la posibilidad de configurar programas o períodos mediante una interfaz de línea de comandos o el período de mantenimiento. [AWS Systems Manager](#) Gracias a estas características, puede gestionar sus instancias de forma más eficaz y precisa, realizar un seguimiento y distribuir los costos entre distintos proyectos o equipos.

Casos prácticos

Pensemos en el ejemplo de una empresa que utiliza Instance Scheduler AWS en un entorno de producción para detener automáticamente las instancias fuera del horario laboral todos los días. Si esta empresa deja todas sus instancias en ejecución a pleno rendimiento, puede ahorrar hasta un 70 % en los costos de las instancias que solo son necesarias durante el horario laboral habitual. En el siguiente gráfico se muestra cómo se reduce la utilización semanal de 168 horas a 50 horas.

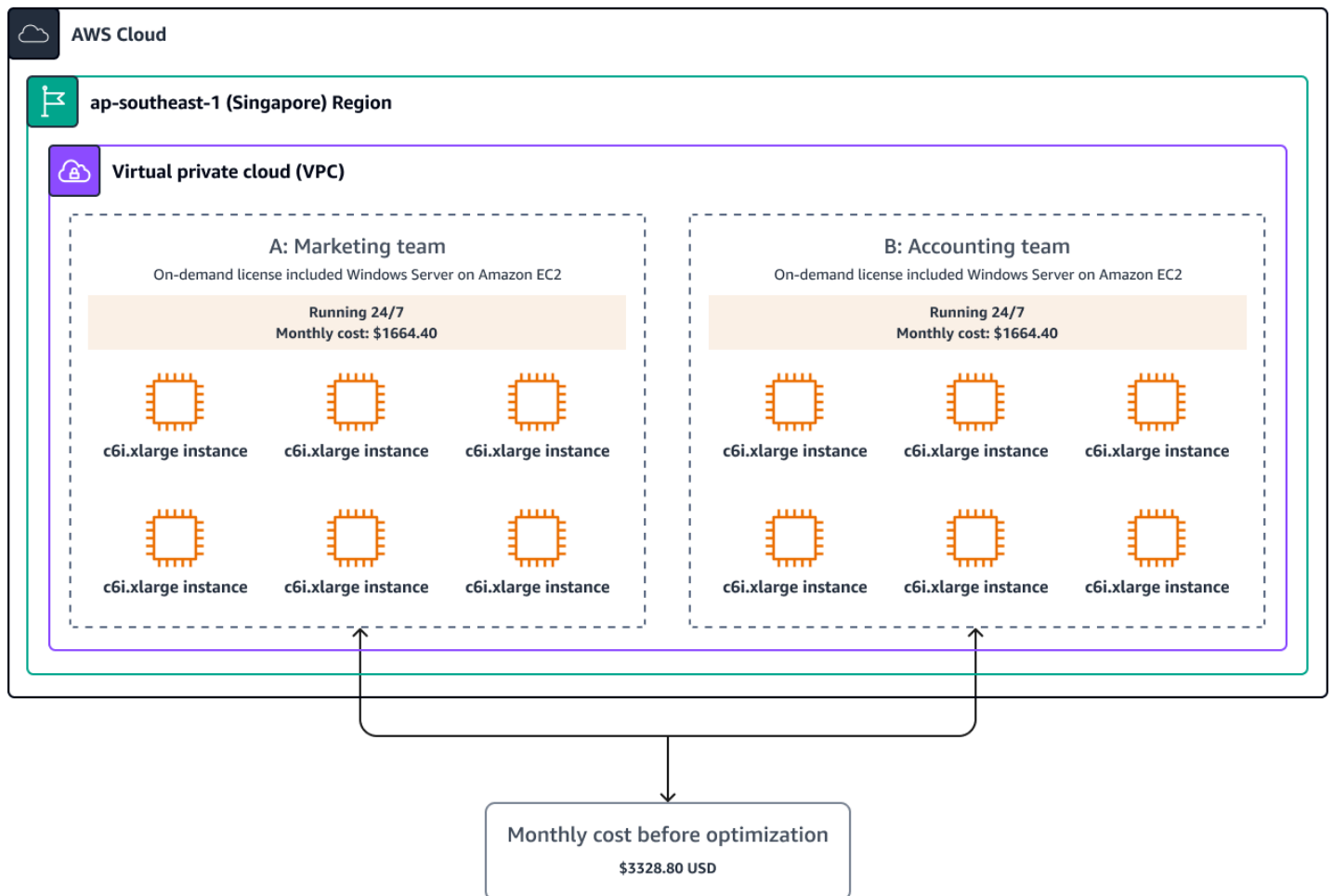


Vayamos con otro ejemplo. La empresa eléctrica Jamaica Public Service Company Limited (JPS) migró su base de datos a Amazon RDS. Ahora, JPS usa Amazon EC2 para alojar servicios de API y poner en marcha otras aplicaciones. Para JPS, Instance Scheduler on AWS se convirtió en la herramienta clave para gestionar los entornos no productivos. JPS utilizó el programador de instancias AWS para reducir los costos de desarrollo y administrar las instancias de EC2 en función de las necesidades del equipo y los horarios de trabajo. Esto ayudó a JPS a reducir los costos en un 40 %. Para obtener más información, consulte el AWS caso práctico [El servicio público de Jamaica migra eficientemente a la nube y reduce los costos en un 40% con AWS Instance Scheduler](#).

Escenario de optimización de costos

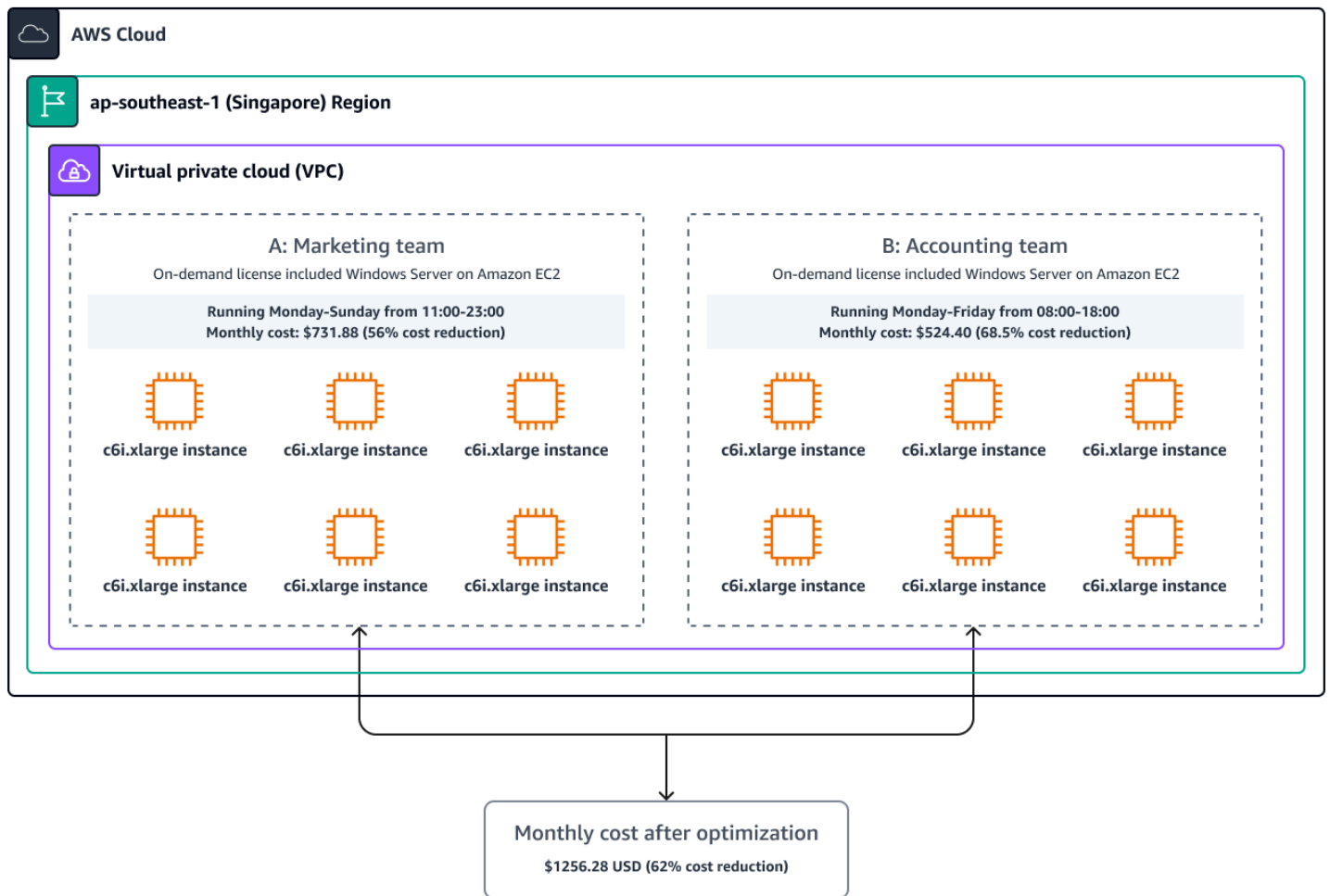
El siguiente escenario de ejemplo ayuda a ilustrar las ventajas económicas de usar Instance Scheduler en AWS. En este escenario, una importante empresa minorista de Singapur implementa dos entornos Windows en Amazon EC2. El equipo de marketing utiliza el primer entorno, denominado “carga de trabajo A”, para analizar en tiempo real las transacciones en las tiendas mientras las tiendas están abiertas. El segundo entorno, denominado “carga de trabajo B”, está reservado para el equipo de contabilidad, que solo trabaja durante el horario laboral habitual. El cronograma operativo actual de ambos entornos (24 horas al día, 7 días a la semana) no es ideal dados los patrones de uso actuales y requiere una optimización para reducir los costos operativos de la empresa.

En el siguiente diagrama, se muestra el costo mensual antes de la optimización.



Por ejemplo, hay 31 días en el mes de marzo, de los cuales, 23 son días laborables. Si el equipo de marketing usa Instance Scheduler activado AWS y opera sus instancias solo cuando es necesario (es decir, 321 horas al mes en lugar de 730 horas al mes), podría ahorrarse 932,52\$ al mes. Esto equivale a una reducción del 56 % en los costos operativos. El equipo de contabilidad también puede disfrutar de importantes ventajas, ya que el tiempo de uso de las instancias se reduce de 730 horas al mes a 230 horas. Esto se traduce en una reducción de 1140 USD, es decir, un 68,5 %. La empresa podría ahorrar un total combinado de 2072,52 USD al mes (lo que equivale a una reducción del 62 %), o 24 870,24 USD al año.

En el siguiente diagrama, se muestra el costo mensual después de la optimización.



Note

Los precios de este ejemplo se determinaron utilizando [Calculadora de precios de AWS](#) en marzo de 2023.

Recomendaciones de optimización de costos

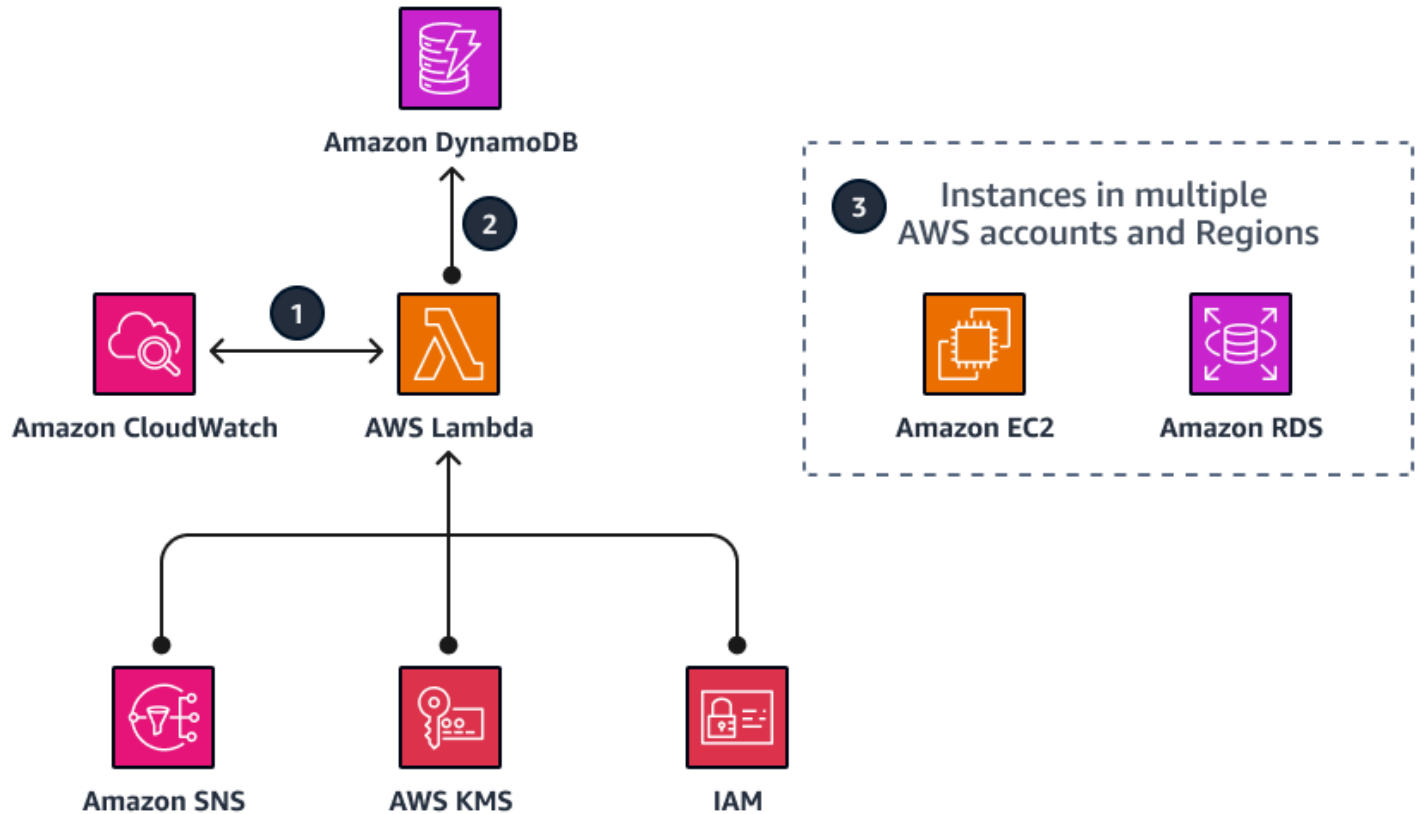
En esta sección se explica cómo implementar y configurar Programador de instancias en AWS basándonos en el escenario de ejemplo descrito en la sección anterior, Escenario de optimización de costos. Le recomendamos que siga los siguientes pasos para optimizar sus costes mediante el uso del Programador de instancias en: AWS

1. Lanzamiento de la pila de Programador de instancias
2. Configuración de periodos

3. Configuración de programaciones

4. Etiqueta de instancias .

El siguiente diagrama de arquitectura muestra lo que crea en la pila de Nube de AWS Instance Scheduler.



En el diagrama, se muestran los siguientes pasos del flujo de trabajo.

1. Una AWS CloudFormation plantilla configura un CloudWatch evento de Amazon en un intervalo que tú definas. Este evento invoca una AWS Lambda función. Durante la configuración, se definen las cuentas Regiones de AWS y. También debe definir una etiqueta personalizada que Instance Scheduler AWS utiliza para asociar los cronogramas con las instancias de Amazon EC2, las instancias de Amazon RDS y los clústeres aplicables.
2. Los valores de configuración de la programación se almacenan en Amazon DynamoDB y la función de Lambda los recupera cada vez que se pone en marcha. A continuación, puede aplicar la etiqueta personalizada a las instancias correspondientes.
3. Durante la configuración inicial de Programador de instancias, debe definir una clave de etiqueta para identificar las instancias de Amazon EC2 y Amazon RDS correspondientes. Al crear un

cronograma, el nombre que especifique se usa como el valor de etiqueta que identifica el cronograma que desea aplicar al recurso etiquetado.

Lanzamiento de la pila de Programador de instancias

En esta sección, se muestra cómo lanzar la CloudFormation pila del programador de instancias.
AWS

Note

Eres responsable del coste de lo que Servicios de AWS se utilice mientras se ejecuta Instance Scheduler en él. AWS En enero de 2023, el costo de poner en marcha esta solución con la configuración predeterminada en la región us-east-1 era de aproximadamente 9,90 USD al mes en el caso de los cargos de Lambda, o menos si tiene un crédito de uso mensual del nivel gratuito de Lambda. Para obtener más información, consulte la sección de costos de la [guía de AWS implementación del programador de instancias](#) en la biblioteca de AWS soluciones.

Para lanzar la pila de Programador de instancias, siga los pasos que se describen a continuación.

1. Inicie sesión [Consola de administración de AWS](#) y elija [Launch solution](#) (plantilla descargable) para lanzar la `instance-scheduler-on-aws.template` CloudFormation plantilla.

Note

También puede [descargar la plantilla](#) para usarla como punto de partida para su propia implementación.

2. La plantilla se lanza en la región Este de EE. UU. (Norte de Virginia) de forma predeterminada. Para lanzar Programador de instancias en una región diferente, utilice el selector de regiones de la barra de navegación de la consola.

Note

(En este ejemplo se utiliza la región Asia-Pacífico (Singapur).

3. En la página Crear pila, en la sección Requisito previo: preparar la plantilla, compruebe que esté seleccionada la opción La plantilla está lista. En la sección Origen de la plantilla, compruebe que esté seleccionada la opción URL de Amazon S3.
4. Verifique que la URL de la plantilla correcta aparezca en el cuadro de texto URL de Amazon S3 y seleccione Siguiente.
5. En la página Especificar los detalles de la pila, especifique un nombre para la pila. Para obtener información sobre las limitaciones de nombres de caracteres, consulte [los límites de IAM y STS](#) en la AWS Identity and Access Management documentación (IAM). El nombre de la pila para el ejemplo de esta guía es MyInstanceScheduler.

Note

El nombre de la pila no puede tener más de 28 caracteres.

6. En Parámetros, revise los parámetros de la plantilla y modifíquelos según sea necesario.
7. Elija Siguiente. En la página Configurar opciones de pila, elija Siguiente.
8. En la página Revisar, revise y confirme la configuración. Seleccione la casilla para aceptar que la plantilla creará recursos de IAM.
9. Elija Crear para implementar la pila.

Configuración de periodos

Tras implementar la CloudFormation plantilla, la solución crea una tabla de DynamoDB que contiene ejemplos de reglas y programas de períodos que puede utilizar como referencia para crear sus propias reglas y programas de períodos personalizados. Para ver un ejemplo de configuración de periodos, consulte [Sample schedules](#) en la documentación de Programador de instancias en AWS .


Para completar el paso de este escenario, debe generar periodos que correspondan a cada carga de trabajo y que satisfagan sus necesidades específicas. Por ejemplo:

```
Period 1 (Workload A):  
  Name: retail-hours  
  Days: Monday to Sunday  
  Hours: 1100 - 2300  
Period 2 (Workload B):  
  Name: office-hours  
  Days: Monday to Friday
```

Hours: 0800 - 1800

Para configurar periodos, siga los pasos que se describen a continuación:

1. Inicie sesión en la consola de [DynamoDB](#) y asegúrese de estar en la misma región en la que lanzó la plantilla para CloudFormation el programador de instancias. AWS
2. En el panel de navegación, elija Tablas y, a continuación, seleccione la tabla denominada. ConfigTable
3. Elija Explore los elementos de la tabla.
4. Para crear un periodo para el horario laboral, seleccione periodo en el elemento office-hours.
5. En la página Editar elemento, cambie el valor de begintime a 0800 y el de endtime a 1800. Deje el valor predeterminado para los días de semana.

 Note

Los valores de begintime y endtime determinan cuándo se deben iniciar y detener las instancias, mientras que el valor de weekdays determina a qué días de la semana se aplica esta programación (de lunes a viernes, en este ejemplo).

6. Seleccione Save changes (Guardar cambios).
7. Para duplicar el periodo de office-hours y usarlo para crear un nuevo periodo para el horario comercial, seleccione periodo para el elemento office-hours. A continuación, en el menú Acciones, seleccione Elemento duplicado.
8. Modifique los atributos para que se ajusten a sus necesidades. Los siguientes atributos se utilizan para cumplir los requisitos del escenario de ejemplo:

```
type: period
name: retail-hours
begintime: 11:00
description: Retail hours
endtime: 23:00
weekdays: mon-sun
```

9. Seleccione Crear elemento.
- 10 En ConfigTableDynamoDB, identifique los dos períodos que acaba de crear que aparecen en las listas de elementos.

Configuración de programaciones

En el contexto de Instance Scheduler on AWS, los cronogramas se refieren a la aplicación de uno o más períodos y a la zona horaria correspondiente. Luego, estas programaciones se asignan a sus instancias como etiquetas. En esta sección, se muestra cómo crear dos programaciones (que se muestran a continuación) para adaptarse a los diferentes patrones de tiempo de las dos cargas de trabajo de ejemplo y, a continuación, asociar los programas a los periodos que creó en la sección anterior.

```
Schedule 1:  
  Name: singapore-office-hours  
  Period: office-hours  
  Timezone: Asia/Singapore  
Schedule 2:  
  Name: singapore-retail-hours  
  Period: retail-hours  
  Timezone: Asia/Singapore
```

Para crear y configurar las programaciones, siga los pasos que se describen a continuación:

1. Inicie sesión en la consola de [DynamoDB](#) y asegúrese de estar en la misma región en la que lanzó la plantilla para CloudFormation el programador de instancias. AWS
2. En el panel de navegación, elija Tablas y, a continuación, seleccione la tabla denominada. ConfigTable
3. Elija Explore los elementos de la tabla.
4. Para duplicar el horario de oficina del Reino Unido y usarlo para crear un nuevo horario para su horario de oficina (horario de oficina de Singapur, por ejemplo), seleccione horario para uk-office-hours el elemento. A continuación, en el menú Acciones, seleccione Elemento duplicado.
5. Modifique los atributos para que se ajusten a sus necesidades. Los siguientes atributos se utilizan para cumplir los requisitos del escenario de ejemplo:

```
type: schedule  
name: singapore-office-hours  
description: Office hours in Singapore  
periods: office-hours  
timezone: Asia/Singapore
```

6. Seleccione Crear elemento.

7. Repita los pasos del 4 al 6 para crear una programación para el horario comercial de Singapur con los siguientes valores de atributo:

```
type: schedule
name: singapore-retail-hours
description: Retail hours in Singapore
periods: retail-hours
timezone: Asia/Singapore
```

8. En ConfigTableDynamoDB, identifique los dos programas y los dos períodos que creó.

Etiqueta de instancias .

Después de definir las programaciones, debe usar etiquetas para asignar las programaciones a las instancias específicas que desee usar. Puede usar el editor de etiquetas de [Grupos de recursos de AWS](#) para generar y asignar etiquetas a sus instancias de Amazon EC2.

1. Inicie sesión en [Consola de administración de AWS](#) y asegúrese de que se encuentra en la misma región en la que lanzó la CloudFormation plantilla anteriormente.
2. Abra la consola de [Grupos de recursos](#). En el panel de navegación, amplíe Etiquetado y, a continuación, seleccione Editor de etiquetas.
3. En la sección Buscar recursos para etiquetar, en Regiones, elija sus regiones. En Tipos de recursos, elija Amazon EC2 o Amazon RDS. Este escenario se centra en las instancias de Amazon EC2 de la carga de trabajo A. El equipo de marketing utiliza la carga de trabajo A en la región de Singapur. Los recursos de esta carga de trabajo ya están etiquetados con una clave Departamento y un valor Marketing. Puede usar esta etiqueta para buscar las instancias.
4. Seleccione Buscar recursos.
5. Seleccione las instancias de la lista de resultados de la búsqueda que desee incluir en la programación y, a continuación, elija Administrar las etiquetas de los recursos seleccionados.
6. En la sección Editar las etiquetas de todos los recursos seleccionados, elija Agregar etiqueta para agregar las etiquetas de programación de Programador de instancias a las instancias de EC2. Puede usar las claves y los valores de las etiquetas que coincidan con programación (que ya se creó en DynamoDB).
7. En Clave de etiqueta, agregue Programación. En Valor de etiqueta, introduce singapore-retail-hours.
8. Seleccione Revisar y aplicar cambios en la etiqueta.

9. Para aplicar la etiqueta a todas las instancias de EC2 que haya seleccionado, elija Aplicar cambios a todos los elementos seleccionados.
10. Repita los pasos del 3 al 9 para cualquier otra programación que quiera aplicar.

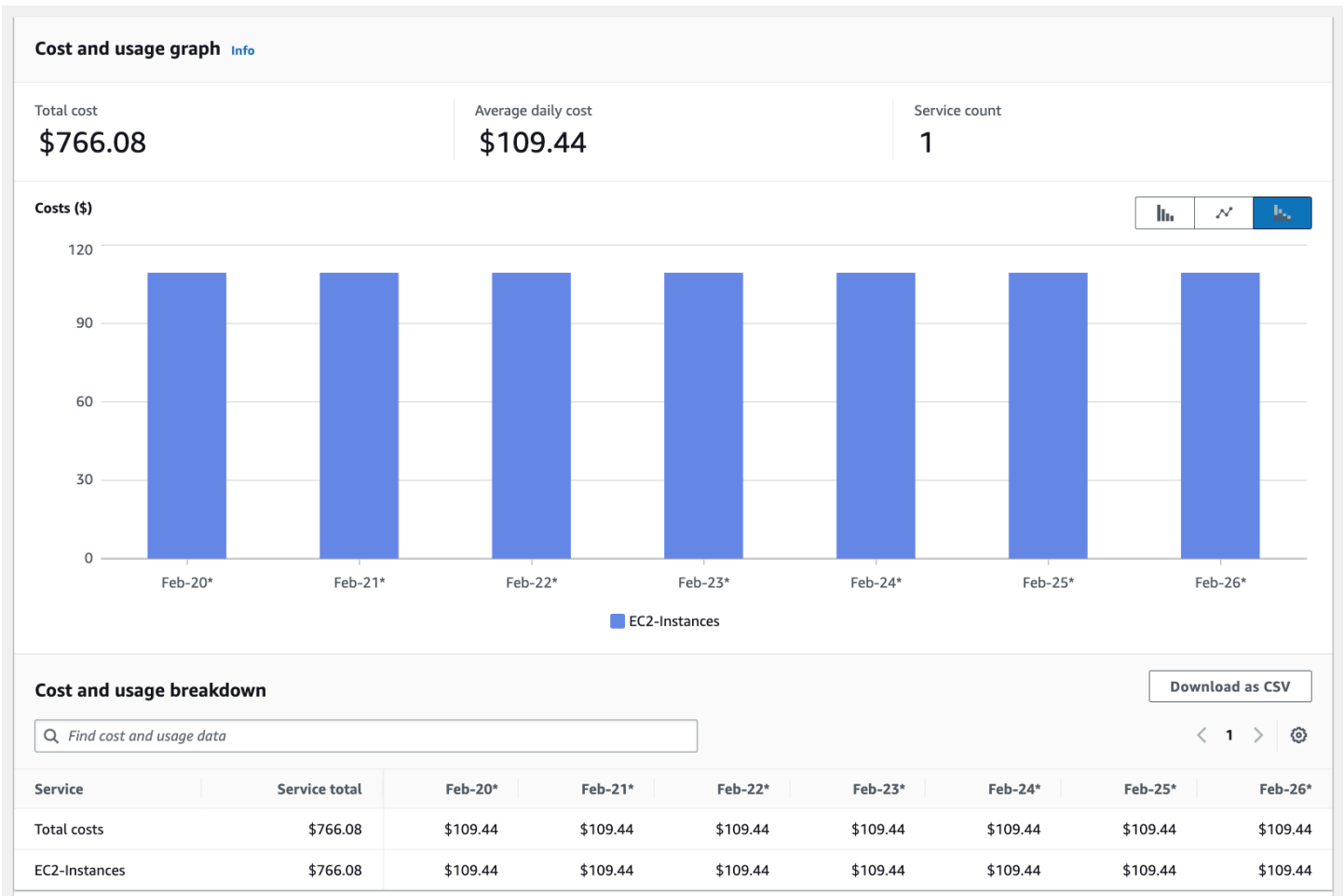
Validar los resultados

Le recomendamos que lo utilice [AWS Cost Explorer](#) para medir los beneficios económicos de usar Instance Scheduler en AWS. Puede usar Explorador de costos para hacer lo siguiente:

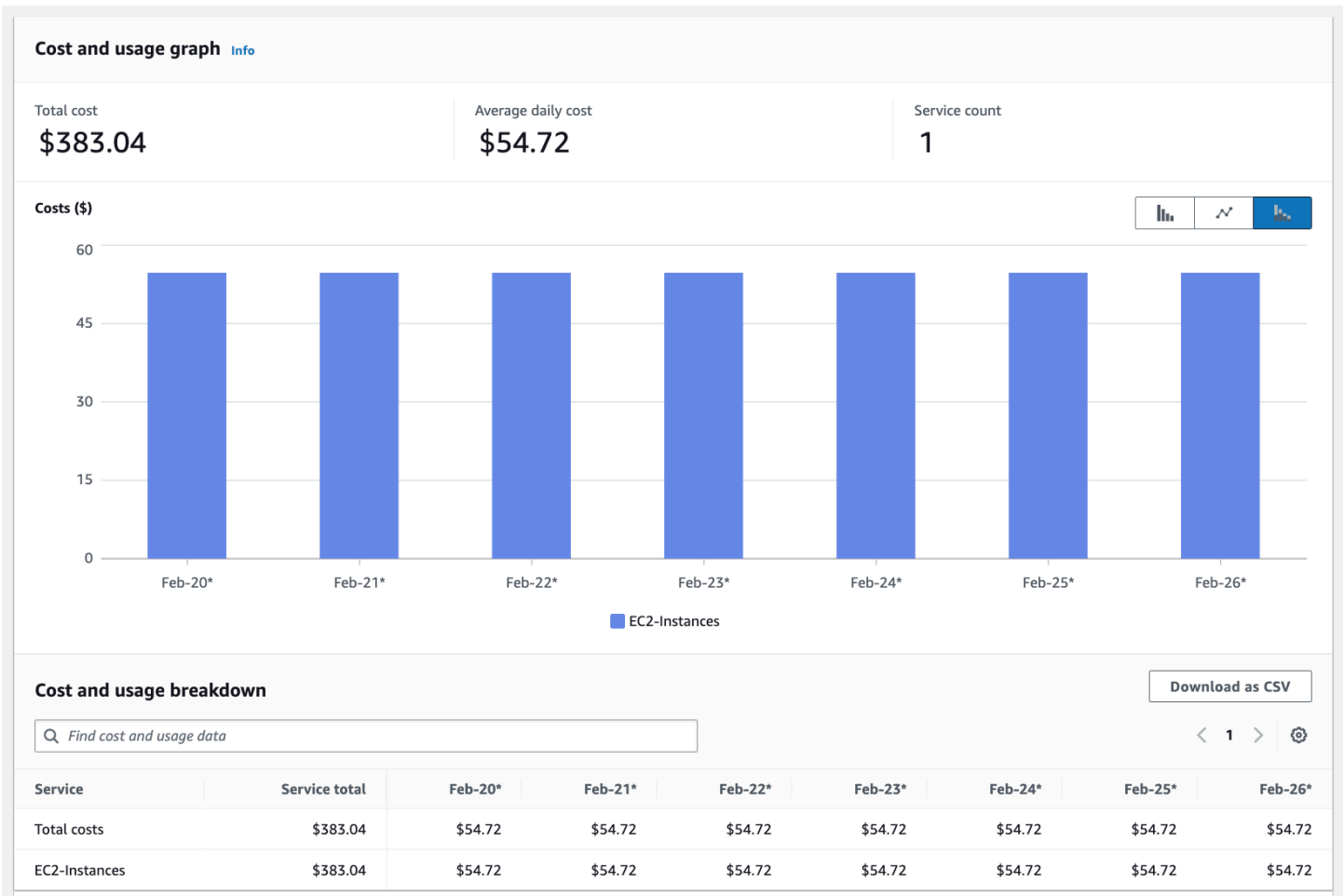
- Ver y analizar los costos asociados a sus instancias de EC2, lo que incluye las instancias administradas por Programador de instancias.
- Filtrar la vista de Explorador de costos por etiquetas para poder centrarse en cargas de trabajo específicas y obtener una visión pormenorizada de los ahorros de costos que se obtienen al utilizar Programador de instancias.
- Obtener información sobre el impacto financiero del uso de Programador de instancias.
- Identifique oportunidades para una mayor optimización de los costos y tome decisiones basadas en datos para optimizar sus AWS gastos.

Los siguientes gráficos ilustran el costo de operar las cargas de trabajo A y B durante un periodo de siete días (de lunes a domingo) antes de la optimización mediante Programador de instancias.

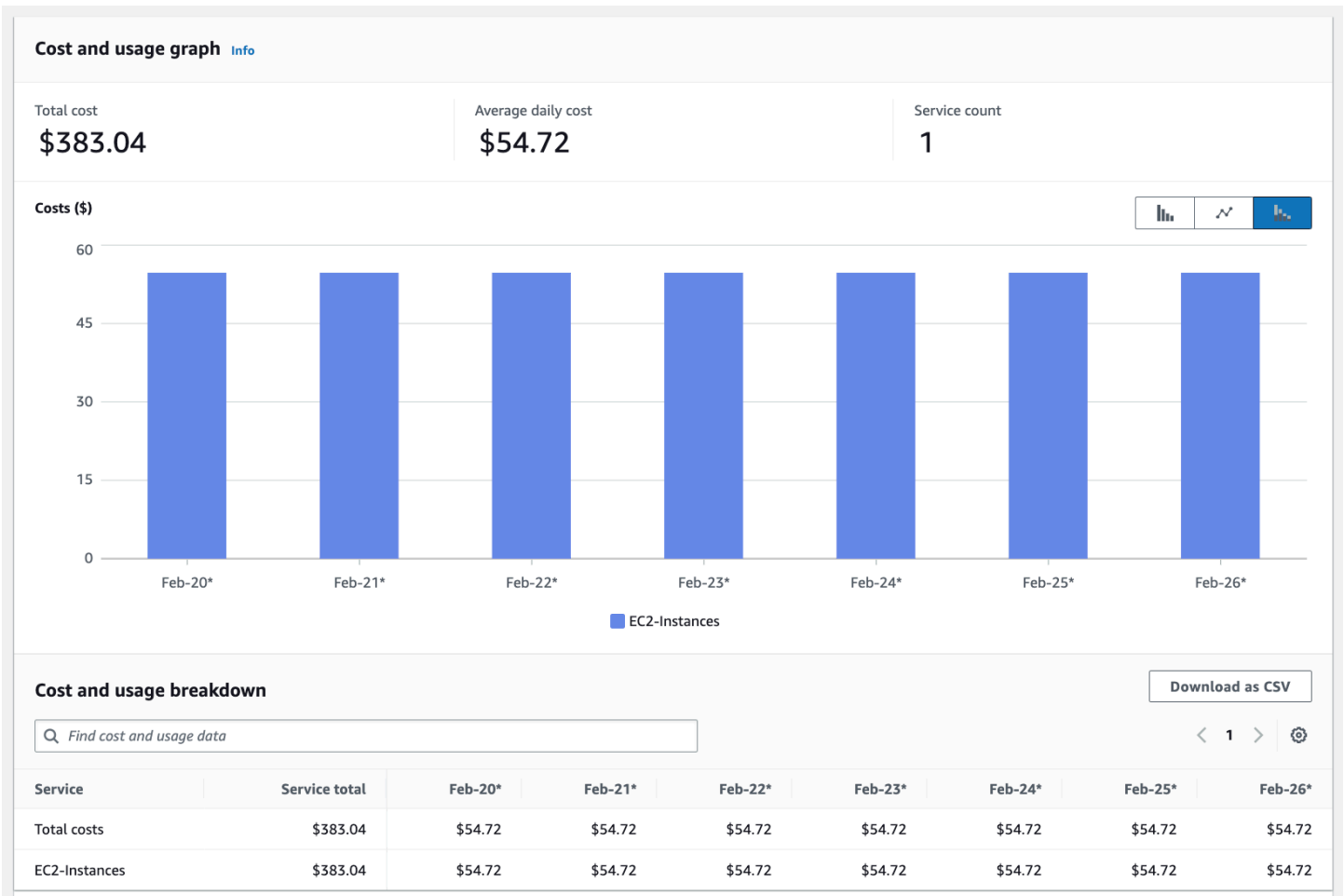
Gastos totales combinados de las cargas de trabajo A y B



Gastos de la carga de trabajo A

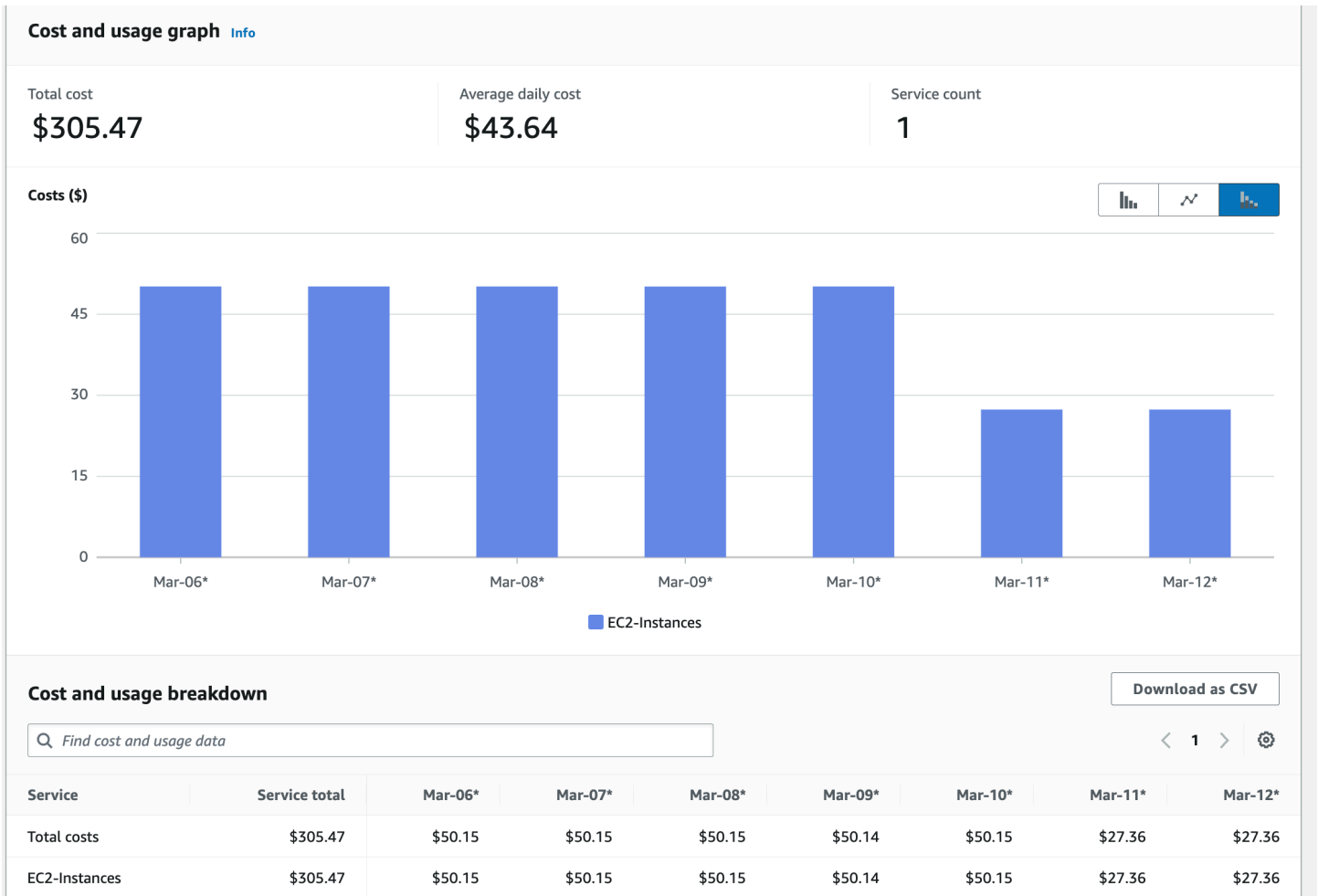


Gastos de la carga de trabajo B

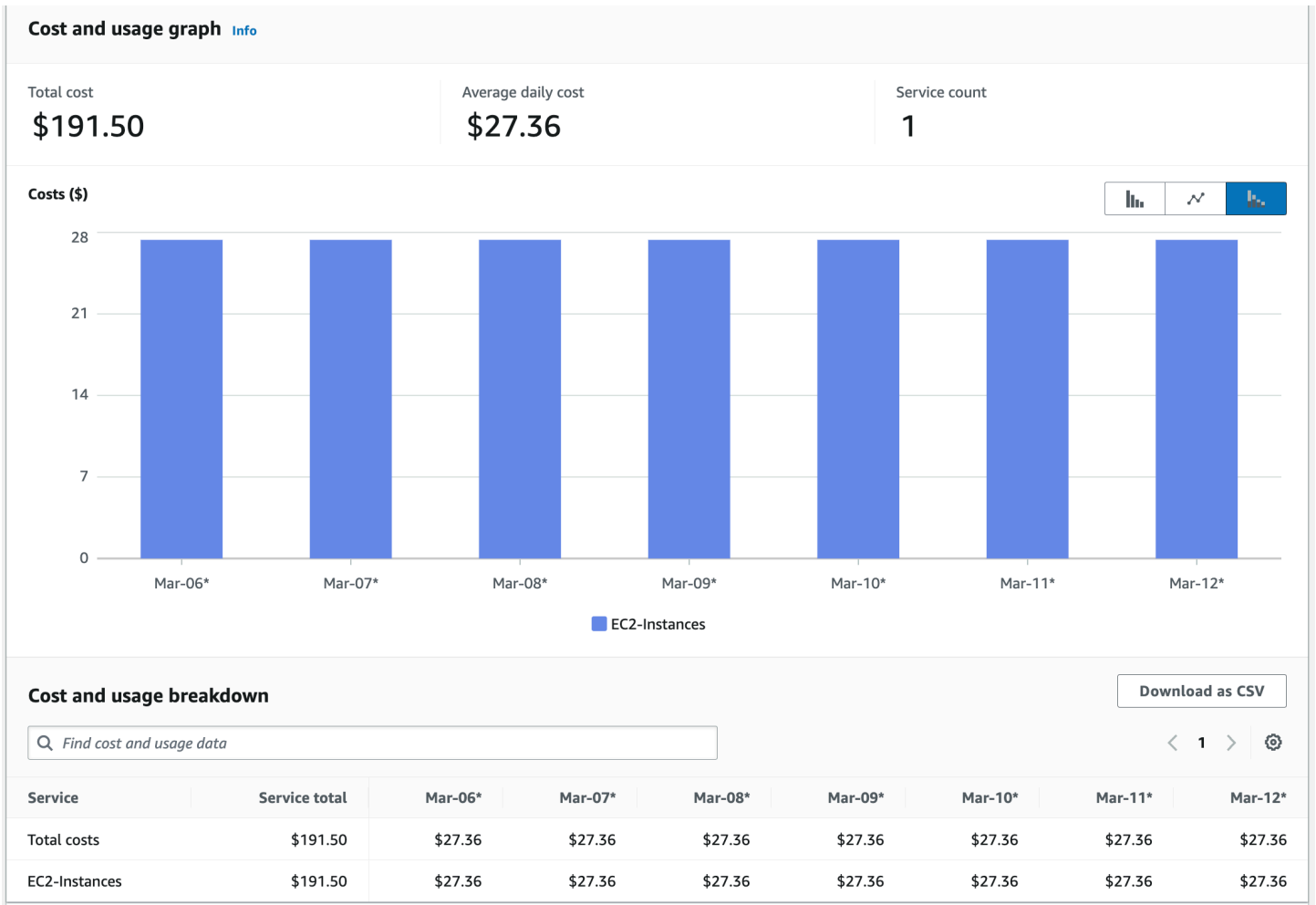


En este escenario, Explorador de costos muestra las reducciones de costos que se obtienen al implementar Programador de instancias en AWS. Los siguientes gráficos ilustran los costos operativos de las cargas de trabajo A y B durante un periodo de siete días (de lunes a domingo) después de la optimización.

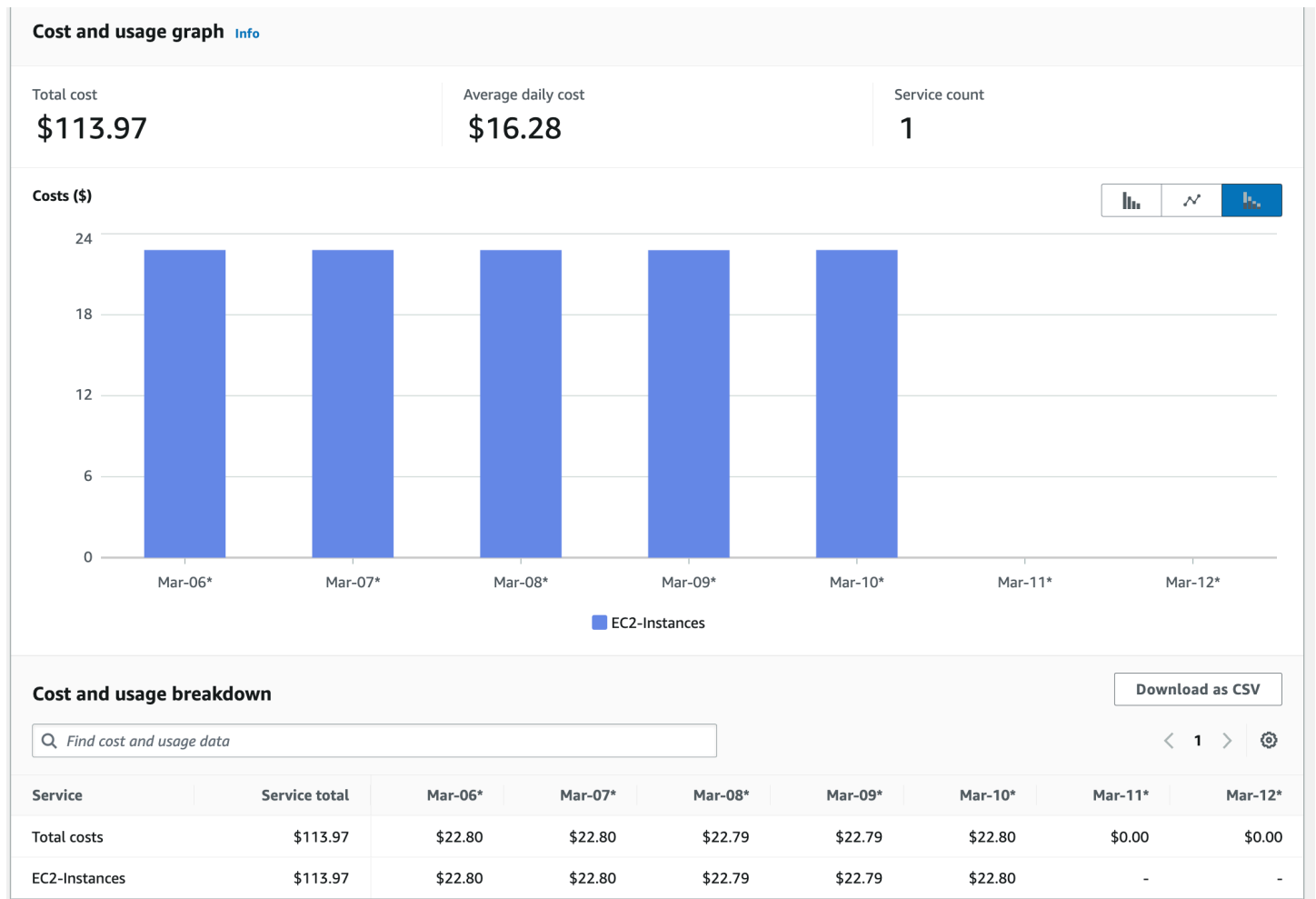
Gastos totales combinados de las cargas de trabajo A y B



Gastos de la carga de trabajo A



Gastos de la carga de trabajo B



Recursos adicionales

- [Automatice el inicio y la detención de AWS instancias](#) (véase el programador de instancias en la documentación AWS)
- [Volver a lo básico: uso de un programador de instancias para controlar los costos de recursos de Amazon EC2 y Amazon RDS](#) () YouTube
- [Etiquetar sus AWS recursos \(Guía del usuario sobre el etiquetado AWS de recursos\)](#)
- [Analizar sus costos con AWS Cost Explorer\(documentación\)](#)Administración de facturación y costos de AWS

Dimensionamiento correcto de las cargas de trabajo de Windows

Descripción general de

El dimensionamiento correcto es una de las herramientas de ahorro de costes más potentes. AWS ofrece varios métodos para recopilar información sobre el tamaño correcto, desde revisar las posibles cargas de trabajo mediante una [evaluación de AWS optimización y licencias \(AWS OLA\)](#) hasta revisar las cargas de trabajo existentes mediante el uso de [AWS Cost Explorer](#)

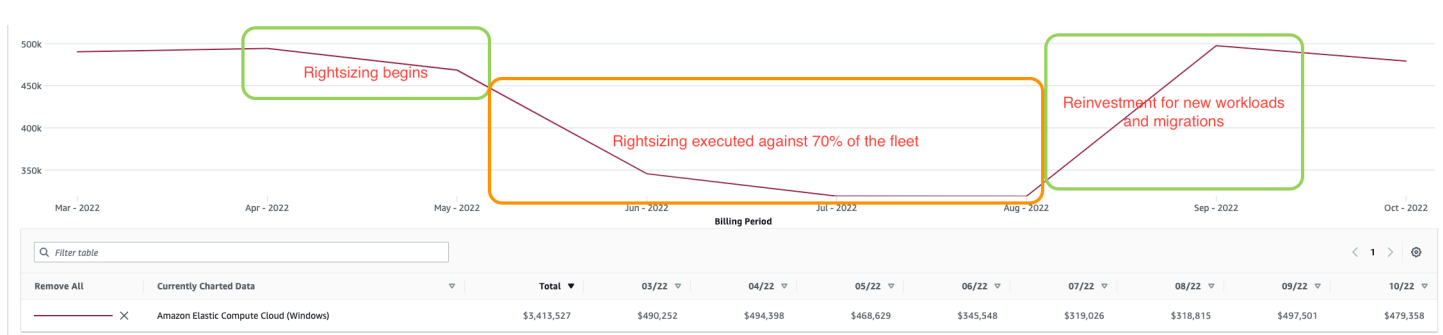
En esta sección se muestra cómo utilizar [AWS Compute Optimizer](#) para identificar oportunidades de dimensionamiento correcto de Amazon EC2. Compute Optimizer ayuda a evitar el sobreaprovisionamiento y el subaprovisionamiento de los siguientes tipos de recursos: AWS

- Tipos de instancias de [Amazon Elastic Compute Cloud \(Amazon EC2\)](#)
- [Volúmenes de Amazon Elastic Block Store \(Amazon EBS\)](#)
- Servicios de [Amazon Elastic Container Service \(Amazon ECS\)](#) en AWS Fargate
- [AWS Lambda](#) funciones basadas en los datos de uso proporcionados por [Amazon CloudWatch](#)

Escenario de optimización de costos

Calcular la eficacia del dimensionamiento correcto puede ser un desafío, ya que este proceso puede llevarse a cabo para una aplicación específica, u un equipo o toda una organización. Por ejemplo, pensemos en una organización que migra varios miles de instancias a cargas de trabajo de Windows AWS, en las que el 90 por ciento de su flota está compuesta por cargas de trabajo de Windows. La organización puede emplear Compute Optimizer para analizar su flota y detectar un exceso significativo de aprovisionamiento en sus cuentas y Regiones de AWS. Luego, pueden usar [Automatización de AWS Systems Manager](#) para ajustar el tamaño de su flota mediante varios periodos de mantenimiento. Como resultado, la organización logra dimensionar correctamente el tipo de instancia del 70 % de su flota y obtener un ahorro del 35 %.

El siguiente panel ilustra los ahorros logrados durante varios meses, ya que esta organización de ejemplo implementó estratégicamente las recomendaciones de dimensionamiento correcto de Compute Optimizer. Su objetivo era operar sus cargas de trabajo existentes de la manera más eficiente posible para reanudar una migración estancada desde un centro de datos de ubicación cuyo contrato estaba a punto de expirar.



Recomendaciones de optimización de costos

Le recomendamos que siga estos pasos para optimizar sus costos mediante Compute Optimizer:

- Activación de Compute Optimizer
- Activación de la recopilación de métricas de memoria para los nodos de Windows
- Consumir las recomendaciones de Compute Optimizer
- Etiquetado de las instancias que se deben redimensionar
- Habilite la etiqueta de asignación de costos para que funcione con AWS las herramientas de facturación
- Implemente las recomendaciones de tamaño adecuadas con Automation AWS Systems Manager
- Métodos alternativos de redimensionamiento
- Revisión del antes y el después de los costos en Explorador de costos

Activación de Compute Optimizer

Puede activar [Compute Optimizer](#) par toda la organización o para una única cuenta en AWS Organizations. La configuración de toda la organización proporciona informes continuos para las instancias nuevas y existentes de toda la flota y para todas las cuentas de miembros. Esto permite que el dimensionamiento correcto sea una actividad recurrente en lugar de una point-in-time actividad.

Nivel de organización

Para la mayoría de las organizaciones, la forma más eficiente de usar Compute Optimizer es hacerlo a un nivel organizacional. Esto ofrece información de varias cuentas y regiones de la organización y centraliza los datos en un único origen para su revisión. Para activar Compute Optimizer por organización, haga lo siguiente:

1. Inicie sesión en su [cuenta de administración de Organizations](#) con un rol que tenga los [permisos necesarios](#) y seleccione la opción de inclusión para todas las cuentas de esta organización. Su organización debe tener [habilitadas todas las características](#).
2. Después de activar la cuenta de administración, puede iniciar sesión en ella, ver todas las otras cuentas de miembros y consultar sus recomendaciones.

Note

Se recomienda configurar una [cuenta de administración delegada](#) para Compute Optimizer. Esto le permite ejercer el principio de privilegio mínimo. De esta forma, puede minimizar el acceso a la cuenta de administración de la organización y, al mismo tiempo, seguir proporcionando acceso al servicio de toda la organización.

Nivel de cuenta única

Si quiere trabajar con una cuenta con costos elevados, pero no tiene acceso a AWS Organizations, puede activar Compute Optimizer para esa cuenta y región. Para obtener más información sobre el proceso de inclusión, consulte [Getting started with AWS Compute Optimizer](#) en la documentación de Compute Optimizer.

Activación de la recopilación de métricas de memoria para los nodos de Windows


Las métricas de memoria proporcionan a Compute Optimizer las métricas esenciales necesarias para hacer recomendaciones bien informadas sobre el dimensionamiento correcto de su organización. Antes de hacer una recomendación, se analiza la CPU, la memoria, la red y el almacenamiento.

Para pasar las métricas de memoria de las instancias EC2 de Windows a Compute Optimizer, debes habilitar CloudWatch el agente y configurar las métricas de memoria para que se recopilen cada 60 segundos. El uso de métricas de memoria con CloudWatch

Habilite el CloudWatch agente y configure las métricas de memoria

Descargue el [ComputeOptimizearchivo.yml](#). Puede usar este archivo para activar la recopilación de memoria en todas las instancias de su cuenta. El archivo de plantilla genera los siguientes componentes:

- [AWS Systems Manager Almacén de parámetros](#): almacena la configuración del CloudWatch agente necesaria para recopilar las métricas de memoria.
- AWS Identity and Access Management Función (IAM) con [políticas AWS administradas AWS Systems Manager](#) adjuntas: esto es para el documento de automatización de Systems Manager.
- [AWS Systems Manager documentos](#): esto instala y configura el CloudWatch agente (sustituyendo cualquier configuración existente CloudWatch).
- AWS Systems Manager Asociación [de administradores de estados](#): permite que los documentos de Systems Manager se ejecuten en todas las instancias de su cuenta.

 Important

La ejecución de esta plantilla sobrescribe cualquier CloudWatch configuración existente en las instancias.

Después, haga lo siguiente:

1. Inicie sesión en la [CloudFormation consola Consola de administración de AWS](#) y ábrala.
2. En el panel de navegación, seleccione Stacks (Pilas).
3. Elija Create stack (Crear pila) y, a continuación, seleccione With existing resources (import resources) (Con recursos existentes (importar recursos)).
4. Elija Siguiente.
5. En Recurso de plantilla, elija Cargar un archivo de plantilla.
6. Elija archivo y cargue el archivo `ComputeOptimize.yml`.
7. Elija Siguiente.
8. En la página Especificar los detalles de la pila, en el campo Nombre de la pila, ingrese un nombre para esta pila y, a continuación, elija Siguiente.
9. En la página Identificar recursos, ingrese los valores de los identificadores de los recursos que va a importar.
10. Seleccione Importar recursos.
11. Una vez implementada la pila, seleccione la pestaña Salidas para buscar la clave, el valor y la descripción de la asociación.

Supervisión del progreso de la asociación

1. Una vez completada la implementación de la CloudFormation pila, abra la [consola de Systems Manager](#).
2. En el panel de navegación, en Administración de nodos, seleccione State Manager.
3. En la página Asociaciones, elija el ID de su asociación.
4. Elija la pestaña Execution history (Historial de ejecución).
5. En la columna ID de ejecución, elija el ID de ejecución de su asociación. El estado debe ser Éxito.

Vea las métricas en CloudWatch

Te recomendamos que esperes al menos cinco minutos para que se completen las métricas.

CloudWatch

1. Abra la [consola de CloudWatch](#).
2. En el panel de navegación, expanda la sección Métricas y seleccione Todas las métricas.
3. Confirma que las métricas aparecen en el CWAgentespacio de nombres.

Note

Para aplicar la configuración a cualquier instancia nueva, vuelva a poner en marcha la asociación.

Consumir las recomendaciones de Compute Optimizer

Supongamos que nos centramos en realizar los cambios de tamaño en una sola cuenta y una sola región. En este ejemplo, Compute Optimizer se activa en un nivel organizativo para todas las cuentas. Tenga en cuenta que el dimensionamiento es un proceso disruptivo que, en la mayoría de los casos, los propietarios de las aplicaciones llevan a cabo con precisión durante un periodo de mantenimiento programado que dura varias semanas.

Si accede a Compute Optimizer desde la cuenta de administración de una organización (como se muestra en los pasos siguientes), puede elegir la cuenta que desea investigar. En este ejemplo, hay seis instancias que se ponen en marcha en una sola cuenta en la región de us-east-1. Las seis instancias están sobreprovisionadas. El objetivo es cambiar el tamaño de las instancias según las recomendaciones de Compute Optimizer.

Identificación de las instancias sobreprovisionadas y exportación de los detalles de las recomendaciones

1. Inicia sesión en la consola de [Compute Optimizer Consola de administración de AWS](#) y ábrela.
2. En el panel de navegación, elija Panel.
3. En el cuadro de búsqueda de la página Panel, ingrese Región=Este de EE. UU. (Norte de Virginia). Luego, ingrese Resultados=Sobreprovisionados. Estos filtros le permiten ver todas las instancias sobreprovisionadas en la región us-east-1.
4. Para ver las recomendaciones detalladas sobre las instancias de EC2 sobreprovisionadas, desplácese hacia abajo hasta la tarjeta instancias de EC2 y, a continuación, seleccione Ver recomendaciones.
5. Elija Exportar y guarde el archivo para usarlo en el futuro.
6. En Bucket de S3, ingrese el nombre del bucket de Amazon S3 que desea que sea el destino del archivo de exportación.

Note

Para guardar las recomendaciones y poder revisarlas en el futuro, debe tener un bucket de S3 disponible para que Compute Optimizer escriba en cada región. Para obtener más información, consulte [Amazon S3 bucket policy for AWS Compute Optimizer](#) en la documentación de Compute Optimizer.

7. En la sección Exportar filtros, seleccione la casilla Incluir recomendaciones para todas las cuentas de miembros de la organización.
8. En Tipo de recurso, elija Instancias de EC2.
9. En la sección Columnas para incluir, marque la casilla Seleccionar todo.
10. Seleccione Exportar.

Selección de instancias en función de las recomendaciones

Las recomendaciones de instancias se basan en las métricas de rendimiento recopiladas y analizadas por Compute Optimizer. Es esencial conocer las cargas de trabajo que se ponen en marcha en la instancia para asegurarse de elegir la mejor instancia. En este ejemplo se presupone que puede elegir entre la última generación de instancias [R6i](#), [R5](#) y [T3](#) de Amazon EC2. Las instancias T3 son ampliables y tienen capacidades de ancho de banda de la red más bajas. Las

instancias R5 y R6 tienen el mismo costo por hora y son prácticamente idénticas. Sin embargo, la instancia R6 tiene una mayor capacidad de ancho de banda de la red, cuenta con la última generación de procesadores Intel y ofrece el mismo espacio de computación que la R5. En este ejemplo, R6 es la mejor opción para redimensionar.

1. En la consola de [Compute Optimizer](#), seleccione Recomendaciones de instancias de EC2 en la barra de navegación. En esta página, se muestra una comparación del tipo de instancia actual con las opciones recomendadas para reemplazarla.
2. Para obtener el ID de la instancia que quiere redimensionar, abra la consola de [Amazon S3](#) desde la cuenta de administración en AWS Organizations.
3. En el panel de navegación, elija Buckets y, a continuación, elija el bucket que utiliza para almacenar los resultados exportados.
4. En la pestaña Objetos, seleccione el archivo de exportación de la lista de objetos y, a continuación, seleccione Descargar.
5. Para extraer la información de la instancia del archivo, puede utilizar el botón Texto a columnas de la pestaña Datos de Microsoft Excel.

Note


IDs Las instancias se representan como Amazon Resource Names (ARNs). Asegúrese de establecer el delimitador en “/” y de extraer el ID de la instancia. Como alternativa, puede escribir un script o usar un entorno de desarrollo integrado (IDE) para recortar el ARN.

6. En Excel, filtre la columna de búsqueda para ver solo las instancias OVER_PROVISIONED. Estas son las instancias que debe redimensionar.
7. Guarda la instancia IDs en un editor de texto para poder acceder a ella fácilmente más adelante.

Etiquetado de las instancias que se deben redimensionar

Etiquetar las cargas de trabajo es una herramienta poderosa para organizar sus recursos en AWS. Las etiquetas le permiten obtener una visibilidad detallada de los costos y facilitar la redistribución de gastos. Para obtener más información sobre las estrategias y los métodos para añadir etiquetas a AWS los recursos, consulta el AWS documento técnico [Mejores prácticas para AWS etiquetar los recursos](#). Para este ejemplo, puede usar [Editor de etiquetas de AWS](#) para ajustar las etiquetas de las instancias sobrepromovidas que quiera redimensionar durante un periodo de mantenimiento. También puede usar esta etiqueta para ver los costos antes y después del cambio.

1. Inicia sesión en la [Grupos de recursos de AWS consola](#) de la cuenta que contiene las instancias a las que se va a cambiar el tamaño Consola de administración de AWS y ábrela.
2. En la barra de navegación, en la sección Etiquetado, seleccione Editor de etiquetas.
3. En Regiones, seleccione la región de destino.
4. Para los tipos de recursos, elija `AWS::EC2::Instance`.
5. Seleccione Buscar recursos.
6. En la página Resultados de la búsqueda de recursos, seleccione todas las instancias que quiera redimensionar y, a continuación, seleccione Administrar las etiquetas de los recursos seleccionados.
7. Seleccione Agregar etiqueta.
8. En Clave de etiqueta, ingrese Dimensionamiento correcto. En Valor de etiqueta, ingrese habilitado. A continuación, seleccione Revisar y aplicar los cambios de etiquetas.

 Note

Puede incluir metadatos adicionales, como el equipo o la unidad de negocio, para facilitar el filtrado más adelante en Explorador de costos.

Después de crear y aplicar etiquetas definidas por el usuario a los recursos, las etiquetas pueden tardar hasta 24 horas en aparecer en la página de etiquetas de asignación de costos para su activación. Después de seleccionar las etiquetas para la activación, pueden tardar hasta 24 horas más en activarse.

Los usuarios avanzados pueden usar [AWS CloudShell](#) dentro de la cuenta y la región de destino para etiquetar varias instancias. Por ejemplo:

```
bash
#!/bin/bash
# Set variables
TAG_KEY="rightsizing"
TAG_VALUE="type-m5"
# Get a list of instance IDs
INSTANCE_IDS=$(aws ec2 describe-instances --query
  "Reservations[].Instances[].InstanceId" --output text)
# Loop through each instance ID and add the tag
for INSTANCE_ID in $INSTANCE_IDS; do
  aws ec2 create-tags --resources $INSTANCE_ID --tags Key=$TAG_KEY,Value=$TAG_VALUE
```

done

Habilite la etiqueta de asignación de costos para que funcione con las herramientas AWS de facturación

Recomendamos activar la etiqueta de asignación de costos definida por el usuario. Esto permite reconocer y filtrar la etiqueta Rightsizing en las herramientas de AWS facturación (por ejemplo, Cost Explorer y). AWS Cost and Usage Report Si no activa esta opción, la opción de filtrado de etiquetas y los datos no estarán disponibles. Para obtener más información sobre el uso de etiquetas de asignación de costos, consulte [Activating user-defined cost allocation tags](#) en la documentación de Administración de facturación y costos de AWS .

1. [Inicie sesión en la consola Consola de administración de AWS y ábrala.AWS Billing](#)
2. En el panel de navegación, en la sección Facturación, elija Etiquetas de asignación de costos.
3. En la pestaña Etiquetas de asignación de costos definidas por el usuario, ingrese Dimensionamiento correcto.
4. Seleccione la clave de etiqueta Dimensionamiento correcto y, a continuación, elija Activar.

Transcurridas 24 horas, la etiqueta debería aparecer en Explorador de costos.

Implementación de las recomendaciones de dimensionamiento correcto con Automatización de Systems Manager

El redimensionamiento es un escenario que requiere detener e iniciar una instancia. En este escenario, es posible que tenga que gestionar esta interrupción en un periodo de mantenimiento y que necesite que diferentes equipos se encarguen de su propio redimensionamiento. Antes de cambiar un tipo de instancia, consulte [Cambios de tipos de instancias de Amazon EC2](#) en la documentación de Amazon EC2.

Los pasos de ejemplo de esta sección implementan las recomendaciones de tamaño correcto por cuenta y región mediante un documento de automatización de Systems Manager denominado [AWS-ResizeInstance](#). Esta estrategia la habitual en la mayoría de las organizaciones, ya que estas suelen requerir diferentes tipos de instancias para diferentes propósitos. También puede usar el mismo documento de automatización AWS-ResizeInstance para segmentar las implementaciones de una o varias cuentas.

1. Inicie sesión en la [consola de Systems Manager Consola de administración de AWS y ábrala](#).

2. En el panel de navegación, en la sección Recursos compartidos, elija Documentos.
3. En la barra de búsqueda, escribe AWS- yResizeInstance, a continuación, selecciona AWS- en los resultados ResizeInstance de la búsqueda.
4. Elija Ejecutar automatización.
5. En la página Ejecutar el runbook de automatización, seleccione Ejecución simple.
6. En la sección Parámetros de entrada, introduzca InstanceId y InstanceType. No cambie ningún otro valor.
7. Seleccione Ejecutar, y, a continuación, espere a que la automatización complete los pasos necesarios para cambiar el tipo de instancia.

Métodos alternativos de redimensionamiento

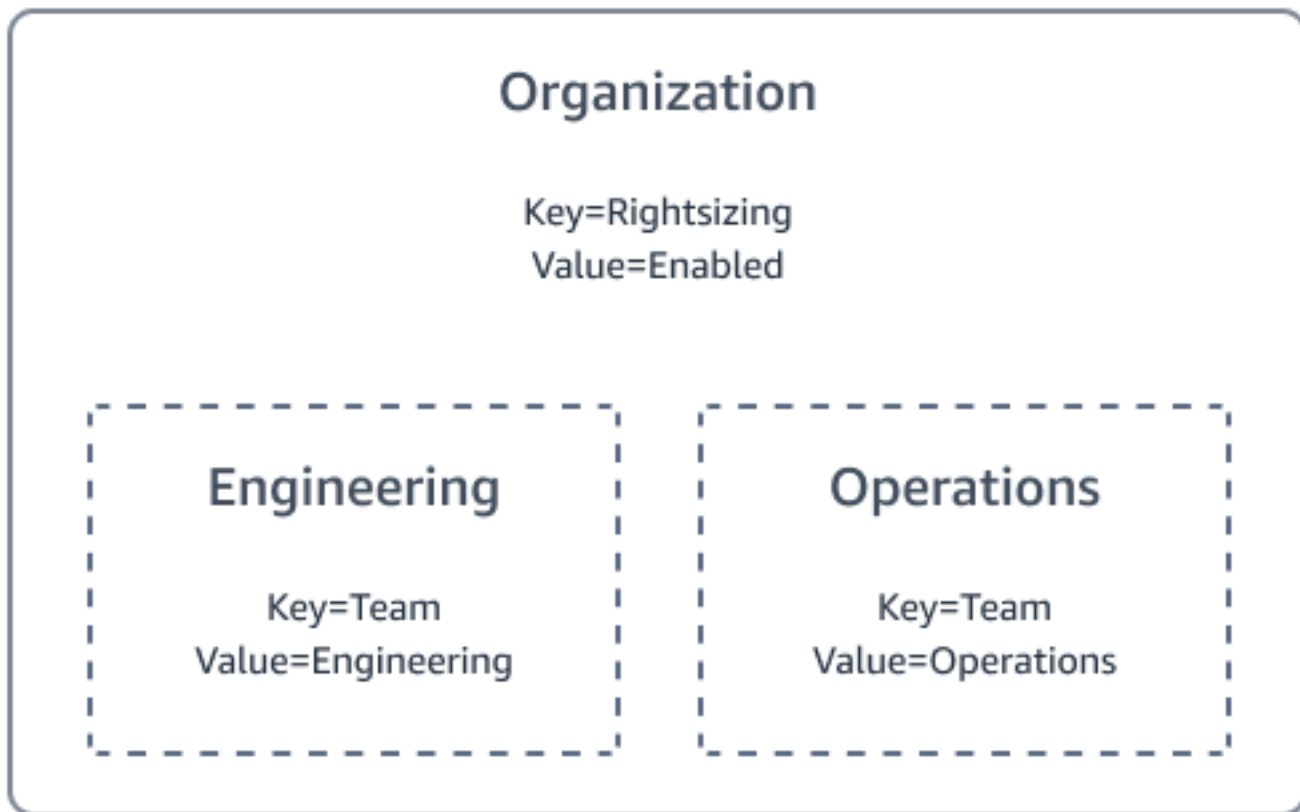
Si utiliza una plantilla de lanzamiento para implementar sus instancias, puede actualizar la plantilla con el tipo de instancia del tamaño correcto y, a continuación, actualizar la instancia para reemplazar las instancias por la versión correctamente dimensionada.

Si planea usar el proceso de dimensionamiento correcto en varias cuentas y regiones, debe crear un documento personalizado de Automatización de Systems Manager. Este documento le permite introducir varias instancias como parámetro y trasladar las instancias de destino al mismo tipo de instancia de destino (por ejemplo, todas las instancias en transición a t3a.medium, independientemente del tipo de instancia de origen).

Revisión del antes y el después de los costos en Explorador de costos

Una vez que haya dimensionado correctamente sus recursos, puede usar Explorador de costos para ver los costos anteriores y posteriores usando la etiqueta Dimensionamiento correcto. Recuerde que puede usar [etiquetas de recursos](#) para hacer un seguimiento de los costos. Al usar varias capas de etiquetas, puede lograr una visibilidad detallada de sus costos. En el ejemplo que se describe en esta guía, la etiqueta Dimensionamiento correcto se utiliza para aplicar una etiqueta genérica a todas las instancias objetivo. A continuación, se utiliza una etiqueta de equipo para organizar aún más los recursos. El siguiente paso consiste en ingresar etiquetas de aplicación para ver mejor el impacto en los costos derivados del funcionamiento de una aplicación específica.

En el siguiente diagrama, se muestra la estructura de etiquetas de una organización.



Imagine una empresa que dimensiona correctamente los servidores web de producción del equipo de operaciones. En Explorador de costos, la etiqueta Dimensionamiento correcto está habilitada, mientras que la etiqueta Equipo es operaciones. En este ejemplo, el dimensionamiento correcto reduce los costos operativos de 0,89 USD a 0,28 USD por hora. Suponiendo que un mes tiene 744 horas, el costo anual antes del dimensionamiento correcto es de 7945,92 USD. Tras redimensionar, el costo anual se reduce a 2499,84 USD. Esto se traduce en una disminución del 68,5 % en los costos anuales de la carga de trabajo. Imagine el impacto que puede tener esto en una organización grande. Tenga en cuenta que esto se hace en un entorno de muestra y las instancias están en su mayoría inactivas. En un entorno de producción, puede obtener ahorros de entre un 10 % y un 35 %.

Ahora, considere el impacto que tendría el dimensionamiento correcto del host bastión del equipo de ingeniería. En Explorador de costos, la etiqueta Dimensionamiento correcto está habilitada, mientras que la etiqueta Equipo es ingeniería. En este ejemplo, el dimensionamiento correcto reduce los costos de 0,75 USD a 0,44 USD por hora. Suponiendo que un mes tiene 744 horas, el costo anual antes del dimensionamiento correcto es de 6696,00 USD. Tras redimensionar, el costo anual se reduce a 3928,32 USD.

Si utilizas varias etiquetas, puede filtrar los datos para obtener detalles pormenorizados de los costos. En este ejemplo, la etiqueta Equipo reduce el ruido para que pueda ver el impacto por equipo. Como la etiqueta Dimensionamiento correcto está habilitada, también puede filtrar por cualquier instancia que tenga esa etiqueta habilitada o sin ningún valor. Esto puede proporcionar una visión global de sus esfuerzos de redimensionamiento, especialmente si se ven en la cuenta de administración (pagadora) en Explorador de costos. Esta vista le permite ver todas las cuentas e instancias.

Imagine una cuenta única en la que la etiqueta Dimensionamiento correcto esté habilitada. Los costos operativos se reducen de 1,64 USD la hora a 0,72 USD la hora. Suponiendo que un mes tiene 744 horas, el costo anual antes del dimensionamiento correcto es de 14 641,92 USD. Tras redimensionar, el costo anual se reduce a 6428,16 USD. Esto se traduce en una disminución del 56 % en los costos de computación de esta cuenta.

Antes de empezar con el proceso de dimensionamiento correcto, tenga en cuenta lo siguiente:

- AWS ofrece muchas opciones para reducir los costes. Esto incluye [AWS OLA](#), en el que se AWS revisan las instancias locales antes de mudarlas a AWS ellas. El AWS OLA también le proporciona recomendaciones sobre el tamaño correcto y orientación sobre licencias.
- Complete todos los procesos de dimensionamiento antes de adquirir [Savings Plans](#). Esto puede ser de ayuda para evitar compra en exceso cuando se contrate Savings Plans.

Recomendaciones

A continuación, recomendamos los pasos próximos:

1. Revise la situación actual y considere la posibilidad de convertir los volúmenes gp2 de Amazon EBS en volúmenes gp3.
2. Revise [Savings Plans](#)

Recursos adicionales

- [AWS Compute Optimizer](#) (documentación de AWS)
- [Mejores prácticas para etiquetar los AWS recursos \(documentos técnicos\)](#)AWS
- [¿Cómo recopilar datos de AWS Compute Optimizer y AWS Trusted Advisor entre sus AWS Organizations](#) () YouTube

- [Optimizing performance and reducing licensing costs: Leveraging AWS Compute Optimizer for Amazon EC2 SQL Server instances](#) (blog de cargas de trabajo de Microsoft en AWS)

Selección del tipo de instancia adecuado para las cargas de trabajo de Windows

Descripción general de

Una diferencia importante entre las cargas de trabajo que funcionan en la nube y las que funcionan en las instalaciones es la práctica del sobreaprovisionamiento. Al comprar hardware físico para uso en las instalaciones, se realiza un gasto en capital que se prevé que dure un periodo predeterminado, normalmente de 3 a 5 años. Para adaptarse al crecimiento previsto durante la vida útil del hardware, éste se adquiere con más recursos de los que su carga de trabajo requiere actualmente. En consecuencia, el hardware físico suele estar sobreaprovisionado, lo que supera con creces las necesidades de la carga de trabajo real.

La tecnología de máquinas virtuales (VM) surgió como un medio eficaz de utilizar los recursos de hardware sobrantes. Los administradores proporcionaron v CPUs y RAM en exceso VMs , lo que permitió al hipervisor gestionar el uso de los recursos físicos entre los servidores ocupados e inactivos mediante la asignación de los recursos no utilizados a cada máquina virtual. Al administrar VMs, los recursos de vCPU y RAM asignados a cada máquina virtual funcionaban más como reguladores de recursos que como indicadores del uso real. La sobreasignación de recursos de máquinas virtuales podría superar fácilmente el triple de los recursos de computación disponibles.

[Amazon Elastic Compute Cloud \(Amazon EC2\)](#) evita el VMs aprovisionamiento excesivo del hardware subyacente, ya que no es necesario. La computación en la nube es un gasto operativo, no un gasto en capital, y solo se paga por lo que se usa. Si su carga de trabajo requiere más recursos en el futuro, aproviónelos cuando realmente los necesite, en lugar de hacerlo de forma preventiva.

Existen cientos de opciones para elegir los [tipos de instancias de Amazon EC2](#) correctos. Si planea migrar una carga de trabajo de Windows a la nube, AWS ofrece un [AWS OLA](#) para ayudarlo a comprender mejor su carga de trabajo actual y proporcionarle un ejemplo de su rendimiento en ella. AWS El análisis de AWS OLA tiene como objetivo hacer coincidir el tipo y el tamaño de la instancia de EC2 adecuados con el uso real en las instalaciones.

Si ya tiene cargas de trabajo funcionando en Amazon EC2 y busca estrategias de optimización de costos, esta sección de la guía le permitirá identificar las diferencias entre las instancias de Amazon EC2 y su aplicabilidad a las cargas de trabajo típicas de Windows.

Recomendaciones de optimización de costos

Para optimizar los costos de los tipos de instancias de EC2, le recomendamos que haga lo siguiente:

- Elección de la familia de instancias adecuada para una carga de trabajo
- Descripción de las variaciones de precios entre las arquitecturas de procesadores
- Comprenda las diferencias de precio y rendimiento de todas las generaciones de EC2.
- Migre a instancias más nuevas.
- Uso de instancias ampliables

Elección de la familia de instancias adecuada para una carga de trabajo

Es importante elegir la familia de instancias adecuada para la carga de trabajo.

Las instancias de Amazon EC2 se dividen en los siguientes grupos:

- Fin general
- Optimizada para computación
- Optimizada para memoria
- Computación acelerada
- Optimizada para almacenamiento
- Optimizadas para HPC

La mayoría de las cargas de trabajo de Windows se dividen en las categorías siguientes:

- Fin general
- Optimizada para computación
- Optimizada para memoria

Para simplificarlo aún más, piense en una instancia de EC2 básica en cada categoría:

- Optimizadas para computación: C6i
- De uso general: M6i
- Optimizada para memoria: R6i

La generación anterior de instancias de EC2 presentaba pequeñas diferencias en los tipos de procesadores. Por ejemplo, las instancias C5 optimizadas para computación tienen procesadores más rápidos que las instancias M5 de uso general o las instancias R5 optimizadas para memoria. Todas las instancias de EC2 de última generación (C6i, M6i, R6i, C6a, M6a y R6a) utilizan el mismo procesador en todas las familias de instancias. Dado que el procesador es el mismo en la última generación de instancias, la diferencia de precio entre las familias de instancias ahora depende más de la cantidad de RAM. Cuanta más RAM tenga una instancia, más cara será.

El siguiente ejemplo ilustra el precio por hora de una instancia de 4 vCPU basada en Intel que funciona en la región us-east-1.

Instancia	v CPUs	RAM	Precio por hora
c6i.xlarge	4	8	0,17\$
m6i.xlarge	4	16	0,19\$
r6i.xlarge	4	32	0,25 DÓLARES

Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

Instancias ampliables

Si bien es una práctica recomendada en la computación en la nube desactivar los recursos computacionales no utilizados para evitar cargos, no todas las cargas de trabajo se pueden desactivar y activar cada vez que se necesitan. Algunas cargas de trabajo permanecen inactivas durante periodos prolongados, pero deben estar accesibles las 24 horas del día.

Las instancias ampliables (T3) ofrecen una forma de mantener activas las cargas de trabajo con picos de actividad o de bajo consumo durante todo el día y, al mismo tiempo, mantener bajos los costos de procesamiento. Las instancias ampliables de EC2 tienen una cantidad máxima de recursos de vCPU que la instancia puede utilizar durante periodos breves. Estas instancias utilizan un sistema basado en [créditos de CPU de ampliación](#). Estos créditos se acumulan durante los periodos de inactividad a lo largo del día. Las instancias Burstable ofrecen vCPU-to-RAM proporciones variables,

lo que las convierte en alternativas a las instancias optimizadas para la computación en algunos casos y a otras instancias de uso general en otros.

En el siguiente ejemplo, se muestra el precio por hora de una instancia T3 (es decir, una instancia ampliable) que funciona en la región us-east-1.

Instancia	v CPUs	RAM (GB)	Precio por hora
t3.nano	2	0,5	0,0052\$
t3.micro	2.	1	0,0104\$
t3.small	2	2.	0,0208\$
t3.medium	2	4	0,0416\$
t3.large	2	8	0,0832\$
t3.xlarge	4	16	0,1664\$
t3.2xlarge	8	32	0,3328\$

Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

Descripción de las variaciones de precios entre las arquitecturas de procesadores

Los procesadores [Intel](#) han sido el estándar para las instancias de EC2 desde su creación. Las generaciones anteriores de instancias de EC2, como C5, M5 y R5, no indicaban el uso de un procesador Intel en la arquitectura (ya que Intel era la opción predeterminada). Las generaciones más recientes de instancias de EC2, como C6i, M6i y R6i, incluyen una “i” para indicar el uso de un procesador Intel.

El cambio en la anotación del procesador en la arquitectura se debe a la aparición de otros procesadores. El procesador más comparable a Intel es [AMD](#) (indicado con una “a”). Los procesadores AMD EPYC utilizan la misma arquitectura x86 y ofrecen un rendimiento similar al de los procesadores Intel, pero a un precio inferior. Como se demuestra en los siguientes ejemplos

de precios, las instancias AMD de EC2 ofrecen un descuento de aproximadamente un 10 % en los costos de computación en comparación con sus homólogos de Intel.

Instancia Intel	Precio por hora	Instancia AMD	Precio	% de diferencia
c6i.xlarge	0,17\$	c6a.xlarge	0,153\$	10%
m6i.xlarge	0,192\$	m6a.xlarge	0,1728\$	10%
r6i.xlarge	0,252\$	r6a.xlarge	0,2268\$	10%

Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

La tercera opción principal son los [procesadores AWS Graviton](#) (indicados con una “g”) en las instancias de EC2. Diseñados por AWS, los procesadores Graviton ofrecen la mejor relación precio-rendimiento en Amazon EC2. Los procesadores Graviton actuales no solo son un 20 % más baratos que sus homólogos de Intel, sino que también ofrecen un aumento del rendimiento del 20 % o más. Se espera que la próxima generación de procesadores Graviton amplíe aún más esta diferencia de rendimiento, y las pruebas muestran un aumento adicional del rendimiento del 25 %.

Windows Server no funciona con los procesadores Graviton, que se basan en la arquitectura ARM. De hecho, Windows Server solo funciona con procesadores x86. Aunque no es posible obtener una mejora del 40 % en la relación precio-rendimiento al usar instancias basadas en Graviton para Windows Server, puede usar procesadores Graviton con determinadas cargas de trabajo de Microsoft. Por ejemplo, las [versiones más recientes de .NET](#) pueden ponerse en marcha en Linux. Esto significa que estas cargas de trabajo pueden utilizar procesadores ARM y beneficiarse de instancias de EC2 de Graviton más rápidas y asequibles.

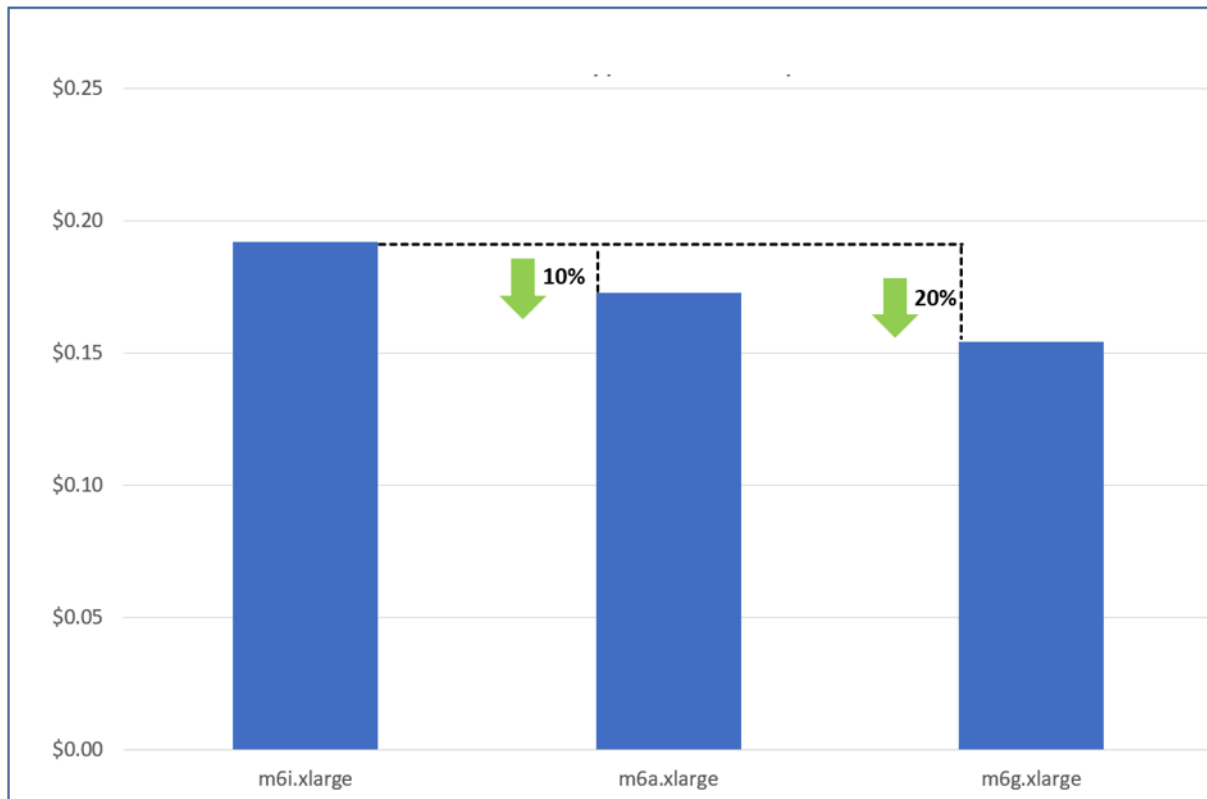
En el siguiente ejemplo, se muestra el precio por hora de una instancia de Graviton que funciona en la región us-east-1.

Instancia Intel	Precio por hora	Instancia Graviton	Precio por hora	% de diferencia
c6i.xlarge	0,17\$	c6g.xlarge	0,136\$	20%
m6i.xlarge	0,192\$	m6g.xlarge	0,154\$	20%
r6i.xlarge	0,252\$	r6g.xlarge	0,2016\$	20%

Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

En el siguiente gráfico se comparan los precios de las instancias de la serie M.



Diferencias de precio y rendimiento entre las generaciones de EC2

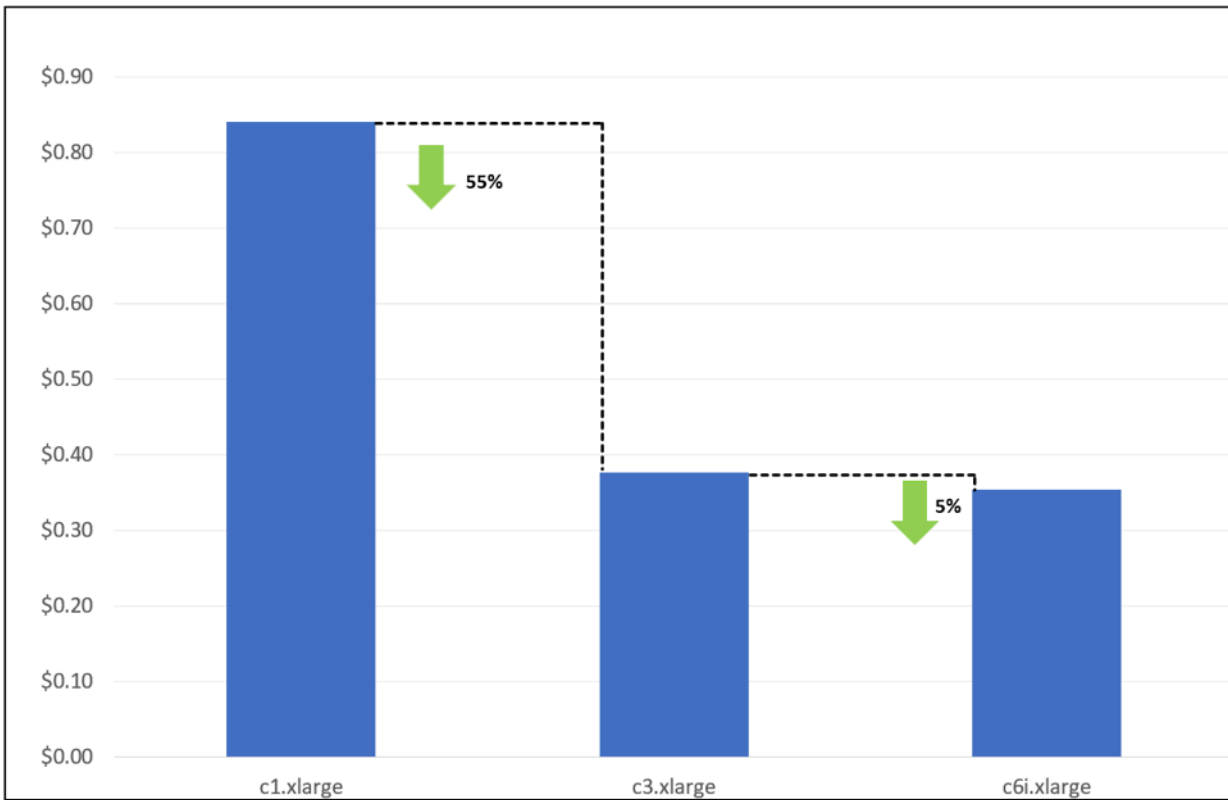
Una de las características más constantes de Amazon EC2 es que cada nueva generación ofrece una mejor relación precio-rendimiento que su predecesora. Como se muestra en la siguiente tabla, el precio de las instancias de EC2 de nueva generación disminuye con cada versión posterior.

Instancia optimizada para la computación	Precio por hora	Instancia de uso general	Precio por hora	Instancia optimizada para memoria	Precio por hora
C1.xlarge	0,52\$	M1.xlarge	0,35\$	r1.xlarge	n/a
C3.xlarge	0,21\$	M3.xlarge	0,266\$	r3.xlarge	0,333\$
C5.xlarge	0,17\$	M5.xlarge	0,192\$	r5.xlarge	0,252\$

Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

En el siguiente cuadro se comparan los costos de las distintas generaciones de instancias de la serie C.



Sin embargo, la 6.^a generación de instancias tiene el mismo precio que la 5.^a generación, como se muestra en la siguiente tabla.

Instancia optimizada para la computación	Precio por hora	Instancia de uso general	Precio por hora	Instancia optimizada para memoria	Precio por hora
C5.xlarge	0,17\$	M5.xlarge	0,192\$	r5.xlarge	0,252\$
C6i.xlarge	0,17\$	M6i.xlarge	0,192\$	r6i.xlarge	0,252\$

Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

A pesar de tener el mismo costo, la nueva generación ofrece una relación precio-rendimiento superior gracias a los procesadores más rápidos, la mejora del rendimiento de la red y el aumento del rendimiento y las IOPS de Amazon Elastic Block Store (Amazon EBS).

Una de las mejoras en cuanto a relación precio-rendimiento más importantes es la mejora de la [instancia X2i](#). Esta generación de instancias ofrece una relación precio-rendimiento hasta un 55 % superior a la de la generación anterior. Como se muestra en la siguiente tabla, la instancia x2iedn supone una mejora en todos los aspectos del rendimiento (todo al mismo precio que la generación anterior).

Instancia	Precio por hora	v CPUs	RAM	Velocidad del procesador	Almacenamiento de la instancia	Red	Rendimiento de Amazon E	IOPS de EBS
x1e.2xlarge	1,66\$	8	244	2.3 GHz	SSD de 237 GB	10 Gbps	125 MB/s	7400
x1iedn.2xlarge	1,66\$	8	256	3.5 GHz	SSD de 240 GB NVMe	25 Gbps	2500 MB/s	65000

Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

Ejemplos de escenarios de

Imagine una empresa de análisis que hace un seguimiento de vehículos de reparto y desea mejorar el rendimiento de SQL Server. Una vez que una pyme de MACO analiza los problemas de rendimiento de esta empresa, la empresa pasa de las instancias x1e.2xlarge a las instancias x2iedn.xlarge. El tamaño de la nueva instancia es más pequeño, pero las mejoras introducidas en las instancias x2 permiten aumentar el rendimiento y la optimización de SQL Server mediante el uso de extensiones de grupos de búferes. Esto permite a la empresa pasar de la edición Enterprise de SQL Server a la edición Standard de SQL Server. También permite a la empresa reducir sus licencias de SQL Server de 8 CPUs a 4 V. CPUs


Antes de la optimización:

Server	Instancia de EC2	Edición de SQL Server	Costo mensual
Prod DB1	x1e.2xlarge	Enterprise	3.918,64 DÓLARES
Prod DB2	x1e.2xlarge	Enterprise	3.918,64 DÓLARES
Total			7.837,28 DÓLARES

Tras la optimización:

Server	Instancia de EC2	Edición de SQL Server	Costo mensual
Prod DB1	x2iedn.xlarge	Standard	1.215,00\$
Orguloso DB2	x2iedn.xlarge	Standard	1.215,00\$
Total			2.430,00 DÓLARES

En conjunto, el cambio de las instancias x1e.2xlarge a las instancias x2iedn.xlarge permite a la empresa, en el escenario del ejemplo, ahorrar 5407 USD al mes en sus servidores de bases de datos de producción. Esto reduce el costo total de la carga de trabajo en un 69 %.

 Note

Los precios se basan en los precios por hora solicitados en la región us-east-1.

Migre a instancias más nuevas.

Las generaciones anteriores de Amazon EC2 utilizan el hipervisor Xen, mientras que las generaciones más recientes utilizan [AWS Nitro System](#). Nitro System entrega casi todos los recursos de computación y memoria del hardware host a sus instancias. Esto se traduce en un rendimiento general mejorado. Al [migrar de instancias basadas en Xen a instancias basadas](#)

en Nitro, se deben tener en cuenta aspectos especiales. Por ejemplo, [AWS Windows AMIs](#) está configurado con la configuración y las personalizaciones predeterminadas que utilizan los medios de instalación de Microsoft. Las personalizaciones incluyen controladores y configuraciones compatibles con los tipos de instancia de última generación ([instancias basadas en Nitro System](#)).

Si va a lanzar instancias desde Windows personalizado AMIs o desde Windows AMIs proporcionado por Amazon que se crearon antes de agosto de 2018, le recomendamos que complete los pasos de [Migración a tipos de instancias de última generación](#) que se indican en la documentación de Amazon EC2.

Uso de instancias ampliables

Si bien las instancias ampliables son una buena forma de ahorrar costos en computación, le recomendamos que las evite en los siguientes casos:

- Las [especificaciones mínimas de Windows Server](#) con la experiencia de escritorio requieren 2 GB de RAM. Evite usar instancias t3.micro o t3.nano con Windows Server porque carecen de la cantidad mínima de RAM.
- Si su carga de trabajo tiene picos de actividad, pero no permanece inactiva el tiempo suficiente para acumular créditos de ampliación, usar instancias de EC2 normales es más eficiente que usar instancias ampliables. Le recomendamos que [supervise los créditos de CPU](#) para verificarlo.
- En la mayoría de los casos, le recomendamos que evite el uso de instancias ampliables con SQL Server. La licencia de SQL Server se basa en el número de v CPUs asignado a una instancia. Si SQL Server está inactivo la mayor parte del día, tendrá que pagar por licencias de SQL que no utilice en su totalidad. En estos escenarios, le recomendamos unificar varias instancias de SQL Server en un servidor más grande.

Siguientes pasos

Le recomendamos que siga estos pasos para optimizar los costos de las instancias de Windows de Amazon EC2:

- Utilice la instancia de EC2 de última generación para obtener la mejor relación precio-rendimiento.
- Utilice instancias de EC2 con procesadores AMD para reducir los costos de computación en un 10 %.
- Maximice la utilización de los recursos eligiendo un tipo de instancia de EC2 que se adapte a su carga de trabajo.

La siguiente tabla muestra ejemplos de puntos de partida comunes para las cargas de trabajo de Windows. Hay opciones adicionales disponibles, como volúmenes de almacenamiento de instancias para mejorar las cargas de trabajo de SQL Server o instancias EC2 con proporciones mucho mayores vCPU-to-RAM. Le recomendamos que pruebe sus cargas de trabajo minuciosamente y que utilice herramientas de supervisión, AWS Compute Optimizer , para realizar los ajustes necesarios.

Carga de trabajo	Típico	Opcional
Active Directory	T3, M6i	R6i
Servidores de archivos	T3, M6i	C6i
Servidores web	T3, C6i	M6i, R6i
SQL Server	R6i	x2iedn, X2iezn

Si debe cambiar el tipo de instancia de EC2, el proceso normalmente implica un simple reinicio del servidor. Para obtener más información, consulte [Cambios de tipos de instancias de Amazon EC2](#) en la documentación de Amazon EC2.

Antes de cambiar el tipo de instancia, le recomendamos que considere las siguientes acciones:

- Debe detener las instancias con copias de seguridad de Amazon EBS para poder cambiar el tipo de instancia. Planifique el tiempo de inactividad mientras la instancia esté detenida. El detenimiento y el cambio de tipo de instancia puede tardar unos minutos y el tiempo que se tarda en reiniciar la instancia es variable, en función de los scripts de inicio de la aplicación. Para obtener más información, consulte [Detención e iniciación de una instancia de Amazon EC2](#) en la documentación de Amazon EC2.
- Cuando detiene e inicia una instancia, la AWS mueve a un nuevo hardware. Si la instancia tiene una IPv4 dirección pública, AWS libera la dirección y proporciona a la instancia una nueva IPv4 dirección pública. Si necesitas una IPv4 dirección pública que no cambie, usa una [dirección IP elástica](#).
- No puede cambiar el tipo de instancia ni su tamaño cuando la [hibernación](#) se encuentra habilitada.
- No puede cambiar el tipo de instancia de una [instancia de spot](#).
- Si la instancia pertenece a un grupo de escalado automático, Amazon EC2 Auto Scaling indica que la instancia detenida está mal estado y podría terminarla e iniciar una instancia de sustitución. Para evitar esto, puede suspender los procesos de escalado del grupo mientras cambia el tipo

de instancia. Para obtener más información, consulte [Suspend and resume a process for an Auto Scaling group](#) en la documentación de Amazon EC2 Auto Scaling.

- Cuando cambias el tipo de instancia de una NVMe instancia con volúmenes de almacenes de instancias, la instancia actualizada puede tener volúmenes de almacenes de instancias adicionales, ya que todos los volúmenes de almacenes de NVMe instancias están disponibles incluso si no están especificados en Amazon Machine Image (AMI) o en la asignación de dispositivos de bloques de instancias. De lo contrario, la instancia actualizada tiene la misma cantidad de volúmenes de almacén de instancias que especificó cuando lanzó la instancia original.

Recursos adicionales

- [Tipos de instancias de Amazon EC2 \(documentación\)](#) AWS
- [Optimización y evaluación de licencias de AWS](#) (documentación de AWS)

Uso de licencias propias para las cargas de trabajo de Windows y SQL Server

Descripción general de

Si tiene inversiones importantes en cargas de trabajo de Microsoft y en acuerdos de licencia empresarial existentes, puede elegir entre varias AWS opciones para respaldar estas cargas de trabajo, incluidas las opciones de [licencia incluida \(proporcionada por AWS\)](#) y [Bring Your Own License \(BYOL\)](#). Puede utilizar los [hosts dedicados de Amazon EC2](#) para aprovechar al máximo los acuerdos de licencia de Microsoft existentes y migrar Windows Server a AWS. Esto puede ahorrarle hasta un 50 % en los costos de las instancias de Amazon EC2. Dado que las licencias de Windows representan aproximadamente la mitad de los costes de las instancias, instalar Windows Server AWS en hosts dedicados puede suponer un importante ahorro de costes. Dado que Windows Server no puede tener un [arrendamiento predeterminado \(compartido\)](#), los hosts dedicados son la opción ideal si desea utilizar sus licencias actuales de Windows Server en AWS ellas.

Los hosts dedicados no son solo para las instancias BYOL de Windows Server. También le ofrecen la flexibilidad necesaria para adaptar sus licencias en las instalaciones a las cargas de trabajo de SQL Server existentes. Los hosts dedicados exponen los núcleos físicos del servidor subyacente y le permiten otorgar la licencia de SQL Server a los núcleos físicos. Esto no es posible en el arrendamiento predeterminado (compartido), donde las licencias de SQL Server se basan en la

cantidad de unidades virtuales CPUs asignadas a la instancia. Esta función le permite licenciar cargas de trabajo de SQL Server AWS de forma coherente con su estrategia de licencias locales. Por lo tanto, puede ahorrar hasta un 50 % en los costos asociados al uso de licencias de SQL Server en comparación con la tenencia (uso compartido) por defecto, además de ahorrar en los costos de las instancias, si utiliza licencias de Windows aptas. Para obtener más información, consulte la sección [Explicación de las licencias de SQL Server](#) de esta guía.

Hosts dedicados de Amazon EC2

Un host dedicado de Amazon EC2 es básicamente el mismo host EC2 que AWS utiliza para ejecutar sus ofertas informáticas de EC2. La diferencia es que estos hosts están totalmente dedicados a un solo cliente y proporcionan acceso exclusivo a la infraestructura física subyacente. Puede usar hosts dedicados para poner en marcha sus instancias en un hardware totalmente dedicado a su uso, en lugar de compartir recursos con otros clientes de AWS . Esto le da un mayor control sobre los recursos en la nube y le permite reducir los costos al traer sus propias licencias de software, como las licencias de Windows Server y SQL Server, a AWS.

Tenga en cuenta lo siguiente:

- Un host dedicado es un servidor físico completamente dedicado a un solo cliente. Obtiene visibilidad de los sockets y los núcleos físicos del host dedicado para poder cumplir con los requisitos de cumplimiento de las licencias, como los acuerdos de licencia de software por socket, por núcleo o por VM.
- Los hosts dedicados que admiten varios tamaños de instancias de la misma familia de instancias se denominan 'hosts dedicados heterogéneos'. Estas [familias de instancias](#) incluyen T3, A1, C5, M5, R5, C5n, R5n y M5n. Otras familias de instancias admiten solo un tamaño de instancia en el mismo host dedicado. Se denominan "hosts dedicados homogéneos".
- Los hosts dedicados se facturan por host. Esto significa que se le cobrará por host dedicado, independientemente del número de instancias que lo usen. Los precios de los hosts dedicados varían según la familia de instancias, la región y las opciones de pago seleccionadas. Puede elegir la configuración óptima para su carga de trabajo a fin de lograr los resultados de rendimiento y costos deseados.

Este diagrama ilustra las diferencias entre las instancias de tenencia compartida y los hosts dedicados.



Hosts dedicados homogéneos

Imagine un escenario en el que se utilice un host dedicado M6i. Los hosts dedicados M6i y R6i tienen dos sockets, 64 núcleos físicos y admiten tipos de instancias del mismo tamaño. Se denominan “hosts dedicados homogéneos”. Esto significa que la cantidad de instancias que puede lanzar en un único host dedicado M6i depende del tamaño de la instancia.

Por ejemplo:

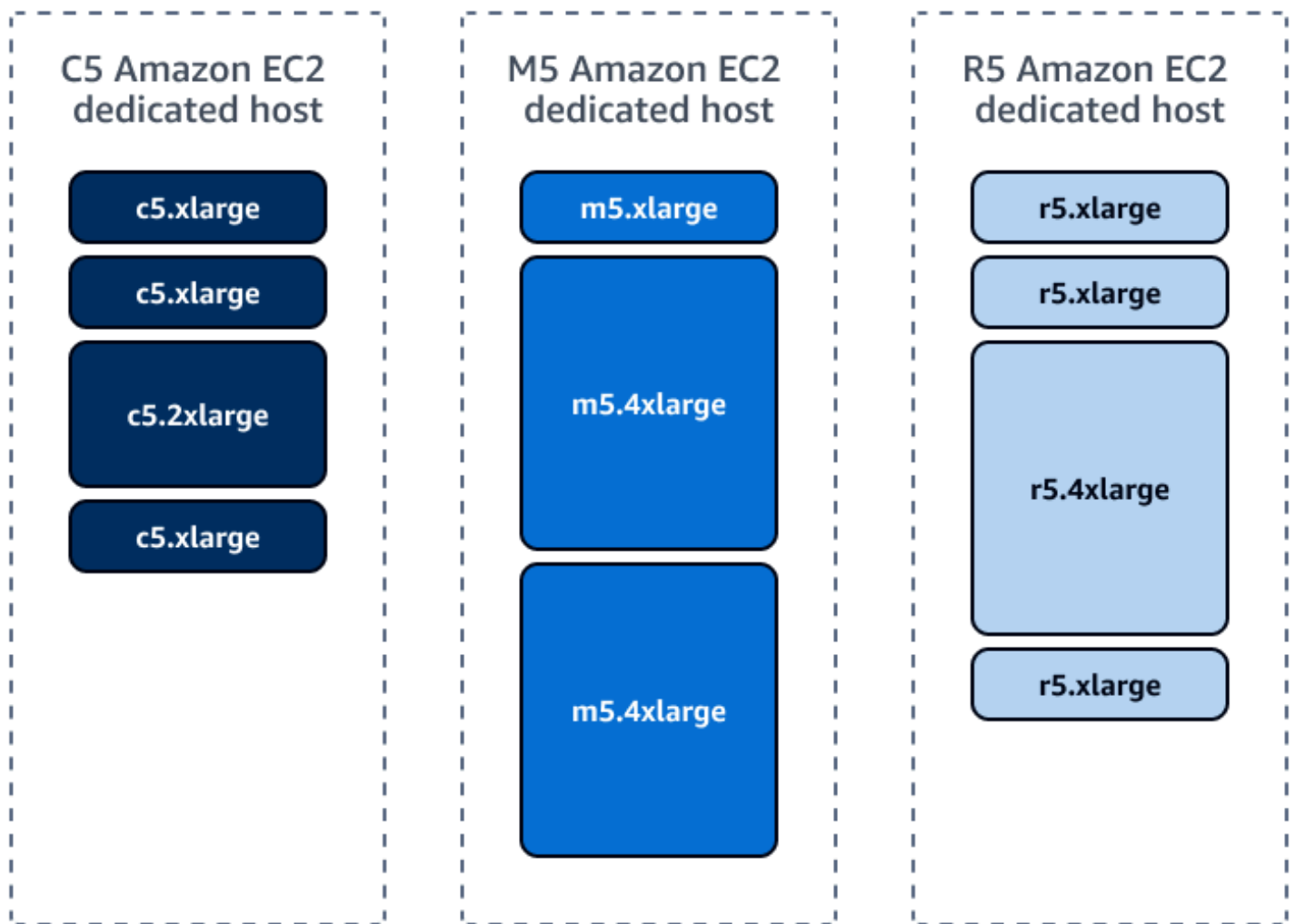
- En el caso de xlarge (4 vCPUs), puede lanzar un máximo de 32 instancias m6i.xlarge en este host dedicado.
- En el caso de 8xlarge (32 vCPUs), puede lanzar un máximo de 4 instancias m6i.8xlarge en este host dedicado.
- En el caso de Metal (128 vCPUs), puede lanzar un máximo de 1 instancia m6i.metal en este host dedicado.

En el siguiente diagrama, se muestran las opciones de host dedicado para las instancias M6.



Hosts dedicados heterogéneos

Los hosts dedicados que admiten varios tamaños de instancia en el mismo host se denominan “hosts dedicados heterogéneos de Amazon EC2”. El siguiente diagrama muestra un ejemplo de hosts dedicados C5, M5 y R5 con varios tamaños de instancia, como 2xlarge, xlarge y 4xlarge.



Administración de hosts dedicados

Le recomendamos que tenga en cuenta lo siguiente en relación con la administración de hosts dedicados de Amazon EC2:

- Para aprovechar al máximo los hosts dedicados, puede [compartir un único host entre varias cuentas de su organización](#). El uso compartido de hosts permite la optimización de los recursos y puede suponer un ahorro de costos al utilizar todos los espacios disponibles en el host. Al compartir un host dedicado entre las unidades de negocio, puede centralizar su infraestructura de TI y mejorar la utilización de los recursos a la par que mantiene la separación entre las cargas de trabajo. Si forma parte de una organización AWS Organizations y el uso compartido está habilitado dentro de su organización, los consumidores de su organización tendrán acceso automático al host dedicado compartido. De lo contrario, los consumidores reciben una invitación para unirse al

recurso compartido y se les concede acceso al host dedicado compartido después de aceptar la invitación.

- Puede poner en marcha Windows Server 2022 en hosts dedicados con el modelo de licencia incluida, ya que Windows Server 2019 es la última versión en la que se puede utilizar BYOL. Si quiere usar Windows Server 2022 en hosts dedicados, debe usar instancias de Windows Server 2022 con licencia incluida.
- [AWS License Manager](#) es una solución integral para administrar las licencias de software de varios proveedores en todos los entornos locales AWS y locales. Si [usa License Manager](#), puede obtener una mayor visibilidad y control sobre cómo se usan las licencias de software, lo que se traduce en ahorros de costos y una mejora del cumplimiento normativo. Puede usar License Manager para establecer reglas que emulen sus condiciones de licencia exclusivas. License Manager le permite implementar esas reglas y evitar el uso indebido de las licencias. Esto puede reducir el riesgo de incumplimiento y mejorar los procesos de administración de licencias.
- Puede usar License Manager para automatizar la ubicación, la liberación y la recuperación de los hosts mediante [grupos de recursos de hosts](#). Esto puede aumentar la productividad y reducir la sobrecarga administrativa. License Manager también proporciona una vista centralizada del uso de las licencias en los entornos locales AWS y en función de las normas de licencia, lo que facilita la gestión de las compras incrementales de licencias, el cumplimiento y las auditorías de los proveedores en toda la organización. Además, License Manager se integra con AWS Organizations y AWS Resource Access Manager (AWS RAM) para compartir las configuraciones de licencia entre cuentas y regiones. Esto le permite crear informes para todo su entorno en función de un cronograma y gestionar las normas de licencia de forma centralizada y unificada en una Cuenta de AWS. En última instancia, esto puede mejorar la gobernanza y reducir la complejidad.
- Al diseñar un sistema de alta disponibilidad para hosts dedicados en una sola región, asegúrese de haber asignado un mínimo de dos hosts dedicados en un mínimo de dos zonas de disponibilidad para las cargas de trabajo críticas para la producción. Para obtener más información, consulte la implementación de referencia en [Amazon EC2 Dedicated Hosts for Microsoft Windows on AWS](#).
- Cada familia de instancias en un host dedicado cuenta con un límite en el número de instancias que se pueden poner en marcha para cada tamaño de instancia. Para obtener más información, consulte [Tabla de configuración de host dedicado](#) en la documentación de Amazon EC2.

AWS opciones de licencia

Las licencias se clasifican en las siguientes categorías principales:

- **Licencia incluida:** esta opción de licencia le permite comprar y usar licencias bajo demanda, pagando únicamente por lo que use. Es ideal para casos de uso en los que busque flexibilidad en el uso de las licencias y desee evitar costos iniciales. Puede elegir entre varios productos de Windows Server, SQL Server y otros productos de Microsoft.
- **Productos BYOL con movilidad de licencias:** si ya dispone de licencias y desea usarlas en la nube, esta opción de licencia le permite llevar sus propias licencias a la nube gracias al [programa Movilidad de licencias de Microsoft](#). Los productos con movilidad de licencias, como SQL Server con Software Assurance (SA), se pueden usar en régimen de tenencia compartida o dedicada. Esto reduce los costes de las AWS instancias.
- **Productos BYOL sin movilidad de licencias:** para los productos de Microsoft, como Windows Server, que carecen de movilidad de licencias, AWS ofrece opciones específicas para usar estos productos en la nube. Además, los hosts dedicados permiten conceder licencias a los núcleos físicos, lo que permite ahorrar un 50 % o más en las licencias necesarias para poner en marcha sus cargas de trabajo. Los hosts dedicados son una opción excelente para cargas de trabajo estables y predecibles que están en marcha la mayor parte del tiempo.

Uso de licencias de Windows Server

Incorporar sus propias licencias de Windows es una de las estrategias más eficaces para la optimización de las licencias, ya que le permite aprovechar las inversiones existentes y reducir sus AWS gastos. Los escenarios BYOL específicos no requieren beneficios de SA ni de Movilidad de licencias, pero siempre es necesaria una infraestructura dedicada de Amazon EC2. Para cumplir los requisitos, debe haber adquirido las licencias perpetuas antes del 1 de octubre de 2019 o haberlas agregado durante un proceso de compensación en el marco de una inscripción empresarial activa en vigor antes del 1 de octubre de 2019. En estos escenarios BYOL específicos, puede actualizar solo las licencias a las versiones disponibles antes del 1 de octubre de 2019. Por ejemplo, si dejó de usar SA en 2017, solo tiene los derechos de implementación hasta Windows Server 2016, no hasta 2019. Sin embargo, 2019 es la última versión válida para BYOL. AWS Para obtener más información, consulte [Licencias: Windows Server](#) en la AWS documentación.

El uso de sus propias licencias puede afectar significativamente al costo de poner en marcha las cargas de trabajo de Microsoft en AWS. Si usa sus propias licencias, no tiene que pagar costos

adicionales por el uso de licencias para las instancias activas en la nube, lo que puede suponer un ahorro de costos considerable.

En la siguiente tabla se muestra el costo mensual bajo demanda que supone poner en marcha una única instancia c5.xlarge las 24 horas del día, los 7 días de la semana, en varias configuraciones.

Configuración	Costo mensual (USD)
Windows Server + Edición Enterprise de SQL Server	1353,00 USD (licencia incluida)
Windows Server + Edición Standard de SQL Server	609,00 USD (licencia incluida)
Solo Windows Server	259,00 USD (licencia incluida)
Solo computación (Linux)	127,00\$

Puede usar sus licencias actuales para reducir los costos de las licencias y ahorrar dinero en su factura total AWS .

Para poder optar por la estrategia BYOL en los hosts dedicados de Amazon EC2, debe traer sus propias licencias de software, como las de Windows Server y SQL Server. BYOL le permite utilizar sus licencias existentes AWS y puede suponer un ahorro de costes. Para usar sus propias licencias, debe disponer de los derechos de licencia del proveedor de software y también proporcionar el medio o la imagen de instalación del software. El medio o la imagen de instalación se pueden usar para lanzar instancias en hosts dedicados. Para obtener más información sobre la creación de una AMI BYOL, consulte [Cómo crear Windows Server Bring-Your-Own-License AMIs desde un entorno local con VM Import/Export en](#) el blog Microsoft Workloads on. AWS

Note

Un tipo de licencia configurado como Automático equivale a una [opción con licencia incluida de AWS](#). Esta opción puede provocar gastos por pedidos bajo demanda no deseados. Tendrá que cambiar de [tipo de licencia](#).

Escenarios de optimización de costos

El dimensionamiento correcto y la optimización de las licencias son un componente clave de la optimización de costos en AWS. Si implementa las estrategias correctas, puede reducir los costos asociados al uso de licencias, garantizar el cumplimiento normativo y obtener el mejor valor posible de su inversión en licencias mediante el uso de los hosts dedicados de Amazon EC2 y la opción BYOL.

En esta sección se abordan los siguientes dos escenarios de ejemplo:

- Ahorro de costos con hosts dedicados (T3)
- Comparación de la tenencia compartida con hosts dedicados con BYOL de SQL Server
- Implementaciones de SQL Server de alta disponibilidad

Ahorro de costos con hosts dedicados (T3)

Los hosts dedicados T3 se diferencian de otros hosts dedicados de Amazon EC2 que tradicionalmente proporcionan recursos de CPU fijos. Los hosts dedicados T3, por el contrario, admiten instancias ampliables que son capaces de compartir los recursos de la CPU, proporcionar un rendimiento básico de la CPU y ampliarse cuando sea necesario. Compartir los recursos de la CPU, también conocido como “sobresuscripción”, es lo que permite que un único host dedicado T3 admita hasta cuatro veces más instancias que los hosts dedicados de uso general comparables.

Los hosts dedicados T3 reducen el TCO, pues ofrecen una densidad de instancias superior a la de cualquier otro host dedicado de Amazon EC2. Las instancias T3 Burstable le permiten consolidar un mayor número de instancias con un uso low-to-moderate medio de la CPU en un menor número de hosts que nunca. Los hosts dedicados T3 también ofrecen tamaños de instancia más pequeños en un mayor número de combinaciones de vCPU y memoria que otros hosts dedicados de Amazon EC2. Los tamaños de instancia más pequeños pueden contribuir a reducir el TCO y ayudar a ofrecer índices de unificación equivalentes o superiores a los de los hosts en las instalaciones.

Los hosts dedicados T3 son los más adecuados para ejecutar software BYOL con uso de low-to-moderate CPU y licencias de software elegibles por socket, por núcleo o por máquina virtual, incluidas las bases de datos Microsoft Windows Desktop, Windows Server, SQL Server y Oracle.

Uso de hosts dedicados T3 para reducir las licencias de la edición Datacenter de Windows Server (por núcleo)

En los entornos locales, está aprovechando el hecho de que puede sobresuscribir fácilmente sus VMware hosts físicos y lograr altos niveles de CPUs consolidación.

Considere el siguiente ejemplo. Actualmente utiliza VMware hosts de 10 x 36 núcleos y 384 GB de RAM en un entorno local. Además, cada host usa máquinas virtuales Windows Server de 96x2 vCPU y 4 GB de RAM con un uso medio de CPU bajo.

Ahora puede lograr niveles de consolidación mucho más altos si mueve sus máquinas virtuales a hosts dedicados T3, que tienen el doble de RAM en comparación con los hosts locales actuales. VMware Puede poner en marcha la misma cantidad de servidores en hosts dedicados T3 con un 50 % menos de costo de alojamiento. Esto puede ser útil para reducir los costos asociados al uso de licencias de Windows Server en un 33 %. En la siguiente tabla se destacan los ahorros derivados del uso de hosts dedicados T3.

	Hosts locales VMware	Hosts dedicados T3	Ahorros
Servidores físicos	10	5	
Núcleos físicos por host	36	48	
RAM por host (GB)	384	768	
2 vCPU, 4 GB de RAM por host VMs	96	192	
Número total de VMs	960	960	
Número total de licencias de la edición Datacenter de Windows Server (por núcleo) = (número de servidores * recuento de núcleos físicos)	$10 * 36 = 360$	$5 * 48 = 240$	33%

Comparación de la tenencia compartida con hosts dedicados con BYOL de SQL Server

Eche un vistazo a este ejemplo práctico para ver el valor de los hosts dedicados de Amazon EC2. En este escenario, una organización pone en marcha una carga de trabajo de SQL Server en un entorno en las instalaciones con 240 núcleos y desea implementar la misma carga de trabajo de forma rentable en AWS. Si esta organización usa sus propias licencias (BYOL), seguirá pagando por SA, y reducir el número de núcleos repercute directamente en sus costos.

En el siguiente diagrama se comparan los AWS ahorros entre los derechos de Microsoft y SQL Server.

Microsoft entitlements (Enterprise Agreements)		SQL Server savings with AWS	
	Number of cores	AWS shared vCPUs	AWS BYOL/Dedicated Hosts cores
SQL Server Enterprise edition	208	120	96
SQL Server Standard edition	32	20	-
Total SA cost	\$341,000	\$197,418	\$151,355

Al dimensionar correctamente las instancias en régimen de arrendamiento AWS compartido, puede reducir las licencias de SQL Server a 140 núcleos. Esto se traduce en un costo de SA de 197 000 USD.

Los hosts dedicados de Amazon EC2 le permiten otorgar la licencia de SQL Server a los núcleos físicos. Esto no es posible en la tenencia compartida, donde las licencias de SQL Server se basan en la cantidad de v CPUs asignada a la instancia. Por lo tanto, si utiliza dos hosts dedicados R5 con 48 núcleos cada uno, solo necesitará cubrir 96 núcleos en lugar de los 140 v que CPUs se requieren en un arrendamiento compartido. Al implementar hosts dedicados R5 y licenciar los núcleos físicos de las cargas, puede reducir la cantidad requerida de licencias de la edición Enterprise de SQL Server a 96 núcleos. Esto significa que puede implementar hasta 192 núcleos (teniendo en cuenta el hipersubprocesamiento) de cargas de trabajo de SQL Server y, al mismo tiempo, cumplir con los requisitos de licencia y lograr importantes ahorros de costos.

En este caso, la organización paga aproximadamente 341 000 USD al año en concepto de costos de SA. Tras dimensionar correctamente en un régimen de tenencia compartida, se reducen los costos a 197 000 USD con 140 vCPU. Los hosts dedicados de Amazon EC2 reducen aún más los costos hasta alcanzar los 151 000 USD (una disminución de aproximadamente un 56 %).

Implementaciones de SQL Server de alta disponibilidad

En este ejemplo, se analiza cómo el costo puede influir en una implementación de SQL Server en AWS teniendo en cuenta diversos factores relativos a las licencias. Supongamos que una organización necesita implementar seis servidores SQL Server Enterprise AWS para admitir tres aplicaciones. Estos servidores requieren una alta disponibilidad y tienen 16 v CPUs y 256 GB de RAM cada uno. Consulte los detalles del escenario siguiente:

- Servidor: SQL Server
- Edición de sistema operativo: edición Datacenter de Windows Server 2019
- Edición de SQL Server: edición Enterprise de SQL Server 2019
- vCPU: 16
- Memoria (GB): 256
- Cantidad: 6

Para optimizar los costos AWS sin sacrificar el rendimiento, le recomendamos que asigne el tamaño adecuado a las instancias en función del uso de la CPU, la memoria, la red y el disco (IOPS/BW). Tras dimensionar correctamente las cargas de trabajo, colóquelas en el tipo de instancia x2iedn.4xlarge, que ofrece 16 v. CPUs Sin embargo, este tipo de instancia también incluye el doble de la memoria necesaria para las cargas de trabajo. Aún es posible realizar una mayor optimización.

Escenario 1

Una organización implementa seis servidores SQL Server Enterprise en régimen de arrendamiento AWS compartido mediante la opción de licencia incluida para Windows y SQL Server. Con esta opción, el costo de las licencias de Windows y SQL Server se incorpora al precio de la instancia. Consulte los detalles del escenario siguiente:

- Tenencia compartida (instancia): x2iedn.4xlarge
- Costo por hora (USD): 10,0705 USD
- Costo mensual por unidad (USD): 7351,47 USD
- Número de servidores: 6
- CPU: 16
- Memoria: 512
- Costo mensual de 6 servidores: 44 108 USD

Escenario 2

Una organización tiene SA y BYOL para SQL Server en régimen de tenencia compartida. Esto significa que la organización usa la opción con licencia incluida para Windows, pero proporciona sus propias licencias de SQL Server en función del número de v asignado a la instancia. CPUs Como la organización tiene seis servidores SQL Server Enterprise con 16 v CPUs cada uno, CPUs se requiere un total de 96 v. Consulte los detalles del escenario siguiente:

- Tenencia compartida (instancia): x2iedn.4xlarge
- Costo por hora (USD): 4,0705 USD
- Costo mensual por unidad (USD): 2971,47 USD
- Número de servidores: 6
- CPU: 16
- Memoria: 512
- Núcleos BYOL: 96
- Costo mensual de 6 servidores: 17 828 USD

Al usar sus propias licencias de SQL Server con SA, la organización en este escenario puede ahorrar costos en comparación con el uso de la opción de SQL Server con licencia incluida. El ahorro de costos preciso depende del precio y las condiciones del acuerdo de licencia específico. En este escenario, AWS los costes se reducen en 26 280\$ al mes al adquirir licencias de SQL Server Enterprise. AWS

Escenario 3

Una organización cuenta con BYOL para Windows y SQL Server en hosts dedicados de Amazon EC2. Esto significa que la organización asignará las licencias a los núcleos físicos, lo que le permitirá licenciar únicamente los núcleos físicos del host La concesión de licencias a los núcleos físicos le permite implementar el número máximo de instancias sin que ello afecte a las licencias requeridas. Este modelo de licencia se usa comúnmente con la edición Datacenter de Windows Server y la edición Enterprise de SQL Server.

En este escenario se utilizan dos hosts dedicados X2iezn de Amazon EC2. Cada host tiene 24 núcleos físicos y 48 V. CPUs Esto proporciona una capacidad adecuada para los seis servidores SQL Server Enterprise con 16 v CPUs y 256 GB de RAM cada uno. Consulte los detalles del escenario siguiente:

- Número de hosts dedicados: 2
- Familia de instancias: x2iezn
- Costo por hora (USD): 11,009 USD
- Costo mensual por unidad (USD): 8036 USD
- Núcleos físicos: 48
- vCPU disponible: 96
- Licencias básicas de Windows Server requeridas: 24
- Licencias requeridas para los núcleos de la edición Enterprise de SQL Server: 24
- Costo mensual: 16 073 USD

El costo total de dos hosts dedicados Amazon EC2 de la familia X2iezn es de 16 073 USD al mes. Para obtener más información sobre los precios, consulte la Calculadora de precios de AWS [estimación](#) de este escenario. En este escenario, la organización puede ahorrar 1755,65 USD al mes si usa sus propias licencias de Windows. Si utiliza hosts dedicados de Amazon EC2, también puede reducir el número de licencias de SQL Server necesarias. En el régimen de tenencia compartida, necesitaría 96 licencias de la edición Enterprise de SQL Server para respaldar los seis servidores con la edición Enterprise de SQL Server con 16 vCPU cada uno. Sin embargo, al utilizar hosts dedicados de Amazon EC2 y otorgar licencias a los núcleos físicos, puede reducir el número de licencias necesarias a 48 núcleos.

En los siguientes detalles se comparan los costos del ejemplo 3 y se muestra cuánto puede ahorrar si implementa cargas de trabajo en hosts dedicados de Amazon EC2 con la opción BYOL en comparación con otros escenarios.

- Servidor en las instalaciones: SQL Server
- vCPU: 16
- Memoria: 256 GB
- Número de servidores: 6
- Costo mensual del escenario 1: Windows (licencia incluida) + Edición Enterprise de SQL Server (licencia incluida): 44 108 USD
- Costo mensual del escenario 2: Windows (licencia incluida) + Edición Enterprise de SQL Server (BYOL): 17 828 USD
- Costo mensual del escenario 3: Windows (licencia incluida) + Edición Enterprise de SQL Server (BYOL) en un host dedicado de Amazon EC2: 16 073 USD

Note

El costo se basa en los precios bajo demanda. Puede reducir aún más los costos mediante el uso de Savings Plans o instancias reservadas dedicadas. Estas opciones ofrecen un modelo de precios flexible con un ahorro de costos significativo en comparación con los precios bajo demanda. Estos planes ofrecen compromisos de uno o tres años. Para obtener más información, consulte la sección [Optimización del gasto de Windows en Amazon EC2](#) de esta guía.

Tenga en cuenta las siguientes opciones de pago para los hosts dedicados de Amazon EC2:

- [Hosts dedicados](#) (documentación de Amazon EC2)
- [Reservas de hosts dedicados](#) (documentación de Amazon EC2)
- [Savings Plans](#) (documentación de Amazon EC2)

[Calculadora de precios de AWS](#) ahora es compatible con los precios de los hosts dedicados. Esto puede ser de ayuda para elegir el host dedicado subyacente adecuado.

Recomendaciones de optimización de costos

Le recomendamos que siga estos pasos para optimizar sus costos mediante AWS Cost Explorer:

1. [Habilitar Cost Explorer](#).
2. Utilice Explorador de costos para [ver y analizar los costos y el uso](#) de sus implementaciones en hosts dedicados de Amazon EC2.
3. Compruebe que esté usando el sistema BYOL. Puede ver los siguientes detalles de la plataforma y valores de las operaciones de uso en las páginas de instancias o AMI de la consola de Amazon EC2, o en la respuesta que devuelven los comandos `describe-images` o `describe-instances`.
 - Detalles de la plataforma: Windows, operación de uso ::0002 RunInstances (licencia incluida)
 - Detalles de la plataforma: Windows BYOL, Operación de uso: :0800 RunInstances

Recursos adicionales

- [Eligible license types for license type conversion](#) (documentación de AWS License Manager)

- [AWS License Manager y taller exclusivo para anfitriones](#) (taller)AWS License Manager
- [Hosts dedicados de Amazon EC2](#) (documentación de FAQs AWS)
- [Cómo crear Windows Server Bring-Your-Own-License AMLs de forma local con VM Import/Export](#) (Microsoft Workloads en el blog) AWS
- [Importación/exportación de máquinas virtuales \(documentación\)](#)AWS
- [Amazon Web Services y Microsoft: Preguntas frecuentes](#) (AWS documentación)
- [Conversiones de tipos de licencia en License Manager](#) (AWS License Manager documentación)
- [Implementación de SQL Server de alta disponibilidad en hosts dedicados de Amazon EC2](#)AWS (blog sobre operaciones y migraciones en la nube)

Optimización del gasto de Windows en Amazon EC2

Descripción general de

Una de las principales preocupaciones a la hora de migrar servidores son los costes de infraestructura. AWS Es cierto que uno de los beneficios de la nube es pagar por los recursos que pida bajo demanda, pero hay cargas de trabajo de producción que deben estar disponibles las 24 horas del día, los 7 días de la semana, los 365 días del año. Los [Savings Plans](#) están diseñados para ahorrar dinero en el AWS uso de estado estable en todas las instancias EC2, y AWS Lambda. AWS Fargate

Los Savings Plans ofrecen un modelo de precios flexible y pueden ayudarle a reducir los precios del uso de Amazon EC2, Fargate, Lambda y SageMaker Amazon AI a cambio de un compromiso con una cantidad constante de uso (por ejemplo, 10\$ la hora). Se comprometes a una cantidad constante de gastos computacionales por hora durante uno o tres años y, a cambio, recibe un descuento por ese uso.

Puede elegir entre tres opciones de pago diferentes con Savings Plans:

- La opción sin pago por adelantado no requiere ningún pago por adelantado y su compromiso se cobra únicamente de forma mensual.
- La opción pago parcial por adelantado ofrece precios más bajos en los Savings Plans. Se le cobrará al menos la mitad de su compromiso por adelantado y el resto se cobrará mensualmente.
- La opción pago total por adelantado ofrece los precios más bajos y todo el compromiso se cobra en un solo pago.

Puede realizar el seguimiento de los vencimientos de sus planes de Savings Plans, así como de los planes futuros que se vayan a adquirir en AWS Cost Explorer. Puede usar las alertas de Savings Plans para recibir alertas anticipadas por correo electrónico 1, 7, 30 o 60 días antes de la fecha de vencimiento de su plan de ahorro o cuando haya un compromiso en espera para su compra. Estas notificaciones también le alertan sobre la fecha de vencimiento. Puede enviar notificaciones a un máximo de 10 destinatarios por correo electrónico.

Descripción de Savings Plans

Cada tipo de uso de computación tiene una tarifa bajo demanda y una tarifa de Savings Plans. Si asume un compromiso de 10 USD por hora de computación, obtendrá precios de Savings Plans para todo el uso de hasta 10 USD de acuerdo con la tarifa de Savings Plans. Cualquier uso que supere este compromiso se cobrará según las tarifas bajo demanda normales. Puede empezar a utilizar Savings Plans con Explorador de costos en Consola de administración de AWS.

Puede comprometerse fácilmente con Savings Plans utilizando las recomendaciones que se proporcionan en [Explorador de costos](#) para obtener los mayores ahorros. El compromiso por hora recomendado se basa en su historial de uso bajo demanda y en el tipo de plan, duración y opción de pago que haya elegido. Savings Plans se aplica primero a la cuenta que compró el plan y, después, se comparte con otras cuentas de la familia de facturación unificada.

Note

La opción de compartir en Savings Plans AWS Organizations está habilitada de forma predeterminada. Puedes rechazar esta opción en la AWS Billing consola de la cuenta del pagador. Puedes visitar la página de [recomendaciones](#) para ver los Savings Plans que AWS recomiendan ayudarte a ahorrar en el uso que cumpla con los requisitos. Estas recomendaciones se pueden actualizar en cualquier momento para facilitarle la compra de los planes de Savings Plans óptimos que le permitan ahorrar.

Savings Plans para computación

Savings Plans para computación ofrece el máximo de flexibilidad y le permite reducir los costos. Estos planes se aplican automáticamente al uso de instancias de EC2, independientemente de la familia de instancias, el tamaño, la zona de disponibilidad, la región, el sistema operativo o la tenencia. También se aplican al uso de Fargate y Lambda. Por ejemplo, con Savings Plans para computación, puede cambiar de instancias C4 a M5, cambiar una carga de trabajo de la región UE

(Irlanda) a la región UE (Londres) o transferir una carga de trabajo de EC2 a Fargate o Lambda en cualquier momento. Seguirá pagando automáticamente el precio de Savings Plans.

Savings Plans para instancias de EC2

Savings Plans para instancias de EC2 ofrece los descuentos más significativos a cambio de comprometerse a usar familias de instancias individuales en una región (por ejemplo, comprometerse a un nivel constante de uso de M5 en la región Norte de Virginia). Esto le ofrece descuentos automáticamente en el precio bajo demanda de la familia de instancias seleccionada en esa región, independientemente de la zona de disponibilidad, el tamaño, el sistema operativo o la tenencia. Savings Plans para instancias de EC2 le permite cambiar su uso entre las instancias de una familia de esa región. Por ejemplo, puede pasar de c5.xlarge con Windows a c5.2xlarge con Linux y beneficiarse automáticamente de los precios de Savings Plans.

Tanto Savings Plans para instancias de EC2 como Savings Plans para computación se aplican a las instancias de EC2 que forman parte de los clústeres de Amazon EMR, Amazon Elastic Kubernetes Service (Amazon EKS) y Amazon Elastic Container Service (Amazon ECS). Savings Plans no incluye los cargos de Amazon EMR, Amazon EKS y Amazon ECS, pero sí los de las instancias de EC2 subyacentes. Los Savings Plans para instancias de EC2 se aplican antes que los Savings Plans para computación porque los Savings Plans para computación pueden aplicarse de manera más amplia.

Note

No puede cambiar fácilmente un plan de Savings Plans después de haber asumido un compromiso. Le recomendamos que planifique detenidamente antes de comprometerse con cualquiera de las opciones de Savings Plans. Savings Plans ofrece precios más bajos en comparación con los precios bajo demanda a cambio de un compromiso y no se pueden cancelar durante la vigencia de dicho compromiso.

Ejemplo de compromiso por hora

Si adquiere un plan de Savings Plans, asume un compromiso monetario por hora durante el plazo del plan. Si asume un compromiso de 10 USD por hora de computación, la tarifa de Savings Plans se aplicará de forma automática a todo el uso que alcance esos 10 USD por hora. Cualquier uso que supere este compromiso se cobrará según las tarifas bajo demanda habituales. Puede usar la herramienta de recomendaciones de compra de Savings Plans en Explorador de costos para obtener compromisos recomendados que puedan maximizar sus ahorros. El compromiso financiero por

hora de un plan específico no se puede modificar durante la vigencia del plan. Si desea aumentar el compromiso después de analizar el uso, puede adquirir un plan de Savings Plans adicional para abordar el exceso de uso.

Beneficios de Savings Plans

En comparación con las instancias reservadas, los planes de Savings Plans ofrecen un modelo de precios más flexible que le permite ahorrar dinero y, al mismo tiempo, aprovechar la amplia selección de opciones computacionales de Savings Plans. Savings Plans ofrece descuentos, incluso aunque sus necesidades computacionales cambien. Esto puede ser útil para seguir el ritmo de su dinámico y cambiante entorno sin incurrir en gastos de administración adicionales. Estos son algunos otros beneficios de usar Savings Plans:

- Fácil de usar: reciba descuentos automáticos a cambio de un compromiso monetario.
- Flexibilidad: un compromiso único que se aplica a varios tipos de uso.
- Ahorros potenciales: hay varias formas de ahorrar. Considere los siguientes ejemplos:
 - Ahorro del 60 % en cargas de trabajo de Windows Server con Savings Plans para computación ([d2.8xlarge, 3 años, pago total adelantado, Windows, tenencia compartida, us-east-2](#))
 - Ahorro del 73 % en cargas de trabajo de Windows Server con Savings Plans para instancias de EC2 ([d2.8xlarge, 3 años, pago total por adelantado, Windows, tenencia compartida, us-east-2](#))
 - Entre un 28 % y un 41 % de ahorro en tipos de instancias no exóticas ([familia t3, 3 años, pago total por adelantado, Windows, tenencia compartida, us-east-2](#))
 - Entre un 25 % y un 40 % de ahorro promedio en servidores Windows

Note

Savings Plans para instancias de EC2 ofrece un descuento mayor que Savings Plans para computación debido a su flexibilidad reducida. Se compromete a hacer un uso determinado a cambio de un precio reducido.

Cada tipo de uso de computación tiene una tarifa de Savings Plans y una tarifa bajo demanda. La siguiente tabla muestra los planes de Savings Plans y las tarifas bajo demanda para cada tipo de sistema operativo. Se le cobrarán las tarifas de Savings Plans por el uso al que se haya comprometido y cualquier uso que supere el compromiso se cobrará según las tarifas bajo demanda habituales.

Nombre de instancia	Tarifa de Savings Plan	Ahorros bajo demanda	Tarifa bajo demanda	Sistema operativo	Region	Opción de pago	Duración del compromiso
x2iedn.xlarge	0,32\$	61%	0,83\$	Linux	Este de EE. UU. (Norte de Virginia)	Sin pago inicial	3
x2iedn.xlarge	2,01 DÓLARES	50%	1,02 DÓLARES	Windows	Este de EE. UU. (Norte de Virginia)	Sin pago inicial	3
x2iedn.xlarge	1,02\$	20%	2,52 DÓLARES	Windows con licencia incluida + Edición Enterprise de SQL Server	Este de EE. UU. (Norte de Virginia)	Sin pago inicial	3
x2iedn.xlarge	0,32\$	61%	0,83\$	BYOL	Este de EE. UU. (Norte de Virginia)	Sin pago inicial	3

Savings Plans incluye el sistema operativo y ofrece un descuento aparte para BYOL. Todo esto aparece desglosado en la [calculadora de Savings Plans para computación](#).

Modelo de precios de instancias reservadas

AWS tiene otro modelo de precios basado en el compromiso conocido como instancias reservadas. Este modelo puede resultar problemático si su proceso cambia después de haber asumido un compromiso, lo que provoca que las instancias reservadas no se utilicen. Los planes de

Savings Plans están diseñados para ofrecer reducciones de costos similares a las de las [instancias reservadas estándar y convertibles](#), pero con una flexibilidad mucho mayor. Savings Plans para computación ofrece precios más bajos en el uso de instancias de EC2, independientemente de la familia de instancias, el tamaño, el sistema operativo, la tenencia o la región. También ofrece la máxima flexibilidad.

La siguiente tabla puede ser útil para elegir entre Savings Plans o instancias reservadas.


	Reserved Instance	Savings Plans para instancias de EC2	Savings Plans para computación
Descuento promedio de 1 año	Hasta un 38 %	Hasta un 29 %	Hasta un 29 %
Descuento promedio de 3 años	Hasta un 58 %	Hasta un 73 %	Hasta un 60 %
Familia de instancias	Fixed	Fixed	Flexible
Tamaño de instancia	Fija (no Linux)	Flexible	Flexible
Geography	1 región	1 región	Flexible
Sistema operativo	Fixed	Flexible	Flexible
Servicio	Amazon EC2 o Amazon RDS	Amazon EC2	Amazon EC2, Fargate o Lambda
Opciones de pago	Pago total, parcial, sin pago por adelantado	Pago total, parcial, sin pago por adelantado	Pago total, parcial, sin pago por adelantado
Límites de instancias	20 por zona de disponibilidad	Sin límite	Sin límite

Note

Savings Plans funciona ofreciéndole un descuento basado en un compromiso monetario por hora. El compromiso financiero por hora no se puede cancelar ni cambiar durante la vigencia de su plan, pero puede adquirir planes de Savings Plans adicionales para abordar el uso

adicional. Esto le permite mantener un compromiso horario constante a medida que su flota crece.

Puede utilizar herramientas como [AWS Cost Explorer](#) o [Nube de AWS Intelligence Dashboards](#) para hacer un seguimiento de su compromiso. Explorador de costos proporciona una línea de objetivo de cobertura que puede ayudar a su organización a planificar su estrategia de cobertura de Savings Plans. Si el 75 % de su carga de trabajo es estable, el 75 % es un buen objetivo. Esto deja el 25 % para gastos bajo demanda o variables en función de las cargas de trabajo dinámicas. Si necesita aumentar esa cobertura al 85 %, puede adquirir otro compromiso de Savings Plans para aumentar el compromiso monetario por hora.

 Note

Le recomendamos que adquiera planes de Savings Plans en lugar de instancias reservadas, pero los dos modelos de compromiso pueden funcionar juntos si ya adquirió instancias reservadas.

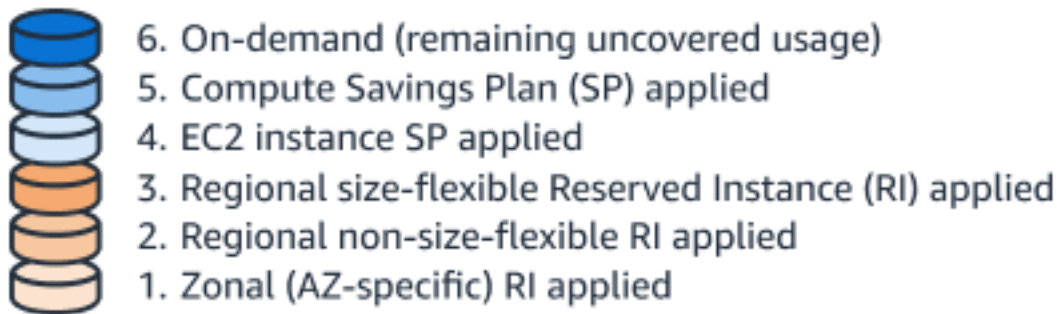
Imagine que compró una instancia reservada, pero quiere empezar a probar una opción de Savings Plans. Es lógico que esta combinación se aplique a su facturación final. Esta es una jerarquía que puede aplicar a sus Cuentas de AWS:

1. La instancia reservada zonal se aplica a la cuenta que la posee. Si a una instancia reservada le quedan horas, se aplica al resto de la organización.
2. Las instancias reservadas regionales para Windows de tamaño no flexible se aplican únicamente al uso coincidente de la cuenta titular. Todo lo que quede se transfiere al resto de la organización.
3. Las instancias reservadas regionales con tamaño flexible se aplican a la cuenta propietaria (primero a la instancia más pequeña de la familia y luego a las instancias más grandes) y, después, al resto de la organización.
4. Las instancias reservadas regionales se aplican a cualquier reserva de capacidad bajo demanda no utilizada.
5. Savings Plans para instancias de EC2 se aplica a la cuenta que adquirió el plan.
6. Savings Plans para computación se aplica a la cuenta que adquirió el plan.

Note

Los descuentos comienzan con el uso que se traduce en el descuento más alto y luego van bajando hasta el descuento mínimo. Tradicionalmente, las instancias de Windows tenían un potencial de descuento más bajo que las de Linux para los tipos de instancias más comunes (por ejemplo, T3, M6 y C5). Esto significa que las instancias de Linux se benefician más que las instancias de Windows en la mayoría de los casos.

El siguiente gráfico muestra el precio después de dividir las instancias reservadas de Savings Plans. Tanto los planes Savings Plans para instancias de EC2 como Savings Plans para computación se aplican primero a las instancias en marcha y, después, a las reservas de capacidad bajo demanda no utilizadas.



Escenarios de optimización de costos

En esta sección se describen los escenarios de optimización de costos para los hosts dedicados de Amazon EC2 y las instancias de Amazon EC2 que utilizan un modelo de facturación con licencia incluida.

Hosts dedicados de Amazon EC2

Imagine que va a migrar sus cargas de trabajo de Windows en las instalaciones a AWS. Su centro de datos tiene los siguientes servidores:

- Dos servidores con 16 vCPU y 128 GB de RAM
- Dos servidores con 32 vCPU y 164 GB de RAM
- Un servidor con 8 vCPU y 64 GB de RAM
- 16 servidores con vCPU y 32 GB de RAM

Además, suponga que puede traer su propia licencia AWS porque tiene suficientes licencias para traerla. En la siguiente tabla, se muestran las instancias de servidor que puede utilizar en AWS.

Tipo de instancia	CPU	RAM	Amount
r5.4xlarge	16	128	2
r5.8xlarge	32	256	2
r5.2xlarge	8	64	1
r5.xlarge	4	32	16
			21

Un análisis muestra que estas 21 máquinas virtuales se pueden distribuir en dos hosts dedicados con un host de la familia de instancias R5. En la siguiente tabla, se muestra el costo de estos dos hosts dedicados.

Escenario de host dedicado bajo demanda	Pago por adelantado	1 mes	1 año	3 años	Calculadora de precios de AWS
Bajo demanda	Ninguno	10.123 DÓLARES	121.475 DÓLARES	364.392 DÓLARES	Calculadora de precios de AWS estimación
Savings Plans de 1 año	Ninguno	7.447 DÓLARES	89.362 DÓLARES	–	Calculadora de precios de AWS estimación
Savings Plans de 3 años	Ninguno	5.476 DÓLARES	65.712 DÓLARES	197.128 DÓLARES	Calculadora de precios

Escenario de host dedicado bajo demanda	Pago por adelantado	1 mes	1 año	3 años	Calculadora de precios de AWS
					de AWS estimación
Savings Plans de 3 años con pago por adelantado	84.438 DÓLARES	2.755 DÓLARES	117.499 DÓLARES	183.618 DÓLARES	Calculadora de precios de AWS estimación

Si tiene los servidores a los que quiere migrar AWS, el precio final de los Savings Plans de 1 año es de 89.362\$, en lugar de los 121.475\$ del precio bajo demanda. Esto representa un descuento del 26,5 % después de un año. Si está pensando en quedarse en casa AWS por un período más largo, puede elegir los planes Savings Plans de 3 años para obtener ahorros de costos aún mayores. Al cabo de tres años, pagará 197 128 USD en lugar de 364 392 USD. Esto se traduce en un ahorro del 46 % del monto total al cabo de tres años.

Instancias de Amazon EC2 con licencias incluidas

Considere un escenario en el que vaya a migrar una sola aplicación de tres niveles y desee utilizar las licencias que proporciona. AWS AWS Además, suponga que su aplicación funciona con los siguientes servidores:

- Dos servidores web con dos v CPUs y 4 GB de RAM
- Dos servidores de aplicaciones con ocho v CPUs y 16 GB de RAM
- Dos servidores de bases de datos con 16 v CPUs y 64 GB de RAM (con la edición SQL Server Standard)

En la siguiente tabla, se muestran las instancias de servidor que puede utilizar en AWS.

Tipo de instancia	CPU	RAM	Amount
c5.large	2	4	2

Tipo de instancia	CPU	RAM	Amount
c5.2xlarge	8	16	2
r5.2xlarge	8	64	2
			6 servidores

En la siguiente tabla se muestra el costo de estos servidores en AWS.

Licencia incluida por AWS	Pago por adelantado	1 mes	1 año	3 años	Calculadora de precios de AWS
Bajo demanda	Ninguno	3.912 DÓLARES	46.950 DÓLARES	140.849 DÓLARES	Calculadora de precios de AWS estimación
Savings Plans de 1 año	Ninguno	3.466 DÓLARES	41.952 DÓLARES		Calculadora de precios de AWS estimación
Savings Plans de 3 años sin pago por adelantado	Ninguno	3.189 DÓLARES	38.264 DÓLARES	114.804 DÓLARES	Calculadora de precios de AWS estimación
Savings Plans de 3 años con pago por adelantado	112.110 DÓLARES	Ninguno	Ninguna	Ninguno	Calculadora de precios de AWS estimación

Si desea utilizar estos servidores para entornos de producción (24 horas al día, 7 días a la semana) con precios bajo demanda, tendrá que pagar un costo mensual de 3912 USD. El pago de este costo mensual equivale a 46 950 USD al cabo de un año y a un total de 140 849 USD al cabo de tres años.

Si elige los planes de Savings Plans de 1 año sin pago por adelantado, el costo mensual se reduce a 3466 USD. Al final del primer año, paga 41 952 USD. Se trata de un descuento total del 11 %. Si elige los planes de Savings Plans de 3 años sin pago por adelantado, el costo mensual se reduce a 3189 USD. Al cabo de tres años, pagará 114 804 USD. Eso supone un ahorro del 18,5 %.

Recomendaciones de optimización de costos

Ambos escenarios le permiten ahorrar dinero a la hora de planificar y predecir sus cargas de trabajo en AWS. Es importante reconocer que el descuento en el segundo escenario es menor en comparación con el primer escenario. En el segundo escenario, el precio de la licencia está incluido en el precio del servidor en la nube. AWS no ofrece descuentos en el precio de la licencia, pero siempre puedes traer tus licencias (en situaciones específicas) y siempre AWS puedes garantizar el mejor compute/instance precio.

Le recomendamos que haga lo siguiente para controlar su AWS gasto en recursos informáticos y de instancia:

- Acceso a las recomendaciones
- Personalizar las recomendaciones según sus necesidades
- Revisión del compromiso por hora

Acceso a las recomendaciones

Puede utilizar la [consola de Amazon EC2](#) para acceder a las recomendaciones de Savings Plans. Puede incluso descargar sus recomendaciones de Savings Plans en formato CSV para revisarlas más adelante. Para obtener más información, consulte [Monitoring your Savings Plans](#) en la documentación de Savings Plans.

Personalización de las recomendaciones según sus necesidades

Abra la [consola de Amazon EC2](#), expanda la sección Instancias y, a continuación, seleccione Savings Plans. En esta página, se muestran los precios de las instancias y los cálculos antes y después de hacer una recomendación. También puede ajustar los siguientes factores de las recomendaciones:

- Vigencia: por ejemplo, de 1 a 3 años
- Opción de pago: pago por adelantado, pago parcial por adelantado o sin pago por adelantado
- Historial: por ejemplo, los últimos 7, 30 o 60 días

Revisión del compromiso por hora

Con el mismo ejemplo, supongamos que tiene una instancia que funciona las 24 horas del día, los 7 días de la semana. La recomendación es utilizar Savings Plans. Según el tamaño, tiene un precio bajo demanda de 120 USD/hora. Tiene la opción de asumir un compromiso de 90 USD por hora, pero esto puede variar en función de la región, la instancia y la opción de compra. En este ejemplo, puede ahorrar un 25 % en comparación con el costo bajo demanda. También puede hacer un seguimiento de su uso y cobertura, si están por debajo del umbral que definió, y configurar una alerta cuando se acabe el presupuesto.

Revisión de recomendaciones

Le recomendamos que revise detenidamente las recomendaciones de Savings Plans. AWS no cambiará nada sin su permiso. Son solo recomendaciones y depende de usted aplicarlas o no.

Compra de un plan

Abra la [consola de Amazon EC2](#), expanda la sección Instancias y, a continuación, seleccione Savings Plans. Luego, elija Adquiera Savings Plans. Según sus requisitos, puede seleccionar las siguientes opciones: vigencia, región, familia de instancias, compromiso por hora, opción de pago e incluso fecha de inicio. Puede elegir entre Compute Savings Plans, EC2 Instance Savings Plans SageMaker y AI Savings Plans. Para obtener más información, consulte [Purchasing Savings Plans](#) en la documentación de Savings Plans.

Generación de un informe de uso

Después de adquirir un plan de Savings Plans, puede obtener un informe de uso. El informe le permite comprobar su utilización, comprobar si el plan adquirido es suficiente para abordar y maximizar el descuento, y cancelar o agregar nuevos descuentos. Este informe se puede exportar a otros formatos, como CSV. Para obtener más información, consulte [Using the utilization report](#) en la documentación de Savings Plans.

Cumplimiento las prácticas recomendadas de compra

Le recomendamos que siga estas prácticas recomendadas antes de adquirir un plan de Savings Plans:

- Use [AWS Trusted Advisor](#) para eliminar los recursos de EC2 inactivos.
- Dimensione correctamente antes de adquirir planes de Savings Plans.

- Establezca una tarifa por hora que mantenga de manera constante durante 30 a 60 días.
- Adquiera un compromiso para abordar la mayor parte de la tarifa por hora constante que su organización pueda asumir. Tenga en cuenta las fluctuaciones de la demanda o de la temporada.
- Elija un presupuesto de Savings Plans revisado trimestralmente para mantener una tarifa constante (por ejemplo, un objetivo de cobertura del 70 % para la cobertura de Savings Plans). Si la tarifa cae por debajo de la cobertura deseada, adquiera un plan de Savings Plans adicional como compensación para cumplir con su objetivo de cobertura.

Recursos adicionales

- [Savings Plans for Amazon EC2 Reserved Instances \(documentos técnicos\)](#) AWS
- [Entender cómo se aplican los Savings Plans a su AWS consumo](#) (documentación de Savings Plans)
- [Anunciamos la facturación por segundo de las instancias EC2 de Windows Server y SQL Server](#) (AWS documentación)
- [AWS Serie sobre optimización de costes: vídeo sobre Savings Plans | Amazon Web Services](#) (YouTube)

Supervise los costes mediante herramientas AWS

Descripción general de

La visibilidad de los costos es un factor clave para optimizar los costos en AWS. AWS tiene una serie de herramientas que puede utilizar para visualizar los costos y crear alertas en respuesta a esos costos. Estas incluyen herramientas que, por ejemplo AWS Budgets, te ayudan a llevar un registro de tus gastos y a reportarlos. En esta sección se describen formas específicas de supervisar sus gastos de Windows en AWS , de modo que pueda hacer un seguimiento de sus necesidades presupuestarias y reaccionar en consecuencia. Esto incluye agregar las etiquetas necesarias a sus recursos de Windows EC2. Estas etiquetas le permiten supervisar correctamente Windows EC2 y otros servicios de Microsoft con AWS Budgets.

Al monitorear los gastos y crear alertas con AWS herramientas, puede estar más informado sobre los gastos actuales, los gastos proyectados y las anomalías en los gastos. Si utiliza [Savings Plans](#) para reducir el precio por hora de sus instancias de EC2, le recomendamos que consulte la utilización y la cobertura generales de Savings Plans. Esto puede ser útil para no dejar de ahorrar. Puede usarlo

AWS Cost Explorer para ver el inventario de Savings Plans y obtener recomendaciones de planes de ahorro adicionales en función del uso anterior. También puede hacer un seguimiento de gastos específicos con [AWS Budgets](#) y configurando [AWS Cost Anomaly Detection](#).

Recomendaciones de optimización de costos

Le recomendamos que siga los siguientes pasos para optimizar sus costes mediante AWS Budgets el uso de Cost Explorer y la detección de anomalías:

- Etiquetado de los recursos de Windows EC2
- Configure alertas mediante AWS Budgets
- Activación de Detección de anomalías en los costos
- Obtenga un análisis de los gastos en tiempo real.
- Vea los gastos de los servicios de Windows con licencia incluida mediante Explorador de costos.

Etiquetado de los recursos de Windows EC2

Para supervisar sus AWS gastos de forma eficaz, debe establecer una [estrategia de etiquetado](#) para las cargas de trabajo que desee supervisar. Esto es importante para poder agrupar los recursos de forma categórica y recibir notificaciones sobre gastos específicos, a diferencia de los gastos de uso general. Puede utilizar recursos de etiquetado que no solo ayudan a reducir los costos, sino que también se pueden utilizar para otros fines, como [Automatización de AWS Systems Manager](#). Además, le recomendamos que implemente algunos controles de gestión para las [etiquetas necesarias](#).

Para realizar un seguimiento de sus gastos en AWS Budgets Cost Explorer y Cost Anomaly Detection, debe asegurarse de que las etiquetas adecuadas estén colocadas. Puede usar las etiquetas a fin de establecer un presupuesto específico para los elementos que coincidan con esas etiquetas, de modo que reciba alertas cuando aumenten los gastos.

Por ejemplo, puede usar una etiqueta sencilla, como Clave=SO Valor=Windows. De este modo, todas las instancias de Windows forman un grupo para facilitar el seguimiento de sus gastos. También puede utilizar etiquetas para otros elementos, como Systems Manager. Tras crear una etiqueta, debe activarla para hacer un seguimiento de los costos. Considere la posibilidad de añadir una [AWS Config regla que supervise las etiquetas](#) adjuntas a determinados recursos. AWS Config puede avisarle si hay recursos en ejecución que no contengan las etiquetas adecuadas, lo que le proporcionará una representación precisa de sus gastos en EC2 en Windows.


Una vez asignadas las etiquetas, podrá crear un presupuesto personalizado en AWS Billing. Así podrá ver sus gastos en Windows EC2. Puede establecer un presupuesto diario o mensual.

Configure alertas mediante AWS Budgets

En este escenario de ejemplo, crea un presupuesto diario para Windows EC2. Es un presupuesto recurrente que utiliza la opción de ajuste automático para hacer un seguimiento de sus gastos y ajustar el presupuesto en consecuencia. Si tiene un entorno estático, puede usar un presupuesto fijo en su lugar. Asegúrese de elegir un intervalo de tiempo de referencia (por ejemplo, 30 días).

1. Inicie sesión en la [AWS Cost Management consola Consola de administración de AWS y ábrala](#).
2. En el panel de navegación, elija Budgets (Presupuestos).
3. En la parte superior de la página, elija Create budget.
4. En Configuración del presupuesto, seleccione Personalización (avanzada).
5. En Tipos de presupuesto, seleccione Presupuesto de costos. A continuación, elija Siguiente.
6. En Detalles, en Nombre del presupuesto, ingrese el nombre del presupuesto. Por ejemplo, Gasto en Windows EC2.
7. En Establecer importe del presupuesto, en Periodo, seleccione Diario.
8. En Tipo de renovación de presupuesto, elija Presupuesto periódico para un presupuesto que se restablezca una vez finalizado su periodo.
9. En Fecha de inicio, elija la fecha o el periodo de inicio para comenzar el seguimiento del monto presupuestado.
10. En Método de presupuestos, seleccione Ajuste automático (Nuevo).
11. En Intervalo de tiempo de referencia, seleccione Intervalo personalizado y, a continuación, ingrese 30 días.
12. Elija Siguiente.
13. En la sección Alcance del presupuesto, selecciona Filtrar dimensiones de AWS coste específicas. Aquí es donde se utilizan las etiquetas para crear las dimensiones adecuadas. AWS Budgets no admite el tipo de plataforma como opción en sus filtros. Por este motivo, debe aplicar las etiquetas SO.
14. Elija Agregar filtro y, a continuación, seleccione la opción Etiqueta en Dimensiones.
15. Elija la etiqueta SO y, a continuación, elija el valor Windows para crear un presupuesto para la etiqueta.
16. Elija Siguiente.

- 17 En la página Configurar alertas, seleccione Agregar un umbral de alerta. Aquí puede configurar dos alertas: una para un umbral del 50 % y otra para un umbral del 100 %. Si se supera el umbral del 50 % antes de la mitad del mes, se emitirá una advertencia. De esta forma, podrá comprobar si sus gastos son superiores a los esperados y reaccionar antes de que acabe el mes.
- 18 En Umbral, ingrese 50 y seleccione % del importe presupuestado.
- 19 En Desencadenador, elija Real.
- 20 En Destinatarios de correo electrónico, ingrese una dirección de correo electrónico. Agregue otra alerta para un umbral de 100.

 Note

En este ejemplo, se utiliza una notificación por correo electrónico para la alerta, pero también se pueden utilizar otras estrategias, como [Slack](#).

Activación de Detección de anomalías en los costos

Puede usar sus etiquetas de costos para configurar alertas de gastos que sean una anomalía. Por ejemplo, puede usar [AWS Cost Anomaly Detection](#) para crear monitoreos y recibir alertas cuando el sistema detecte gastos anormales en la cuenta.

Para configurar un monitoreo y alertas para la etiqueta Clave=SO y Valor=Windows que creó anteriormente, haga lo siguiente:

1. Inicie sesión en la [AWS Cost Management consola Consola de administración de AWS y ábrala](#).
2. En el panel de navegación, elija Detección de anomalías en los costos.
3. Seleccione la pestaña Monitoreos de costos y, a continuación, seleccione Crear monitoreo.
4. En el paso 1, elija Etiqueta de asignación de costos como tipo de monitoreo.
5. En Clave de etiqueta de asignación de costos, elija Gastos de Windows EC2.
6. En Valor de etiqueta de asignación de costos, elija Windows.
7. En Asigne un nombre al monitoreo, ingrese Gastos de Windows EC2.
8. Elija Siguiente.
9. Para crear una suscripción para las alertas, seleccione Crear una nueva suscripción. Si tiene suscripciones existentes, seleccione Choose an existing subscription (Elegir una suscripción existente).

- 10 En Nombre de la suscripción, ingrese Anomalía en los gastos de Windows EC2.
- 11 En Frecuencia de las alertas, elija Resúmenes diarios.
- 12 En Destinatarios de la alerta, ingrese su dirección de correo electrónico.
- 13 Elija Agregar umbral. En Umbral, ingrese 10 y, a continuación, seleccione porcentaje por encima de la velocidad esperada.
- 14 Elija Crear monitor.

Visión en tiempo real de los gastos

Una alerta es una herramienta útil para supervisar los gastos de Windows EC2, pero debe utilizar Explorador de costos si quiere tener una visión en tiempo real de los gastos. Vea este video para saber cómo el Explorador de costos le permite analizar y reducir los costos de EC2. Para obtener más información, vea el vídeo [AWS Supports You | Understanding and Reducing Your EC2 Costs](#) que aparece YouTube aquí.

Visualización de los gastos de los servicios de Windows con licencia incluida

Puede ver los gastos de Windows EC2 en su cuenta mediante Explorador de costos. Para ver los gastos de los servicios de Windows con licencia incluida, debe configurar los siguientes [filtros](#) correctos en Explorador de costos:

- En Plataforma, elija Windows (Amazon VPC). Para el funcionamiento de la API, elija RunInstance: 0002. Este es el AWS Billing código para las instancias EC2 de Windows con licencia incluida.
- Si quiere ver sus gastos en instancias BYOL, cambie: 0002 a: 0800. RunInstance RunInstance Este es el código de facturación BYOL de Windows EC2.

Con toda esta información en Explorador de costos, puede filtrar rápidamente sus costos para determinar exactamente lo que gasta en Windows EC2. Si quieres profundizar aún más en tus AWS gastos, puedes filtrar los gastos AWS Cost and Usage Report a nivel de instancia individual. También puede generar informes que se pueden visualizar en Amazon Quick y crear paneles personalizados.

Para obtener más información, vea el vídeo [Le AWS apoyamos: cómo visualizar sus informes de costos y uso](#) en YouTube

Recursos adicionales

- [Cómo configurar las etiquetas necesarias con AWS Config](#) (AWS Config documentación)

- [AWS Budgets Tutorial: Configurar alertas para AWS Billing | Amazon Web Services](#) (YouTube)
- [AWS Cost and Usage Report Biblioteca de consultas](#) (AWS Well-Architected Labs)

SQL Server

Los clientes han estado ejecutando cargas de trabajo de Microsoft AWS durante más de 15 años, más tiempo que cualquier otro proveedor de nube. Esto se debe en gran parte a que AWS tiene más experiencia con las aplicaciones de Microsoft en la nube y ofrece la mejor plataforma para Windows Server y Microsoft SQL Server en las siguientes áreas:

- Mayor rendimiento y fiabilidad
- Mejores servicios de seguridad e identidad
- Mayor soporte para migraciones
- Las capacidades más amplias y detalladas
- Menor costo total de propiedad (TCO)
- Opciones de licencia flexibles

AWS es compatible con todo lo necesario para crear y ejecutar aplicaciones de Windows que se basan en SQL Server, incluidas Active Directory, .NET, SQL Server, el escritorio como servicio de Windows y todas las versiones compatibles de Windows Server. Con una experiencia comprobada, AWS puede ayudarlo a levantar y cambiar, refactorizar o incluso modernizar fácilmente sus cargas de trabajo de Windows.

Esta sección de la guía abarca los siguientes temas:

- [Elección de una solución de alta disponibilidad y recuperación ante desastres](#)
- [Explicación de las licencias de SQL Server](#)
- [Selección de la instancia de EC2 adecuada para las cargas de trabajo de SQL Server](#)
- [Unificación de instancias](#)
- [Comparación de las ediciones de SQL Server](#)
- [Evaluación de la edición Developer de SQL Server](#)
- [Evaluación de SQL Server en Linux](#)
- [Optimización de las estrategias de copia de seguridad de SQL Server](#)
- [Modernización de las bases de datos de SQL Server](#)
- [Optimización del almacenamiento para SQL Server](#)
- [Optimización de las licencias de SQL Server con Compute Optimizer](#)
- [Optimización del tamaño de SQL Server con Compute Optimizer](#)

- [Revise Trusted Advisor las recomendaciones para las cargas de trabajo de SQL Server](#)

Elección de una solución de alta disponibilidad y recuperación ante desastres

Descripción general de

Le recomendamos que diseñe una arquitectura para la implementación de SQL Server en AWS que se adapte a las necesidades de su empresa y, al mismo tiempo, cumpla con sus [objetivos de recuperación ante desastres](#), lo que abarca el objetivo de tiempo de recuperación (RTO) y el objetivo de punto de recuperación (RPO). Las siguientes soluciones pueden serle útiles para diseñar la arquitectura adecuada para SQL Server en Amazon Elastic Compute Cloud (Amazon EC2) y, al mismo tiempo, para optimizar los costos de las cargas de trabajo de SQL Server.

- Grupos de disponibilidad de SQL Server Always On: los grupos de disponibilidad de SQL Server Always On proporcionan alta disponibilidad y recuperación ante desastres (HA/DR) solutions for SQL Server databases. An availability group consists of a set of user databases that fail over together. Always On availability groups also provide redundancy at the database level, but don't require shared storage—each replica has its own local storage. You can deploy this feature as an HA/DRsolución). Para obtener más información, consulte [What is an Always On availability group?](#) en la documentación de Microsoft.
- Instancias de clúster de conmutación por error (FCI) de SQL Server Always On: SQL Server Always On FCIs utiliza el clúster de conmutación por error de Windows Server (WSFC) para proporcionar alta disponibilidad a nivel de instancia de SQL Server. FCIs requieren almacenamiento compartido para alojar bases de datos. Puede utilizar el almacenamiento en bloques compartido o el almacenamiento de archivos compartido. Por ejemplo, puede utilizar Amazon FSx for Windows File Server o Amazon FSx for NetApp ONTAP como solución de almacenamiento compartido con varias zonas de disponibilidad. Para obtener más información, consulte [Always On Failover Cluster Instances \(SQL Server\)](#) en la documentación de Microsoft.
- SIOS DataKeeper: SIOS DataKeeper puede ayudarlo a cumplir con los requisitos de alta disponibilidad y recuperación ante desastres al habilitar una FCI de SQL Server que abarque tanto las zonas de disponibilidad como las. Regiones de AWS SIOS DataKeeper crea una SAN virtual agrupada mediante volúmenes locales de Amazon Elastic Block Store (Amazon EBS) y utiliza la replicación sincrónica entre las zonas de disponibilidad para alta disponibilidad, a la vez que utiliza la replicación asíncrona entre regiones y para la recuperación ante desastres. Para obtener

más información, consulte la sección [High Availability Protection for Windows Applications](#) en la documentación de SIOS.

- **Grupos de disponibilidad distribuida:** un grupo de disponibilidad distribuida es un tipo especial de grupo de disponibilidad que abarca dos grupos de disponibilidad Always On independientes. Un grupo de disponibilidad puede residir en dos regiones distintas (por ejemplo, us-east-1 y us-west-1). Puede pensar en un grupo de disponibilidad distribuido como un grupo de disponibilidad de grupos de disponibilidad porque los grupos de disponibilidad Always On subyacentes están configurados en dos clústeres de WSFC diferentes. Se requiere la edición Enterprise de SQL Server para implementar grupos de disponibilidad distribuida. Para obtener más información, consulte [Distributed availability groups](#) en la documentación de Microsoft.
- **Envío de registros:** puede implementar el envío de registros para proteger sus bases de datos en varias regiones, en el raro caso de que una región se vea afectada y deje de estar disponible. En función de la frecuencia de envío de las transacciones y los registros, puede lograr el RPO y el RTO en cuestión de minutos. Para obtener más información, consulte [About Log Shipping \(SQL Server\)](#) en la documentación de Microsoft.
- **AWS Elastic Disaster Recovery—** Elastic Disaster Recovery es una aplicación de software como servicio (SaaS) que administra la replicación de servidores desde cualquier infraestructura con AWS fines de recuperación ante desastres. También puede usar Recuperación elástica ante desastres para replicar SQL Server en todas las regiones. Recuperación elástica ante desastres es una solución basada en agentes que replica máquinas virtuales completas, lo que incluye el sistema operativo, todas las aplicaciones instaladas y todas las bases de datos en un área de almacenamiento provisional. Para obtener más información, consulte [What is Elastic Disaster Recovery?](#) en la documentación de Recuperación elástica ante desastres.
- **AWS Database Migration Service (AWS DMS):** AWS DMS admite la migración en tiempo real de datos hacia y desde AWS una región diferente. Puede usar esta característica para configurar una instancia de SQL Server independiente en una región diferente para que sirva como base de datos de la recuperación ante desastres. Para obtener más información, consulte [¿Qué es AWS Database Migration Service?](#) en la AWS DMS documentación.

Grupos de disponibilidad Always On de SQL Server

Si utiliza la edición Enterprise de SQL Server solo para un [grupo de disponibilidad Always On](#) de alta disponibilidad, puede pasarse a la edición Standard de SQL Server aprovechando los grupos de disponibilidad básicos. Puede reducir los costos entre un 65 % y un 75 % si utiliza grupos de disponibilidad básicos en lugar de grupos de disponibilidad Always On.

Note

Para obtener información adicional sobre las diferencias de costo entre las distintas ediciones de SQL Server, consulte la sección [Comparación de las ediciones de SQL Server](#) de esta guía.

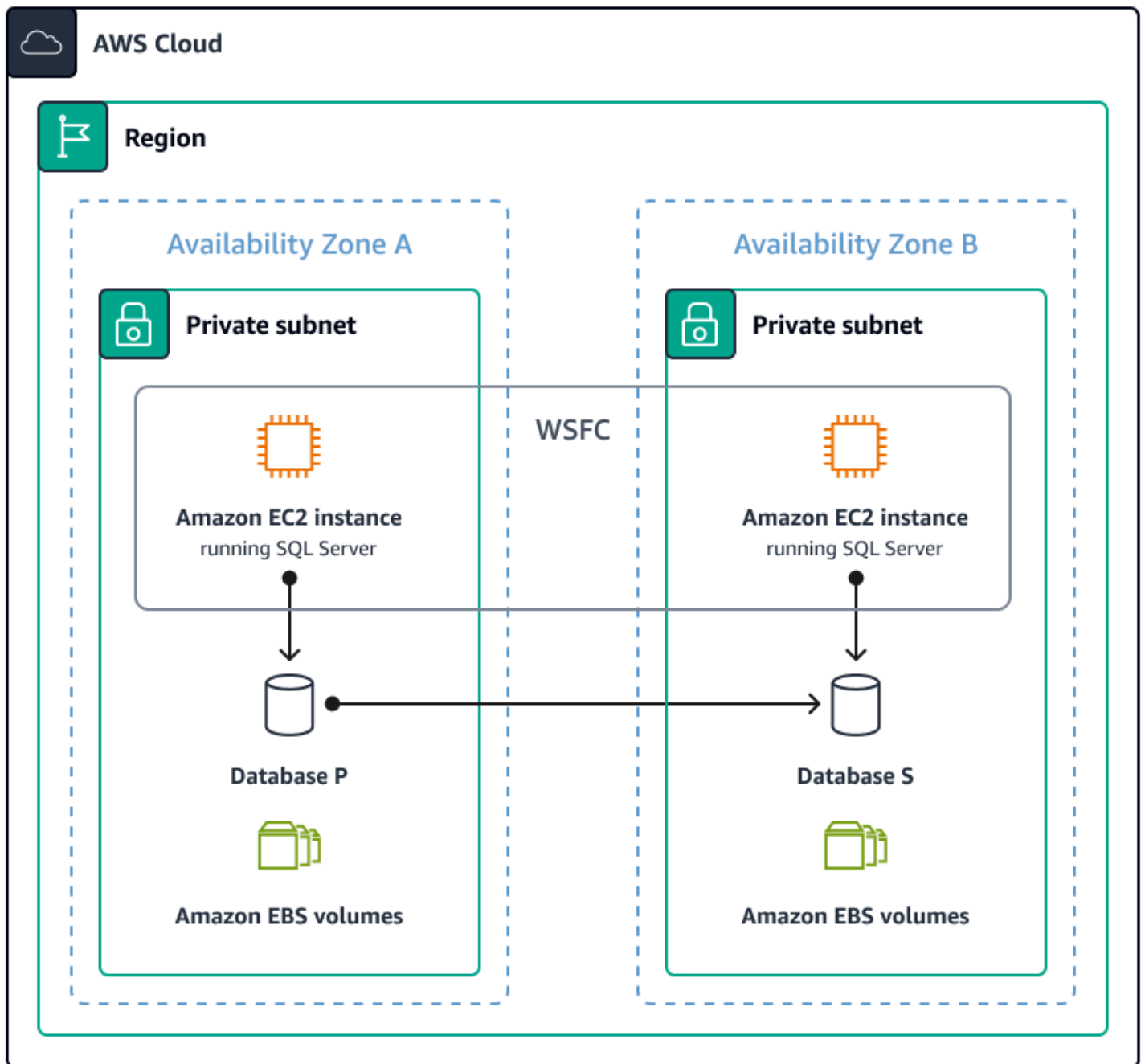
Características

- Disponible en la edición Standard de SQL Server.
- Límite de dos réplicas (principal y secundaria).
- No hay acceso de lectura en la réplica secundaria.
- No hay comprobaciones de integridad en las réplicas secundarias.

Limitaciones

- Solo admite una base de datos de disponibilidad por grupo de disponibilidad.
- Los grupos de disponibilidad básicos no pueden formar parte de un grupo de disponibilidad distribuida,


En el siguiente diagrama, se muestra un ejemplo de arquitectura para una solución de clúster de conmutación por error de Windows Server.



Instancias de clústeres de conmutación por error Always On de SQL Server

Puede utilizar instancias de clúster de conmutación por error (FCIs) para garantizar la continuidad de las operaciones de la base de datos y, al mismo tiempo, minimizar el tiempo de inactividad y reducir el riesgo de pérdida de datos. FCIs ofrece una solución fiable si busca una alta disponibilidad para su base de datos de SQL Server sin una configuración de réplica de lectura.

A diferencia de los grupos de disponibilidad, FCIs puede proporcionar una solución de conmutación por error fiable sin necesidad de la edición SQL Server Enterprise. En su lugar, solo FCIs requieren licencias de la edición estándar de SQL Server. Se puede utilizar FCIs para reducir los costes de licencias de SQL Server entre un 65 y un 75 por ciento.

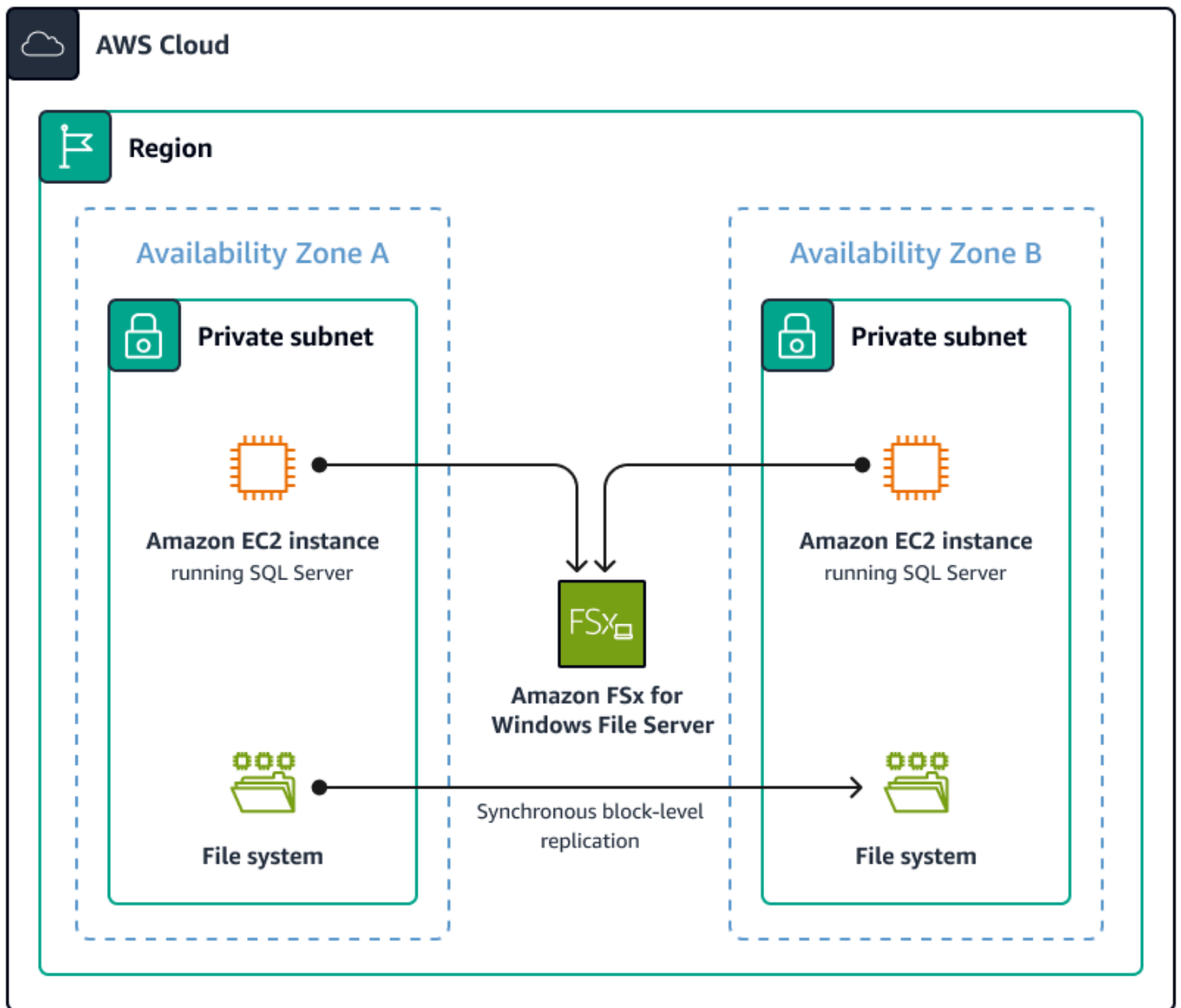
 Note

Para obtener información adicional sobre las diferencias de costo entre las ediciones de SQL Server, consulte la sección [Comparación de las ediciones de SQL Server](#) de esta guía.

Considere lo siguiente:

- Amazon FSx for Windows File Server ofrece una potente solución para cumplir con los requisitos de almacenamiento compartido FCI de SQL Server. Puede utilizar Windows File Server FSx para evitar la necesidad de adquirir una licencia para una solución de replicación del almacenamiento y gestionar el almacenamiento compartido por su cuenta. Esto puede suponer un ahorro significativo de entre un 30 % y un 40 %. Para obtener más información, consulte la publicación [Simplifique las implementaciones de alta disponibilidad de Microsoft SQL Server con Amazon FSx for Windows File Server](#) en el blog sobre AWS almacenamiento.
- Con [Software Assurance benefits summary](#) (PDF descargable) y el modelo traiga su propia licencia (BYOL), podrá aprovechar las ventajas de la conmutación por error pasiva, siempre y cuando el servidor secundario sea pasivo. Esto se traduce en un ahorro de los costos derivados del uso de licencias de SQL, ya que no es necesario proporcionar licencias al nodo pasivo del clúster.

El siguiente diagrama muestra un ejemplo de arquitectura para una FCI de SQL Server utilizada FSx para Windows File Server.




SIOS DataKeeper

Le recomendamos que tenga en cuenta los requisitos de almacenamiento compartido si planea implementar SQL Server FCIs en AWS. Las infraestructuras en las instalaciones tradicionales suelen utilizar una SAN para cumplir con los requisitos de almacenamiento compartido, pero esta no es una opción viable en AWS. Amazon FSx for Windows File Server es la solución de almacenamiento recomendada para SQL Server FCI on AWS, pero tiene limitaciones que impiden añadir servidores de clústeres en diferentes Regiones de AWS ubicaciones.

Puede utilizar [SIOS DataKeeper](#) para crear una FCI de SQL Server que cubra tanto las zonas de disponibilidad como las regiones y, al mismo tiempo, reduzca los costes entre un 58 y un 71 por ciento. SIOS DataKeeper puede ayudarlo a obtener los beneficios de alta disponibilidad de la FCI. Esto convierte a SIOS en DataKeeper una solución rentable y fiable para las organizaciones.

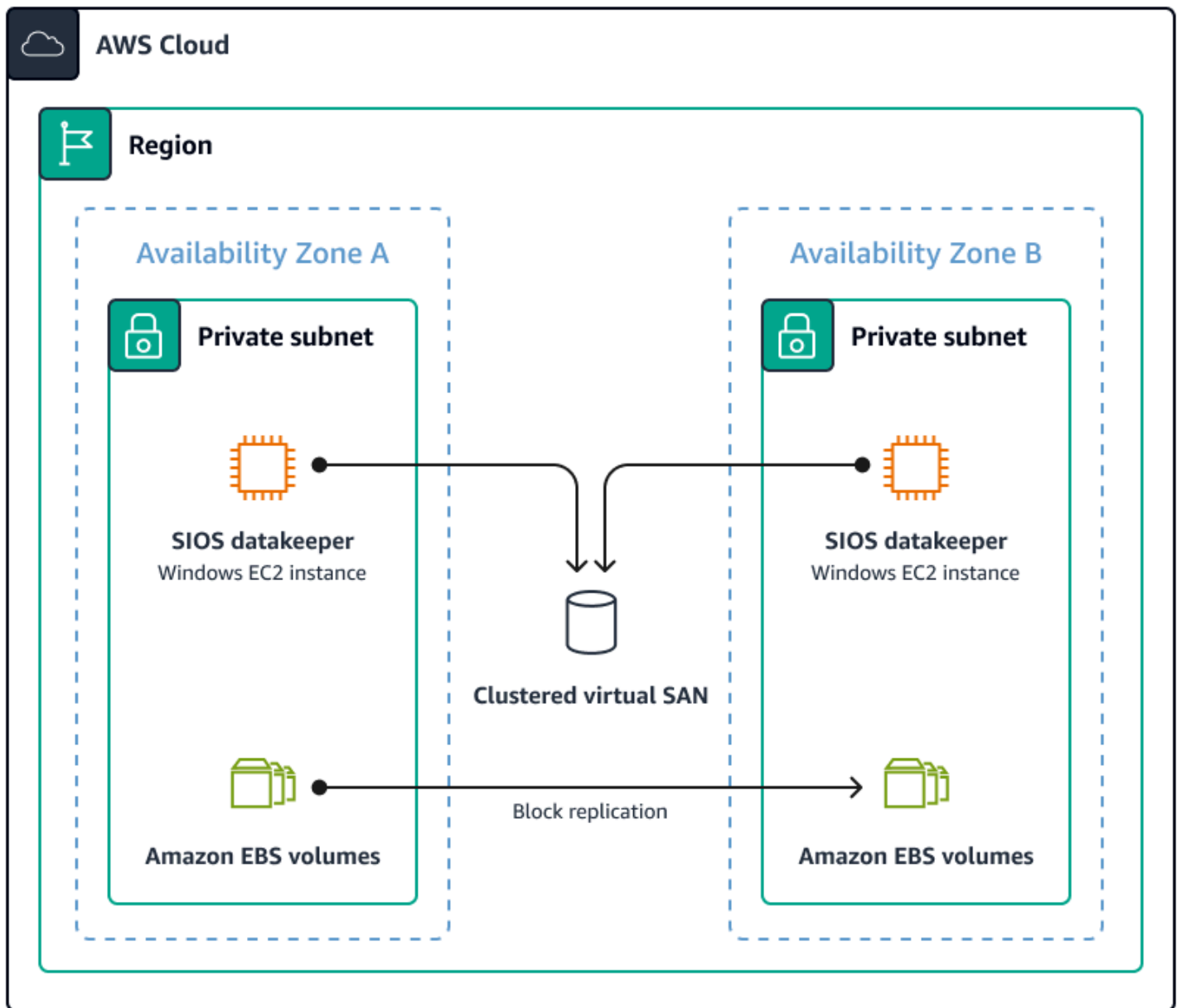
Tenga en cuenta las siguientes ventajas adicionales del uso de SIOS: DataKeeper

- SIOS DataKeeper crea una SAN virtual agrupada mediante volúmenes de EBS locales y utiliza la replicación sincrónica entre las zonas de disponibilidad para lograr una alta disponibilidad. Para la recuperación ante desastres, SIOS DataKeeper utiliza la replicación asíncrona entre regiones.
- SIOS DataKeeper proporciona funciones de agrupamiento en clústeres de clase empresarial mediante el uso de la edición SQL Server Standard. Esto reduce los costos asociados al uso de licencias de SQL Server entre un 65 % y un 75 % en comparación con la implementación de la alta disponibilidad con los grupos de disponibilidad de SQL Server Always On que utilizan la edición Enterprise de SQL Server. Con SIOS DataKeeper, puede crear un entorno de SQL Server rentable, flexible y de alta disponibilidad que satisfaga las necesidades de su organización.

 Note

Para obtener información adicional sobre las diferencias de costo entre las ediciones de SQL Server, consulte la sección [Comparación de las ediciones de SQL Server](#) de esta guía.

En el siguiente diagrama, se muestra un ejemplo de arquitectura para una FCI de SQL Server que utiliza una solución de SAN virtual en clústeres.

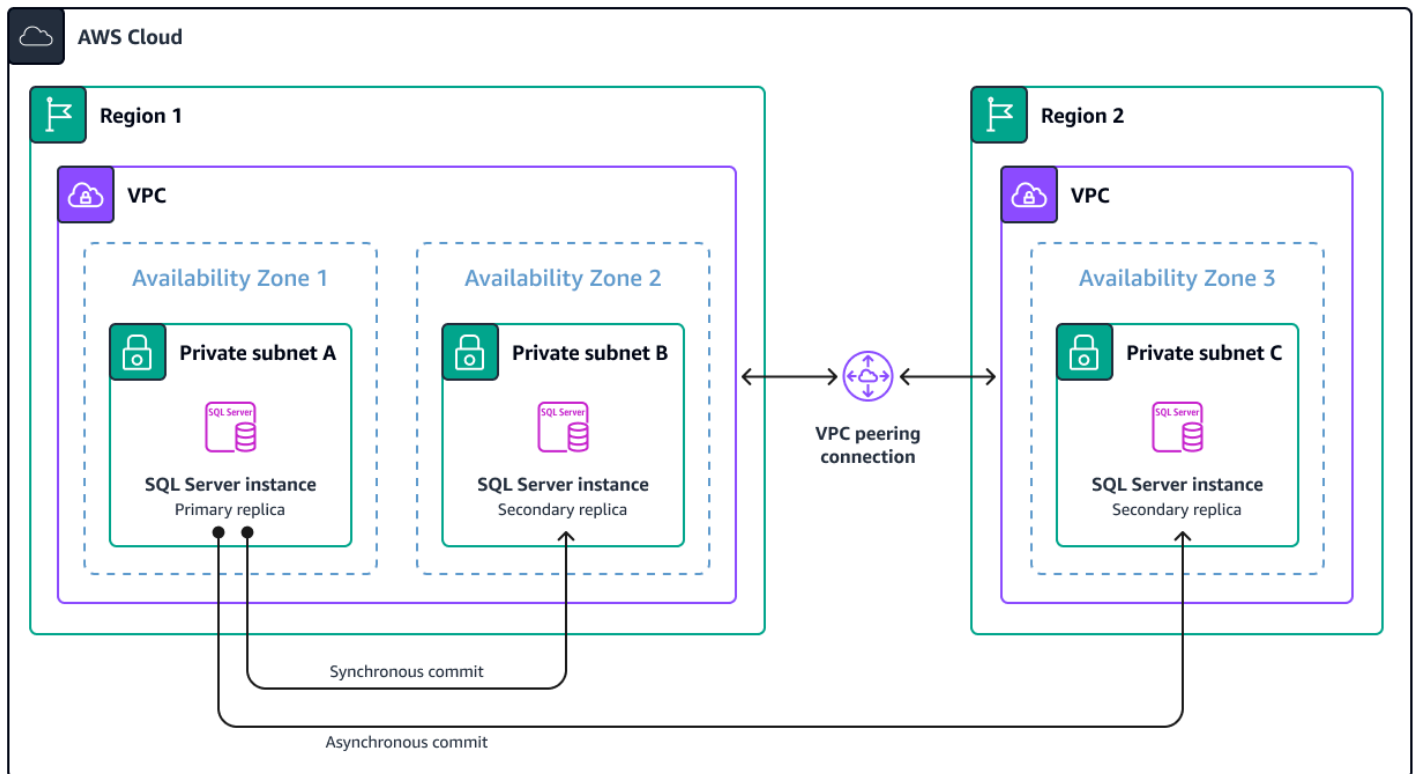


Grupos de disponibilidad AlwaysOn

Puede usar los grupos de disponibilidad Always On para lograr una alta disponibilidad y para fines de recuperación ante desastres. Puede lograr una alta disponibilidad implementando SQL Server en dos zonas de disponibilidad de una región. Puede lograr la recuperación ante desastres ampliando los grupos de disponibilidad a varias regiones.

En el siguiente diagrama, se muestra un ejemplo de arquitectura para una solución basada en grupos de disponibilidad Always On. Las réplicas de la región 1 del diagrama utilizan una confirmación sincrónica, que proporciona una conmutación por error automática del grupo de

disponibilidad. La réplica de la región 2 utiliza una confirmación asíncrona, que requerirá una conmutación por error manual del grupo de disponibilidad.



Grupos de disponibilidad distribuida

En el caso de las implementaciones de SQL Server críticas en las que no se puede comprometer la fiabilidad ni la recuperación ante desastres, recomendamos una estrategia multirregional. La distribución de los grupos de disponibilidad en varias regiones es la solución más resiliente para mantener la continuidad empresarial y minimizar el tiempo de inactividad.

Esta arquitectura aprovecha al máximo las capacidades de Amazon FSx for Windows File Server, incluido el almacenamiento compartido, la replicación sincrónica a nivel de bloque y SQL Server. Estas capacidades le permiten crear un entorno de SQL Server de alta disponibilidad que abarque varias zonas de disponibilidad. Al replicar esta configuración en otra región, obtiene un sistema completamente redundante que puede gestionar incluso las interrupciones más graves. Lo que distingue a esta solución es el nivel de flexibilidad y seguridad que ofrece. La arquitectura agnóstica al dominio de los grupos de disponibilidad distribuida permite que los servidores de clústeres de Windows subyacentes se unan a diferentes dominios de Active Directory, mientras que la autenticación basada en certificados garantiza la máxima protección para los entornos de SQL Server y proporciona estrictos requisitos de RTO y RPO para una estrategia de recuperación

ante desastres multirregional. Para obtener información sobre cómo crear una arquitectura multirregional, consulte [Notas de campo: Creación de una arquitectura multirregional para SQL Server mediante FCI y grupos de disponibilidad distribuida](#) en el blog de arquitectura. AWS

En el siguiente diagrama, se muestra un ejemplo de arquitectura para una solución multirregional que utiliza grupos de disponibilidad distribuida.



Envío de registros

El envío de registros es un método probado, fiable y rentable para proteger sus bases de datos en todas las regiones en caso de que se produzca una interrupción inesperada. Las organizaciones llevan décadas usando el envío de registros para proteger sus datos.

Si implementa el envío de registros AWS, puede lograr el RPO y el RTO en cuestión de minutos, según la frecuencia de las transacciones y las tareas de envío de registros. En el improbable caso de que una región quede inaccesible, el envío de registros mantiene sus datos protegidos y recuperables.

Tenga en cuenta las siguientes ventajas adicionales de utilizar el envío de registros:

- Reduzca los costos y cumpla con los requisitos de su empresa gracias al envío de registros para mejorar la resiliencia de la recuperación ante desastres en todas las regiones. El envío de registros

reduce el costo total de propiedad, ya que solo necesita licencias de la edición Standard o Web de SQL Server.

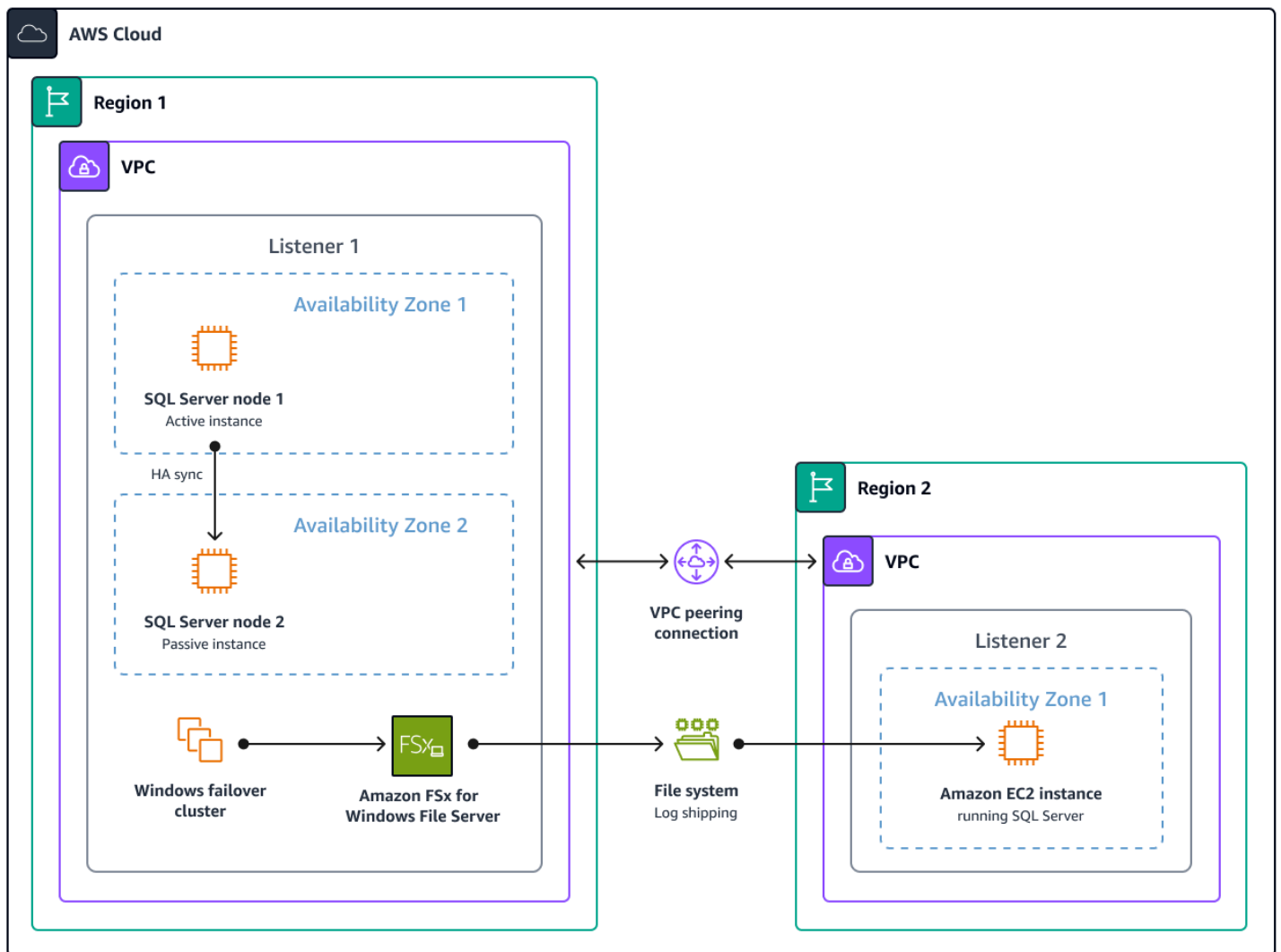
- Elimine los costes de licencia de un recovery/passive servidor en caso de avería mediante el envío de registros con [Software Assurance](#) activo. Solo se necesita una licencia para primary/active SQL Server cuando se utiliza el envío de registros con Software Assurance.
- Reduzca los costos asociados al uso de licencias de SQL Server entre un 65 % y un 75 % al no tener que usar la edición Enterprise de SQL Server para configurar grupos de disponibilidad distribuida entre las regiones. Para ello, utilice la edición Standard de SQL Server y SQL Server FCIs junto con el envío de registros para cumplir con sus requisitos de recuperación ante desastres.

Note

Para obtener información adicional sobre las diferencias de costo entre las ediciones de SQL Server, consulte la sección [Comparación de las ediciones de SQL Server](#) de esta guía.

Para obtener más información, consulte [Ampliar la recuperación ante desastres de SQL Server mediante el envío de registros para la configuración FCI de SQL Server con Amazon FSx para Windows](#) en el blog de AWS arquitectura.

En el siguiente diagrama, se muestra un ejemplo de arquitectura para una solución de envío de registros.

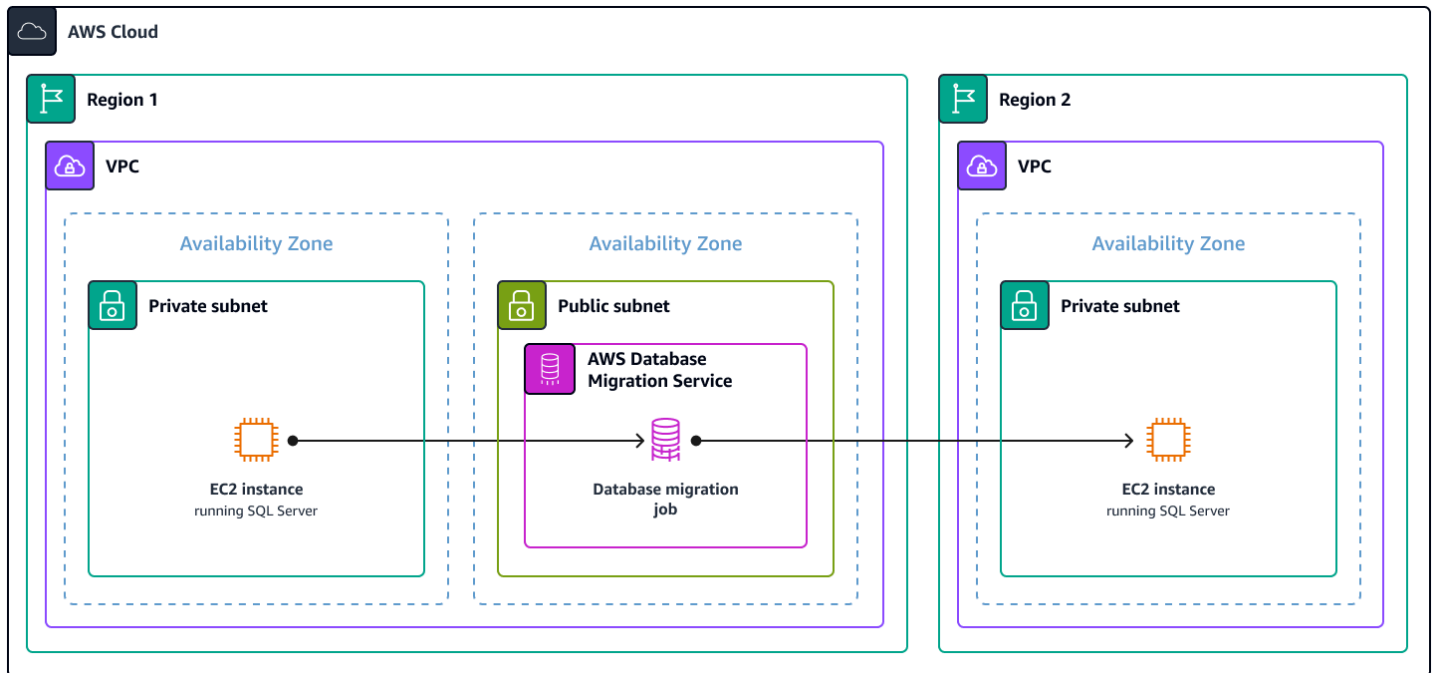


AWS Database Migration Service

Puede usar AWS Database Migration Service (AWS DMS) para diseñar una HA/DR solución en función de las necesidades de su aplicación. AWS DMS le permite copiar fácilmente los datos a una base de datos secundaria de SQL Server en la misma región (HA) o entre regiones (DR). Esta estrategia es técnicamente segura y le permite maximizar su inversión en infraestructura de AWS, así como también optimizar el uso de los recursos.

AWS DMS es un servicio rentable. Solo se le cobrarán los recursos de CPU utilizados durante el proceso de transferencia y cualquier almacenamiento de registros adicional. Esto significa que puede beneficiarse de esta solución sin incurrir en costos adicionales significativos. Puede usarlo AWS DMS para garantizar que sus datos estén disponibles y accesibles y, al mismo tiempo, minimizar los costos asociados con las licencias y el uso de los recursos.

En el siguiente diagrama, se muestra un ejemplo de arquitectura para una solución basada en AWS DMS.



AWS Elastic Disaster Recovery

Algunas organizaciones deben asegurarse de que todas las aplicaciones empresariales críticas cuenten con un plan de recuperación ante desastres. En el pasado, muchas de estas organizaciones hacían grandes inversiones en las soluciones tradicionales de recuperación ante desastres, que requerían la creación previa y el mantenimiento de toda una infraestructura duplicada. Esta estrategia es costosa, laboriosa y difícil de escalar.

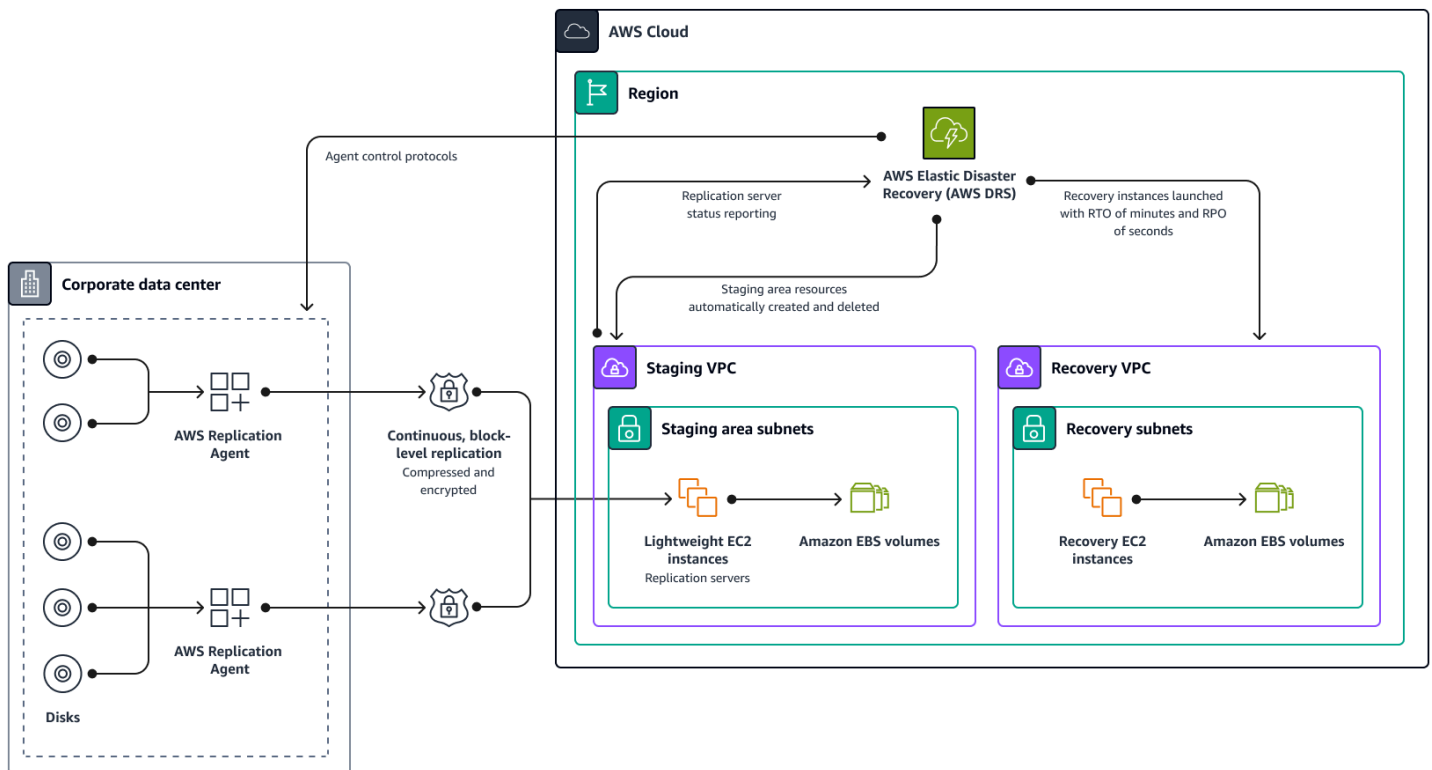
Ahora, puede utilizarla AWS Elastic Disaster Recovery para eliminar la necesidad de crear previamente una infraestructura de recuperación ante desastres. Las máquinas de recuperación ante desastres no se inician en Recuperación elástica ante desastres hasta que son necesarias, por lo que solo paga por lo que usa cuando lo necesitas. Esto significa que puede reducir considerablemente los costos asociados al uso de licencias de software y de recursos de computación de alto rendimiento.

Además, el área de almacenamiento provisional de la solución de recuperación ante desastres contiene volúmenes de Amazon Elastic Block Store (Amazon EBS) de bajo costo. Los volúmenes de EBS reducen aún más el costo de aprovisionar recursos duplicados. Esto le permite reducir los costos generales de recuperación ante desastres y, al mismo tiempo, mantener una solución de

recuperación ante desastres sólida y fiable que cumpla con los requisitos de su empresa. Puedes usar Elastic Disaster Recovery para centrarte en tus actividades comerciales principales y, al mismo tiempo, cuidar la infraestructura subyacente de tu solución de recuperación ante desastres.

En el caso de SQL Server, puede usar Recuperación elástica ante desastres como una opción rentable de recuperación ante desastres. La licencia para el nodo pasivo en una arquitectura de SQL Server altamente disponible y tolerante a errores viene incluida si utiliza y activa Software Assurance. Sin embargo, tendrá que seguir pagando los costos de procesamiento para que el servidor pasivo esté en línea. Con Recuperación elástica ante desastres, el servidor principal puede replicarse en el entorno de recuperación ante desastres sin necesidad de mantener activo Software Assurance y sin tener que pagar los costos de los recursos de computación de la recuperación ante desastres. Esta combinación de ahorros puede reducir los costos de recuperación ante desastres de SQL Server en un 50 % o más.

En el siguiente diagrama, se muestra un ejemplo de arquitectura para una solución basada en Recuperación elástica ante desastres.



Para obtener más información, consulte [Cómo configurar la alta disponibilidad de SQL Server en un sitio de recuperación ante desastres que se restauró mediante AWS Elastic Disaster Recovery](#) el AWS blog Microsoft Workloads on.

Comparación de los costos

En la siguiente tabla se comparan los costes de las HA/DR soluciones incluidas en esta sección. A los efectos de esta comparación, se hacen las siguientes suposiciones:

- Tipo de instancia: r5d.xlarge
- Tipo de licencia: licencia incluida para Windows y SQL Server
- Región: us-east-1

Solución	Alta disponibilidad	Recuperación ante desastres	Enterprise Edition	Standard Edition	Costo
Envío de registros	No	Sí	Sí	Sí	Edición Enterprise de SQL Server: 32 674,8 USD (2 nodos) Edición Standard de SQL Server: 14 804,4 (2 nodos)
Grupos de disponibilidad AlwaysOn	Sí	Sí	Sí	Sí, pero grupos de disponibilidad básicos (2 nodos)	Edición Enterprise de SQL Server: 32 674,8 USD (2 nodos) Edición Standard de SQL Server: 14 804,4 (2 nodos)

Solución	Alta disponibilidad	Recuperación ante desastres	Enterprise Edition	Standard Edition	Costo
Siempre encendido FCIs	Sí	No	Sí	Sí (2 nodos)	Edición Standard de SQL Server: 14 804,4
Grupos de disponibilidad distribuida	Sí	Sí	Sí	No	Edición Enterprise de SQL Server: 65 349,6 USD (4 nodos)

Solución	Alta disponibilidad	Recuperación ante desastres	Enterprise Edition	Standard Edition	Costo
Recuperación de desastres elástica	No	Sí	Sí	Sí	<p>Aproximadamente 107,48 USD al mes para la replicación de 1 instancia y 1 TB de almacenamiento</p> <p>Nota: Recuperación elástica ante desastres se factura por hora y por servidor de replicación. El costo es el mismo, independientemente de la cantidad de discos, el tamaño del almacenamiento, la cantidad de operaciones de simulacro o recuperación iniciadas o la región</p>

Solución	Alta disponibilidad	Recuperación ante desastres	Enterprise Edition	Standard Edition	Costo
					en la que se vaya a replicar.
SIOS DataKeeper	Sí	Sí	Sí	Sí	<p>Grupos de disponibilidad Always On con Software Assurance (2 nodos, 24 núcleos): 213 480 USD</p> <p>Clúster de SQL Server de 2 nodos que se ejecuta en la edición SQL Server Standard con SIOS DataKeeper y Software Assurance : 61.530\$ (2 nodos)</p>


Solución	Alta disponibilidad	Recuperación ante desastres	Enterprise Edition	Standard Edition	Costo
AWS DMS	No	Sí	Sí	Sí	745,38 USD al mes por una instancia r5.xlarge y 1 TB de almacenamiento

Recomendaciones de optimización de costos

Le recomendamos que siga los siguientes pasos para elegir una HA/DR solución que cumpla con los requisitos de su organización:

- Consulte la sección [Selección de la instancia de EC2 adecuada para las cargas de trabajo de SQL Server](#) de esta guía.
- Determine los requisitos de IOPS y rendimiento de sus cargas de trabajo poniendo en marcha contadores de rendimiento durante los picos de carga de trabajo:
 - IOPS = disco reads/sec + escrituras de disco por segundo
 - Rendimiento = lectura de disco + bytes de escritura de disco por segundo bytes/sec
- Utilice los siguientes tipos de volúmenes de almacenamiento para mejorar el rendimiento y ahorrar costos:
 - NVMe almacenamiento de instancias y extensión de grupo de tempdb búferes
 - Volúmenes io2 para archivos de bases de datos
- Use [AWS Trusted Advisor](#) para obtener recomendaciones sobre la optimización de costos para SQL Server en Amazon EC2. No necesita instalar un agente Trusted Advisor para realizar las comprobaciones de optimización de SQL Server. Trusted Advisor inspecciona las configuraciones de instancias de Amazon EC2 SQL Server incluidas en la licencia, como CPUs virtual (CPUsv), versión y edición. A continuación, Trusted Advisor hace recomendaciones basadas en las mejores prácticas.
- Úselo AWS Compute Optimizer para las recomendaciones de tamaño correcto de la instancia Amazon EC2 y Amazon EBS.

- Úselo [Calculadora de precios de AWS](#) para diseñar su HA/DR estrategia de estimación de costos.
- Para determinar si es posible pasar de la edición Enterprise de SQL Server a la edición Standard de SQL Server, utilice la vista de administración dinámica [sys dm_db_persisted_sku_features](#) para identificar las características específicas de la edición que están activas en la base de datos actual.

 Note

Side-by-side Las migraciones son necesarias para realizar cambios en la edición de SQL Server cuando se utilizan instancias EC2 con licencia incluida.

- Realice simulacros de recuperación ante desastres semestrales o anuales para hacer un mejor diseño que pueda recuperar la base de datos con un RTO y un RPO definidos. Esto también puede ser útil para identificar cualquier punto débil en la arquitectura.

Recursos adicionales

- [Simplifique las implementaciones de alta disponibilidad de Microsoft SQL Server con Amazon FSx for Windows File Server](#) (blog sobre AWS almacenamiento)
- [Notas de campo: Creación de una arquitectura multirregional para SQL Server mediante FCI y grupos de disponibilidad distribuida](#) (AWS blog de arquitectura)
- [Diseño una recuperación ante desastres para SQL Server en AWS: Parte 1](#) (blog sobre AWS bases de datos)
- [Alta disponibilidad de Microsoft SQL con Amazon FSx para Windows](#) (YouTube)
- [Maximizing Microsoft SQL Server Performance with Amazon EBS](#) (blog de almacenamiento de AWS)
- [Comparación de sus patrones de almacenamiento local con los servicios AWS de almacenamiento](#) (blog sobre AWS almacenamiento)
- [Planea sustituir el NAS de un centro de datos por Amazon FSx File Gateway](#) (blog sobre AWS almacenamiento)
- [Optimización de los costes de las implementaciones de SQL Server de alta disponibilidad en AWS](#)(blog sobre AWS almacenamiento)
- [How to set up disaster recovery for SQL Server Always On Availability Groups using AWS Elastic Disaster Recovery](#) (Cargas de trabajo de Microsoft en AWS)
- [Cómo configurar la alta disponibilidad de SQL Server en un sitio de recuperación ante desastres que se restauró con AWS Elastic Disaster Recovery](#) (Microsoft Workloads activado AWS)

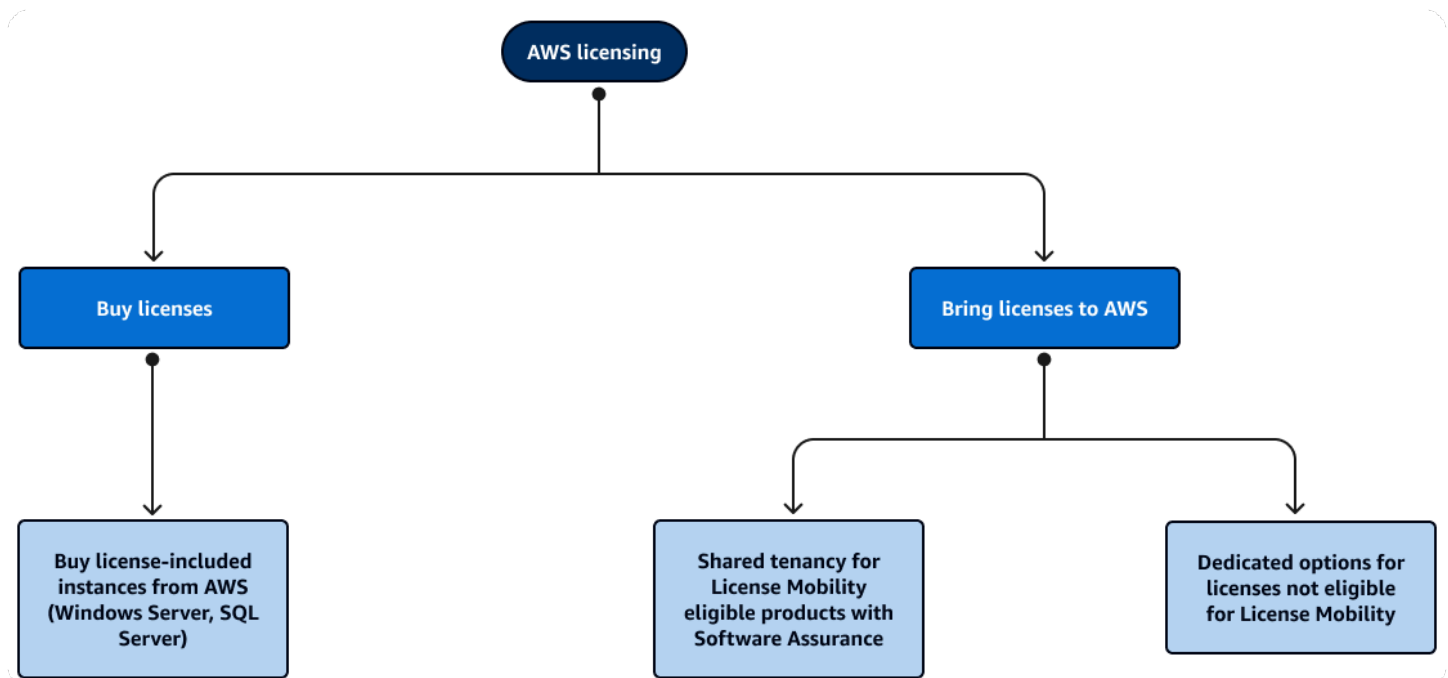
Explicación de las licencias de SQL Server

Descripción general de

A medida que más y más empresas trasladan sus cargas de trabajo a la nube, la optimización de los costos en las plataformas en la nube ahora supone una de las principales prioridades. La concesión de licencias es uno de los costes más importantes asociados a la ejecución de las cargas de trabajo de Microsoft. AWS En esta sección se explica cómo optimizar los costes AWS mediante la optimización de las licencias de Microsoft para SQL Server.

AWS opciones de licenciamiento

AWS ofrece una gama de opciones flexibles de optimización de costes para la concesión de licencias. Estas opciones de licencia están diseñadas para que pueda reducir los costos, mantener el cumplimiento normativo y satisfacer las necesidades de su empresa.



AWS clasifica las licencias en tres tipos principales:

1. Licencia incluida: esta opción de licencia le permite comprar y usar licencias bajo demanda, pagando solo por lo que use. La opción con licencia incluida es ideal para situaciones en las que necesita flexibilidad en el uso de las licencias y desea evitar costos iniciales. Puede elegir entre una gama de productos de Windows Server, SQL Server y otros productos de Microsoft.

2. Productos Bring Your Own License (BYOL) con movilidad de licencias: esta opción de licencia está diseñada para situaciones en las que ya tiene licencias existentes y desea utilizarlas en la nube. AWS permite a los clientes llevar sus propias licencias a la nube a través del programa [License Mobility](#) de Microsoft. Puede utilizar productos con movilidad de licencias, como SQL Server with Software Assurance (SA), en régimen de arrendamiento compartido o dedicado para reducir los costes de las AWS instancias.
3. Productos BYOL sin movilidad de licencias: para los productos de Microsoft que no tienen movilidad de licencias, como Windows Server, AWS ofrece opciones específicas para usar estos productos en la nube. Además, los hosts dedicados ofrecen la oportunidad de conceder licencias para los núcleos físicos. Esto puede ahorrarle un 50 % o más en las licencias necesarias para ejecutar sus cargas de trabajo. Los hosts dedicados son una opción excelente para las cargas de trabajo estables y predecibles que funcionan la mayor parte del tiempo.

El impacto en los costos de usar sus propias licencias

El uso de sus propias licencias puede tener un impacto significativo en el costo de poner en marcha las cargas de trabajo de Microsoft en AWS. Si usa sus propias licencias, no tiene que pagar costos adicionales por el uso de licencias para las instancias activas en la nube. Este ahorro puede ser considerable.

La siguiente comparación muestra el costo mensual bajo demanda que implica poner en marcha una sola instancia c5.xlarge las 24 horas del día, los 7 días de la semana:

- Windows Server + Edición Enterprise de SQL Server: 1353 USD al mes (licencia incluida)
- Windows Server + Edición Standard de SQL Server: 609 USD al mes (licencia incluida)
- Solo para Windows Server: 259 USD al mes (licencia incluida)
- Solo computación (Linux): 127 USD al mes

En última instancia, usar sus propias licencias puede tener un impacto significativo en el costo de poner en marcha las cargas de trabajo de Microsoft en AWS. Si usa sus licencias actuales, puede reducir los costos de las licencias y ahorrar dinero en su AWS factura total.

Optimización de las licencias

Una evaluación de AWS optimización y licencias (AWS OLA) puede ayudarlo a optimizar sus licencias al reducir los costos informáticos y de licencias. AWS La OLA está diseñada para evaluar

sus requisitos de licencia para las cargas de trabajo que se estén ejecutando AWS o para las cargas de trabajo cuya migración esté planificada. AWS OLA ofrece recomendaciones para optimizar el uso de las licencias.

Una de las estrategias clave para optimizar el uso de las licencias es el [dimensionamiento correcto de las instancias](#). El dimensionamiento correcto implica seleccionar el tipo de instancia adecuado para su carga de trabajo en función de sus requisitos de CPU, memoria y almacenamiento. Al elegir el tamaño de instancia adecuado, puede asegurarse de utilizar los recursos de forma rentable. Este ahorro puede ser considerable.

Con las licencias de software de Microsoft, la cantidad de núcleos que usa el software es un factor fundamental a la hora de determinar los costos de las licencias. Por ejemplo, las licencias de Windows Server y SQL Server se suelen conceder según el número de núcleos. Al dimensionar las instancias correctamente, puede reducir la cantidad de núcleos que usa el software de Microsoft y, a su vez, reducir tanto el costo de la instancia como la cantidad de licencias necesarias.

Recomendaciones de optimización de costos

La optimización de las licencias es un componente clave de la optimización de costos en AWS. Al implementar las estrategias adecuadas, puede reducir los costos asociados al uso de licencias, garantizar el cumplimiento normativo y obtener el mejor valor posible de su inversión en licencias. En esta sección se describen varias estrategias para la optimización de las licencias.

Use sus propias licencias de Windows Server aptas

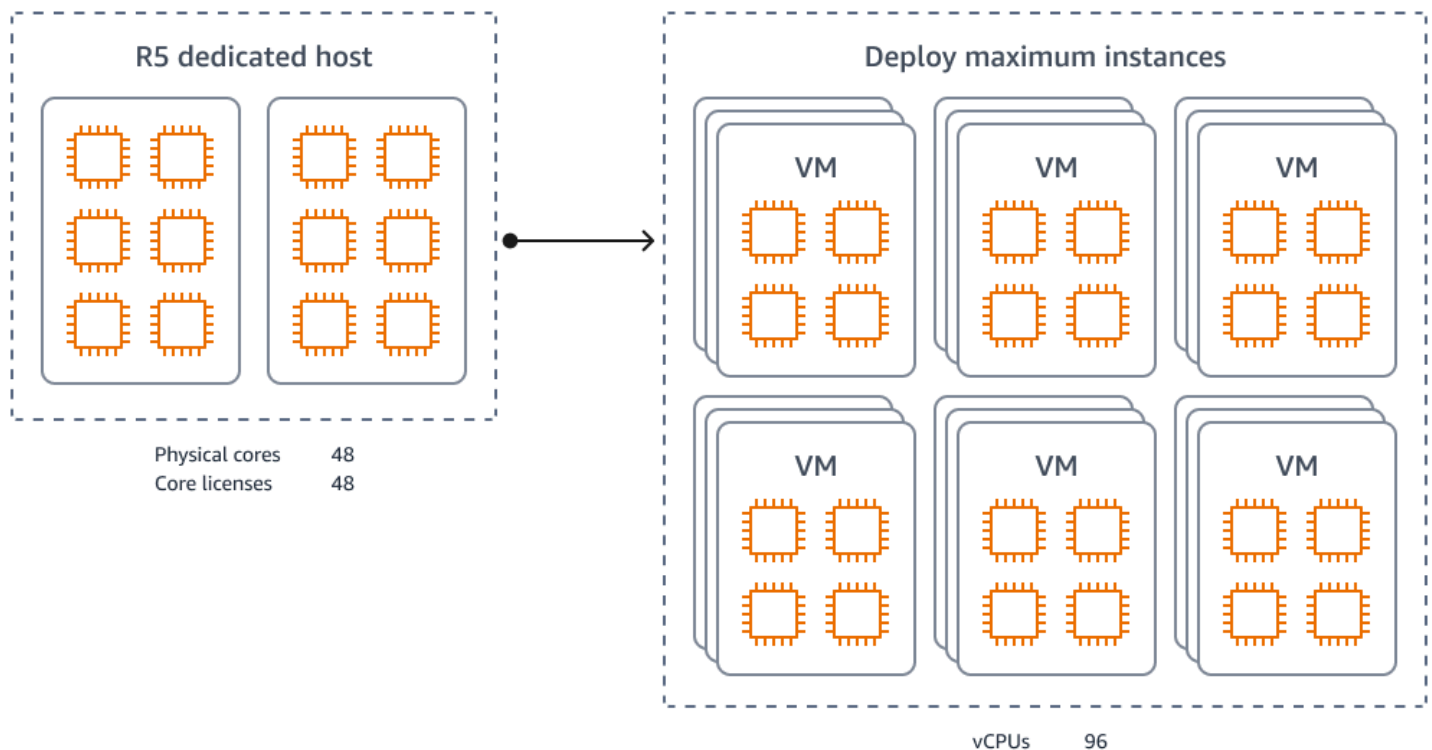
Usar sus propias licencias de Windows Server es una de las estrategias más eficaces para la optimización de las licencias. Esta estrategia le permite aprovechar sus inversiones actuales para reducir sus AWS gastos.

Por ejemplo, puede implementar Windows Server 2019 y versiones anteriores en [hosts dedicados de Amazon EC2](#) si compró las licencias antes del 1 de octubre de 2019 o las compró durante un proceso de “true-up” o compensación en virtud de acuerdos empresariales activos firmados antes de esa fecha. Esta regla se basa en un cambio que Microsoft realizó en 2019 en sus términos y condiciones de licencia para productos sin movilidad de licencia, como Windows Server, cuando se implementaron en [proveedores listados](#) (por ejemplo, Alibaba o Google Cloud). AWS Según las nuevas condiciones, no puedes traer tus propias licencias de Windows Server, AWS sino que debes usar instancias con licencia incluida. Sin embargo, si adquirió licencias perpetuas antes de esa fecha, aún puede implementar esas licencias de Windows Server en hosts dedicados de Amazon EC2.

Licencias de nivel físico

La concesión de licencias para los núcleos físicos le permite licenciar solo los núcleos físicos de un host, de modo que pueda implementar un número máximo de instancias sin que ello afecte al número de licencias necesarias. Esto suele hacerse mediante la edición Datacenter de Windows Server y la edición Enterprise de SQL Server.

Como ejemplo, pensemos en un host dedicado R5 con 48 núcleos, lo que equivale a 96 v. CPUs. Si utiliza la edición Windows Server Datacenter, solo necesitará 48 licencias. Esto le permite implementar una combinación de instancias de hasta 96 vCPUs, como se muestra en el siguiente diagrama.



Esta estrategia puede ser especialmente rentable si tiene suficientes cargas de trabajo para maximizar la cantidad de instancias que puede poner en marcha en un host. Al licenciar núcleos físicos, no tiene que pagar costos adicionales por el uso de licencias para cada instancia y obtener el mejor valor posible de su inversión en licencias.

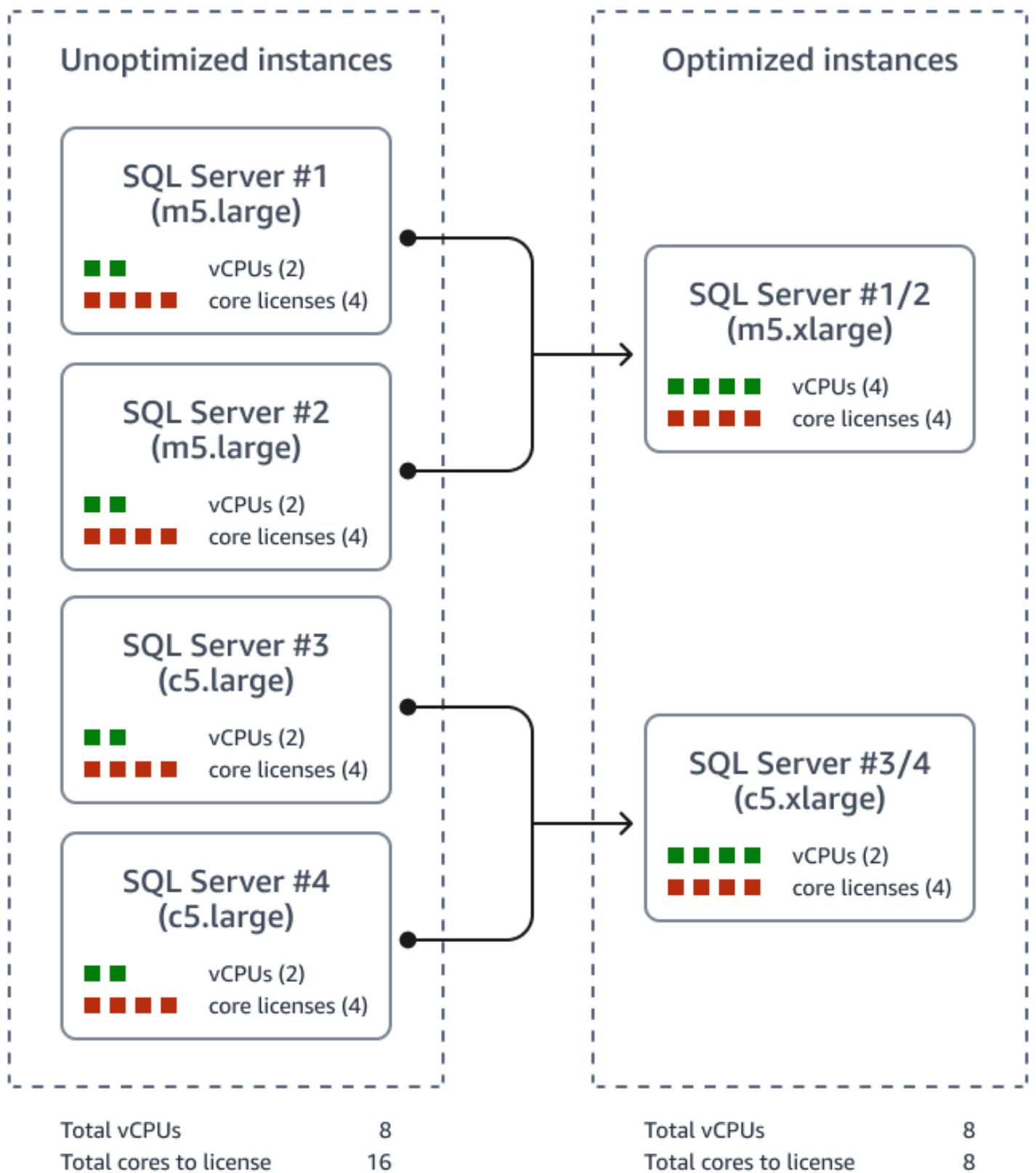
Licencias para núcleos físicos de SQL Server

En la tenencia compartida, las licencias de SQL Server se basan en la cantidad de v CPUs asignada a la instancia. Por el contrario, con los hosts dedicados, puede otorgar la licencia de la edición Enterprise de SQL Server a los núcleos físicos o a la vCPU.

Al igual que en el ejemplo anterior del host dedicado R5, si otorga la licencia de la edición Enterprise de SQL Server a los núcleos físicos, solo necesitará 48 licencias de la edición Enterprise de SQL Server para los hosts. Por el contrario, en la tenencia compartida, donde la única opción es otorgar licencias a cada vCPU, debe tener 96 licencias de la edición Enterprise de SQL Server para la misma carga de trabajo. Por lo tanto, los hosts dedicados pueden ahorrarle hasta un 50 % en los costos asociados al uso de licencias de SQL Server en comparación con la tenencia compartida. Este ahorro se suma al obtenido si usa sus propias licencias de Windows aptas.

Unificación de instancias de SQL Server

La [unificación de SQL Server](#) es el proceso de combinar varias instancias de SQL Server en un servidor. SQL Server requiere un mínimo de cuatro licencias principales por instancia, incluso si la instancia solo tiene dos vCPUs. Esto significa que ejecutar SQL Server en servidores con menos de cuatro núcleos puede provocar que se sobrelicencien estas instancias y se usen más licencias de las necesarias.



Por ejemplo, la consolidación de dos instancias con dos v CPUs cada una en una sola instancia con cuatro v CPUs puede reducir los requisitos de licencia en un 50 por ciento. Esto se debe a que solo se requieren cuatro licencias principales en lugar de ocho.

Para obtener más información sobre la unificación, consulte la sección [Unificación de SQL Server](#) de esta guía.

Cambio a una versión inferior de SQL Server

[Cambiar a una versión inferior de SQL Server](#) puede ser una estrategia clave para optimizar el uso de las licencias y reducir los costos. La edición Enterprise de SQL Server es considerablemente más cara que la edición Standard, por lo que cambiar a una versión inferior puede suponer un importante ahorro de costos.

El cifrado de datos transparente (TDE) y los grupos de disponibilidad Always On son dos características populares de la edición Enterprise de SQL Server. Sin embargo, existen alternativas rentables a estas características que puede considerar si no necesita el conjunto completo de características de la edición Enterprise de SQL Server. Por ejemplo, puede obtener el TDE en la edición Standard de SQL Server a partir de la versión SQL Server 2019. En lugar de los grupos de disponibilidad Always On, puede utilizar clústeres de conmutación por error con almacenamiento compartido FSx para Windows File Server para obtener una alta disponibilidad con la edición SQL Server Standard.

Al pasar de la edición Enterprise de SQL Server a la edición Standard de SQL Server, puede reducir significativamente los costos asociados al uso de licencias. Para obtener más información, consulte la publicación [Cómo optimizar el coste de las implementaciones de SQL Server de alta disponibilidad en el AWS](#) blog sobre almacenamiento. AWS

Además de reducir los costos asociados al uso de licencias, la degradación de las ediciones de SQL Server puede ayudar a reducir sus gastos en Software Assurance y a evitar futuras compensaciones. Si devuelve las licencias no utilizadas, no tiene que pagar costos adicionales por el uso de licencias y podrá sacar el máximo partido posible de su inversión en licencias.

Es importante evaluar cuidadosamente las cargas de trabajo de SQL Server y determinar qué características son fundamentales para las necesidades de su empresa. Para obtener más información, [consulte Evaluación del entorno](#) en Guía AWS prescriptiva y determine si la base de datos de Microsoft SQL Server utiliza funciones específicas de la edición SQL Server Enterprise.

Si elige la edición correcta de SQL Server y utiliza alternativas a las características de la edición Enterprise de SQL Server, puede lograr importantes ahorros de costos y, al mismo tiempo,

mantener el cumplimiento normativo y satisfacer las necesidades de su empresa. Para obtener más información sobre las opciones de cambio a una versión inferior, consulte la sección [Comparación de las ediciones de SQL Server](#) de esta guía.

Uso de la edición Developer de SQL Server en entornos que no son de producción

En entornos que no son de producción, puede implementar ediciones licenciadas de SQL Server, como las ediciones Enterprise o Standard, mediante suscripciones a MSDN en entornos en las instalaciones. Sin embargo, la suscripción a MSDN no incluye Movilidad de licencias. Por lo tanto, si migra a AWS, no podrá transferir esas licencias. En su lugar, debe usar la edición Developer de SQL Server.

La edición Developer de SQL Server es una edición con todas las características de SQL Server que está disponible de forma gratuita. Esta edición está disponible para las versiones 2016 y posteriores de SQL Server. Puede descargarla desde el sitio web de Microsoft. La edición Developer de SQL Server está diseñada para usarse en todos los entornos que no son de producción, como los entornos de desarrollo, pruebas y almacenamiento provisional, siempre que no se conecte a datos de producción en tiempo real.

Si utiliza la edición Developer de SQL Server en entornos que no son de producción, no tiene que pagar costos adicionales por el uso de licencias. Para obtener más información, consulte la sección [Evaluación de la edición Developer de SQL Server](#) de esta guía.

Optimización de la CPU para las cargas de trabajo de SQL Server

En algunos casos, es posible que tengas que elegir un tipo de instancia con CPUs más instancias de las necesarias para tu carga de trabajo debido a otros factores, como los límites de RAM o de red. Sin embargo, AWS ofrece una solución que le ayuda a optimizar los costes de licencia en estas situaciones.

Al igual que la mayoría de los clientes que ofrecen licencias principales de SQL Server, puede deshabilitar el hiperprocesamiento o apagar las CPU de la instancia EC2 para limitar la cantidad de unidades disponibles CPUs para el host. Esta opción le permite aprovechar otras capacidades de la instancia, como la RAM, y, al mismo tiempo, ahorrar costos al no tener que adquirir licencias adicionales.

Por ejemplo, si despliega una instancia r5.4xlarge porque su carga de trabajo requiere 128 GB de memoria pero solo necesita ocho núcleos de SQL Server, puede deshabilitar el hiperprocesamiento de una instancia con solo ocho núcleos activos. CPUs De este modo, puede ahorrar un 50 % en las

licencias de SQL Server necesarias, ya que solo tiene que licenciar los ocho núcleos que se utilizan activamente.

Tipo de instancia	Total v. CPUs	vCPU activa con función de optimización CPUs	Ahorros en licencias de SQL Server
r5.4xlarge	16	8	50%
r5.12xlarge	48	8	83%

La función Optimizar la CPU se puede configurar durante la configuración de lanzamiento de Amazon EC2 o modificando una instancia existente. También se puede aplicar tanto a instancias BYOL como a instancias Amazon EC2 con licencia incluida. Esta flexibilidad le ayuda a ajustar el tamaño de la CPU a las necesidades de su carga de trabajo y, al mismo tiempo, a reducir las licencias. Windows Server SQL Server En el caso de las instancias Amazon EC2 con licencia incluida, la CPUs reducción supone un ahorro instantáneo en los costes de licencia.

Si dimensiona correctamente sus instancias, puede asegurarse de utilizar los tipos de instancias más rentables para sus cargas de trabajo. Al AWS introducir nuevos tipos de instancias, es importante evaluar si estas nuevas instancias pueden satisfacer los requisitos de carga de trabajo con menos núcleos.

Recursos adicionales

- [Amazon Web Services y Microsoft: Preguntas frecuentes](#) (AWS documentación)

Selección de la instancia de EC2 adecuada para las cargas de trabajo de SQL Server

Important

Antes de leer esta sección, le recomendamos que lea las secciones [Explicación de las licencias de SQL Server](#) y [Selección del tipo de instancia adecuado para las cargas de trabajo de Windows](#) de esta guía.

Descripción general de

Microsoft SQL Server lleva más de 15 años ejecutándose en instancias de Amazon Elastic Compute Cloud (Amazon EC2). AWS ha aprovechado esa experiencia y la ha utilizado para ayudar a desarrollar instancias de Amazon EC2 que se adapten a las cargas de trabajo de SQL Server, desde especificaciones mínimas hasta clústeres multirregionales de alto rendimiento.

Elegir la instancia de EC2 correcta para SQL Server depende en gran medida de la carga de trabajo. Comprender cómo se licencia SQL Server, cómo utiliza la memoria y cómo las características de SQL Server se adaptan a las ofertas de Amazon EC2 puede ser útil para dar con la mejor instancia de EC2 para su aplicación.

En esta sección se abordan diversas cargas de trabajo de SQL Server y cómo pueden combinarse con determinadas instancias de EC2 para reducir al mínimo los costos asociados al uso de licencias y de recursos de procesamiento.


Comparación de los costos

Amazon EC2 le permite traer su propia licencia (BYOL) o pagar por uso con las licencias de Windows Server y SQL Server. En el caso de las pay-as-you-go licencias, los costes de las licencias de Windows Server y SQL Server se incluyen en el coste por hora de la instancia EC2. Por ejemplo, puede tener diferentes AMIs precios. El precio de la AMI depende de la edición de SQL Server en la que funcione la AMI.

Los precios de Windows Server y SQL Server no están detallados. No encontrará precios detallados en herramientas como [Calculadora de precios de AWS](#). Si selecciona diferentes combinaciones de ofertas con licencia incluida, se pueden deducir los costos de la licencia, tal y como se muestra en la siguiente tabla.

Instancia de EC2	AMI	Precio de los recursos de computación	Precio de la licencia de Windows	Precio de la licencia de SQL	Precio total
r5.xlarge	Linux (precios de computación)	183,96\$	-	-	183,96 DÓLARES
r5.xlarge	Linux + Edición	183,96 DÓLARES	\$0	\$0	183,96 DÓLARES

Instancia de EC2	AMI	Precio de los recursos de computación	Precio de la licencia de Windows	Precio de la licencia de SQL	Precio total
	Developer de SQL				
r5.xlarge	Windows Server (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	-	318,28 DÓLARES
r5.xlarge	Windows + Edición Developer de SQL	183,96 DÓLARES	134,32 DÓLARES	\$0	318,28 DÓLARES
r5.xlarge	Windows + Edición Web de SQL (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	49,64 DÓLARES	367,92 DÓLARES
r5.xlarge	Windows + Edición Standard de SQL (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	350,4 DÓLARES	668,68 DÓLARES
r5.xlarge	Windows + Edición Enterprise de SQL (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	1095 DÓLARES	1413,28 DÓLARES

 Note

Los precios de la tabla anterior se basan en los precios bajo demanda de la región us-east-1.

El método más rentable para poner en marcha SQL Server consiste en utilizar una edición de nivel inferior hasta que necesite una característica de una edición de nivel superior. Para obtener más información, consulte la sección [Comparación de las ediciones de SQL Server](#) de esta guía. Pasar de la edición Web de SQL Server a la edición Standard de SQL Server multiplica por más de siete el costo adquirir la licencia de SQL Server y por más de tres el costo de pasar de la edición Standard a la edición Enterprise. La disparidad en los costos de las licencias es un factor importante que se debe tener en cuenta; esto se analiza en el resto de la sección.

Escenario de optimización de costos

Imagine un escenario en el que una empresa de análisis que hace un seguimiento de los vehículos de entrega busca mejorar el rendimiento de SQL Server. Una vez que un experto en MACO analiza los problemas de rendimiento de la empresa, la empresa pasa de las instancias x1e.2xlarge a las instancias x2iedn.xlarge. Si bien el tamaño de la instancia es menor, las mejoras introducidas en las instancias x2 mejoran el rendimiento y la optimización de SQL Server mediante el uso de extensiones de grupos de búferes. Esto permitió a la empresa pasar de la edición SQL Server Enterprise a la edición SQL Server Standard y reducir sus licencias de SQL Server de 8 a 4 vCPUs . CPUs

Antes de la optimización:

Server	Instancia de EC2	Edición de SQL Server	Costo mensual
Prod DB1	x1e.2xlarge	Enterprise	3.918,64 DÓLARES
Orguloso DB2	x1e.2xlarge	Enterprise	3.918,64 DÓLARES
Total			7.837,28 DÓLARES

Tras la optimización:

Server	Instancia de EC2	Edición de SQL Server	Costo mensual
Orguloso DB1	x2iedn.xlarge	Standard	1.215,00\$
Orguloso DB2	x2iedn.xlarge	Standard	1.215,00\$

Server	Instancia de EC2	Edición de SQL Server	Costo mensual
Total			2.430,00 DÓLARES

Los cambios combinados de las instancias x1e.2xlarge a las instancias x2iedn.xlarge permitieron al cliente del ejemplo ahorrar 5407 USD al mes en sus servidores de bases de datos de producción. Esto redujo el costo total de la carga de trabajo en un 69 %.

Note

Los precios de la tabla anterior se basan en los precios bajo demanda de la región us-east-1.

Recomendaciones de optimización de costos

instancias optimizadas para memoria

Uno de los aspectos más importantes de SQL Server es comprender que depende de la memoria. SQL Server intenta utilizar toda la RAM disponible que no esté siendo utilizada por el sistema operativo (hasta 2 TB para una instalación predeterminada). Lo hace por motivos de rendimiento. Trabajar con datos en la memoria es mucho más eficaz que tener que extraer datos del disco constantemente, realizar cambios y volver a escribirlos en el disco. En su lugar, SQL Server intenta cargar tantos datos de las bases de datos adjuntas como sea posible y los guarda en la RAM. Los cambios realizados en los datos se producen en la memoria y, posteriormente, se almacenan en el disco.

Note

Para obtener una explicación detallada de cómo SQL Server escribe los cambios, consulte [Writing Pages](#) en la documentación de Microsoft.

Dado que SQL Server funciona mejor con grandes cantidades de RAM, normalmente recomendamos empezar con tipos de instancias [optimizadas para memoria de Amazon EC2](#). Las instancias optimizadas para memoria son versátiles y ofrecen una variedad de opciones diferentes. La familia

R tiene una vCPU-to-RAM proporción de 1 a 8 y cuenta con opciones para procesadores Intel, procesadores AMD, redes mejoradas, rendimiento EBS mejorado, almacenamiento de instancias y velocidad de procesador mejorada. Para cargas de trabajo con un uso intensivo de memoria, también hay una familia X que combina muchas de las mismas opciones y amplía la proporción de 1 a 32. vCPU-to-RAM Gracias a la versatilidad de las instancias optimizadas para memoria, puede aplicarlas a cargas de trabajo de SQL Server de todas las formas y tamaños.

Cargas de trabajo por debajo del mínimo de recursos (menos de 4 v) CPUs

Si bien algunos casos de uso funcionan bien con las instancias ampliables (T3), le recomendamos que, en general, evite el uso de instancias ampliables para las cargas de trabajo de SQL Server. La licencia de SQL Server se basa en la cantidad de v CPUs asignada a una instancia. Si SQL Server está inactivo la mayor parte del día y está adquiriendo créditos de ampliación, tendrá que pagar por las licencias de SQL que no utilice en su totalidad. Además, SQL Server requiere una licencia mínima de 4 núcleos por servidor. Esto significa que si tiene una carga de trabajo de SQL Server que no requiere una potencia de cálculo equivalente a 4 vCPUs , está pagando una licencia de SQL Server que no está utilizando. En estos escenarios, lo mejor sería [unificar varias instancias de SQL Server](#) en un servidor más grande.

Cargas de trabajo con un uso mínimo de recursos (menos de 64 GB de RAM)

Muchas cargas de trabajo de SQL Server de menos de 64 GB de RAM no dan prioridad al alto rendimiento ni a la alta disponibilidad. En este tipo de cargas de trabajo, la edición Web de SQL Server podría ser una buena opción si la aplicación está cubierta por las restricciones de licencia de Microsoft.

Important

La edición Web de SQL Server tiene un caso de uso restringido según los términos de licencia de Microsoft. SQL Server Web Edition solo se puede usar para admitir páginas web, sitios web, aplicaciones web y servicios web públicos y accesibles a través de Internet. No se puede usar para dar soporte a line-of-business aplicaciones (por ejemplo, aplicaciones de administración de relaciones con los clientes, administración de recursos empresariales y otras aplicaciones similares).

La edición Web de SQL Server se amplía hasta 32 V CPUs y 64 GB de RAM y es un 86 por ciento más económica que la edición SQL Server Standard. En el caso de cargas de trabajo con pocos recursos, utilizar una instancia optimizada para memoria de AMD como la r6a, que tiene un precio

de procesamiento un 10 % más económico que su homóloga de Intel, también es una buena forma de reducir al mínimo los costos asociados al uso de licencias de SQL y al uso de recursos de computación.

Cargas de trabajo con un uso medio de recursos (menos de 128 GB de RAM)

La edición Standard de SQL Server se utiliza en la mayoría de las cargas de trabajo de SQL Server de hasta 128 GB de RAM. La edición SQL Server Standard es entre un 65 y un 75 por ciento más económica que la edición SQL Server Enterprise y puede ampliarse hasta 48 vCPUs y 128 GB de RAM. Dado que el límite de 128 GB de RAM suele alcanzarse antes que el límite de 48 vCPU, es el objetivo de la mayoría de los clientes que desean evitar usar la edición Enterprise de SQL Server.

SQL Server tiene una característica llamada [extensión de grupo de búferes](#). Esta característica permite a SQL Server utilizar una parte de un disco como extensión de la RAM. La extensión del grupo de búferes funciona bien cuando se combina con un almacenamiento ultrarrápido, como el que NVMe SSDs se utiliza en el almacenamiento de instancias de [Amazon EC2](#). Las instancias de Amazon EC2 que incluyen almacenamiento de instancias se indican con una “d” en el nombre de la instancia (por ejemplo, r5d, r6id y x2iedn).

Las extensiones de grupos de búferes no son un reemplazo de la RAM normal. Sin embargo, si necesita más de 128 GB de RAM, puede utilizar extensiones de grupos de búferes con instancias de EC2, como r6id.4xlarge y x2iedn.xlarge, para retrasar el tener que adquirir licencias de la edición Enterprise.

Cargas de trabajo de alto rendimiento (más de 128 GB de RAM)

Las cargas de trabajo de SQL Server que requieren un alto rendimiento representan un desafío para la optimización de costos debido a que dependen de una gran cantidad de recursos. Sin embargo, comprender las diferencias entre las instancias de EC2 puede evitar que tome una decisión equivocada.

La siguiente tabla muestra una variedad de instancias de EC2 optimizadas para memoria y sus límites de rendimiento.

	r5b	r6idn	r7iz	x2iedn	x2iezn
Procesador	3.1 GHz	3.5 GHz	3.9 GHz	3.5 GHz	4.5 GHz
	Procesador Intel Xeon de	Procesador Intel Xeon de	Procesador escalable	Procesador Intel Xeon de	Procesador Intel Xeon de

	r5b	r6idn	r7iz	x2iedn	x2iezn
	2. ^a generación	3. ^a generación	Intel Xeon de 4. ^a generación	3. ^a generación	2. ^a generación
Proporción CPU:RAM	1:8	1:8	1:8	1:32	1:32
vCPU máxima	96	128	128	128	48
RAM máxima	768 GB	1024 GB	1024 GB	4096 GB	1536 GB
Almacenamiento de instancias	–	NVMe SSD (4 × 1900 GB)	–	NVMe SSD (2 × 1900 GB)	–
io2 Block Express	Soportado	Soportado	Soportado	compatible	–
IOPS máximas de EBS	260 000	350.000	160 000	260 000	80 000
Rendimiento máximo de EBS	60 Gbps	80 Gbps	40 Gbps	80 Gbps	19 Gbps
Ancho de banda de la red máximo	25 Gbps	200 Gbps	50 Gbps	100 Gbps	100 Gbps

Cada instancia se usa para un propósito diferente. Comprender la carga de trabajo de SQL Server puede ser útil para elegir el tipo de instancia que mejor se adapte a sus necesidades.

Detalles sobre los atributos:

- r5b: el atributo “b” de r5b significa que este tipo de instancia se centra en un alto rendimiento de EBS. En la 5.^a generación de instancias optimizadas para memoria, la r5b era la opción más

popular. Fue el primer tipo de instancia que utilizó los volúmenes de io2 Block Express y alcanzó un máximo de 260 000 IOPS de almacenamiento. El tipo de instancia r5b sigue siendo una alternativa rentable para las necesidades de alto rendimiento de EBS.

- r6idn: la 6.^a generación de instancias optimizadas para memoria ofreció mejoras considerables con respecto a la generación anterior. Las mejoras de rendimiento de EBS derivadas de la r5b van un paso más allá con la r6idn, que eleva el máximo de IOPS a 350 000. La r6idn también tiene un volumen de almacén de instancias para tempdb y las extensiones de grupos de búferes a fin de aumentar aún más el rendimiento de SQL Server.
- x2iedn: la x2iedn es similar a la r6idn. Ofrece niveles similares de EBS mejorado, redes mejoradas y almacenamiento de instancias en NVMe SSD, pero con una vCPU-to-RAM proporción de 1:32 para cargas de trabajo de memoria elevadas y una cantidad de CPU reducida (costes de licencia de SQL Server más bajos).
- x2iezn: el atributo “z” de x2iezn indica que este tipo de instancia se centra en un alto rendimiento del procesador. El procesador Cascade Lake dispone de una frecuencia turbo de hasta 4,5 GHz en todos los núcleos. Le recomendamos que utilice esta instancia EC2, junto con una vCPU-to-RAM proporción de 1:32, en un escenario en el que desee mantener baja la cantidad de vCPU. Esto, a su vez, puede mantener bajos los costos asociados al uso de licencias de SQL Server.
- r7iz: el atributo “z” de r7iz indica que este tipo de instancia se centra en un alto rendimiento del procesador. El procesador Sapphire Rapids tiene una frecuencia turbo integral de hasta 3,9 GHz. Al igual que las instancias x2iezn, el r7iz prioriza el rendimiento del procesador de alta frecuencia, pero con una relación de 1:8. vCPU-to-RAM

Recursos adicionales

- [Instancias Amazon EC2 de uso general \(documentación\)](#) AWS
- [Comparison tool](#) (Vantage)
- [Licencias: SQL Server](#) (AWS documentación)

Unificación de instancias

Esta sección se centra en la técnica de optimización de costos, que consiste en combinar varias instancias de SQL Server en el mismo servidor para minimizar los costos asociados al uso de licencias y maximizar la utilización de los recursos.

Descripción general de

La creación de una instancia forma parte del proceso de instalación del motor de base de datos de SQL Server. La instancia de SQL Server es una instalación completa que contiene sus propios archivos de servidor, inicios de sesión de seguridad y bases de datos del sistema (master, model, msdb y tempdb). Como una instancia tiene todos sus propios archivos y servicios, puede instalar varias instancias de SQL Server en el mismo sistema operativo sin que las instancias interfieran entre sí. Sin embargo, dado que las instancias están instaladas en el mismo servidor, todas comparten los mismos recursos de hardware, como el procesamiento, la memoria y las redes.

Lo habitual es utilizar solo una instancia de SQL Server por servidor en los entornos de producción, de modo que una instancia “ocupada” no utilice en exceso los recursos de hardware compartidos. Asignar a cada instancia de SQL Server su propio sistema operativo, con recursos independientes, constituye un límite mejor que depender únicamente de la gobernanza de recursos. Esto es especialmente cierto en el caso de las cargas de trabajo de SQL Server de alto rendimiento que requieren grandes cantidades de recursos de RAM y CPU.

Sin embargo, no todas las cargas de trabajo de SQL Server utilizan una gran cantidad de recursos. Por ejemplo, algunas organizaciones asignan a cada uno de sus clientes su propia instancia de SQL Server dedicada por motivos de cumplimiento o seguridad. En el caso de clientes más pequeños o clientes que no suelen estar activos, eso significa poner en marcha las instancias de SQL Server con un mínimo de recursos.

Como se indica en [Microsoft SQL Server 2019: Licensing guide](#), cada servidor que use SQL Server debe tener un mínimo de cuatro licencias de CPU. Esto significa que, incluso si ejecuta un servidor con solo dos vCPUs, debe seguir licenciando SQL Server para cuatro versiones CPUs. Según el [precio público de SQL Server de Microsoft](#), esto supone una diferencia de 3.945\$ si utilizas la edición estándar de SQL Server. En el caso de las organizaciones que utilizan varios servidores con una única instancia de SQL Server y utilizan un mínimo de recursos, el costo combinado de tener que licenciar los recursos no utilizados puede ser considerable.

Escenario de optimización de costos

En esta sección se analiza un escenario de ejemplo en el que se compara la diferencia entre poner en marcha cuatro servidores de Windows Server, cada uno con una única instancia de SQL Server, y poner en marcha un único servidor de Windows Server de mayor tamaño que usa varias instancias de SQL Server simultáneamente.

Si cada instancia de SQL Server solo necesita dos v CPUs y 8 GB de RAM, el coste total por servidor es de 7.890\$ para la licencia de SQL Server, además de un coste informático por hora de 0,096\$.

Instancia de EC2	v CPUs	RAM	Precio	v CPUs para licenciar	Costo total de las licencias de SQL Server
m6i.large	2	8	0.096	4	7.890 DÓLARES

Si se amplía esta cifra a cuatro servidores, el costo total de la licencia de SQL Server es de 31 560 USD, con un costo asociado al uso de recursos de computación por hora de 0,384 USD.

Instancia de EC2	v CPUs	RAM	Precio	v CPUs para licenciar	Costo total de las licencias de SQL Server
4 m6i.large	2	32	0.384	16	31.560 DÓLARES

Si combina las cuatro instancias de SQL Server en una sola instancia de EC2, la cantidad total de recursos de computación y de computación permanece igual. Sin embargo, al eliminar los costos innecesarios de licencias de SQL Server, puede reducir el costo total de poner en marcha la carga de trabajo en 15 780 USD.

Instancia de EC2	v CPUs	RAM	Precio	v CPUs para licenciar	Costo total de las licencias de SQL Server
m6i.2xlarge	8	32	0.384	8	15.780 DÓLARES

Note

En las tablas anteriores, los costos asociados al uso de recursos de computación muestran los precios por hora y bajo demanda para los servidores de Amazon EC2 que usan Windows Server en la región us-east-1. Los costos asociados al uso de licencias de la edición Standard de SQL Server se basan en la información de [Precios y licencias de SQL Server 2022](#).

Recomendaciones de optimización de costos

Si está pensando en unificar las instancias de SQL Server, el mayor problema al que se enfrenta es el consumo de recursos de cada una de las instancias que desee unificar. Es importante obtener métricas de rendimiento durante periodos prolongados para comprender mejor los patrones de carga de trabajo de cada servidor. Algunas herramientas comunes para la supervisión del consumo de recursos son [Amazon CloudWatch](#), [Windows Performance Monitor](#) (perfmon) y las [herramientas de supervisión nativas](#) de SQL Server.

Le recomendamos que tenga en cuenta las siguientes preguntas al analizar si las cargas de trabajo de SQL Server se pueden combinar para utilizar los mismos recursos del servidor sin que interfieran entre sí:

- ¿Qué recursos (CPU, memoria y ancho de banda de la red) se consumen durante el estado estable?
- ¿Qué recursos (CPU, memoria y ancho de banda de la red) se consumen durante los picos?
- ¿Con qué frecuencia se producen los picos? ¿Los picos son consistentes?
- ¿Los picos de recursos de un servidor coinciden con los picos de recursos de otro servidor?
- ¿Qué capacidad de IOPS y rendimiento de almacenamiento requiere SQL Server?

Si desea seguir adelante con un plan para combinar instancias de SQL Server, consulte la publicación [Run multiple instances of SQL Server on one Amazon EC2 instance](#) en el blog de operaciones en la nube y migraciones de AWS . En esta publicación, se proporcionan instrucciones sobre cómo realizar los cambios de configuración en SQL Server para agregar instancias adicionales. Antes de empezar, tenga en cuenta las pequeñas diferencias que se producen cuando se instalan varias instancias en el mismo servidor:

- La instancia de base de datos de SQL Server predeterminada se denomina MSSQLSERVER y utiliza el puerto 1433.
- Cada instancia adicional instalada en el mismo servidor es una instancia de base de datos “con nombre”.
- Cada instancia con nombre tiene un nombre de instancia y un puerto únicos.
- El [navegador de SQL Server](#) debe ponerse en marcha para coordinar el tráfico a las instancias con nombre.
- Cada instancia puede usar ubicaciones distintas para los archivos de datos de la base de datos y los inicios de sesión independientes.
- La [configuración máxima de memoria](#) de SQL Server debe configurarse de acuerdo con las necesidades de rendimiento de cada instancia, y su total combinado también debe dejar suficiente memoria para el sistema operativo subyacente.
- Puede utilizar las capacidades de [copia de seguridad y restauración](#) nativas de SQL Server o [AWS DMS](#) para la migración o unificación.

Recursos adicionales

- [Hoja de datos de licencias de SQL Server](#) (blog sobre operaciones y migraciones en la AWS nube)
- Entrada del blog sobre la [configuración de instancias múltiples de SQL Server](#) (blog sobre operaciones y migraciones en la AWS nube)

Comparación de las ediciones de SQL Server

Descripción general de

Las licencias de Microsoft SQL Server son uno de los mayores gastos de un entorno de carga de trabajo de Windows. Los costos asociados al uso de licencias de SQL Server pueden ampliarse fácilmente más allá de los costos derivados del uso de recursos de computación necesarios para poner en marcha la carga de trabajo. Si elige la edición equivocada, podría pagar por características que no utiliza o que ni siquiera necesita. En esta sección se comparan las siguientes ediciones de SQL Server, incluidas sus características y costos relativos:

- Enterprise: la edición Enterprise de SQL Server proporciona capacidades de centro de datos con alto rendimiento, virtualización ilimitada y varias herramientas de inteligencia empresarial (BI).

- **Standard:** la edición Standard de SQL Server proporciona administración de datos básica e inteligencia empresarial para organizaciones y departamentos más pequeños.
- **Web:** la edición web de SQL Server es adecuada para empresas que ofrecen alojamiento web o proveedores de valor agregado web (VAPs). Esta edición ofrece un costo total de propiedad bajo y proporciona capacidades de escalabilidad y capacidad de administración para propiedades web de pequeña y gran escala.

Important

La edición web de SQL Server solo se puede usar para respaldar páginas web, sitios web, aplicaciones web y servicios web públicos y accesibles a través de Internet. No puede utilizar la edición web de SQL Server para dar soporte a line-of-business aplicaciones (como las aplicaciones de gestión de relaciones con los clientes o de gestión de recursos empresariales).

- **Developer:** la edición Developer de SQL Server incluye todas las funciones de la edición Enterprise, pero está destinada únicamente a fines de desarrollo.
- **Express:** la edición Express de SQL Server es una base de datos gratuita y se puede utilizar para aprender o para crear aplicaciones de escritorio. Puede actualizar la edición Express a otras ediciones.

Note


La edición Evaluation de SQL Server está disponible durante un periodo de prueba de 180 días.

Impacto del costo

Puede adquirir licencias de SQL Server de un distribuidor de Microsoft y usarlas en AWS con Software Assurance. Como alternativa, puede usar licencias de SQL Server con un pay-as-you-go modelo que tenga Amazon EC2 incluida en la licencia. AMIs

Si compra licencias de SQL Server a distribuidores de Microsoft, las licencias principales se venden en paquetes de dos y debe adquirir licencias para un mínimo de cuatro núcleos por servidor. La siguiente tabla muestra una comparación entre los costos de las ediciones Enterprise y Standard.

Versión	Edición Enterprise de SQL Server (paquete de 2 núcleos)	Edición Standard de SQL Server (paquete de 2 núcleos)	Ahorros
2022	15.123 DÓLARES	3.945 DÓLARES	74%
2019	13.748 DÓLARES	3.586 DÓLARES	74%

 Note

Los precios de la tabla anterior se basan en los precios públicos de Microsoft para [SQL Server 2022](#) y [SQL Server 2019](#).

La siguiente comparación de costos muestra el alojamiento de diferentes ediciones de SQL Server con Amazon EC2 con licencia incluida. AMIs En esta comparación, SQL Server está alojado en r6i.xlarge (4 vCPU) en la región us-east-1.

Instancia	Costo de los recursos de computación	Costo de las licencias de Windows	Costos de las licencias de SQL Server	Total
R6i.xlarge (Linux)	183,96\$	–	–	183,96 DÓLARES
R6i.xlarge + Windows	183,96 DÓLARES	134,32 DÓLARES	–	318,28 DÓLARES
R6i.xlarge + Edición Web de SQL Server	183,96 DÓLARES	134,32 DÓLARES	49,35 DÓLARES	367,63 DÓLARES
R6i.xlarge + Edición Standard de SQL Server	183,96 DÓLARES	134,32 DÓLARES	350,4 DÓLARES	668,68 DÓLARES

Instancia	Costo de los recursos de computación	Costo de las licencias de Windows	Costos de las licencias de SQL Server	Total
R6i.xlarge + Edición Enterprise de SQL Server	183,96 DÓLARES	134,32 DÓLARES	1.095 DÓLARES	1.413,28 DÓLARES

Puede ahorrar hasta un 95 % en los costos asociados al uso de licencias de SQL Server si selecciona la edición de SQL Server adecuada para su carga de trabajo. En la siguiente tabla se compara el costo de las licencias de SQL Server en las instancias r6i.xlarge.

Edition	% de ahorro
Standard en comparación con Enterprise	68%
Web en comparación con Standard	86%
Web en comparación con Enterprise	95%

En la mayoría de los escenarios, las organizaciones cambian de la edición Enterprise a la Standard, pero hay algunos casos en los que es posible cambiar de la edición Standard o Enterprise a la edición web.

Recomendaciones de optimización de costos

Puede elegir la mejor edición para su carga de trabajo en función de los límites de escalado, la alta disponibilidad, el rendimiento y la seguridad. En la siguiente tabla se muestran las características compatibles con todas las ediciones de SQL Server. Esto puede ser de ayuda para decidir qué edición usar. Esta comparación se aplica a [SQL Server 2016 SP1 y versiones posteriores](#).

Límites de escalado

En la siguiente tabla se comparan los límites de escalado de las distintas ediciones de SQL Server.

Característica	Enterprise Edition	Standard Edition	Edición Web	Edición Express
Capacidad de computación máxima utilizada por una sola instancia de SQL Server Database Engine, SQL Server Analysis Services o SQL Server Reporting Services	Máximo de sistemas operativos	Limitado al menor valor entre 4 sockets o 24 núcleos	Limitado al menor valor entre 4 sockets o 16 núcleos	Limitado al menor valor entre 4 sockets o 4 núcleos
Memoria máxima para el grupo de búferes por instancia de SQL Server Database Engine	Máximo de sistemas operativos	128 GB	64 GB	1410 MB
Capacidad máxima de extensión de grupo de búferes por instancia de SQL Server Database Engine	Hasta 32 veces la memoria máxima configurada	Hasta 4 veces la memoria máxima configurada	N/A	N/A
Tamaño máximo de la base de datos relacional	524 PB	524 PB	524 PB	10 GB
Memoria máxima para las memorias caché de Columnstore o los datos optimizados para memoria	Máximo de sistemas operativos	32 GB	16 GB	352 MB

Si su aplicación requiere menos de 16 núcleos (32 vCPUs) y 64 GB de RAM, puede empezar a evaluarla desde la edición web de SQL Server. Si su carga de trabajo requiere más de 64 GB de memoria u otras opciones de alta disponibilidad, debe actualizar a la edición Standard de SQL Server.

Puede utilizar la edición Web de SQL Server para respaldar páginas web, sitios web, aplicaciones web y servicios web públicos y accesibles a través de Internet, pero no puede utilizar la edición Web de SQL Server para respaldar aplicaciones de línea de negocio. Para obtener más información acerca de los casos de uso de la edición Web de SQL Server, contacte con el [equipo de soporte para licencias de Microsoft](#) o con su distribuidor de Microsoft.

Puede usar la edición SQL Server Standard para cargas de trabajo de hasta 24 núcleos (48 vCPUs) y 128 GB de memoria. Sin embargo, puede usar [extensiones de grupos de búferes](#) para permitir que la edición Standard de SQL Server utilice el [almacenamiento de instancias local](#), como las presentes en las instancias r6id de EC2. Esto amplía hasta cuatro veces la memoria máxima configurada. Esta combinación de características puede retrasar la actualización de un servidor a la edición Enterprise cuando los requisitos de memoria comiencen a aumentar.

Para identificar el uso de la memoria, consulte las páginas de las bases de datos en el grupo de búferes y los contadores de [esperanza de vida útil de la página](#). La esperanza de vida de la página indica cuánto tiempo permanece la página en la memoria antes de volver a guardarse en el disco. El valor predeterminado de este contador es de 300. Si las páginas permanecen en la memoria durante horas o días, existe la posibilidad de que se reduzca la memoria asignada.

Alta disponibilidad

En la siguiente tabla se comparan las capacidades de alta disponibilidad de las distintas ediciones de SQL Server.

Característica	Enterprise Edition	Standard Edition	Edición Web	Edición Express
Soporte para Server Core (1)	Sí	Sí	Sí	Sí
Envío de registros	Sí	Sí	Sí	No

Característica	Enterprise Edition	Standard Edition	Edición Web	Edición Express
Duplicación de bases de datos	Sí	Modo de seguridad COMPLETO	Solo como testigo	Solo como testigo
Compresión de copias de seguridad	Sí	Sí	No	No
Instancias de clústeres de conmutación por error Always On	16 nodos	2 nodos	No	No
Grupos de disponibilidad AlwaysOn	Hasta 8 réplicas secundarias, lo que incluye 2 réplicas secundarias sincrónicas	No	No	No
Grupos de disponibilidad básica	No	2 nodos	No	No
Restauración de páginas y archivos en línea	Sí	No	No	No
Indexación en línea	Sí	No	No	No
Cambio de esquema en línea	Sí	No	No	No

Característica	Enterprise Edition	Standard Edition	Edición Web	Edición Express
Recuperación rápida	Sí	No	No	No
Copias de seguridad duplicadas	Sí	No	No	No
Adición en caliente de memoria y CPU	Sí	No	No	No
Copia de seguridad cifrada	Sí	Sí	No	No
Copia de seguridad híbrida en Microsoft Azure (copia de seguridad en URL)	Sí	Sí	No	No
Servidor de conmutación por error para la recuperación ante desastres	Sí	Sí	No	No
Servidores de conmutación por error para una alta disponibilidad	Sí	Sí	No	No

Otras características comunes

En la siguiente tabla se comparan las características más comunes de las diferentes ediciones de SQL Server. Para obtener una lista completa de características, consulte [Editions and supported features of SQL Server 2019](#) en la documentación de Microsoft.

Característica	Enterprise Edition	Standard Edition	Edición Web	Edición Express
(Rendimiento) Organizador de recursos	Sí	No	No	No
(Seguridad) Cifrado de bases de datos transparente (TDE)	Sí	Sí	No	No
(Seguridad) Administración ampliable de claves (EKM)	Sí	No	No	No
(Replicación) Publicación en Oracle	Sí	No	No	No
(Replicación) Replicación transaccional punto a punto	Sí	No	No	No
Captura de datos de cambios	Sí	Sí	No	No

Edición Developer de SQL Server

Todas las cargas de trabajo que no sean de producción, como los entornos de desarrollo, control de calidad, pruebas, almacenamiento provisional y UAT, pueden utilizar la edición Developer de SQL Server para ahorrar un 100 % en los costos de licencia de SQL Server. Tras [descargar SQL Server](#), puede instalar la edición Developer de SQL Server en las instancias de EC2 mediante una tenencia compartida. No se necesita una infraestructura dedicada para la edición Developer de SQL Server. Para obtener más información, consulte la recomendación de esta guía para la [edición Developer de SQL Server](#).

Cambio de edición

En el caso de las cargas de trabajo existentes, el cambio de una edición a otra requiere pruebas exhaustivas. Se recomienda comprobar las cargas de trabajo que usan las ediciones Enterprise o Standard para comprobar si se utilizan características específicas de cada edición y si hay alguna solución alternativa para esas características. Por ejemplo, si desea comprobar si sus bases de datos utilizan alguna característica de la edición Enterprise, puede poner en marcha [vistas de administración dinámica \(DMV\)](#) en todas las bases de datos, tal y como se muestra en el siguiente comando de ejemplo:

```
SELECT feature_name FROM sys.dm_db_persisted_sku_features; GO
```

Hay algunas características de la edición Enterprise que no se pueden capturar en T-SQL, como la reindexación en línea como parte de las tareas de mantenimiento de SQL. Estas deben verificarse manualmente.

Consideraciones sobre la migración

La forma en que asigne las licencias de SQL Server determinará sus opciones para cambiar de edición. El costo de la licencia de las AMI, como las AMI de SQL Server, está incluido en el precio de la instancia de EC2; el costo de la licencia está vinculado a la AMI. Puede utilizar los [códigos de facturación de AWS](#) para comprobar la versión de SQL Server incluida en la AMI. En el caso de las instancias con licencia incluida de AWS, cambiar la edición de SQL Server en el sistema operativo no cambiará la facturación asociada a la AMI. Debe migrar las bases de datos a una nueva instancia de EC2 con una AMI que use la nueva edición de SQL Server.

Si usa su propia licencia, dispondrá de más flexibilidad. Por lo general, se recomienda migrar a otra instancia de EC2 que use la nueva versión. De este modo, podrá usar una conmutación por recuperación sencilla si algo no sale según lo planeado. Sin embargo, si debe usar el servidor

existente, aún puede realizar una side-by-side instalación de SQL Server y migrar las bases de datos entre instancias. Para obtener información más detallada sobre side-by-side las versiones anteriores, consulte [Actualización y degradación de ediciones en SQL Server en](#) el MSSQLTips sitio web.

Recursos adicionales

- [Editions and supported features of SQL Server 2022](#) (Microsoft Learn)
- [sys.dm_db_persisted_sku_features \(Transact-SQL\)](#) (Microsoft Learn)
- [Which Version of SQL Server Should You Use?](#) (Brent Ozar Unlimited)
- [Calculadora de precios de AWS](#) (AWS)

Evaluación de la edición Developer de SQL Server

Descripción general de

La [edición Developer de SQL Server](#) es una edición gratuita de SQL Server que contiene todas las características de la edición Enterprise y se puede utilizar en cualquier entorno que no sea de producción. En la nube, donde no se pueden usar las licencias de Microsoft Developer Network (MSDN), la edición Developer de SQL Server es una buena forma de ahorrar costos sin tener que obtener licencias para las cargas de trabajo de desarrollo y prueba. Esto es especialmente cierto en el caso de los equipos que gestionan grandes entornos de desarrollo y pruebas y buscan reducir costos innecesarios.

Un entorno de producción se define como un entorno al que acceden los usuarios finales de una aplicación (como un sitio web de Internet) y que se utiliza para algo más que recopilar comentarios o realizar pruebas de aceptación de esa aplicación. Algunos otros ejemplos de escenarios que constituyen entornos de producción son los siguientes:

- Entornos que se conectan a una base de datos de producción.
- Entornos que admiten la recuperación ante desastres o la creación de copias de seguridad para un entorno de producción.
- Entornos que se utilizan para la producción al menos una parte del tiempo, como un servidor que pasa a estar en producción durante los periodos de mayor actividad

Para obtener más información, consulte [Amazon Web Services and Microsoft: Frequently Asked Questions](#) en la documentación de AWS .

Impacto del costo

Si usa la edición Developer de SQL Server para cargas de trabajo que no son de producción, puede ahorrar el 100 % de los costos actuales asociados al uso de licencias de SQL Server para entornos de desarrollo y pruebas.

Versión de SQL Server	Edición Enterprise de SQL Server (paquete de 2 núcleos)	Edición Standard de SQL Server (paquete de 2 núcleos)	Edición Developer de SQL Server
2022	15.123 DÓLARES	3.945 DÓLARES	Free
2019	13.748 DÓLARES	3.586 DÓLARES	Free

Note

Los precios de la tabla anterior se basan en los precios públicos de Microsoft para [SQL Server 2022](#) y [SQL Server 2019](#).

En la siguiente tabla se compara el costo de las diferentes ediciones de SQL Server que funcionan con la versión 4 CPUs y utilizan los precios bajo demanda en la us-east-2 región. Esto se aplica a los escenarios que dependen de instancias con licencia incluida de AWS

Instancia de EC2	AMI	Precio de los recursos de computación	Precio de la licencia de Windows	Precio de la licencia de SQL Server	Precio total
r5.xlarge	Linux (precios de computación)	183,96\$	–	–	183,96 DÓLARES
r5.xlarge	Linux + Edición Developer de SQL Server	183,96 DÓLARES	\$0	\$0	183,96 DÓLARES

Instancia de EC2	AMI	Precio de los recursos de computación	Precio de la licencia de Windows	Precio de la licencia de SQL Server	Precio total
r5.xlarge	Windows Server (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	–	318,28 DÓLARES
r5.xlarge	Windows + Edición Developer de SQL Server	183,96 DÓLARES	134,32 DÓLARES	\$0	318,28 DÓLARES
r5.xlarge	Windows + Edición Web de SQL Server (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	49,64 DÓLARES	367,92 DÓLARES
r5.xlarge	Windows + Edición Standard de SQL Server (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	350,4 DÓLARES	668,68 DÓLARES
r5.xlarge	Windows + Edición Enterprise de SQL Server (licencia incluida)	183,96 DÓLARES	134,32 DÓLARES	1095 DÓLARES	1413,28 DÓLARES

Escenario de optimización de costos

Después de que una empresa de integridad de datos realizara una nueva adquisición, quiso migrar la carga de trabajo recién adquirida de su ubicación actual a un proveedor de alojamiento

administrado para unificarla con las demás cargas de trabajo en la Nube de AWS. Los precios iniciales mostraron que la carga de trabajo de SQL Server de la empresa costaría un 60 por ciento más de funcionamiento que con el AWS proveedor de servicios gestionados actual. Una pyme de MACO evaluó la previsión y descubrió que, en realidad, el cliente estaba pagando las licencias de SQL Server al proveedor de alojamiento administrado para sus entornos de desarrollo y pruebas. Al cambiar las cargas de trabajo no relacionadas con la producción a la edición Developer de SQL Server durante la migración, la empresa redujo la adquisición de licencias de SQL Server en un 40 %.

Licencia de SQL Server incluida en Amazon EC2

Si tiene instancias de SQL Server en EC2 que utilizan la [licencia incluida AMIs](#), no es posible realizar una conversión directa de la edición Enterprise a la edición Developer. Los costos de licencia de las instancias con licencia incluida están sujetos a la AMI. Aunque SQL Server se desinstale desde el sistema operativo, se seguirán cobrando los costos de licencia a la instancia de EC2.

Para pasar a la edición Developer, debe [descargar la edición Developer de SQL Server](#), instalarla en una nueva instancia de EC2 y, a continuación, migrar las bases de datos. Puede migrar bases de datos de SQL Server entre instancias de EC2 mediante diversos métodos. Para obtener más información, consulte [SQL Server database migration methods](#) en la guía *Migrating Microsoft SQL Server databases to the Nube de AWS*. También puede usar la solución [Automated SQL Server Developer](#) para preparar la nueva instancia a la que planea migrar.

Traiga su propia licencia (BYOL) de SQL Server a Amazon EC2

Si tiene instancias de SQL Server que utilizan un BYOL, puede elegir entre las siguientes opciones de conversión o degradación in situ: side-by-side

- Descargue la [edición Developer de SQL Server](#) desde el sitio web de Microsoft. Para obtener instrucciones para la instalación manual o automática, consulte la publicación [Automating SQL Server Developer deployments](#) en el blog de AWS .
- Utilice el [backup y la restauración nativos de SQL Server](#) para migrar bases de datos o detach/attach bases de datos de una instancia de SQL a otra.
- Utilice una [herramienta de automatización](#) para las implementaciones por lotes.

Note

La edición Developer de SQL Server es solo para entornos que no son de producción.

Recursos adicionales

- [Automatización de las implementaciones de SQL Server Developer para implementar SQL Server Developer Edition en EC2 \(blog\)](#) AWS
- [Precios de SQL 2022](#) (Microsoft)
- [Precios de SQL 2019](#) (Microsoft)
- [Licensing options](#) (SQL Server en Amazon EC2)
- [Calculadora de precios de AWS](#) (documentación de SQL Server en Amazon EC2)
- [Microsoft SQL Server 2019 Licensing guide](#) (descarga de Microsoft)
- [SQL Server 2022 Developer edition](#) (descarga de Microsoft)

Evaluación de SQL Server en Linux

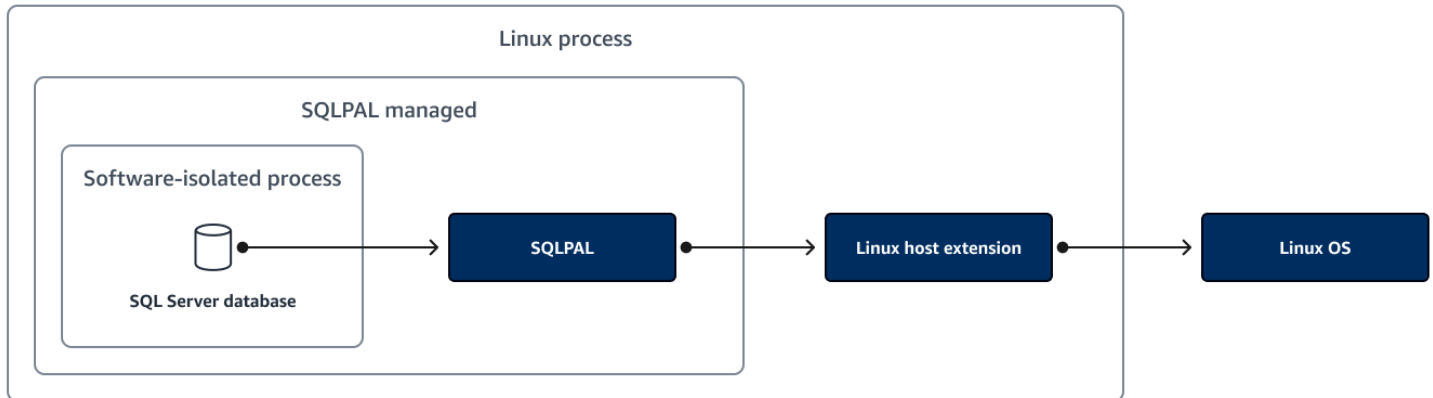
Descripción general de

Desde SQL Server 2017, es posible instalar SQL Server en sistemas operativos Linux. SQL Server en Linux está preparado para empresas y ofrece flexibilidad, alto rendimiento, funciones de seguridad, reducción del coste total de propiedad, HA/DR funciones y una excelente experiencia de usuario. Puede cambiar de SQL Server en Windows Server a SQL Server en Linux para ahorrar en los costos asociados al uso de licencias de Windows Server.

En el caso de Linux, SQL Server está disponible para implementarse en Red Hat Enterprise Linux (RHEL), SUSE Linux Enterprise Server (SLES), Ubuntu y Amazon Linux 2. El motor de base de datos de SQL Server funciona de la misma manera tanto en Windows Server como en Linux, pero hay algunos cambios fundamentales en ciertas tareas al utilizar Linux. Una diferencia clave entre la puesta en marcha de aplicaciones de SQL Server Always On en Linux y Windows está relacionada con los clústeres de conmutación por error. Si implementa grupos de disponibilidad Always On en un host de Windows Server, puede aprovechar los [clústeres de conmutación por error de Windows Server \(WSFC\)](#) y Active Directory como características integradas que admiten los clústeres de conmutación por error. Sin embargo, ni WSFC ni

Active Directory están disponibles en Linux para respaldar los clústeres de conmutación por error. [Si desea lanzar clústeres de conmutación por error para SQL Server en Linux, puede utilizar Pacemaker AWS Launch Wizard para simplificar la configuración del clúster y la instalación de SQL en instancias de Linux. ClusterLabs](#)

SQL Server en Windows y Linux comparten una base de código común. Es decir, el motor principal de SQL Server no se modificó en absoluto para que funcione en Linux. SQL Server introdujo una Platform Abstraction Layer (capa de abstracción de plataforma, SQLPAL), tal y como se muestra en el siguiente diagrama.



La SQLPAL es responsable de la abstracción de las llamadas y la comunicación entre SQL Server y el sistema operativo subyacente. La extensión del host es simplemente una aplicación nativa de Linux. Las funciones de bajo nivel del sistema operativo son llamadas nativas para optimizar la E/S, la memoria y el uso de la CPU. Cuando se inicia la extensión del host, carga e inicializa la SQLPAL, que luego abre SQL Server. La SQLPAL lanza procesos de software aislados que proporcionan la traducción necesaria para el resto del código. La adición de esta nueva capa a la arquitectura de SQL Server significa que las mismas características y beneficios principales de la edición Enterprise que hacen que SQL Server sea tan potente en Windows estén disponibles independientemente del sistema operativo.

Impacto del costo

En el caso de las instancias r5.2xlarge, la reducción del costo asociado al uso de licencias de Windows Server es de unos 268 USD en cada escenario. La reducción supone un porcentaje mayor del costo total del servidor en comparación con el uso de ediciones más baratas de SQL Server. En la tabla siguiente, se muestran los ahorros.

Instancia	Edition	Costo mensual de SQL Server en Windows	Costo mensual de SQL Server en Linux	Ahorros
r5.2xlarge	Web	735 DÓLARES	466 DÓLARES	37%
r5.2xlarge	Standard	1.337 DÓLARES	1.068 DÓLARES	20%
r5.2xlarge	Enterprise	2.826 DÓLARES	2.558 DÓLARES	10%

Note

La estimación de precios de la tabla anterior se basa en los precios bajo demanda de la región us-east-1 y se puede consultar directamente en [Calculadora de precios de AWS](#).

Imagine un escenario en el que el cliente, una pyme, de un ISV busca ahorrar costos en su entorno de desarrollo. Ya utiliza la edición Developer de SQL Server en un conjunto de servidores Windows. Al cambiar de Windows con la edición Developer de SQL Server a Linux con la edición Developer de SQL Server, el cliente de un ISV puede ahorrar un 33 % en su carga de trabajo de desarrollo. La siguiente tabla muestra los siguientes costos estimados para este escenario.

Estimación	Costo mensual
Windows + SQL Server	9.307,72 DÓLARES
Linux + SQL Server	6.218,36 DÓLARES
Ahorro de costos estimado	3089,36 USD (33 %)

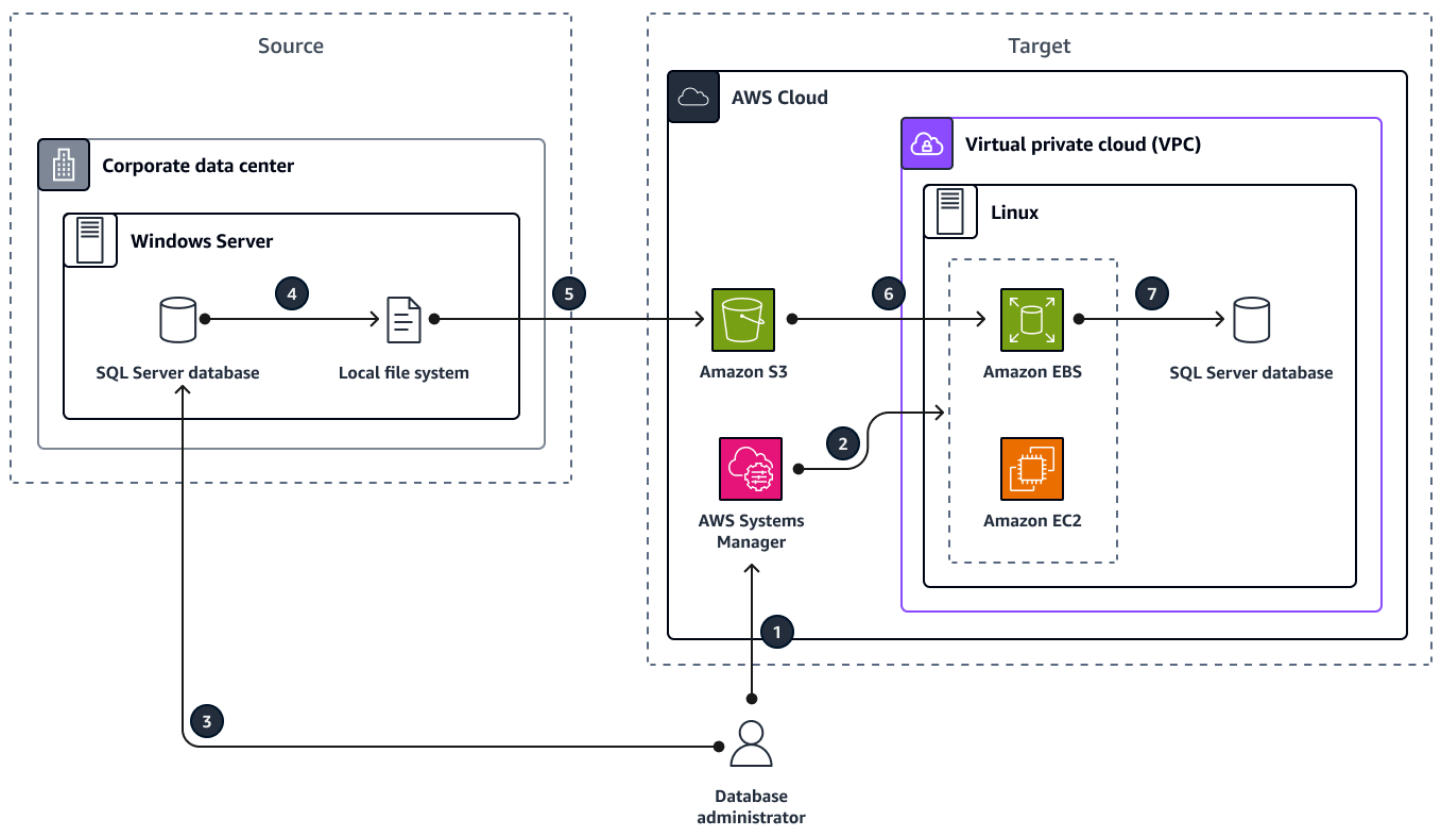
En otro escenario de ejemplo, una empresa migra las instancias de EC2 de SQL Server con licencia incluida de Windows a Linux. La empresa ahorra un total de 300 000 USD al año en costos asociados al uso de licencias de Windows Server, aproximadamente el 20 % de su factura de AWS total.

Recomendaciones de optimización de costos

Le recomendamos que tenga en cuenta lo siguiente:

- A partir de SQL Server 2017, se admite SQL Server en Linux.
- Para poder realizar el cambio, puede utilizar el [asistente para redefinir la plataforma de Windows a Linux para las bases de datos de Microsoft SQL Server](#). El asistente para redefinir la plataforma es una herramienta de creación de scripts útil para mover las cargas de trabajo de SQL Server existentes de los sistemas operativos Windows a Linux comprobando las incompatibilidades comunes, exportando las bases de datos del host de Windows y, a continuación, importándolas a una instancia de EC2 que usa Microsoft SQL Server 2017 en Ubuntu 16.04.
- También puede utilizar las funciones de [copia de seguridad y restauración](#) de SQL Server para cambiar de SQL Server en Windows a Linux.
- Puede implementar fácil y rápidamente SQL Server en Linux o Ubuntu con [AWS Launch Wizard](#). Launch Wizard puede implementar SQL Server en Linux o Ubuntu tanto en escenarios independientes como de alta disponibilidad en función de las necesidades de la aplicación. Para obtener más información, consulte la publicación [Implementación en SQL Server Always on Linux AWS Launch Wizard with](#) en el AWS blog Microsoft Workloads on.

En el siguiente diagrama, se muestra la arquitectura de una solución que utiliza el asistente para redefinir la plataforma de Windows a Linux para las bases de datos de Microsoft SQL Server.



Recursos adicionales

- [Overview of SQL Server on Linux](#) (Microsoft Learn)
- [Installation guide for SQL Server on Linux](#) (Microsoft Learn)
- [Implementación en SQL Server Always en Linux con AWS Launch Wizard](#) (Microsoft Workloads en el AWS blog)
- [SQL Server de alta disponibilidad en Linux](#) (blog de código AWS abierto)

Optimización de las estrategias de copia de seguridad de SQL Server

Descripción general de

La mayoría de las organizaciones buscan la solución adecuada para proteger sus datos de SQL Server en [Amazon EC2](#) para cumplir los requisitos actuales de objetivo de punto de recuperación (RPO), que es el tiempo máximo aceptable desde la última copia de seguridad, y el objetivo de tiempo de recuperación (RTO), que es el tiempo máximo aceptable entre la interrupción

y la restauración del servicio. Cuando usa SQL Server en instancias de EC2, tiene varias opciones para crear copias de seguridad de los datos y también para restaurarlos. Algunas estrategias de copia de seguridad para proteger los datos de SQL Server en Amazon EC2 incluyen las siguientes:

- Copia de seguridad a nivel del servidor mediante instantáneas de [Amazon Elastic Block Store \(Amazon EBS\)](#) habilitadas para Windows Volume Shadow Copy Service (VSS) o [AWS Backup](#)
- Copia de seguridad de bases de datos mediante las capacidades de [copia de seguridad y restauración nativas](#) de SQL Server

Dispone de las siguientes opciones de almacenamiento para [realizar copias de seguridad nativas de las bases de datos](#):

- Una copia de seguridad local con un [volumen de Amazon EBS](#)
- Una copia de seguridad del sistema de archivos en red con [Amazon FSx para Windows File Server](#) o Amazon FSx para NetApp ONTAP
- Una copia de seguridad de red en Amazon Simple Storage Service (Amazon S3) mediante [AWS Storage Gateway](#)
- Copia de seguridad directa en Amazon S3 para SQL Server 2022

Esta sección hace lo siguiente:

- Destaca las características que le permiten ahorrar espacio de almacenamiento.
- Compara los costos entre las diferentes opciones de almacenamiento de backend.
- Proporciona enlaces a documentación detallada para que pueda implementar estas recomendaciones.

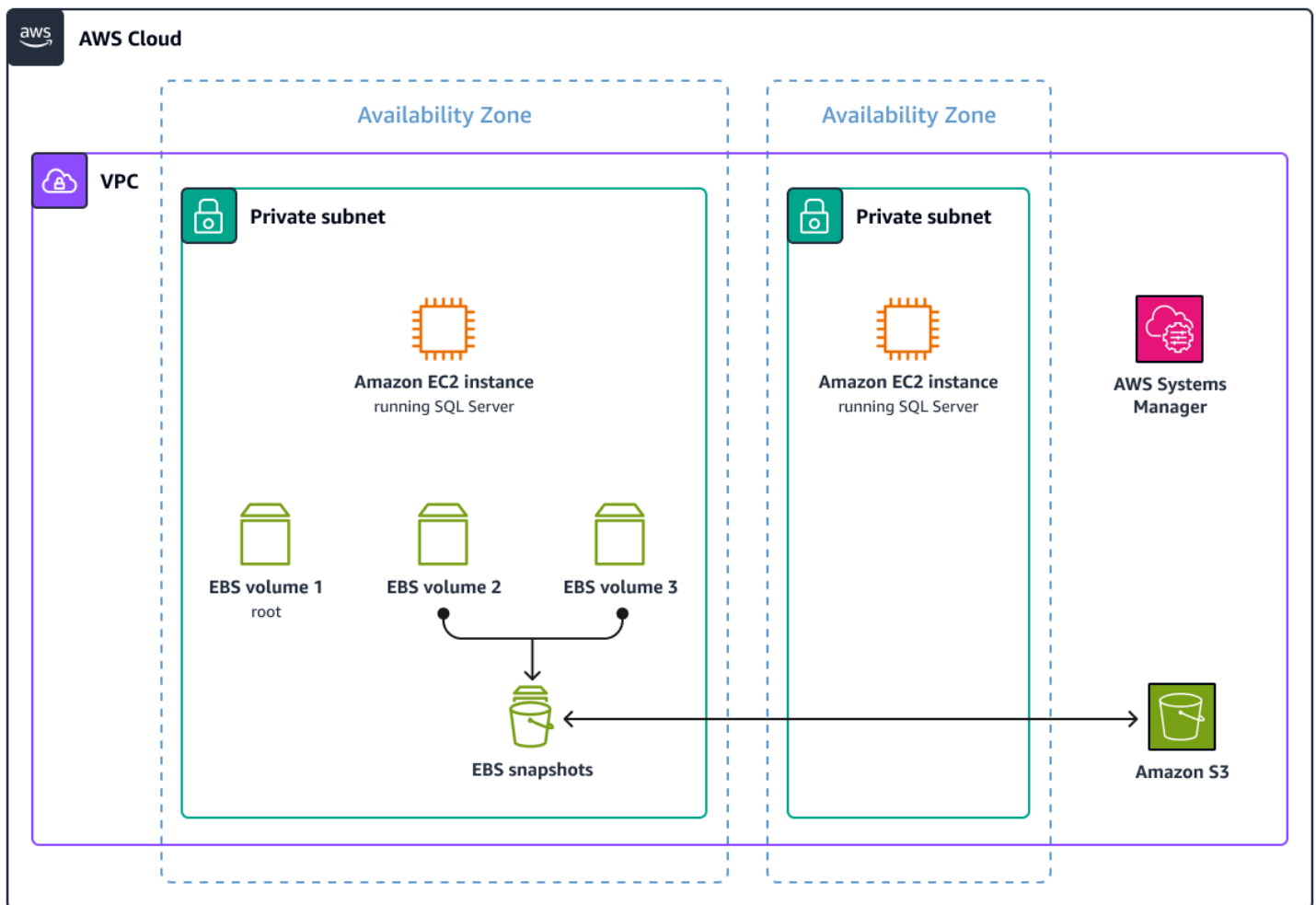
Copia de seguridad de servidores con instantáneas compatibles con VSS

Una arquitectura de instantáneas compatible con VSS utiliza el [comando AWS Systems Manager Run](#) para instalar el agente VSS en las instancias de SQL Server. También puede usar el comando Run para invocar todo el flujo de trabajo consistente en vaciar los búferes del sistema operativo y las aplicaciones en el disco, pausar I/O las operaciones, tomar una point-in-time instantánea de los volúmenes de EBS y, a continuación, reanudar la E/S.

Run Command crea instantáneas automatizadas de todos los volúmenes de EBS conectados a una instancia de destino. También tiene la opción de excluir el volumen raíz, ya que los archivos de

bases de datos de usuarios suelen almacenarse en otros volúmenes. Si divide varios volúmenes de EBS para crear un único sistema de archivos para los archivos de SQL Server, Amazon EBS también admite instantáneas de varios volúmenes coherentes ante bloqueos mediante un único comando de API. Para obtener más información sobre las [instantáneas de EBS compatibles con VSS](#), consulte [Instantáneas de Amazon EBS basadas en VSS de Windows consistentes con la aplicación](#) en la documentación de Amazon EC2.

El siguiente diagrama muestra una arquitectura para realizar copias de seguridad de servidores mediante instantáneas compatibles con VSS.



Tenga en cuenta las siguientes ventajas del uso de instantáneas compatibles con VSS:

- La primera instantánea de una instancia de base de datos contiene los datos de la instantánea de base de datos completa. Las instantáneas posteriores de la misma instancia de base de datos son [graduales](#), lo que significa que solo se guardan los datos que han cambiado después de guardar la última instantánea.

- point-in-time Las instantáneas de EBS proporcionan recuperación.
- Puede [restaurar a una nueva instancia de EC2 de SQL Server a partir de una instantánea](#).
- Si una instancia se cifra con Amazon EBS o si se cifra una base de datos en la instancia mediante un cifrado de datos transparente (TDE), esa instancia o base de datos se restaura de forma automática con el mismo cifrado.
- Puede copiar [copias de seguridad automatizadas entre regiones](#).
- Al restaurar un volumen de EBS a partir de una instantánea, las aplicaciones pueden acceder a él. Esto significa que puede poner SQL Server en línea de forma inmediata después de restaurar uno o más de sus volúmenes de EBS subyacentes a partir de instantáneas.
- De forma predeterminada, los volúmenes restaurados obtienen bloques subyacentes de Amazon S3 la primera vez que una aplicación intenta leerlos. Esto significa que puede haber un retraso en el rendimiento después de restaurar un volumen de EBS a partir de una instantánea. Con el tiempo, el volumen alcanza el rendimiento nominal. Sin embargo, puede evitar ese retraso mediante la [restauración rápida de instantáneas \(FSR\)](#).
- Puede utilizar la [administración del ciclo de vida de las instantáneas de EBS](#).

Tenga en cuenta las siguientes limitaciones del uso de instantáneas compatibles con VSS:

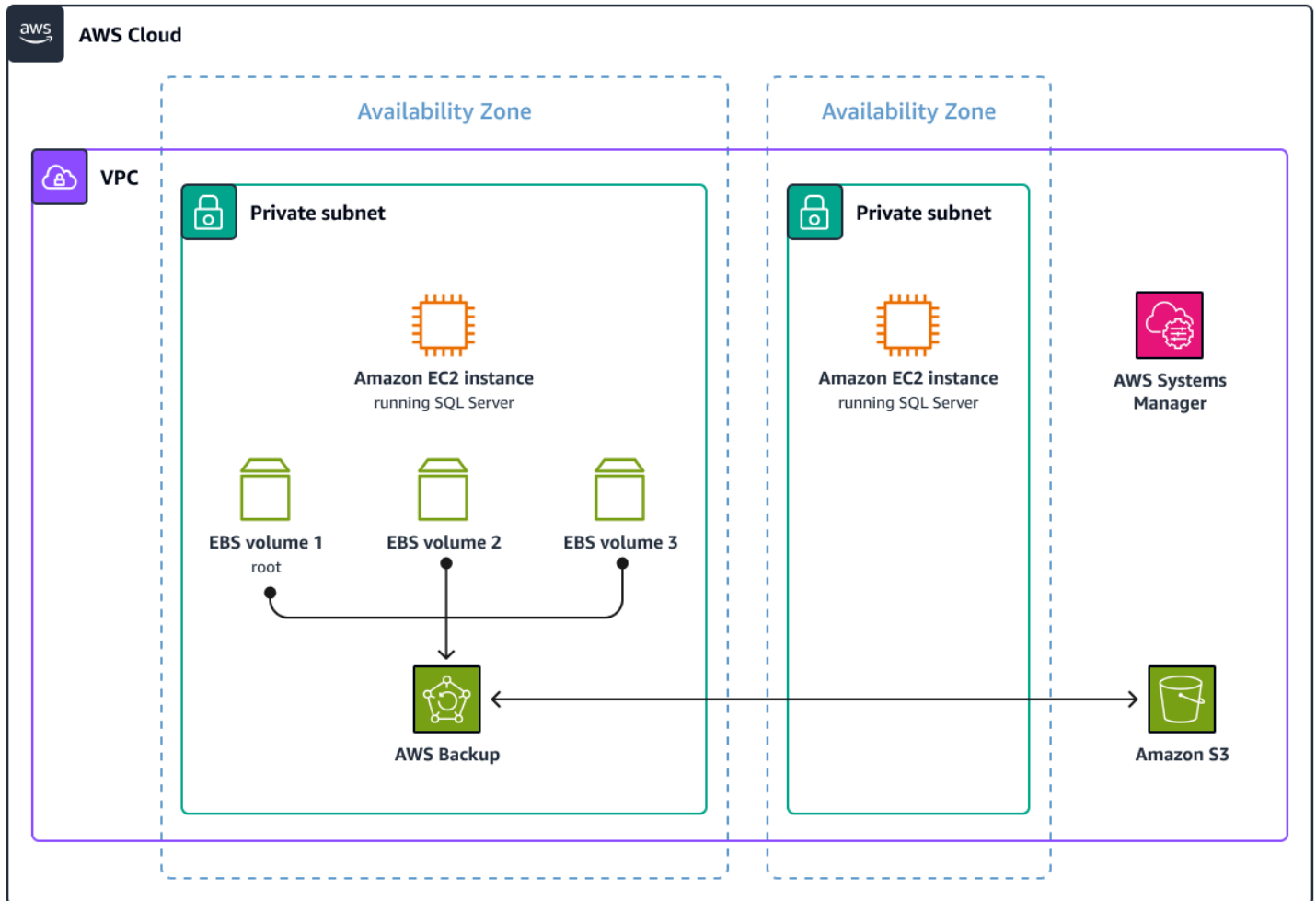
- No puede realizar una point-in-time recuperación entre regiones con una instantánea cifrada para una instancia de SQL Server.
- No puede crear una instantánea cifrada de una instancia sin cifrar.
- No puede restaurar una base de datos individual, ya que la instantánea se crea a partir del volumen de EBS.
- No puede restaurar la instancia en sí misma.
- Una instantánea de una instancia de bases de datos debe cifrarse con la misma clave de AWS Key Management Service (AWS KMS) que la instancia de bases de datos.
- I/O El almacenamiento se suspende durante una fracción de segundo (aproximadamente 10 milisegundos) durante el proceso de copia de seguridad de la instantánea.

Realice copias de seguridad de SQL Server mediante AWS Backup

Se puede utilizar [AWS Backup](#) para centralizar y automatizar la protección de datos en todas las partes de AWS. AWS Backup ofrece una solución rentable, totalmente gestionada y basada en políticas que simplifica la protección de datos a gran escala. AWS Backup también

le ayuda a cumplir con sus obligaciones de conformidad normativa y a cumplir sus objetivos de continuidad empresarial. Además AWS Organizations, le AWS Backup permite implementar de forma centralizada políticas de protección de datos (respaldo) para configurar, administrar y gobernar su actividad de respaldo en todos los recursos Cuentas de AWS y la organización.

En el siguiente diagrama, se muestra la arquitectura de una solución de copia de seguridad y restauración para SQL Server en EC2 con AWS Backup.



Tenga en cuenta las siguientes ventajas de realizar copias de seguridad de SQL Server mediante AWS Backup:

- Puede automatizar la programación de copias de seguridad, la administración de la retención y la administración del ciclo de vida.
- Puede centralizar su estrategia de respaldo en toda su organización, abarcando varias cuentas y Regiones de AWS

- Puede centralizar la supervisión de la actividad de copia de seguridad y las alertas en todos los Servicios de AWS.
- Puede implementar copias de seguridad entre regiones para planificar la recuperación de desastres.
- La solución es compatible con las copias de seguridad entre cuentas.
- Puede realizar copias de seguridad seguras con un cifrado de copias de seguridad secundario.
- Todas las copias de seguridad admiten el cifrado mediante claves de AWS KMS cifrado.
- La solución funciona con TDE.
- Puede restaurar a un punto de recuperación específico desde la consola de AWS Backup .
- Puede hacer una copia de seguridad de una instancia completa de SQL Server, que incluye todas las bases de datos de SQL Server.

Copia de seguridad a nivel de la base de datos

Estos enfoques utilizan la funcionalidad de copia de seguridad nativa de Microsoft SQL Server. Puede realizar copias de seguridad de bases de datos individuales en la instancia de SQL Server y restaurar una base de datos individual.

Cada una de estas opciones de copia de seguridad y restauración nativas de SQL Server también admite lo siguiente:

- Compresión y copia de seguridad de varios archivos
- Copias de seguridad completas, diferenciales y de registros T
- Bases de datos cifradas con TDE

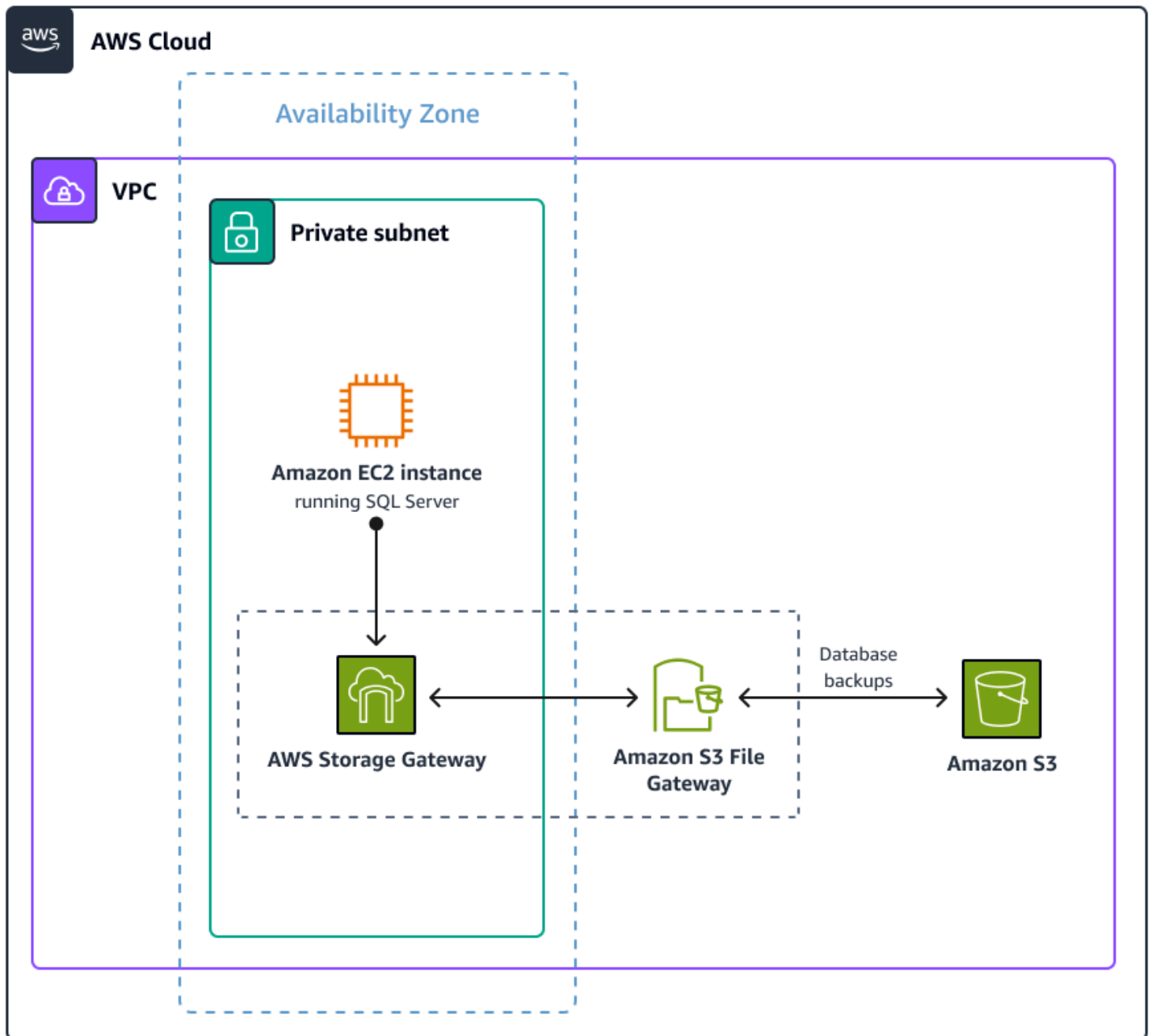
Copia de seguridad y restauración nativas de SQL Server en Amazon S3

SQL Server en Amazon EC2 admite la copia de seguridad y la restauración nativas para las bases de datos de SQL Server. Puede realizar una copia de seguridad de la base de datos de SQL Server y, a continuación, restaurar la copia de seguridad de archivos en una base de datos existente o en una nueva instancia EC2 de SQL Server, Amazon RDS para SQL Server o un servidor en las instalaciones.

Storage Gateway es un servicio de almacenamiento de nube híbrida que brinda acceso a las aplicaciones en las instalaciones a un almacenamiento en la nube prácticamente ilimitado.

Puede utilizar Storage Gateway para realizar copias de seguridad de las bases de datos de Microsoft SQL Server directamente en Amazon S3, lo que reduce el espacio de almacenamiento en las instalaciones y utiliza el almacenamiento duradero, escalable y rentable de Amazon S3.

En el siguiente diagrama, se muestra la arquitectura de una solución de copia de seguridad y restauración nativas con Storage Gateway y Amazon S3.



Tenga en cuenta los siguientes beneficios de usar las capacidades de copia de seguridad nativas de SQL Server con Storage Gateway:

- Puede asignar una puerta de enlace de almacenamiento como un recurso compartido de archivos del bloque de mensajes del servidor (SMB) en la instancia EC2 y enviar la copia de seguridad a Amazon S3.
- La copia de seguridad se envía directamente del bucket de S3 o a través de la memoria caché de archivos de Storage Gateway.
- Se admiten copias de seguridad de varios archivos.

Tenga en cuenta las siguientes limitaciones de las capacidades de copia de seguridad nativas con Storage Gateway:

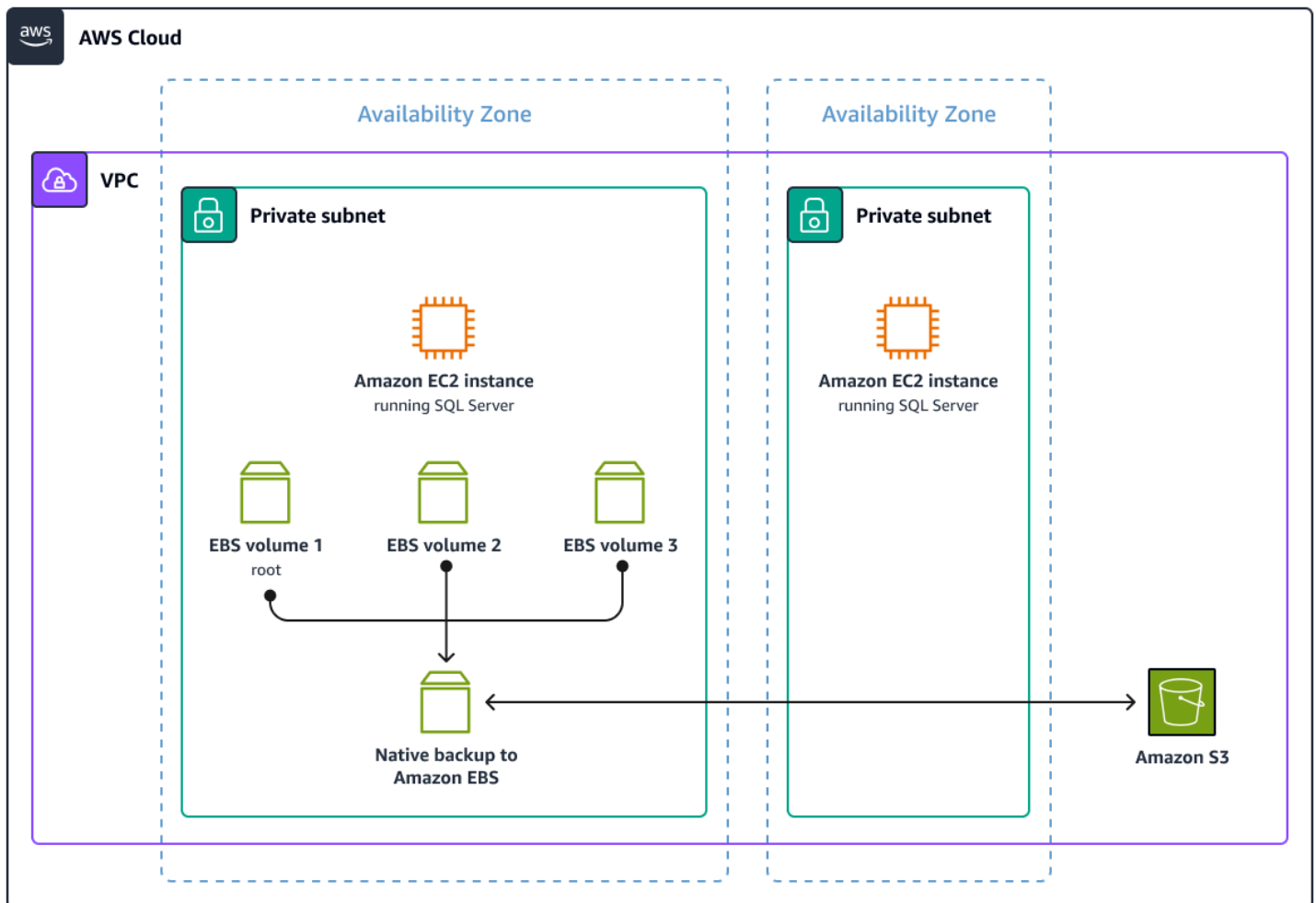
- Debe configurar la copia de seguridad y la restauración para cada base de datos individual.
- Debe administrar la [Política de ciclo de vida de Amazon S3](#) para los archivos de copia de seguridad.

Para obtener más información sobre cómo configurar Storage Gateway, consulte la publicación [Store SQL Server backups in Amazon S3 using AWS Storage Gateway](#) en el blog AWS .

Copias de seguridad nativas de SQL Server en volúmenes de EBS

Puede realizar una copia de seguridad nativa de la base de datos de SQL Server y almacenar el archivo en un volumen de Amazon EBS. Amazon EBS es un servicio de almacenamiento en bloques de alto rendimiento. Los volúmenes de EBS son elásticos, lo que permite el cifrado. Se pueden conectar a una instancia de EC2, así como también desconectarse. Puede hacer una copia de seguridad de SQL Server en una instancia de EC2 en el mismo tipo de volumen de EBS o en un tipo de volumen de EBS diferente. Una ventaja de realizar copias de seguridad en un volumen de EBS diferente es el ahorro de costos.

En el siguiente diagrama, se muestra la arquitectura de una copia de seguridad nativa en un volumen de EBS.



Tenga en cuenta los siguientes beneficios de usar las copias de seguridad nativas de SQL Server en volúmenes de EBS:

- Puede realizar copias de seguridad de bases de datos individuales en una instancia de EC2 de SQL Server y restaurar una base de datos individual, en lugar de tener que restaurar la instancia completa.
- Se admiten copias de seguridad de varios archivos.
- Puede programar trabajos de copia de seguridad mediante el agente de SQL Server y el motor de trabajos de SQL Server.
- Puede obtener beneficios de rendimiento a través de las opciones de equipos. Por ejemplo, puede utilizar los volúmenes de almacenamiento st1 para lograr un mayor rendimiento.

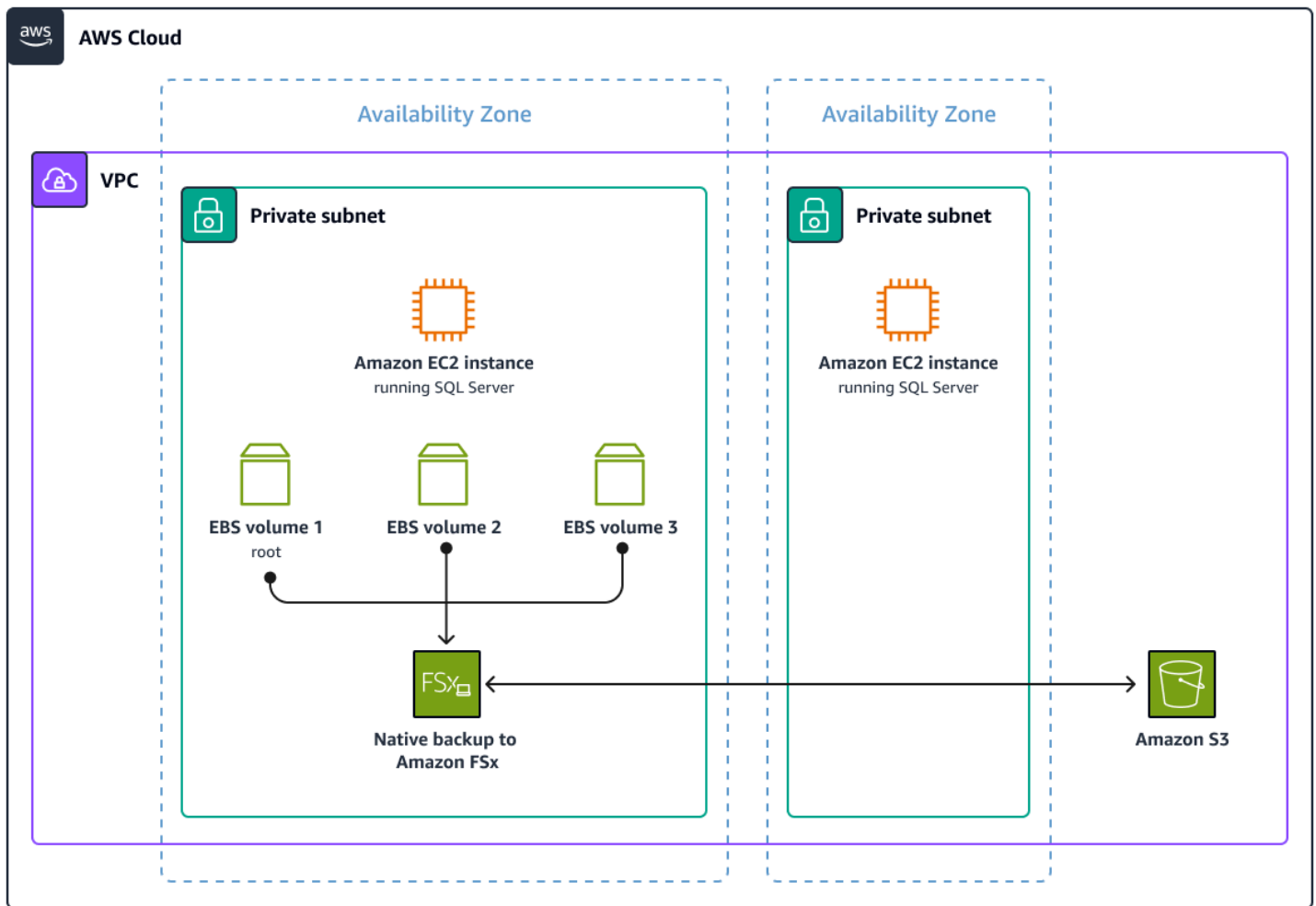
Tenga en cuenta las siguientes limitaciones de usar copias de seguridad nativas en volúmenes de EBS:

- Debe mover de forma manual las copias de seguridad a Amazon S3 desde el volumen de EBS.
- En el caso de las copias de seguridad de gran tamaño, debe administrar el espacio en disco de Amazon EC2.
- En la instancia de EC2, el rendimiento de Amazon EBS puede suponer un cuello de botella.
- Se requiere almacenamiento adicional para almacenar copias de seguridad en Amazon EBS.

Backup nativo de SQL Server en Amazon FSx para Windows File Server

[Amazon FSx for Windows File Server](#) es un sistema de archivos de Windows nativo totalmente gestionado que ofrece hasta 64 TB de almacenamiento y está diseñado para ofrecer un rendimiento rápido, predecible y uniforme. AWS introdujo el [soporte nativo para las implementaciones de sistemas de archivos Multi-AZ](#) en FSx Windows File Server. El soporte nativo facilita la implementación del almacenamiento de archivos de Windows en AWS con alta disponibilidad y redundancia en varias zonas de disponibilidad. AWS también incorporó compatibilidad con los [recursos compartidos de archivos con la disponibilidad continua \(CA\) SMB](#). Puede utilizar el FSx servidor de archivos de Windows como almacenamiento de respaldo para una base de datos de SQL Server.

El siguiente diagrama muestra la arquitectura de una copia de seguridad nativa de SQL Server FSx para Windows File Server.



Tenga en cuenta las siguientes ventajas de utilizar copias de seguridad nativas de SQL Server FSx para Windows File Server:

- Puede hacer una copia de seguridad de la base de datos de SQL Server en un recurso compartido de FSx archivos de Amazon.
- Puede realizar copias de seguridad de bases de datos individuales en una instancia de SQL Server y restaurar una base de datos individual, en lugar de tener que restaurar la instancia completa.
- Se admiten copias de seguridad de varias partes.
- Puede programar trabajos de copia de seguridad mediante el agente de SQL Server y el motor de trabajos.
- Las instancias tienen un mayor ancho de banda de la red, en comparación con Amazon EBS.

Tenga en cuenta las siguientes limitaciones a la hora de utilizar copias de seguridad nativas de SQL Server FSx para Windows File Server:

- Debe mover manualmente las copias de seguridad a Amazon S3 desde Amazon FSx mediante AWS Backup o AWS DataSync.
- Las copias de seguridad de gran tamaño pueden requerir una sobrecarga adicional para la administración del espacio en disco en Amazon FSx.
- En la instancia de EC2, el rendimiento de la red puede suponer un cuello de botella.
- Se necesita espacio de almacenamiento adicional para almacenar las copias de seguridad en el FSx servidor de archivos de Windows.

Copia de seguridad de SQL Server en Amazon FSx para NetApp ONTAP

Las instantáneas FSx para ONTAP siempre son coherentes con los fallos, pero es necesario poner en modo inactivo (o pausar) la base I/O de datos para crear una instantánea coherente con la aplicación. Puede utilizar ONTAP NetApp SnapCenter (una herramienta de organización con complementos para aplicaciones específicas, incluido SQL Server) FSx para crear instantáneas coherentes con las aplicaciones y proteger, replicar y clonar sus bases de datos sin coste adicional.

NetApp SnapCenter

NetApp SnapCenter es una plataforma unificada para una protección de datos coherente con las aplicaciones. SnapCenter se refiere a las instantáneas como copias de seguridad. Esta guía adopta la misma convención de nomenclatura. SnapCenter proporciona un panel único para administrar copias de seguridad, restauraciones y clones coherentes con las aplicaciones. Agregue un SnapCenter complemento para su aplicación de base de datos específica para crear copias de seguridad coherentes con las aplicaciones. El SnapCenter complemento para SQL Server proporciona las siguientes funciones que simplifican el flujo de trabajo de protección de datos.

- Opciones de copia de seguridad y restauración con granularidad para copias de seguridad completas y de registros
- Restauración in situ y restauración en una ubicación alternativa

Para obtener más información al respecto SnapCenter, consulte la publicación [Proteja sus cargas de trabajo de SQL Server NetApp SnapCenter con Amazon FSx for NetApp ONTAP](#) en el blog sobre AWS almacenamiento.

Optimización de costos de las copias de seguridad

Las siguientes opciones pueden ser útiles para reducir el costo de almacenar las copias de seguridad de SQL Server en AWS.

- Active la [compresión de SQL Server](#) al crear el archivo de copia de seguridad y envíe el archivo más pequeño posible al almacenamiento. Por ejemplo, una relación de compresión de 3:1 indica que está ahorrando alrededor del 66 % de espacio en disco. Para realizar consultas en estas columnas, puede usar la siguiente instrucción de Transact-SQL: `SELECT backup_size/compressed_backup_size FROM msdb..backupset;`
- En el caso de las copias de seguridad que enviadas a buckets de S3, active la clase de almacenamiento [Amazon S3 Intelligent-Tiering](#) para reducir los costos de almacenamiento en un 30 %.
- Para las copias de seguridad destinadas a FSx Windows File Server o a ONTAP, utilice una única zona de disponibilidad FSx para ahorrar un 50 por ciento en los costes (en comparación con el uso de varias zonas de disponibilidad). Para obtener información sobre precios, consulte los [precios de Amazon FSx para Windows File Server](#) y [Amazon FSx para NetApp ONTAP](#).
- La opción más eficaz para SQL Server 2022 es hacer una copia de seguridad directa en Amazon S3. Puede ahorrar costos adicionales si evita usar Storage Gateway.

Resultados de las pruebas de rendimiento de las copias de seguridad

En esta sección, se comparan las siguientes opciones en términos de costos y rendimiento de una base de datos de ejemplo de 1 TB, según los resultados de las pruebas comparativas de rendimiento de las soluciones de copia de seguridad explicadas en esta guía.

- Especificaciones de la instancia de EC2: r5d.8xlarge con Windows Server 2019 y la edición Developer de SQL Server 2019
- Especificaciones de la base de datos: 1 TB de tamaño con el TDE desactivado

Las pruebas se realizaron con una instancia r5d.8xlarge y una base de datos de SQL Server de 1 TB como origen. El sistema de origen se configuró de acuerdo con las prácticas recomendadas; la base de datos de origen contenía 4 archivos de datos (250 GB cada uno) y un archivo de registro (50 GB) repartidos en volúmenes gp3 independientes. El comando BACKUP nativo de SQL Server incluye la escritura en 10 archivos de copia de seguridad mediante la compresión, con el fin de optimizar el rendimiento de las copias de seguridad y reducir la cantidad de datos enviados a través de la red y

escritos en el destino. En todos los casos de prueba, el rendimiento del almacenamiento fue el factor limitante.

Existe una variedad casi infinita de configuraciones posibles para este tipo de pruebas. Esta prueba se centró en optimizar el rendimiento, el costo, la escalabilidad y los casos de uso reales. En la siguiente tabla se muestran las métricas de rendimiento que se capturaron para las opciones de copia de seguridad en el destino.

Opciones de copia de seguridad	Nivel	Duración de la puesta en marcha (aproximación)	Velocidad de la copia de seguridad	Costo en USD al mes*
Copia de seguridad nativa en el volumen st1 basado en HDD de EBS, 2 TB	Base de datos	00:30:46 min	554,7 Mbps	92,16 DÓLARES
Copia de seguridad nativa en el volumen gp3 basado en SSD de EBS, 2 TB	Base de datos	00:22:00 min	512 Mbps	193,84 DÓLARES
Respaldo nativo FSx para disco duro Windows File Server, 2 TB @512 Mbps de rendimiento	Base de datos	00:20:58 min	814,0 Mbps	1.146 DÓLARES
Respaldo nativo FSx para SSD Windows File Server, 2 TB	Base de datos	00:20:00 min	814,0 Mbps	1326 DÓLARES

Opciones de copia de seguridad	Nivel	Duración de la puesta en marcha (aproximación)	Velocidad de la copia de seguridad	Costo en USD al mes*
@512 Mbps de rendimiento				
Copia de seguridad nativa en una instancia m6i.4xlarge de S3 File Gateway (16 vCPU, 64 GB) con un volumen gp3 de 2 TB	Base de datos	00:23:20 min	731,5 Mbps	470,42 DÓLARES
Instantánea de EBS con VSS	Volumen de EBS	00:00:02 s 00:00:53 s	N/D (instantánea)	51 DÓLARES
AWS Backup (Respaldo de AMI)	AMI	00:00:04 s 00:08:00 min	N/D (instantánea)	75 DÓLARES
Copia de seguridad nativa de SQL Server directamente en Amazon S3 (SQL Server 2022)	Base de datos	00:12:00 min	731,5 Mbps	Primeros 50 TB al mes, 0,023 USD por GB (23,55 USD al mes)
Respaldo nativo FSx para ONTAP (usando SnapCenter)	Base de datos	–	–	440,20\$

La tabla anterior se basa en los siguientes supuestos:

- Los costos de transferencia de datos y Amazon S3 no están incluidos.
- El precio del almacenamiento está incluido en el precio de la instancia.
- Los costos se calculan en función de la región us-east-1.
- El rendimiento y las IOPS aumentan un 10 %, mientras que las copias de seguridad múltiples tienen una tasa general de cambio del 10 % a lo largo del mes.

Los resultados de las pruebas muestran que la opción más rápida es realizar una copia de seguridad nativa de la base de datos de SQL Server FSx para Windows File Server. Una copia de seguridad en Storage Gateway y en volúmenes de EBS conectados localmente es la opción más rentable, pero tiene un rendimiento más lento. Para los respaldos a nivel de servidor (AMI), recomendamos utilizarlos AWS Backup para obtener un rendimiento, un costo y una capacidad de administración óptimos.

Recomendaciones de optimización de costos

Comprender las posibles soluciones para hacer copias de seguridad de SQL Server en Amazon EC2 es fundamental para proteger los datos y asegurarse de que cumple con las necesidades de copias de seguridad, así como para implementar un plan que le permita recuperarse de eventos críticos. Las diferentes formas de realizar copias de seguridad y restaurar las instancias y bases de datos de SQL Server que se exploran en esta sección pueden ser de ayuda para diseñar una estrategia de copia de seguridad y restauración que proteja los datos y cumpla con los requisitos de su organización.

Esta sección abarca las siguientes opciones de copia de seguridad:

- Compresión
- Amazon S3 Intelligent-Tiering
- Una sola zona de disponibilidad
- Copia de seguridad en URL

La orientación proporcionada para cada una de estas opciones es de carácter general. Si desea implementar alguna de estas recomendaciones en su organización, le recomendamos que se ponga en contacto con su equipo de cuentas. A continuación, el equipo puede contactar con un especialista

en SA de Microsoft para que dirija la conversación. También puede contactarnos enviando un correo electrónico a optimize-microsoft@amazon.com.

En resumen, recomendamos lo siguiente:

- Si utiliza SQL Server 2022, la opción más rentable es realizar copias de seguridad en Amazon S3.
- Si utiliza SQL Server 2019 y ediciones anteriores de SQL Server, considere la posibilidad de realizar copias de seguridad en Storage Gateway con el respaldo de Amazon S3 como la opción más rentable.

Compresión

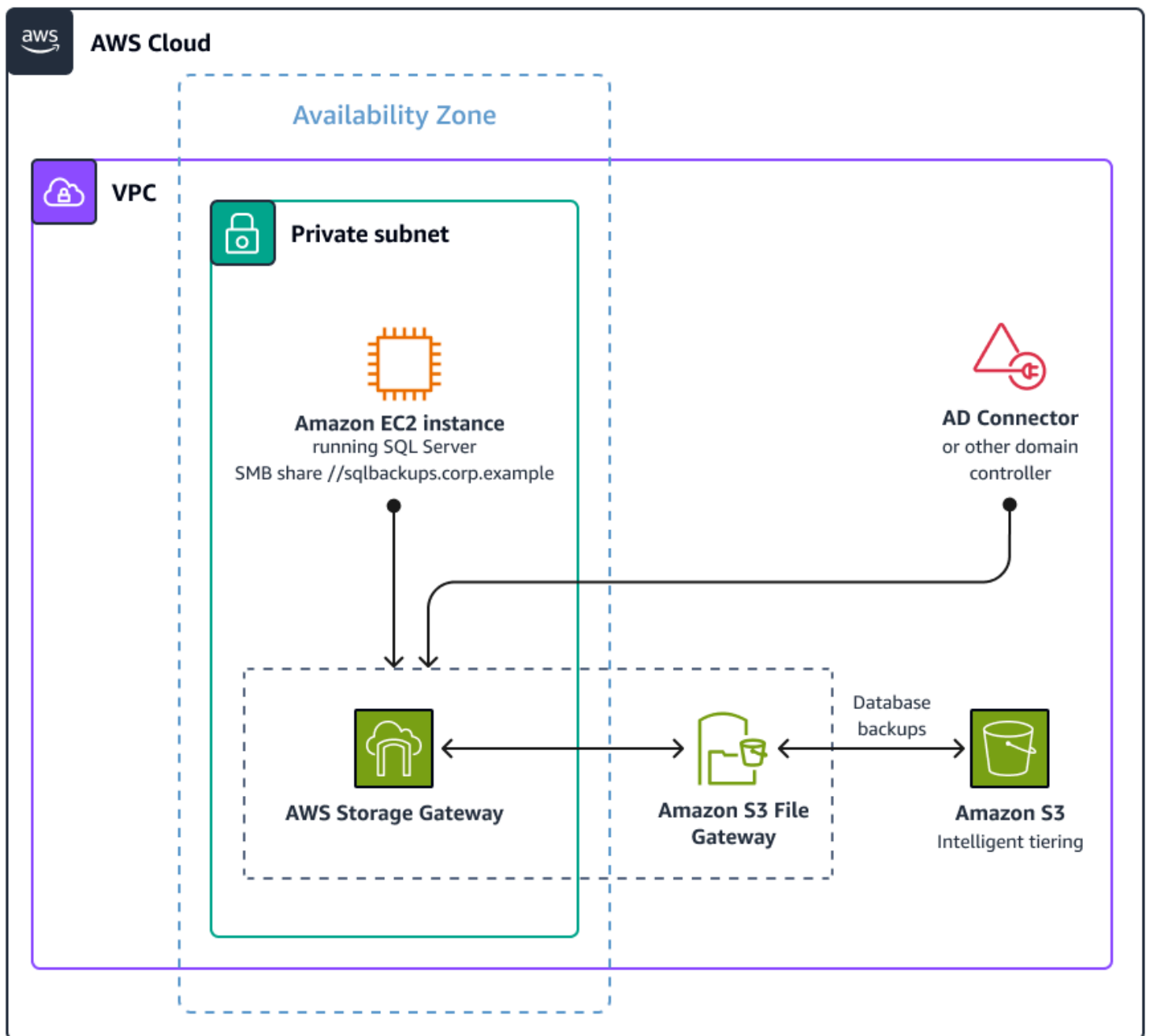
El objetivo de la compresión es que cada copia de seguridad consuma menos espacio de almacenamiento, lo que resulta beneficioso para varias opciones de almacenamiento. Debe activar la compresión de una copia de seguridad de SQL Server para las [instancias de SQL Server](#). El siguiente ejemplo muestra cómo agregar la palabra clave “compression” con una base de datos de copia de seguridad:

```
BACKUP DATABASE <database_name> TO DISK WITH COMPRESSION (ALGORITHM =
QAT_DEFLATE)
```

Amazon S3 Intelligent-Tiering

En el caso de las copias de seguridad enviadas a buckets de Amazon S3, puede activar [Amazon S3 Intelligent-Tiering](#) como [clase de almacenamiento](#) de Amazon S3 File Gateway. Esto reduce los costos de almacenamiento hasta en un 30 %. A continuación, debe montar S3 File Gateway en sus servidores SQL mediante un recurso compartido de archivos SMB que se puede integrar con su [dominio de Active Directory](#). Esto le otorga el control del acceso al recurso compartido, la posibilidad de aprovechar las cuentas de servicio existentes y el acceso a Amazon S3 mediante un protocolo de archivos común centrado en Microsoft. En el caso de las cuentas que quizás no tengan conectividad directa con un controlador de dominio, puede utilizar [Conector Active Directory](#) para facilitar la comunicación con Active Directory en las instalaciones o en la nube. Para configurar los ajustes de Active Directory en la puerta de enlace, debe especificar el conector de Active Directory IPs para que el controlador de dominio envíe las solicitudes por proxy a Active Directory.

En el siguiente diagrama, se muestra una arquitectura para una solución basada en S3 Intelligent-Tiering.



De forma predeterminada, los archivos de copia de seguridad escritos en el bucket de S3 utilizan el nivel Standard. Para convertir los archivos de copia de seguridad del nivel Standard al nivel S3 Intelligent-Tiering, debe [crear una regla de ciclo de vida](#). También puede utilizar la [Consola de administración de AWS](#) para activar S3 Intelligent-Tiering. Para obtener más información, consulte [Introducción al uso de Amazon S3 Intelligent-Tiering](#) en la documentación de AWS .

Una sola zona de disponibilidad

Para crear un sistema de archivos de zona de disponibilidad única, elija la opción Single-AZ al [crear un sistema de archivos FSx para Windows File Server](#). Amazon FSx también realiza copias de seguridad muy duraderas (almacenadas en Amazon S3) de su sistema de archivos a diario mediante el servicio Windows Volume Shadow Copy y le permite realizar copias de seguridad adicionales en cualquier momento. Tenga en cuenta algunos de los problemas relacionados con el uso de una zona de disponibilidad única. Por ejemplo, el recurso compartido de archivos SMB deja de estar accesible si una zona de disponibilidad afectada en la que se aprovisiona el sistema de archivos deja de funcionar durante horas seguidas. Si necesita acceder a los datos, debe restaurarlos a partir de las copias de seguridad de una zona de disponibilidad disponible dentro de la región de origen. Para obtener más información, consulte la sección [Uso de una única zona de disponibilidad](#) de esta guía.

Copia de seguridad en URL

En el caso de SQL Server 2022, la característica de [copia de seguridad en URL](#) permite realizar copias de seguridad directas en Amazon S3. Este es el enfoque de respaldo ideal para ejecutar SQL Server 2022, AWS ya que se obtiene el conjunto completo de funciones de Amazon S3 en la capa de almacenamiento y se elimina el costo del AWS Storage Gateway dispositivo necesario en las versiones anteriores para facilitar esta funcionalidad. Al implementar esta característica, se deben tener en cuenta dos costos principales: los costos de transferencia de datos y la clase de almacenamiento de S3 elegida. Si desea las capacidades de recuperación ante desastres nativas de Amazon S3, debe tener en cuenta que la [replicación entre regiones](#) genera [costos de salida de datos](#) entre regiones. Para obtener más información sobre cómo configurar esta opción, consulte la publicación [Backup SQL Server databases to Amazon S3](#) del blog de cargas de trabajo de Microsoft en AWS .

Recursos adicionales

- [Opciones de backup y restauración para SQL Server en Amazon EC2 \(guíaAWS prescriptiva\)](#)
- [Point-in-time recuperación y copia de seguridad continua para Amazon RDS con AWS Backup \(AWS Storage Blog\)](#)
- [Proteja sus cargas de trabajo de SQL Server NetApp SnapCenter con Amazon FSx for NetApp ONTAP \(blog sobreAWS almacenamiento\)](#)
- [Introducción al uso de Amazon S3 Intelligent-Tiering \(Centro de recursos de AWS introducción\)](#)
- [Estrategias de backup y restauración de Amazon RDS for SQL ServerAWS \(blog sobre bases de datos\)](#)

- [Migración de una base de datos de Microsoft SQL Server local a Amazon EC2 AWS](#) (orientación prescriptiva)
- [Prácticas recomendadas para la implementación de Microsoft SQL Server en Amazon EC2](#) (AWS documento técnico)

Modernización de las bases de datos de SQL Server

Descripción general de

Si está empezando a modernizar las bases de datos heredadas para mejorar la escalabilidad, el rendimiento y la optimización de costos, es posible que tenga que hacer frente a desafíos relacionados con bases de datos comerciales como SQL Server. Las bases de datos comerciales son caras, atan a los clientes y ofrecen condiciones de licencia muy estrictas. En esta sección se proporciona una descripción general de las opciones para migrar y modernizar bases de datos de SQL Server a bases de datos de código abierto, así como información sobre cómo elegir la mejor opción para su carga de trabajo.

Puede refactorizar sus bases de datos de SQL Server para convertirlas en bases de datos de código abierto, como Amazon Aurora PostgreSQL, para ahorrar en los costos asociados al uso de licencias de Windows y SQL Server. Las bases de datos modernas nativas en la nube, como Aurora, combinan la flexibilidad y el bajo costo de las bases de datos de código abierto con las sólidas características empresariales de las bases de datos comerciales. Si tiene cargas de trabajo variables o cargas de trabajo de tenencia múltiple, también puede migrar a [Aurora V2 sin servidor](#). Esto puede reducir los costos hasta en un 90 %, según las características de la carga de trabajo. Además, AWS ofrece capacidades como [Babelfish para Aurora PostgreSQL](#), herramientas como [AWS Schema Conversion Tool \(AWS SCT\)](#) y servicios como [AWS Database Migration Service \(AWS DMS\)](#) para simplificar la migración y la modernización de las bases de datos de SQL Server en AWS.

Ofertas de bases de datos

La migración de SQL Server en Windows a bases de datos de código abierto como Amazon Aurora, Amazon RDS para MySQL o Amazon RDS para PostgreSQL puede ofrecer importantes ahorros de costos sin comprometer el rendimiento ni las características. Considere lo siguiente:

- El cambio de la edición Enterprise de SQL Server en Amazon EC2 a Amazon RDS para PostgreSQL o Amazon RDS para MySQL puede suponer un ahorro de costos de hasta un 80 %.

- El cambio de la edición Enterprise de SQL Server en Amazon EC2 a la edición compatible con PostgreSQL de Amazon Aurora o a la edición compatible con MySQL de Amazon Aurora puede generar ahorros de hasta un 70 % de los costos.

En el caso las cargas de trabajo de bases de datos tradicionales, Amazon RDS para PostgreSQL y Amazon RDS para MySQL abordan los requisitos y proporcionan una solución rentable para las bases de datos relacionales. Aurora agrega numerosas características de disponibilidad y rendimiento que antes estaban limitadas a proveedores comerciales caros. Las características de resiliencia de Aurora tienen un costo adicional. Sin embargo, en comparación con características similares de otros proveedores comerciales, los costos de resiliencia de Aurora siguen siendo más baratos que los que cobra el software comercial por el mismo tipo de características. La arquitectura Aurora de está optimizada para ofrecer mejoras significativas en el rendimiento en comparación con las implementaciones estándar de MySQL y PostgreSQL.

Como Aurora es compatible con bases de datos PostgreSQL y MySQL de código abierto, la portabilidad supone una ventaja adicional. A la hora de decidirse entre Amazon RDS para PostgreSQL, Amazon RDS para MySQL o Aurora, hay que tener en cuenta los requisitos empresariales y las características que necesitamos.

Comparación entre Amazon RDS y Aurora

En la tabla siguiente se resumen las diferencias entre Amazon RDS y Amazon Aurora.

Categoría	Amazon RDS para PostgreSQL y Amazon RDS para MySQL	Aurora PostgreSQL o Aurora MySQL
Desempeño	Buen rendimiento	Rendimiento mejor o 3 veces mayor
Conmutación por error	Normalmente entre 60 y 120 segundos*	Normalmente 30 segundos
Escalabilidad	Hasta 5 réplicas de lectura Retraso en segundos	Hasta 15 réplicas de lectura Retraso en milisegundos
Almacenamiento	Hasta 64 TB	Hasta 128 TB

Categoría	Amazon RDS para PostgreSQL y Amazon RDS para MySQL	Aurora PostgreSQL o Aurora MySQL
Alta disponibilidad (almacenamiento)	Multi-AZ con una o dos zonas en espera, cada una con copia de la base de datos	De forma predeterminada, 6 copias de datos en 3 zonas de disponibilidad
Copia de seguridad	Copias de seguridad diarias de instantáneas y registros	Copia de seguridad asíncrona y continua en Amazon S3
Innovaciones con Aurora	N/D	100 GB Clonación rápida de base de datos
	Réplicas de lectura de escalado automático	
	Administración de planes de consultas	
	Aurora sin servidor	
	Réplicas de lectura entre regiones con base de datos global	
	Administración de la memoria caché de clústeres**	
	Consulta paralela	
	Secuencias de actividades de la base de datos	

* Las transacciones grandes pueden aumentar los tiempos de conmutación por error.

** Disponible en Aurora PostgreSQL.

La siguiente tabla muestra el costo mensual estimado de los diferentes servicios de bases de datos explicados en esta sección.

Servicios de bases de datos	Costo en USD al mes*	Calculadora de precios de AWS (requiere Cuenta de AWS)
Edición Enterprise de Amazon RDS para SQL Server	3.750 dólares	Estimación
Edición Standard de Amazon RDS para SQL Server	2.318 DÓLARES	Estimación
Edición Enterprise de SQL Server en Amazon EC2	2.835 DÓLARES	Estimación
Edición Standard de SQL Server en Amazon EC2	1.345 DÓLARES	Estimación
Amazon RDS para PostgreSQL	742 DÓLARES	Estimación
Amazon RDS para MySQL	712 DÓLARES	Estimación
Aurora PostgreSQL	1.032 DÓLARES	Estimación
Aurora MySQL	1.031 DÓLARES	Estimación

El precio del almacenamiento está incluido en el precio de la instancia. Los costos se calculan en función de la región us-east-1. El rendimiento y las IOPS son suposiciones. Los cálculos son para las instancias r6i.2xlarge y r6g.2xlarge.

Recomendaciones de optimización de costos

Las migraciones de bases de datos heterogéneas suelen requerir convertir el esquema de la base de datos del motor de base de datos de origen al de destino y migrar los datos de la base de datos de origen a la base de datos de destino. El primer paso hacia la migración es evaluar y convertir el esquema del servidor SQL y los objetos de código para el motor de base de datos de destino.

Puede utilizar [AWS Schema Conversion Tool \(AWS SCT\)](#) para evaluar la compatibilidad de la base de datos con varias opciones de bases de datos de código abierto de destino, como Amazon RDS para MySQL o Amazon RDS para PostgreSQL, Aurora MySQL y PostgreSQL. También puede utilizar la herramienta Babelfish Compass para evaluar la compatibilidad con Babelfish para Aurora PostgreSQL. Esto convierte a Compass AWS SCT y a Compass en herramientas poderosas para comprender el trabajo inicial que implica antes de decidir una estrategia de migración. Si decide continuar, AWS SCT automatiza los cambios necesarios en el esquema. La filosofía central de Babelfish Compass es permitir que la base de datos SQL se migre a Aurora sin modificaciones o con muy pocas modificaciones. Compass evaluará la base de datos SQL existente para determinar si esto se puede lograr. De esta forma, se puede saber el resultado antes de dedicar ningún esfuerzo a migrar datos de SQL Server a Aurora.

AWS SCT automatiza la conversión y migración del esquema y el código de la base de datos al motor de base de datos de destino. Puede utilizar Babelfish para Aurora PostgreSQL para migrar la base de datos y su aplicación de SQL Server a Aurora PostgreSQL sin cambiar el esquema o haciendo cambios mínimos. Esto puede acelerar sus migraciones.

Una vez migrado el esquema, puede utilizarlo AWS DMS para migrar los datos. AWS DMS puede realizar una carga completa de datos y replicar los cambios para realizar la migración con un tiempo de inactividad mínimo.

En esta sección se analizan las siguientes herramientas con más detalle:

- AWS Schema Conversion Tool
- Babelfish para Aurora PostgreSQL
- Babelfish Compass
- AWS Database Migration Service

AWS Schema Conversion Tool

Puede utilizarlas AWS SCT para evaluar sus bases de datos de SQL Server existentes y evaluar la compatibilidad con Amazon RDS o Aurora. Para simplificar el proceso de migración, también puede convertir el esquema de un motor de base de datos a otro en una migración de base de datos heterogénea. AWS SCT Se puede utilizar AWS SCT para evaluar la aplicación y convertir el código de la aplicación incrustado para aplicaciones escritas en C#, C++, Java y otros lenguajes. Para obtener más información, consulte [Converting application SQL using AWS SCT](#) en la documentación de AWS SCT .

AWS SCT es una AWS herramienta gratuita que admite muchas [fuentes](#) de bases de datos. Para utilizarla AWS SCT, hay que dirigirla a la base de datos de origen y, a continuación, realizar una evaluación. Luego, [AWS SCT](#) evalúa el esquema y genera el informe de evaluación. Los informes de evaluación incluyen un resumen ejecutivo, la complejidad de la migración y el esfuerzo requerido, los motores de bases de datos de destino adecuados y recomendaciones para la conversión. Para descargarla AWS SCT, consulte [Instalación, verificación y actualización AWS SCT](#) en la AWS SCT documentación.

La siguiente tabla muestra un ejemplo de resumen ejecutivo generado por AWS SCT para que pueda ver lo complejo que es cambiar la base de datos a diferentes plataformas de destino.

Plataforma de destino	Cambios automáticos o mínimos			Acciones complejas			
	Objetos de almacenamiento	Objetos de código	Funciones de conversión	Objetos de almacenamiento		Objetos de código	
Amazon RDS para MySQL	60 (98 %)	8 (35 %)	42	1 (2 %)	1	15 (65 %)	56
Amazon Aurora MySQL-Compatible Edition	60 (98 %)	8 (35 %)	42	1 (2 %)	1	15 (65 %)	56
Amazon RDS para PostgreSQL	60 (98 %)	12 (52 %)	54	1 (2 %)	1	11 (48 %)	26

Edición de Amazon Aurora compatible con PostgreSQL	60 (98 %)	12 (52 %)	54	1 (2 %)	1	11 (48 %)	26
Amazon RDS para MariaDB	60 (98 %)	7 (30 %)	42	1 (2 %)	1	16 (70 %)	58
Amazon Redshift	61 (100 %)	9 (39 %)	124	0 (0 %)	0	14 (61 %)	25
AWS Glue	0 (0 %)	17 (100 %)	0	0 (0 %)	0	0 (0 %)	0
Babelfish	59 (97 %)	10 (45 %)	20	2 (3 %)	2	12 (55 %)	30

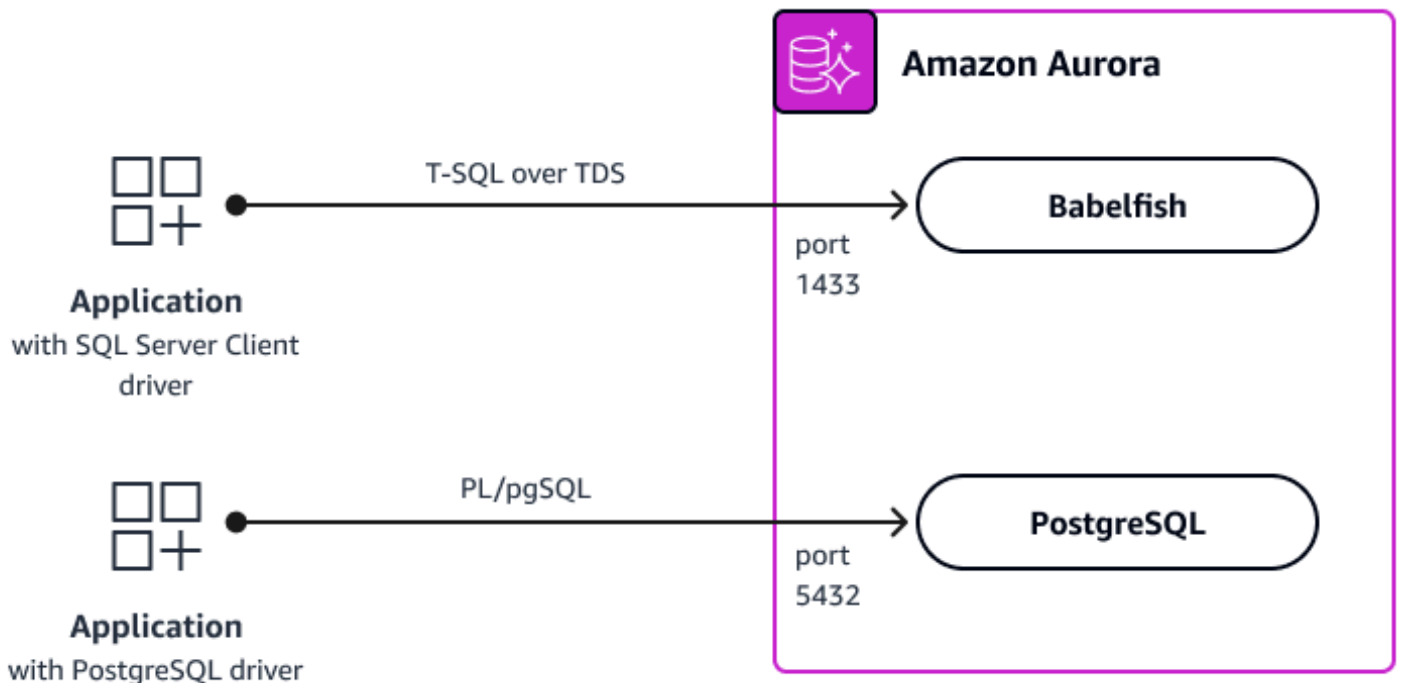
Un AWS SCT informe también proporciona detalles sobre los elementos del esquema que no se pueden convertir automáticamente. Puede cerrar las brechas AWS SCT de conversión y optimizar los esquemas de destino consultando los manuales de [AWS migración](#). Existen muchos manuales de procedimientos de migración de bases de datos útiles para facilitar las migraciones heterogéneas.

Babelfish para Aurora PostgreSQL

Babelfish para Aurora PostgreSQL amplía su clúster de base de datos de Aurora PostgreSQL porque puede aceptar conexiones de bases de datos de clientes de SQL Server. Con Babelfish, las aplicaciones que se crearon originalmente para SQL Server pueden funcionar directamente con Aurora PostgreSQL haciendo pocos cambios en el código en comparación con una migración tradicional, y sin tener que cambiar los controladores de base de datos. Babelfish hace que Aurora PostgreSQL sea bilingüe para que Aurora PostgreSQL pueda funcionar tanto con T-SQL como con los lenguajes PL/pgSQL. Babelfish minimiza los esfuerzos de migración de SQL Server a Aurora PostgreSQL. Esto acelera las migraciones, minimiza el riesgo y reduce los costos de migración de manera significativa. Puede seguir utilizando T-SQL después de completar la

migración, pero también existe la [opción de utilizar las herramientas nativas de PostgreSQL](#) para el desarrollo.

El siguiente diagrama ilustra cómo una aplicación que usa T-SQL se conecta al puerto predeterminado 1433 de SQL Server y usa el traductor Babelfish para comunicarse con la base de datos de Aurora PostgreSQL, mientras que una aplicación que usa PL/pgSQL puede conectarse directa y simultáneamente a la base de datos de Aurora PostgreSQL mediante el puerto predeterminado 5432 en Aurora PostgreSQL.



Babelfish no admite determinadas características de T-SQL de SQL Server. Por este motivo, Amazon proporciona herramientas de evaluación para line-by-line analizar tus sentencias SQL y determinar si alguna de ellas no es compatible con Babelfish.

Hay dos opciones para las evaluaciones de Babelfish. AWS SCT puede evaluar la compatibilidad de su base de datos de SQL Server con Babelfish. Otra opción es la herramienta Babelfish Compass, que es una solución recomendada porque la herramienta Compass se actualiza en función de las nuevas versiones de Babelfish para Aurora PostgreSQL.

Babelfish Compass

[Babelfish Compass](#) es una herramienta que se puede descargar gratuitamente y que se adhiere a la última versión de Babelfish para Aurora PostgreSQL. Por el contrario, AWS SCT admitirá las versiones más nuevas de Babelfish después de un tiempo. [Babelfish Compass](#) usa el esquema

de base de datos de SQL Server. También puede extraer el esquema de la base de datos de SQL Server de origen mediante herramientas como SQL Server Management Studio (SSMS). A continuación, puede poner en marcha el esquema usando Babelfish Compass. Esto genera el informe que detalla la compatibilidad del esquema de SQL Server con Babelfish y si es necesario realizar algún cambio antes de la migración. La herramienta Babelfish Compass también puede automatizar muchos de estos cambios y, en última instancia, acelerar las migraciones.

Una vez finalizada la evaluación y los cambios, puede migrar el esquema a Aurora PostgreSQL mediante herramientas nativas de SQL Server, como SSMS o sqlcmd. Para obtener instrucciones, consulte la publicación [Migrate from SQL Server to Amazon Aurora using Babelfish](#) en el blog de bases de datos de AWS .

AWS Database Migration Service

Una vez migrado el esquema, puedes usar AWS Database Migration Service (AWS DMS) para migrar los datos AWS con un tiempo de inactividad mínimo. AWS DMS no solo carga todos los datos, sino que también replica los cambios del origen al destino mientras el sistema de origen está en funcionamiento. Una vez sincronizadas las bases de datos de origen y destino, la actividad de transición puede tener lugar en la que la aplicación apunta a la base de datos de destino que finaliza la migración. AWS DMS actualmente solo realiza una carga de datos completa con Babelfish para un destino de Aurora PostgreSQL y no replica los cambios. Para obtener más información, consulte [Uso de Babelfish como](#) objetivo en la documentación. AWS Database Migration Service AWS DMS

AWS DMS puede realizar migraciones homogéneas (en el mismo motor de base de datos) y heterogéneas (en diferentes motores de base de datos). AWS DMS admite muchos motores de bases de datos de origen y destino. Para obtener más información, consulte la publicación [Migración de la base de datos de SQL Server a Amazon RDS for SQL Server AWS DMS](#) mediante AWS el blog de bases de datos.

Recursos adicionales

- [Adiós Microsoft SQL Server, hola Babelfish \(blog de AWS noticias\)](#)
- [Convierta esquemas de bases de datos y SQL de aplicaciones mediante la AWS Schema Conversion Tool CLI \(blog de AWS bases de datos\)](#)
- [Migre SQL Server a Amazon Aurora PostgreSQL con las prácticas recomendadas y las lecciones aprendidas sobre el terreno AWS \(blog sobre bases de datos\)](#)
- [Valide los objetos de la base de datos después de la migración de Microsoft SQL Server a Amazon RDS para PostgreSQL y Amazon Aurora PostgreSQL \(blog sobre bases de datos\)AWS](#)

Optimización del almacenamiento para SQL Server

Descripción general de

Esta sección se centra en las optimizaciones de costos del almacenamiento en SSD de Amazon Elastic Block Store (Amazon EBS) para cargas de trabajo de EC2.

Dispone de una amplia variedad de opciones de almacenamiento para implementar y ejecutar cargas de trabajo de SQL Server. AWS La selección del almacenamiento adecuado debe basarse en el propósito, la arquitectura, la durabilidad, el rendimiento, la capacidad y el costo. AWS los clientes que ejecutan cargas de trabajo de SQL Server suelen utilizar una combinación de almacenamiento de Amazon EBS FSx, NVMe Amazon y Amazon Simple Storage Service (Amazon S3).

Amazon EBS es un almacenamiento conectado a la red vinculado a instancias de computación de EC2 y que se utiliza para almacenar y procesar archivos generales del sistema operativo, las aplicaciones, las bases de datos y las copias de seguridad. El almacenamiento en unidades de estado sólido (SSD) de Amazon EBS incluye SSD de uso general (gp2 y gp3) y SSD de IOPS aprovisionadas (io1, io2 e io2Bx). Considere lo siguiente:

- Algunas instancias de EC2, como la r5d, tienen una conexión NVMe SSDs física local a la instancia host. Estos volúmenes proporcionan un almacenamiento por bloques que se suele utilizar para tempdb y la extensión de grupo de búferes de SQL Server.
- Amazon FSx for Windows File Server es un servicio de almacenamiento de archivos totalmente gestionado, mientras que Amazon FSx NetApp para ONTAP es un almacenamiento compartido totalmente gestionado NetApp basado en el popular sistema de archivos ONTAP. Amazon FSx se utiliza con frecuencia para poner en marcha cargas de trabajo de SQL Server en una configuración de instancias de clústeres de conmutación por error (FCI) de SQL Server de alta disponibilidad. Esta solución aloja archivos de registro y datos de SQL Server, lo que reduce los requisitos de rendimiento de EBS en las instancias de EC2.
- Amazon S3 es un servicio de almacenamiento de objetos que ofrece características de escalabilidad, disponibilidad de datos, seguridad y rendimiento líderes del sector. Puede almacenar archivos de backup nativos de SQL Server AMIs, instantáneas de EBS, registros de aplicaciones y mucho más en Amazon S3.

Tipos de almacenamiento en SSD, rendimiento y costo para Amazon EBS

Los costos de almacenamiento en SSD de Amazon EBS suelen ser más altos a medida que aumentan la durabilidad y el rendimiento. Actualmente, se ofrecen cinco volúmenes de almacenamiento, cada uno con sus [propias métricas de rendimiento únicas](#). Para obtener un resumen de los casos de uso y las características de los volúmenes respaldados por SSD, consulte la tabla de la sección [Solid state drive \(SSD\) volumes](#) de la documentación de Amazon EBS.

Puedes usar Amazon CloudWatch para monitorear el rendimiento de las SSD, capturar datos de tendencias y configurar alarmas cuando se alcancen ciertos umbrales. Si está ejecutando cargas de trabajo de SQL Server AWS, considere la posibilidad de habilitar una [supervisión detallada](#) y de implementar [métricas CloudWatch personalizadas para capturar métricas](#) detalladas de rendimiento del volumen, como la latencia del disco, las IOPS, el rendimiento, la longitud de la cola de discos, la capacidad utilizada frente a la capacidad libre, etc. Puede utilizar estas métricas de CloudWatch rendimiento para identificar el almacenamiento subaprovisionado y sobreaprovisionado y proporcionar puntos de datos históricos para definir con precisión los requisitos de almacenamiento.

Los costos de almacenamiento en SSD de Amazon EBS también varían en función de la capacidad asignada. La siguiente tabla muestra una comparación entre los diferentes tipos de volumen. Todos los tipos de volumen tienen 1 TB de capacidad y configuraciones de rendimiento similares.

Tipo de volumen	IOPS máximas (E/S de 16 KiB)	Rendimiento máximo (E/S de 128 KiB)	Precio por 1 TB	Porcentaje de ahorro de costos
gp2	3000	250	102,40\$	
gp3	3000	250	86,92 DÓLARES	15%
io1	16,000	500	1.168 DÓLARES	
io2	16,000	500	1.168 DÓLARES	
gp3	16,000	500	146,92 DÓLARES	87%
io2bx	16,000	4.000	1.168 DÓLARES	

Tipo de volumen	IOPS máximas (E/S de 16 KiB)	Rendimiento máximo (E/S de 128 KiB)	Precio por 1 TB	Porcentaje de ahorro de costos
gp3	16,000	1 000	181,92 DÓLARES	84%

Note

Las métricas de rendimiento y costo de la tabla anterior son por volumen y se basan en una [estimación](#) de la Calculadora de precios de AWS. Cuenta de AWS Se requiere un número para acceder a la estimación incluida en Calculadora de precios de AWS.

Los volúmenes gp3 basados en SSD de Amazon EBS ofrecen un rendimiento excelente a un bajo costo. Puede ahorrar hasta un 87 por ciento si elige un volumen gp3 en lugar de los volúmenes io1 o io2 para cargas de trabajo que requieren menos de 16 000 IOPS y un rendimiento de 500. MiBps

Los volúmenes io2 Block Express (io2Bx) ofrecen un rendimiento superior al de los volúmenes io2 normales. Con 16 000 IOPS, los volúmenes io1 o io2 solo tienen un rendimiento de 500, mientras que los volúmenes de io2Bx se pueden configurar con un MiBps rendimiento de hasta 4000. MiBps En comparación con los volúmenes io1 e io2, los volúmenes io2Bx ofrecen un rendimiento más de cuatro veces superior, entre 16 000 y 64 000 IOPS, exactamente al mismo precio. Los volúmenes io2 normales se pueden convertir en volúmenes io2BX conectándolas a instancias de EC2 compatibles con io2BX. Para obtener una lista de instancias de EC2 compatibles con io2BX, consulte [Provisioned IOPS SSD volumes](#) en la documentación de Amazon EBS. Antes de implementar un nuevo almacenamiento, puede usar la [Calculadora de precios de AWS](#) para estimar su costo mensual y comprender el impacto en el costo en función de las compensaciones entre durabilidad, rendimiento y capacidad.

Optimización general de los costos de SSD para Amazon EBS

Le recomendamos que evalúe lo que va a almacenar y que se asegure de utilizar el tipo y la clase de almacenamiento correctos. Por ejemplo, Amazon S3 ofrece un precio excelente, políticas de ciclo de vida integradas y opciones de replicación ideales para las copias de seguridad de SQL Server. SQL Server 2022 tiene la capacidad de realizar copias de seguridad directamente en Amazon S3, mientras que las versiones anteriores de SQL Server se basan en copias de seguridad locales

nativas. Si utiliza versiones anteriores de SQL Server, considere la posibilidad de realizar copias de seguridad en volúmenes HDD de Amazon EBS y, a continuación, llevar la copia de seguridad a Amazon S3. Esta solución puede ahorrar un 53 % en comparación con el uso de volúmenes gp3 para las copias de seguridad.

La siguiente tabla muestra la diferencia de precio de 1 TB de almacenamiento cuando se usa un volumen gp3 de Amazon EBS, un volumen st1 HDD de Amazon EBS o Amazon S3.

Tipo de almacenamiento	Capacidad	Precio al mes
EBS gp3 500 MiBps	1 TB	96,92 DÓLARES
EBS st1 Burst 500 MiBps		46,08 DÓLARES
S3 Standard		23,55 DÓLARES
S3 Standard (acceso poco frecuente)		12,80 DÓLARES
S3 Glacier Deep Archive		1,03\$

Note

Las métricas de costos de la tabla anterior se basan en una [estimación](#) de la Calculadora de precios de AWS. Cuenta de AWS Se requiere un número para acceder a la estimación del Calculadora de precios de AWS.

Le recomendamos que tenga en cuenta lo siguiente:

- Permita una supervisión detallada e implemente métricas CloudWatch personalizadas para capturar con precisión sus requisitos de rendimiento de almacenamiento.
- Actualice el almacenamiento de Amazon EBS de gp2 a gp3 para reducir costos, aumentar la flexibilidad y mejorar el rendimiento.
- Actualice el almacenamiento de Amazon EBS de io1 a io2 para aumentar la durabilidad y la flexibilidad del rendimiento.
- Utilice io2BX en lugar de io1 o io2 cuando sea posible para aumentar la durabilidad y el rendimiento.

- Considere un mix-and-match enfoque al elegir el almacenamiento que ayude a reducir los requisitos de capacidad y el costo de los volúmenes de alto rendimiento. Por ejemplo, podría utilizar volúmenes gp3 de bajo costo para el volumen raíz (sistema operativo), la instalación de SQL Server, las bases de datos del sistema (excepto tempdb) y las bases de datos de usuarios de menor rendimiento. Esto podría ayudar a reducir la capacidad y el costo de un volumen io2, que puede dedicarse a bases de datos de usuarios de alto rendimiento.
- Si aloja bases de datos de SQL Server AWS, le recomendamos que utilice varios archivos de datos de SQL Server por base de datos. Esto brinda la oportunidad de distribuir las read/write cargas de trabajo en varios volúmenes, lo que reduce los requisitos de rendimiento y capacidad por volumen y, en consecuencia, reduce los costos.
- Incluso si las cargas de trabajo de producción requieren un almacenamiento de mayor rendimiento, como io1 o io2/io2BX, puede usar los volúmenes gp3 para las cargas de trabajo que no son de producción para reducir los costos.
- Supervise y analice la utilización del almacenamiento a lo largo del tiempo para identificar fácilmente picos de uso y costos inesperados.
- Use [AWS Compute Optimizer](#) para obtener recomendaciones sobre cómo aumentar o reducir los volúmenes de EBS en función de la utilización real.
- Utilice la elasticidad de AWS para ajustar las necesidades de rendimiento y capacidad de sus volúmenes de SSD para Amazon EBS. A diferencia de los entornos en las instalaciones, no es necesario aprovisionar en exceso el rendimiento y la capacidad del almacenamiento para futuras cargas de trabajo. Puede migrar sus cargas de trabajo de SQL Server existentes AWS y ajustar el rendimiento o la capacidad según sea necesario y, al mismo tiempo, mantener sus bases de datos en línea.

Recursos adicionales

- [Amazon EBS volume types](#) (documentación de Amazon EBS)
- [Amazon Elastic Block Store \(Amazon EBS\)](#) (documentación de Amazon EBS)
- [Provisioned IOPS SSD volumes](#) (documentación de Amazon EBS)
- [Volúmenes del almacén de instancias SSD](#) (documentación de Amazon EC2)
- [CloudWatch Métricas de Amazon para Amazon EBS](#) (documentación de Amazon EBS)
- [Specifications for Amazon EC2 storage optimized instances](#) (documentación de Amazon EC2)
- [Proteja sus cargas de trabajo de SQL Server NetApp SnapCenter con Amazon FSx for NetApp ONTAP \(blog sobre AWS almacenamiento\)](#)

- [Preguntas frecuentes sobre Amazon EC2 \(página AWS del producto\)](#)

Optimización de las licencias de SQL Server con Compute Optimizer

Guía sobre cómo optimizar las licencias de SQL Server mediante el uso AWS Compute Optimizer.

Descripción general de

[AWS Compute Optimizer](#) puede recomendar oportunidades de optimización de licencias para cargas de trabajo de Microsoft SQL Server en Amazon Elastic Compute Cloud (Amazon EC2). Compute Optimizer puede proporcionar recomendaciones automatizadas para reducir los costos asociados al uso de licencias. Las recomendaciones de Compute Optimizer aparecen junto a cada una de las instancias de EC2 con licencias de Microsoft SQL Server. La información que se proporciona incluye las oportunidades de ahorro recomendadas, los precios bajo demanda de las instancias de EC2 y los precios de cada hora con traiga su propia licencia (BYOL). Esta información puede ser útil para decidir si debe adquirir la licencia de una versión inferior.

Compute Optimizer detecta automáticamente las instancias de SQL Server en Amazon EC2 según el tipo de carga de trabajo inferida. Para ver las recomendaciones de licencias, puedes seleccionar las instancias de SQL Server en Compute Optimizer y, a continuación, autenticarte con [Amazon CloudWatch Application Insights](#) con tus credenciales de base de datos de solo lectura. Compute Optimizer analiza si utiliza alguna de las características de la edición Enterprise de SQL Server. Si no se utilizan características de la edición Enterprise, Compute Optimizer le recomendará que cambie a la edición Standard para reducir los costos asociados al uso de licencias.

También puede usar Compute Optimizer para recibir recomendaciones de tamaño para las instancias de Amazon EC2 que usen cargas de trabajo de SQL Server. Para obtener más información, consulte [Optimización del tamaño de SQL Server con Compute Optimizer](#) en esta guía.

Recomendaciones de optimización de costos

Las recomendaciones de licencia de Compute Optimizer pueden ser útiles para evaluar las características que utiliza en Microsoft SQL Server y a elegir la edición más rentable para sus cargas de trabajo. La edición Enterprise de SQL Server es considerablemente más cara que la edición Standard. Para obtener más información, consulte [Comparación de las ediciones de SQL Server](#) en esta guía y [Precios y licencias de SQL Server 2022](#) en el sitio web de Microsoft. Invertir tiempo

en configurar Compute Optimizer para evaluar su flota de SQL Server y obtener recomendaciones puede reducir drásticamente los costos asociados al uso de licencias.

La página Detalles de la licencia proporciona la siguiente información:


- Utilice la tabla para comparar las configuraciones de licencia actuales (como la edición, el modelo y la cantidad de núcleos de instancia) con las recomendaciones de Compute Optimizer.
- Utilice los gráficos de uso para comprobar el número de características de la edición Enterprise que se utilizaron durante el periodo de análisis.

Para obtener más información, consulte [Viewing details of a commercial software license recommendation](#) en la documentación de Compute Optimizer.

Configuración de Compute Optimizer

Compute Optimizer analiza las licencias de software comercial mediante la métrica `mssql_enterprise_features_used`. Para obtener más información sobre esta métrica, consulte [Metrics for commercial software licenses](#).

1. Asegúrese de que el usuario cuente con los permisos necesarios para darse de alta en Compute Optimizer. Para obtener más información, consulte los siguientes temas:
 - [Política de suscripción a Compute Optimizer](#)
 - [Policies to grant access to Compute Optimizer for standalone Cuentas de AWS](#)
 - [Políticas para conceder acceso a Compute Optimizer para una cuenta de administración de una organización](#)
2. Adjunta las funciones y la política de instancia requeridas para CloudWatch Application Insights. Para obtener instrucciones, consulte [Policies to enable commercial software license recommendations](#).
3. Habilite CloudWatch Application Insights mediante sus credenciales de base de datos de Microsoft SQL Server. Para obtener instrucciones, consulte [Configurar la aplicación para la supervisión](#) en la CloudWatch documentación.

 Note


Para generar recomendaciones de licencias de software comercial, se requieren al menos 30 horas consecutivas de datos CloudWatch métricos. Para obtener más información, consulte [los requisitos de CloudWatch métricas](#).

4. Utilice la siguiente consulta SQL para configurar el acceso con privilegios mínimos para CloudWatch Application Insights.

```
GRANT VIEW SERVER STATE TO [LOGIN];  
GRANT VIEW ANY DEFINITION TO [LOGIN];
```

Esto habilita un nuevo servicio, SQL. PrometheusSqlExporter

5. Desde la cuenta de administración de destino Cuenta de AWS o de la organización, activa Compute Optimizer. Para obtener instrucciones, consulte [Opting in your account](#).

 Note

Una vez haya realizado la activación, los resultados y las recomendaciones de optimización pueden tardar hasta 24 horas en generarse.

6. En la [consola de Compute Optimizer](#), seleccione Licencias en el panel de navegación.
7. En la columna Resultados, busque cualquier instancia con el resultado Métricas insuficientes. Compute Optimizer devuelve este resultado si detecta que CloudWatch Application Insights no está habilitada o tiene permisos insuficientes. Para obtener más información, consulte [Finding reasons](#). Haga lo siguiente para resolver estos resultados:
 - a. Seleccione la instancia.
 - b. Agregue un secreto.
 - c. Confirme que la política y el rol de instancia estén asociados.
 - d. Elija Habilitar las recomendaciones de licencias.
8. En la columna Resultados, busque cualquier instancia con el resultado No optimizadas. Compute Optimizer devuelve este resultado si detecta que su infraestructura de Amazon EC2 no utiliza ninguna de las características de licencia de Microsoft SQL Server por las que está pagando. Para obtener más información, consulte [Finding reasons](#). Haga lo siguiente para resolver estos resultados:

- a. Seleccione la instancia.
- b. Compare la edición de la licencia actual con la edición de la licencia recomendada.
- c. Revise el gráfico de uso de las licencias actuales.
- d. Si quiere cambiar de licencia y usar la de una edición inferior, elija Implementar la recomendación.
- e. Revise los requisitos y siga las instrucciones para cambiar de licencia y usar la de una versión inferior. Si quieres automatizar el proceso, consulta [Cómo cambiar la edición Enterprise de SQL Server mediante AWS Systems Manager Document para reducir los costes](#) (AWS blog).

Recursos adicionales

- [Reduzca los costos de licencias de Microsoft SQL Server con AWS Compute Optimizer](#) (AWS Blog)
- [¿Qué es AWS Compute Optimizer?](#) (AWS documentación)
- [Viewing commercial software license recommendations](#) (documentación de AWS)
- [Downgrade your Microsoft SQL Server edition](#) (documentación de AWS)
- [Microsoft SQL Server on AWS](#) (AWS)
- [Microsoft Licensing on AWS](#) (AWS)
- [Precios de Microsoft SQL Server 2019](#) (Microsoft)
- [Precios de Microsoft SQL Server 2022](#) (Microsoft)

Optimización del tamaño de SQL Server con Compute Optimizer

Descripción general de

[AWS Compute Optimizer](#) ayuda a los administradores de bases de datos (DBA) a detectar las cargas de trabajo de Microsoft SQL Server en Amazon Elastic Compute Cloud (Amazon EC2) y a ajustar el tamaño de las instancias de EC2 para reducir los costos de licencia hasta en un 25 %. La función de [tipo de carga de trabajo inferida](#) de Compute Optimizer utiliza el aprendizaje automático (ML) y detecta automáticamente las aplicaciones que podrían estar ejecutándose en AWS tus recursos. Compute Optimizer incluye soporte para SQL Server como un tipo de carga de trabajo inferida. Al utilizar la característica de tipo de carga de trabajo inferida, puede identificar oportunidades de ahorro de costos en función de la carga de trabajo específica que esté en marcha en sus instancias de Amazon EC2.

Con esta característica, puede clasificar las oportunidades de ahorro de costos según los tipos de carga de trabajo inferida compatibles, como SQL Server. Compute Optimizer puede detectar automáticamente las instancias de EC2 de SQL Server que están sobreaprovisionadas. Puede cambiar a la consola de EC2 para reducir el tamaño de la instancia, lo que ayuda a reducir los costos asociados a la infraestructura y al uso de licencias.

También puede usar Compute Optimizer para obtener recomendaciones de licencias de SQL Server. Para obtener más información, consulte [Optimización de las licencias de SQL Server con Compute Optimizer](#) en esta guía.

Configuración de Compute Optimizer

Para obtener instrucciones sobre cómo usar Compute Optimizer con cargas de trabajo inferidas de SQL Server, consulte [Optimización del rendimiento y reducción de los costos de licencia: aprovechamiento de las instancias de SQL Server de AWS Compute Optimizer Amazon EC2 \(blog\)](#). AWS Puede activarlo para cuentas independientes, cuentas que sean miembros de una organización y cuentas de administración de una organización. En el caso de las cuentas independientes y de miembros, la opción activa Compute Optimizer únicamente para esa cuenta. En el caso de una cuenta de administración de la organización, puede activar Compute Optimizer solo para esa cuenta o para todas las cuentas de miembros de la organización.

El proceso de suscripción de Compute Optimizer crea automáticamente un rol vinculado al servicio AWS Identity and Access Management (IAM). Para obtener más información, consulte [Uso de roles vinculados a servicios para AWS Compute Optimizer](#).

Compute Optimizer analiza los recursos en función de CloudWatch las métricas de Amazon, como la CPU, las E/S, la red y el uso de Amazon Elastic Block Store (Amazon EBS). Para generar recomendaciones, se requieren al menos 30 horas consecutivas de datos de CloudWatch métricas en los últimos 14 días. Si activa la característica de métricas de infraestructura mejoradas, las métricas de uso se ampliarán a 93 días. Para obtener más información, consulta [los requisitos de CloudWatch métricas](#) y [las métricas de infraestructura mejorada](#) en la documentación de Compute Optimizer.

Compute Optimizer ofrece opciones y los ahorros asociados a cada opción en función de la vCPU, la memoria, el almacenamiento, la red, el riesgo y el esfuerzo de migración. Puedes usar el panel de CloudWatch métricas para analizar los datos que se utilizan para hacer la recomendación. Con estos datos, puede dimensionar correctamente las instancias de EC2 que usan cargas de trabajo de SQL Server. Para obtener más información sobre los tipos de instancias de Amazon EC2, consulte [Cambios de tipos de instancias de Amazon EC2](#) en la documentación de Amazon EC2.

Recursos adicionales

- [AWS Compute Optimizer identifica y filtra las cargas de trabajo de Microsoft SQL Server](#) (AWS)
- [Optimización del rendimiento y reducción de los costes de licencias: aprovechamiento de AWS Compute Optimizer las instancias AWS de SQL Server de Amazon EC2](#) (blog)
- [¿Qué es? AWS Compute Optimizer](#) (AWS documentación)
- [Visualización de las recomendaciones de instancias de EC2](#) (AWS documentación)

Revise Trusted Advisor las recomendaciones para las cargas de trabajo de SQL Server

Descripción general de

[AWS Trusted Advisor](#) proporciona recomendaciones que le permiten seguir las prácticas recomendadas de AWS. Al analizar el uso, la configuración y los gastos, Trusted Advisor ofrece recomendaciones prácticas para reducir los costos, mejorar la disponibilidad y el rendimiento del sistema o ayudar a cerrar las brechas de seguridad. Esta sección se centra en las Trusted Advisor comprobaciones que pueden ayudarlo a reducir los costos de operación de las cargas de trabajo de SQL Server en Nube de AWS.

Recomendaciones de optimización de costos

Trusted Advisor proporciona recomendaciones que le ayudan a optimizar sus cargas de trabajo de SQL Server en Amazon Elastic Compute Cloud (Amazon EC2). Las comprobaciones inspeccionan las cargas de trabajo de SQL Server y enumeran automáticamente las instancias que tienen que optimizarse. Poner en práctica Trusted Advisor las recomendaciones puede reducir los costos y mejorar la postura de seguridad de su organización.

Las siguientes son Trusted Advisor comprobaciones que se centran en Microsoft SQL Server:

- [Instancias de Amazon EC2 sobreprovisionadas para Microsoft SQL Server](#): esta comprobación analiza las instancias de Amazon EC2 que usan SQL Server y le avisa si una instancia supera el límite de vCPU del software de SQL Server. Por ejemplo, una instancia con la edición SQL Server Standard puede utilizar hasta 48 vCPUs. Una instancia con SQL Server Web puede utilizar hasta 32 vCPUs.

Edition	Mín. de vCPU	Máx. de vCPU
Web	4	32
Standard	4	48
Enterprise	4	Límites del SO

- [Unificación de instancias de Amazon EC2 para Microsoft SQL Server](#): esta verificación analiza las instancias de Amazon EC2 y le avisa si la instancia tiene menos de la cantidad mínima de licencias de SQL Server. Puede consolidar instancias de SQL Server más pequeñas para ayudar a reducir los costos. Si tiene muchas instancias pequeñas de SQL Server con licencia incluida, considere la posibilidad de unificarlas. Según la [guía de licencias de Microsoft SQL Server 2019](#), SQL Server requiere un mínimo de 4 vCPU con licencia por instancia. Si unifica estas bases de datos, puede ahorrar en los costos asociados al uso de licencias. Puede tomar su decisión en función del número de bases de datos de la instancia, el tamaño máximo de la base de datos y el tamaño total de las bases de datos. La unificación es compatible con las ediciones Web, Standard y Enterprise de SQL Server. Para obtener más información, consulte [Consolidating SQL Server Databases](#) (publicación del blog de Microsoft).

AWS no recomienda colocar bases de datos de producción de gran tamaño en un solo servidor. Sin embargo, puede unificar las más pequeñas que se utilizan en entornos que no sean de producción, es el caso de los entornos de desarrollo, pruebas y almacenamiento provisional. Esto depende del uso actual de SQL Server; si tiene bases de datos de bajo uso, puede unificarlas en un mismo servidor.

Configuración Trusted Advisor

Haga lo siguiente para evaluar los registros centrados en SQL Server Trusted Advisor.

1. Inicie sesión en Consola de administración de AWS.
2. Abra la [consola de AWS Trusted Advisor](#).
3. En el panel de navegación, en Recomendaciones, elija Optimización de costos.
4. En la lista Comprobaciones de optimización de costos, revise el estado de las comprobaciones de unificación de las instancias de Amazon EC2 para Microsoft SQL Server y de sobreaprovisionamiento de las instancias de Amazon EC2 para Microsoft SQL Server.

- Los símbolos de comprobación verdes indican que las instancias de Amazon EC2 están configuradas de forma óptima.
 - Los símbolos de alerta de color naranja indican que hay oportunidades de mejora.
5. Seleccione una comprobación para ver sus detalles y recomendaciones.
 6. Siga las instrucciones de la comprobación para optimizar las instancias de Amazon EC2 que usan cargas de trabajo de SQL Server.
 7. Supervise sus instancias con regularidad y actualice las comprobaciones periódicamente.

Recursos adicionales

- [Trusted Advisor comprobar la referencia](#) (AWS documentación)
- [Microsoft SQL Server on AWS](#) (AWS)
- [Microsoft Licensing on AWS](#) (AWS)
- [Precios de SQL Server 2019](#) (Microsoft)
- [AWS Launch Wizard para SQL Server](#) (AWS documentación)

Contenedores

La modernización es un proceso de transformación que ofrece muchas opciones, como descomponer monolitos en microservicios, rediseñar la arquitectura de las aplicaciones para que se basen en eventos mediante el uso de funciones sin servidor (AWS Lambda) y reutilizar bases de datos de SQL Server en Amazon Aurora o bases de datos administradas y personalizadas. Las vías de modernización para redefinir la plataforma de las aplicaciones de .NET Framework a contenedores de Linux y Windows requieren menos esfuerzo que otras opciones de modernización. Los contenedores ofrecen las siguientes ventajas:

- **Aceleración de la innovación:** el uso de contenedores facilita la automatización de las fases del ciclo de vida del desarrollo, que incluyen la creación, las pruebas y la implementación de aplicaciones. Al automatizar estos procesos, los equipos de desarrollo y operaciones disponen de más tiempo para centrarse en la innovación.
- **Reducción del costo total de propiedad (TCO):** el uso de contenedores también puede hacer que dependa menos de las herramientas de administración de licencias y protección de puntos de conexión. Como los contenedores son unidades de computación efímeras, puede automatizar y simplificar las tareas de administración, como la aplicación de parches, el escalado y las copias de seguridad y la restauración. Esto reduce el costo total de propiedad de la administración y del funcionamiento de las cargas de trabajo basadas en contenedores. Por último, los contenedores son más eficientes que las máquinas virtuales, ya que puede usarlos para maximizar la ubicación de las aplicaciones gracias a un mejor aislamiento. Esto aumenta la utilización de los recursos de infraestructura de la aplicación.
- **Mejora de la utilización de los recursos:** los contenedores son más eficientes que las máquinas virtuales porque puede usarlos para maximizar la ubicación de las aplicaciones. Esto aumenta la utilización de los recursos de infraestructura de la aplicación gracias a un mejor aislamiento.
- **Cierre la brecha de habilidades:** AWS ofrece días de inmersión para capacitar a sus equipos de desarrollo en tecnología y DevOps prácticas de contenedores.

Esta sección abarca los siguientes temas:

- [Transferencia de aplicaciones de Windows a contenedores](#)
- [Optimice los costos de AWS Fargate las tareas en Amazon ECS](#)
- [Mejora la visibilidad de sus costos en Amazon EKS.](#)
- [Redefinición de la plataforma de las aplicaciones de Windows con App2Container](#)

Para obtener información sobre licencias, consulte la sección Licensing de [Amazon Web Services and Microsoft FAQs](#) o envíe sus preguntas por correo electrónico a microsoft@amazon.com.

Transferencia de aplicaciones de Windows a contenedores

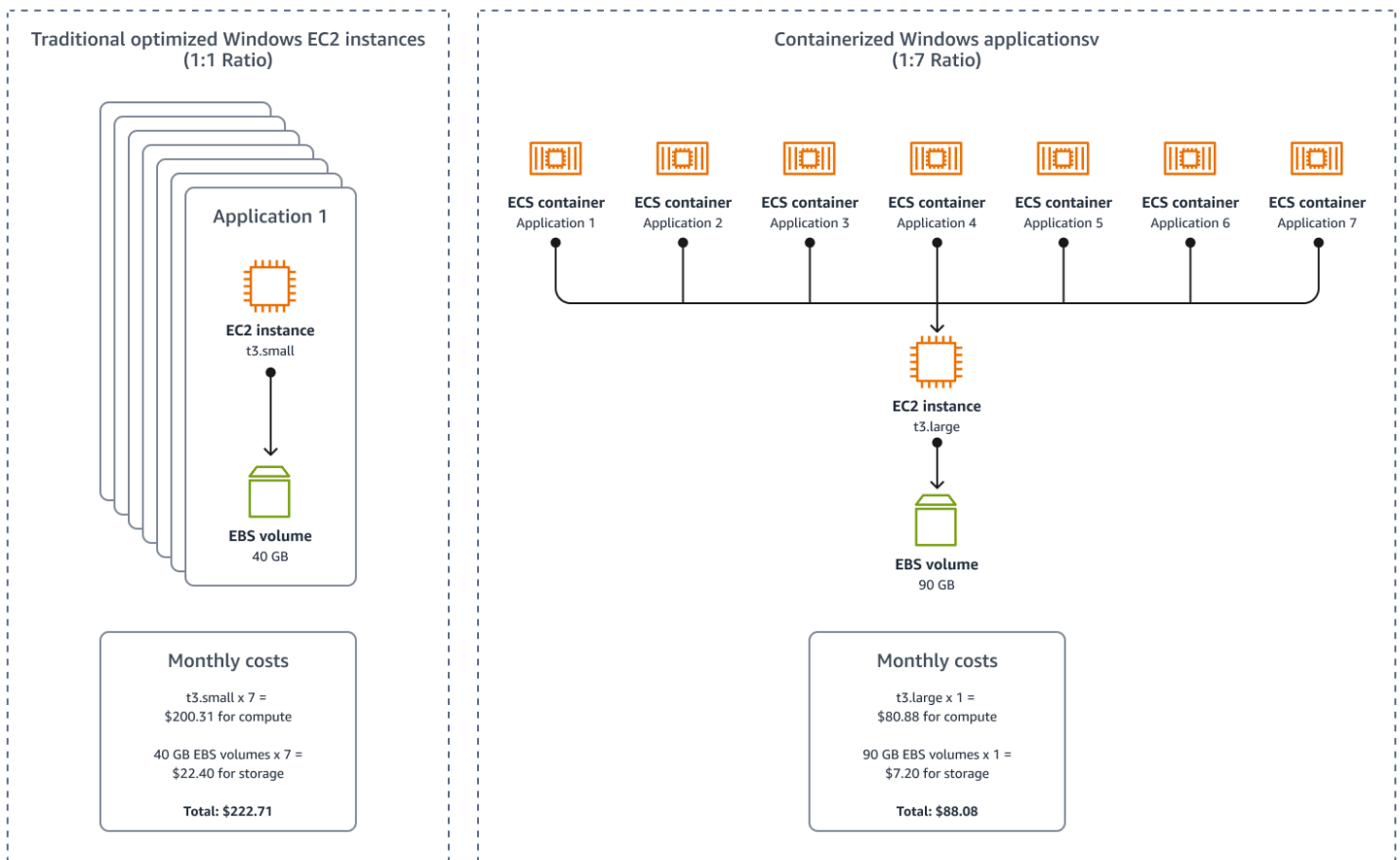
Descripción general de

Según la [encuesta anual de 2021 de la CNCF](#), el 96 % de las organizaciones utilizan o se plantean utilizar contenedores para modernizar su infraestructura. Esto se debe a que los contenedores pueden ayudar a su organización a reducir riesgos, mejorar la eficiencia operativa y acelerar los procesos, además de aportar agilidad. También puede usar contenedores para reducir el costo asociado a la puesta en marcha de sus aplicaciones. En esta sección se ofrecen recomendaciones para ejecutar contenedores de forma rentable en todos los servicios de AWS contenedores, incluidos [Amazon Elastic Container Service \(Amazon ECS\)](#), [Amazon Elastic Kubernetes Service \(Amazon EKS\)](#) y [AWS Fargate](#)

Beneficios en términos de costos

La siguiente infografía muestra los ahorros de costos que una empresa puede lograr al consolidar sus aplicaciones de ASP.NET Framework en instancias de Amazon Elastic Compute Cloud (Amazon EC2) según [AWS una recomendación de evaluación de optimización y licencias \(OLA\)](#).AWS

La siguiente infografía muestra los ahorros adicionales que se pueden lograr al transferir las aplicaciones a un contenedor de Windows.



La AWS OLA recomendó que la empresa realizara una transición a instancias t3.small individuales. La empresa podría lograr estos ahorros poniendo en marcha siete aplicaciones de ASP.NET en servidores en las instalaciones, como se muestra en el siguiente análisis de utilización del rendimiento.

Server name	Storage	Operating system	On-premises CPU AVG utilization	On-premises CPU peak utilization	On-premises RAM (GB)	On-premises RAM AVG utilization (GB)	On-premises RAM peak utilization (GB)	Instance size	vCPU	RAM (GB)
1 AppServer01	60	Windows Server 2012	7.00%	17.00%	8	13.50%	17.10%	t3.small	2	2
2 AppServer02	39	Windows Server 2012	20.07%	22.00%	16	7.50%	12.40%	t3.small	2	2
3 AppServer03	39	Windows Server 2012	24.00%	25.50%	16	8.80%	11.90%	t3.small	2	2
4 AppServer04	4	Windows Server 2012	21.40%	24.00%	16	7.80%	10.70%	t3.small	2	2
5 AppServer05	40	Windows Server 2012	21.30%	23.00%	16	8.20%	12.00%	t3.small	2	2
6 AppServer06	39	Windows Server 2012	21.50%	23.50%	16	7.90%	10.90%	t3.small	2	2
7 AppServer07	39	Windows Server 2012	21.60%	22.90%	16	8.40%	11.50%	t3.small	2	2

Un análisis más detallado reveló que la empresa podría ahorrar aún más en costos poniendo en marcha sus cargas de trabajo en contenedores. Los contenedores disminuyen la carga que el sistema operativo ejerce sobre recursos como la CPU, la RAM y el almacenamiento en disco (se explica en la siguiente sección). En este escenario, la empresa podría unificar las siete aplicaciones

en una sola instancia t3.large y aún disponer de 3 GB de RAM libres. La migración a contenedores puede ayudar a la empresa a lograr un ahorro medio del 64 % en costos de computación y almacenamiento mediante el uso de contenedores en lugar de Amazon EC2.

Recomendaciones de optimización de costos

En la siguiente sección, se ofrecen recomendaciones para optimizar los costos mediante la unificación de aplicaciones y el uso de contenedores.

Reducción del espacio de Windows en Amazon EC2

Los contenedores de Windows pueden reducir el espacio que ocupa Windows en Amazon EC2, pues le permite unificar más aplicaciones en menos instancias de EC2. Suponga, por ejemplo, que tiene 500 aplicaciones ASP.NET. Si pone en marcha un núcleo por aplicación para Windows en Amazon EC2, eso equivale a 500 instancias de Windows (t3.small). Si asume una proporción de 1:7 (que puede aumentar considerablemente según el tipo y el tamaño de la instancia de EC2) por utilizar contenedores de Windows (con t3.large), solo necesitará aproximadamente 71 instancias de Windows. Esto representa una disminución del 85,8 % en el espacio ocupado por Windows en Amazon EC2.

Reducción de los costos asociados a las licencias de Windows

Si licencia una instancia de Windows, no tiene que licenciar los contenedores que se ponen en marcha en esa instancia. Como resultado, la unificación de las aplicaciones de ASP.NET mediante contenedores de Windows puede reducir considerablemente los costos asociados a las licencias de Windows.

Reducción del espacio de almacenamiento

Cada vez que lanza una nueva instancia de EC2, crea y paga un nuevo volumen de Amazon Elastic Block Store (Amazon EBS) para alojar el sistema operativo. Cuando este se escala, el costo también aumenta. Si usa contenedores, puede reducir los costos de almacenamiento porque todos los contenedores comparten el mismo sistema operativo base. Además, los contenedores utilizan el concepto de capas para reutilizar partes inmutables de una imagen de contenedor en todos los contenedores en marcha basados en esa imagen. En el ejemplo anterior, todos los contenedores usan .NET Framework y, por lo tanto, todos comparten la capa intermedia e inmutable de ASP.NET Framework.

Migre los end-of-support servidores a contenedores

El soporte para Windows Server 2012 y Windows Server 2012 R2 finalizó el 10 de octubre de 2023. Puede migrar las aplicaciones que se ponen en marcha en Windows Server 2012 o versiones anteriores agrupándolas en contenedores para que funcionen en nuevos sistemas operativos. De esta forma, evita poner en marcha sus aplicaciones en sistemas operativos que no cumplen con los requisitos y, al mismo tiempo, aprovecha la rentabilidad, la reducción de riesgos, la eficiencia operativa, la velocidad y la agilidad que ofrecen los contenedores.

Una advertencia a tener en cuenta a la hora de adoptar este enfoque es si la aplicación requiere requisitos específicos APIs relacionados con la versión del sistema operativo que se utiliza actualmente (COM Interop, por ejemplo). En este caso, debe probar a actualizar la aplicación a una versión más reciente de Windows. Los contenedores de Windows adaptan su imagen de contenedor base (por ejemplo, Windows Server 2019) a la del sistema operativo del host del contenedor (por ejemplo, Windows Server 2019). Las pruebas y la transferencia a contenedores pueden facilitar las actualizaciones del sistema operativo en el futuro cambiando la imagen base de su Dockerfile e implementándola en un conjunto nuevo de hosts que usen la última versión de Windows.

Eliminación de las licencias y las herramientas de administración de terceros

La administración de su flota de servidores requiere el uso de varias herramientas de operación del sistema de terceros para administrar los parches y la configuración. Esto puede hacer que la administración de la infraestructura sea compleja y, a menudo, incurra en costos asociados al uso de licencias de terceros. Si utilizas contenedores AWS, no necesitas administrar nada desde el punto de vista del sistema operativo. El tiempo de ejecución de los contenedores administra los contenedores. Esto significa que el host subyacente es efímero y se puede reemplazar fácilmente. Puede poner en marcha sus contenedores sin necesidad de administrar directamente el host del contenedor. Además, puedes usar herramientas gratuitas, como acceder fácilmente AWS Systems Manager Session Manager al host y solucionar problemas.

Mejora del control y de la portabilidad

Los contenedores le proporcionan un mayor control de los recursos del servidor, como la CPU y la RAM, que el que obtiene al trabajar directamente con instancias de EC2. En el caso de las instancias de EC2, puede controlar la CPU y la RAM seleccionando una familia de instancias, un tipo de instancia y las [opciones de CPU](#). Sin embargo, con los contenedores, puede definir exactamente la cantidad de CPU o RAM que desea asignar a un contenedor en las definiciones de tareas de ECS o a los [pods de Amazon EKS](#). De hecho, le recomendamos que [especifique la memoria y la CPU de](#)

[los contenedores](#) para los contenedores de Windows. Este nivel de granularidad aporta beneficios en términos de costos. Observe el siguiente código de ejemplo:

```
json
{
  "taskDefinitionArn": "arn:aws:ecs:us-east-1:123456789012:task-definition/demo-
service:1",
  "containerDefinitions": [
    {
      "name": "demo-service",
      "image": "mcr.microsoft.com/dotnet/framework/samples:aspnetapp-
windowsservercore-ltsc2019",
      "cpu": 512,
      "memory": 512,
      "links": [],
      "portMappings": [
        {
          "containerPort": 80,
          "hostPort": 0,
          "protocol": "tcp"
        }
      ],
    }
  ],
}
```

Aceleración de la innovación

El uso de contenedores facilita la automatización de las fases del ciclo de vida del desarrollo, que incluyen la creación, las pruebas y la implementación de aplicaciones. Si automatiza estos procesos, sus equipos de desarrollo y operaciones dispondrán de más tiempo para centrarse en la innovación.

Reducción del TCO

El uso de contenedores suele reducir la dependencia que se tiene de las herramientas de administración de licencias y protección de puntos de conexión. Como los contenedores son unidades de computación efímeras, puede automatizar y simplificar las tareas de administración, como la aplicación de parches, el escalado y las copias de seguridad y la restauración. Esto puede reducir el costo total de propiedad de la administración y del funcionamiento de las cargas de trabajo basadas en contenedores. Los contenedores son más eficientes en comparación que las máquinas virtuales porque permiten maximizar la ubicación de las aplicaciones para aumentar la utilización de los recursos de infraestructura de dichas aplicaciones.

Cierre de brecha de habilidades

AWS ofrece programas y jornadas de inmersión para capacitar a los equipos de desarrollo de clientes en materia de contenedores y DevOps tecnología. Esto incluye capacitación y asesoramiento prácticos.

Refactorización a .NET 5+ y uso contenedores de Linux

Si bien puede reducir los costos transfiriendo sus aplicaciones de .NET Framework a contenedores, puede ahorrar aún más si refactoriza las aplicaciones de .NET heredadas para convertirlas en alternativas nativas en la nube en AWS.

Eliminación de los costos asociados al uso de licencias

La refactorización de su aplicación de .NET Framework en Windows a .NET Core en Linux supone un ahorro de costos de aproximadamente un 45 %.

Acceso a las mejoras más recientes

Al refactorizar sus aplicaciones de .NET Framework en Windows a .NET Core en Linux, tendrá acceso a las mejoras más recientes, como Graviton2. Graviton2 ofrece una relación rendimiento-precio un 40 % superior que las instancias comparables.

Mejora de la seguridad y del rendimiento

La refactorización de la aplicación de .NET Framework en Windows a .NET Core en contenedores de Linux mejora la seguridad y el rendimiento. Esto se debe a que obtiene los parches de seguridad más recientes, se beneficia del aislamiento de los contenedores y tiene acceso a nuevas características.

Utilice contenedores de Windows en lugar de poner en marcha muchas aplicaciones en una sola instancia de IIS

Tenga en cuenta las siguientes ventajas del uso de contenedores de Windows en lugar de poner en marcha varias aplicaciones en una instancia de EC2 de Windows con Internet Information Services (IIS):

- **Seguridad:** los contenedores proporcionan un nivel de seguridad incorporado que no se puede lograr mediante el aislamiento de IIS. Si un sitio web o aplicación de IIS se ve comprometido, todos los demás sitios alojados quedan expuestos y se vuelven vulnerables. La fuga de contenedores

es poco frecuente y es una vulnerabilidad más difícil de explotar que hacerse con el control de un servidor a través de una vulnerabilidad web.

- **Flexibilidad:** la capacidad de poner en marcha contenedores de forma aislada durante el proceso y disponer de su propia instancia se traduce en opciones de red más detalladas. Los contenedores también ofrecen métodos de distribución complejos en muchas instancias de EC2. Estas ventajas no se obtienen cuando se unifican las aplicaciones en una sola instancia de IIS.
- **Sobrecarga administrativa:** la indicación del nombre del servidor (SNI) genera una sobrecarga que requiere administración y automatización. Además, debe lidiar con las operaciones de administración habituales del sistema operativo, como la aplicación de parches, la solución de problemas de BSOD (si no se implementa el escalado automático), la protección de puntos de conexión, etc. Configurar sitios de IIS de acuerdo con las [prácticas recomendadas en materia de seguridad](#) es una actividad continua y laboriosa. Es posible que incluso tenga que configurar [niveles de confianza](#), lo que también aumenta la sobrecarga administrativa. Los contenedores están diseñados para no tener estado y ser inmutables. En última instancia, sus implementaciones son más rápidas, seguras y repetibles si en su lugar utiliza contenedores de Windows.

Siguientes pasos

Invertir en una infraestructura moderna para ejecutar sus cargas de trabajo heredadas aporta enormes beneficios a su organización. AWS Los servicios de contenedores facilitan la administración de la infraestructura subyacente, ya sea en las instalaciones o en la nube, para que pueda centrarse en la innovación y en las necesidades de su empresa. Casi el 80 por ciento de todos los contenedores de la nube funcionan AWS actualmente. AWS proporciona un amplio conjunto de servicios de contenedores para prácticamente todos los casos de uso. Para empezar, consulte [Contenedores en AWS](#).

Recursos adicionales

- [Optimice los costes de las cargas de trabajo en contenedores con los proveedores de capacidad de ECS y las instancias puntuales de EC2 \(blog\)](#)AWS
- [Cost Optimization Checklist for Amazon ECS and AWS Fargate](#) (Blog de AWS)
- [Amazon EKS en AWS Graviton2 está disponible de forma general: consideraciones sobre aplicaciones de arquitectura múltiple](#) (Blog)AWS
- [Optimización de costos para Kubernetes](#) en (Blog) AWSAWS
- [Optimización de los costes informáticos de Kubernetes](#) con la consolidación de Karpenter (Blog)AWS

Optimice los costos de AWS Fargate las tareas en Amazon ECS

Descripción general de

El dimensionamiento correcto de AWS Fargate las tareas es un paso importante para la optimización de costos. Con demasiada frecuencia, las aplicaciones se crean asignando tamaños arbitrarios a las tareas de Fargate y nunca se revisan. Esto puede provocar un aprovisionamiento excesivo de las tareas de Fargate y gastos innecesarios. En esta sección, se explica cómo usar [AWS Compute Optimizer](#) a fin de obtener recomendaciones prácticas para optimizar la CPU y la memoria de las tareas para los servicios de Amazon Elastic Container Service (Amazon ECS) que se ponen en marcha en Fargate. Compute Optimizer también cuantifica el impacto en los costos derivados de la adopción de estas recomendaciones. Esto le permite priorizar sus esfuerzos de optimización en función del tamaño de la oportunidad de ahorro. Las recomendaciones de Compute Optimizer proporcionan configuraciones de CPU y memoria por contenedor para reducir el tamaño de las tareas.

Beneficios en términos de costos

El dimensionamiento correcto de las tareas de Amazon ECS en Fargate puede reducir los costos entre un 30 % y un 70 % en el caso de las tareas de larga duración. Si no revisa las métricas de rendimiento de las aplicaciones para dimensionar correctamente el tamaño de las tareas, puede acabar aplicando la misma lógica utilizada en las instancias de computación de EC2 en el dimensionamiento de contenedores. Esto da lugar a tareas de Fargate sobredimensionadas que aumentan los costos por recursos inactivos. Puede usar Compute Optimizer para detectar las oportunidades de dimensionamiento adecuadas de forma reactiva. Lo ideal es que el propietario de la aplicación revise las métricas de rendimiento específicas de dicha aplicación y elimine la sobrecarga del sistema operativo para garantizar que el tamaño de la tarea sea el adecuado. Para obtener más información, consulte la sección [Transferencia de aplicaciones de Windows a contenedores](#) de esta guía.

Recomendaciones de optimización de costos

En esta sección, se ofrecen recomendaciones para usar Compute Optimizer para dimensionar correctamente las tareas de Amazon ECS en Fargate.

Como parte del proceso de optimización de costos, le recomendamos que haga lo siguiente:

- Activación de Compute Optimizer

- Consumo de resultados de Compute Optimizer
- Etiquetar las tareas para que tengan el tamaño correcto
- Habilite la etiqueta de asignación de costos para que funcione con las herramientas AWS de facturación
- Implementar recomendaciones para asignar el tamaño correcto
- Revisión del antes y el después de los costos en Explorador de costos

Activación de Compute Optimizer

Puede activar [AWS Compute Optimizer](#) por organización o cuenta única en AWS Organizations. La configuración de toda la organización proporciona informes continuos para las instancias nuevas y existentes de toda la flota y para todas las cuentas de miembros. Esto permite que el dimensionamiento correcto sea una actividad recurrente en lugar de una point-in-time actividad.

Nivel de organización

Para la mayoría de las organizaciones, la forma más eficiente de usar Compute Optimizer es hacerlo a un nivel organizacional. Esto ofrece información de varias cuentas y regiones de la organización y centraliza los datos en un único origen para su revisión. Para activar Compute Optimizer por organización, haga lo siguiente:

1. Inicie sesión en su [cuenta de administración de AWS Organizations](#) con un rol que tenga los [permisos necesarios](#) y seleccione la opción de inclusión para todas las cuentas de esta organización. Su organización debe tener [habilitadas todas las características](#).
2. Después de activar la cuenta de administración, puede iniciar sesión en ella, ver todas las otras cuentas de miembros y consultar sus recomendaciones.

Note

Se recomienda configurar una [cuenta de administración delegada](#) para Compute Optimizer. Esto le permite aplicar el principio de privilegio mínimo, lo que minimiza el acceso a la cuenta de administración de AWS Organizations y, al mismo tiempo, proporciona acceso al servicio de toda la organización.

Nivel de cuenta única

Si quiere trabajar con una cuenta con costos elevados, pero no tiene acceso a AWS Organizations, puede activar Compute Optimizer para esa cuenta y región. Para obtener más información sobre el proceso de suscripción, consulta [Cómo empezar](#) con. AWS Compute Optimizer

Note

Las recomendaciones se actualizan a diario y pueden tardar hasta 12 horas en generarse. Tenga en cuenta que Compute Optimizer requiere 24 horas de métricas en los últimos 14 días para generar recomendaciones para Amazon ECS en Fargate. Para obtener más información, consulte [Requirements for Amazon ECS services on Fargate](#) en la documentación de Compute Optimizer.

Compute Optimizer analiza automáticamente las siguientes métricas de uso de CloudWatch Amazon y Amazon ECS para sus servicios de Amazon ECS en Fargate:

- `CPUUtilization`: el porcentaje de capacidad de la CPU que se utiliza en el servicio.
- `MemoryUtilization`: el porcentaje de memoria que se utiliza en el servicio.


Consumo de resultados de Compute Optimizer

Supongamos que nos centramos en realizar los cambios de tamaño en una sola cuenta y una sola región. En este ejemplo, Compute Optimizer se activa en un nivel organizativo para todas las cuentas. Tenga en cuenta que el dimensionamiento es un proceso disruptivo que, en la mayoría de los casos, los propietarios de las aplicaciones llevan a cabo con precisión durante un periodo de mantenimiento programado que dura varias semanas.

Si accede a Compute Optimizer desde la cuenta de administración de una organización (como se muestra en los pasos siguientes), puede elegir la cuenta que desea investigar. En este ejemplo, una tarea se pone en marcha en una sola cuenta que está sobreaprovisionada en us-east-1. El objetivo es centrarse en cambiar el tamaño al tamaño recomendado para el servicio de Amazon ECS.

1. Abra la [consola de Compute Optimizer](#).
2. En la página Panel, filtre por Resultados=Sobrep provisionados para ver todos los servicios de Amazon ECS en Fargate.

3. Para revisar las recomendaciones detalladas sobre Servicios de ECS sobreaprovisionados en Fargate, desplácese hacia abajo y seleccione Ver recomendaciones.
4. Elija Exportar y guarde el archivo para usarlo en el futuro.

 Note

Para guardar las recomendaciones y poder revisarlas en el futuro, debe tener un bucket de S3 disponible para que Compute Optimizer escriba en cada región. Para obtener más información, consulte [Amazon S3 bucket policy for AWS Compute Optimizer](#) en la documentación de Compute Optimizer.

Para ver las recomendaciones de Compute Optimizer, haga lo siguiente:

1. En la [consola de Compute Optimizer](#), vaya a la página Exportar recomendaciones.
2. En Destino del bucket de S3, elija el bucket de S3.
3. En la sección Exportar filtros, en Tipo de recurso, elija Servicios de ECS en Fargate.
4. En la página Recomendaciones para los servicios de ECS en Fargate, elija uno de los servicios de ECS en Fargate y consulte las recomendaciones de CPU y memoria de Compute Optimizer. Por ejemplo, consulte las recomendaciones de las secciones Comparar la configuración actual con el tamaño recomendado para la tarea y Comparar la configuración actual con el tamaño recomendado para el contenedor.

Para obtener la lista de servicios de ECS para Fargate que tienen que dimensionarse correctamente, haga lo siguiente:

1. Abra la [consola de Amazon S3](#).
2. En el panel de navegación, elija Buckets y, a continuación, seleccione el bucket en el que exportó los resultados.
3. En la pestaña Objetos, seleccione su objeto y elija Descargar.
4. En los resultados descargados, filtre la columna de búsqueda para ver solo los servicios OVER_PROVISIONED de Amazon ECS en Fargate. Ahora verá los servicios de Amazon ECS que tiene que dimensionar correctamente.
5. Guarde las definiciones de las tareas en un editor de texto para usarlas más adelante.

Dimensionamiento correcto de tareas con etiquetas

Etiquetar las cargas de trabajo es una herramienta poderosa para organizar sus recursos en AWS. Puede usar etiquetas para obtener una visibilidad detallada de los costos y redistribuir los gastos. Existen muchos métodos y estrategias para agregar etiquetas a los recursos de AWS con el objetivo de gestionar la redistribución de gastos la automatización. Para obtener más información, consulte el AWS documento técnico [Mejores prácticas](#) para etiquetar los recursos. AWS En el siguiente ejemplo se utiliza [AWS CloudShell](#) para etiquetar todas las tareas que forman parte de cualquier servicio de Amazon ECS dentro de la cuenta de destino y Región de AWS.

```
#!/bin/bash
# Set variables
TAG_KEY="rightsizing"
TAG_VALUE="enabled"
# Get a list of ECS Clusters
ClustersArns=$( aws ecs list-clusters -query 'clusterArns' -output text)
for ClustersArn in $ClustersArns; do
  ServiceArns=$( aws ecs list-services -cluster $ClustersArn -query 'serviceArns' -
output text)
  for ServiceArn in $ServiceArns; do
    TasksArns=$( aws ecs list-tasks -cluster $ClustersArn -service-name $ServiceArn -
query 'taskArns' -output text)
    for TasksArn in $TasksArns; do
      aws ecs tag-resource -resource-arn $TasksArn -tags key=$TAG_KEY,value=$TAG_VALUE
    done
  done
done
```

El siguiente ejemplo de código muestra cómo activar la [propagación de etiquetas](#) a todos los servicios de Amazon ECS.

```
#!/bin/bash
# Set variables
TAG_KEY="rightsizing"
TAG_VALUE="enabled"
# Get a list of ECS Clusters
ClustersArns=$(aws ecs list-clusters --query 'clusterArns' --output text)
for ClustersArn in $ClustersArns; do
  ServiceArns=$(aws ecs list-services --cluster $ClustersArn --query 'serviceArns' --
output text)
  for ServiceArn in $ServiceArns; do
```

```
aws ecs update-service --cluster $ClustersArn --service $ServiceArn --propagate-tags
SERVICE &>/dev/null
aws ecs tag-resource --resource-arn $ServiceArn --tags key=$TAG_KEY,value=$TAG_VALUE
done
done
```

Habilite la etiqueta de asignación de costos para que funcione con las herramientas de facturación AWS

Recomendamos activar la etiqueta de asignación de costos definida por el usuario. Esto permite reconocer y filtrar la etiqueta Rightsizing en las herramientas de AWS facturación (por ejemplo, y). AWS Cost Explorer AWS Cost and Usage Report Si no activa esta opción, la opción de filtrado de etiquetas y los datos no estarán disponibles. Para obtener más información sobre el uso de etiquetas de asignación de costos, consulte [Activating user-defined cost allocation tags](#) en la documentación de Administración de facturación y costos de AWS .

Tras esperar 24 horas, podrá ver la etiqueta en Explorador de costos antes de implementar las recomendaciones de dimensionamiento correcto en la siguiente sección. Para ello, busque la etiqueta Dimensionamiento correcto en Explorador de costos.

Implementar recomendaciones para asignar el tamaño correcto

Compute Optimizer proporcionará recomendaciones de tamaño para las tareas o el contenedor. Para implementar recomendaciones de dimensionamiento correcto, haga lo siguiente.

1. Abra la [consola de Amazon ECS](#).
2. En la barra de navegación, seleccione la Región que contiene la definición de tarea.
3. En el panel de navegación, elija Task Definitions (Definiciones de tareas).
4. En la página Task definitions (Definiciones de tareas), elija la tarea y, a continuación, elija Create new revision (Crear nueva revisión).
5. En la página Create new task definition revision (Crear nueva revisión de definición de tarea), realice cambios. Para actualizar la recomendación de tamaño del contenedor, actualice `cpu` y `memory` en el bloque `containerDefinitions` de la definición de la [tarea de ECS](#). Por ejemplo:

```
"containerDefinitions": [  
  {  
    "name": "your-container-name",  
    "image": "your-image",  
    "cpu": 1024,
```

```
"memory": 2048,  
  }  
],
```

6. Verifique la información y, luego, seleccione Create (Crear).

Para actualizar el servicio de Amazon ECS, haga lo siguiente:

1. Abra la [consola de Amazon ECS](#).
2. En la página Clusters (Clústeres), seleccione el clúster.
3. En la página Cluster overview (Información general del clúster), seleccione el servicio y, a continuación, elija Update (Actualizar).
4. En Task definition (Definición de tareas), elija la familia y la revisión de definiciones de tareas que se va a utilizar.

Para los operadores avanzados, puede utilizar CloudShell para actualizar el servicio Amazon ECS.

Por ejemplo:

```
bash  
#!/bin/bash  
# Set variables  
ClustersName="workshop-cluster"  
ServiceName="lab7-fargate-service"  
TaskDefinition="lab7-fargate-demo:3"  
# update the service  
aws ecs update-service --cluster $ClustersName --service $ServiceName --task-definition  
$TaskDefinition
```

Revisión del antes y el después de los costos

Una vez que haya dimensionado correctamente sus recursos, puede usar Explorador de costos para ver los costos anteriores y posteriores usando la etiqueta Dimensionamiento correcto. Recuerde que puede usar [etiquetas de recursos](#) para hacer un seguimiento de los costos. Al usar varias capas de etiquetas, puede lograr una visibilidad detallada de sus costos. En el ejemplo que se describe en esta guía, la etiqueta Dimensionamiento correcto se utiliza para aplicar una etiqueta genérica a todas las instancias objetivo. A continuación, se utiliza la etiqueta equipo para organizar aún más los recursos. El siguiente paso consiste en ingresar etiquetas de aplicación para ver mejor el impacto en los costos derivados del funcionamiento de una aplicación específica.

Piense en la reducción de costos que se puede lograr al usar la etiqueta Dimensionamiento correcto en una única cuenta. En este ejemplo, los costos operativos pasan de 30,26 USD al día a 7,56 USD al día. Suponiendo que un mes tiene 744 horas, el costo anual antes del dimensionamiento correcto es de 11 044,9 USD. Tras redimensionar, el costo anual se reduce a 2759,4 USD. Esto se traduce en una disminución del 75 % en los costos de computación de esta cuenta. Imagine el impacto que puede tener esto en una organización grande.

Antes de empezar con el proceso de dimensionamiento correcto, tenga en cuenta lo siguiente:

- AWS ofrece muchas opciones para reducir los costos. Esto incluye [AWS OLA](#), en el que se AWS revisan las instancias locales antes de mudarlas a AWS ellas. El AWS OLA también le proporciona recomendaciones sobre el tamaño correcto y orientación sobre licencias.
- Complete todos los procesos de dimensionamiento antes de adquirir [Savings Plans](#). Esto puede ser de ayuda para evitar compra en exceso cuando se contrate Savings Plans.

Siguientes pasos

A continuación, recomendamos los pasos próximos:

1. Revise la situación actual y considere la posibilidad de convertir los volúmenes gp2 de Amazon EBS en volúmenes gp3.
2. Revise [Savings Plans](#)

Recursos adicionales

- [Introducción a Compute Optimizer \(documentación\)](#)AWS
- [Prácticas recomendadas para etiquetar AWS recursos \(documentos técnicos\)](#)AWS
- [Windows Containers está activado AWS](#) (AWS Workshop Studio)

Mejora la visibilidad de sus costos en Amazon EKS.

Descripción general de

Se necesita una visión holística para supervisar de manera efectiva el costo de una implementación de Kubernetes. El único costo fijo y conocido es el del plano de control de Amazon Elastic Kubernetes Service (Amazon EKS). Esto incluye todos los demás componentes que

forman parte de la implementación, desde el computación y el almacenamiento hasta las redes, y se trata de un monto que varía en función de las necesidades de la aplicación.

Puede usar [Kubecost](#) para analizar el costo de su infraestructura de Kubernetes, desde los [espacios de nombres](#) y los [servicios](#) hasta los [pods](#) individuales, y luego mostrar los datos en un panel. Kubecost muestra los costos dentro del clúster, como el cómputo y el almacenamiento, y los costos, out-of-cluster como los depósitos de [Amazon Simple Storage Service \(Amazon S3\)](#) y las instancias del Amazon Relational Database Service ([Amazon RDS](#)). Kubecost recomendará el tamaño correcto en función de estos datos y mostrará las alertas críticas que puedan afectar al sistema. Kubecost se puede [integrar](#) con [AWS Cost and Usage Report](#) para ver los ahorros de [Savings Plans para computación](#), [Instancias reservadas](#) y otros programas de descuentos.

Beneficios en términos de costos

Kubecost proporciona informes y paneles que visualizan el costo de las implementaciones de Amazon EKS. Le permite hacer un desglose de cada uno de los distintos componentes del clúster, como los controladores, los servicios, los nodos, los pods y los volúmenes. Esto le proporciona una visión holística de las aplicaciones que se ponen en marcha en un entorno Amazon EKS. Al activar esta visibilidad, puede seguir las recomendaciones de Kubecost o ver los costos de cada aplicación de forma pormenorizada. El dimensionamiento correcto de un grupo de nodos de Amazon EKS ofrece los mismos ahorros potenciales que las instancias de EC2 estándar. Si puede dimensionar correctamente sus contenedores y nodos, podrá eliminar el exceso de computación tanto del tamaño de la instancia que se necesita para poner en marcha el contenedor como del número de instancias de EC2 requeridas en el grupo de escalado automático.

Recomendaciones de optimización de costos

Para aprovechar Kubecost, le recomendamos que haga lo siguiente:

1. Implementación de Kubecost en el entorno
2. Obtención de un desglose detallado de los costos de las aplicaciones de Windows
3. Dimensionamiento correcto de los nodos del clúster
4. Dimensionamiento correcto de las solicitudes de contenedor
5. Administración de los nodos infrautilizados
6. Corrección de las cargas de trabajo abandonadas
7. Toma de medidas basadas en las recomendaciones
8. Actualización de los nodos autoadministrados

Implementación de Kubecost en el entorno

El [taller Finhack de Amazon EKS](#) le enseña cómo implementar un entorno de Amazon EKS configurado para usar Kubecost en una AWS cuenta propia. Esto le permite obtener experiencia práctica con la tecnología. Si le interesa organizar este taller en su organización, contacte con su equipo de cuentas.

Para implementar Kubecost en su clúster de Amazon EKS con [Helm](#), consulte la publicación en el AWS blog sobre la [colaboración con Kubecost para ofrecer un monitoreo de costos a los clientes de EKS](#). AWS Como alternativa, puede consultar la [documentación oficial de Kubecost](#) para obtener instrucciones sobre cómo instalar y configurar Kubecost. Para obtener información sobre el soporte de Kubecost para los nodos de Windows, consulte [Windows Node Support](#) en la documentación de Kubecost.

Obtención de un desglose detallado de los costos de las aplicaciones de Windows

Si bien puede ahorrar costos de manera significativa usando [instancias de spot de Amazon EC2](#), también puede beneficiarse del hecho de que las cargas de trabajo de Windows suelen tener estado. El uso de instancias de spot depende de la aplicación y le recomendamos que compruebe si son aplicables a su caso de uso.

Para obtener un desglose detallado de los costos de sus aplicaciones de Windows, [inicie sesión en Kubecost](#). En el panel de navegación, elija Ahorro.

Dimensionamiento correcto de los nodos del clúster

En [Kubecost](#), seleccione Ahorro en la barra de navegación y, a continuación, elija Dimensionar correctamente el nodo del clúster.

Imagine que Kubecost indica que el clúster está sobreaprovisionado tanto en términos de vCPU como de RAM. En la tabla siguiente, se muestran los detalles y recomendaciones de Kubecost.

	Actuales	Recomendación: simple	Recomendación: compleja
Recuento total	3462,57 USD al mes	137,24 USD al mes	303,68 USD al mes
Recuentos de nodos	4	5	4
CPU	74 VCPUs	10 VCPUs	8 VCPUs

	Actuales	Recomendación: simple	Recomendación: compleja
RAM	152 GB	20 GB	18 GB
Desglose de instancia	2 c5.xlarge + 2 más	5 t3a.medium	2 c5n.large + 1 más

Como se describe en la publicación del blog de Kubecost, [Find an optimal set of nodes for a Kubernetes cluster](#), la opción simple utiliza un solo grupo de nodos, mientras que la compleja utiliza una estrategia de grupo de varios nodos. El botón Aprenda a adoptar permite cambiar el tamaño del clúster con un solo clic. Requiere la instalación del [controlador de clústeres de Kubecost](#).

Si utiliza [nodos de Windows autoadministrados](#) que no se crearon con [eksctl](#), consulte [Actualización de una pila de nodos de AWS CloudFormation](#). Estas instrucciones le muestran cómo cambiar el tipo de instancia en la plantilla de lanzamiento de Amazon EC2 utilizada por el [grupo de escalado automático](#).

Dimensionamiento correcto de las solicitudes de contenedor

En [Kubecost](#), seleccione Ahorro en la barra de navegación y, a continuación, vaya a la página Solicitar recomendaciones de dimensionamiento correcto. En esta página se muestra la [eficiencia](#) de los pods, las recomendaciones de dimensionamiento correcto y el ahorro de costos previsto. Puede usar el botón Personalizar para filtrar por clúster, nodo, espacio de nombres/controlador, etc.

Supongamos que Kubecost calculó que algunos de sus pods están sobreaprovisionados en términos de CPU y RAM (memoria). A continuación, Kubecost recomienda que defina nuevos valores de CPU y RAM para conseguir el ahorro mensual previsto. Para cambiar los valores de CPU y RAM, debe actualizar el archivo de [manifiesto de la implementación](#).

Administración de los nodos infrautilizados

En [Kubecost](#), seleccione Ahorro en la barra de navegación y, a continuación, elija Administrar nodos infrautilizados.

Imagine que la página muestra que un nodo del clúster está infrautilizado en términos de CPU y RAM (memoria) y, por lo tanto, puede vaciarse y terminarse o redimensionarse. Si elige los nodos que no pasan las comprobaciones de nodos y pods, obtendrá más información sobre los motivos por los que no se pueden vaciar.

Corrección de las cargas de trabajo abandonadas

En [Kubecost](#), seleccione Ahorro en la barra de navegación y, a continuación, elija la página Cargas de trabajo abandonadas. En este ejemplo, filtra por un espacio de nombres denominado windows. En esta página se muestran los pods que no alcanzaron el umbral de tráfico y que se consideran abandonados. Los pods deben enviar o recibir cierta cantidad de tráfico de red durante el periodo definido.

Tras considerar detenidamente el abandono de uno o más pods, puede ahorrar costos reduciendo el número de réplicas, eliminando la implementación, redimensionándola para que consuma menos recursos o notificando al propietario de la aplicación que cree que la implementación está abandonada.

Toma de medidas basadas en las recomendaciones

En la sección sobre el Dimensionamiento correcto de los nodos del clúster, Kubecost analiza el uso de los nodos de trabajo del clúster y hace recomendaciones sobre el tamaño correcto de los nodos para reducir los costos. Hay dos tipos de grupos de nodos que se pueden usar con Amazon EKS: [autoadministrados](#) y [administrados](#).

Actualización de los nodos autoadministrados


Para obtener información sobre la actualización de los nodos autoadministrados, consulte [Actualización de los nodos autoadministrados para un clúster](#) en la documentación de Amazon EKS. Indica que los grupos de nodos creados con `eksctl` no se pueden actualizar y deben migrarse a un nuevo grupo de nodos con la nueva configuración.

Por ejemplo, supongamos que tiene un grupo de nodos de Windows llamado `ng-windows-m5-2xlarge` (que usa una instancia de EC2 `m5.2xlarge`) y desea migrar los pods a un [nuevo grupo de nodos](#) llamado `ng-windows-t3-large` (que está respaldado por una instancia de EC2 `t3.large` para ahorrar costos).

Para migrar a un nuevo grupo de nodos cuando utilice grupos de nodos implementados por `eksctl`, haga lo siguiente:

1. Para buscar el nodo en el que se encuentra actualmente el pod, use el comando `kubectl describe pod <pod_name> -n <namespace>`.
2. Ejecute el comando `kubectl describe node <node_name>`. El resultado muestra que el nodo está en marcha en una instancia `m5.2xlarge`. También coincide con el nombre del grupo de nodos (`ng-windows-m5-2xlarge`).

3. Para cambiar la implementación para usar el grupo de nodos `ng-windows-t3-large`, elimine el grupo de nodos `ng-windows-m5-2xlarge` y ponga en marcha `kubectl describe svc,deploy,pod -n windows`. La implementación comienza a reimplementarse inmediatamente cuando se elimina su grupo de nodos.

 Note

El servicio se interrumpirá cuando elimine el grupo de nodos.

4. Use el comando `kubectl describe svc,deploy,pod -n windows` nuevamente al cabo de unos minutos. El resultado muestra que todos los pods vuelven a tener el estado En ejecución.
5. Para mostrar que los pods están en marcha ahora en el grupo de nodos `ng-windows-t3-large`, vuelva a usar los comandos `kubectl describe pod <pod_name> -n <namespace>` y `kubectl describe node <node_name>`.

Métodos de redimensionamiento alternativos

Este método se aplica a cualquier combinación de grupos de nodos autoadministrados o administrados. La publicación del blog [Seamlessly migrate workloads from EKS self-managed node group to EKS-managed node groups](#) proporciona orientación sobre cómo migrar las cargas de trabajo de un grupo de nodos con un tipo de instancia sobredimensionado a un grupo de nodos con el tamaño adecuado sin ningún tiempo de inactividad.

Siguientes pasos

Kubecost facilita la visualización del costo de sus entornos de Amazon EKS. La profunda integración de Kubecost con Kubernetes AWS APIs puede ayudarte a encontrar posibles ahorros de costes. Puedes ver esta información como recomendaciones en el panel Ahorro de Kubecost. Kubecost también puede implementar algunas de estas recomendaciones a con su [característica de controlador de clústeres](#).

Le recomendamos que revise la step-by-step implementación en la entrada del blog Containers [AWS y Kubecost colabore para ofrecer un seguimiento de los costes a los clientes de EKS](#). AWS

Recursos adicionales

- [Amazon EKS Workshop](#) (Amazon EKS Workshop)

- [AWS y Kubecost colaboran para ofrecer un seguimiento de los costes a los clientes de EKS](#) (blog)AWS
- Taller [Amazon EKS Finhack \(AWS Workshop Studio\)](#)
- [Contenedores de Windows activados AWS](#) (AWS Workshop Studio)

Redefinición de la plataforma de las aplicaciones de Windows con App2Container

Descripción general de

[AWS App2Container](#) es una herramienta de línea de comandos para migrar y modernizar aplicaciones web de Java y .NET a contenedores. App2Container analiza y crea un inventario de todas las aplicaciones que se ejecutan en instancias bare metal, máquinas virtuales, instancias de Amazon Elastic Compute Cloud (Amazon EC2) u otros proveedores de servicios en la nube. Seleccione la aplicación que desee almacenar en contenedores. App2Container empaqueta los artefactos y dependencias de la aplicación en imágenes de contenedor, configura los puertos de red y genera los artefactos de implementación necesarios de Amazon Elastic Container Service (Amazon ECS) y Amazon Elastic Kubernetes Service (Amazon EKS), que son plantillas de infraestructura como código (IaC). App2Container aprovisiona la infraestructura en la nube y las canalizaciones de CI/CD necesarias para implementar la aplicación almacenada en un contenedor en un entorno de producción. Para obtener más información, consulte [How App2Container works](#) en la documentación de App2Container.

Con App2Container, puede migrar AWS y modernizar sus aplicaciones como contenedores y, al mismo tiempo, estandarizar la implementación y las operaciones de sus aplicaciones. Puede usar App2Container para crear rápidamente una prueba de concepto (PoC) o acelerar el implementación de cargas de trabajo de producción en contenedores.

Al usar aplicaciones de Windows, debe tener en cuenta varios aspectos. App2Container admite el almacenamiento en contenedores de aplicaciones ASP.NET implementadas en Microsoft Internet Information Services (IIS), lo que incluye las aplicaciones de Windows Communication Foundation (WCF) alojadas en IIS que se ponen en marcha en Windows Server 2016, Windows Server 2019 o Windows Server Core 2004. Para obtener más información, consulte [Supported applications for Windows](#) en la documentación de App2Container. App2Container utiliza Windows Server Core como imagen base para los artefactos del contenedor

adaptando la versión del contenedor de Windows Server Core con la versión del sistema operativo (SO) del servidor en el que se usan los comandos de almacenamiento en contenedores. Esta estrategia desacopla la aplicación del sistema operativo subyacente para que pueda actualizar el SO sin realizar una migración tradicional.

Si utiliza una máquina de trabajo para alojar la aplicación en un contenedor, la imagen base del contenedor, como el canal de servicio a largo plazo (LTSC) de Windows Server 2019, se ajusta al SO de la máquina de trabajo, como Windows Server 2019. Si lleva a cabo el almacenamiento en contenedores directamente en los servidores de aplicaciones, la versión se alinea con el sistema operativo del servidor de aplicaciones. Si sus aplicaciones funcionan en Windows Server 2008 o 2012 R2, puede seguir utilizando App2Container configurando una máquina de trabajo para realizar los pasos de almacenamiento en contenedores e implementación. App2Container no admite aplicaciones que funcionen en sistemas operativos cliente Windows, como Windows 7 o Windows 10. App2Container es compatible con los marcos Tomcat, ToMEE y JBoss (en modo independiente) para procesos de Java. Para obtener más información, consulte [App2Container compatibility](#).

Beneficios en términos de costos

La creación de contenedores y la consolidación de las aplicaciones pueden suponer un [ahorro de cómputo de hasta un 60%](#) en comparación con un patrón de diseño de despliegue de servidores. one-application-to-one App2Container ayuda a agilizar el proceso de almacenamiento en contenedores de las aplicaciones. Las siguientes son algunas de las ventajas de utilizar App2Container para sus procesos de modernización:

- App2Container se ofrece sin cargo adicional.
- App2Container admite múltiples aplicaciones en una imagen de contenedor.
- Utilice App2Container para almacenar sus aplicaciones antiguas de .NET en contenedores si usa un sistema operativo que esté a punto de quedarse sin soporte. Puede pasarse a un sistema operativo más nuevo, evitar tener que pagar por un soporte adicional y reducir los riesgos de seguridad.
- Los contenedores son un método eficaz y rentable de empaquetar sus aplicaciones .NET. Consulte las ventajas de los contenedores en [MACO Recommendation - Moving to containers](#).
- La unificación de aplicaciones y la organización en contenedores ayudan a reducir el consumo de recursos de computación, almacenamiento y licencias, ya que los recursos de computación se usan de forma más eficiente.

- Almacenar en contenedores puede reducir la sobrecarga operativa y los costos asociados a la infraestructura, así como aumentar la portabilidad del desarrollo y la agilidad de la implementación.

Recomendaciones de optimización de costos

Para obtener instrucciones sobre cómo usar App2Container, consulte [Getting started with AWS App2Container](#). Para obtener más información acerca de los comandos de App2Container, consulte [App2Container command reference](#).

Siguientes pasos

App2Container puede acelerar el proceso de almacenamiento de aplicaciones en contenedores y su implementación en Amazon EKS o Amazon ECS. La implementación de aplicaciones en contenedores reduce los costos derivados de la computación, las redes y el almacenamiento, además de reducir la sobrecarga operativa de los equipos que gestionan las aplicaciones.

[Para obtener una experiencia práctica con App2Container, consulte el taller sobre cómo modernizar con él. AWS App2Container](#) Si quieres tener una experiencia de aprendizaje profunda, pide a tu equipo de AWS cuentas que organice una jornada de inmersión en App2Container.

Recursos adicionales

- [Contenerización de aplicaciones Windows complejas de varios niveles mediante](#) (entrada del blog) AWS App2ContainerAWS
- [Contenerización de aplicaciones ASP.NET heredadas mediante](#) (entrada de blog) AWS App2ContainerAWS
- Aplicaciones compatibles con [App2Container](#) (documentación)AWS
- [Modernice con AWS App2Container Workshop \(Workshop Studio\)](#)AWS
- [AWS App2Container FAQs](#)(AWS sitio web)

Almacenamiento

Elegir el almacenamiento adecuado para sus cargas de trabajo de Microsoft es una decisión arquitectónica fundamental. Como parte del proceso de toma de decisiones, le recomendamos que desarrolle un plan de almacenamiento y determine los requisitos funcionales de sus aplicaciones y servicios. En este capítulo se ofrece información general sobre las siguientes opciones de almacenamiento que puede barajar en la planificación.

Secciones:

- [Amazon EBS](#)
- [Amazon FSx](#)
- [AWS Storage Gateway](#)

Amazon EBS

Amazon Elastic Block Store (Amazon EBS) es un servicio de almacenamiento en bloque completamente administrado que ofrece volúmenes de almacenamiento por bloques persistente para utilizarlos con instancias de Amazon Elastic Compute Cloud (Amazon EC2). Puede aprovechar varias características de Amazon EBS para administrar y optimizar eficazmente los recursos de almacenamiento para las cargas de trabajo de Windows en la nube. Por ejemplo, puede usar Amazon EBS para aprovisionar la cantidad exacta de IOPS y el rendimiento que necesita para su carga de trabajo, seleccionar entre una variedad de tipos de volumen que se adapten a los requisitos de sus cargas de trabajo y utilizar herramientas para identificar y eliminar los recursos de almacenamiento desperdiciados. Este control pormenorizado del rendimiento y el uso del almacenamiento le permite optimizar sus recursos de almacenamiento a la par que evita costos innecesarios.

Esta sección abarca los siguientes temas:

- [Migración de volúmenes de Amazon EBS de gp2 a gp3](#)
- [Edición de instantáneas de Amazon EBS](#)
- [Eliminación de volúmenes de Amazon EBS no conectados](#)

Migración de volúmenes de Amazon EBS de gp2 a gp3

Descripción general de

Una unidad de estado sólido (SSD) es la opción de almacenamiento estándar para cargas de trabajo de producción y de alto rendimiento. Amazon EBS ofrece un [volumen SSD de uso general](#) para cargas de trabajo de rendimiento medio a alto. El estándar en muchos Servicios de AWS (incluido Amazon EC2) es [gp2](#), la segunda generación de estos volúmenes SSD de uso general. La tercera generación de dispositivos de uso general SSDs, denominada [gp3](#), se lanzó en diciembre de 2020.

La gama gp3 introdujo mejoras significativas en cuanto a la personalización del rendimiento con respecto a la generación anterior. En el caso de los volúmenes gp2 de Amazon EBS, el rendimiento está estrechamente relacionado con el tamaño del volumen. Por cada 1 GB de capacidad, los volúmenes gp2 obtienen un rendimiento de 3 IOPS. Es decir, un volumen gp2 de 2000 GB es capaz de generar 6000 IOPS. En el caso de los volúmenes gp3, el rendimiento se puede personalizar independientemente de la capacidad de almacenamiento. Esto permite que incluso los volúmenes de pequeña capacidad alcancen capacidades de rendimiento de hasta 80 000 IOPS y un rendimiento de 2000 MB/s.

Otro cambio importante de los volúmenes gp3 es el rendimiento básico de las IOPS. Los volúmenes gp3 ofrecen un rendimiento inicial de 3000 IOPS. En comparación, los volúmenes gp2 deben alcanzar un tamaño de 1 TiB antes de lograr la misma capacidad de rendimiento. En lo que respecta a Windows Server, que normalmente tiene una unidad C: con un tamaño muy inferior a 1 TiB, pasar de gp2 a gp3 supone una mejora significativa del rendimiento.

Por último, el precio de los volúmenes gp3 es una de las mayores mejoras en comparación con los volúmenes gp2. Los volúmenes gp3 ofrecen capacidades de rendimiento mejoradas a un costo un 20 % inferior al de los volúmenes gp2.

Impacto del costo

Dado que es posible escalar el rendimiento de forma independiente de la capacidad, es importante comprender cómo se cobran las IOPS y el rendimiento adicionales. En cuanto a los volúmenes gp2, el precio se basa en la capacidad aprovisionada, con un costo de 0,10 USD por GiB al mes. En el caso de los volúmenes gp3, el precio es similar al de los [volúmenes SSD de IOPS aprovisionadas](#) de alto rendimiento, que tienen un costo por capacidad y otro por IOPS y rendimiento adicionales.

Como se indica en la siguiente tabla, los volúmenes de gp3 tienen un precio de capacidad de 0,08 USD por GiB al mes (un 20 por ciento menos que el de gp2) y unos costes independientes de IOPS

de 0,005 USD por mes aprovisionadas (más de 3000 000) y 0,04 USD por mes aprovisionado de más de 125 USD por rendimiento. MiBs MiBs

	gp3	gp2
Tamaño del volumen	1 GiB — 64 TiB	1 GiB – 16 TiB
IOPS estándar	3000	3 IOPS/GiB (mínimo 100 IOPS) a un máximo de 16 000 IOPS Los volúmenes de menos de 1 TiB pueden alcanzar ráfagas de hasta 3000 IOPS
Máximo de IOPS por volumen	80 000	16,000
Rendimiento de referencia	125 MiBs	El límite de rendimiento está entre 128 y 250 MiBs MiBs, según el tamaño del volumen
Rendimiento máximo por volumen	2.000 MiBs	250 MiBs
Precio	0,08 USD/GiB al mes 3000 IOPS gratuitas y 0,005 USD por IOPS aprovisionada al mes cuando se superen las 3000 125\$ MiBs gratis y 0,04\$ con aprovisionamiento (más de MiBs 125\$ al mes) MiBs	0,10 USD/GiB al mes

⚠ Important

Aunque los volúmenes gp3 tienen costos independientes para la capacidad y el rendimiento, siguen siendo más baratos que los volúmenes gp2 si se configuran con los mismos niveles de rendimiento.

En las siguientes tablas se muestran ejemplos de los ahorros de costos que se pueden lograr al convertir los volúmenes gp2 en gp3 con distintas configuraciones de capacidad y rendimiento.

Configuración de gp2 de ejemplo

Tamaño del volumen (GiB)	Máx. de IOPS	Rendimiento () MiBs	Costo (USD/mes)
30	3 000	128	3,00\$
100	3 000	128	10,00 DÓLARES
500	3 000	250	50,00 DÓLARES
1 000	3 000	250	100,00 DÓLARES
2000	6000	250	200,00 DÓLARES
6000	16 000	250	600,00 DÓLARES

Configuración de gp3 (estándar) de ejemplo

Máx. de IOPS	Rendimiento () MiBs	Costo (USD/mes)	Reducción de costos (en comparación con gp2)
3 000	125	2,40\$	20%
3 000	125	8,00 DÓLARES	20%
3 000	125	40,00 DÓLARES	20%

Máx. de IOPS	Rendimiento () MiBs	Costo (USD/mes)	Reducción de costos (en comparación con gp2)
3 000	125	80,00 DÓLARES	20%
3 000	125	160,00 DÓLARES	20%
3 000	125	480,00 DÓLARES	20%

Configuración de gp3 (igual que gp2) de ejemplo

Máx. de IOPS	Rendimiento () MiBs	Costo (USD/mes)	Reducción de costos (en comparación con gp2)
3 000	128	2,52\$	16%
3 000	128	8,12 DÓLARES	19%
3 000	250	45,00 DÓLARES	10%
3 000	250	85,00 DÓLARES	15%
6000	250	180,00 DÓLARES	10%
16 000	250	550,00 DÓLARES	8%

Para analizar los costos, consulte la sección Calculadora de ahorro de costos de la migración de gp2 a gp3 de EBS en [Recursos de Amazon EBS](#). Puede descargar la calculadora y utilizarla para saber cuánto puede ahorrar gracias a la migración de sus volúmenes gp2 a gp3.

Recomendaciones de optimización de costos

Para obtener instrucciones sobre cómo completar el proceso de migración, consulte la publicación [Migrar sus volúmenes de Amazon EBS de gp2 a gp3 y ahorrar hasta un 20% en costes en el AWS blog sobre almacenamiento](#).

Recursos adicionales

- [Migre sus volúmenes de Amazon EBS de gp2 a gp3 y ahorre hasta un 20% en costes](#) (AWS blog de almacenamiento)
- [Cree una regla AWS Config personalizada para optimizar los tipos de volúmenes de Amazon EBS](#) ([blog de operaciones y migraciones AWS en la nube](#))
- [Controle sus AWS costes mediante la eliminación de los volúmenes de Amazon EBS no utilizados](#) (blog de operaciones y migraciones en la AWS nube)
- [Utilidad de migración de Amazon EBS](#) () GitHub
- [Finding savings from 2020 re:Invent announcements](#) (administración financiera en la nube de AWS)
- [Taller de optimización de costos](#) (AWS Well-Architected Labs)
- [Calculadora de ahorro de costos de la migración de gp2 a gp3 de EBS](#) (descarga)

Edición de instantáneas de Amazon EBS

Descripción general de

Eliminar los volúmenes de EBS y administrar la retención y el archivo de las instantáneas es un aspecto importante para controlar los costos desde el principio. Puede realizar copias de seguridad de los datos de sus volúmenes de EBS en Amazon Simple Storage Service (Amazon S3) realizando instantáneas. point-in-time Las instantáneas son copias de seguridad incrementales, por lo que solo guardan los bloques que cambiaron en el dispositivo después de la instantánea más reciente. Esto disminuye el tiempo necesario para crearlo y ahorra costos de almacenamiento, ya que no se duplican los datos. Cada instantánea contiene toda la información necesaria para restaurar los datos (del momento en que se creó la instantánea) en un volumen de EBS nuevo.

Los cargos por las instantáneas de EBS se calculan por gigabytes consumidos al mes. Se le facturará según el tamaño de la instantánea y el tiempo que la conserve. Los precios varían en función del nivel de almacenamiento. En el caso del [nivel Standard](#), solo se facturarán los bloques modificados que se almacenen. En el caso del nivel Archivo, se facturarán todos los bloques de instantáneas que se almacenen. También se le facturará la recuperación de instantáneas del [nivel Archivo](#). A continuación se presentan algunos ejemplos de cada nivel de almacenamiento:

- Nivel Standard: tiene un volumen que almacena 100 GB de datos. Se le facturarán los 100 GB completos de datos de la primera instantánea (instantánea A). Cuando cree la siguiente

instantánea (instantánea B), dispondrá de 105 GB de datos. En ese caso, solo se le facturarán los 5 GB de almacenamiento adicionales para la instantánea B incremental.

- Nivel Archivo: archiva la instantánea B. La instantánea pasa luego al nivel Archive, y se le facturará el bloque completo de instantáneas de 105 GB.

Puede usar [Amazon Data Lifecycle Manager](#) como ayuda para configurar un ciclo de vida que le permita retener y administrar las instantáneas según lo programado.

Impacto del costo

Los cargos por los volúmenes e instantáneas de EBS se gestionan por separado. Las instantáneas de EBS se facturan a una tarifa inferior a la de los volúmenes de EBS activos. Cuando una instancia finaliza, el valor del [DeleteOnTermination atributo](#) de cada volumen de EBS adjunto determina si se debe conservar o eliminar el volumen. De manera predeterminada, el atributo `DeleteOnTermination` del volumen raíz se establece en `True`. Su valor es `False` para todos los demás tipos de volumen. Esto crea situaciones en las que el operador tiene la intención de eliminar una instancia de EC2, pero deja atrás los volúmenes que se agregaron a la instancia además del volumen raíz. Para obtener instrucciones sobre cómo comprobar los volúmenes (y sus instantáneas asociadas) que ya no necesita, consulte [View information about an Amazon EBS volume](#) en la documentación de Amazon EBS.

De forma predeterminada, al crear una instantánea, se almacena en el nivel Standard de instantáneas de Amazon EBS (nivel estándar). Las instantáneas almacenadas en el nivel estándar son progresivas. Eso significa que solo los bloques del volumen que han cambiado después de la última instantánea se guardan. [Archivo de instantáneas de Amazon EBS](#) es un nuevo nivel de almacenamiento que puede utilizar como almacenamiento a bajo costo y largo plazo de las instantáneas a las que se acceda con poca frecuencia y que no necesiten recuperación rápida o frecuente. La diferencia de precio entre el nivel estándar y el de archivo es significativa y debería ser un factor clave que considerar a la hora de configurar su estrategia de instantáneas. El archivo de instantáneas de Amazon EBS ofrece una reducción de hasta el 75 % menos de los costos de almacenamiento de instantáneas para las instantáneas que planea almacenar durante 90 días o más y a las que rara vez necesita acceder.

Almacenamiento de instantáneas de Amazon EBS	Costo
Standard	0,05 USD/GB al mes

Almacenamiento de instantáneas de Amazon EBS	Costo
Archivado	0,0125 USD/GB al mes

En entornos más pequeños, es posible que el ahorro de costos no sea significativo. Los ahorros son más significativos a gran escala, cuando hay varias cuentas y miles de instancias de EC2 con instantáneas TBs de EBS almacenadas incluso cuando se han eliminado los volúmenes de EBS.

En la siguiente tabla se comparan los niveles estándar y de archivo por mes con un uso de tan solo 50 TB. Incluso a esta escala más baja, se siguen ahorrando miles de dólares al año.

Almacenamiento de instantáneas de Amazon EBS	Costo por mes	Costo por año
Standard (50 TB)	312,50 DÓLARES	3.750 DÓLARES
Archivo (50 TB)	78,13 DÓLARES	937,60 DÓLARES
	Ahorros anuales	2.812,40 DÓLARES

Recomendaciones de optimización de costos

Eliminar una instantánea no reduce necesariamente los costos de almacenamiento de datos de la organización. Es posible que otras instantáneas hagan referencia a los datos de esta, y los datos a los que se hace referencia se conservan siempre. Por ejemplo, al tomar la primera snapshot de un volumen con 10 GiB de datos, el tamaño de la snapshot también es de 10 GiB. Como las instantáneas son incrementales, la segunda instantánea que tome del mismo volumen contiene solamente los bloques de datos que hayan cambiado desde la primera instantánea. Además, la segunda snapshot hace referencia a los datos de la primera snapshot. Si modifica 4 GiB de datos y crea una segunda instantánea, el tamaño de esta es de 4 GiB. Asimismo, la segunda snapshot hace referencia a los 6 GiB que no han cambiado de la primera snapshot. Para obtener más información, consulte [¿Por qué mis costes de almacenamiento no se redujeron después de eliminar una instantánea de mi volumen de EBS y, a continuación, eliminar el propio volumen?](#) en el AWS Knowledge Center.

Considere lo siguiente:

- No se le facturan las instantáneas que otra persona Cuenta de AWS posea y comparta con tu cuenta. Solo se le facturará cuando copie la instantánea compartida en su cuenta. También se le facturarán los volúmenes de EBS que cree a partir de la instantánea compartida.
- Si otra instantánea (instantánea B) hace referencia a una instantánea (instantánea A), es posible que eliminar la instantánea B no reduzca los costos de almacenamiento. Cuando se elimina una instantánea, solo se borran los datos que son únicos de dicha instantánea. Los datos a los que hacen referencia otras instantáneas se conservan y se le facturarán estos datos referenciados. Para eliminar una instantánea incremental, consulte [Incremental snapshot deletion](#) en la documentación de Amazon EBS.

La limpieza de las instantáneas es una práctica operativa estándar cuando se ponen en marcha cargas de trabajo en AWS. Con el tiempo, las instantáneas pueden suponer un gasto significativo por datos que no necesita.

Recursos adicionales

- [Controle sus AWS costes mediante la eliminación de los volúmenes de Amazon EBS no utilizados](#) (blog de operaciones y migraciones en la AWS nube)
- [Delete an Amazon EBS snapshot](#) (documentación de Amazon EBS)
- [Taller de optimización de costos](#) (AWS Well-Architected Labs)
- [Automatically archive Amazon EBS Snapshots with Amazon Data Lifecycle Manager](#) (blog de almacenamiento de AWS)

Eliminación de volúmenes de Amazon EBS no conectados

Descripción general de

Los volúmenes de EBS independientes (huérfanos) pueden generar costes de almacenamiento innecesarios en su entorno. AWS Es esencial incorporar la revisión y la eliminación periódicas de los volúmenes de EBS no utilizados y no utilizados como parte de la higiene del entorno. AWS Se recomienda contar con un proceso para revisar continuamente el uso de los volúmenes de EBS. Puede usar [AWS Compute Optimizer](#) para revisar las instancias que no se usen lo suficiente. Esta sección le permite identificar, administrar y eliminar los volúmenes de EBS que están desconectados o no se usen lo suficiente.

Amazon EBS

[Amazon Elastic Block Store \(Amazon EBS\)](#) es un servicio de bloques que ofrece volúmenes de almacenamiento para instancias de Amazon Elastic Compute Cloud (Amazon EC2). EBS proporciona almacenamiento persistente, con la flexibilidad de conectarse a instancias de EC2 y desconectarse de ellas. Esto significa que el ciclo de vida de los volúmenes de EBS se mantiene incluso si se termina una instancia de EC2. El [DeleteOnTermination](#) atributo es una función que controla si se deben conservar o eliminar los volúmenes de EBS adjuntos al finalizar la instancia. De manera predeterminada, el atributo se establece en `True` para el volumen raíz, lo que provoca la eliminación. Su valor es `False` para otros volúmenes, lo que permite su conservación.

Impacto del costo

Los volúmenes de EBS no conectados, también denominados “volúmenes huérfanos” o “no utilizados”, incurren en los mismos cargos que los volúmenes conectados, según el tamaño y el tipo de almacenamiento aprovisionados. Si bien el costo promedio de los cargos de Amazon EBS puede parecer mínimo, 0,10 USD por GB al mes, es fundamental reconocer que la acumulación de volúmenes de EBS no utilizados puede generar costos significativos con el tiempo.

Por ejemplo, piense en las consecuencias de conservar 50 volúmenes de EBS no utilizados, cada uno aprovisionado con un tamaño de almacenamiento de 100 GB, como se muestra en la siguiente tabla.

Número de volúmenes de almacenamiento	Tipo de volumen	Tamaño	Costo total mensual
50 volúmenes	gp2 (0,10 USD)	100 GB	100 GB x 50 volúmenes de EBS a 0,10 USD por GB/mes = 500,00 USD

El escenario de la tabla anterior supone una reducción de costos de aproximadamente 500 dólares al mes o 6000 dólares al año. Se trata de un paso eficaz para reducir los costos. Asegúrese de incorporar la eliminación de los volúmenes de EBS independientes como práctica habitual en la higiene de su AWS entorno.

Recomendaciones de optimización de costos

Se puede utilizar AWS para automatizar fácilmente la eliminación de volúmenes de EBS no adjuntos. Por ejemplo, puede usar AWS Lambda Amazon CloudWatch y AWS Systems Manager para definir los criterios para eliminar los volúmenes no adjuntos en función de la antigüedad, las etiquetas y otras especificaciones. AWS Config También puede utilizarlos Servicios de AWS para automatizar el proceso de limpieza a escala.

Para evitar consecuencias imprevistas, le recomendamos que actúe con la diligencia debida antes de eliminar los volúmenes de EBS no conectados.

Administración de volúmenes de EBS no conectados

Recomendamos que siga estas prácticas recomendadas:

- Cumpla con los requisitos de cumplimiento: compruebe que la eliminación de los volúmenes de EBS no conectados se adhiera a los requisitos de gobernanza y cumplimiento de su organización.
- Establezca políticas de copia de seguridad y retención de datos: antes de eliminar un volumen de EBS no conectado, haga una copia de seguridad de los datos importantes en otro repositorio de almacenamiento (por ejemplo, [Amazon S3](#)). En el caso de la retención de datos, las [instantáneas de Amazon EBS](#) son una forma más rentable de retener datos que los volúmenes de EBS, y pueden restaurar el volumen si es necesario en el futuro. Para obtener más información sobre la administración eficaz de las instantáneas, consulte la sección [Edición de instantáneas de Amazon EBS](#) de esta guía.
- Compruebe si hay dependencias: compruebe si hay dependencias entre los volúmenes de EBS no conectados y otros recursos de AWS - Puede utilizar [Consola de administración de AWS o una API](#) para recopilar información descriptiva sobre sus volúmenes de EBS, como el tamaño, el estado y los recursos asociados. Este es un paso importante para evitar que se eliminen recursos temporalmente no conectados.
- Cree una política de retención: establezca un periodo de retención para los volúmenes de EBS no conectados. Esto puede ayudarle a identificar el momento adecuado para eliminar los volúmenes no adjuntos, garantizando así que su AWS entorno siga optimizado. Por ejemplo, puede crear una EventBridge regla de [Amazon](#) para iniciar una función Lambda de forma programada. La función Lambda puede usar el AWS SDK para identificar activamente cualquier volumen de EBS no adjunto, aplicar un mecanismo de etiquetado para facilitar el seguimiento y enviar notificaciones cuando un volumen de EBS no adjunto alcance o supere un umbral definido.
- Etiquete volúmenes de EBS no conectados: [etiquetar](#) los volúmenes de EBS es una práctica útil que puede ayudar a organizar e identificar los volúmenes en función de atributos como el entorno,

la aplicación o el propietario. Esto puede resultar especialmente útil a la hora de decidir qué volúmenes no conectados se van a eliminar, ya que permite identificar rápidamente los volúmenes que ya no se necesitan en función de sus etiquetas.

- Garantice una eliminación segura: comprobar cuándo se conectó por última vez un volumen de EBS puede ayudar a determinar si es seguro eliminarlo. Para obtener más información, consulte [¿Cómo uso AWS CLI los comandos para enumerar el historial de adjuntos o separaciones de un volumen específico de Amazon EBS?](#) en el Centro de conocimiento AWS .
- Identifique los volúmenes de EBS infrutilizados: identificar y eliminar los volúmenes de EBS infrutilizados es una práctica muy recomendable para reducir los costos de almacenamiento y mantener un entorno optimizado. AWS AWS Trusted Advisor y [AWS Compute Optimizer](#) puede ayudarlo a identificar los volúmenes de EBS infrutilizados y a ofrecer recomendaciones para reducir los costos y mejorar la eficiencia. Por ejemplo, consulte [Configurar la automatización para optimizar los volúmenes de EBS con AWS Trusted Advisor](#) (GitHub), [Establecer un panel de Trusted Advisor organización \(TAO\) \(AWS Workshop Studio\)](#) y Optimizar los [costes de los volúmenes de Amazon EBS mediante AWS Compute Optimizer](#) (AWS Storage Blog).

Automatización de la limpieza de volúmenes de EBS no conectados

Le recomendamos que considere la posibilidad de utilizar las siguientes herramientas para automatizar la limpieza de volúmenes de EBS no conectados:

- [AWS APIs \(DescribeVolumes\)](#) — Puede filtrar y buscar volúmenes de EBS independientes mediante AWS SDKs o el AWS Command Line Interface (AWS CLI). Puede ahorrar tiempo y esfuerzo automatizando este proceso con un script o una [función de Lambda](#) que se ponga en marcha cuando lo programe. Un [ejemplo de script GitHub muestra](#) cómo funciona esto. El script usa Lambda para analizar los AWS CloudTrail registros e identificar los volúmenes de EBS no conectados.
- [AWS Systems Manager Automatización](#): esto le permite automatizar las tareas rutinarias de mantenimiento y corrección en su infraestructura. Para empezar, [cree un manual de procedimientos de automatización](#) en el que se defina una serie de pasos que deben ponerse en marcha en un orden específico. Por ejemplo, puede crear un manual de procedimientos que cree primero una instantánea del volumen de EBS no conectado y, a continuación, elimine el volumen en sí. Esto puede ser de ayuda para automatizar tareas que, de hacerse manualmente, serían lentas y propensas a errores.
- [AWS Config](#)— Esto le permite evaluar, auditar y realizar un seguimiento de los cambios en sus AWS recursos a lo largo del tiempo. Al capturar los cambios de configuración, puede utilizarlos

AWS Config para evaluar el cumplimiento, la gobernanza y la utilización de los recursos en su entorno. Por ejemplo, AWS Config puede identificar los [volúmenes de EBS no utilizados](#). Además, puede asociar la AWS Systems Manager automatización AWS Config para corregir automáticamente la eliminación de los volúmenes de EBS no utilizados.

Recursos adicionales

- [Elimine los volúmenes no utilizados de Amazon Elastic Block Store \(Amazon EBS\) AWS Config utilizando AWS Systems Manager](#) y AWS (Guía prescriptiva)
- [Controle sus AWS costes mediante la eliminación de los volúmenes de Amazon EBS no utilizados](#) (blog de operaciones y migraciones en la AWS nube)
- [AWSConfigRemediation-DeleteUnusedEBSVolume](#)(Referencia del AWS Systems Manager manual de automatización)

Amazon FSx

Amazon FSx for Windows File Server es un servicio de almacenamiento de archivos totalmente gestionado que está optimizado para las cargas de trabajo de Windows. Le proporciona una solución simple y escalable para poner en marcha sus aplicaciones y cargas de trabajo basadas en Windows, sin necesidad de administrar una infraestructura de almacenamiento compleja. Puede utilizar Windows File Server FSx para aprovisionar y acceder fácilmente al almacenamiento de archivos compartidos que admite sus aplicaciones de Windows de forma nativa, incluidas Microsoft SQL Server SharePoint, Microsoft y aplicaciones.NET personalizadas. Además, FSx para Windows, File Server le ayuda a administrar los costos al ofrecer opciones de precios flexibles, como cuotas de almacenamiento, pay-as-you-go y deduplicación automática de datos para reducir el espacio de almacenamiento y optimizar el rendimiento y los costos.

Esta sección abarca los siguientes temas:

- [Elección del almacenamiento de archivos SMB adecuado](#)
- [Habilite la deduplicación de datos en Amazon FSx](#)
- [Conozca la fragmentación de datos en Windows FSx File Server](#)
- [Comprenda el uso del volumen del disco duro en Amazon FSx](#)
- [Uso de una única zona de disponibilidad](#)

Elección del almacenamiento de archivos SMB adecuado

Descripción general de

AWS ofrece una variedad de servicios de almacenamiento totalmente gestionados que le proporcionan las amplias prestaciones de los servicios de archivos líderes del sector y, al mismo tiempo, combinan las últimas innovaciones de AWS infraestructura y seguridad. Puede incorporar los servicios de AWS en los flujos de trabajo de infraestructura como código (IaC) e integrarlos con los servicios de computación, supervisión y protección de datos de AWS. Para las cargas de trabajo de Windows, puede elegir entre dos servicios de archivos totalmente gestionados que se pueden utilizar para satisfacer las necesidades de sus aplicaciones: FSx para Windows File Server y Amazon FSx para NetApp ONTAP.

FSx para Windows File Server

Amazon FSx for Windows File Server proporciona almacenamiento compartido totalmente gestionado integrado en Windows Server y ofrece una amplia gama de funciones de acceso, gestión y administración de datos. FSx para Windows File Server se integra fácilmente en los entornos de Windows porque es un servicio nativo de Windows. Se recomienda utilizar el servidor de archivos de Windows FSx para compartir usuarios y grupos, las instancias de clúster de conmutación por error Always On para SQL Server, las aplicaciones de Windows y la infraestructura de escritorios virtuales (VDI). FSx para Windows File Server también se integra bien con Amazon FSx File Gateway, Amazon Kendra, los registros de auditoría de Amazon S3 y Amazon Data Firehose.

FSx para ONTAP

FSx for ONTAP se basa en NetApp en el sistema de archivos ONTAP patentado por ONTAP. Requiere cierto nivel de capacitación y se recomienda principalmente a los usuarios locales existentes. NetApp Entre los casos de uso habituales se incluyen los recursos compartidos de usuarios y grupos, las instancias de clúster de conmutación por error Always On para SQL Server y las aplicaciones de Windows. FSx para ONTAP, admite varios protocolos, sistemas de archivos de más de 64 TB (a escala PB sin un servidor de espacio de nombres DFS), la clonación, la replicación, las instantáneas, la compresión (eficiencia del almacenamiento) y la organización inteligente de los datos en niveles.

Impacto del costo

FSx para Windows File Server

FSx para Windows File Server fue la primera solución de almacenamiento compartido AWS para implementar instancias de clúster de conmutación por error para SQL Server. Con FSx el servidor

de archivos de Windows, puede lanzar instancias de clúster de conmutación por error utilizando las licencias de la edición SQL Standard. Sin embargo, esto le impide confiar en los grupos de disponibilidad Always On, que requieren licencias de la edición Enterprise de SQL Server. Al cambiar de la edición Enterprise de SQL Server a la edición Standard de SQL Server, podría ahorrar entre un 65 % y un 75 % en sus [licencias de SQL Server](#).

Puede utilizar FSx for Windows File Server para instancias de clústeres de conmutación por error a fin de reducir la carga de almacenamiento del I/O almacenamiento de EBS típico. Al pasarse I/O a FSx for Windows File Server, podría reducir las instancias EC2, que dependen de un alto rendimiento de Amazon EBS y de IOPS, sin que ello afecte al rendimiento del almacenamiento.

FSx para ONTAP

Puede usar FSx for ONTAP para ejecutar su clúster de conmutación por error de Microsoft en el protocolo de bloques iSCSI y beneficiarse de la inicialización instantánea de archivos de SQL Server, el uso de la replicación entre regiones, la compatibilidad con antivirus y la clonación. SnapMirror Si crea varias copias de bases de datos para realizar pruebas, la clonación puede suponer una diferencia significativa tanto en el consumo de espacio como en la rapidez con la que se pueden crear esas copias de bases de datos. Además, puede utilizar ONTAP NetApp SnapCenter para gestionar las funciones de copia de seguridad, restauración y clonación con sus instancias EC2 para SQL Server. FSx FSx for ONTAP también ofrece una organización automática en niveles desde unidades SSD hasta un almacenamiento agrupado de capacidad de bajo coste, lo que permite combinar rendimiento y rentabilidad.

FSx para ONTAP NetApp es compatible con el sistema de archivos (ONTAP), a diferencia del servidor de archivos FSx de Windows, que admite un sistema de archivos NTFS nativo de Windows. El tamaño mínimo FSx para ONTAP es de 1024 GB, mientras que FSx para Windows File Server puede empezar con tan solo 32 GB.

Integración con el sistema de archivos distribuido de Microsoft

FSx para Windows File Server y FSx ONTAP, se integran con el [sistema de archivos distribuido \(DFS\)](#) de Microsoft para lograr una integración perfecta en las implementaciones existentes. Tenga en cuenta los siguientes aspectos al planificar la arquitectura:

- FSx para Windows File Server y FSx ONTAP, admiten [espacios de nombres DFS \(DFS\) en ambos tipos de implementación](#) (varias zonas de disponibilidad y zonas de disponibilidad únicas).
- Solo FSx para Windows, el servidor de archivos admite la [replicación DFS \(DFSR\)](#) y solo cuando se utilizan zonas de disponibilidad únicas.

Recomendaciones de optimización de costos

Tanto para Windows File Server como FSx para ONTAP, el rendimiento depende en gran medida de la configuración, al igual que su precio. FSx para Windows File Server, el precio depende principalmente de la capacidad y el tipo de almacenamiento, la capacidad de rendimiento, las copias de seguridad y los datos transferidos. Con ONTAP, usted paga FSx por el almacenamiento en SSD, las IOPS de SSD, el uso del conjunto de capacidades, la capacidad de rendimiento y las copias de seguridad.

Servicio de archivos	Costo de 5 TB de almacenamiento	Configuración	Region
FSx para Windows File Server	982,78\$	Una sola zona de disponibilidad SSD (15 000 IOPS) 32 MBps 5 TB de copia de seguridad (sin ahorros en la deduplicación)	Este de EE. UU. (Norte de Virginia)
FSx para ONTAP	979,28 DÓLARES	Una sola zona de disponibilidad SSD al 100 % Nivel de capacidad de lectura y escritura de 15 000 15 000 IOPS de SSD 128 MBps 5 TB de copia de seguridad (sin ahorros en la deduplicación)	Este de EE. UU. (Norte de Virginia)

Tenga en cuenta lo siguiente:

- La deduplicación y la compresión le permiten almacenar más datos en dispositivos físicos mediante la reducción del tamaño de los datos, pero pagará por el almacenamiento en unidades de estado sólido (SSD) o unidades de disco duro (HDD) aprovisionadas.
- Puede utilizar ONTAP FSx para organizar sus datos en niveles. Es extremadamente raro que se pueda acceder al 100 % de los datos con regularidad y que se requiera almacenamiento en SSD. Puede mover los datos fríos y a los que se accede con poca frecuencia a un nivel de capacidad para ahorrar costos.
- Los precios que se mencionan aquí se calculan en base al 100 % de los datos en el nivel SSD y 15 000 IOPS en el nivel SSD.

Copia de seguridad

De forma predeterminada, tanto FSx para ONTAP como FSx para Windows File Server almacenan sus copias de seguridad totalmente gestionadas en Amazon S3. Sin embargo, FSx para ONTAP hay una opción adicional para el uso de copias de seguridad SnapVault, que permite configurar las copias de seguridad para que residan en el nivel de capacidad. La copia de seguridad SnapVault es un mecanismo autogestionado que resulta más rentable que la opción de copia de seguridad totalmente gestionada por defecto. La opción de copia de seguridad completamente administrada cuesta 0,05 USD por GB al mes. El coste de la SnapVault copia de seguridad de ONTAP (unidad SSD FSx por unidad de almacenamiento con capacidad de 10:1) cuesta 0,03221 USD (0,9 x 0,0219+0,1 x 0,125).

Tenga en cuenta lo siguiente:

- AWS Las copias de seguridad gestionadas ofrecen una granularidad de una hora. [SnapVault](#) permite trabajar en tan solo cinco minutos.
- Puede usar NetApp las herramientas (como la CLI y la API) para configurar SnapVault las relaciones y la replicación de instantáneas.
- Habilite la política de all estratificación en un SnapVault volumen para usar el nivel de capacidad como almacenamiento para los datos de respaldo.
- SnapVault los destinos pueden estar en el mismo lugar Región de AWS, entre regiones o locales. Esto se suele realizar en un destino de copia de seguridad del sistema de archivos en una única o en varias zonas de disponibilidad. En comparación, AWS Backup está respaldado por la resiliencia regional de Amazon S3.

Dimensionamiento correcto

También puede ahorrar costos y aprovechar al máximo su sistema de archivos si dimensiona correctamente y evita el aprovisionamiento excesivo.

Para dimensionar correctamente, haga lo siguiente:

1. Identifique sus necesidades actuales en función de los datos. En las cargas de trabajo típicas de Windows, puede utilizar herramientas integradas del sistema operativo, como [Monitor de rendimiento](#).
2. En Monitor de rendimiento, utilice los siguientes contadores para evaluar sus necesidades de rendimiento actuales. El intervalo de captura es de un segundo, con un tamaño de registro máximo de 1000 MB y la sobrescritura activada.

```
Logman.exe create counter PerfLog-Short -o "c:\perflogs\PerfLog-Long.blg" -f bincirc  
-v mmddhhmm -max 1024 -c "\\LogicalDisk(*)\*" "\\Memory\*" "\.NET CLR Memory(*)\*"  
"\Cache\*" "\\Network Interface(*)\*" "\\Paging File(*)\*" "\\PhysicalDisk(*)\*"  
"\Processor(*)\*" "\\Processor Information(*)\*" "\\Process(*)\*" "\\Thread(*)\*"  
"\Redirector\*" "\\Server\*" "\\System\*" "\\Server Work Queues(*)\*" "\\Terminal  
Services\*" -si 00:00:01
```

3. Para iniciar la captura de registros, use el comando `logman start PerfLog-Short`. Para detener la captura de registros, use el comando `logman stop PerfLog-Short`.

Note

Puede encontrar los archivos de registro de rendimiento en `:\perflogs`, en el servidor que pone en marcha la captura. Para obtener más información, consulte [Windows Performance Monitor Overview](#) en la documentación de Microsoft.

4. Tras identificar la configuración correcta, comprueba si tu estimación es correcta en el sistema de FSx archivos de Amazon mediante herramientas de stress de disco como Microsoft [DISKSPD](#).
5. Si está conforme con el rendimiento, pase al recurso compartido de archivos.

Recomendamos una estrategia conservadora en cuanto a la capacidad de almacenamiento, ya que solo se puede escalar verticalmente. La capacidad de rendimiento se puede escalar o reducir verticalmente según sea necesario.

Recursos adicionales

- [Amazon FSx for NetApp ONTAP FAQs](#) (AWS sitio web)
- [Optimización del rendimiento del servidor de archivos de Amazon FSx para Windows con nuevas métricas](#) (blog sobre AWS almacenamiento)

Habilite la deduplicación de datos en Amazon FSx

Descripción general de

La deduplicación de datos es una característica que le permite almacenar sus datos de manera más eficiente y con menos requisitos de capacidad. Implica buscar y eliminar la duplicación en los datos sin comprometer su fidelidad ni integridad. La deduplicación de datos utiliza la fragmentación y la compresión de archivos secundarios de tamaño variable, lo que ofrece ratios de optimización de 2:1 para servidores de archivos generales y de hasta 20:1 para datos de virtualización. La deduplicación de datos es mucho más eficaz que la compresión NTFS. La arquitectura de deduplicación incorpora resiliencia ante errores de hardware, pues cada dato y metadato se somete a una validación completa mediante sumas de verificación. Esta validación incluye comprobar la redundancia de los metadatos y de los fragmentos de datos a los que se accede con mayor frecuencia.

FSx para Windows File Server es totalmente compatible con la deduplicación de datos. Su uso puede suponer un ahorro medio del 50 % al 60 % en el caso de los recursos compartidos de archivos de uso general. En los recursos compartidos, los ahorros oscilan entre el 30 % y el 50 % en los documentos de usuario y entre el 70 % y el 80 % en los conjuntos de datos de desarrollo de software. Es importante comprender que el ahorro de almacenamiento que puede lograr con la deduplicación de datos depende de la naturaleza del conjunto de datos, incluida la cantidad de duplicados que existan entre los archivos. La deduplicación no es una buena opción si los datos almacenados son de naturaleza dinámica.

Impacto del costo

Para hacer frente al crecimiento del almacenamiento de datos en la empresa, los administradores unifican los servidores y hacen que el escalado de la capacidad y la optimización de los datos sean objetivos clave. La configuración predeterminada de la deduplicación de datos puede suponer un ahorro inmediato. Los administradores también pueden refinar la configuración para obtener beneficios adicionales. Por ejemplo, puede configurar la deduplicación para que se ejecute solo en determinados tipos de archivos o puede crear un cronograma de trabajo personalizado.

En términos generales, la deduplicación tiene tres tipos de tareas: optimización, recopilación de elementos no utilizados y limpieza. Tenga en cuenta que no se liberará espacio hasta que ponga en marcha un recopilación de elementos no utilizados después de la optimización. Puede programar el trabajo o ponerlo en marcha manualmente. Todos los ajustes disponibles al programar un trabajo de deduplicación de datos también están disponibles al iniciar un trabajo manualmente (excepto en el caso de los trabajos específicos de la programación).

A pesar de que la deduplicación solo supone un 25 por ciento de ahorro efectivo, Windows File Server supone un importante ahorro FSx de costes. Estos ahorros previstos se basan en una [estimación](#) de Calculadora de precios de AWS.

Recomendaciones de optimización de costos

La deduplicación activada FSx para los sistemas de archivos del servidor de archivos de Windows no está habilitada de forma predeterminada. Para habilitar la deduplicación mediante la [administración remota activada PowerShell](#), debe ejecutar el `Enable-FSxDedup` comando y, a continuación, usarlo para establecer la configuración. `Set-FSxDedupConfiguration` Para obtener más información, consulte [Administrar sistemas de archivos](#) en la documentación del servidor FSx de archivos de Windows.

Para activar la deduplicación, use el comando siguiente:

```
PS C:\Users\Admin> Invoke-Command -ComputerName amznfsxxxxxxx.corp.example.com -  
ConfigurationName FSxRemoteAdmin -ScriptBlock {Enable-FsxDedup }
```

Para comprobar la configuración de la deduplicación, use el comando siguiente:

```
Invoke-Command -ComputerName amznfsxxxxxxx.corp.example.com -ConfigurationName  
FSxRemoteAdmin -ScriptBlock {  
Set-FSxDedupSchedule -Name "CustomOptimization" -Type Optimization -Days  
Mon,Tues,Wed,Sat -Start 09:00 -DurationHours 7  
}
```

Al ejecutar el PowerShell `Measure-DedupFileMetadata` cmdlet, puede determinar cuánto espacio potencial de disco se puede recuperar en un volumen si elimina un grupo de carpetas, una sola carpeta o un solo archivo y, a continuación, ejecuta un trabajo de recolección de elementos no utilizados. En concreto, el valor `DedupDistinctSize` indica cuánto espacio se recupera si se eliminan esos archivos. Los archivos suelen tener fragmentos que se comparten en otras carpetas, por lo que el motor de deduplicación calcula qué fragmentos son únicos y se eliminarían tras un trabajo de recopilación de elementos no utilizados.

Las [programaciones predeterminadas de los trabajos de deduplicación de datos](#) están diseñadas para funcionar bien con las cargas de trabajo recomendadas y ser lo menos intrusivas posible (esto no incluye el trabajo de optimización de prioridades que se activa para el tipo de uso de copia de seguridad). Si las cargas de trabajo requieren recursos de gran tamaño, le recomendamos programar los trabajos para que se pongan en marcha únicamente durante las horas de inactividad o para reducir o aumentar la cantidad de recursos del sistema que puede consumir un trabajo de deduplicación de datos.

De forma predeterminada, la deduplicación de datos utiliza el 25 % de la memoria disponible. Sin embargo, esto se puede aumentar con `-memory switch`. En el caso de los trabajos de optimización, le recomendamos que establezca un rango entre 15 y 50. En el caso de los trabajos programados, puede utilizar un mayor consumo de memoria. Por ejemplo, con los trabajos de recopilación de elementos no utilizados y limpieza (que normalmente se programan para que se pongan en marcha fuera del horario laboral), puede establecer un consumo de memoria más alto (por ejemplo, 50).

Para obtener información adicional sobre la configuración de deduplicación de datos, consulte [Reducir los costos de almacenamiento mediante la deduplicación de datos](#) en la documentación del servidor de archivos FSx de Windows.

Recursos adicionales

- [Understanding Data Deduplication](#) (documentación de Microsoft)
- [Reducir los costos de almacenamiento mediante la deduplicación de datos](#) (FSx para ver la documentación del servidor de archivos de Windows)

Conozca la fragmentación de datos en Windows FSx File Server

Descripción general de

FSx para Windows File Server, el rendimiento depende de la configuración. Este se basa principalmente en el tipo de almacenamiento, la capacidad de almacenamiento y la configuración del rendimiento. La capacidad de rendimiento que seleccione determina los recursos de rendimiento disponibles para el servidor de archivos, incluidos I/O los límites de la red, la CPU y la memoria y los I/O límites de disco impuestos por el servidor de archivos. La capacidad y el tipo de almacenamiento que seleccione determinan los recursos de rendimiento disponibles para los volúmenes de almacenamiento, es decir, los I/O límites de disco impuestos por los discos de almacenamiento. Además del rendimiento, las opciones de configuración también influyen en el costo. FSx para

Windows File Server, el precio depende principalmente de la capacidad y el tipo de almacenamiento, la capacidad de rendimiento, las copias de seguridad y los datos transferidos.

Cuando el almacenamiento de archivos es relativamente grande y se requieren altos niveles de rendimiento, la partición de datos puede ser útil. La partición de datos implica [dividir los datos de los archivos](#) en conjuntos de datos más pequeños y almacenarlos en diferentes sistemas de archivos. Las aplicaciones que acceden a los datos desde varias instancias pueden alcanzar niveles de rendimiento altos al leer y escribir en estos fragmentos en paralelo. Al mismo tiempo, aún puede presentar una vista unificada en un espacio de nombres común para las aplicaciones. Además, esto también puede ayudar a escalar el almacenamiento de datos de archivos (hasta cientos de petabytes) más allá de lo que admite cada sistema de archivos (64 TB) en el caso de conjuntos de datos de archivos grandes.

Impacto del costo

En el caso de conjuntos de datos de gran tamaño, suele ser más eficaz implementar varios sistemas de archivos pequeños FSx para Windows File Server, en lugar de una unidad SSD grande para lograr el mismo nivel de rendimiento. El uso de una combinación FSx de los tipos de almacenamiento en disco duro y SSD del servidor de archivos de Windows permite ahorrar costes y adaptar la carga de trabajo al mejor subsistema de disco subyacente. En las siguientes tablas, puede ver la diferencia entre un único sistema de archivos de 17 TB y varios sistemas de archivos más pequeños que suman la misma capacidad.

Sistema de archivos de gran tamaño en SSD con múltiples cargas de trabajo

Nombre del servidor	Costo	Configuración	Region
Servidor FSx de archivos Amazon para Windows	5716 USD	SSD de 17 TB Deduplicación del 30 % 256 Mbps Copia de seguridad de 17 TB	Este de EE. UU. (Norte de Virginia)

Carga de trabajo particionada mediante DFSN

Nombre del servidor	Costo	Configuración	Region	Share
Servidor FSx de archivos Amazon para Windows	1024 USD	SSD de 2 TB Deduplicación del 20 % 128 Mbps Copia de seguridad de 2 TB Multi-AZ	Este de EE. UU. (Norte de Virginia)	Recurso compartido 1
Servidor FSx de archivos Amazon para Windows	2132 USD	SSD de 5 TB Deduplicación del 30 % 256 Mbps Copia de seguridad de 5 TB Multi-AZ	Este de EE. UU. (Norte de Virginia)	Recurso compartido 2
Servidor FSx de archivos Amazon para Windows	1036 USD	DISCO DURO DE 10 TB Deduplicación del 40 % 128 Mbps Copia de seguridad de 10 TB	Este de EE. UU. (Norte de Virginia)	Recurso compartido 3

Nombre del servidor	Costo	Configuración	Region	Share
		Multi-AZ		
Instancias de EC2 de Windows (DFSN)	27 USD	t3a.medium 2 v CPUs Memoria de 4 GiB	Este de EE. UU. (Norte de Virginia)	Instancias (DFSN)

El costo anual de un sistema de archivos de gran tamaño en SSD es de 68 592 USD. El costo anual de una carga de trabajo particionada es de 50 640 USD. En este ejemplo, se puede lograr un ahorro del 26 % y, al mismo tiempo, adaptar la carga de trabajo al almacenamiento en backend adecuado. Para obtener más información sobre la estimación de precios, consulte la estimación de [Calculadora de precios de AWS](#).

Recomendaciones de optimización de costos

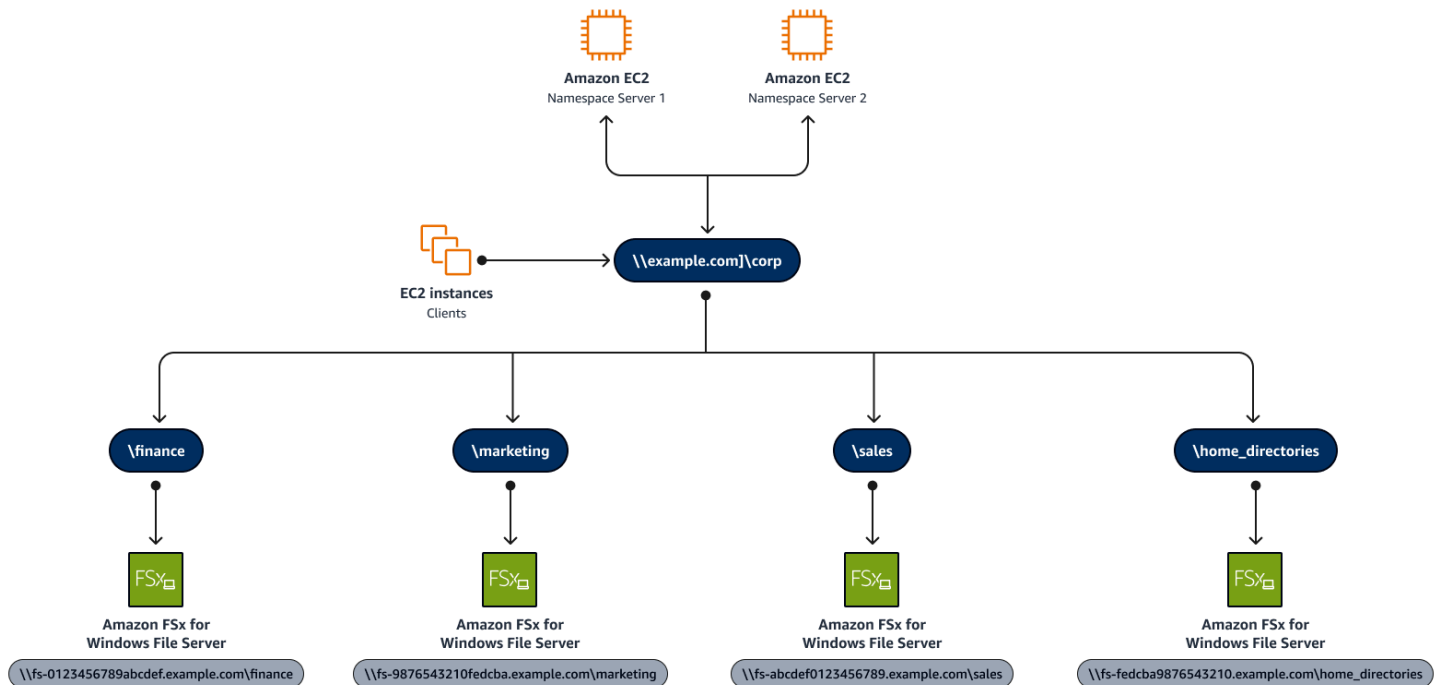
Para implementar una solución de deduplicación de datos, debe configurar un espacio de [nombres DFS de Microsoft](#) en función del tipo de datos, el I/O tamaño y el patrón de acceso. I/O Cada espacio de nombres admite hasta 50 000 recursos compartidos de archivos y cientos de petabytes de capacidad de almacenamiento en total.

Lo más eficiente es elegir una partición que distribuya las operaciones de E/S de manera uniforme entre todos los sistemas de archivos que vaya a utilizar. Supervisar la carga de trabajo le permitirá lograr una optimización adicional o reducir los costos. Si necesita ayuda para medir la información de rendimiento del sistema de FSx archivos de Amazon, consulte el [rendimiento del servidor FSx de archivos de Windows en la documentación del servidor](#) FSx de archivos de Windows.

Tras elegir una estrategia de partición, puede agrupar los sistemas de archivos para acceder fácilmente a sus recursos compartidos mediante espacios de nombres DFS. Esto permite a los usuarios ver un sistema de archivos homogéneo, cuando en realidad están accediendo a varios sistemas de archivos diferentes con casos de uso específicos. Es importante crear los recursos compartidos con una convención de nomenclatura adecuada para que los usuarios finales puedan descifrar fácilmente la carga de trabajo para la que están diseñados los recursos compartidos. También es importante etiquetar los recursos compartidos de producción y los que no son de

producción, de modo que los usuarios finales no coloquen por error los archivos en el sistema de archivos incorrecto.

El siguiente diagrama muestra cómo se puede utilizar un único espacio de nombres DFS como punto de acceso para varios sistemas de archivos de Amazon FSx .



Tenga en cuenta lo siguiente:

- Puede añadir los recursos compartidos existentes FSx para el servidor de archivos de Windows a un árbol de DFS.
- No se puede añadir Amazon FSx a la raíz de la ruta de acceso compartido de DFS. Tiene una única subcarpeta.
- Debe implementar una instancia de EC2 para alojar la configuración del espacio de nombres DFS.

Para obtener más información sobre la configuración de DFS-N, consulte [DFS Namespaces overview](#) en la documentación de Microsoft. Para obtener más información sobre el uso de los espacios de nombres DFS, consulte el vídeo [Uso de los espacios de nombres DFS con Amazon FSx para Windows File Server](#) sobre YouTube.

Recursos adicionales

- [Agrupación de varios sistemas de archivos con espacios de nombres DFS](#) (documentación de Amazon) FSx

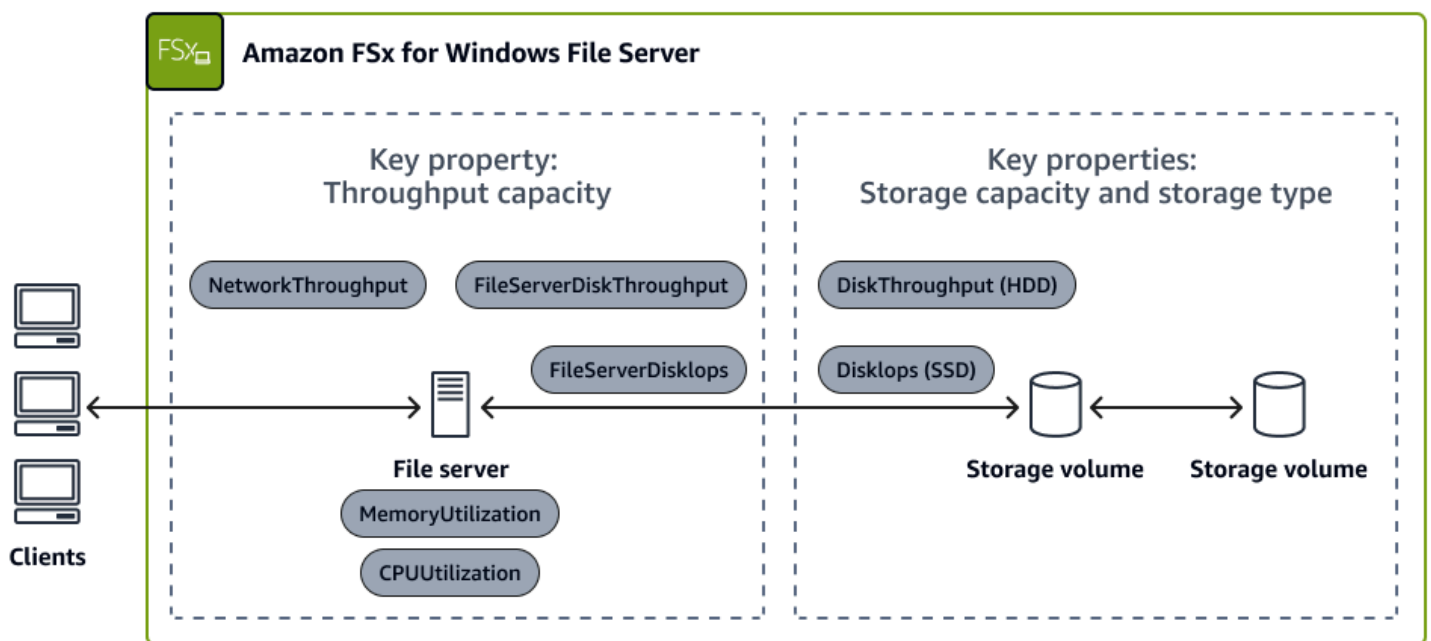
- [Tutorial 6: Ampliar el rendimiento con fragmentos](#) (documentación de Amazon FSx)
- [Uso de espacios de nombres DFS con Amazon FSx for Windows File Server](#) (Labs)AWS

Comprenda el uso del volumen del disco duro en Amazon FSx

Descripción general de

Amazon FSx para Windows File Server ofrece la flexibilidad de elegir el rendimiento independientemente de la capacidad del sistema de archivos. Hay dos configuraciones de capacidad disponibles: unidad de disco duro (HDD) y unidad de estado sólido (SSD).

En el siguiente diagrama se muestra la relación entre el rendimiento y la configuración de almacenamiento.



Con el almacenamiento basado en disco duro, recibe una base de 12 IOPS con 80 IOPS de disco en ráfaga (por IOPs TiB de almacenamiento) y un rendimiento de 12 IOPS de referencia Megabytes/second con 80 ráfagas (Megabytes/second por TiB de almacenamiento). Por ejemplo, si su recurso compartido tiene un tamaño de 50 TB, obtendrá $50 * 12 = 600$ inicialmente tanto para el rendimiento como para las IOPS.

El servidor de archivos de Amazon FSx para Windows proporciona 80 ráfagas de IOPS. Los créditos de ampliación se recargan automáticamente cuando la utilización está por debajo de la tasa de referencia y se consumen automáticamente cuando la utilización está por encima de dicha tasa.

Por ejemplo, si su carga de trabajo solo utiliza 10 IOPS/TB durante una hora (2 por IOPS/TB below your baseline rate), you can then utilize 14 IOPS/TB (2 IOPS/TB encima de la línea base) durante la siguiente hora antes de volver a quedarse sin créditos de ráfaga.

Para las operaciones de archivos, Amazon FSx for Windows File Server proporciona latencias consistentes de menos de milisegundos con almacenamiento SSD y latencias de milisegundos de un solo dígito con almacenamiento en disco duro. Para todos los sistemas de archivos, incluidos aquellos con almacenamiento en disco duro, Amazon FSx para Windows File Server proporciona una caché rápida (en memoria) en el servidor de archivos, de modo que puedes obtener un alto rendimiento y latencias de menos de milisegundos para los datos a los que se accede activamente, independientemente del tipo de almacenamiento.

Cuando corresponda, el uso del almacenamiento en HDD puede ayudar a reducir el costo de la capacidad total de almacenamiento y proporcionar una plataforma de almacenamiento fiable para sus necesidades.

Impacto del costo

El rendimiento del servidor de archivos de Amazon FSx para Windows depende de tres factores: la capacidad de almacenamiento, el tipo de almacenamiento y el rendimiento. El I/O rendimiento de la red y el tamaño de la caché en memoria vienen determinados únicamente por la capacidad de rendimiento, mientras que el I/O rendimiento del disco viene determinado por una combinación de la capacidad de rendimiento, el tipo de almacenamiento y la capacidad de almacenamiento.

Si bien el SSD se recomienda para cargas de trabajo I/O intensivas, hay una variedad de cargas de trabajo cuyas necesidades se pueden satisfacer con las especificaciones de rendimiento del disco duro. El almacenamiento en el disco duro está diseñado para una amplia gama de cargas de trabajo, incluidos los directorios principales, los recursos compartidos de usuarios y departamentos y los sistemas de administración de contenido. Por ejemplo, si sus usuarios solo necesitan un acceso de baja latencia a los datos necesarios para los proyectos actuales, el acceso a la mayoría de los datos almacenados será poco frecuente.

Puede utilizar [Calculadora de precios de AWS](#) para comparar una unidad SSD de 20 TB con un sistema de archivos HDD en us-east-1. Como se muestra en la siguiente tabla, aunque no se ahorre en la deduplicación, la diferencia de costos es significativa cuando se comparan los sistemas de archivos HDD con los sistemas de archivos SSD.

Configuración FSx del sistema de archivos de Amazon	Costos mensuales
SSD Multi-AZ de 20 TB (us-east-1)	4.699,30\$
HDD Multi-AZ de 20 TB (us-east-1)	542,88 DÓLARES
Ahorros mensuales estimados	4.156,42\$

Note

FSx Para obtener más ahorros en Windows File Server, consulte la FSx sección [Habilitar la deduplicación de datos en Amazon](#) de esta guía.

Si identifica correctamente sus necesidades de rendimiento, puede seleccionar el almacenamiento adecuado para la carga de trabajo y reducir los costos.

Recomendaciones de optimización de costos

Si decide utilizar el almacenamiento en HDD, pruebe el sistema de archivos para asegurarse de que cumple sus requisitos de rendimiento. El almacenamiento en HDD tiene un costo menor en comparación con el almacenamiento en SSD, pero con niveles más bajos de rendimiento del disco e IOPS de disco por unidad de almacenamiento. Puede ser adecuado para recursos compartidos de usuarios de uso general y directorios principales con I/O requisitos bajos, sistemas de administración de contenido de gran tamaño en los que los datos se recuperan con poca frecuencia o conjuntos de datos con un número reducido de archivos grandes.

No se puede cambiar el tipo de almacenamiento de un sistema de archivos ya existente. Para convertir el tipo de almacenamiento de un sistema de archivos de Amazon FSx para Windows File Server, debe realizar una copia de seguridad del sistema de archivos existente y restaurarlo en un nuevo sistema de archivos con el tipo de almacenamiento deseado. Si desea convertir un sistema de archivos SSD existente en un sistema de archivos HDD, tenga en cuenta que el almacenamiento en HDD tiene una capacidad mínima mucho mayor (2 TB).

Para restaurar una copia de seguridad con un tipo de almacenamiento diferente, haga lo siguiente:

1. [Haga una copia de seguridad del sistema de archivos existente.](#)

2. [Cree un nuevo sistema de FSx archivos de Amazon](#) con el tipo de almacenamiento HDD.
3. Restaure la copia de seguridad en el nuevo sistema de archivos con el tipo de almacenamiento deseado.
4. Compruebe que el nuevo sistema de archivos tenga el tipo de almacenamiento correcto y que sus datos estén intactos.

Antes de pasar los cambios a producción, te recomendamos que analices el rendimiento de tu sistema de FSx archivos de Amazon y compruebes que el cambio es aceptable. Para obtener más información, consulta la publicación [Optimización del rendimiento del servidor de archivos de Amazon FSx para Windows con nuevas métricas](#) en el blog sobre AWS almacenamiento.

Recursos adicionales

- [Optimización de costes con Amazon FSx](#) (FSx documentación de Amazon)

Uso de una única zona de disponibilidad

Descripción general de

En esta sección se explica cuándo es más beneficioso utilizar una implementación de zona de disponibilidad única de [Amazon FSx for Windows File Server](#). Abarca situaciones en las que el traslado a una única zona de disponibilidad reduce los costes y, al mismo tiempo, le permite utilizar Amazon FSx for Windows File Server como su servicio de almacenamiento de archivos gestionado. Le recomendamos que implemente una única zona de disponibilidad para Amazon FSx para las cargas de trabajo de producción. Esto puede ser útil para garantizar la redundancia de varias zonas de disponibilidad.

Impacto del costo

Un sistema de archivos en una única zona de disponibilidad ofrece una reducción de costos de aproximadamente un 40 % en comparación con una implementación en varias zonas de disponibilidad. Con un sistema de archivos en varias zonas de disponibilidad, paga 0,230 USD por GB al mes en SSD y 0,025 USD por GB al mes en HDD, frente a los 0,130 USD por GB al mes en SSD y 0,013 USD por GB al mes en HDD con un sistema de archivos en una única zona de disponibilidad. Puede ver una comparación de los costos y crear sus propias estimaciones con la [Calculadora de precios de AWS](#).

En el caso de un sistema de archivos de 10 TB, esto puede suponer la diferencia entre pagar aproximadamente 1200 USD al mes si se usan varias zonas de disponibilidad o 680 USD al mes si se usa una sola zona de disponibilidad. En este [ejemplo](#), se utiliza un sistema de archivos de 10 TB FSx para Windows File Server con SSD. Se calcula un ahorro del 50 % si se deduplica. A grandes rasgos, una única zona de disponibilidad tiene un costo de entrada más bajo, pero hay ciertas limitaciones, como verá en la siguiente sección.

Recomendaciones de optimización de costos

Implementaciones en una única zona de disponibilidad

Para asegurarse de que una única zona de disponibilidad es la adecuada, tenga en cuenta la propia SLAs para los datos que se almacenan en el FSx servidor de archivos de Windows. Esto implica comprender si tiene que proporcionárselo SLAs a sus clientes (internos y externos) y si los tres niveles de disponibilidad de una FSx sola zona de disponibilidad de Amazon le permitirán cumplirlos SLAs. FSx en el caso de Windows, el servidor de archivos con una única zona de disponibilidad sigue teniendo un tiempo de actividad del 99,9 por ciento. El SLA de Amazon FSx para múltiples zonas de disponibilidad es superior al 99,99 por ciento. En el caso de las cargas de trabajo críticas, le recomendamos que utilice varias zonas de disponibilidad en lugar de una única zona de disponibilidad, incluso si esto supone un costo adicional.

Las implementaciones en una única zona de disponibilidad son ideales para cargas de trabajo como las copias de seguridad de bases de datos de SQL Server. Pueden proporcionar almacenamiento de bajo costo con un nivel HDD y, al mismo tiempo, proporcionarle un tiempo de actividad constante. Si necesita mayor disponibilidad para una carga de trabajo de producción, como servidores SQL de alta disponibilidad o acceso a aplicaciones de producción, una única zona de disponibilidad no es la opción adecuada para sus cargas de trabajo. En el caso de las copias de seguridad, las pruebas no relacionadas con la producción y los entornos de desarrollo, la implementación de una FSx única zona de disponibilidad de Amazon puede reducir los costes operativos.

Un caso de uso en el que un sistema de archivos de zona de disponibilidad FSx única de Amazon funciona bien es en una situación de producción en la que se utilizan varios sistemas de archivos de zona de disponibilidad FSx única de Amazon, como almacenamiento por servidor en un clúster de SQL Server de alta disponibilidad mediante grupos de disponibilidad Always On. Para obtener más información, consulte la sección [Optimización de los costes de las implementaciones de SQL Server de alta disponibilidad publicada en el AWS blog sobre AWS almacenamiento](#).

Replicación multirregional

Una posible opción para reducir los costos con un único sistema de archivos de zona de disponibilidad (uno en el que solo funcione un sistema de archivos de zona de disponibilidad) es si desea aprovechar la replicación multirregional con Amazon FSx. Puede implementar [sistemas de archivos en una única zona de disponibilidad](#) que admitan el uso con la característica DFS-R nativa de Microsoft. DFS-R tiene la capacidad de replicar automáticamente los datos entre regiones y varios sitios. Para obtener más información sobre la configuración del DFS-R mediante Amazon FSx, consulte Uso de la [replicación del sistema de archivos distribuido de Microsoft](#) en la documentación de Amazon FSx.

Otra alternativa para ahorrar costes en varias regiones es utilizar. AWS Storage Gateway Esto le permite implementar un [Amazon FSx File Gateway](#) en otra región para el acceso multirregional de Amazon FSx. Para más información, consulte la sección [AWS Storage Gateway](#) de esta guía.

Si usa varias regiones, debe tener en cuenta el costo de transferencia de datos del tráfico de datos entre regiones. El tráfico que fluye entre regiones conlleva un cargo de 0,02 USD por GB. Por lo tanto, si sus datos cambian de forma constante en grandes volúmenes, esto aumentará el costo total. Por [ejemplo](#), 1 TB de datos transferidos equivale aproximadamente a 20,48 USD.

período de mantenimiento

El período de mantenimiento es una consideración clave si utilizas una zona de disponibilidad única con Amazon FSx. Durante el período de mantenimiento, el sistema de FSx archivos de Amazon no estará disponible durante unos 20 minutos, debido a la aplicación rutinaria de parches de software para el servidor Windows subyacente. Si utilizas el sistema de archivos para realizar copias de seguridad nocturnas, ajusta el período de FSx mantenimiento de Amazon en consecuencia para evitar interrupciones durante la copia de seguridad. Puedes ajustar la [ventana de mantenimiento](#) después de crear tu sistema de FSx archivos de Amazon.

Recursos adicionales

- [Disponibilidad y durabilidad: sistemas de archivos Single-AZ y Multi-AZ](#) (documentación de Amazon FSx)
- [Precios de servidores de archivos de Amazon FSx para Windows](#) (AWS sitio web)

AWS Storage Gateway

AWS Storage Gateway es un servicio de almacenamiento en la nube híbrida que conecta los entornos locales con el almacenamiento en AWS la nube. Le permite integrar sin problemas su infraestructura local existente AWS, lo que le permite almacenar y recuperar datos de la nube y ejecutar aplicaciones en un entorno híbrido. En el caso de las cargas de trabajo de Windows, puede usar Storage Gateway para almacenar los datos y acceder a ellos mediante protocolos nativos de Windows, como SMB y NFS. Puede usar Storage Gateway para reducir los costos asociados con la ejecución de cargas de trabajo de Windows AWS mediante el uso de hardware y software locales como puente hacia la nube. Esto le permite aprovechar la escalabilidad y la rentabilidad que supone AWS no tener que realizar cambios significativos en su infraestructura existente.

Bajo el paraguas de Storage Gateway, encontrará Amazon S3 File Gateway, Amazon FSx File Gateway, Tape Gateway y Volume Gateway. S3 File Gateway y FSx File Gateway se utilizan con mayor frecuencia con las cargas de trabajo de Microsoft.

Puerta de enlace de archivo de Amazon S3

[Amazon S3 File Gateway](#) le permite almacenar sus archivos en Amazon S3 y, al mismo tiempo, proporciona acceso a sus usuarios mediante el uso de recursos compartidos SMB tradicionales. Esto proporciona una interfaz de usuario familiar y permite reducir los costos almacenando los datos en Amazon S3 y aprovechando los distintos niveles de almacenamiento de Amazon S3. Puede implementar Storage Gateway con S3 Intelligent Tiering para transferir automáticamente los archivos del ciclo de vida a los niveles de almacenamiento de menor costo para reducir ahorrar aún más. Recomendamos S3 File Gateway si busca escalabilidad horizontal, acceso de solo lectura, lecturas rápidas y repetidas (desde la memoria caché) y volcados de bases de datos. Por lo general, no se recomienda para escrituras, edición de archivos o recursos compartidos departamentales de alto rendimiento o alta disponibilidad.

Amazon FSx File Gateway

[Amazon FSx File Gateway](#) también puede ofrecer un ahorro de costes al trabajar con los sistemas de archivos de Amazon FSx Windows. Puede instalar un FSx File Gateway para proporcionar acceso localizado a un sistema de FSx archivos de Amazon en otra región para evitar los costes de tener dos sistemas de archivos independientes. Esto también puede resultar útil si tiene varios servidores de archivos en las instalaciones y desea unificarlos para evitar tener que pagar por varios dispositivos de hardware.

Impacto del costo

Amazon S3 File Gateway

Configurar S3 File Gateway es fácil porque puede usar el asistente de inicialización de Storage Gateway. Puede implementar la puerta de enlace en cuestión de minutos mediante una instancia EC2 en su AWS entorno. Una vez configurada la puerta de enlace, puede configurar los recursos compartidos de Storage Gateway para que sean accesibles a través de los protocolos SMB y NFS. En el caso de las cargas de trabajo típicas de Windows, también puede usar esta configuración para aprovechar un entorno de Active Directory y establecer permisos en los recursos compartidos de archivos. Puede integrar eficazmente Storage Gateway en su uso habitual, ya que funcionará como un recurso compartido de archivos típico de Windows. Los archivos y las carpetas se almacenan como objetos y las listas de control de acceso NTFS (ACLs) como metadatos.

En la siguiente tabla se comparan los costos de 10 TB de almacenamiento con las tres opciones de almacenamiento disponibles:

- FSx para el servidor de archivos de Windows
- Puerta de enlace de archivo de Amazon S3
- Amazon Elastic Block Store (Amazon EBS)

El precio de tener 10 TB de almacenamiento es considerablemente más económico si utiliza Amazon S3, ya que puede particionar los datos en varios niveles de uso. En la estimación de precios, se utiliza S3 Intelligent Tiering por su flexibilidad. Esto incluye el 80 % en S3 Standard, el 10 % en Infrequent Access y el 10 % en Amazon Glacier. Si bien puede utilizar Amazon Glacier, es importante establecer las reglas de ciclo de vida adecuadas para garantizar que no sea necesario acceder inmediatamente a los archivos que se trasladen a Amazon Glacier. Amazon Glacier es exclusivamente para uso de archivos, no para uso de acceso normal.

Sistemas de almacenamiento	Costo de 10 TB de almacenamiento	Region
FSx para Windows File Server (suponiendo un ahorro del 50% en la deduplicación)	683,20 USD (SSD)	Este de EE. UU. (Norte de Virginia)

Sistemas de almacenamiento	Costo de 10 TB de almacenamiento	Region
Amazon S3 File Gateway	449,51 USD (Intelligent Tiering)	Este de EE. UU. (Norte de Virginia)
Amazon EBS	1.335,69 USD GP3	Este de EE. UU. (Norte de Virginia)

Considere lo siguiente:

- En Amazon Glacier, recibe I/O errores genéricos a menos que utilice la [RestoreObjectAPI](#) para restaurar el objeto en Amazon S3. Te recomendamos que utilices una notificación para este I/O error mediante Amazon CloudWatch Events. De esta forma, su equipo de operaciones podrá reaccionar si un usuario recibe este error en un archivo al que podría tener que acceder. Para obtener más información sobre estos errores, consulte [Error: InaccessibleStorageClass](#) en la documentación de Amazon S3 File Gateway.
- Además de la limitación de acceso de Amazon Glacier, [solo se ACLs permiten 10 por objeto o carpeta](#) en Storage Gateway. Antes de decidir usar Storage Gateway, asegúrese de no necesitar más de 10 entradas de ACL.

Amazon FSx File Gateway

Al igual que una puerta de enlace de archivos de Amazon S3, una puerta de enlace de FSx archivos proporciona acceso a un sistema de archivos que conserva los datos a largo plazo. En Amazon S3 File Gateway, los datos residen en Amazon S3. En el caso de FSx File Gateway, los datos residen en FSx el servidor de archivos de Windows. Si bien las opciones Multi-AZ están disponibles FSx para Windows File Server, no hay ninguna opción multirregional. Si tiene una empresa global o una oficina remota, es posible que necesite una plataforma de almacenamiento compartido que esté geográficamente más cerca del usuario final para evitar la latencia. Si tuviera que implementar otro sistema de FSx archivos de Amazon, esto añadiría el coste de un sistema de archivos Amazon FSx for Windows File Server completamente nuevo y el almacenamiento necesario. Para evitar la creación de un sistema de archivos completamente nuevo y la duplicación de costes, puede implementar FSx File Gateway en la región secundaria. Esto proporciona a los usuarios acceso localizado a los archivos y, al mismo tiempo, permite reducir los costos generales.

Sistemas de almacenamiento	Costo de 10 TB de almacenamiento	Region
Servidor FSx de archivos Amazon para Windows	683,20 USD (SSD)	Este de EE. UU. (Norte de Virginia)
Amazon FSx File Gateway	503,70 USD (puerta de enlace única)	Este de EE. UU. (Norte de Virginia)

Note

Los precios de la tabla anterior se basan en los [precios de Storage Gateway](#).

Tenga en cuenta lo siguiente:

- FSx File Gateway puede ayudarle a ahorrar aproximadamente 180\$ al mes (o 2100\$ al año) en cargas de trabajo multirregionales.
- Los gastos de transferencia de datos son mucho más bajos con FSx File Gateway, ya que solo necesita almacenar en caché los archivos a los que se accede con regularidad y no una copia secundaria completa.
- Aunque puede tener dos implementaciones de FSx Windows File Server en distintas regiones y mantenerlas actualizadas con AWS Backup o AWS DataSync, ninguna de las dos opciones es prácticamente en tiempo real.

Recomendaciones de optimización de costos

Amazon S3 File Gateway

S3 File Gateway ofrece una opción económica para almacenar archivos, pero hay que tener en cuenta algunos aspectos relacionados con la forma en que implementa y utiliza el sistema de archivos. Por ejemplo, S3 File Gateway requiere el uso de una máquina virtual para poner en marcha el software de Storage Gateway. En AWS, Storage Gateway se implementa en Amazon EC2 mediante una instancia m5.xlarge, de forma predeterminada. Si desea reducir los costos de almacenamiento local, puede implementar Storage Gateway como un dispositivo virtual en plataformas de virtualización como VMware Hyper-V.

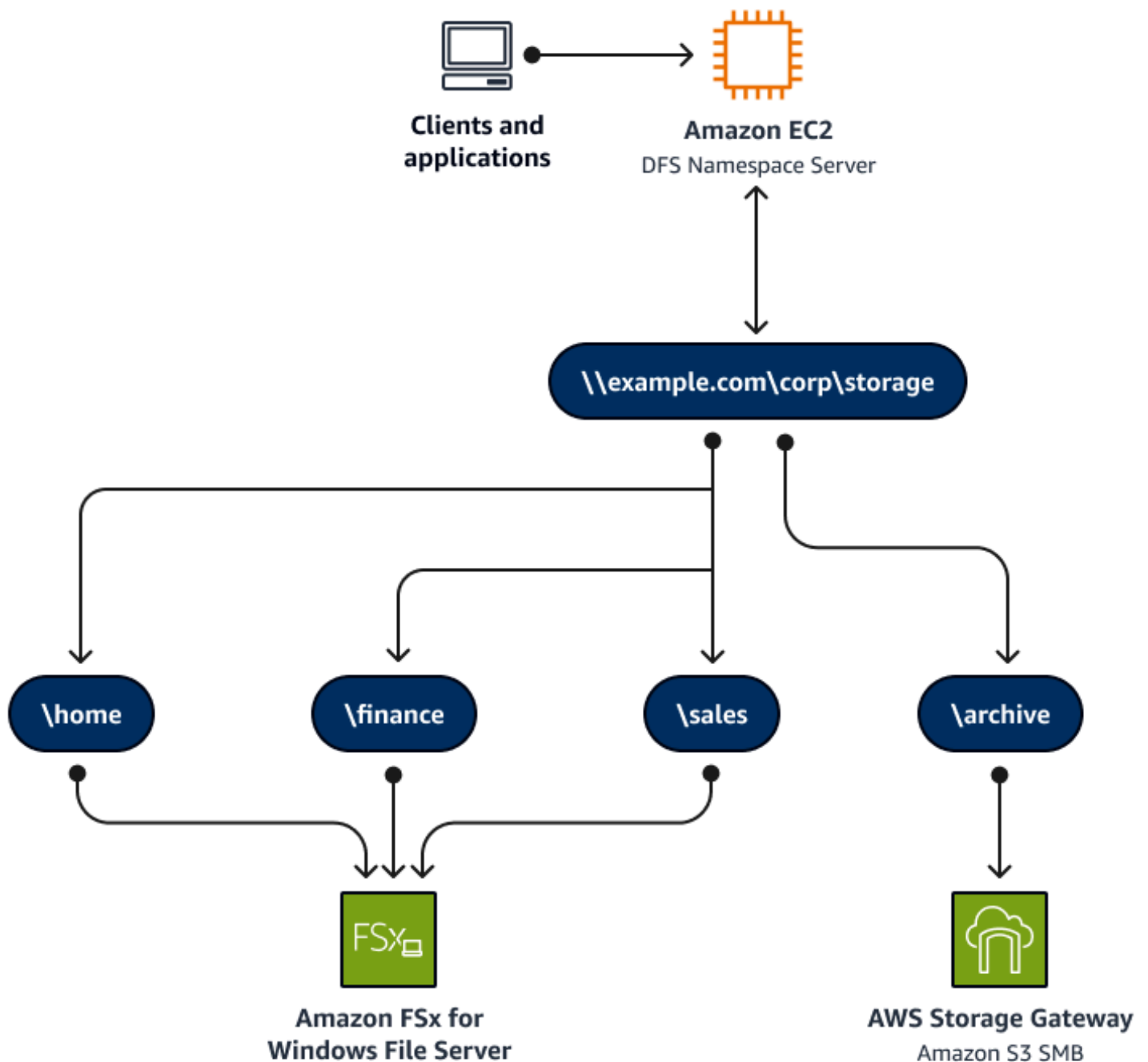
Consideraciones sobre la alta disponibilidad

La puesta en marcha de Storage Gateway es un único punto de error para el acceso a los archivos. Para evitar tiempos de inactividad innecesarios, le recomendamos que implemente un control de acceso estricto que permita a los usuarios realizar cambios o detener e iniciar la instancia de Storage Gateway. Además, para las implementaciones en marcha AWS, resulta beneficioso utilizar Amazon Data Lifecycle Manager para crear instantáneas de enrutamiento que permitan recuperar rápidamente la implementación de Storage Gateway. Si ejecuta Storage Gateway de forma local VMware, puede configurarlo para una [alta disponibilidad](#).

Uso de varios sistemas de archivos

Separar las cargas de trabajo de archivos de uso diario de las cargas de trabajo de archivo puede ayudar a evitar costos de almacenamiento innecesarios. Storage Gateway tiene la capacidad de implementarse junto con un sistema de archivos FSx para Windows File Server. Al usar los [espacios de nombres DFS](#), puede presentar su almacenamiento principal de uso diario que se ejecuta en Windows File Server y su almacenamiento que se ejecuta en FSx Amazon S3 (al que se accede a través de Storage Gateway).

En el siguiente diagrama, se muestra cómo se puede utilizar un único espacio de nombres DFS como punto de acceso frontend para distintas opciones de almacenamiento de backend.



Los clientes se redirigen a una estructura de carpetas, como \\ejemplo.com\almacenamiento. Este directorio principal contiene los subdirectorios. Un sistema de archivos FSx para servidores de archivos de Windows contiene los recursos compartidos de archivos a los que se accede de forma habitual. Puede usar un recurso compartido de archivos creado en Storage Gateway para archivar datos. Los usuarios pueden archivar los elementos manualmente en la carpeta de archivo, aunque

también puede crear un proceso para automatizar el traslado de algunos archivos de los recursos compartidos de archivos habituales a la carpeta de archivos.

Considere lo siguiente:

- Revise sus requisitos de almacenamiento y proporcione el [almacenamiento adecuado para la memoria caché](#).
- Añada su puerta de enlace a la configuración de Active Directory y utilice [Windows estándar ACLs para acceder a los archivos](#).

FSx File Gateway

La implementación de FSx File Gateway es similar a la implementación de S3 File Gateway, pero es aún más fácil si usa el asistente de inicio. Para obtener instrucciones detalladas, consulte el [paso 3: Crear y activar un Amazon FSx File Gateway](#) en la documentación de Amazon FSx File Gateway. Tras implementar FSx File Gateway en su entorno, podrá asociarlo a sus sistemas de FSx archivos de Amazon existentes y obtener acceso a sus archivos.

El almacenamiento es la consideración principal a la hora de implementar FSx File Gateway. El almacenamiento predeterminado proporciona 150 GB, que es una cantidad decente de espacio para almacenar archivos en caché. La creación de alertas de supervisión cuando haya poco espacio libre puede ser útil para ajustar el tamaño del almacenamiento sin asignar recursos en exceso.

Recursos adicionales

- [AWS Storage Gateway recursos](#) (AWS documentación)

Active Directory

Amazon Elastic Compute Cloud (Amazon EC2) con Windows Server es un entorno seguro, fiable y de alto rendimiento para implementar aplicaciones y cargas de trabajo basadas en Windows. Puede aprovisionar instancias rápidamente y escalarlas o reducirlas verticalmente según sea necesario, pagando solo por lo que utilice. Los servicios de Active Directory se utilizan como origen principal de administración de identidades en los entornos de Windows Server.

Esta sección abarca los siguientes temas:

- [Active Directory autoadministrado en Amazon EC2](#)
- [AWS Managed Microsoft AD](#)
- [Conector de AD](#)

Active Directory autoadministrado en Amazon EC2

Descripción general de

En esta sección se proporcionan recomendaciones para reducir el costo de la puesta en marcha de Active Directory en Amazon Elastic Compute Cloud (Amazon EC2). El objetivo principal es asegurarse de que puede dimensionar los controladores de dominio de Active Directory de forma adecuada y utilizar la flexibilidad que ofrecen para adaptarse Nube de AWS a su entorno según sea necesario. AWS puede ayudarle a detener fácilmente una instancia y redimensionarla para adaptarla a sus necesidades cambiantes, o a reducir el tamaño de la instancia si la amplía demasiado rápido. Elegir el tamaño y el tipo de instancia correctos puede generar ahorros significativos.

Impacto del costo

En la siguiente tabla, se muestra la diferencia entre elegir una instancia de la familia de instancias ampliables y una instancia de uso general. Esta opción puede ahorrarle una cantidad considerable de dinero cada mes. La planificación y el dimensionamiento adecuados de la instancia pueden ser de ayuda para gestionar los costos.

Tipo de instancia	Number of instances (Número de instancias)	vCPU	Memoria	Costo
t3a.medium	2	2.	8	81,76 USD/mes
m5a.large	2	2.	8	259,88 USD/mes

[Para obtener más información sobre los costos, consulta la Calculadora de precios de AWS estimación.](#)

Un ahorro de 178,12 USD al mes se traduce en un ahorro de más de 2000 USD al año por el uso de controladores de dominio. Recuerde que se trata de un espacio reducido de tan solo dos controladores de dominio en una cuenta. A escala, con varias cuentas y controladores de dominio adicionales, estos ahorros pueden traducirse en una reducción de costos significativa.

Recomendaciones de optimización de costos

Microsoft proporciona [recomendaciones de planificación de la capacidad](#) para implementar su entorno de Active Directory. Le recomendamos que tenga en cuenta los siguientes componentes principales al planificar o escalar su entorno de Active Directory:

- Memoria
- Network
- Almacenamiento
- Procesador

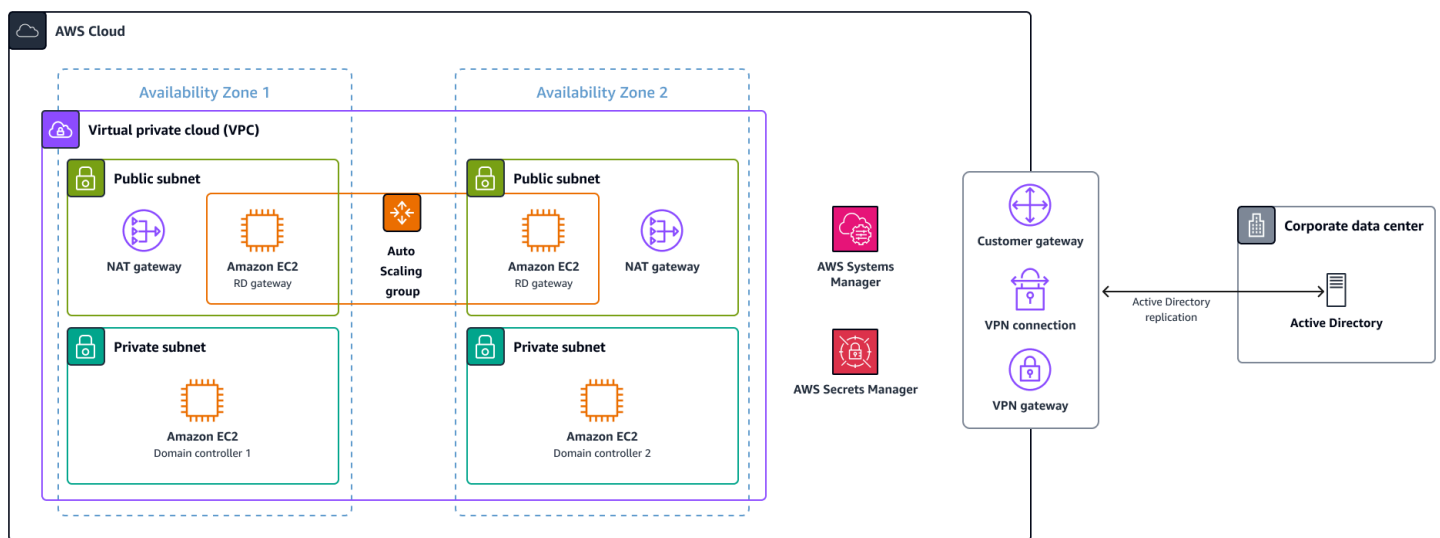
Teniendo en cuenta estos componentes principales, puede seleccionar un tipo de instancia que se adapte a su entorno de Active Directory en AWS. En esta sección se describen algunos ejemplos de escenarios de AWS implementación desde Active Directory. Estos escenarios dejan claro que no es necesario replicar el entorno local si no se planea gestionar el mismo número de usuarios y ordenadores que en el entorno local. AWS

En la siguiente tabla, se destacan los componentes importantes relacionados con la vCPU, la memoria y el disco para su AWS espacio físico.

Componente	Estimaciones
Almacenamiento o tamaño de la base de datos	De 40 a 60 KB por usuario
RAM	Tamaño de base de datos Recomendaciones del sistema operativo base Aplicaciones de terceros
Network	1 GB
CPU	1000 usuarios simultáneos para cada núcleo

Escenario de implementación híbrido

En el siguiente diagrama, se muestra un ejemplo de arquitectura para una implementación híbrida de Active Directory.



Como se muestra en el diagrama, normalmente cuenta con un espacio en las instalaciones y, luego, lo amplía a la Nube de AWS. En las fases iniciales de una migración, normalmente no tendrá todos los usuarios y servidores implementados en AWS. Por eso es importante implementar inicialmente un espacio de menor tamaño para ahorrar dinero al migrar.

Si va a mantener un espacio en las instalaciones en el que los servidores y los usuarios se autentican en dichas instalaciones, no necesitará el mismo espacio para los controladores de

dominio de AWS. Si sigue las prácticas recomendadas de Active Directory, puede implementar los [sitios y servicios de Active Directory](#) adecuados para autenticar a los usuarios y los equipos en su espacio en las instalaciones y hacer que solo el espacio en AWS se autentique con los controladores de dominio de AWS. Esto le permite evitar sobredimensionar el espacio que ocupa Active Directory AWS al limitar el uso únicamente a AWS los recursos y no a toda la infraestructura local. Para obtener orientación sobre cómo diseñar una configuración híbrida, consulte [Proper placement of domain controllers and site considerations](#) en la documentación de Microsoft.

Optimice para una AWS migración ajustando el tamaño

Si va a implementar una nueva instancia de Active Directory para sus usuarios o planea migrar completamente a ella AWS para su infraestructura de Active Directory, le recomendamos que planifique el tamaño según las recomendaciones de Microsoft en cuanto a vCPU, memoria y espacio en disco para las instancias seleccionadas en la tabla anterior.

Si se trata de un espacio nuevo, puede empezar poco a poco y aprovechar la posibilidad de [cambiar fácilmente los tipos de instancias](#) para cambiar el tamaño del entorno según vaya creciendo en AWS. En la sección [Windows en Amazon EC2](#) de esta guía se muestra cómo supervisar y revisar el uso de la CPU y la memoria en AWS. De esta forma, sabrá cuándo debe aumentar el tamaño de la instancia de EC2.

Si va a migrar por completo su entorno en las instalaciones de Active Directory a AWS, puede implementar los mismos planes de tamaño para garantizar un rendimiento adecuado. Antes de duplicar lo que ya tiene instalado en las instalaciones AWS, le recomendamos que realice una revisión exhaustiva del entorno de Active Directory. Esto puede ser útil para evitar el sobreaprovisionamiento. Asegúrese de usar el Monitor de rendimiento para recopilar información sobre la cantidad de tráfico y la utilización de los controladores de dominio actuales. Esto le permitirá comprender el uso general para poder ajustar el tamaño y, en última instancia, reducir los costos.

Optimice Active Directory en AWS

Si está ejecutando Active Directory AWS, también es importante que supervise continuamente la utilización y que cambie el tamaño de las instancias según sea necesario para reducir sus gastos. Puede AWS Compute Optimizer utilizarlos para obtener información sobre los recursos que utiliza AWS. Para obtener información sobre el uso de Compute Optimizer para ajustar el tamaño de las cargas de trabajo de Windows, consulte la sección [Windows en Amazon EC2](#) de esta guía. Para obtener un análisis más exhaustivo, puede utilizar Monitor de rendimiento para supervisar el uso de los controladores de dominio de Active Directory, evaluar el rendimiento y, a continuación, cambiar el tamaño en consecuencia.

También se puede utilizar CloudWatch para supervisar el rendimiento de los controladores de dominio. Para optimizar tus controladores de dominio (ampliándolos o reduciéndolos), puedes usar las métricas disponibles CloudWatch para ayudarte a tomar las decisiones correctas. Puede usar el CloudWatch agente para configurar métricas personalizadas del Monitor de rendimiento que se enviarán para la recopilación de datos. Para obtener instrucciones, consulte [¿Cómo puedo usar el CloudWatch agente para ver las métricas del Monitor de rendimiento en un servidor Windows?](#) en el Centro de AWS conocimiento.

Tras implementar el CloudWatch agente, puede configurar las siguientes métricas en el archivo de configuración del agente que se encuentra en `metrics_collected`:

Categoría métrica	Nombre de métrica
Base de datos a instancias (NTDSA)	% de aciertos de la caché de la base de datos
Latencia media de lecturas de la base de datos de E/S	
I/O database reads/sec	
Latencia media de escrituras de registros de E/S	
DirectoryServices (NTDS)	Tiempo de enlace de LDAP
Operaciones de replicación pendientes de DRA	
Sincronizaciones de replicación pendientes de DRA	
DNS	Consultas recursivas por segundo
Error de consulta recursiva por segundo	
Consultas de TCP recibidas por segundo	
Consultas totales recibidas por segundo	
Respuestas totales enviadas por segundo	

Categoría métrica	Nombre de métrica
Consultas de UDP recibidas por segundo	
LogicalDisk	Longitud de la cola de disco
% de espacio libre	
Memoria	% de bytes confirmados en uso
Tiempo de conservación medio de la caché en espera a largo plazo (s)	
Interfaz de red	Bytes enviados por segundo
Bytes recibidos por segundo	
Ancho de banda actual	
NTDS	Retraso de fila estimado de ATQ
Latencia de solicitudes de ATQ	
Lecturas del directorio DS por segundo	
Búsquedas en el directorio DS por segundo	
Escrituras en el directorio DS por segundo	
Sesiones de clientes LDAP	
Búsquedas LDAP por segundo	
Enlaces LDAP correctos por segundo	
Procesador	% de tiempo de procesador
Estadísticas de seguridad para todo el sistema	Autenticaciones de Kerberos
Autenticaciones de NTLM	

Recursos adicionales

- [Active Directory Domain Services en AWS: Guía de implementación de soluciones para socios](#) (AWS documentación)
- [Capacity planning for Active Directory Domain Services](#) (documentación de Microsoft)
- [Consideraciones de diseño para ejecutar Active Directory en instancias EC2 \(documentos técnicos AWS\)](#)

AWS Managed Microsoft AD

Descripción general de

AWS Directory Service for Microsoft Active Directory, también conocido como AWS Managed Microsoft AD, funciona con un Active Directory de Windows Server y es administrado por AWS. Se puede utilizar AWS Managed Microsoft AD para migrar una amplia gama de aplicaciones compatibles con Active Directory a la Nube de AWS. AWS Managed Microsoft AD funciona con una variedad de aplicaciones y servicios nativos de Active Directory. También es compatible con [aplicaciones y servicios administrados por AWS](#). Si bien el servicio y sus mecanismos de facturación no AWS Managed Microsoft AD disponen de muchas herramientas de optimización de costes, existen algunos principios de diseño que pueden ayudarle a mantener los costes al mínimo.

Impacto del costo

Dado que AWS Managed Microsoft AD se trata de un servicio gestionado basado en el presente SKUs, el dimensionamiento es un proceso relativamente sencillo. Actualmente hay dos tamaños SKUs disponibles: las ediciones Standard y Enterprise. Otras SKUs incluyen el uso compartido de directorios, la adición de controladores de dominio adicionales (incluidas regiones adicionales) y la transferencia de datos entre regiones.

Recomendaciones de optimización de costos

Existen diferencias entre la Edición AWS Managed Microsoft AD Estándar y la Edición AWS Managed Microsoft AD Empresarial. La edición Enterprise admite hasta 500 000 objetos de Active Directory y 500 recursos compartidos entre cuentas (límite flexible); ofrece también soporte para varias regiones. La edición Standard admite hasta 30 000 objetos de Active Directory y 5 recursos compartidos entre cuentas (el límite flexible es de aproximadamente 25 como máximo); no ofrece soporte para varias regiones.

Note

Los límites superiores de los objetos de Active Directory son aproximados. Su directorio podría admitir más o menos objetos en función del tamaño de estos y del comportamiento y las necesidades de rendimiento de sus aplicaciones.

Las preguntas que hay que tener en cuenta antes de seleccionar el tipo de directorio son las siguientes:

- ¿Se requiere soporte para varias regiones?
- ¿Se va a compartir el directorio con más de 25 cuentas?
- ¿El número de objetos de Active Directory será superior a 30 000?

Si la respuesta a alguna de las preguntas anteriores es afirmativa, se requiere la edición Enterprise. Si la respuesta a todas las preguntas es negativa, le recomendamos que empiece con la edición Standard.

Note

Puede pasar un directorio de la edición Standard a la edición Enterprise, pero no al revés. Implementar la edición Standard no es una decisión irreversible. Si desea actualizar su directorio a Enterprise Edition, póngase en contacto con AWS.

Si comparte directorios en la edición Enterprise de AWS Managed Microsoft AD, hay un costo por cada recurso compartido. Este es inferior al costo de implementar un directorio en cada cuenta, pero tenga en cuenta que los costos derivados del uso compartido pueden aumentar si no se controlan. Le recomendamos que solo comparta directorios con cuentas que contengan Amazon Relational Database Service (Amazon RDS) y FSx Amazon for Windows File Server, ya que solo esos servicios admiten esta función. Tenga en cuenta que tiene la opción de integrar FSx Windows File Server con su Active Directory autogestionado, incluido un. AWS Managed Microsoft AD Si solo FSx se requiere Amazon en otra cuenta, puedes realizar un FSx despliegue autogestionado de Amazon AWS Managed Microsoft AD sin necesidad de compartir el directorio.

Al decidir cuándo implementar controladores de dominio adicionales, tenga en cuenta que AWS Managed Microsoft AD solo admite dos subredes en zonas de disponibilidad independientes en la

misma VPC. Agregar controladores de dominio adicionales no permite agregar más subredes. Para determinar si debes añadir controladores de dominio adicionales debido a problemas de rendimiento, consulta las [métricas de rendimiento de los controladores de dominio en CloudWatch](#). Esto le indica si uno o todos los controladores de dominio están sobrecargados. Si llega a la conclusión de que solo hay un controlador de dominio sobrecargado, agregar controladores de dominio adicionales no aliviará la carga y tendrá que profundizar en las aplicaciones y no en el equilibrio de carga entre los controladores de dominio disponibles actualmente. Si todos los controladores de dominio se utilizan de forma intensiva, agregar un controlador de dominio adicional podría reducir la carga de los controladores de dominio existentes. Para obtener instrucciones sobre cómo automatizar el escalado, consulte [Cómo automatizar el AWS Managed Microsoft AD escalado en función de las métricas](#) de uso en el blog AWS de seguridad.

Si amplió su directorio a varias regiones, le recomendamos que no utilice el directorio compartido NETLOGON ni SYSVOL para almacenar archivos. Todos los controladores de dominio replican el contenido de esos recursos compartidos. Al no utilizar recursos compartidos para el almacenamiento de archivos, los costos de transferencia de datos se reducen al mínimo.

También tiene la opción de suscribirse a un acuerdo empresarial con AWS. Los contratos empresariales le ofrecen la opción de personalizar los contratos que mejor se adapten a sus necesidades. Para obtener más información, consulte [Enterprise Customers](#).

Recursos adicionales

- [AWS Managed Microsoft AD cuotas](#) (AWS Directory Service documentación)
- [AWS Directory Service Precios](#) (AWS sitio web)
- [Active Directory Domain Services on AWS](#) (documentos técnicos de AWS)

Conector de AD

Descripción general de

[AD Connector](#) es un servicio de proxy que proporciona una forma sencilla de conectar tu Microsoft Active Directory local existente a [AWS aplicaciones](#) compatibles, como Amazon WorkSpaces, Amazon Quick y una unión de dominios integrada para instancias de Amazon Elastic Compute Cloud (Amazon EC2), sin almacenar en caché ninguna información en la nube. Con el conector de AD, puede agregar una cuenta de servicio a su directorio de Active Directory. El conector de AD elimina la necesidad de sincronizar directorios y los costos y dificultades que conlleva alojar una

infraestructura de federación. Si bien el conector de AD no cuenta con muchas herramientas de optimización de costos debido a la naturaleza del servicio y sus mecanismos de facturación, puede seguir las recomendaciones de diseño de esta sección para mantener los costos al mínimo.

Impacto del costo

AD Connector es un servicio gestionado basado en un ajuste preestablecido SKUs. Esto hace que el dimensionamiento sea un proceso sencillo. Hay dos tamaños SKUs disponibles: tamaño pequeño y grande. Puede usar [Calculadora de precios de AWS](#) para hacer una estimación de los costos relacionados con el conector de AD.

Recomendaciones de optimización de costos

Aparte de los recursos de computación del backend, no hay diferencia entre el tamaño pequeño y el grande del conector.

Las preguntas que hay que tener en cuenta antes de seleccionar el tipo de directorio son las siguientes:

- ¿Hay un gran número (más de 10 000) de usuarios activos que utilizan AWS aplicaciones integradas con el AD Connector?
- ¿El usuario pertenece a numerosos grupos, a grupos con muchos niveles de anidamiento o a grupos con anidamiento circular?

Si la respuesta a ambas preguntas es negativa, le recomendamos que comience con el tamaño pequeño. Si responde afirmativamente a alguna de las preguntas anteriores, puede valer la pena considerar un tamaño grande. Puede empezar con un conector de AD de tamaño pequeño y, si el directorio se deteriora debido al rendimiento, puede solicitar la mejora a un tamaño más grande.

Note

Puede pasar de un conector de AD pequeño a uno grande, pero no al revés.

La mayoría de los problemas de rendimiento no están relacionados con el conector de AD, sino con la sobrecarga de los controladores de dominio de Active Directory en las instalaciones debido a que muchos usuarios pertenecen a numerosos grupos, a grupos con muchos niveles de anidamiento o a grupos con anidamiento circular

También tiene la opción de suscribirse a un acuerdo empresarial con. AWS Los contratos empresariales le ofrecen la opción de personalizar los contratos que mejor se adapten a sus necesidades. Para obtener más información, consulte [Enterprise Customers](#).

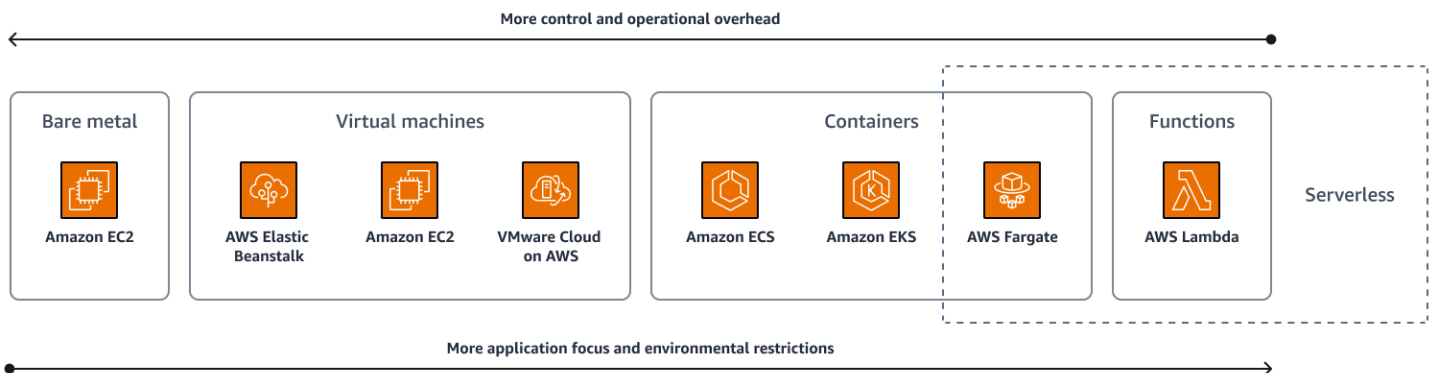
Recursos adicionales

- [Cuotas de AD Connector](#) (AWS Directory Service documentación)
- [Precios de otros tipos de directorios](#) (AWS sitio web)
- [Active Directory Domain Services on AWS](#) (documentos técnicos de AWS)

.NET

Desarrollar e implementar aplicaciones .NET es un aspecto clave para poder alcanzar la escala y la agilidad que ofrece la computación en la nube. Para muchas aplicaciones de .NET heredadas, la opción de procesamiento más adecuada para ejecutar aplicaciones AWS es usar máquinas virtuales, ya sea a través de Amazon Elastic Compute Cloud (Amazon EC2) AWS Elastic Beanstalk o Amazon. También es posible poner en marcha aplicaciones .NET en contenedores de Windows y Linux.

La introducción de .NET core le permite diseñar aplicaciones .NET modernas que aprovechan todas las ventajas de la nube. Las aplicaciones modernas pueden utilizar el conjunto tradicional de opciones informáticas y, además, orientarse a varios tipos de entornos sin servidor, incluidos AWS Fargate o AWS Lambda. .NET 6 y versiones posteriores ahora ofrecen un alojamiento eficiente de cargas de trabajo en instancias ARM64 de EC2, como las familias Graviton2 de EC2. Esto permite el acceso a la última generación de procesadores disponibles en Amazon EC2, lo que significa que sus aplicaciones se pueden alojar en sistemas de computación especializados para su tipo de carga de trabajo, como la codificación de video, los servidores web y la computación de alto rendimiento (HPC)



En esta sección se proporcionan recomendaciones para que pueda adaptar sus aplicaciones .NET de modo que pueda aprovechar las ventajas de la nube, pero centrándose en la rentabilidad.

Esta sección abarca los siguientes temas:

- [Refactorización para usar las versiones más modernas de .NET y migración a Linux](#)
- [Almacenamiento de aplicaciones .NET en contenedores](#)
- [Uso de instancias y contenedores Graviton](#)
- [Escalado dinámico para aplicaciones estáticas de .NET Framework](#)
- [Uso del almacenamiento en caché para reducir la demanda de la base de datos](#)

- [Uso de .NET sin servidor](#)
- [Uso de bases de datos personalizadas](#)

Refactorización para usar las versiones más modernas de .NET y migración a Linux

Descripción general de

La modernización de las aplicaciones heredadas de .NET Framework puede ser útil para mejorar la seguridad, el rendimiento y la escalabilidad. Una forma eficaz de modernizar las aplicaciones de .NET Framework es migrarlas a una versión moderna de .NET (6 o superior). Estas son algunas de las principales ventajas de migrar estas aplicaciones a .NET de código abierto:

- Reducir los costos asociados al uso de licencias de Windows usando un sistema operativo Linux.
- Aprovechar la disponibilidad de los lenguajes modernos.
- Obtener un rendimiento optimizado para Linux.

Muchas organizaciones siguen utilizando versiones anteriores de .NET Framework. Esto puede plantear riesgos de seguridad, ya que Microsoft ya no aborda las vulnerabilidades de las versiones anteriores. Microsoft ha dejado de dar soporte a las versiones recientes de .NET Framework 4.5.2, 4.6 y 4.6.1. Es muy importante evaluar los riesgos y beneficios de seguir utilizando las versiones anteriores del marco. Para reducir el riesgo y los costos, puede valer la pena invertir tiempo y esfuerzo en la refactorización a una versión moderna de .NET.

Impacto del costo

Considere utilizar un tipo de instancia de EC2 de uso general (m5), que ofrezca un conjunto equilibrado de recursos de computación, de memoria y de red. Estas instancias son adecuadas para una variedad de aplicaciones, como servidores web, bases de datos medianas y repositorios de código fuente.

Por ejemplo, una instancia m5.xlarge bajo demanda con 4 v CPUs y 16 GB de memoria en Windows Server (licencia incluida) en el este de EE. UU. (Virginia del Norte) cuesta 274,48\$ al mes. Los mismos recursos en un servidor Linux cuestan 140,16 USD al mes. En este ejemplo, el costo se reduce en un 49 % al migrar la aplicación de .NET Framework a una versión moderna de .NET y poner en marcha la aplicación en un servidor Linux. El costo puede variar en función de las opciones

(por ejemplo, el tipo de instancia, el sistema operativo o el almacenamiento) que elija al seleccionar una [instancia de EC2](#). Puede optimizar aún más los costos mediante [Savings Plans](#) o [instancias reservadas](#). Para obtener más información, utilice la [Calculadora de precios de AWS](#) para hacer una previsión de los costos. En el caso de las instancias incluidas en Windows, el costo de la licencia es de [0,046 USD por vCPU por hora](#), independientemente del modelo de precios.

La migración de estas aplicaciones .NET Framework a la versión moderna de .NET requiere un esfuerzo por parte del equipo de desarrollo. Debe evaluar sus aplicaciones y sus dependencias para comprobar si son compatibles con la versión de la plataforma de destino. [AWS Asistente de portabilidad para .NET](#) es una herramienta de asistencia que analiza las aplicaciones de .NET Framework y evalúa compatibilidad con .NET, lo que le permite portar con mayor rapidez sus aplicaciones para que sean compatibles con Linux. Asistente de portabilidad para .NET identifica las incompatibilidades con .NET, busca sustitutos conocidos y hace una evaluación de la compatibilidad detallada. Tras migrar la solución, debe realizar cambios manuales en el código para que el proyecto se compile correctamente con las dependencias. Esto reduce el esfuerzo manual que implica la modernización de las aplicaciones a Linux. Si su aplicación es compatible con los procesadores ARM, al migrar a Linux podrá utilizar instancias Graviton. Esto puede ser útil para reducir aún más los costos en un 20 %. Para obtener más información, consulte [Cómo potenciar .NET 5 con AWS Graviton2: puntos de referencia](#) en el blog sobre informática. AWS

Existen otras herramientas, como [Kit de herramientas de AWS para la refactorización de .NET](#) y [Asistente para la actualización de .NET](#), que pueden ser de ayuda para migrar aplicaciones heredadas de .NET Framework a la versión moderna de .NET.

Recomendaciones de optimización de costos

Para migrar aplicaciones de .NET Framework, haga lo siguiente:

1. Requisitos previos: para utilizar Asistente de portabilidad para .NET, debe instalar .NET 5 o una versión posterior en la máquina en la que va a analizar el código fuente de la aplicación. Los recursos de la máquina deben tener una velocidad de GHz procesamiento mínima de 1,8, 4 GB de memoria y 5 GB de espacio de almacenamiento. Para obtener más información, consulte [Prerequisites](#) en la documentación de Asistente de portabilidad para .NET.
2. Evaluación: descargue Asistente de portabilidad para .NET como un [archivo procesable](#) (de descarga). Puede descargar e instalar la herramienta en su máquina para iniciar la evaluación de sus aplicaciones. La página de evaluación contiene proyectos y paquetes transferidos APIs que son incompatibles con la versión moderna de .NET. Por este motivo, se producen errores de compilación en la solución después de la evaluación. Puede ver o descargar los resultados de

la evaluación en un archivo CSV. Para obtener más información, consulte [Port a solution](#) en la documentación de Asistente de portabilidad para .NET.

3. Refactorización: después de evaluar la aplicación, puede migrar sus proyectos a la versión del marco de trabajo de destino. Al migrar una solución, Asistente de portabilidad modificará los archivos del proyecto y parte del código. Puede consultar los registros para revisar los cambios en el código fuente. En la mayoría de los casos, el código requerirá un esfuerzo adicional para completar la migración y las pruebas de cara a la producción. Según la aplicación, algunos de los cambios pueden incluir el marco de la entidad, la identidad y la autenticación. Para obtener más información, consulte [Port a solution](#) en la documentación de Asistente de portabilidad para .NET.

Este es el primer paso para modernizar sus aplicaciones a contenedores. La modernización de las aplicaciones de .NET Framework y su almacenamiento en contenedores de Linux puede implicar una serie de factores empresariales y técnicos. Uno de los factores clave es la reducción del costo total de propiedad al pasar de un sistema operativo Windows a uno Linux. Esto reduce los costos de licencia al migrar la aplicación a una versión multiplataforma de .NET y a contenedores para optimizar la utilización de los recursos.

Cuando la aplicación se haya migrado a Linux, podrá utilizarla [AWS App2Container](#) para almacenarla en contenedores. App2Container utiliza Amazon ECS o Amazon EKS como servicios de punto de conexión en los que puede implementar directamente. App2Container proporciona todos los artefactos de infraestructura como código (IaC) necesarios para implementar y almacenar en contenedores sus aplicaciones cuantas veces haga falta.

Consideraciones y recursos adicionales

- Si tiene aplicaciones creadas en VB.NET (un marco heredado de 2002) y desea migrarlas a .NET 6, consulte la publicación [Cómo portar aplicaciones antiguas de VB.NET a .NET 6.0 con Porting Assistant para .NET](#) en el blog Microsoft Workloads on. AWS
- Si tiene aplicaciones heredadas en Windows Communication Foundation (WCF) y desea ponerlas en marcha en una versión moderna de .NET, puede adoptar CoreWCF. Para obtener más información, consulte la publicación [Modernización de aplicaciones WCF heredadas a CoreWCF mediante Porting Assistant para .NET](#) en el blog Microsoft Workloads on. AWS
- Puede agregar un asistente de portabilidad como una extensión a su IDE de Visual Studio. Esto le permite realizar todas las tareas necesarias para convertir el código sin necesidad de cambiar entre el IDE y la herramienta Asistente de portabilidad para .NET. Para obtener más información, consulte la publicación sobre cómo [acelerar la modernización de las aplicaciones .NET con la](#)

[extensión IDE de Porting Assistant para .NET Visual Studio](#) en el AWS blog Microsoft Workloads on.

- [Asistente de portabilidad para .NET de AWS es ahora una herramienta de código abierto](#) con los componentes de análisis de código fuente y compatibilidad empleados en la evaluación. Esto puede animar a sus desarrolladores a utilizar y compartir los conocimientos y las prácticas recomendadas de portabilidad de .NET.
- Puede migrar aplicaciones de .NET Framework a la versión moderna de .NET en Linux mediante el AWS kit de herramientas para la refactorización de .NET. Para obtener más información, consulte la publicación [Acelere la modernización de .NET con el AWS kit de herramientas para la refactorización de .NET](#) en el blog Microsoft Workloads on. AWS
- Puede [acelerar la contenedorización y la migración de las aplicaciones de ASP.NET Core para utilizarlas](#). AWS AWS App2Container

Almacenamiento de aplicaciones .NET en contenedores

Descripción general de

Los contenedores son una forma ligera y eficiente de empaquetar e implementar aplicaciones de forma coherente y reproducible. En esta sección se explica cómo puede utilizar AWS Fargate, un servicio de contenedores sin servidor, para reducir los costos de sus aplicaciones .NET y, al mismo tiempo, tener una infraestructura escalable y fiable.

Impacto del costo

Algunos factores que influyen en la eficacia del uso de contenedores para ahorrar costos incluyen el tamaño y la complejidad de la aplicación, la cantidad de aplicaciones que deben implementarse y el nivel de tráfico y demanda de las aplicaciones. En el caso de las aplicaciones pequeñas o sencillas, es posible que los contenedores no supongan un ahorro de costos significativo en comparación con las estrategias de infraestructura tradicionales, ya que la sobrecarga que supone gestionar los contenedores y los servicios asociados puede, de hecho, aumentar los costos. Sin embargo, en el caso de aplicaciones más grandes o complejas, el uso de contenedores puede suponer un ahorro de costos, ya que mejora la utilización de los recursos y reduce la cantidad de instancias necesarias.

Le recomendamos que tenga en cuenta las siguientes consideraciones al utilizar contenedores para ahorrar costos:

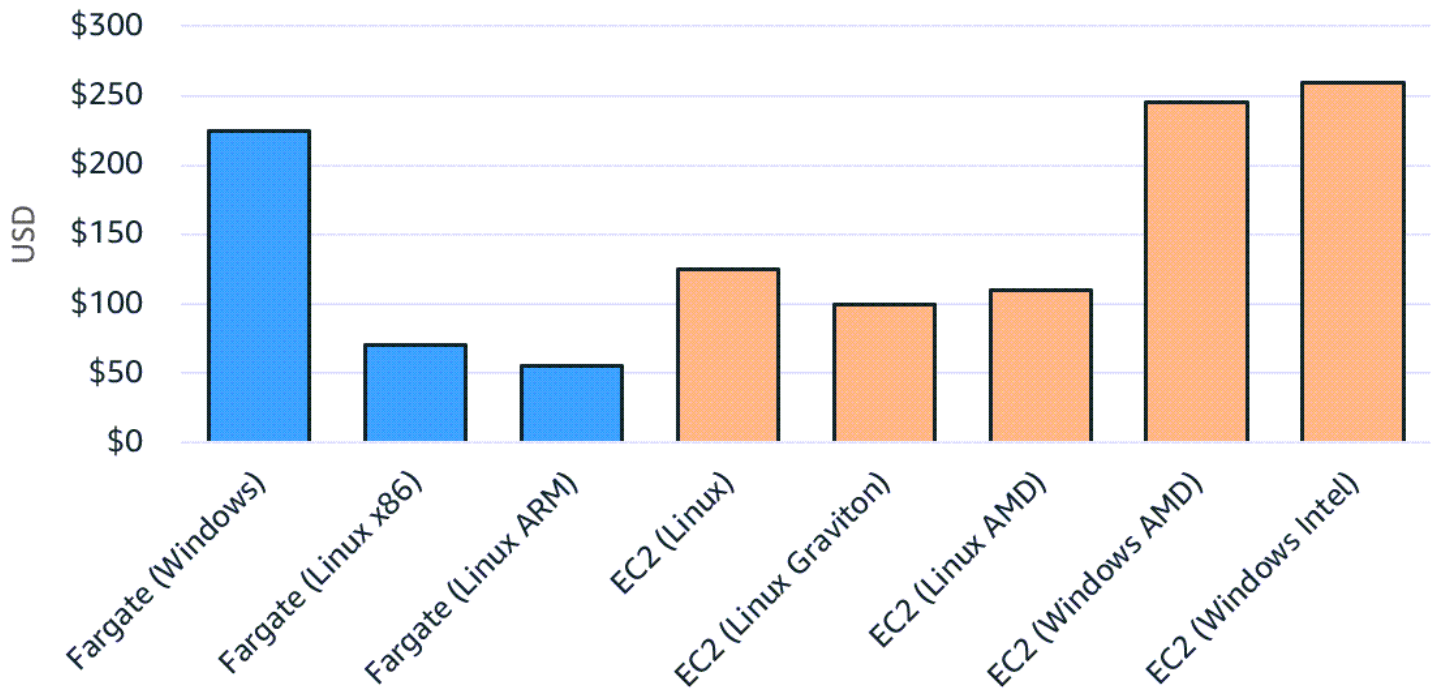
- **Tamaño y complejidad de las aplicaciones:** las aplicaciones más grandes y complejas son más adecuadas para el almacenamiento en contenedores porque suelen requerir más recursos y pueden beneficiarse más de un mejor uso de dichos recursos.
- **Cantidad de aplicaciones:** cuantas más aplicaciones deba implementar su organización, mayor será el ahorro de costos que se logrará al almacenar en contenedores.
- **Tráfico y demanda:** las aplicaciones que experimentan una demanda y un tráfico elevados pueden beneficiarse de la escalabilidad y la elasticidad que ofrecen los contenedores. Esto puede generar ahorros de costos.

Las diferentes arquitecturas y sistemas operativos influyen en los costos de los contenedores. Si utiliza contenedores de Windows, es posible que los costos no disminuyan debido a aspectos relativos al uso de licencias. Los costos asociados al uso de licencias son más bajos o incluso inexistentes con los contenedores de Linux. AWS Fargate En el siguiente gráfico se utiliza una configuración básica de la región EE.UU. Este (Ohio) con los siguientes ajustes: 30 tareas al mes, ejecutadas durante 12 horas cada una, con 4 V CPUs y 8 GB de memoria asignados.

Puede elegir entre dos plataformas informáticas principales para ejecutar sus contenedores AWS: [hosts de contenedores basados en EC2 y sistemas sin servidor](#) o [AWS Fargate](#) Si utiliza Amazon Elastic Container Service (Amazon ECS) en lugar de Fargate, debe mantener instancias en funcionamiento para que el motor de ubicación pueda crear contenedores cuando sea necesario. Si usa Fargate en su lugar, solo se aprovisiona la capacidad de computación necesaria.

En el siguiente gráfico se muestra la diferencia entre contenedores equivalentes que utilizan Fargate y Amazon EC2. Gracias a la flexibilidad de Fargate, las tareas de una aplicación pueden estar en funcionamiento durante 12 horas al día y no consumir recursos fuera de ese horario. Sin embargo, en el caso de Amazon ECS, debe controlar la capacidad de computación mediante un [grupo de escalado automático](#) de instancias de EC2. Esto puede hacer que la capacidad funcione las 24 horas del día, lo que, en última instancia, puede aumentar los costos.

Monthly costs of Fargate and Amazon EC2



Recomendaciones de optimización de costos

Uso de contenedores de Linux en lugar de Windows

Puede lograr ahorros significativos si utiliza contenedores de Linux en lugar de contenedores de Windows. Por ejemplo, puede ahorrar aproximadamente un 45 % en los costos de procesamiento si usa .NET Core en EC2 Linux en lugar de .NET Framework en EC2 Windows. Puede obtener un ahorro adicional del 40 por ciento si utiliza la arquitectura ARM (AWS Graviton) en lugar de la x86.

Si planea poner en marcha contenedores basados en Linux para las aplicaciones de .NET Framework existentes, debe migrar estas aplicaciones a versiones modernas y multiplataforma de .NET (por ejemplo, [.NET 6.0](#)) para poder utilizar contenedores de Linux. Una consideración importante es sopesar el costo de la refactorización y compararlo con los ahorros que se obtienen al reducir el costo de los contenedores de Linux. Para obtener más información sobre la migración de sus aplicaciones a versiones modernas de .NET, consulte [Porting Assistant for .NET](#) en la documentación de AWS .

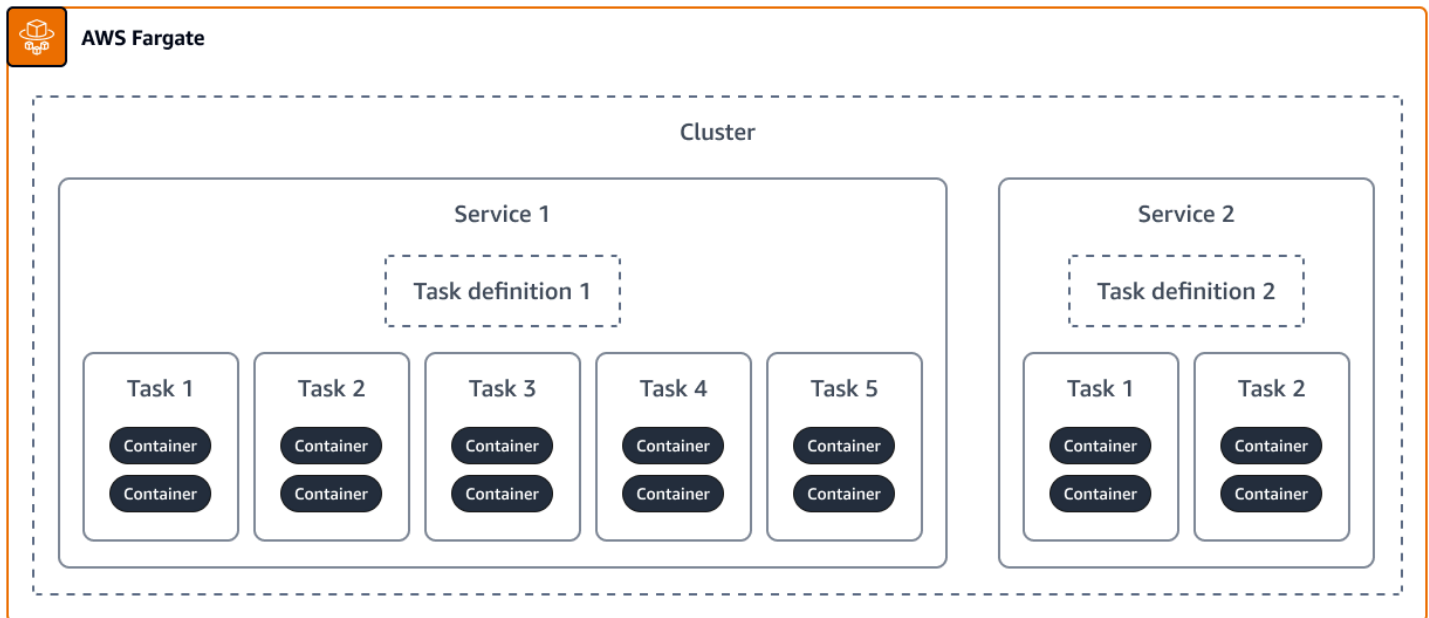
Otra ventaja de migrar a versiones modernas de .NET Framework (es decir, evitar .NET Framework) es que hay más oportunidades de modernización disponibles. Por ejemplo, puede considerar la

documentación de Savings Plans y en la sección [Optimizar el gasto para Windows en Amazon EC2](#) de esta guía.

Dimensionamiento correcto de las tareas de Fargate

Es importante asegurarse de que las tareas de Fargate tengan el tamaño correcto para lograr la mayor optimización posible de los costos. Con frecuencia, los desarrolladores no disponen de toda la información de uso necesaria a la hora de definir inicialmente las configuraciones de las tareas de Fargate utilizadas en sus aplicaciones. Esto puede provocar un aprovisionamiento excesivo de tareas y, posteriormente, generar gastos innecesarios. Para evitarlo, le recomendamos que cargue las aplicaciones de prueba que funcionan en Fargate para comprender cómo funciona la configuración de una tarea específica en diferentes escenarios de uso. Puede utilizar los resultados de las pruebas de carga, la vCPU, la asignación de memoria de las tareas y las políticas de escalado automático para dar con el equilibrio adecuado entre rendimiento y costo.

En el siguiente diagrama, se muestra cómo Compute Optimizer genera recomendaciones de tamaño para la tarea y el contenedor.



Un enfoque consiste en utilizar una herramienta de pruebas de carga, como la que se describe en [Distributed Load Testing on AWS](#), para establecer una línea base para la utilización de la vCPU y la memoria. Después de poner en marcha la prueba de carga para simular una carga de aplicación típica, puede refinar la configuración de la vCPU y la memoria para la tarea hasta alcanzar el valor de uso de referencia.

Recursos adicionales

- [Cost Optimization Checklist for Amazon ECS and AWS Fargate](#) (publicación del blog de contenedores de AWS)
- [Optimización teórica de costes mediante el tipo de lanzamiento de Amazon ECS: Fargate frente a EC2](#) (entrada del blog de AWS Containers)
- [Asistente de portabilidad para .NET \(documentación\)](#) AWS
- [Las pruebas de carga distribuidas están AWS](#) activas (biblioteca de AWS soluciones)
- [AWS Compute Optimizer lanza el soporte para los servicios de Amazon ECS en AWS Fargate](#) (entrada del blog AWS Cloud Financial Management)

Uso de instancias y contenedores Graviton

Descripción general de

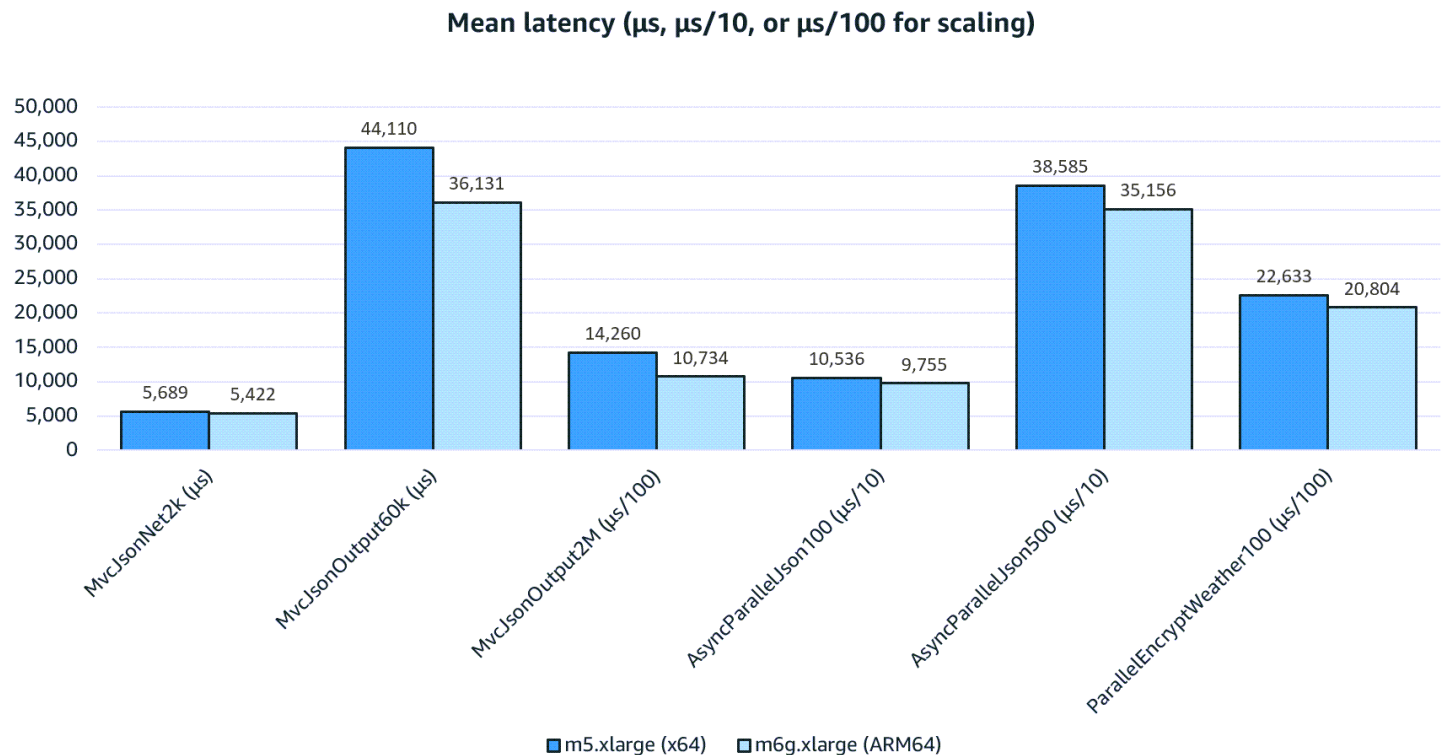
AWS Las instancias Graviton funcionan con procesadores ARM diseñados para ofrecer la mejor AWS relación precio-rendimiento para las cargas de trabajo en la nube que se ejecutan en Amazon Elastic Compute Cloud (Amazon EC2), incluidos los contenedores que se ejecutan en ellas. AWS Actualmente hay tres generaciones de Graviton disponibles para su uso en Amazon EC2. Esta guía se centra en el uso de Graviton 2 y 3 con aplicaciones .NET, ya que supone un importante ahorro de costos si usa las versiones más recientes de Graviton. Tenga en cuenta que las instancias Graviton solo funcionan en el sistema operativo Linux. Como resultado, las instancias Graviton son una buena opción para .NET en Linux, pero no son una opción para el sistema operativo Windows ni para las aplicaciones heredadas de .NET Framework.

Graviton 3 es un 60 % más eficiente que las instancias de EC2 comparables, con un rendimiento hasta un 40 % mejor. Esta guía se centra en los beneficios económicos del uso de Graviton, pero es importante tener en cuenta que Graviton ofrece otras dos ventajas: la mejora del rendimiento y una mayor sostenibilidad ambiental.

Impacto del costo

Puede ahorrar hasta un 45 % si empieza a usar Graviton. Tras refactorizar cualquier aplicación heredada de .NET Framework para convertirla a una versión moderna de .NET, podrá utilizar instancias Graviton. La migración a Graviton es una técnica eficaz de optimización de costos para los desarrolladores de .NET.

El ejemplo de la siguiente tabla muestra las posibles mejoras de rendimiento que se pueden lograr al migrar a las instancias Graviton.



Para obtener un desglose completo y una explicación del enfoque de evaluación comparativa utilizado para crear los resultados del diagrama anterior, consulte [Potenciar .NET 5 con AWS Graviton2: Benchmarks in the Compute Blog. AWS](#)

Una de las razones de la mejora en la eficiencia es que el término vCPU tiene significados distintos en x86 y en Graviton. En la arquitectura x86, una vCPU es un núcleo lógico que se logra mediante hipersubprocesamiento. En Graviton, una vCPU equivale a un núcleo físico que permite que la vCPU se comprometa plenamente con la carga de trabajo.

El resultado con Graviton2 es una relación precio-rendimiento un 40 % superior a la de las instancias x86 o x64 comparables. Graviton3 ofrece lo siguiente en comparación con Graviton2:

- Un perfil de rendimiento mejorado con un rendimiento hasta un 25 % mejor.
- Hasta el doble de rendimiento en operaciones de coma flotante.
- Un rendimiento el doble de rápido en cargas de trabajo criptográficas.
- Un rendimiento tres veces mejor en tareas de machine learning.

Además, Graviton3 es la primera instancia en la nube que incluye memoria. DDR5

En las tablas siguientes se muestra una comparación del ahorro de costos logrado al usar instancias basadas en Graviton y al usar instancias equivalentes basadas en x86.

En esta tabla se muestra que el uso de Graviton supone un 19,20 % de ahorro.

Tipo de instancia	Arquitectura	vCPU	Memoria (GB)	Costo por hora (bajo demanda)
t4g.xlarge	ARM	4	16	0,1344\$
t3.xlarge	x86	4	16	0,1664\$

En esta tabla se muestra que el uso de Graviton supone un 14,99 % de ahorro.

Tipo de instancia	Arquitectura	vCPU	Memoria (GB)	Costo por hora (bajo demanda)
c7g.4xlarge	ARM	16	32	0,5781\$
c6i.4xlarge	x86	16	32	0,6800\$

Es importante probar el perfil de rendimiento de la aplicación si planea usar Graviton. Graviton no reemplaza la necesidad de aplicar buenas prácticas de desarrollo de software. Puede utilizar las pruebas para comprobar si está aprovechando al máximo sus recursos de computación subyacentes.

Recomendaciones de optimización de costos

Hay varias formas de aprovechar los procesadores o instancias Graviton. En esta sección, se explican los cambios necesarios para pasar de utilizar una máquina con arquitectura x86 a instancias Graviton (ARM).

Cambio de la configuración del tiempo de ejecución en Lambda

Le recomendamos que modifique la configuración del tiempo de ejecución. AWS Lambda
Para obtener más información, consulte [Modificación del entorno de tiempo de ejecución](#) en la

documentación de Lambda. Dado que .NET es un lenguaje compilado, debe seguir un proceso de compilación para que funcione. Para ver un ejemplo de cómo hacerlo, consulte [.NET en Graviton](#) in GitHub.

Contenedores

En el caso de una carga de trabajo almacenada en un contenedor, cree una imagen de contenedor multiarquitectura. Para ello, especifique varias arquitecturas en el comando de compilación de Docker. Por ejemplo:

```
docker buildx build -t "myImageName:latest" --platform linux/amd64,linux/arm64 --push .
```

También puedes usar una herramienta como la que te ayude AWS Cloud Development Kit (AWS CDK) a [organizar la construcción](#). Para ver ejemplos de Docker, consulte [Building Multi-Arch Images for Arm and x86 with Docker Desktops](#) en la documentación de Docker.

Amazon EC2

Para migrar a ARM desde x86 o x64, debe centrarse en la arquitectura ARM en el paso de compilación. En Visual Studio, puede crear una ARM64 CPU. Para obtener instrucciones, consulte [To configure a project to target Arm64 and other platforms](#) en la documentación de Microsoft.

Si utiliza la CLI de .NET, obtendrá una compilación válida para Graviton al poner en marcha la compilación en una máquina ARM. Para ver una demostración, vea [Acelerar el rendimiento de .NET 6 con Arm64 en AWS Graviton2](#) o más. YouTube Los problemas de dependencia provocarán errores en el tiempo de compilación que luego se podrán solucionar de forma individual. Mientras haya bibliotecas ARM para cualquier dependencia, la transición debería ser relativamente sencilla.

Recursos adicionales

- [Cómo crear sus contenedores para ARM y ahorrar con las instancias Graviton y Spot en Amazon ECS](#) (AWS blog)
- [AWS Lambda Funciones impulsadas por el procesador AWS Graviton2: ejecute sus funciones en Arm y obtenga una relación precio-rendimiento hasta un 34% superior](#) (blog)AWS
- [Migración de AWS Lambda funciones a procesadores AWS Graviton2 basados en ARM](#) (blog)AWS
- [Cree e implemente aplicaciones web.NET en clústeres Amazon ECS de AWS Graviton 2 con tecnología ARM mediante AWS CDK](#) (blog)AWS

- [Graviton Fast Start: un nuevo programa que le ayudará a trasladar sus cargas de trabajo a Graviton](#) (blog) AWS AWS
- [Cómo potenciar .NET 5 con AWS Graviton2](#): puntos de referencia (blog)AWS

Escalado dinámico para aplicaciones estáticas de .NET Framework

Descripción general de

Una de las principales ventajas de usar la nube para las aplicaciones es la elasticidad, es decir, la capacidad de escalar o reducir horizontalmente los recursos de computación en función de la demanda. Esto le permite pagar solo por la capacidad de computación que necesita, en lugar de aprovisionarla para los picos de uso. El Cyber Monday, día en el que los minoristas que venden por Internet pueden lograr rápidamente mucho más tráfico de lo normal (por ejemplo, [incrementos de miles por ciento en cuestión de minutos](#)), es un buen ejemplo de elasticidad.

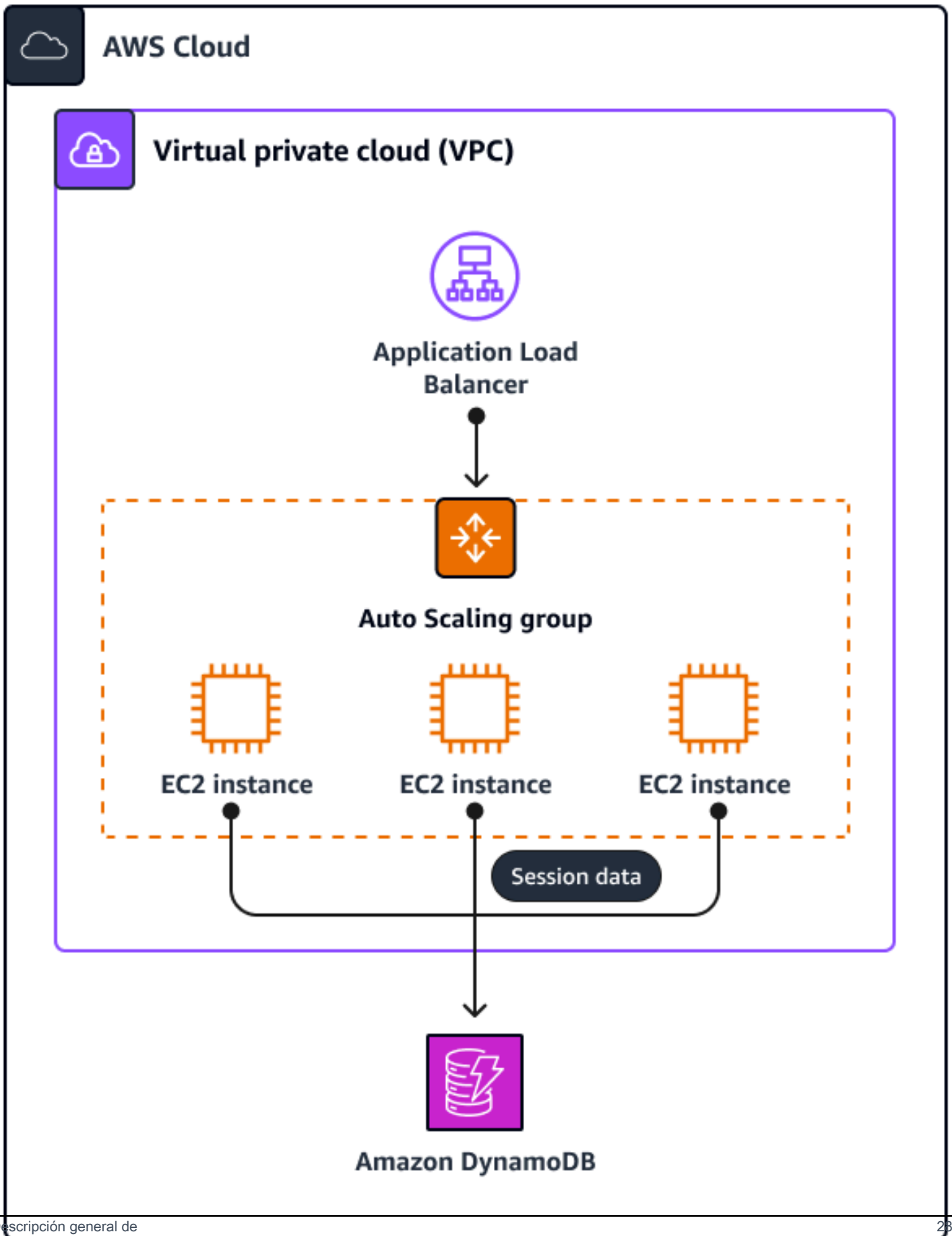
Si va a migrar aplicaciones web de .NET heredadas a la nube (por ejemplo, aplicaciones de ASP.NET Framework que funcionan en IIS), la capacidad de escalar rápidamente los conjuntos de servidores con equilibrio de carga puede resultar difícil o imposible debido a la naturaleza activa de la aplicación. Los datos de la sesión del usuario se almacenan en la memoria de la aplicación, normalmente mediante el [estado de la sesión de ASP.NET](#) o con variables estáticas que contienen datos de varias solicitudes que deben conservarse. La afinidad entre las sesiones de los usuarios se suele mantener mediante sesiones persistentes del equilibrador de carga.

Esto supone un desafío desde el punto de vista operativo. Cuando se requiere una mayor capacidad, debe aprovisionar y agregar servidores de forma intencionada. Este proceso puede ser lento. Dejar los nodos fuera de servicio, ya sea para aplicar parches o ante fallos inesperados, puede resultar problemático para la experiencia del usuario final, ya que todos los usuarios asociados a los nodos afectados pierden su estado. En el mejor de los casos, esto requeriría que los usuarios inicien sesión de nuevo.

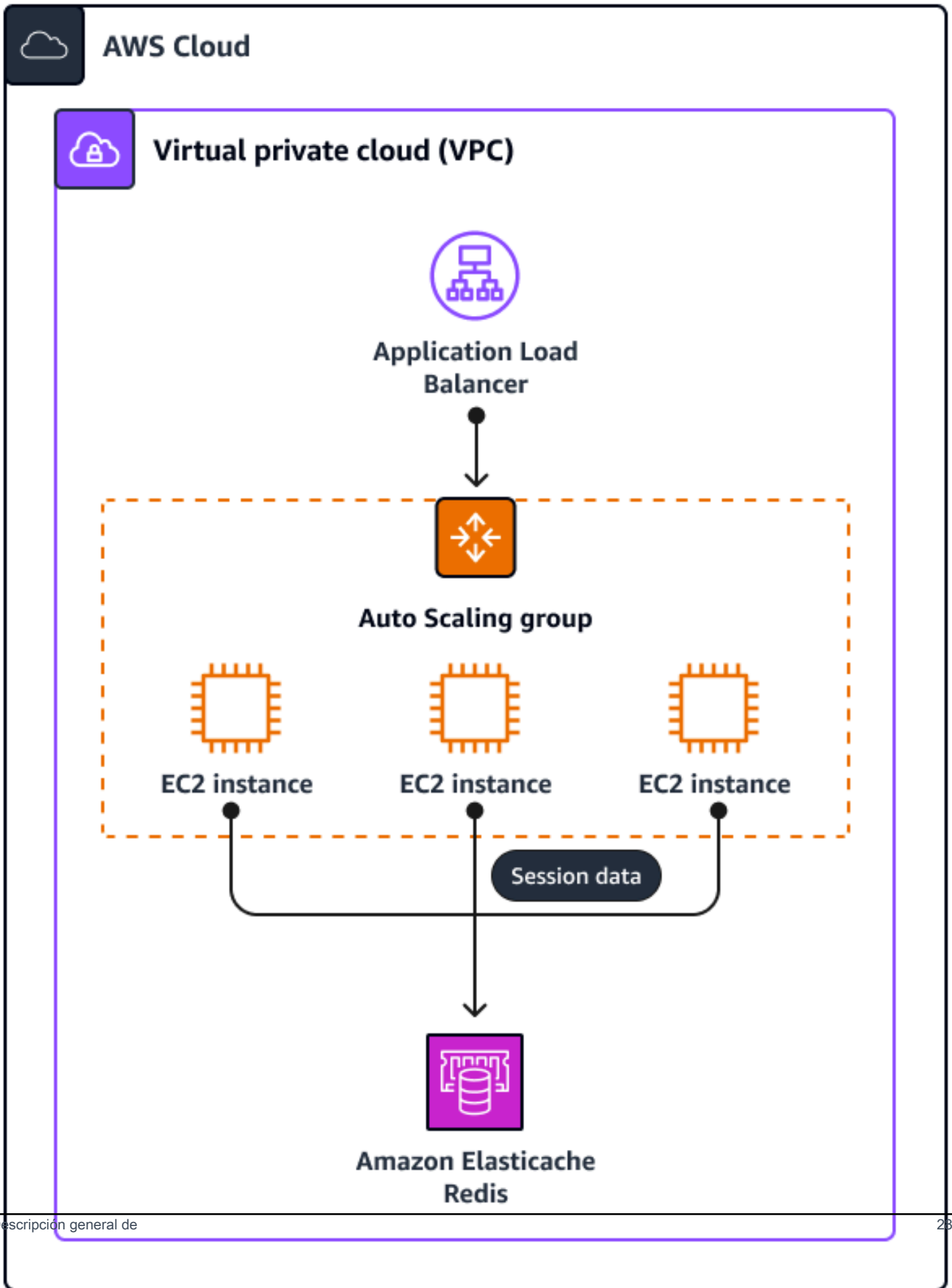
Al centralizar el estado de la sesión de las aplicaciones de ASP.NET y aplicar reglas de escalado automático a las aplicaciones de ASP.NET heredadas, puede aprovechar la elasticidad de la nube y, potencialmente, aprovechar los ahorros de costos al poner en marcha las aplicaciones. Por ejemplo, ahorra costos gracias a la escalabilidad de los recursos de computación, pero también puede elegir entre los diferentes modelos de precios disponibles, como reducir el [uso de instancias reservadas](#) y adherirse a los [precios de las instancias de spot de Amazon](#).

Dos técnicas habituales incluyen el uso de [Amazon DynamoDB como proveedor de estado de sesión](#) y el uso de [ElastiCache Amazon \(Redis OSS\) como](#) almacén de sesiones de ASP.NET.

En el siguiente diagrama, se muestra una arquitectura que utiliza DynamoDB como proveedor de estado de sesión.



El siguiente diagrama muestra una arquitectura que utiliza ElastiCache (Redis OSS) como proveedor de estado de sesión.



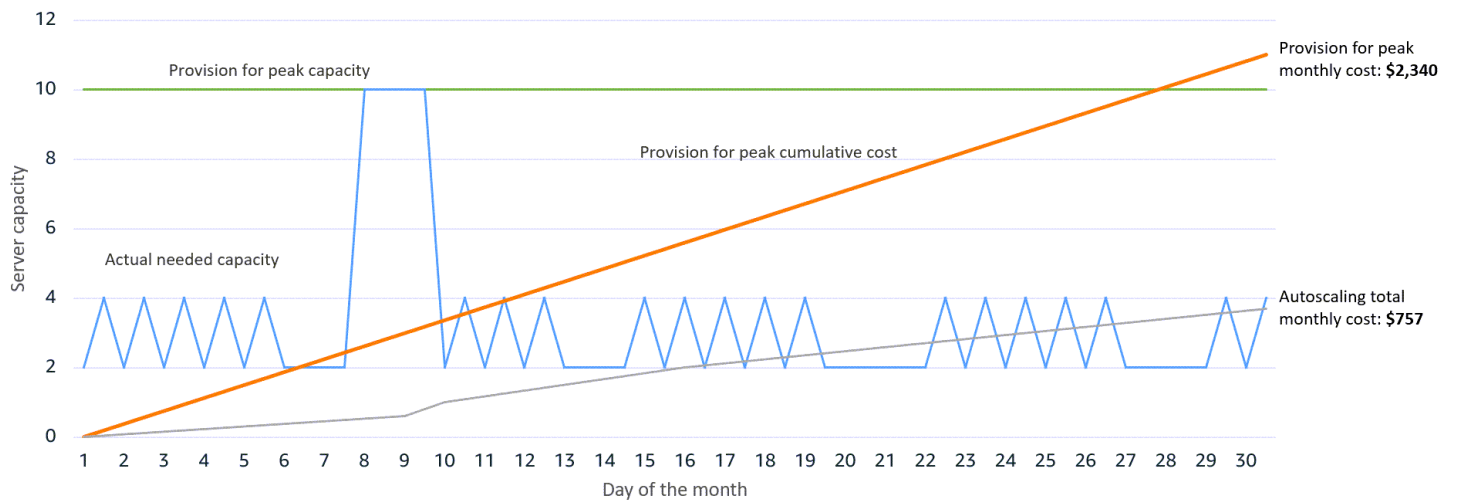
Impacto del costo

Para ver las ventajas de escalar una aplicación de producción, le recomendamos que modele la demanda real. En esta sección se hacen las siguientes suposiciones para modelar una aplicación de muestra:

- Las instancias que se agregan y se eliminan de la rotación son idénticas y no se introduce ninguna variación en el tamaño de las instancias.
- Siempre hay dos servidores activos como mínimo para mantener la alta disponibilidad de la aplicación.
- La cantidad de servidores se amplía linealmente con el tráfico (es decir, el doble de tráfico requerirá el doble de procesamiento).
- El tráfico se modela a lo largo de un mes en incrementos de 6 horas, con una variación intradía y un pico de tráfico anormal (por ejemplo, una venta promocional) en un día en el que el tráfico se multiplica por 10. El tráfico de fin de semana se modela en función de la utilización básica.
- El tráfico nocturno se modela en función del uso básico, mientras que el tráfico de los días laborables se modela en función del cuádruple del uso.
- Los precios de las instancias reservadas se basan en los precios del plan de un año, sin pagos iniciales. Los precios diurnos normales se basan en los precios bajo demanda, mientras que para la demanda puntual se utilizan los precios de las instancias de spot.

El siguiente diagrama ilustra cómo este modelo aprovecha la elasticidad de una aplicación .NET en lugar de aprovisionar para hacer frente a los picos de uso. Esto se traduce en un ahorro de aproximadamente el 68 %.

Comparison of cumulative costs for peak provisioning and autoscaling



Si utiliza DynamoDB como mecanismo de almacenamiento del estado de la sesión, utilice los siguientes parámetros:

Storage: 20GB
 Session Reads: 40 million
 Session Writes: 20 million
 Pricing Model: On demand

El costo mensual estimado de este servicio es de aproximadamente 35 USD al mes.

Si utiliza ElastiCache (Redis OSS) como mecanismo de almacenamiento del estado de la sesión, utilice los siguientes parámetros:

Number of Nodes: 3
 Node size: cache.t4g.medium
 Pricing Model: 1y reserved

El costo mensual estimado de este servicio es de aproximadamente 91 USD al mes.

Recomendaciones de optimización de costos

El primer paso es implementar el estado de la sesión en una aplicación .NET heredada. Si lo utiliza ElastiCache como mecanismo de almacenamiento de estado, siga las instrucciones que aparecen en el blog de herramientas para AWS desarrolladores [ElastiCache como almacén de sesiones de](#)

[ASP.NET](#). Si utiliza DynamoDB, siga las instrucciones de [¿Qué es? de AWS SDK para .NET la documentación](#). SDK para .NET

Si la aplicación utiliza la InProcesión para empezar, asegúrese de que todos los objetos que planea almacenar en la sesión se puedan serializar. Para ello, utilice el atributo `SerializableAttribute` para decorar las clases cuyas instancias se almacenarán en la sesión. Por ejemplo:

```
[Serializable()]
public class TestSimpleObject {
    public string SessionProperty {get;set;}
}
```

Además, la `MachineKey` de .NET debe ser la misma en todos los servidores en uso. Este suele ser el caso cuando las instancias se crean a partir de una imagen de máquina de Amazon (AMI) común. Por ejemplo:

```
<machineKey
    validationKey="some long hashed value"
    decryptionKey="another long hashed value"
    validation="SHA1"/>
```

Sin embargo, es importante asegurarse de que, si se cambia una imagen base, se configure con la misma imagen de máquina .NET (ya sea a nivel de IIS o de servidor). [Para obtener más información, consulte `System.Web.Security.MachineKey` Propiedad](#) en la documentación de Microsoft.

Por último, debe definir el mecanismo para agregar servidores a un grupo de escalado automático en respuesta a un evento de escalado. Existen muchas formas de hacerlo. Recomendamos los siguientes métodos para implementar sin problemas las aplicaciones .NET Framework en una instancia de EC2 de un grupo de escalado automático:

- Utilice [Generador de imágenes de EC2](#) para configurar una AMI que contenga el servidor y la aplicación completamente configurados. A continuación, puede usar esta AMI para configurar la [plantilla de lanzamiento de su grupo de escalado automático](#).
- Se utiliza [AWS CodeDeploy](#) para implementar la aplicación. CodeDeploy permite la integración directa con [Amazon EC2 Auto Scaling](#). Esto proporciona una alternativa a la creación de una AMI nueva para cada versión de la aplicación.

Recursos adicionales

- [Create images with EC2 Image Builder](#) (documentación de Generador de imágenes de EC2)
- [Implementación de aplicaciones web.NET AWS CodeDeploy mediante Visual Studio Team Services](#) (blog sobre herramientas para AWS desarrolladores)

Uso del almacenamiento en caché para reducir la demanda de la base de datos

Descripción general de

Puede utilizar el almacenamiento en caché como una estrategia eficaz para reducir los costos de sus aplicaciones .NET. Muchas aplicaciones utilizan bases de datos en el backend, como SQL Server, cuando las aplicaciones requieren un acceso frecuente a los datos. El costo de mantener estos servicios de backend para hacer frente a la demanda puede ser elevado, pero puede utilizar una estrategia de almacenamiento en caché eficaz para reducir la carga de las bases de datos de backend reduciendo los requisitos de tamaño y escalado. Esto puede ser útil a la hora de reducir los costos y mejorar el rendimiento de sus aplicaciones.

El almacenamiento en caché es una técnica útil para ahorrar costos relacionados con las cargas de trabajo de lectura intensiva que utilizan recursos más caros, como SQL Server. Es importante utilizar la técnica adecuada para la carga de trabajo. Por ejemplo, el almacenamiento en caché local no es escalable y requiere que mantenga una caché local para cada instancia de una aplicación. Debe sopesar el impacto en el rendimiento y compararlo con los posibles costos de modo que el menor costo del origen de datos subyacente compense cualquier costo adicional relacionado con el mecanismo de almacenamiento en caché.

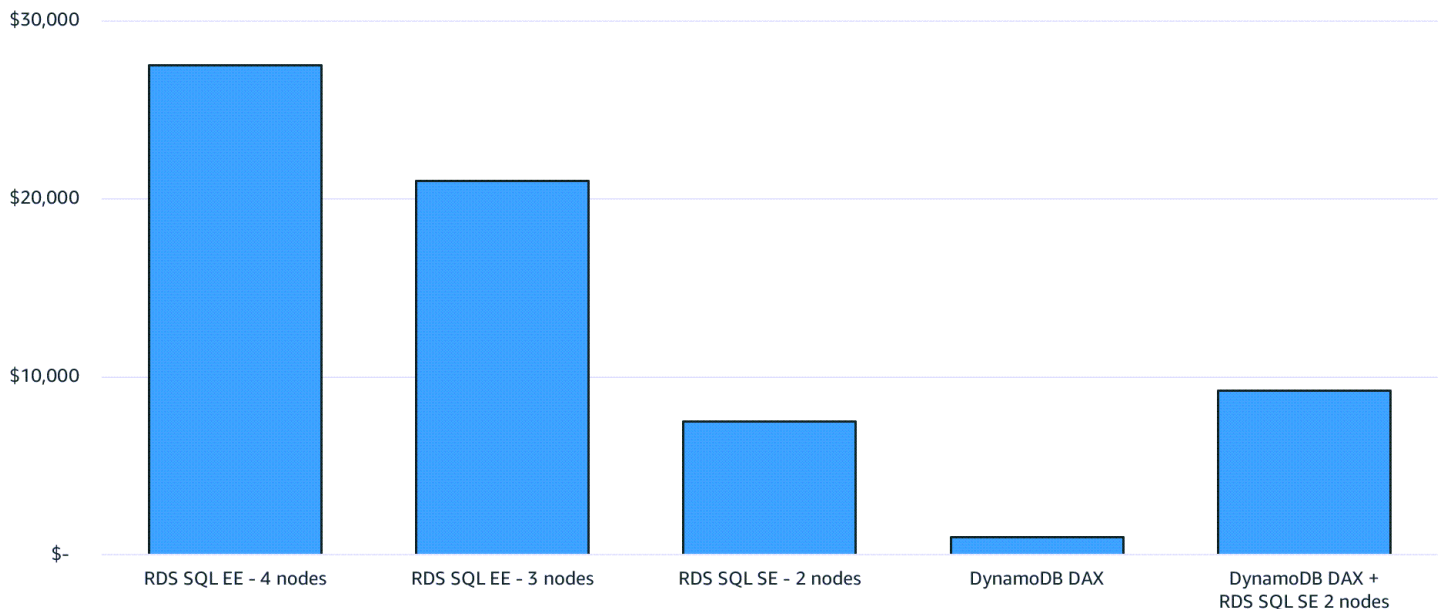
Impacto del costo

SQL Server requiere que tenga en cuenta las solicitudes de lectura al dimensionar la base de datos. Esto podría afectar a los costos, ya que es posible que tenga que introducir réplicas de lectura para adaptarse a la carga. Si utiliza réplicas de lectura, debe tener en cuenta que solo están disponibles en la edición Enterprise de SQL Server. Esta edición requiere una licencia más cara que la edición Standard de SQL Server.

El siguiente diagrama está diseñado para explicar la eficacia del almacenamiento en caché. Muestra Amazon RDS para SQL Server con cuatro nodos db.m4.2xlarge que usan la edición Enterprise de

SQL Server. Se implementa en una configuración Multi-AZ con una réplica de lectura. El tráfico de lectura exclusivo (por ejemplo, las consultas SELECT) se envía a las réplicas de lectura. En comparación, Amazon DynamoDB utiliza un clúster de Acelerador de DynamoDB (DAX) r4.2xlarge de dos nodos.

En el siguiente gráfico se muestran los resultados obtenidos al no necesitar réplicas de lectura dedicadas que gestionen un tráfico de lectura elevado.



Puede ahorrar bastante si usa el almacenamiento en caché local sin réplicas de lectura o combina el uso de DAX con SQL Server en Amazon RDS como capa de almacenamiento en caché. Esta capa se descarga de SQL Server y reduce el tamaño de los recursos de SQL Server necesarios para poner en marcha la base de datos.

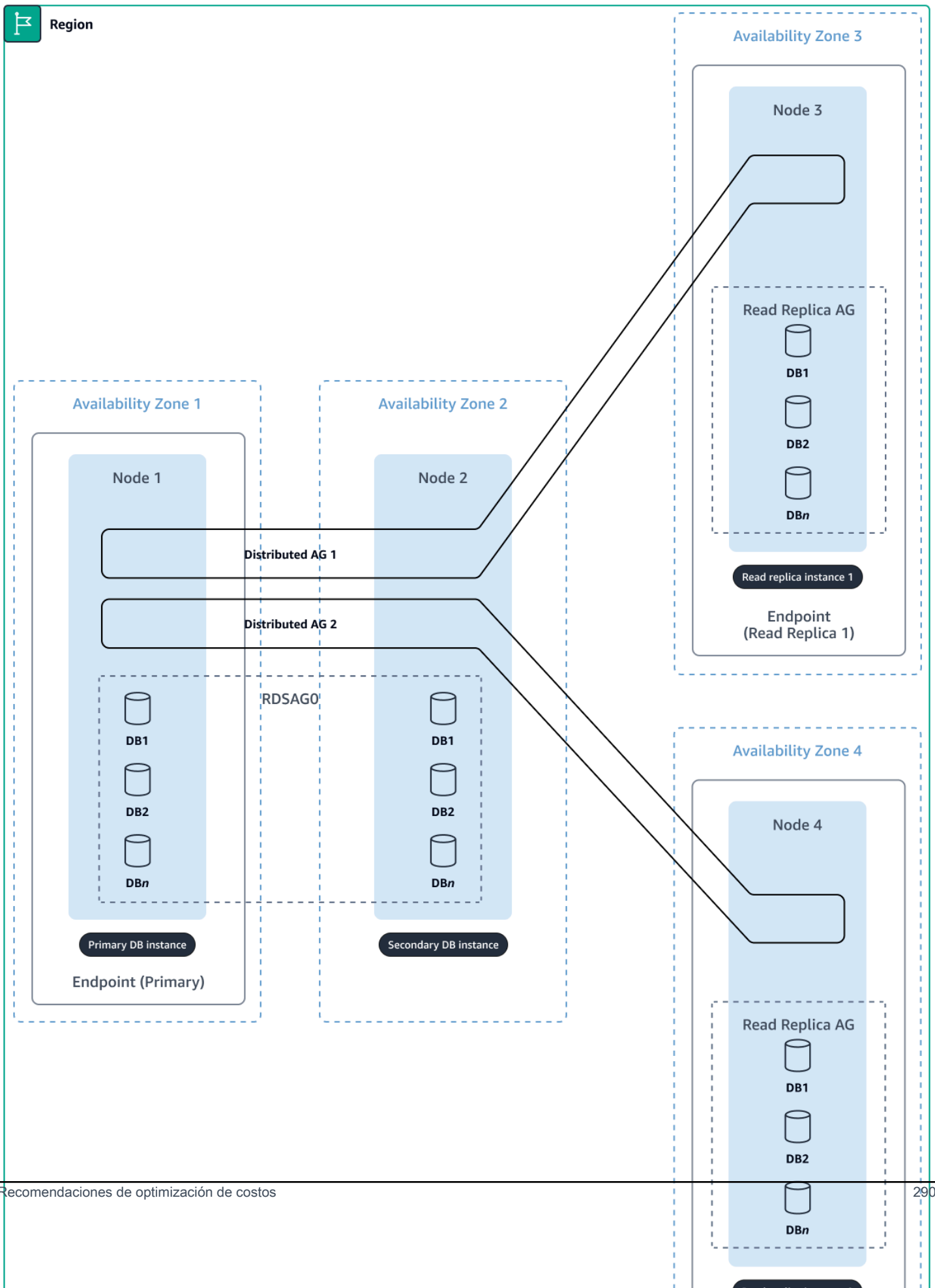
Recomendaciones de optimización de costos

Almacenamiento en la caché local

El almacenamiento en caché local es una de las formas más utilizadas para almacenar en caché el contenido de las aplicaciones alojadas tanto en las instalaciones como en la nube. Esto se debe a que su implementación es relativamente fácil e intuitiva. El almacenamiento en caché local implica tomar contenido de una base de datos u otro origen y almacenarlo en caché local en la memoria o en el disco para un acceso más rápido. Esta estrategia, aunque es fácil de implementar, no es ideal para algunos casos de uso. Por ejemplo, esto incluye casos de uso en los que el contenido almacenado en caché debe persistir en el tiempo, como cuando se requiere preservar el estado de

la aplicación o el estado del usuario. Otro caso de uso es presenta cuando es necesario acceder al contenido almacenado en caché desde otras instancias de la aplicación.

El siguiente diagrama muestra un clúster de SQL Server de alta disponibilidad con cuatro nodos y dos réplicas de lectura.



Con el almacenamiento en caché local, es posible que tenga que equilibrar la carga del tráfico entre varias instancias de EC2. Cada instancia debe mantener su propia caché local. Si la caché almacena información con estado, es necesario que se realicen confirmaciones periódicas en la base de datos y es posible que sea necesario redirigir a los usuarios a la misma instancia para cada solicitud posterior (sesión persistente). Esto supone un desafío cuando se intenta escalar las aplicaciones, ya que algunas instancias podrían estar excediendo su capacidad de uso, mientras que otras no se usan lo suficiente debido a la distribución desigual del tráfico.

Puede utilizar el almacenamiento en caché local, ya sea en memoria o mediante almacenamiento local, para las aplicaciones .NET. Para ello, puede agregar la funcionalidad de almacenar objetos en el disco y recuperarlos cuando sea necesario, o bien consultar datos de la base de datos y conservarlos en la memoria. Para realizar un almacenamiento en caché local en memoria y un almacenamiento local de los datos de SQL Server en C#, por ejemplo, puede utilizar una combinación de bibliotecas `MemoryCache` y `LiteDB`. `MemoryCache` proporciona almacenamiento en caché en memoria, mientras que `LiteDB` es una base de datos NoSQL integrada basada en disco que es rápida y ligera.

Para realizar el almacenamiento en caché en memoria, utilice la biblioteca de .NET `System.Runtime.MemoryCache`. El siguiente ejemplo de código muestra cómo utilizar la clase `System.Runtime.Caching.MemoryCache` para almacenar datos en caché en memoria. Esta clase proporciona una forma de almacenar datos temporalmente en la memoria de la aplicación. Esto puede ayudar a mejorar el rendimiento de una aplicación, pues reduce la necesidad de obtener datos de un recurso más caro, como una base de datos o una API.

El código funciona de la siguiente manera:

1. Se crea una instancia estática privada de `MemoryCache` denominada `_memoryCache`. La memoria caché recibe un nombre (`dataCache`) para identificarla. A continuación, la memoria caché almacena y recupera los datos.
2. El método `GetData` es un método genérico que utiliza dos argumentos: una clave `string` y un delegado `Func<T>` llamado `getData`. La clave se usa para identificar los datos en caché, mientras que el delegado `getData` representa la lógica de recuperación de datos que se pone en marcha cuando los datos no están presentes en la memoria caché.
3. Lo primero que hace el método es comprobar si los datos están presentes en la memoria caché mediante el método `_memoryCache.Contains(key)`. Si los datos están en la memoria caché, el método los recupera usando `_memoryCache.Get(key)` y convierte los datos en el tipo `T` esperado.

4. Si los datos no están en la memoria caché, el método llama al delegado `getData` para que busque los datos. A continuación, agrega los datos a la memoria caché usando `_memoryCache.Add(key, data, DateTimeOffset.Now.AddMinutes(10))`. Esta llamada especifica que la entrada de la memoria caché debe vencer después de 10 minutos, momento en el que los datos se eliminan automáticamente de la memoria caché.
5. El método `ClearCache` toma una clave `string` como argumento y elimina los datos asociados a esa clave de la memoria caché mediante `_memoryCache.Remove(key)`.

```
using System;
using System.Runtime.Caching;

public class InMemoryCache
{
    private static MemoryCache _memoryCache = new MemoryCache("dataCache");

    public static T GetData<T>(string key, Func<T> getData)
    {
        if (_memoryCache.Contains(key))
        {
            return (T)_memoryCache.Get(key);
        }

        T data = getData();
        _memoryCache.Add(key, data, DateTimeOffset.Now.AddMinutes(10));

        return data;
    }

    public static void ClearCache(string key)
    {
        _memoryCache.Remove(key);
    }
}
```

Puede utilizar el siguiente código:

```
public class Program
{
    public static void Main()
    {
        string cacheKey = "sample_data";
    }
}
```

```
Func<string> getSampleData = () =>
{
    // Replace this with your data retrieval logic
    return "Sample data";
};

string data = InMemoryCache.GetData(cacheKey, getSampleData);
Console.WriteLine("Data: " + data);
}
}
```

El siguiente ejemplo muestra cómo utilizar [LiteDB](#) para almacenar datos en caché en un almacenamiento local. Puede usar LiteDB como alternativa o complemento al almacenamiento en caché en memoria. El siguiente código muestra cómo utilizar la biblioteca LiteDB para almacenar datos en caché en el almacenamiento local. La clase `LocalStorageCache` contiene las funciones principales para administrar la memoria caché.

```
using System;
using LiteDB;

public class LocalStorageCache
{
    private static string _liteDbPath = @"Filename=LocalCache.db";

    public static T GetData<T>(string key, Func<T> getData)
    {
        using (var db = new LiteDatabase(_liteDbPath))
        {
            var collection = db.GetCollection<T>("cache");
            var item = collection.FindOne(Query.EQ("_id", key));

            if (item != null)
            {
                return item;
            }
        }

        T data = getData();

        using (var db = new LiteDatabase(_liteDbPath))
        {
            var collection = db.GetCollection<T>("cache");
```

```
        collection.Upsert(new BsonValue(key), data);
    }

    return data;
}

public static void ClearCache(string key)
{
    using (var db = new LiteDatabase(_liteDbPath))
    {
        var collection = db.GetCollection("cache");
        collection.Delete(key);
    }
}

public class Program
{
    public static void Main()
    {
        string cacheKey = "sample_data";

        Func<string> getSampleData = () =>
        {
            // Replace this with your data retrieval logic
            return "Sample data";
        };

        string data = LocalStorageCache.GetData(cacheKey, getSampleData);
        Console.WriteLine("Data: " + data);
    }
}
```

Si tiene una memoria caché estática o archivos estáticos que no cambian con frecuencia, también puede almacenar estos archivos en un almacén de objetos de Amazon Simple Storage Service (Amazon S3). La aplicación puede recuperar el archivo de la memoria caché estática al iniciarse para usarse localmente. Para obtener más información sobre cómo recuperar archivos de Amazon S3 mediante .NET, consulte [Descarga de objetos](#) en la documentación de Amazon S3.

Almacenamiento en caché con DAX

Puede utilizar una capa de almacenamiento en caché que se pueda compartir entre todas las instancias de la aplicación. [Acelerador de DynamoDB \(DAX\)](#) es una caché en memoria altamente disponible y completamente administrada para DynamoDB que multiplica el rendimiento por 10. Puede utilizar DAX para reducir los costos reduciendo la necesidad de asignar más unidades de capacidad de lectura de las necesarias en las tablas de DynamoDB. Esto resulta especialmente útil para cargas de trabajo con muchas solicitudes de lectura y que requieren lecturas repetidas de claves individuales.

DynamoDB cuenta con precios bajo demanda o en función de la capacidad aprovisionada, por lo que el número de lecturas y escrituras al mes se refleja en el costo. Si tiene cargas de trabajo con muchas solicitudes de lectura, los clústeres de DAX pueden ser de ayuda para ahorrar, pues reducen el número de lecturas en las tablas de DynamoDB. Para obtener instrucciones sobre cómo configurar DAX, consulte [Aceleración en memoria con Acelerador de DynamoDB \(DAX\)](#) en la documentación de DynamoDB. Para obtener información sobre la integración de aplicaciones de .NET, consulte [Integrating Amazon DynamoDB DAX en su aplicación ASP.NET en YouTube](#)

Recursos adicionales

- [Aceleración en memoria con Acelerador de DynamoDB \(DAX\): Amazon DynamoDB](#) (documentación de DynamoDB)
- [Integración de Amazon DynamoDB DAX en su aplicación ASP.NET \(\) YouTube](#)
- [Descarga de objetos](#) (documentación de Amazon S3)

Uso de .NET sin servidor

Descripción general de

La computación sin servidor se convirtió en una estrategia popular para crear e implementar aplicaciones. Esto se debe principalmente a la escalabilidad y la agilidad que ofrece la estrategia sin servidor a la hora de crear una arquitectura moderna. Sin embargo, es importante tener en cuenta el impacto en los costos de la computación sin servidor en algunos escenarios.

Lambda es una plataforma de computación sin servidor que permite a los desarrolladores usar código sin necesidad de servidores dedicados. Lambda es una opción particularmente atractiva para los desarrolladores de .NET que buscan reducir los costos asociados a la infraestructura.

Con Lambda, los desarrolladores de .NET pueden desarrollar e implementar aplicaciones que son altamente escalables y potencialmente rentables. Al utilizar una estrategia sin servidor, los desarrolladores ya no aprovisionan servidores para gestionar las solicitudes de las aplicaciones. En su lugar, los desarrolladores pueden crear funciones que se pongan en marcha bajo demanda. Esto hace que una estrategia sin servidor sea más escalable, administrable y potencialmente más rentable que usar, administrar y escalar máquinas virtuales. Como resultado, solo paga por los recursos utilizados por la aplicación, sin tener que preocuparse por los recursos que no se usan lo suficiente o los costos derivados del mantenimiento del servidor.

Los desarrolladores pueden usar versiones modernas de .NET multiplataforma para crear aplicaciones sin servidor que sean rápidas, eficientes y rentables. .NET Core y las versiones más recientes son un marco gratuito y de código abierto ideal para plataformas sin servidor que las versiones anteriores de .NET Framework. Esto permite a los desarrolladores reducir el tiempo de desarrollo y aumentar el rendimiento de las aplicaciones. Las versiones modernas de .NET también son compatibles con varios lenguajes de programación, como C# y F#. Por este motivo, es una opción atractiva para los desarrolladores que desean crear arquitecturas modernas en la nube.

En esta sección se explica cómo puede ahorrar costos usando Lambda como opción sin servidor. Para optimizar aún más los costos, puede refinar los perfiles de ejecución de sus funciones de Lambda, dimensionar correctamente la asignación de memoria de sus funciones de Lambda, utilizar la [compilación nativa anticipada](#) y comenzar a usar funciones basadas en Graviton.

Impacto del costo

La medida en que pueda reducir los costos depende de varios factores, como el número de ejecuciones que ejecutarán sus funciones sin servidor, además de la cantidad de memoria asignada y la duración de cada función. AWS Lambda ofrece una capa gratuita, que incluye un millón de solicitudes gratuitas al mes y 400 000 GB por segundo de tiempo de procesamiento al mes. Puede reducir considerablemente los costos mensuales de las cargas de trabajo que se encuentren en los límites del nivel gratuito o próximos a ellos.

El uso de un equilibrador de carga con funciones de Lambda como destino también puede suponer costos adicionales. Esto se calcula como la cantidad de datos procesados por el equilibrador de carga para los [destinos de Lambda](#).

Recomendaciones de optimización de costos

Dimensionamiento correcto de las funciones de Lambda

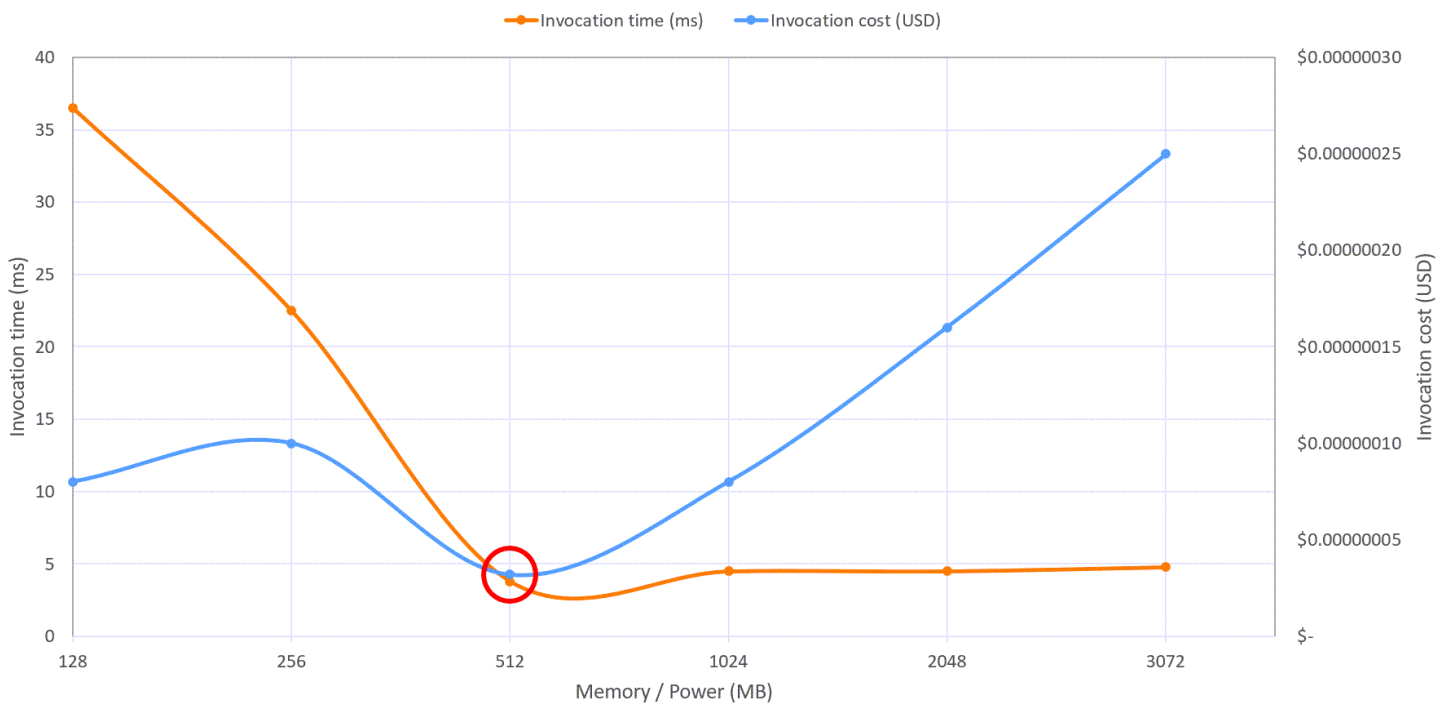
El dimensionamiento correcto es una práctica esencial para la optimización de costos en las funciones de Lambda basadas en .NET. Este proceso implica identificar la configuración de memoria óptima que proporcione un equilibrio entre rendimiento y costo, sin necesidad de cambiar el código.

Al configurar la memoria para una función de Lambda, que va desde 128 MB hasta 10 240 MB, también se ajusta la cantidad de vCPU disponible durante la invocación. Esto permite que las aplicaciones vinculadas a la memoria o a la CPU accedan a recursos adicionales durante la puesta en marcha, lo que se traduce en una posible reducción de la duración de la invocación y del costo total.

Sin embargo, identificar la configuración óptima para las funciones de Lambda basadas en .NET puede ser un proceso manual y laborioso, especialmente si los cambios son frecuentes. La herramienta [AWS Lambda Power Tuning](#) puede serle de ayuda a la hora de identificar la configuración adecuada, pues analiza un conjunto de configuraciones de memoria y las compara con un ejemplo de carga útil.

Por ejemplo, aumentar la memoria de una función de Lambda basada en .NET puede mejorar el tiempo total de invocación y reducir los costos sin repercutir en el rendimiento. La configuración de memoria óptima de una función puede variar. La herramienta AWS Lambda Power Tuning puede ayudar a identificar la configuración más rentable para cada función.

En el siguiente gráfico de ejemplo, el tiempo total de invocación mejora a medida que aumenta la memoria de esta función de Lambda. Esto se traduce en una reducción del costo total de la puesta en marcha sin que ello repercuta en el rendimiento original de la función. En esta función, la configuración de memoria óptima es de 512 MB, ya que esta opción permite una utilización de los recursos más eficiente teniendo en cuenta el costo total de cada invocación. Esto varía según la función y, al usar la herramienta en sus funciones de Lambda, puede averiguar si el dimensionamiento correcto genera beneficios.



Le recomendamos que complete este ejercicio con regularidad como parte de cualquier prueba de integración cuando se publiquen nuevas actualizaciones. Si se actualiza con poca frecuencia, realice este ejercicio periódicamente para asegurarse de que las funciones estén ajustadas y tengan el tamaño adecuado. Cuando haya identificado la configuración de memoria adecuada de las funciones de Lambda, puede agregar el dimensionamiento correcto a sus procesos. La herramienta AWS Lambda Power Tuning genera resultados programáticos que pueden utilizar sus flujos de trabajo de CI/CD durante la publicación del nuevo código. Esto le permite automatizar la configuración de la memoria.

Puede descargar la herramienta [AWS Lambda Power Tuning](#) de forma gratuita. Para obtener instrucciones sobre cómo utilizar la herramienta, consulte [Cómo ejecutar la máquina de estados en GitHub](#)

Lambda también admite la compilación nativa anticipada, lo que permite precompilar las aplicaciones .NET. Reducir los tiempos de ejecución de las funciones .NET puede generar ahorro de costos. Para obtener más información sobre la creación de funciones con compilación nativa anticipada, consulte [Funciones .NET con compilación nativa anticipada](#) en la documentación de Lambda.

Evitar el tiempo de espera de inactividad

La duración de la función de Lambda es una dimensión que se utiliza para calcular la facturación. Cuando el código de la función realiza una llamada de bloqueo, se le facturará el tiempo que espere

hasta recibir una respuesta. Este tiempo de espera puede aumentar cuando las funciones de Lambda están encadenadas o una función actúa como orquestadora de otras funciones. Si tiene flujos de trabajo como operaciones por lotes o sistemas de entrega de pedidos, esto agrega una sobrecarga de administración. Además, puede que no sea posible completar toda la lógica del flujo de trabajo ni la gestión de errores dentro del tiempo de espera máximo de Lambda de 15 minutos.

En lugar de utilizar esta lógica en el código de la función, le recomendamos que rediseñe la solución para utilizar [AWS Step Functions](#) como orquestador del flujo de trabajo. Cuando utilice un flujo de trabajo estándar, se le facturará por cada transición de [estado](#) del flujo de trabajo y no por la duración total del flujo de trabajo. Además, puede es posible gestionar los reintentos, las condiciones de espera, los flujos de error y las funciones de [callback](#) dentro del estado, permitiendo así que las funciones de Lambda se dediquen únicamente a la lógica empresarial. Para obtener más información, [consulte Optimización de AWS Lambda los costes, segunda parte](#), en el blog de AWS informática.

Uso de funciones basadas en Graviton

Las funciones de Lambda impulsadas por los procesadores Graviton2 de última generación ya están disponibles de forma general. Las funciones Graviton2, que utilizan una arquitectura de procesador basada en ARM, están diseñadas para ofrecer hasta un 19 % más de rendimiento a un costo un 20 % menor para varias cargas de trabajo sin servidor. Con una latencia más baja y un mejor rendimiento, las funciones impulsadas por los procesadores Graviton2 son ideales para poner en marcha aplicaciones esenciales sin servidor.

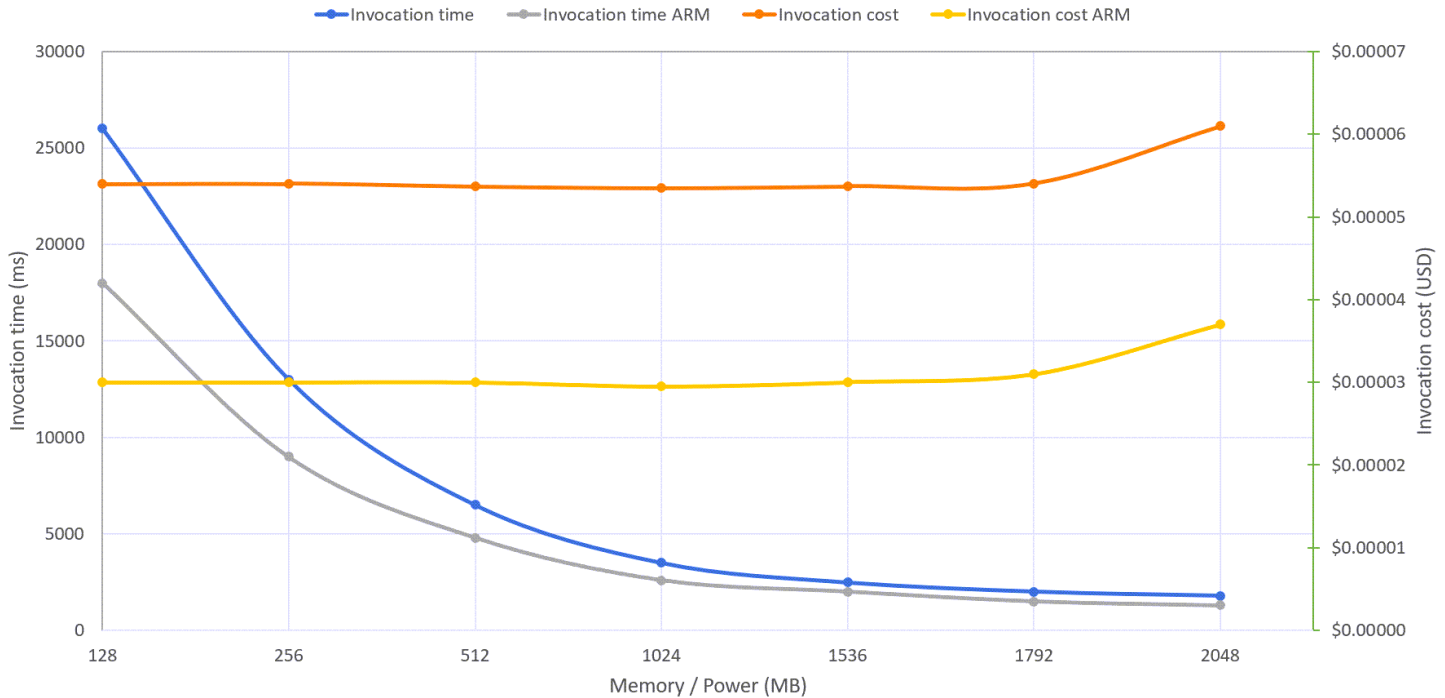
La migración a funciones de Lambda basadas en Graviton puede ser una opción rentable para los desarrolladores de .NET que busquen optimizar los costos asociados al uso de Lambda. Las funciones basadas en Graviton utilizan procesadores basados en ARM en lugar de los procesadores x86 tradicionales. Esto puede suponer un importante ahorro de costos sin sacrificar el rendimiento.

Si bien empezar a usar funciones basadas en Graviton tiene varias ventajas, también hay varios desafíos y consideraciones que le recomendamos que tenga en cuenta. Por ejemplo, las funciones basadas en Graviton requieren el uso de Amazon Linux 2, que puede no ser compatible con todas las aplicaciones .NET. Además, es posible que haya problemas de compatibilidad con bibliotecas o dependencias de terceros que no sean compatibles con los procesadores basados en ARM.

Si tiene en marcha aplicaciones de .NET Framework y desea aprovechar las ventajas de la tecnología sin servidor con Lambda, puede considerar la posibilidad de migrar las aplicaciones a una versión moderna de .NET mediante [Asistente de portabilidad para .NET](#). Esto le permite acelerar la

migración de aplicaciones .NET heredadas a las versiones modernas de .NET, lo que permitirá que la aplicación funcione en Linux.

En la siguiente tabla se comparan los resultados de las arquitecturas x86 y ARM/Graviton2 de una función que calcula números primos.



La función utiliza un único subproceso. La duración más baja para ambas arquitecturas se indica cuando la memoria es de 1,8 GB. En los casos en los que la memoria es superior, las funciones de Lambda tienen acceso a más de 1 vCPU, pero, en este caso, la función no puede utilizar la potencia adicional. Por este mismo motivo, los costos se mantienen estables con una memoria de hasta 1,8 GB. Con más memoria, los costos aumentan porque esta carga de trabajo no supone ningún beneficio de rendimiento adicional. Es evidente que el procesador Graviton2 proporciona un mejor rendimiento y unos costos más bajos para esta función que requiere un uso intensivo de recursos de computación.

Para que la función utilice un procesador basado en ARM con Graviton, haga lo siguiente:

1. Inicie sesión en la Consola de administración de AWS consola [Lambda](#) y ábrala.
2. Seleccione Creación de función.
3. En Function name (Nombre de función), escriba un nombre.
4. Para Runtime, elija .NET 6 (C#/PowerShell).
5. En Arquitectura, seleccione amd64.

6. Realice las configuraciones adicionales que sean necesarias y, a continuación, elija Crear función.

Recursos adicionales

- [Funciones Lambda como objetivos \(documentación\)](#) AWS
- [Optimización de los AWS Lambda costes y el rendimiento mediante AWS Compute Optimizer](#) (AWS Compute Blog)
- [Optimización de AWS Lambda los costes: primera parte](#) (blog AWS sobre informática)
- [Optimización de AWS Lambda los costes: segunda parte](#) (blog AWS sobre informática)
- [Creación de aplicaciones.NET sin servidor AWS Lambda mediante .NET 7](#) (AWS Compute Blog)

Uso de bases de datos personalizadas

Descripción general de

Uno de los aspectos más costosos de la puesta en marcha de cargas de trabajo basadas en Microsoft es el uso de licencias de bases de datos comerciales, como SQL Server. Las empresas suelen estandarizar SQL Server como la plataforma de base de datos preferida y esto queda arraigado en la cultura de desarrollo de la organización. Los desarrolladores suelen elegir un modelo relacional basado en SQL Server, independientemente del caso de uso. Los motivos pueden ser:

- La empresa ya tiene and/or licencias de instancias de SQL Server disponibles.
- Los equipos se han acostumbrado al modelo de programación SQL mediante el uso de bibliotecas compartidas y la lógica empresarial. ORMs
- La gerencia no conoce las alternativas.
- Los desarrolladores no conocen las alternativas.

Las bases de datos personalizadas pueden adaptarse a los patrones de acceso a los datos de su caso de uso. Cada vez más empresas adoptan estas bases de datos conforme avanzan hacia arquitecturas modernas (como los microservicios) y el alcance de cada aplicación se vuelve más limitado.

Una base de datos personalizada no excluye un modelo relacional ni requiere un modelo NoSQL (no relacional). De hecho, una base de datos relacional se considera personalizada cuando se selecciona en respuesta a las necesidades específicas de una carga de trabajo. El uso de

bases de datos personalizadas puede ayudar a los equipos a reducir los costos asociados a sus aplicaciones .NET y, al mismo tiempo, obtener las ventajas estándar de la nube, como la escalabilidad, la resiliencia y la reducción del trabajo pesado que no aporta valor.

La siguiente tabla muestra las bases de datos especialmente diseñadas que ofrece. AWS

Base de datos	Tipo	Características
Amazon Aurora PostgreSQL o Amazon Aurora MySQL	Relacional	<p>Casos de uso en los que los datos tienen una estructura fija</p> <p>Las bases de datos relacionales mantienen de forma natural la coherencia de datos en las transacciones ACID.</p>
Amazon DynamoDB	Par clave-valor	<p>Base de datos NoSQL que almacena datos mediante una estructura de datos de tabla hash.</p> <p>Almacenamiento y recuperación de datos no estructurados de alto rendimiento.</p> <p>Algunos casos de uso son los perfiles de usuario, el estado de la sesión y los datos del carrito de compras.</p>
Amazon ElastiCache	En memoria	<p>Base de datos NoSQL de alto rendimiento que almacena datos no estructurados en la memoria con un tiempo de acceso inferior a un milisegundo.</p> <p>Se utiliza para datos efímeros a los que se accede con</p>

Base de datos	Tipo	Características
		<p>frecuencia, como sesiones de usuario, y como capa de almacenamiento en caché delante de otros almacenes de datos más lentos.</p> <p>Incluye soporte para ElastiCache (Redis OSS) y ElastiCache (Memcached)</p>
Amazon MemoryDB	En memoria con durabilidad	Base de datos personalizada compatible con Redis con almacenamiento duradero.
Amazon Timestream	Serie temporal	<p>Base de datos diseñada para la ingesta de datos de alto rendimiento en orden temporal.</p> <p>Algunos casos de uso son las aplicaciones de Internet de las cosas (IoT) y el almacenamiento de métricas o datos de telemetría.</p>
Amazon DocumentDB	Documento	<p>Base de datos NoSQL que almacena datos sin una estructura prescrita ni relaciones forzadas con otros datos.</p> <p>Suele utilizarse para las cargas de trabajo con muchas solicitudes de lectura, como los catálogos de productos.</p>

Base de datos	Tipo	Características
Amazon Neptune	Gráfico	<p>Base de datos NoSQL que contiene datos y una representación de las conexiones entre los elementos de datos.</p> <p>Algunos casos de uso son la detección de fraudes, los motores de recomendación y las aplicaciones de redes sociales.</p>
Amazon Keyspaces	Columna ancha	<p>Base de datos distribuida de alto rendimiento basada en Apache Cassandra.</p> <p>Algunos casos de uso son las aplicaciones de IoT, el procesamiento de eventos y las aplicaciones de juegos.</p>

Uno de los principales impulsores de la adopción de bases de datos personalizadas es la eliminación de las licencias comerciales. Sin embargo, la capacidad de escalado automático de bases de datos como [DynamoDB](#) (lo que abarca el [modo bajo demanda](#)), [Aurora](#), [Amazon Neptune](#) y [Amazon Keyspaces](#) le permite aprovisionar capacidad para casos normales en lugar de para picos de uso. Las bases de datos personalizadas, como Timestream, no tienen servidor y se escalan automáticamente para satisfacer la demanda sin necesidad de hacer un aprovisionamiento previo.

AWS ofrece [Babelfish para Aurora PostgreSQL](#) si desea utilizar una base de datos relacional especialmente diseñada y compatible con código abierto, pero no puede o no quiere realizar cambios significativos en el código de su aplicación. En algunos casos, Babelfish le permite utilizar código de acceso a SQL Server existente, prácticamente sin cambios.

Al elegir una base de datos relacional personalizada para sus aplicaciones, es importante retener las mismas características (o funcionalmente equivalentes) que necesita. Esta recomendación aborda

las bases de datos personalizadas como almacén de datos principal para las aplicaciones. Ciertas aplicaciones, como el almacenamiento en caché, se tratan en otras recomendaciones.

Impacto del costo

Si bien es poco probable que la adopción de bases de datos diseñadas específicamente para las cargas de trabajo de .NET afecte directamente a la computación, puede influir en el consumo/costo directamente en el costo de los servicios de bases de datos que consumen las aplicaciones .NET. De hecho, el ahorro de costos puede ser un objetivo secundario si se compara con los beneficios adicionales de agilidad, escalabilidad, resiliencia y durabilidad de los datos.

En esta guía no se explica el proceso completo de elección de una base de datos personalizada para las aplicaciones ni de rediseño de una estrategia de datos para utilizarlas de forma eficaz. Para obtener más información, consulte [Purpose-built databases](#) en el directorio de tutoriales de AWS .

En las tablas siguientes se muestran varios ejemplos de cómo la sustitución de SQL Server por una base de datos personalizada puede alterar los costos de las aplicaciones. Tenga en cuenta que se trata simplemente de estimaciones aproximadas. Se requieren valores de referencia y la optimización de las cargas de trabajo reales para calcular el costo de producción exacto.

Estas son algunas estimaciones de bases de datos personalizadas que incluyen computación bajo demanda y SSD de 100 GB para bases de datos de instancia única en us-east-1. Los costos asociados al uso de licencias contemplan la licencia de SQL Server más la garantía del software.

En la siguiente tabla se muestran los costos estimados de los ejemplos de bases de datos comerciales:

Motor de base de datos	Modelo de licencia	Tipo instancia /especificaciones	AWS coste de cómputo más almacenamiento	Costo de la licencia	Costo total mensual
Edición Standard de SQL Server en Amazon EC2	Licencia incluida	r6i.2xlarge (8 CPU/64 GB de RAM)	1.345,36 DÓLARES	0,00\$	1.345,36 DÓLARES

Motor de base de datos	Modelo de licencia	Tipo instancia /especificaciones	AWS coste de cómputo más almacenamiento	Costo de la licencia	Costo total mensual
Edición Enterprise de SQL Server en Amazon EC2	Licencia incluida	r6i.2xlarge (8 CPU/64 GB de RAM)	2.834,56 DÓLARES	0,00\$	2.834,56 DÓLARES
Edición Standard de SQL Server en Amazon EC2	BYOL	r6i.2xlarge (8 CPU/64 GB de RAM)	644,56 DÓLARES	456,00 DÓLARES	1.100,56 DÓLARES
Edición Enterprise de SQL Server en Amazon EC2	BYOL	r6i.2xlarge (8 CPU/64 GB de RAM)	644,56 DÓLARES	1.750,00 DÓLARES	2.394,56 DÓLARES
Edición Standard de SQL Server en Amazon RDS		db.r6i.2xlarge (8 CPU/64 GB de RAM)	2.318,30 DÓLARES	0,00\$	2.318,30 DÓLARES
Edición Enterprise de SQL Server en Amazon RDS		db.r6i.2xlarge (8 CPU/64 GB de RAM)	3.750,56 DÓLARES	0,00\$	3.750,56 DÓLARES

En la siguiente tabla se muestran los costos estimados de los ejemplos de bases de datos personalizadas:

Motor de base de datos	Tipo instancia/ especificaciones	AWS coste de cómputo más almacenamiento	Costo de la licencia	Costo total mensual
PostgreSQL de Amazon Aurora	r6g.2xlarge (8 CPU/64 GB de RAM)	855,87 DÓLARES	0,00\$	855,87 DÓLARES
DynamoDB	Base aprovisio nada: 100 WCU/400 RCU	72,00 DÓLARES		72,00 DÓLARES
Amazon DocumentDB	db.r6i.2xlarge (8 CPU/64 GB de RAM)	778,60 DÓLARES		778,60 DÓLARES

Important

La tabla se basa en los costos estimados del uso de licencias de SQL Server con Software Assurance durante los tres primeros años de la compra. En el caso de la edición Standard de SQL Server: 4100 USD, paquete de 2 núcleos, 3 años. En el caso de la edición Enterprise de SQL Server: 15 700 USD, paquete de 2 núcleos, 3 años.

Le recomendamos que considere los posibles costos antes de adoptar bases de datos personalizadas. Por ejemplo, el costo de actualizar las aplicaciones para usar una base de datos personalizada varía según la complejidad de la aplicación y de la base de datos de origen. Asegúrese de tener en cuenta el costo total de propiedad al planificar este cambio de arquitectura. Esto incluye refactorizar las aplicaciones, formar al personal en nuevas tecnologías y planificar cuidadosamente el rendimiento y el consumo previstos de cada carga de trabajo. A partir de ahí, puede determinar si la inversión compensa el ahorro de costos. En la mayoría de los casos, el mantenimiento de un end-of-support producto supone un riesgo para la seguridad y la conformidad, y el coste de remediarlo vale la pena tanto el esfuerzo como la inversión inicial.

Recomendaciones de optimización de costos

En lo que respecta a las aplicaciones .NET que acceden a SQL Server, existen bibliotecas que reemplazan las bases de datos relacionales personalizadas. Puede implementar estas bibliotecas en su aplicación para reemplazar una funcionalidad similar de una aplicación de SQL Server.

En la siguiente tabla se muestran algunas bibliotecas que se pueden usar en muchos escenarios comunes.

Library	Base de datos	Reemplazo de	Marco compatible
Npgsql Entity Framework Core Provider	PostgreSQL de Amazon Aurora	Entity Framework Core SQL Server Provider	Versiones modernas de .NET
Npgsql Entity Framework 6 Provider	PostgreSQL de Amazon Aurora	Entity Framework 6,0 SQL Server Provider	.NET Framework
Npgsql (biblioteca PostgreSQL compatible con ADO.NET)	PostgreSQL de Amazon Aurora	ADO.NET	.NET .NET Framework/Modern
MySQL Entity Framework Core Provider	Amazon Aurora MySQL	Entity Framework Core SQL Server Provider	Versiones modernas de .NET
Pomelo.EntityFrameworkCore.MySql	Amazon Aurora MySQL	Entity Framework Core SQL Server Provider	Versiones modernas de .NET

La [conexión a Amazon Aurora PostgreSQL con Babelfish](#) no requiere ningún tipo de código especial. Sin embargo, todo el código debe probarse minuciosamente antes de su uso.

Otras bases de datos personalizadas tienen bibliotecas para acceder a bibliotecas compatibles con .NET que le permiten acceder a bases de datos personalizadas. Entre los ejemplos se incluyen:

- [Uso de bases de datos NoSQL de Amazon DynamoDB \(documentación\)](#) AWS SDK para .NET

- [MongoDB C# Driver](#) (documentación de MongoDB)
- [.NET](#) (documentación de Timestream)
- [Using a Cassandra .NET Core client driver to access Amazon Keyspaces programmatically](#) (documentación de Amazon Keyspaces)
- [Using .NET to connect to a Neptune DB instance](#) (documentación de Neptune)

Si migra a bases de datos diseñadas específicamente, puede utilizar estas herramientas AWS para facilitar el proceso de migración:

- [AWS Schema Conversion Tool \(AWS SCT\)](#) puede ser útil para transformar los esquemas de SQL Server para Amazon Aurora y Amazon DynamoDB.
- [AWS Database Migration Service \(AWS DMS\)](#) puede ser útil para migrar datos, de forma puntual o continua, de SQL Server a Aurora o DynamoDB.
- [Babelfish Compass](#) puede ser útil para comprobar la compatibilidad de su base de datos de SQL Server para utilizarla con Babelfish para Aurora PostgreSQL.

Recursos adicionales

- [Guidance for migrating SQL Server to Amazon Aurora PostgreSQL](#) (blog de bases de datos de AWS)
- [Jornada de inmersión en la modernización de las aplicaciones de Babelfish](#) (Workshop Studio)AWS
- [Día de inmersión en .NET](#) (AWS taller de estudio)
- [Cómo empezar a usar Amazon Timestream con .NET](#) () GitHub
- [Bases de datos especialmente diseñadas para aplicaciones de .NET modernas](#) en (presentación) AWSAWS

Pasos a seguir a continuación

Cuando termine de revisar esta guía, recomendamos que siga estos pasos para implementar MACO:

1. Contacte con un experto en MACO. Un experto en MACO puede responder sus preguntas y solucionar sus dudas. Si ya trabaja con un equipo de cuentas de AWS, contacte con este y solicite la ayuda de un experto en MACO. Si no tienes un equipo de cuentas, contacte con optimize-microsoft@amazon.com.
2. Aplique las recomendaciones. Aplique las recomendaciones, las prácticas recomendadas y las estrategias que aprendió en esta guía y tras hablar con un experto en MACO.
3. Haga un seguimiento de los cambios en los costos. Etiquete sus cargas de trabajo y utilice servicios como AWS Cost Explorer y AWS Budgets para hacer un seguimiento, una supervisión y un control detallados de los costos.

Historial de documentos

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

Cambio	Descripción	Fecha
Actualizaciones de SQL Server	Hemos actualizado la sección Optimizar la CPU para las cargas de trabajo de SQL Server para añadir más información sobre la función Optimizar la CPU para las EC2 instancias de Amazon .	22 de octubre de 2025
Actualizaciones de SQL Server	Actualizamos la sección Optimización de la CPU para las cargas de trabajo de SQL Server .	25 de octubre de 2024
Actualizaciones de SQL Server y Contenedores	Añadimos las secciones Optimizar el tamaño de SQL Server mediante Compute Optimizer , Trusted Advisor Revisar las recomendaciones para las cargas de trabajo de SQL Server y Cambiar la plataforma de las aplicaciones de Windows con App2Container .	29 de junio de 2024
Optimización de las licencias de SQL Server	Agregamos la sección Optimización de las licencias de SQL Server con Compute Optimizer .	22 de mayo de 2024
Publicación inicial	—	21 de diciembre de 2023

AWS Glosario de orientación prescriptiva

Los siguientes son términos de uso común en las estrategias, guías y patrones proporcionados por la Guía AWS prescriptiva. Para sugerir entradas, utilice el enlace [Enviar comentarios](#) al final del glosario.

Números

Las 7 R

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

- **Refactorizar/rediseñar:** traslade una aplicación y modifique su arquitectura mediante el máximo aprovechamiento de las características nativas en la nube para mejorar la agilidad, el rendimiento y la escalabilidad. Por lo general, esto implica trasladar el sistema operativo y la base de datos. Ejemplo: Migrar la base de datos de Oracle en las instalaciones a Amazon Aurora PostgreSQL-Compatible Edition.
- **Redefinir la plataforma (transportar y redefinir):** traslade una aplicación a la nube e introduzca algún nivel de optimización para aprovechar las capacidades de la nube. Ejemplo: Migrar la base de datos Oracle en las instalaciones a Amazon Relational Database Service (Amazon RDS) para Oracle en la nube de Nube de AWS.
- **Recomprar (readquirir):** cambie a un producto diferente, lo cual se suele llevar a cabo al pasar de una licencia tradicional a un modelo SaaS. Ejemplo: Migrar el sistema de administración de las relaciones con los clientes (CRM) a Salesforce.com.
- **Volver a alojar (migrar mediante lift-and-shift):** traslade una aplicación a la nube sin realizar cambios para aprovechar las capacidades de la nube. Ejemplo: Migrar la base de datos de Oracle en las instalaciones a Oracle en una instancia de EC2 en la Nube de AWS.
- **Reubicar:** (migrar el hipervisor mediante lift and shift): traslade la infraestructura a la nube sin comprar equipo nuevo, reescribir aplicaciones o modificar las operaciones actuales. Los servidores se migran de una plataforma en las instalaciones a un servicio en la nube para la misma plataforma. Ejemplo: migrar una Microsoft Hyper-V aplicación a AWS.
- **Retener (revisitar):** conserve las aplicaciones en el entorno de origen. Estas pueden incluir las aplicaciones que requieren una refactorización importante, que desee posponer para más adelante, y las aplicaciones heredadas que desee retener, ya que no hay ninguna justificación empresarial para migrarlas.

- Retirar: retire o elimine las aplicaciones que ya no sean necesarias en un entorno de origen.

A

ABAC

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

servicios abstractos

Consulte [servicios administrados](#).

ACID

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

migración activa-activa

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

migración activa-pasiva

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

función de agregación

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

IA

Consulte [inteligencia artificial](#).

AIOps

Consulte [operaciones de inteligencia artificial](#)

anonimización

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

antipatrones

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

control de aplicaciones

Enfoque de seguridad que permite usar de manera exclusiva aplicaciones aprobadas para ayudar a proteger un sistema contra el malware.

cartera de aplicaciones

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

inteligencia artificial (IA)

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

operaciones de inteligencia artificial (AIOps)

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

cifrado asimétrico

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

atomicidad, consistencia, aislamiento, durabilidad (ACID)

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

control de acceso basado en atributos (ABAC)

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

origen de datos fidedigno

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

Zona de disponibilidad

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

AWS Marco de adopción de la nube (AWS CAF)

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

AWS Marco de calificación de la carga de trabajo (AWS WQF)

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

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

B

bot malicioso

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

BCP

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

gráfico de comportamiento

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

sistema big-endian

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

clasificación binaria

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

filtro de floración

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

implementación azul/verde

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

bot

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

botnet

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

branch

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

acceso de emergencia

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

estrategia de implementación sobre infraestructura existente

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

caché de búfer

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

capacidad empresarial

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

planificación de la continuidad del negocio (BCP)

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

C

CAF

Consulte [AWS Cloud Adoption Framework](#).

implementación canario

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

CCoE

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

CDC

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

captura de datos de cambio (CDC)

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

ingeniería del caos

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

CI/CD

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

clasificación

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

cifrado del cliente

Cifrado de datos localmente, antes de que el objetivo los Servicio de AWS reciba.

Centro de excelencia en la nube (CCoE)

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

computación en la nube

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

modelo operativo en la nube

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

etapas de adopción de la nube

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

- Proyecto: ejecución de algunos proyectos relacionados con la nube con fines de prueba de concepto y aprendizaje
- Fundamento: realizar inversiones fundamentales para escalar su adopción de la nube (p. ej., crear una landing zone, definir una CCoE, establecer un modelo de operaciones)

- Migración: migración de aplicaciones individuales
- Reinención: optimización de productos y servicios e innovación en la nube

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

CMDB

Consulte [base de datos de administración de configuración](#).

repositorio de código

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

caché en frío

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

datos fríos

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

visión artificial (CV)

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

deriva de configuración

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

base de datos de administración de configuración (CMDB)

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

paquete de conformidad

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

integración y entrega continuas (CI/CD)

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

CV

Consulte [visión artificial](#).

D

datos en reposo

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

clasificación de datos

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

deriva de datos

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

datos en tránsito

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

mallado de datos

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

minimización de datos

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

perímetro de datos

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

preprocesamiento de datos

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

procedencia de los datos

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

titular de los datos

Persona cuyos datos se recopilan y procesan.

almacenamiento de datos

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

lenguaje de definición de datos (DDL)

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

lenguaje de manipulación de datos (DML)

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

DDL

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

conjunto profundo

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

aprendizaje profundo

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

defense-in-depth

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

administrador delegado

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

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

Implementación

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

entorno de desarrollo

Consulte [entorno](#).

control de detección

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

asignación de flujos de valor para el desarrollo (DVSM)

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

gemelo digital

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

tabla de dimensiones

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

desastre

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

recuperación de desastres (DR)

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

DML

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

diseño basado en el dominio

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

DR

Consulte [recuperación ante desastres](#).

Detección de desviaciones

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

DVSM

Consulte [asignación de flujos de valor para el desarrollo](#).

E

EDA

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

EDI

Consulte [intercambio electrónico de datos](#).

computación en la periferia

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

intercambio electrónico de datos (EDI)

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

cifrado

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

clave de cifrado

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

endianidad

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

punto de conexión

Consulte [punto de conexión de servicio](#).

servicio de punto de conexión

Servicio que puede alojar en una nube privada virtual (VPC) para compartir con otros usuarios. Puede crear un servicio de punto final AWS PrivateLink y conceder permisos a otras Cuentas de AWS o a responsables AWS Identity and Access Management (de IAM). Estas cuentas o

entidades principales pueden conectarse a su servicio de punto de conexión de forma privada mediante la creación de puntos de conexión de VPC de interfaz. Para obtener más información, consulte [Creación de un servicio de punto de conexión](#) en la documentación de Amazon Virtual Private Cloud (Amazon VPC).

planificación de recursos empresariales (ERP)

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

cifrado de sobre

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

entorno

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

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

epopeya

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

ERP

Consulte [planificación de recursos empresariales](#).

análisis de datos de tipo exploratorio (EDA)

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

F

tabla de hechos

Tabla central de un [esquema en estrella](#). Almacena datos cuantitativos sobre operaciones empresariales. Por lo general, una tabla de hechos contiene dos tipos de columnas: las que contienen medidas y las que contienen una clave externa para una tabla de dimensiones.

Fail Fast

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

límite de aislamiento de errores

En el Nube de AWS, un límite, como una zona de disponibilidad Región de AWS, un plano de control o un plano de datos, que limita el efecto de una falla y ayuda a mejorar la resiliencia de las cargas de trabajo. Para más información, consulte [AWS Fault Isolation Boundaries](#).

rama de característica

Consulte [rama](#).

características

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

importancia de las características

La importancia que tiene una característica para las predicciones de un modelo. Por lo general, esto se expresa como una puntuación numérica que se puede calcular mediante diversas

técnicas, como las explicaciones aditivas de Shapley (SHAP) y los gradientes integrados. Para obtener más información, consulte [Interpretabilidad del modelo de aprendizaje automático](#) con AWS

transformación de funciones

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

peticiones con pocos pasos

Proporcionar a un [LLM](#) una pequeña cantidad de ejemplos que demuestren la tarea y el resultado deseado antes de pedirle que lleve a cabo una tarea similar. Esta técnica es una aplicación del aprendizaje contextual, mediante el que los modelos aprenden a partir de ejemplos (pasos) incrustados en las peticiones. La técnica de peticiones con pocos pasos puede ser eficaz para las tareas que requieren un formato, un razonamiento o un conocimiento del dominio específicos. Consulte también [peticiones desde cero](#).

FGAC

Consulte [control de acceso detallado](#).

control de acceso preciso (FGAC)

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

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

FM

Consulte [modelo fundacional](#).

Modelo fundacional (FM)

Una gran red neuronal de aprendizaje profundo que se ha estado entrenando con conjuntos de datos masivos de datos generalizados y sin etiquetar. FMs son capaces de realizar una amplia variedad de tareas generales, como comprender el lenguaje, generar texto e imágenes

y conversar en lenguaje natural. Para más información, consulte [¿Qué son los modelos fundacionales?](#)

G

IA generativa

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

bloqueo geográfico

Consulte [restricciones geográficas](#).

restricciones geográficas (bloqueo geográfico)

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

Flujo de trabajo de Gitflow

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

imagen dorada

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

estrategia de implementación desde cero

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

barrera de protección

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

H

HA

Consulte [alta disponibilidad](#).

migración heterogénea de bases de datos

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

alta disponibilidad (HA)

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

modernización histórica

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

datos de reserva

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

migración homogénea de bases de datos

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

datos recientes

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

hotfix

Una solución urgente para un problema crítico en un entorno de producción. Debido a su urgencia, una revisión suele realizarse fuera del flujo de trabajo de DevOps publicación típico.

periodo de hiperatención

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

I

IaC

Consulte [infraestructura como código](#).

políticas basadas en identidades

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

aplicación inactiva

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

IloT

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

infraestructura inmutable

Modelo que implementa una nueva infraestructura para las cargas de trabajo de producción en lugar de actualizar o modificar la infraestructura existente o aplicarle revisiones. Las infraestructuras inmutables son de manera intrínseca más coherentes, fiables y predecibles que las [infraestructuras mutables](#). Para más información, consulte la práctica recomendada [Implementación mediante una infraestructura inmutable](#) en el Marco de AWS Well-Architected.

VPC entrante (de entrada)

En una arquitectura de AWS cuentas múltiples, una VPC que acepta, inspecciona y enruta las conexiones de red desde fuera de una aplicación. La [arquitectura AWS de referencia de seguridad](#) recomienda configurar la cuenta de red con entradas, salidas e inspección VPCs para proteger la interfaz bidireccional entre la aplicación y el resto de Internet.

migración gradual

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

Industria 4.0

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

infraestructura

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

infraestructura como código (IaC)

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

Internet de las cosas industrial (T) Ilo

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

VPC de inspección

En una arquitectura de AWS cuentas múltiples, una VPC centralizada que gestiona las inspecciones del tráfico de red VPCs entre Internet y las redes locales (en una misma o Regiones de AWS diferente). La [arquitectura AWS de referencia de seguridad](#) recomienda configurar su cuenta de red con entrada, salida e inspección VPCs para proteger la interfaz bidireccional entre la aplicación e Internet en general.

Internet de las cosas (IoT)

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

interpretabilidad

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

IoT

Consulte [Internet de las cosas](#).

biblioteca de información de TI (ITIL)

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

administración de servicios de TI (ITSM)

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

ITIL

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

ITSM

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

L

control de acceso basado en etiquetas (LBAC)

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

zona de aterrizaje

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

modelo de lenguaje de gran tamaño (LLM)

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

migración grande

Migración de 300 servidores o más.

LBAC

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

privilegio mínimo

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

migrar mediante lift-and-shift

Consulte [Las 7 R](#).

sistema little-endian

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

LLM

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

entornos inferiores

Consulte [entorno](#).

M

machine learning (ML)

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

rama principal

Consulte [rama](#).

malware

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

Servicios administrados

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

sistema de ejecución de fabricación (MES)

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

MAP

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

mecanismo

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

cuenta de miembro

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

MES

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

Message Queuing Telemetry Transport (MQTT)

[Un protocolo de comunicación ligero machine-to-machine \(M2M\), basado en el patrón de publicación/suscripción, para dispositivos de IoT con recursos limitados.](#)

microservicio

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

arquitectura de microservicios

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

Programa de aceleración de la migración (MAP)

Un AWS programa que proporciona soporte de consultoría, formación y servicios para ayudar a las organizaciones a crear una base operativa sólida para migrar a la nube y para ayudar a compensar el costo inicial de las migraciones. El MAP incluye una metodología de migración para ejecutar las migraciones antiguas de forma metódica y un conjunto de herramientas para automatizar y acelerar los escenarios de migración más comunes.

migración a escala

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

fábrica de migración

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

metadatos de migración

Información sobre la aplicación y el servidor que se necesita para completar la migración. Cada patrón de migración requiere un conjunto diferente de metadatos de migración. Algunos ejemplos de metadatos de migración son la subred de destino, el grupo de seguridad y AWS la cuenta.

patrón de migración

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

Migration Portfolio Assessment (MPA)

Herramienta en línea que proporciona información a fin de validar los argumentos comerciales necesarios para migrar a la Nube de AWS. La MPA ofrece una evaluación detallada de la cartera

(adecuación del tamaño de los servidores, precios, comparaciones del costo total de propiedad, análisis de los costos de migración), así como una planificación de la migración (análisis y recopilación de datos de aplicaciones, agrupación de aplicaciones, priorización de la migración y planificación de oleadas). La [herramienta MPA](#) (requiere iniciar sesión) está disponible de forma gratuita para todos los AWS consultores y consultores de los socios de APN.

Evaluación de la preparación para la migración (MRA)

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

estrategia de migración

Enfoque utilizado para migrar una carga de trabajo a la Nube de AWS. Para más información, consulte la entrada [Las 7 R](#) de este glosario y también [Mobilize your organization to accelerate large-scale migrations](#).

ML

Consulte [machine learning](#).

modernización

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

evaluación de la preparación para la modernización

Evaluación que ayuda a determinar la preparación para la modernización de las aplicaciones de una organización; identifica los beneficios, los riesgos y las dependencias; y determina qué tan bien la organización puede soportar el estado futuro de esas aplicaciones. El resultado de la evaluación es un esquema de la arquitectura objetivo, una hoja de ruta que detalla las fases de desarrollo y los hitos del proceso de modernización y un plan de acción para abordar las brechas identificadas. Para más información, consulte [Evaluating modernization readiness for applications in the Nube de AWS](#).

aplicaciones monolíticas (monolitos)

Aplicaciones que se ejecutan como un único servicio con procesos estrechamente acoplados. Las aplicaciones monolíticas presentan varios inconvenientes. Si una característica de la

aplicación experimenta un aumento en la demanda, se debe escalar toda la arquitectura. Agregar o mejorar las características de una aplicación monolítica también se vuelve más complejo a medida que crece la base de código. Para solucionar problemas con la aplicación, puede utilizar una arquitectura de microservicios. Para obtener más información, consulte [Descomposición de monolitos en microservicios](#).

MPA

Consulte [Migration Portfolio Assessment](#).

MQTT

Consulte [Message Queuing Telemetry Transport](#).

clasificación multiclase

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

infraestructura mutable

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

O

OAC

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

OAI

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

OCM

Consulte [administración del cambio organizacional](#).

migración fuera de línea

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

OI

Consulte [integración de operaciones](#).

OLA

Consulte [acuerdo de nivel operativo](#).

migración en línea

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

OPC-UA

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

Open Process Communications: arquitectura unificada (OPC-UA)

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

acuerdo de nivel operativo (OLA)

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

revisión de la preparación operativa (ORR)

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

tecnología operativa (TO)

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

integración de operaciones (OI)

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

registro de seguimiento organizativo

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

administración del cambio organizacional (OCM)

Marco para administrar las transformaciones empresariales importantes y disruptivas desde la perspectiva de las personas, la cultura y el liderazgo. La OCM ayuda a las empresas a prepararse para nuevos sistemas y estrategias y a realizar la transición a ellos, al acelerar la adopción de cambios, abordar los problemas de transición e impulsar cambios culturales y organizacionales. En la estrategia de AWS migración, este marco se denomina aceleración de personal, debido a la velocidad de cambio que requieren los proyectos de adopción de la nube. Para obtener más información, consulte la [Guía de OCM](#).

control de acceso de origen (OAC)

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

identidad de acceso de origen (OAI)

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

ORR

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

OT

Consulte [tecnología operativa](#).

VPC saliente (de salida)

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

P

límite de permisos

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

información de identificación personal (PII)

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

PII

Consulte [información de identificación personal](#).

manual de estrategias

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

PLC

Consulte [controlador lógico programable](#).

PLM

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

policy

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

persistencia políglota

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

evaluación de cartera

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

predicate

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

inserción de predicados

Técnica de optimización de consultas en bases de datos que filtra los datos de la consulta antes de transferirlos. Esta técnica reduce la cantidad de datos de la base de datos relacional que se tienen que recuperar y procesar. Además, mejora el rendimiento de las consultas.

control preventivo

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

entidad principal

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

Privacidad desde el diseño

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

zonas alojadas privadas

Un contenedor que contiene información sobre cómo desea que Amazon Route 53 responda a las consultas de DNS de un dominio y sus subdominios dentro de uno o más VPCs. Para obtener más información, consulte [Uso de zonas alojadas privadas](#) en la documentación de Route 53.

control proactivo

[Control de seguridad](#) que se diseñó para evitar la implementación de recursos que no cumplan con la normativa. Estos controles analizan los recursos antes de aprovisionarlos. Si el recurso no cumple con los requisitos del control, no se aprovisiona. Para obtener más información, consulte la [guía de referencia de controles](#) en la AWS Control Tower documentación y consulte [Controles proactivos](#) en la sección Implementación de controles de seguridad en AWS.

administración del ciclo de vida del producto (PLM)

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

entorno de producción

Consulte [entorno](#).

controlador lógico programable (PLC)

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

encadenamiento de peticiones

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

seudonimización

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

publish/subscribe (pub/sub)

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

Q

plan de consulta

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

regresión del plan de consulta

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

R

Matriz RACI

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

RAG

Consulte [generación aumentada por recuperación](#).

ransomware

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

Matriz RASCI

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

RCAC

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

réplica de lectura

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

rediseñar

Consulte [Las 7 R](#).

objetivo de punto de recuperación (RPO)

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

objetivo de tiempo de recuperación (RTO)

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

refactorizar

Consulte [Las 7 R](#).

Region

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

regresión

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

volver a alojar

Consulte [Las 7 R](#).

versión

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

reubicar

Consulte [Las 7 R](#).

redefinir la plataforma

Consulte [Las 7 R](#).

recomprar

Consulte [Las 7 R](#).

resiliencia

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

política basada en recursos

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

matriz responsable, confiable, consultada e informada (RACI)

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

control receptivo

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

retain

Consulte [Las 7 R](#).

retirar

Consulte [Las 7 R](#).

Generación aumentada de recuperación (RAG)

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

rotación

Proceso mediante el que periódicamente se actualiza un [secreto](#) para que resulte más difícil que un atacante pueda acceder a las credenciales.

control de acceso por filas y columnas (RCAC)

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

RPO

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

RTO

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

manual de procedimientos

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

S

SAML 2.0

Un estándar abierto que utilizan muchos proveedores de identidad (IdPs). Esta función permite el inicio de sesión único (SSO) federado, de modo que los usuarios pueden iniciar sesión Consola de administración de AWS o llamar a las operaciones de la AWS API sin tener que crear un

usuario en IAM para todos los miembros de la organización. Para obtener más información sobre la federación basada en SAML 2.0, consulte [Acerca de la federación basada en SAML 2.0](#) en la documentación de IAM.

SCADA

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

SCP

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

secreta

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

seguridad desde el diseño

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

control de seguridad

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

refuerzo de la seguridad

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

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

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

automatización de la respuesta de seguridad

Acción predefinida y programada que está diseñada para responder automáticamente a un evento de seguridad o corregirlo. Estas automatizaciones sirven como controles de seguridad [preventivos o adaptables](#) que le ayudan a implementar las mejores prácticas AWS de seguridad. La modificación de un grupo de seguridad de VPC, la aplicación de revisiones a una instancia de Amazon EC2 o la rotación de credenciales son algunos ejemplos de acciones de respuesta automatizadas.

cifrado del servidor

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

política de control de servicio (SCP)

Política que proporciona un control centralizado de los permisos de todas las cuentas de una organización en AWS Organizations. SCPs defina barreras o establezca límites a las acciones que un administrador puede delegar en usuarios o roles. Puede utilizarlas SCPs como listas de permitidos o rechazados para especificar qué servicios o acciones están permitidos o prohibidos. Para obtener más información, consulte [las políticas de control de servicios](#) en la AWS Organizations documentación.

punto de enlace de servicio

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

acuerdo de nivel de servicio (SLA)

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

indicador de nivel de servicio (SLI)

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

objetivo de nivel de servicio (SLO)

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

modelo de responsabilidad compartida

Un modelo que describe la responsabilidad con AWS la que compartes la seguridad y el cumplimiento de la nube. AWS es responsable de la seguridad de la nube, mientras que usted es responsable de la seguridad en la nube. Para obtener más información, consulte el [Modelo de responsabilidad compartida](#).

SIEM

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

único punto de error (SPOF)

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

SLA

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

SLI

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

SLO

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

split-and-seed modelo

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

SPOF

Consulte [único punto de error](#).

esquema en estrella

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

patrón de higo estrangulador

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

subred

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

control de supervisión y adquisición de datos (SCADA)

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

cifrado simétrico

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

pruebas sintéticas

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

petición del sistema

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

T

etiquetas

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

variable de destino

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

lista de tareas

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

entorno de prueba

Consulte [entorno](#).

entrenamiento

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

puerta de enlace de tránsito

Un centro de tránsito de red que puede usar para interconectar sus redes con VPCs las locales. Para obtener más información, consulte [Qué es una pasarela de tránsito](#) en la AWS Transit Gateway documentación.

flujo de trabajo basado en enlaces troncales

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

acceso de confianza

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

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

ajuste

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

equipo de dos pizzas

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

U

incertidumbre

Un concepto que hace referencia a información imprecisa, incompleta o desconocida que puede socavar la fiabilidad de los modelos predictivos de ML. Hay dos tipos de incertidumbre: la incertidumbre epistémica se debe a datos limitados e incompletos, mientras que la incertidumbre aleatoria se debe al ruido y la aleatoriedad inherentes a los datos.

tareas indiferenciadas

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

entornos superiores

Consulte [entorno](#).

V

succión

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

control de versión

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

Emparejamiento de VPC

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

vulnerabilidad

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

W

caché caliente

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

datos templados

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

función de ventana

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

carga de trabajo

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

flujo de trabajo

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

WORM

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

WQF

Consulte [AWS Workload Qualification Framework](#).

escritura única y lectura múltiple (WORM)

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

Z

ataque de día cero

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

vulnerabilidad de día cero

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

peticiones desde cero

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

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

aplicación zombi

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

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