

---

# AWSAuto Scaling

Guía del usuario



## AWSAuto Scaling: Guía del usuario

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

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

## Table of Contents

¿Qué es AWS Auto Scaling? .....	1
Características de AWS Auto Scaling .....	1
Precios .....	2
Cómo empezar .....	2
Servicios relacionados .....	2
Cómo funcionan los planes de escalado .....	4
Introducción .....	6
Paso 1: Buscar los recursos escalables .....	6
Requisitos previos para el plan de escalado de muestra .....	7
Añadir el grupo de Auto Scaling al plan de escalado de muestra .....	7
Obtenga más información sobre cómo detectar sus recursos escalables .....	8
Paso 2: Especificación de la estrategia de escalado .....	9
Paso 3: Configuración de opciones avanzadas (opcional) .....	11
Configuración general .....	11
Configuración de escalado dinámico .....	13
Configuración de escalado predictivo .....	13
Paso 4: Creación del plan de escalado .....	14
(Opcional) Ver la información de escalado de un recurso .....	15
Paso 5: Eliminar recursos .....	17
Eliminar el grupo de Auto Scaling .....	17
Paso 6: Pasos siguientes .....	17
Prácticas recomendadas .....	19
Otras consideraciones .....	19
Cómo evitar el error <code>ActiveWithProblems</code> .....	20
Seguridad .....	21
Protección de los datos .....	21
Identity and Access Management .....	22
Control de acceso .....	22
Cómo funciona AWS Auto Scaling con IAM .....	23
Roles vinculados a servicios .....	25
Ejemplos de políticas basadas en identidades .....	27
Registro y monitorización .....	31
Validación de la conformidad .....	32
Resiliencia .....	32
Seguridad de la infraestructura .....	33
Uso de puntos de enlace de la VPC para conectividad privada .....	33
Creación de un punto de enlace de la VPC de tipo interfaz .....	33
Crear una política de punto de enlace de la VPC .....	34
Migración de puntos de enlace .....	34
Cuotas de servicio .....	36
Recursos .....	37
Historial de revisión .....	38
.....	xi

# ¿Qué es AWS Auto Scaling?

AWS Auto Scaling le permite configurar el escalado automático de los recursos de AWS que forman parte de la aplicación en cuestión de minutos. La consola de AWS Auto Scaling ofrece una interfaz de usuario único que permite usar las características de escalado automático de varios servicios de AWS. Puede configurar el escalado automático para recursos individuales o para aplicaciones completas.

Con AWS Auto Scaling, la configuración y la administración del escalado de los recursos se lleva a cabo mediante un plan de escalado. El plan de escalado utiliza el escalado dinámico y el escalado predictivo para escalar automáticamente los recursos de la aplicación. De este modo, se garantiza que se añade la potencia de cómputo necesaria para satisfacer la carga en la aplicación y, a continuación, se elimina cuando ya no es necesaria. El plan de escalado le permite elegir las estrategias de escalado para definir cómo optimizar la utilización de recursos. Puede optimizar la disponibilidad, los costos o mantener un equilibrio entre ambos. También puede crear estrategias de escalado personalizadas.

AWS Auto Scaling resulta útil para aplicaciones que experimentan variaciones diarias o semanales en el flujo de tráfico, incluidas las siguientes:

- Aplicaciones de tráfico cíclico; por ejemplo, las que experimentan un uso elevado de recursos durante las horas laborables y un uso reducido de recursos por la noche
- Patrones de carga de trabajo de activación y desactivación, como procesamiento por lotes, pruebas o análisis periódicos
- Aplicaciones de patrones de tráfico variables, como campañas de marketing con periodos de picos de crecimiento

## Características de AWS Auto Scaling

Utilice AWS Auto Scaling para escalar automáticamente los siguientes recursos:

- Grupos de : Amazon EC2 Auto Scaling. Lance o termine instancias EC2 en un grupo de Auto Scaling.
- Solicitudes de flota de spot de Amazon EC2: Lance o termine instancias desde una solicitud de flota de spot, o reemplace automáticamente las instancias que se interrumpen por motivos de precio o de capacidad.
- Amazon ECS: Ajuste el recuento deseado del servicio ECS hacia arriba o hacia abajo en respuesta a las variaciones de carga.
- Amazon DynamoDB: Describe cómo habilitar una tabla de DynamoDB o un índice secundario global para aumentar o reducir su capacidad de lectura y escritura provisionada para administrar los aumentos de tráfico sin que se aplique una limitación controlada.
- Amazon Aurora: Ajusta dinámicamente el número de réplicas de lectura de Aurora aprovisionadas para un clúster de base de datos Aurora para controlar los cambios en las conexiones activas o en la carga de trabajo.

Las características de escalado disponibles actualmente son el escalado dinámico y el escalado predictivo.

El escalado dinámico crea políticas de escalado de seguimiento de destino para los recursos escalables de la aplicación. Esto permite al plan de escalado añadir y eliminar capacidad para cada recurso según sea necesario con el fin de mantener la utilización de recursos en el valor de destino especificado. Las métricas de escalado predeterminadas proporcionadas están basadas en las métricas más utilizadas para el escalado automático.

Funcionamiento del escalado predictivo:

- **Previsión de la carga:** AWS Auto Scaling analiza hasta 14 días de historial para una métrica de carga especificada y prevé la demanda futura para los dos días siguientes. Estos datos están disponibles en intervalos de una hora y se actualizan todos los días.
- **Acciones de escalado programadas:** AWS Auto Scaling programa las acciones de escalado que añaden y eliminan capacidad de recursos de forma proactiva para reflejar la previsión de carga. A la hora programada, AWS Auto Scaling actualiza la capacidad mínima del recurso con el valor especificado por la acción de escalado programada. El objetivo es mantener la utilización de recursos en el valor de destino especificado por la estrategia de escalado. Si la aplicación necesita más capacidad de la prevista, el escalado dinámico está disponible para añadir capacidad adicional.
- **Comportamiento de capacidad máxima:** Cada recurso tiene un límite de capacidad mínima y máxima entre el que se espera que se encuentre el valor especificado por la acción de escalado programada. Sin embargo, es posible controlar si la aplicación puede añadir recursos más allá de su capacidad máxima en el caso de que la capacidad prevista sea superior a la capacidad máxima.

En la actualidad, el escalado predictivo solo está disponible para los grupos de Amazon EC2 Auto Scaling.

## Precios

Las características de AWS Auto Scaling las habilitan las métricas y las alarmas de Amazon CloudWatch. Las características se proporcionan sin cargo adicional más allá de las cuotas de servicio para CloudWatch y el resto de recursos de AWS que se utilicen.

## Cómo empezar

Para obtener una introducción a AWS Auto Scaling, recomendamos que consulte la siguiente información:

- [Cómo funcionan los planes de escalado \(p. 4\)](#) — este tema presenta los conceptos de estrategias de escalado, escalado dinámico y escalado predictivo para ayudarle a familiarizarse con AWS Auto Scaling.
- [AWS Auto Scaling FAQs](#) Las preguntas frecuentes de la página del producto proporcionan información sobre los beneficios de este servicio.—
- [Regiones y puntos de enlace de](#) en la AWS General Reference—Esta tabla muestra la disponibilidad regional de AWS Auto Scaling.
- [Guía del usuario de Amazon EC2 Auto Scaling](#) — esta guía le muestra cómo crear y administrar los grupos de Auto Scaling que debe utilizar al escalar la flota de instancias Amazon EC2.
- [Guía del usuario de Auto Scaling de aplicaciones](#) — esta guía le facilita temas y recursos relacionados con el escalado automático de recursos más allá de Amazon EC2. Siempre que necesite más información específica sobre el escalado de un recurso o servicio escalable distinto de Amazon EC2, puede obtener acceso a la documentación técnica de esta guía.

Para empezar, complete el tutorial de introducción de AWS Auto Scaling que se encuentra en [Introducción a AWS Auto Scaling \(p. 6\)](#).

## Servicios relacionados

[AWS CloudFormation](#) le permite utilizar plantillas, que son archivos de texto con formato JSON o YAML, para modelar y aprovisionar una colección de recursos de AWS relacionados. Puede utilizar las plantillas de muestra de AWS CloudFormation o crear las suyas propias para crear los recursos de AWS, así

como cualquier dependencia asociada o parámetros de tiempo de ejecución necesarios para ejecutar su aplicación. También puede crear plantillas de planes de escalado mediante AWS CloudFormation.

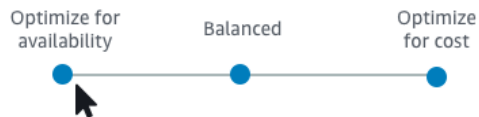
[Amazon CloudWatch](#) es un servicio de monitorización para los recursos de la nube de AWS y las aplicaciones que se ejecutan en AWS. CloudWatch le permite recopilar y realizar el seguimiento de métricas y archivos de registro, así como reaccionar automáticamente a los cambios en las aplicaciones mediante alarmas. También puede publicar sus propias métricas personalizadas en CloudWatch utilizando la AWS CLI o una API.

# Cómo funcionan los planes de escalado

Un plan de escalado es el componente principal de AWS Auto Scaling. Es donde se configura un conjunto de instrucciones para escalar los recursos. Si trabaja con recursos de AWS CloudFormation o añade etiquetas a recursos de AWS, puede configurar planes de escalado para diferentes conjuntos de recursos, por aplicación. AWS Auto Scaling proporciona recomendaciones para definir estrategias de escalado personalizadas para cada recurso. Después de crear el plan de escalado, AWS Auto Scaling combina los métodos de escalado dinámico y escalado predictivo para ayudarle a desarrollar su propia estrategia de escalado.

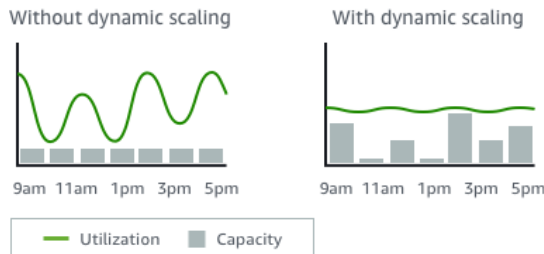
¿Qué es una estrategia de escalado?

La estrategia de escalado indica a AWS Auto Scaling cómo optimizar la utilización de los recursos en el plan de escalado. Puede optimizar la disponibilidad, los costos o mantener un equilibrio entre ambos. Otra opción consiste en crear una estrategia propia, de acuerdo con las métricas y umbrales que se definan. Puede establecer estrategias independientes para cada recurso o tipo de recurso.



¿Qué es el escalado dinámico?

El escalado dinámico crea políticas de escalado de seguimiento de destino para los recursos del plan de escalado. Estas políticas de escalado ajustan la capacidad de los recursos en respuesta a los cambios en la utilización de los recursos. El objetivo es proporcionar suficiente capacidad para mantener la utilización de recursos en el valor de destino especificado por la estrategia de escalado. Se asemeja a los termostatos que se utilizan para mantener la temperatura del hogar. Se elige la temperatura y el termostato hace el resto.

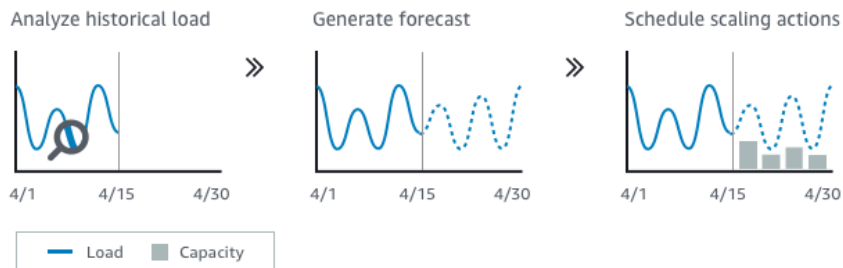


Por ejemplo, puede configurar el plan de escalado de forma que mantenga el número de tareas que ejecuta el servicio de ECS al 75 por ciento de CPU. Cuando la utilización de CPU del servicio sube por encima del 75 por ciento (lo que significa que se está utilizando más del 75 por ciento de la CPU que se ha reservado para el servicio), esto indica a la política de escalado que debe añadir otra tarea al servicio para ayudar con el aumento de carga.

¿Qué es el escalado predictivo?

El escalado predictivo utiliza el aprendizaje automático para analizar la carga de trabajo histórica de cada uno de los recursos y prevé de forma periódica la carga futura para los próximos dos días. Esto es similar

a cómo funcionan las previsiones meteorológicas. Utilizando la previsión, el escalado predictivo genera acciones de escalado programadas para garantizar que la capacidad de los recursos esté disponible antes de que la aplicación la necesite. Al igual que el escalado dinámico, el escalado predictivo mantiene la utilización en el valor objetivo especificado por la estrategia de escalado.



Por ejemplo, puede habilitar el escalado predictivo y configurar la estrategia de escalado de forma que mantenga la utilización media de la CPU del grupo de Auto Scaling en el 50 por ciento. Su previsión indica que se van a producir picos de tráfico todos los días a las 8 de la mañana. El plan de escalado crea las acciones de escalado programadas futuras para asegurarse de que el grupo de Auto Scaling esté listo para atender el tráfico de antemano. Esto ayuda a mantener constante el rendimiento de la aplicación, siempre con el objetivo de tener la capacidad necesaria de mantener la utilización de recursos tan cercana al 50 por ciento como sea posible en todo momento.



# Introducción a AWS Auto Scaling

En esta sección, se describen los pasos para comenzar a utilizar AWS Auto Scaling. Puede utilizar la Consola de administración de AWS para crear su primer plan de escalado. A continuación, aprenderá los conceptos básicos de la creación de planes de escalado con el escalado predictivo y el escalado dinámico habilitados.

Antes de crear un plan de escalado para utilizarlo con una aplicación, revise esta a fondo mientras se ejecuta en la nube de AWS. Tome nota de lo siguiente:

- Si dispone de políticas de escalado existentes creadas desde otras consolas. Puede reemplazar las políticas de escalado existentes o mantenerlas (sin que se le permita realizar cambios en sus valores) al crear el plan de escalado.
- La utilización de destino que tiene sentido para cada recurso escalable de la aplicación en función del recurso en su conjunto. Por ejemplo, se espera que la cantidad de CPU que las instancias EC2 de un grupo de Auto Scaling utilicen en comparación con su CPU disponible. O bien, para un servicio como DynamoDB que utiliza un modelo de desempeño provisionado, la cantidad de actividad de lectura y escritura que se espera que utilice una tabla o un índice en comparación con el desempeño disponible. En otras palabras, la proporción entre capacidad consumida y provisionada. Puede cambiar el objetivo de utilización en cualquier momento una vez que haya creado el plan de escalado.
- Cuánto tiempo tarda en lanzar y configurar un servidor. Esta información le ayudará a configurar una ventana para que cada instancia EC2 se prepare después del lanzamiento para garantizar que no lancemos un nuevo servidor mientras que el anterior se sigue lanzando.
- Si el historial de métricas es lo suficientemente largo para utilizarlo con el escalado predictivo (si se utilizan grupos de Auto Scaling recién creados). En general, si se dispone de 14 días completos de datos históricos, esto se traducirá en unas previsiones más precisas. El valor mínimo es de 24 horas.

Cuanto mejor conozca la aplicación, mayor será la eficacia del plan de escalado.

## Temas

- [Paso 1: Buscar los recursos escalables \(p. 6\)](#)
- [Paso 2: Especificación de la estrategia de escalado \(p. 9\)](#)
- [Paso 3: Configuración de opciones avanzadas \(opcional\) \(p. 11\)](#)
- [Paso 4: Creación del plan de escalado \(p. 14\)](#)
- [Paso 5: Eliminar recursos \(p. 17\)](#)
- [Paso 6: Pasos siguientes \(p. 17\)](#)

## Paso 1: Buscar los recursos escalables

En la sección de introducción, creará un plan de escalado y obtendrá una introducción práctica al uso de AWS Auto Scaling a través de la consola de administración de AWS.

Al crear un plan de escalado desde la consola, AWS Auto Scaling le ayuda a encontrar los recursos escalables como primer paso. Hay tres formas de localizar los recursos para un nuevo plan de escalado desde la consola:

- Puede elegir una pila de AWS CloudFormation para que la consola de AWS Auto Scaling la utilice para detectar automáticamente sus recursos escalables.

- Puede elegir un conjunto de etiquetas para que la consola de AWS Auto Scaling las utilice para detectar automáticamente sus recursos escalables.
- Puede elegir uno o varios grupos de Amazon EC2 Auto Scaling para añadirlos a su plan de escalado.

Si este es su primer plan de escalado, le recomendamos que empiece eligiendo la tercera opción y creando un plan de escalado de muestra mediante un grupo de Auto Scaling de EC2.

## Requisitos previos para el plan de escalado de muestra

Para obtener un tutorial sencillo para utilizar la consola con el fin de crear un plan de escalado, le recomendamos que comience por crear un grupo de Auto Scaling y, a continuación, crear el plan de escalado y añadir el grupo de Auto Scaling. Mediante el uso de un grupo de Auto Scaling, puede activar la característica de escalado predictivo y escalado dinámico. Debe habilitar ambas características para utilizar el conjunto completo de características disponibles en el plan de escalado.

Empiece con la creación de un grupo de Auto Scaling, si aún no tiene uno. Para obtener más información, consulte [Introducción a Amazon EC2 Auto Scaling](#) en la Guía del usuario de Amazon EC2 Auto Scaling. Si crea un nuevo grupo, puede eliminarlo después. En cuanto se elimine el grupo, dejarán de acumularse cargos por las instancias Amazon EC2 que este ejecuta.

Configure el grupo de Auto Scaling como se indica a continuación para garantizar que el plan de escalado funciona según lo previsto:

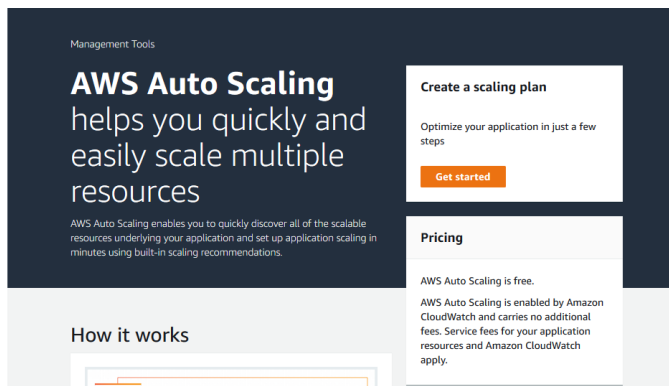
- En la plantilla de lanzamiento o la configuración de lanzamiento que asocia al grupo de Auto Scaling, habilite la monitorización detallada para obtener datos de métricas de CloudWatch para instancias EC2 con una frecuencia de 1 minuto. Se aplican cargos adicionales. Para obtener más información, consulte [Configurar la monitorización de instancias de Auto Scaling en la Guía del usuario de Amazon EC2 Auto Scaling](#).
- Describe cómo habilitar las métricas de grupo de Auto Scaling para obtener datos agregados para el grupo de instancias en CloudWatch. Para obtener más información, consulte [Habilitar métricas de grupo de Auto Scaling en la Guía del usuario de Amazon EC2 Auto Scaling](#).
- Si utiliza un tipo de instancia T2 o T3, utilice una plantilla de lanzamiento para configurar las instancias como `unlimited`, de modo que puedan mantener el alto rendimiento de la CPU mientras realiza las pruebas. Podrían aplicarse cargos adicionales. Para obtener más información, consulte [Uso de un grupo de Auto Scaling para lanzar una instancia de rendimiento ampliable como ilimitada](#) en la Guía del usuario de Amazon EC2 para instancias de Linux.

## Añadir el grupo de Auto Scaling al plan de escalado de muestra

Ahora que ha creado su grupo de Auto Scaling, está listo para crear su plan de escalado de muestra mediante la Consola de administración de AWS.

Para añadir un grupo de Auto Scaling a un nuevo plan de escalado

1. Abra la consola de AWS Auto Scaling en <https://console.aws.amazon.com/autoscaling/>.
2. En la barra de navegación de la parte superior de la pantalla, elija la misma región que utilizó cuando creó el grupo de Auto Scaling.
3. En la página de bienvenida, elija Get started (Introducción).



4. En la página Find scalable resources (Buscar recursos escalables), elija Search by CloudFormation stack (Buscar por pila de), Search by tag (Buscar por etiqueta) o Choose EC2 Auto Scaling groups (Elegir grupos de Auto Scaling de EC2).

#### Note

En este tutorial se supone que está eligiendo un grupo de Auto Scaling. Posteriormente, puede utilizar este mismo procedimiento para crear un plan de escalado mediante la opción Search by CloudFormation stack o Search by tag.

- Si eligió Search by CloudFormation stack (Buscar por pila de ), elija la pila de AWS CloudFormation que desea utilizar.
- Si ha elegido Search by tag (Buscar por etiqueta), para cada etiqueta, elija una clave de etiqueta en Key (Clave) y valores de etiqueta en Value (Valor). Para añadir etiquetas, elija Add another row (Añadir otra fila). Para eliminar etiquetas, elija Remove (Quitar).
- Si ha elegido Choose EC2 Auto Scaling groups (Elegir grupos de Auto Scaling de EC2), en Auto Scaling groups (Grupos de Auto Scaling), elija una o varios grupos de Auto Scaling.

5. Elija Next (Siguiente) para agregar el grupo de Auto Scaling al plan de escalado y para continuar con el siguiente paso.

Si elige la opción Search by CloudFormation stack o Search by tag, al elegir Next, los recursos escalables asociados con la pila o el conjunto de etiquetas están disponibles para el plan de escalado. Cuando defina el plan de escalado, puede elegir cuáles de estos recursos desea incluir o excluir.

## Obtenga más información sobre cómo detectar sus recursos escalables

Si ya ha creado un plan de escalado de ejemplo y desea crear más, la siguiente información explica los escenarios para utilizar una pila de CloudFormation o un conjunto de etiquetas con más detalle. Puede

utilizar esta sección para decidir si desea elegir la opción Search by CloudFormation stack o Search by tag para detectar los recursos escalables cuando utilice la consola para crear el plan de escalado.

#### Detección de recursos escalables con una pila de CloudFormation

Al utilizar CloudFormation, trabaja con pilas para aprovisionar recursos. Todos los recursos de una pila se definen mediante la plantilla de la pila. El plan de escalado añade una capa de orquestación encima de la pila que facilita la configuración de escalado para varios recursos. Sin un plan de escalado, tendría que configurar el escalado para cada recurso escalable por separado. Esto implica averiguar el orden de aprovisionamiento de los recursos y las políticas de escalado, así como conocer las peculiaridades de cómo funcionan estas dependencias.

En la consola de AWS Auto Scaling, puede seleccionar una pila existente para buscar los recursos que se pueden configurar para el escalado automático. AWS Auto Scaling solo busca los recursos que están definidos en la pila seleccionada. No busca en las pilas anidadas.

Para que sus servicios de ECS se puedan detectar en una pila de CloudFormation, la consola de AWS Auto Scaling debe saber qué clúster de ECS está ejecutando el servicio. Esto requiere que sus servicios de ECS estén en la misma pila de CloudFormation que el clúster de ECS que está ejecutando el servicio. De lo contrario, deben formar parte del clúster predeterminado. Para que se identifique correctamente, el nombre del servicio ECS también debe ser único en cada uno de estos clústeres de ECS.

Para obtener más información acerca de CloudFormation, consulte [¿Qué es AWS CloudFormation?](#) en la Guía del usuario de AWS CloudFormation.

#### Detección de recursos escalables mediante etiquetas

Las etiquetas proporcionan metadatos que se pueden utilizar para detectar recursos escalables relacionados en la consola de AWS Auto Scaling mediante filtros de etiquetas.

Utilice etiquetas para buscar cualquiera de los siguientes recursos:

- Clústeres de base de datos de Aurora
- Grupos de Auto Scaling
- Tablas e índices secundarios globales de DynamoDB

Cuando busca por más de una etiqueta, cada recurso debe tener todas las etiquetas indicadas para que se pueda encontrar.

Las etiquetas se pueden asignar de diferentes formas. Para obtener más información, consulte [Etiquetado de recursos de AWS](#) en la AWS General Reference.

## Paso 2: Especificación de la estrategia de escalado

Utilice el siguiente procedimiento para especificar estrategias de escalado para los recursos detectados en el paso anterior.

Para cada tipo de recurso, AWS Auto Scaling elige la métrica que se usa habitualmente para determinar la cantidad del recurso que está en uso en un momento dado. Usted elige la estrategia de escalado más apropiada para optimizar el rendimiento de la aplicación en función de esta métrica. Cuando habilita la característica de escalado dinámico y la característica de escalado predictivo, la estrategia de escalado se comparte entre ellas. Para obtener más información, consulte [Cómo funcionan los planes de escalado](#) (p. 4).

Las siguientes estrategias de escalado están disponibles:

- **Optimize for availability (Optimizar para disponibilidad)** — AWS Auto Scaling escala el recurso de forma ascendente o descendente para mantener la utilización del recurso al 40 por ciento. Esta opción es útil cuando la aplicación tiene necesidades de escalado urgentes y a veces impredecibles.
- **Balance availability and cost (Equilibrar disponibilidad y costo)** — AWS Auto Scaling escala el recurso de forma ascendente o descendente para mantener la utilización del recurso al 50 por ciento. Esta opción le ayuda a mantener una alta disponibilidad y a reducir los costos al mismo tiempo.
- **Optimize for cost (Optimizar para costo)** — AWS Auto Scaling escala el recurso de forma ascendente o descendente para mantener la utilización del recurso al 70 por ciento. Esta opción es útil para reducir los costos si la aplicación es capaz de mantener una capacidad del búfer reducida cuando se producen cambios inesperados en la demanda.

Por ejemplo, el plan de escalado configura su grupo de Auto Scaling para añadir o eliminar instancias Amazon EC2 en función de la cantidad media de CPU utilizada para todas las instancias del grupo. Puede optimizar estos recursos para factores como la disponibilidad, el costo o una combinación de ambos cambiando la estrategia de escalado.

También puede configurar una estrategia personalizada si la estrategia que viene de serie no satisface sus necesidades. Con una estrategia personalizada, puede cambiar el valor de utilización objetivo, elegir una métrica diferente o ambas cosas.

### Important

Si no tiene experiencia, realice únicamente el primer paso del siguiente procedimiento y, a continuación, seleccione Next (Siguiente) para continuar. (Puede omitir el resto del procedimiento, ya que el tutorial se centra en cómo usar la estrategia de escalado predeterminada, Optimize for availability (Optimizar para disponibilidad), que mantiene la utilización media de la CPU de su grupo de Auto Scaling en un 40 por ciento).

### Para especificar estrategias de escalado

1. En la página Specify scaling strategy (Especificar estrategia de escalado), en Scaling plan details (Detalles del plan de escalado), Name (Nombre), escriba un nombre para el plan de escalado. El nombre del plan de escalado debe ser único en el conjunto de planes de escalado de la región de AWS, puede tener un máximo de 128 caracteres y no debe contener barras verticales "|", barras diagonales "/" ni caracteres de dos puntos ":".
2. Para cada tipo de recurso, proporcione las siguientes instrucciones de escalado.
  - a. Para Scaling strategy (Estrategia de escalado), elija una de estas opciones: : optimiza para la disponibilidad, equilibra la disponibilidad y el costo, optimiza para el costo o personalizada.

**Auto Scaling groups (1)**  Include in scaling plan

Specify a scaling strategy for 1 Auto Scaling group.

**Scaling strategy**  
The strategy defines the scaling metric and target value used to scale your resources.

**Optimize for availability**  
Keep the average CPU utilization of your Auto Scaling groups at 40% to provide high availability and ensure capacity to absorb spikes in demand.

**Balance availability and cost**  
Keep the average CPU utilization of your Auto Scaling groups at 50% to provide optimal availability and reduce costs.

**Optimize for cost**  
Keep the average CPU utilization of your Auto Scaling groups at 70% to ensure lower costs.

**Custom**  
Choose your own scaling metric, target value, and other settings.

**Enable predictive scaling**  
Support your scaling strategy by continually forecasting load and proactively scheduling capacity ahead of when you need it. [Info](#)

**Enable dynamic scaling**  
Support your scaling strategy by creating target tracking scaling policies to monitor your scaling metric and increase or decrease capacity as you need it. [Info](#)

► [Configuration details](#)

- b. Si ha elegido Custom (Personalizada) en el paso anterior, elija la configuración personalizada en Configuration details (Detalles de configuración). Aquí puede encontrar la lista de métricas disponibles (si procede), así como gráficos relacionados basados en datos de CloudWatch. El historial de métricas recientes es el principal objetivo de los gráficos.

- En Scaling metric (Métrica de escalado), elija la métrica de escalado que desee. Si no hay otras métricas predefinidas disponibles, esta opción no tiene una lista desplegable.
  - En Target value (Valor de destino), elija el valor de objetivo de utilización que desee.
  - En Load metric (Métrica de carga) (solo para los grupos de Auto Scaling), elija una métrica de carga apropiada para utilizarla en el escalado predictivo.
  - En Replace external scaling policies (Reemplazar políticas de escalado externas), elija si desea eliminar las políticas de escalado creadas desde fuera del plan de escalado (por ejemplo, desde otras consolas) y sustituirlas por políticas de escalado de seguimiento de destino nuevas creadas por el plan de escalado.
- c. (Opcional) El escalado predictivo está habilitado de forma predeterminada en los grupos de Auto Scaling. Para deshabilitar el escalado predictivo en los grupos de Auto Scaling, anule la selección de Enable predictive scaling (Habilitar escalado predictivo).
  - d. (Opcional) El escalado dinámico está habilitado de forma predeterminada para todos los tipos de recursos. Para deshabilitar el escalado dinámico para un tipo de recurso, desactive la opción Enable dynamic scaling (Habilitar escalado dinámico).
  - e. (Opcional) De forma predeterminada, cuando se especifica un origen de aplicación del que se detectan varios recursos escalables, todos los tipos de recursos se incluyen automáticamente en el plan de escalado. Para omitir un tipo de recurso del plan de escalado, desactive la opción Include in scaling plan (Incluir en el plan de escalado).
3. Cuando haya terminado, elija Next (Siguiendo).

## Paso 3: Configuración de opciones avanzadas (opcional)

Ahora que ha especificado la estrategia de escalado que se va a utilizar para cada tipo de recurso, puede elegir personalizar cualquiera de los valores predeterminados de cada recurso mediante el paso Configure advanced settings (Configurar opciones avanzadas). Para cada tipo de recurso, existen varios grupos de ajustes que se pueden personalizar. No obstante, en la mayoría de los casos, la configuración predeterminada debería ser óptima, con la posible excepción de los valores de capacidad mínima y capacidad máxima, que deben ajustarse con cuidado.

Omita este procedimiento si desea mantener la configuración predeterminada. Puede cambiar esta configuración en cualquier momento editando el plan de escalado.

### Important

En el tutorial dirigido a los usuarios sin experiencia, vamos a realizar algunos cambios para actualizar la capacidad máxima del grupo de Auto Scaling y habilitar el escalado predictivo en modo de solo previsión. Aunque no necesita personalizar todos los ajustes del tutorial, vamos a examinar brevemente también los ajustes de cada sección.

## Configuración general

Utilice este procedimiento para ver y personalizar los ajustes que ha especificado en el paso anterior para cada recurso. También puede personalizar la capacidad mínima y la capacidad máxima de cada recurso.

Para ver y personalizar la configuración general

1. En la página Configure advanced settings (Configurar opciones avanzadas), elija la flecha situada a la izquierda de cualquier sección para expandirla. Para el tutorial, expanda la sección Auto Scaling groups (Grupos de Auto Scaling).
2. En la tabla que se muestra, seleccione el grupo de Auto Scaling que va a utilizar en este tutorial.

3. Deje la opción **Include in scaling plan** (Incluir en plan de escalado) seleccionada. Si esta opción no está seleccionada, el recurso se omite del plan de escalado. Si no incluye al menos un recurso, no se puede crear el plan de escalado.
4. Para ampliar la vista y ver los detalles de la sección **Configuración general**, seleccione la flecha que aparece a la izquierda del título de la sección.
5. Puede seleccionar opciones para los siguientes elementos. En este tutorial, busque la opción **Maximum capacity** (Capacidad máxima) y especifique un valor de 3 en lugar del valor actual.
  - **Scaling strategy** (Estrategia de escalado) — le permite optimizar la disponibilidad, el costo, mantener un equilibrio entre ambos o especificar una estrategia personalizada.
  - **Enable dynamic scaling** (Habilitar escalado dinámico) — si esta opción está desactivada, el recurso seleccionado no se puede escalar mediante una configuración de escalado de seguimiento de destino.
  - **Enable predictive scaling** (Habilitar escalado predictivo) (solo para los grupos de Auto Scaling) — si se desactiva esta opción, el grupo de escalado no se puede escalar con el escalado predictivo.
  - **Scaling metric** (Métrica de escalado) — especifica la métrica de escalado que se debe utilizar. Si elige **Custom** (Personalizada), puede especificar una métrica personalizada en lugar de las métricas predefinidas que están disponibles en la consola. Para obtener más información, consulte el siguiente tema de esta sección.
  - **Target value** (Valor de destino) — especifica el valor del objetivo de utilización que se va a utilizar.
  - **Load metric** (Métrica de carga) (solo para los grupos de Auto Scaling) — especifica la métrica de carga que se va a utilizar. Si elige **Custom** (Personalizada), puede especificar una métrica personalizada en lugar de las métricas predefinidas que están disponibles en la consola. Para obtener más información, consulte el siguiente tema de esta sección.
  - **Minimum capacity** (Capacidad mínima) — especifica la capacidad mínima del recurso. AWS Auto Scaling garantiza que el recurso nunca quede por debajo de este tamaño.
  - **Maximum capacity** (Capacidad máxima) — especifica la capacidad máxima del recurso. AWS Auto Scaling garantiza que el recurso nunca supere este tamaño.

#### Note

Cuando se utiliza el escalado predictivo, tiene la opción de elegir un comportamiento de capacidad máxima diferente en función de la capacidad de previsión. Esta opción está en la sección **Predictive scaling settings** (Configuración del escalado predictivo).

## Métricas personalizadas

AWS Auto Scaling proporciona las métricas más utilizadas para el escalado automático. Sin embargo, dependiendo de sus necesidades, es posible que prefiera obtener datos de métricas diferentes a las de la consola. Amazon CloudWatch tiene muchas métricas distintas entre las que puede elegir. CloudWatch también le permite publicar sus propias métricas.

El JSON se utiliza para especificar una métrica personalizada de CloudWatch. Para que pueda seguir estas instrucciones, le recomendamos que se familiarice con la [Guía del usuario de Amazon CloudWatch](#).

Para especificar una métrica personalizada, debe crear una carga con formato JSON utilizando un conjunto de parámetros obligatorios a partir de una plantilla. Los valores de cada parámetro se añaden desde CloudWatch. Proporcionamos la plantilla como parte de las opciones personalizadas de **Scaling metric** (Métrica de escalado) y **Load metric** (Cargar métrica) en la configuración avanzada del plan de escalado.

JSON representa los datos de dos formas:

- Un objeto, que es una colección sin ordenar de pares de nombre-valor. Un objeto está delimitado por una llave de apertura (`{`) y una llave de cierre (`}`). Cada par de nombre-valor comienza por el nombre, seguido de dos puntos, seguido del valor. Los pares de nombre-valor están separados por comas.

- Una matriz, que es una colección ordenada de valores. Una matriz está delimitada por un corchete de apertura ([]) y un corchete de cierre (]). Los elementos de la matriz están separados por comas.

A continuación, se muestra un ejemplo de una plantilla JSON con valores de muestra para cada parámetro:

```
{
  "MetricName": "MyBackendCPU",
  "Namespace": "MyNamespace",
  "Dimensions": [
    {
      "Name": "MyOptionalMetricDimensionName",
      "Value": "MyOptionalMetricDimensionValue"
    }
  ],
  "Statistic": "Sum"
}
```

Para obtener más información, consulte [Especificación de métrica de escalado personalizada](#) y [Especificación de métrica de carga personalizada](#) en la Referencia de la API de AWS Auto Scaling.

## Configuración de escalado dinámico

Utilice este procedimiento para ver y personalizar la configuración de la política de escalado de seguimiento de destino que crea AWS Auto Scaling.

Para ver y personalizar la configuración de escalado dinámico

1. Para ampliar la vista y ver los detalles de la sección Configuración del escalado dinámico, seleccione la flecha que aparece a la izquierda del título de la sección.
2. Puede seleccionar opciones para los siguientes elementos. Sin embargo, la configuración predeterminada es idónea para este tutorial.
  - Replace external scaling policies (Reemplazar políticas de escalado externas) — si se desactiva esta opción, se mantienen las políticas de escalado existentes creadas desde fuera del plan de escalado y no se crean políticas nuevas.
  - Disable scale-in (Deshabilitar escalado descendente) — si esta opción está desactivada, se permite el escalado descendente para reducir la capacidad actual del recurso cuando la métrica especificada es inferior al valor de destino.
  - Cooldown (Recuperación) — crea periodos de recuperación de escalado ascendente y escalado descendente. El periodo de recuperación es la cantidad de tiempo que la política de escalado espera a que surta efecto una actividad de escalado anterior. Para obtener más información, consulte [Periodo de recuperación](#) en la Guía del usuario de Auto Scaling de aplicaciones. (Esta configuración no se muestra si el recurso es un grupo de Auto Scaling).
  - Instance warmup (Preparación de las instancias) — [solo para los grupos de Auto Scaling] permite controlar la cantidad de tiempo que transcurre antes de que la instancia que se acaba de lanzar comience a contribuir a las métricas de CloudWatch. Para obtener más información, consulte [Preparación de las instancias](#) en la Guía del usuario de Amazon EC2 Auto Scaling.

## Configuración de escalado predictivo

Si el recurso es un grupo de Auto Scaling, utilice este procedimiento para ver y personalizar la configuración que AWS Auto Scaling utiliza para el escalado predictivo.



Para ver y personalizar la configuración de escalado predictivo

1. Para ampliar la vista y ver los detalles de la sección Configuración del escalado predictivo, seleccione la flecha que aparece a la izquierda del título de la sección.
2. Puede seleccionar opciones para los siguientes elementos. En este tutorial cambie el valor de Predictive scaling mode (Modo de escalado predictivo) a Forecast only (Solo previsión).
  - Predictive scaling mode (Modo de escalado predictivo) — especifica el modo de escalado. El valor predeterminado es Forecast and scale (Previsión y escalado). Si lo cambia a Forecast only (Solo previsión), el plan de escalado prevé capacidad futura, pero no aplica las acciones de escalado.
  - Pre-launch instances (Lanzamiento previo de instancias) — ajusta las acciones de escalado para que se ejecuten antes al realizar el escalado ascendente. Por ejemplo, la previsión establece que se debe agregar capacidad a las 10:00 a.m. y el tiempo de búfer es de 5 minutos (300 segundos). El tiempo de ejecución de la acción de escalado correspondiente será entonces las 9:55 a. m. Esto resulta útil para los grupos de Auto Scaling, en los que pueden transcurrir varios minutos desde que se lanza una instancia hasta que entra en servicio. El tiempo real puede variar, ya que depende de varios factores, como el tamaño de la instancia y si hay scripts de inicio que deben completarse. El valor predeterminado es de 300 segundos.
  - Max Capacity Behavior (Comportamiento de capacidad máxima) — controla si el recurso seleccionado se puede escalar por encima de la capacidad máxima cuando la capacidad prevista se acerca a la capacidad máxima especificada actualmente o la supera. El valor predeterminado es Enforce the maximum capacity setting (Aplicar la configuración de capacidad máxima).
    - Enforce the maximum capacity setting (Aplicar la configuración de capacidad máxima) — AWS Auto Scaling no puede escalar la capacidad del recurso por encima de la capacidad máxima. La capacidad máxima se aplica como un límite invariable.
    - Set the maximum capacity to equal forecast capacity (Definir la capacidad máxima igual que la capacidad de previsión) — AWS Auto Scaling puede escalar la capacidad por encima de la capacidad máxima para que sea igual a la capacidad máxima, pero no mayor.
    - Increase maximum capacity above forecast capacity (Aumentar la capacidad máxima por encima de la capacidad de previsión) — AWS Auto Scaling puede escalar la capacidad del recurso por encima de la capacidad máxima en un valor de búfer especificado. El objetivo es dar a la política de escalado de seguimiento de destino capacidad adicional si se produce tráfico inesperado.
  - Max capacity behavior buffer (Búfer de comportamiento de capacidad máxima) — si ha elegido Increase maximum capacity above forecast capacity (Aumentar la capacidad máxima por encima de la capacidad de previsión), elija el tamaño del búfer de capacidad que desea utilizar cuando la capacidad de previsión se aproxime a la capacidad máxima o la supere. El valor se especifica como un porcentaje en relación con la capacidad prevista. Por ejemplo, con un búfer del 10 %, si la capacidad de previsión es 50 y la capacidad máxima es 40, la capacidad máxima efectiva es 55.
3. Cuando haya terminado de personalizar la configuración, elija Next (Siguiendo).

#### Note

Para revertir cualquiera de sus cambios, seleccione los recursos y elija Revert to original (Volver al original). Esta opción restablece los recursos seleccionados a su último estado conocido dentro del plan de escalado.

## Paso 4: Creación del plan de escalado

En la página Review and create (Revisar y crear), revise los detalles de su plan de escalado y elija Create scaling plan (Crear plan de escalado). Se le dirigirá a una página que muestra el estado del plan de escalado. El plan de escalado puede tardar algún tiempo en terminar de crearse mientras se actualizan los recursos.

Con el escalado predictivo, AWS Auto Scaling analiza el historial de la métrica de carga especificada de los últimos 14 días (se necesitan un mínimo de 24 horas de datos) para generar una previsión de dos días. A

continuación, programa acciones de escalado para ajustar la capacidad de los recursos con respecto a la previsión para cada hora del periodo de la previsión.

Una vez que se ha terminado de crear el plan de escalado, consulte los detalles del plan eligiendo su nombre en la pantalla Scaling plans (Planes de escalado).

## (Opcional) Ver la información de escalado de un recurso

Utilice este procedimiento para ver la información de escalado creada para un recurso.

Los datos se presentan de las siguientes formas:

- Mediante gráficos que muestran datos del historial de métricas recientes de CloudWatch.
- Gráficos de escalado predictivo que muestran las previsiones de carga y las previsiones de capacidad basados en los datos de AWS Auto Scaling.
- Una tabla que incluye todas las acciones de escalado predictivo programadas para el recurso.

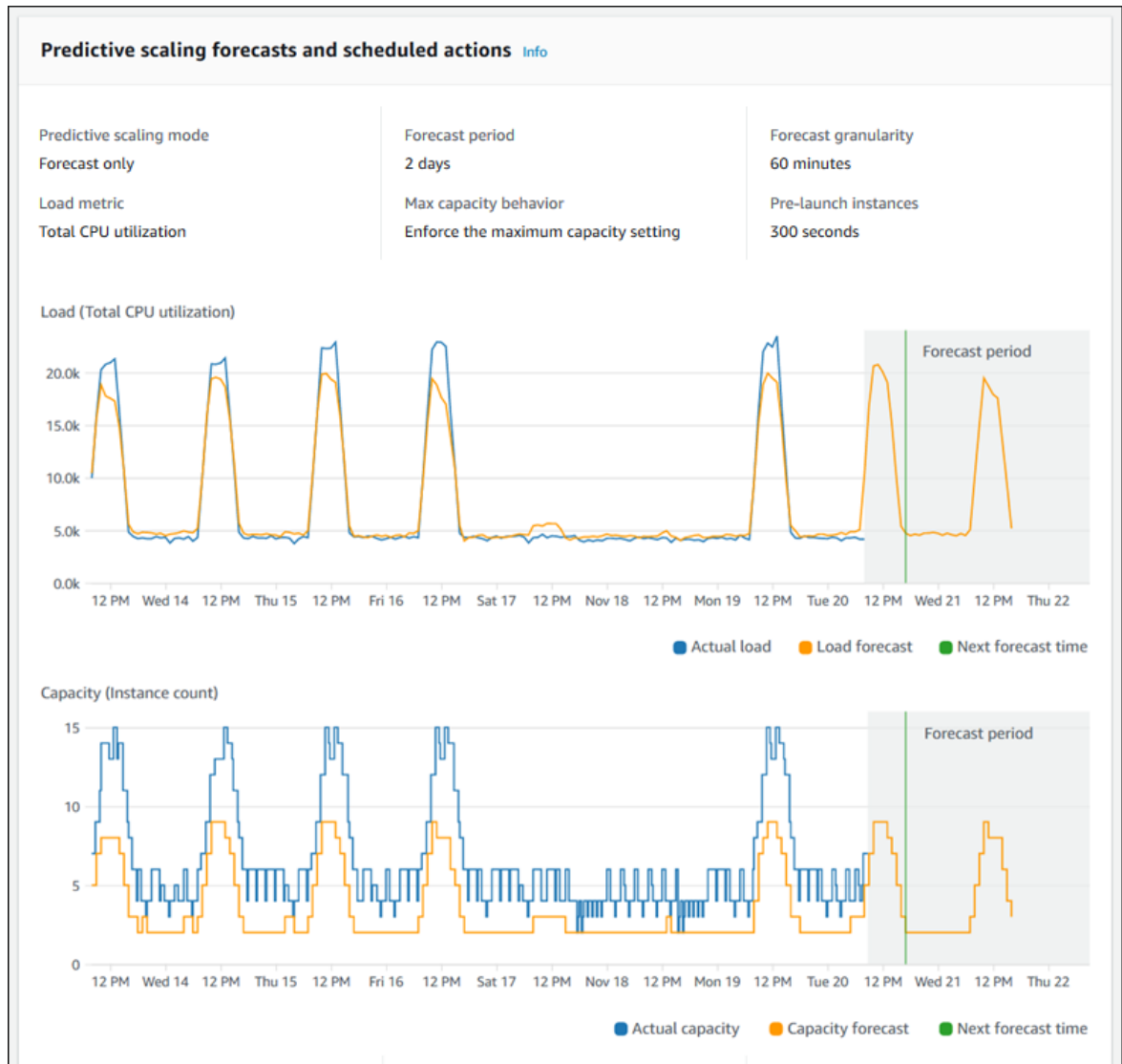
Para ver la información de escalado de un recurso

1. Abra la consola de AWS Auto Scaling en <https://console.aws.amazon.com/autoscaling/>.
2. En la página Scaling plans (Planes de escalado), elija el plan de escalado.
3. En la página Scaling plan details (Detalles del plan de escalado), elija el recurso que desea ver.

## Monitoreo y evaluación de previsiones

Cuando su plan de escalado esté en funcionamiento, puede monitorizar la previsión de carga, la capacidad de previsión y las acciones de escalado para examinar el rendimiento del escalado predictivo. Todos estos datos están disponibles en la consola de AWS Auto Scaling para todos los grupos de Auto Scaling con escalado predictivo habilitado. Tenga en cuenta que el plan de escalado requiere al menos 24 horas de datos de carga históricos para crear la previsión inicial.

En el siguiente ejemplo, en el lado izquierdo de cada gráfico se muestra la pauta histórica. En el lado derecho del gráfico se muestra la previsión de carga generada por el plan de escalado para el periodo de la previsión. Ambos valores reales y previstos (en azul y naranja) se representan gráficamente.



AWS Auto Scaling aprende a partir de los datos de forma automática. En primer lugar, realiza una previsión de carga. A continuación, se utiliza una previsión de capacidad de cálculo para determinar determina la cantidad mínima de instancias que se necesitan para la aplicación. En función de la capacidad prevista, AWS Auto Scaling programa acciones de escalado que escalan el grupo de Auto Scaling antes de los cambios de carga previstos. Si el escalado dinámico está habilitado (opción recomendada), el grupo de Auto Scaling puede escalar capacidad adicional (o eliminarla) en función del uso que se realice en cada momento de la capacidad del grupo de instancias.

Al evaluar el nivel de rendimiento del escalado predictivo, monitorice el grado de ajuste de los valores reales y los previstos a lo largo del tiempo. Al crear un plan de escalado, AWS Auto Scaling presenta los gráficos en función de los datos reales más recientes. También facilita una previsión inicial durante las siguientes 48 horas. Sin embargo, al crear el plan de escalado, existen muy pocos datos de previsión para comparar con los datos reales. Espere hasta que el plan de escalado haya obtenido valores de previsión durante algunos periodos antes de comparar los valores de previsión históricos con los reales. Tras realizar previsiones diarias durante algunos días, tendrá una muestra más grande de valores de previsión con los que comparar los valores reales.

Para los patrones que se presentan a diario, el intervalo temporal entre la creación del plan de escalado y la evaluación de la efectividad de la previsión puede ser breve, de solo unos pocos días. Sin embargo, este

plazo temporal es insuficiente para evaluar la previsión sobre la base de un cambio de patrón reciente. Por ejemplo, supongamos que está analizando la previsión de un grupo de Auto Scaling que ha iniciado una nueva campaña de marketing en la última semana. La campaña hace aumentar significativamente el tráfico web durante los dos mismos días cada semana. En situaciones como esta, recomendamos esperar a que el grupo recopile nuevos datos durante una semana completa o un par de ellas antes de evaluar la efectividad de la previsión. La misma recomendación se aplica a un nuevo grupo de Auto Scaling que acabe de empezar a recopilar datos de métricas.

Si los valores reales y de previsión no coinciden tras monitorizarlos durante un periodo de tiempo apropiado, debería también tener en cuenta si la métrica de carga que ha elegido es adecuada. Para que sea eficaz, la métrica de carga debe representar una medición de confianza y precisa de la carga total en todas las instancias del grupo de Auto Scaling. La métrica de carga es clave para el escalado predictivo. Si elige una métrica de carga que no sea óptima, esta podría impedir que el escalado predictivo realizase previsiones exactas de carga y capacidad y que programase los ajustes de capacidad correctos para el grupo de Auto Scaling.

## Paso 5: Eliminar recursos

Una vez que haya completado el tutorial de introducción, puede que desee conservar el plan de escalado. No obstante, si no está utilizando el plan de escalado de forma activa, debería plantearse su eliminación para que la cuenta no incurra en cargos innecesarios.

Cuando se elimina un plan de escalado, se eliminan las políticas de escalado de seguimiento de destino, las alarmas de CloudWatch asociadas y las acciones de escalado predictivo que AWS Auto Scaling ha creado en su nombre.

La eliminación de un plan de escalado no elimina su pila de AWS CloudFormation, el grupo de Auto Scaling ni los recursos escalables.

Para eliminar un plan de escalado

1. Abra la consola de AWS Auto Scaling en <https://console.aws.amazon.com/autoscaling/>.
2. En la página Scaling plans (Planes de escalado), seleccione el plan de escalado que creó para este tutorial y elija Delete (Eliminar).
3. Cuando se le pida confirmación, seleccione Delete (Eliminar).

Tras eliminar su plan de escalado, sus recursos no vuelven a su capacidad original. Por ejemplo, si el grupo de Auto Scaling se ajusta a 10 instancias al eliminar el plan de escalado, el grupo seguirá ajustándose a 10 instancias tras eliminarse el plan de escalado. Puede actualizar la capacidad de recursos específicos obteniendo acceso a la consola para cada uno de los recursos.

## Eliminar el grupo de Auto Scaling

Para evitar que su cuenta acumule cargos adicionales de Amazon EC2, debe eliminar el grupo de Auto Scaling que ha creado para este tutorial.

Para obtener instrucciones paso a paso, consulte [Eliminar el grupo de Auto Scaling en la Guía del usuario de Amazon EC2 Auto Scaling](#).

## Paso 6: Pasos siguientes

Ahora que ya se ha familiarizado con AWS Auto Scaling y algunas de sus características, es posible que desee intentar crear su propia plantilla de plan de escalado mediante AWS CloudFormation.

Una plantilla de AWS CloudFormation es un archivo de texto con formato JSON o YAML que describe la infraestructura de AWS necesaria para ejecutar una aplicación o un servicio, junto con las interconexiones entre los componentes de la infraestructura. Con AWS CloudFormation, implementa y administra una colección asociada de recursos como una pila. AWS CloudFormation está disponible sin que suponga ningún cargo adicional y solo paga por los recursos de AWS necesarios para ejecutar sus aplicaciones. Los recursos pueden ser cualquier recurso de AWS que defina en la plantilla. Para obtener más información, consulte [Conceptos de AWS CloudFormation en la Guía del usuario de AWS CloudFormation](#).

En la Guía del usuario de AWS CloudFormation, hemos proporcionado una plantilla sencilla para comenzar. La plantilla de ejemplo está disponible como un ejemplo en la sección [AWS::AutoScalingPlans::ScalingPlan](#) de la documentación de referencias de plantillas de AWS CloudFormation. La plantilla de ejemplo crea un plan de escalado para un solo grupo de Auto Scaling y habilita el escalado predictivo y el escalado dinámico.

Para obtener más información, consulte [Introducción a AWS CloudFormation](#) en la Guía del usuario de AWS CloudFormation.

# Prácticas recomendadas para planes de escalado de AWS Auto Scaling

Las siguientes prácticas recomendadas pueden ayudarle a sacar el máximo partido de los planes de escalado:

- Siempre que sea posible, debe escalar basándose en métricas de instancias Amazon EC2 cuya frecuencia sea de 1 minuto, ya que esto garantiza una respuesta más rápida a los cambios de utilización. Realizar el escalado basándose en métricas con una frecuencia de 5 minutos puede generar un tiempo de respuesta más lento y hacer que el escalado se realice con datos de métricas que están obsoletos. De forma predeterminada, las instancias EC2 tienen habilitada la monitorización básica, lo que significa que los datos de las métricas de las instancias están disponibles en intervalos de 5 minutos. Abonando una cantidad adicional, puede habilitar la monitorización detallada para que los datos de las métricas de las instancias se reciban con una frecuencia de 1 minuto. Para obtener más información, consulte [Configurar la monitorización de instancias de Auto Scaling en la Guía del usuario de Amazon EC2 Auto Scaling](#).
- También recomendamos habilitar las métricas de grupo de Auto Scaling. De lo contrario, los datos de la capacidad real no se muestran en los gráficos de previsión de la capacidad que están disponibles al finalizar el asistente Crear plan de escalado. Para habilitar las métricas de grupo de Auto Scaling, abra un grupo de Auto Scaling en la consola de Amazon EC2 y, en la pestaña Monitoring (Monitorización), elija Enable Group Metrics Collection (Habilitar la recopilación de métricas de grupo). Estas métricas describen el grupo en lugar de cualquiera de sus instancias. Para obtener más información, consulte [Habilitar métricas de grupo de Auto Scaling en la Guía del usuario de Amazon EC2 Auto Scaling](#).
- Compruebe el tipo de instancia que utiliza el grupo de Auto Scaling. Las instancias Amazon EC2 de rendimiento ampliable, que son las instancias T3 y T2, se han diseñado para ofrecer un rendimiento de CPU básico con posibilidad de ampliarlo a un nivel superior si así lo exige la carga de trabajo. Dependiendo del objetivo de utilización especificado por el plan de escalado, se podría correr el riesgo de sobrepasar la base de referencia y quedarse sin créditos de CPU, lo que limitaría el rendimiento. Para obtener más información, consulte [Créditos de CPU y rendimiento de referencia para las instancias de rendimiento ampliable](#). Para configurar estas instancias como unlimited, consulte [Uso de un grupo de Auto Scaling para lanzar una instancia de rendimiento ampliable como ilimitada](#) en la Guía del usuario de Amazon EC2 para instancias de Linux.

## Otras consideraciones

Tenga en cuenta las siguientes consideraciones adicionales:

- El escalado predictivo utiliza previsiones de carga de trabajo para programar la capacidad en el futuro. La calidad de las previsiones varía en función de las fluctuaciones cíclicas de la carga de trabajo y de la aplicabilidad del modelo de previsión entrenado. El escalado predictivo se puede ejecutar en modo de solo previsión para evaluar la calidad de las previsiones y las acciones de escalado creadas por las previsiones. Puede establecer el modo de escalado predictivo en Forecast only (Solo previsión) cuando cree el plan de escalado y después cambiarlo a Forecast and scale (Previsión y escalado) cuando termine de evaluar la calidad de la previsión. Para obtener más información, consulte [Configuración de escalado predictivo \(p. 13\)](#) y [Monitoreo y evaluación de previsiones \(p. 15\)](#).
- Si decide especificar métricas diferentes para el escalado predictivo, debe asegurarse de que la métrica de escalado y la de carga guarden una estrecha relación. El valor de la métrica debe aumentar y disminuir proporcionalmente al número de instancias del grupo de Auto Scaling. De esta forma, se asegurará de que los datos de las métricas se puedan utilizar para ampliar o reducir proporcionalmente

el número de instancias. Por ejemplo, la métrica de carga es el número total de solicitudes y la métrica de escalado es el uso medio de la CPU. Si el número total de solicitudes aumenta en un 50 por ciento, el uso medio de la CPU también debería aumentar en un 50 por ciento, siempre que la capacidad no cambie.

- Antes de crear su plan de escalado, debe eliminar todas las acciones de escalado programadas anteriormente que ya no necesite obteniendo acceso a las consolas en las que fueron creadas. AWS Auto Scaling no crea una acción de escalado predictivo que se solape con una acción de escalado programada existente.
- La configuración personalizada para la capacidad mínima y máxima, junto con los demás ajustes que se utilizan para el escalado dinámico, se muestran en las otras consolas. Sin embargo, le recomendamos que, después de crear un plan de escalado, no modifique estos ajustes desde las otras consolas, ya que el plan de escalado no recibe estas actualizaciones de esas otras consolas.
- El plan de escalado puede contener recursos de varios servicios, pero cada recurso solo puede estar en un plan de escalado a la vez.

## Cómo evitar el errorActiveWithProblems

Puede producirse un error "ActiveWithProblems" cuando se crea un plan de escalado o se añaden recursos a un plan de escalado. El error aparece cuando el plan de escalado está activo, pero no se ha podido aplicar la configuración de escalado de uno o varios recursos.

Normalmente, esto sucede porque un recurso ya tiene una política de escalado o porque un grupo de Auto Scaling no cumple los requisitos mínimos del escalado predictivo.

Si alguno de sus recursos ya tiene políticas de escalado de varias consolas de AWS, AWS Auto Scaling no sobrescribe estas otras políticas de escalado ni crea otras nuevas de forma predeterminada. Si lo desea, puede eliminar las políticas de escalado existentes y reemplazarlas por políticas de escalado de seguimiento de destino creadas desde la consola de AWS Auto Scaling. Para ello, habilite la opción `Replace external scaling policies` (Reemplazar políticas de escalado externas) para cada recurso que tenga políticas de escalado que desee sobrescribir.

Con el escalado predictivo, se recomienda esperar 24 horas después de crear un nuevo grupo de Auto Scaling para configurar el escalado predictivo. Como mínimo, debe disponerse de 24 horas de datos históricos para generar la previsión inicial. Si el grupo tiene menos de 24 horas de datos históricos y se habilita el escalado predictivo, esto impide que el plan de escalado pueda generar una previsión hasta que el siguiente periodo de previsión después del grupo haya recopilado la cantidad necesaria de datos. Sin embargo, también puede editar y guardar el plan de escalado para reiniciar el proceso de previsión en cuanto estén disponibles las 24 horas de datos.

# Seguridad en AWS Auto Scaling

La seguridad en la nube de AWS es la mayor prioridad. Como cliente de AWS, se beneficiará de una arquitectura de red y un centro de datos diseñados para satisfacer los requisitos de seguridad de las organizaciones más exigentes.

La seguridad es una responsabilidad compartida entre AWS y usted. El [modelo de responsabilidad compartida](#) describe esto como seguridad de la nube y la seguridad en la nube:

- Seguridad de la nube – AWS es responsable de proteger la infraestructura que funciona AWS servicios en el AWS Nube. AWS también le ofrece servicios que puede utilizar de forma segura. Los auditores de terceros comprueban y verifican con regularidad la efectividad de nuestra seguridad como parte de la [AWS programas de cumplimiento](#). Para conocer los programas de cumplimiento aplicables a AWS Auto Scaling, ver [Servicios AWS en alcance por programa de cumplimiento](#).
- Seguridad en la nube – Su responsabilidad está determinada por el AWS el servicio que utiliza. También es responsable de otros factores, incluida la confidencialidad de los datos, los requisitos de la empresa y la legislación y los reglamentos aplicables.

Esta documentación le ayudará a conocer cómo puede aplicar el modelo de responsabilidad compartida cuando utilice AWS Auto Scaling. En los siguientes temas, se le mostrará cómo configurar AWS Auto Scaling para satisfacer sus objetivos de seguridad y conformidad. También aprenderá a utilizar otros servicios AWS que le ayuden a controlar y asegurar su AWS Auto Scaling recursos.

Temas:

- [AWS Auto Scaling y protección de datos \(p. 21\)](#)
- [Identity and Access Management para AWS Auto Scaling \(p. 22\)](#)
- [Registro y monitoreo en AWS Auto Scaling \(p. 31\)](#)
- [Validación de la conformidad de AWS Auto Scaling \(p. 32\)](#)
- [Resiliencia en AWS Auto Scaling \(p. 32\)](#)
- [Seguridad de la infraestructura de AWS Auto Scaling \(p. 33\)](#)
- [AWS Auto Scaling y puntos de enlace de la VPC de tipo interfaz \(p. 33\)](#)

## AWS Auto Scaling y protección de datos

El [modelo de responsabilidad compartida](#) de AWS se aplica a la protección de datos en AWS Auto Scaling. Como se describe en este modelo, AWS es responsable de proteger la infraestructura global que ejecuta toda la nube de AWS. Usted es responsable de mantener el control sobre el contenido alojado en esta infraestructura. Este contenido incluye la configuración de seguridad y las tareas de administración de los servicios de AWS que utiliza. Para obtener más información acerca de la privacidad de datos, consulte las [Preguntas frecuentes sobre privacidad de datos](#). Para obtener más información sobre la protección de datos en Europa, consulte la publicación de blog [AWS Shared Responsibility Model and GDPR](#) en el blog de seguridad de AWS.

Para fines de protección de datos, recomendamos proteger las credenciales de cuenta de AWS y configurar cuentas de usuario individuales con AWS Identity and Access Management (IAM). De esta manera, solo se otorgan a cada usuario los permisos necesarios para cumplir con sus obligaciones laborales. También le recomendamos proteger sus datos de las siguientes formas:

- Utilice la autenticación multifactor (MFA) con cada cuenta.



- Utilice SSL/TLS para comunicarse con los recursos de AWS. Le recomendamos TLS 1.2 o una versión posterior.
- Configure la API y el registro de actividad del usuario con AWS CloudTrail.
- Utilice las soluciones de cifrado de AWS, junto con todos los controles de seguridad predeterminados dentro de los servicios de AWS.
- Utilice los servicios de seguridad administrados avanzados como, por ejemplo, Amazon Macie, que ayudan a detectar y proteger los datos personales almacenados en Amazon S3.
- Si necesita módulos criptográficos validados FIPS 140-2 al acceder a AWS a través de una interfaz de línea de comandos o una API, utilice un punto de enlace de FIPS. Para obtener más información acerca de los puntos de enlace de FIPS disponibles, consulte [Estándar de procesamiento de la información federal \(FIPS\) 140-2](#).

Le recomendamos encarecidamente que nunca introduzca información de identificación confidencial, como, por ejemplo, números de cuenta de sus clientes, en los campos de formato libre, como el campo Name (Nombre). No debe introducir esta información cuando trabaje con AWS Auto Scaling u otros servicios de AWS a través de la consola, la API, la AWS CLI de AWS o los SDK de AWS. Cualquier dato que escriba en AWS Auto Scaling o en otros servicios se puede incluir en los registros de diagnóstico. Cuando proporcione una URL a un servidor externo, no incluya información de credenciales en la URL para validar la solicitud para ese servidor.

## Identity and Access Management para AWS Auto Scaling

AWS Identity and Access Management (IAM) es un servicio de AWS que ayuda a un administrador a controlar de forma segura el acceso a los recursos de AWS. Los administradores de IAM controlan quién puede ser autenticado (iniciar sesión) y estar autorizado (tener permisos) para utilizar los recursos de AWS Auto Scaling. IAM es un servicio de AWS que se puede utilizar sin costo adicional.

Para usar AWS Auto Scaling, necesita una cuenta de AWS y credenciales de AWS. Para aumentar la seguridad de su cuenta de AWS, le recomendamos que utilice un usuario de IAM para proporcionar credenciales de acceso en lugar de usar las credenciales de su cuenta de AWS. Para obtener más información, consulte [Credenciales de usuario raíz de la cuenta de AWS frente a credenciales de usuario de IAM](#) en la Referencia general de AWS e [IAM Best Practices](#) en la Guía del usuario de IAM.

Para obtener información general sobre los usuarios de IAM y por qué son importantes para la seguridad de su cuenta, consulte [Credenciales de seguridad de AWS](#) en la Referencia general de AWS.

Para obtener más información acerca de cómo trabajar con IAM, consulte la [Guía del usuario de IAM](#).

### Control de acceso

Aunque disponga de credenciales válidas para autenticar las solicitudes, si no tiene permisos, no podrá crear recursos de AWS Auto Scaling ni obtener acceso a ellos. Por ejemplo, debe tener permisos para crear planes de escalado, configurar el escalado predictivo, etc.

En las secciones siguientes se incluye información sobre cómo un administrador de IAM puede utilizar IAM para ayudar a proteger sus recursos de AWS controlando quién puede realizar acciones de AWS Auto Scaling.

#### Temas

- [Cómo funciona AWS Auto Scaling con IAM \(p. 23\)](#)
- [Roles vinculados a servicios de AWS Auto Scaling \(p. 25\)](#)

- [Ejemplos de políticas basadas en identidades de AWS Auto Scaling \(p. 27\)](#)

## Cómo funciona AWS Auto Scaling con IAM

Antes de utilizar IAM para administrar el acceso a AWS Auto Scaling, debe saber qué características de IAM están disponibles para utilizarse con AWS Auto Scaling. Para obtener una perspectiva general de cómo funcionan AWS Auto Scaling y otros servicios de AWS con IAM, consulte [Servicios de AWS que funcionan con IAM](#) en la Guía del usuario de IAM.

### Temas

- [Políticas basadas en identidades de AWS Auto Scaling \(p. 23\)](#)
- [Políticas basadas en recursos de AWS Auto Scaling \(p. 24\)](#)
- [Listas de control de acceso \(ACL\) \(p. 24\)](#)
- [Autorización basada en etiquetas de AWS Auto Scaling \(p. 24\)](#)
- [Roles de IAM de AWS Auto Scaling: \(p. 24\)](#)

## Políticas basadas en identidades de AWS Auto Scaling

Con las políticas basadas en identidades de IAM, puede especificar las acciones permitidas o denegadas, así como los recursos y las condiciones en las que se permiten o deniegan las acciones. AWS Auto Scaling admite acciones, claves de condiciones y recursos específicos. Para obtener más información acerca de los elementos que utiliza en una política de JSON, consulte [Referencia de los elementos de las políticas de JSON de IAM](#) en la Guía del usuario de IAM.

### Actions

Administrators can use AWS JSON policies to specify who has access to what. That is, which principal can perform actions on what resources, and under what conditions.

El elemento `Action` de una política JSON describe las acciones que puede utilizar para permitir o denegar el acceso en una política. Las acciones de la política generalmente tienen el mismo nombre que la operación de API de AWS asociada. Hay algunas excepciones, como acciones de solo permiso que no tienen una operación de API coincidente. También hay algunas operaciones que requieren varias acciones en una política. Estas acciones adicionales se denominan acciones dependientes.

Incluya acciones en una política para conceder permisos para realizar la operación asociada.

Las acciones de políticas de AWS Auto Scaling utilizan el siguiente prefijo antes de la acción: `autoscaling-plans:`. Las instrucciones de política deben incluir un elemento `Action` o `NotAction`. AWS Auto Scaling define su propio conjunto de acciones, que describen las tareas que se pueden realizar con este servicio.

Para especificar varias acciones en una única instrucción, sepárelas con comas como se muestra en el siguiente ejemplo.

```
"Action": [
    "autoscaling-plans:DescribeScalingPlans",
    "autoscaling-plans:DescribeScalingPlanResources"
```

Puede utilizar caracteres comodín para especificar varias acciones (\*). Por ejemplo, para especificar todas las acciones que comiencen con la palabra `Describe`, incluya la siguiente acción.

```
"Action": "autoscaling-plans:Describe*"
```

Para ver la lista de las acciones de AWS Auto Scaling, consulte [Acciones](#) en la Referencia de la API de AWS Auto Scaling.

## Resources

El elemento `Resource` especifica el objeto u objetos a los que se aplica la acción.

AWS Auto Scaling no tiene recursos definidos por el servicio que se pueden utilizar como el elemento `Resource` de una declaración política de IAM. Por lo tanto, no hay nombres de recurso de Amazon (ARN) de AWS Auto Scaling que se puedan utilizar en una política de IAM. Para controlar el acceso a las acciones de AWS Auto Scaling, utilice siempre un \* (asterisco) como recurso al escribir una política de IAM.

## Claves de condición

El elemento `Condition` (o `Condition` bloque de ) permite especificar condiciones en las que entra en vigor una instrucción. Por ejemplo, es posible que desee que solo se aplique una política después de una fecha específica. Para expresar condiciones, se usan claves de condición predefinidas.

AWS Auto Scaling no proporciona ninguna clave de condición específica del servicio, pero sí admite el uso de algunas claves de condición globales. Para ver todas las claves de condición globales de AWS, consulte [Claves de contexto de condición globales de AWS](#) en la Guía del usuario de IAM.

El elemento `Condition` es opcional.

## Examples

Para ver ejemplos de políticas de AWS Auto Scaling basadas en identidades, consulte [Ejemplos de políticas basadas en identidades de AWS Auto Scaling \(p. 27\)](#).

## Políticas basadas en recursos de AWS Auto Scaling

Otros servicios de AWS, como Amazon Simple Storage Service, admiten políticas de permisos basadas en recursos. Por ejemplo, puede asociar una política de permisos a un bucket de S3 para administrar los permisos de acceso a dicho bucket.

AWS Auto Scaling no admite políticas basadas en recursos.

## Listas de control de acceso (ACL)

AWS Auto Scaling no admite las listas de control de acceso (ACL).

## Autorización basada en etiquetas de AWS Auto Scaling

AWS Auto Scaling no tiene recursos definidos por servicios que se puedan etiquetar. Por lo tanto, no permite el control del acceso basado en etiquetas.

## Roles de IAM de AWS Auto Scaling:

Un [rol de IAM](#) es una entidad de la cuenta de AWS que dispone de permisos específicos.

## Uso de credenciales temporales con AWS Auto Scaling

Puede utilizar credenciales temporales para iniciar sesión con federación, adoptar un rol de IAM o adoptar un rol de acceso entre cuentas. Las credenciales de seguridad temporales se obtienen llamando a operaciones de la API de AWS STS, como [AssumeRole](#) o [GetFederationToken](#).

AWS Auto Scaling admite el uso de credenciales temporales.

## Roles vinculados a servicios

AWS Auto Scaling admite los roles vinculados a servicios.

En función de los recursos que agregue a su plan de escalado, se crea automáticamente un rol vinculado a servicios para Amazon EC2 Auto Scaling y un rol vinculado a servicios por tipo de recurso para Auto Scaling de aplicaciones.

Si habilita el escalado predictivo, AWS crea automáticamente el rol vinculado a servicios de AWS Auto Scaling.

Para obtener más información, consulte la documentación específica del rol correspondiente:

- [Roles vinculados a servicios para Amazon EC2 Auto Scaling](#) en la Guía del usuario de Amazon EC2 Auto Scaling
- [Roles vinculados a servicios para Auto Scaling de aplicaciones](#) en la Guía del usuario de Auto Scaling de aplicaciones
- [Roles vinculados a servicios de AWS Auto Scaling \(p. 25\)](#) en esta guía

Puede utilizar el siguiente procedimiento para determinar si su cuenta ya tiene un rol vinculado a servicio.

Para determinar si ya existe un rol vinculado a servicio

1. Abra la consola de IAM en <https://console.aws.amazon.com/iam/>.
2. Seleccione Roles en el panel de navegación.
3. Busque en la lista "AWSServiceRole" los roles vinculados a servicios que existen en su cuenta. Busque el nombre del rol vinculado a servicio que desea comprobar.

## Roles de servicio

AWS Auto Scaling no tiene roles de servicio.

## Roles vinculados a servicios de AWS Auto Scaling

### Important

Para obtener información completa acerca de los roles vinculados a servicios que se necesitan para usar planes de escalado, consulte [Roles vinculados a servicios \(p. 25\)](#).

AWS Auto Scaling utiliza roles vinculados a servicios para los permisos que necesita para llamar a otros servicios de AWS en su nombre. Un rol vinculado a un servicio es un tipo único de rol de AWS Identity and Access Management (IAM) que está vinculado directamente a un servicio de AWS.

Los roles vinculados a servicios proporcionan una manera segura de delegar permisos en los servicios de AWS, ya que solo el servicio vinculado puede asumir un rol vinculado a servicio. Para obtener más información, consulte [Usar roles vinculados a servicios](#) en la Guía del usuario de IAM.

En las secciones siguientes, se describe cómo se crean y administran roles vinculados a servicio de AWS Auto Scaling. Comience configurando permisos que permitan a una entidad de IAM (como un usuario, un grupo o un rol) crear, editar o eliminar un rol vinculado a servicios.

## Permisos concedidos por el rol vinculado a servicios

AWS Auto Scaling utiliza el rol vinculado a servicio `AWSServiceRoleForAutoScalingPlans_EC2AutoScaling` para administrar el escalado predictivo de los grupos de Amazon EC2 Auto Scaling en su nombre. Este rol está predefinido con permisos para poder realizar llamadas en su nombre:

- `cloudwatch:GetMetricData`
- `autoscaling:DescribeAutoScalingGroups`
- `autoscaling:DescribeScheduledActions`
- `autoscaling:BatchPutScheduledUpdateGroupAction`
- `autoscaling:BatchDeleteScheduledAction`

Para asumir este rol se confía en el servicio `autoscaling-plans.amazonaws.com`.

## Creación del rol vinculado al servicio (automático)

AWS Auto Scaling crea el rol `_EC2AutoScalingAWSServiceRoleForAutoScalingPlans` automáticamente la primera vez que se crea un plan de escalado con el escalado predictivo habilitado.

### Important

Compruebe que ha habilitado los permisos de IAM que permiten a una entidad de IAM (como un usuario, un grupo o un rol) crear un rol vinculado al servicio. De lo contrario, la creación automática no se lleva a cabo. Para obtener más información, consulte [Permisos de roles vinculados a servicios](#) en la Guía del usuario de IAM o la información sobre [Permisos necesarios para crear un rol vinculado a un servicio](#) (p. 30) en esta guía.

## Creación del rol vinculado al servicio (manual)

Para crear el rol vinculado a servicio manualmente, puede utilizar la consola de IAM, la AWS CLI o la API de IAM. Para obtener más información, consulte [Crear un rol vinculado a un servicio](#) en la Guía del usuario de IAM.

Para crear un rol vinculado a un servicio (CLI de AWS)

Utilice el comando `create-service-linked-role` de la CLI para crear el rol vinculado a servicio de AWS Auto Scaling.

```
aws iam create-service-linked-role --aws-service-name autoscaling-plans.amazonaws.com
```

## Editar el rol vinculado a servicios

Con el rol `AWSServiceRoleForAutoScalingPlans_EC2AutoScaling` creado por AWS Auto Scaling, puede editar únicamente su descripción y no sus permisos. Para obtener más información, consulte [Editar un rol vinculado a un servicio](#) en la Guía del usuario de IAM.

## Eliminar el rol vinculado a un servicio

Si ya no utiliza planes de escalado, le recomendamos que elimine el rol vinculado a servicio. Solo puede eliminar un rol vinculado a un servicio después de eliminar sus recursos de AWS relacionados. Si un rol vinculado a un servicio se utiliza con varios planes de escalado, debe eliminar todos los planes de escalado que tengan habilitado el escalado predictivo para poder eliminar el rol. Esto protege los planes de escalado al impedirle que elimine accidentalmente los permisos que necesita para administrarlos. Para obtener más información, consulte [Paso 5: Eliminar recursos](#) (p. 17).

Puede utilizar IAM para eliminar el rol vinculado a un servicio. Para obtener más información, consulte [Eliminar un rol vinculado a un servicio](#) en la Guía del usuario de IAM.

Después de eliminar el rol vinculado al servicio `AWSServiceRoleForAutoScalingPlans_EC2AutoScaling`, AWS Auto Scaling lo volverá a crear cuando se cree un plan de escalado con el escalado predictivo habilitado.

## Regiones admitidas para el rol vinculado a servicio

AWS Auto Scaling admite el uso de roles vinculados a servicios en todas las regiones de AWS en las que el servicio está disponible.

## Ejemplos de políticas basadas en identidades de AWS Auto Scaling

De forma predeterminada, un nuevo usuario de IAM no tiene permisos para realizar ninguna actividad. Los administradores de IAM deben crear políticas de IAM que concedan permisos a los usuarios y los roles para realizar acciones de AWS Auto Scaling, como configurar políticas de escalado. El administrador debe asociar esas políticas a los usuarios o roles de IAM que necesiten esos permisos.

Para obtener información sobre cómo crear una política de IAM mediante estos documentos de políticas JSON de ejemplo, consulte [Creación de políticas en la pestaña JSON](#) en la Guía del usuario de IAM.

Si es la primera vez que crea una política, le recomendamos que en primer lugar cree un usuario de IAM en su cuenta y después asocie las políticas al usuario. Puede utilizar la consola para comprobar los efectos de cada política a medida que asocia la política al usuario.

### Temas

- [Prácticas recomendadas relativas a políticas](#) (p. 27)
- [Permitir a los usuarios crear planes de escalado](#) (p. 28)
- [Permitir a los usuarios habilitar el escalado predictivo](#) (p. 28)
- [Permisos necesarios adicionales](#) (p. 29)
- [Permisos necesarios para crear un rol vinculado a un servicio](#) (p. 30)

## Prácticas recomendadas relativas a políticas

Las políticas basadas en identidad son muy eficaces. Determinan si alguien puede crear, acceder o eliminar los recursos de AWS Auto Scaling de su cuenta. Estas acciones pueden generar costes adicionales para su cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidad:

- **Introducción sobre el uso de políticas administradas de AWS:** para comenzar a utilizar AWS Auto Scaling rápidamente, utilice las políticas administradas de AWS para proporcionar a los empleados los permisos necesarios. Estas políticas ya están disponibles en su cuenta y las mantiene y actualiza AWS. Para obtener más información, consulte [Introducción al uso de permisos con políticas administradas de AWS](#) en la Guía del usuario de IAM.
- **Conceder privilegios mínimos** – Al crear políticas personalizadas, conceda solo los permisos necesarios para llevar a cabo una tarea. Comience con un conjunto mínimo de permisos y conceda permisos adicionales según sea necesario. Por lo general, es más seguro que comenzar con permisos que son demasiado tolerantes e intentar hacerlos más severos más adelante. Para obtener más información, consulte la sección [Otorgar privilegios mínimos](#) en la Guía del usuario de IAM.
- **Habilitar MFA para operaciones confidenciales** – Para mayor seguridad, obligue a los usuarios de IAM a que utilicen la autenticación multifactor (MFA) para acceder a recursos u operaciones de API confidenciales. Para obtener más información, consulte [Uso de la autenticación multifactor \(MFA\) en AWS](#) en la Guía del usuario de IAM.
- **Utilizar condiciones de política para mayor seguridad** – En la medida en que sea práctico, defina las condiciones en las que sus políticas basadas en identidad permitan el acceso a un recurso. Por ejemplo, puede escribir condiciones para especificar un rango de direcciones IP permitidas desde el que debe proceder una solicitud. También puede escribir condiciones para permitir solicitudes solo en un intervalo

de hora o fecha especificado o para solicitar el uso de SSL o MFA. Para obtener más información, consulte [Elementos de la política de JSON de IAM: condición](#) en la Guía del usuario de IAM.

## Permitir a los usuarios crear planes de escalado

A continuación, se incluye un ejemplo de una política de permisos que permite a los usuarios crear planes de escalado.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "autoscaling-plans:*",
        "cloudwatch:PutMetricAlarm",
        "cloudwatch>DeleteAlarms",
        "cloudwatch:DescribeAlarms",
        "cloudformation:ListStackResources"
      ],
      "Resource": "*"
    }
  ]
}
```

Para que un usuario trabaje con un plan de escalado, debe tener permisos adicionales que le permitan trabajar con determinados recursos de su cuenta de AWS. Estos permisos se enumeran en [Permisos necesarios adicionales \(p. 29\)](#).

Cada usuario de la consola también necesita permisos que le permitan descubrir los recursos escalables de su cuenta de AWS y ver gráficos de datos de métricas de CloudWatch desde la consola de AWS Auto Scaling. El conjunto adicional de permisos necesarios para trabajar con la consola de AWS Auto Scaling se enumera a continuación:

- `cloudformation:ListStacks`: Para enumerar pilas de .
- `tag:GetTagKeys`: Buscar recursos escalables que contengan determinadas claves de etiqueta.
- `tag:GetTagValues`: para buscar recursos que contienen determinados valores de etiqueta.
- `autoscaling:DescribeTags`: Buscar grupos de Auto Scaling que contengan determinadas etiquetas.
- `cloudwatch:GetMetricData`: para ver datos en gráficos de métricas de recursos de AWS.

## Permitir a los usuarios habilitar el escalado predictivo

A continuación, se incluye un ejemplo de una política de permisos que permite a los usuarios habilitar el escalado predictivo. Estos permisos amplían las características de los planes de escalado que se configuran para escalar los grupos de Auto Scaling.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "cloudwatch:GetMetricData",
        "autoscaling:DescribeAutoScalingGroups",
        "autoscaling:DescribeScheduledActions",
        "autoscaling:BatchPutScheduledUpdateGroupAction",

```

```
        "autoscaling:BatchDeleteScheduledAction"  
      ],  
      "Resource": "*"   
    }  
  ]  
}
```

## Permisos necesarios adicionales

Al conceder permisos para AWS Auto Scaling, debe decidir para qué recursos los usuarios obtienen permisos. Según los escenarios que desee admitir, puede especificar las acciones siguientes en el elemento `Action` de una declaración de política de IAM.

### Grupos de Auto Scaling

Para agregar grupos de Auto Scaling a un plan de escalado, los usuarios deben tener los siguientes permisos de Amazon EC2 Auto Scaling:

- `autoscaling:UpdateAutoScalingGroup`
- `autoscaling:DescribeAutoScalingGroups`
- `autoscaling:PutScalingPolicy`
- `autoscaling:DescribePolicies`
- `autoscaling>DeletePolicy`

### Servicios de ECS

Para agregar servicios de ECS a un plan de escalado, los usuarios deben tener los siguientes permisos de Amazon ECS y Auto Scaling de aplicaciones:

- `ecs:DescribeServices`
- `ecs:UpdateService`
- `application-autoscaling:RegisterScalableTarget`
- `application-autoscaling:DescribeScalableTargets`
- `application-autoscaling:DeregisterScalableTarget`
- `application-autoscaling:PutScalingPolicy`
- `application-autoscaling:DescribeScalingPolicies`
- `application-autoscaling>DeleteScalingPolicy`

### Flota de subasta

Para agregar flotas de spot a un plan de escalado, los usuarios deben tener los siguientes permisos de Amazon EC2 y Auto Scaling de aplicaciones:

- `ec2:DescribeSpotFleetRequests`
- `ec2:ModifySpotFleetRequest`
- `application-autoscaling:RegisterScalableTarget`
- `application-autoscaling:DescribeScalableTargets`
- `application-autoscaling:DeregisterScalableTarget`
- `application-autoscaling:PutScalingPolicy`
- `application-autoscaling:DescribeScalingPolicies`



- `application-autoscaling:DeleteScalingPolicy`

### Tablas o índices globales de DynamoDB

Para agregar tablas de DynamoDB o índices globales a un plan de escalado, los usuarios deben tener los siguientes permisos de DynamoDB y Auto Scaling de aplicaciones:

- `dynamodb:DescribeTable`
- `dynamodb:UpdateTable`
- `application-autoscaling:RegisterScalableTarget`
- `application-autoscaling:DescribeScalableTargets`
- `application-autoscaling:DeregisterScalableTarget`
- `application-autoscaling:PutScalingPolicy`
- `application-autoscaling:DescribeScalingPolicies`
- `application-autoscaling:DeleteScalingPolicy`

### Clústeres de Aurora DB

Para agregar clústeres de Aurora DB a un plan de escalado, los usuarios deben tener los siguientes permisos de Amazon Aurora y Auto Scaling de aplicaciones:

- `rds:AddTagsToResource`
- `rds>CreateDBInstance`
- `rds>DeleteDBInstance`
- `rds:DescribeDBClusters`
- `rds:DescribeDBInstances`
- `application-autoscaling:RegisterScalableTarget`
- `application-autoscaling:DescribeScalableTargets`
- `application-autoscaling:DeregisterScalableTarget`
- `application-autoscaling:PutScalingPolicy`
- `application-autoscaling:DescribeScalingPolicies`
- `application-autoscaling:DeleteScalingPolicy`

## Permisos necesarios para crear un rol vinculado a un servicio

AWS Auto Scaling requiere permisos para crear un rol vinculado a servicio la primera vez que un usuario de su cuenta de AWS crea un plan de escalado con el escalado predictivo habilitado. Si el rol vinculado al servicio aún no existe, AWS Auto Scaling lo crea en su cuenta. El rol vinculado al servicio concede permisos a AWS Auto Scaling para que pueda llamar a otros servicios en su nombre.

Para que la creación automática de roles se realice correctamente, los usuarios deben disponer de permisos para la acción `iam:CreateServiceLinkedRole`.

```
"Action": "iam:CreateServiceLinkedRole"
```

A continuación, se muestra un ejemplo de una política de permisos que permite a un usuario crear un rol vinculado a servicio de AWS Auto Scaling.

```
{
```

```
"Version": "2012-10-17",
"Statement": [
  {
    "Effect": "Allow",
    "Action": "iam:CreateServiceLinkedRole",
    "Resource": "arn:aws:iam::*:role/aws-service-role/autoscaling-
plans.amazonaws.com/AWSServiceRoleForAutoScalingPlans_EC2AutoScaling",
    "Condition": {
      "StringLike": {
        "iam:AWSServiceName": "autoscaling-plans.amazonaws.com"
      }
    }
  }
]
```

Para obtener más información, consulte [Roles vinculados a servicios de AWS Auto Scaling \(p. 25\)](#).

## Registro y monitoreo en AWS Auto Scaling

La monitorización es una parte importante del mantenimiento de la fiabilidad, la disponibilidad y el rendimiento de AWS Auto Scaling y de las demás soluciones de AWS. Debe recopilar datos de monitorización de todas las partes de su solución de AWS para que le resulte más sencillo depurar un error que se produce en distintas partes del código, en caso de que ocurra. AWS proporciona las siguientes herramientas para registrar y monitorizar las actividades, así como para realizar acciones automáticas cuando corresponda:

### Alarmas de Amazon CloudWatch

Para detectar el comportamiento incorrecto de las aplicaciones, CloudWatch le ayuda a monitorizar automáticamente determinadas métricas de sus recursos de AWS. Puede configurar una alarma de CloudWatch y una notificación de Amazon SNS que envíe un correo electrónico cuando el valor de una métrica no sea el esperado o cuando se detecten determinadas anomalías. Por ejemplo, puede recibir una notificación cuando la actividad de la red de repente sea superior o inferior al valor esperado de una métrica. Para obtener más información, consulte [Guía del usuario de Amazon CloudWatch](#).

### AWS CloudTrail

AWS CloudTrail captura las llamadas a la API y los eventos relacionados realizados por su cuenta de AWS o en su nombre. A continuación, entrega los archivos de registro al bucket de Amazon S3 que especifique. También pueden identificar qué usuarios y cuentas llamaron a AWS, la dirección IP de origen desde la que se realizaron las llamadas y el momento en que se efectuaron las llamadas. Para obtener más información, consulte [AWS CloudTrail User Guide](#). Para obtener información sobre las llamadas a la API de AWS Auto Scaling registradas por CloudTrail, consulte [Registro de llamadas a la API de AWS Auto Scaling con CloudTrail](#).

### Amazon CloudWatch Logs

Amazon CloudWatch Logs permite monitorizar, almacenar y obtener acceso a sus archivos de registro desde instancias Amazon EC2, CloudTrail y otros orígenes. CloudWatch Logs puede monitorizar información en los archivos de registro y enviarle una notificación cuando se alcancen determinados umbrales. También se pueden archivar los datos de los registros en un almacenamiento de larga duración. Para obtener más información, consulte [Amazon CloudWatch Logs User Guide](#).

### Amazon EventBridge

EventBridge proporciona un flujo casi en tiempo real de eventos del sistema que describen los cambios realizados en los recursos de AWS. AWS Auto Scaling actualmente no emite eventos. Sin embargo, puede escribir reglas que desencadenen acciones automatizadas en otros servicios de AWS

como resultado de llamadas a la API realizadas por AWS Auto Scaling. Para obtener más información, consulte [Creación de una regla de EventBridge que se dispare en función de una llamada a la API de AWS utilizando AWS CloudTrail](#) en la Guía del usuario de Amazon EventBridge.

#### Temas relacionados

- [Monitorización de sus instancias y grupos de Auto Scaling](#) en la Guía del usuario de Amazon EC2 Auto Scaling
- [Auto Scaling de aplicaciones Monitorización de](#) en la Guía del usuario de Auto Scaling de aplicaciones

## Validación de la conformidad de AWS Auto Scaling

Hay auditores externos que evalúan la seguridad y la conformidad de los servicios de Amazon Web Services (AWS) en diferentes programas de conformidad de AWS. Esto incluye SOC, PCI, FedRAMP, HIPAA y otros.

Para obtener una lista de servicios de AWS en el ámbito de programas de conformidad específicos, consulte [Servicios de AWS en el ámbito del programa de conformidad](#). Para obtener información general, consulte [Programas de conformidad de AWS](#).

Puede descargar los informes de auditoría de terceros utilizando AWS Artifact. Para obtener más información, consulte [Descarga de informes en AWS Artifact](#).

Su responsabilidad en el ámbito de la conformidad al usar AWS Auto Scaling viene determinada por la confidencialidad de los datos, los objetivos de conformidad de su empresa y las leyes y regulaciones aplicables. AWS proporciona los siguientes recursos para ayudarlo con los requisitos de conformidad:

- [Guías de inicio rápido de seguridad y conformidad de](#) Guías de inicio rápido de seguridad y conformidad Estas guías de implementación tratan consideraciones sobre arquitectura y ofrecen pasos para implementar los entornos de referencia centrados en la seguridad y la conformidad en –AWS
- [Documento técnico sobre arquitectura para seguridad y conformidad de HIPAA](#) – este documento técnico describe cómo las empresas pueden utilizar AWS para crear aplicaciones conformes con HIPAA.
- [Recursos de conformidad de AWS](#) recursos de conformidad de –
- [Evaluateing resources with rules](#) en la AWS Config Guía para desarrolladores de – el servicio AWS Config evalúa en qué medida las configuraciones de los recursos cumplen las prácticas internas, las directrices del sector y las normativas.
- [AWS Security Hub](#) – Este servicio de AWS ofrece una vista integral de su estado de seguridad en AWS que le ayuda a comprobar la conformidad con las normas del sector de seguridad y las prácticas recomendadas.

## Resiliencia en AWS Auto Scaling

La infraestructura global de AWS está conformada por regiones y zonas de disponibilidad de AWS.

Las regiones de AWS proporcionan varias zonas de disponibilidad físicamente independientes y aisladas que se encuentran conectadas mediante redes con un alto nivel de rendimiento y redundancia, además de baja latencia.

Con las zonas de disponibilidad, puede diseñar y utilizar aplicaciones y bases de datos que realizan una conmutación por error automática entre las zonas sin interrupciones. Las zonas de disponibilidad tienen una mayor disponibilidad, tolerancia a errores y escalabilidad que las infraestructuras tradicionales de centros de datos únicos o múltiples.

Para obtener más información sobre las zonas de disponibilidad y las regiones de AWS, consulte [Infraestructura global de AWS](#).

## Seguridad de la infraestructura de AWS Auto Scaling

Al ser un servicio administrado, AWS Auto Scaling está protegido por los procedimientos de seguridad de red globales de AWS que se describen en la [Amazon Web Services: Documento técnico Información general sobre procesos de seguridad de](#) .

Puede utilizar llamadas a la API publicadas en AWS para acceder a AWS Auto Scaling a través de la red. Los clientes deben ser compatibles con Transport Layer Security (TLS) 1.0 o una versión posterior. Le recomendamos TLS 1.2 o una versión posterior. Los clientes también deben ser compatibles con conjuntos de cifrado con confidencialidad directa total (PFS) tales como Ephemeral Diffie-Hellman (DHE) o Elliptic Curve Ephemeral Diffie-Hellman (ECDHE). La mayoría de los sistemas modernos como Java 7 y posteriores son compatibles con estos modos.

Además, las solicitudes deben estar firmadas mediante un ID de clave de acceso y una clave de acceso secreta que esté asociada a una entidad principal de IAM. También puede utilizar [AWS Security Token Service](#) (AWS STS) para generar credenciales de seguridad temporales para firmar solicitudes.

## AWS Auto Scaling y puntos de enlace de la VPC de tipo interfaz

Puede establecer una conexión privada entre su nube virtual privada (VPC) y la API de AWS Auto Scaling creando un punto de enlace de la VPC de tipo interfaz. Puede utilizar esta conexión para llamar a la API de AWS Auto Scaling desde su VPC sin enviar tráfico por Internet. El punto de enlace proporciona una conectividad de confianza y escalable a la API de AWS Auto Scaling. Lo hace sin necesidad de una gateway de Internet, instancia NAT o conexión de VPN.

Los puntos de enlace de la VPC de tipo interfaz utilizan la tecnología de AWS PrivateLink, una característica que permite la comunicación privada entre los servicios de AWS mediante direcciones IP privadas. Para obtener más información, consulte [AWS PrivateLink](#) .

### Creación de un punto de enlace de la VPC de tipo interfaz

Puede crear un punto de enlace de la VPC para el servicio de AWS Auto Scaling mediante la consola de Amazon VPC o AWS Command Line Interface (AWS CLI). Cree un punto de enlace para AWS Auto Scaling utilizando el siguiente nombre de servicio:

- `com.amazonaws.region.autoscaling-plans` — Crea un punto de enlace para las operaciones de la AWS Auto Scaling API de .

Para obtener más información, consulte [Creación de un punto de enlace de interfaz](#) en la Guía del usuario de Amazon VPC.

Habilite el DNS privado para el punto de enlace que va a realizar solicitudes de API al servicio compatible mediante su nombre de host DNS predeterminado (por ejemplo, `autoscaling-plans.us-`

east-1.amazonaws.com). Cuando se crea un punto de enlace para los servicios de AWS, esta configuración se habilita de forma predeterminada. Para obtener más información, consulte [Acceso a un servicio a través de un punto de enlace de interfaz](#) en la Guía del usuario de Amazon VPC.

No es necesario cambiar ninguna AWS Auto Scaling configuración. AWS Auto Scaling llama a otros servicios de AWS mediante puntos de enlace de servicio o puntos de enlace de la VPC de tipo interfaz privados, los que estén en uso.

## Crear una política de punto de enlace de la VPC

Puede asociar una política a su punto de enlace de la VPC para controlar el acceso a la API de AWS Auto Scaling. La política específica:

- La entidad principal que puede realizar acciones.
- Las acciones que se pueden realizar.
- El recurso en el que se pueden realizar las acciones.

En el ejemplo siguiente se muestra una política de punto de enlace de la VPC que deniega a todos los usuarios el permiso para eliminar un plan de escalado a través del punto de enlace. La política de ejemplo también concede permiso a todos los usuarios para realizar todas las demás acciones.

```
{
  "Statement": [
    {
      "Action": "*",
      "Effect": "Allow",
      "Resource": "*",
      "Principal": "*"
    },
    {
      "Action": "autoscaling-plans:DeleteScalingPlan",
      "Effect": "Deny",
      "Resource": "*",
      "Principal": "*"
    }
  ]
}
```

Para obtener más información, consulte [Uso de políticas de punto de enlace de la VPC](#) en la Guía del usuario de Amazon VPC.

## Migración de puntos de enlace

El 22 de noviembre de 2019, AWS Auto Scaling introdujo `autoscaling-plans.region.amazonaws.com` como el nuevo nombre de host DNS predeterminado y punto de enlace para las llamadas a la API de AWS Auto Scaling. El nuevo punto de enlace es compatible con la última versión de la CLI de AWS y SDKs. Si aún no lo ha hecho, instale la última CLI de AWS y SDKs para utilizar el nuevo punto de enlace. Para actualizar la AWS CLI, consulte [Instalación de la CLI de AWS mediante pip](#) en la AWS Command Line Interface Guía del usuario. Para obtener más información acerca de AWS SDKs, consulte [Herramientas para Amazon Web Services](#).

### Important

Para la compatibilidad con versiones anteriores, el punto de enlace `autoscaling.region.amazonaws.com` existente seguirá siendo compatible con las llamadas a la API de AWS Auto Scaling. Para configurar el `autoscaling.region.amazonaws.com` punto de enlace como un punto de enlace de la VPC de tipo interfaz privado, consulte [Amazon](#)

[EC2 Auto Scaling y Puntos de enlace de la VPC de tipo interfaz](#) en la Guía del usuario de Amazon EC2 Auto Scaling .

Punto de enlace al que se debe llamar cuando se utiliza la CLI o la API de AWS

Para la versión actual de AWS Auto Scaling, sus llamadas a la API de AWS Auto Scaling se dirigen automáticamente al punto de enlace `autoscaling-plans.region.amazonaws.com` en lugar de a `autoscaling.region.amazonaws.com`.

Puede llamar al nuevo punto de enlace en la CLI utilizando el siguiente parámetro con cada comando para especificar el punto de enlace: `--endpoint-url https://autoscaling-plans.region.amazonaws.com`.

Aunque no se recomienda, también puede llamar al punto de enlace anterior en la CLI utilizando el siguiente parámetro con cada comando para especificar el punto de enlace: `--endpoint-url https://autoscaling.region.amazonaws.com`.

Para los distintos SDKs utilizados para llamar a la APIs, consulte la documentación del SDK de interés para aprender a dirigir las solicitudes a un punto de enlace específico. Para obtener más información, consulte [Herramientas para Amazon Web Services](#).

# Cuotas de servicio de AWS Auto Scaling

Su cuenta de AWS tiene las siguientes cuotas predeterminadas para AWS Auto Scaling (anteriormente se denominaban "límites").

Para solicitar un aumento, utilice el [formulario de límites de Auto Scaling](#). No olvide especificar el tipo de recurso con la solicitud de aumento de la cuota; por ejemplo, Amazon EC2 Auto Scaling, Amazon ECS o DynamoDB.

Cuotas predeterminadas por región y cuenta

Elemento	Valor predeterminado
Número máximo de recursos escalables por tipo de recurso	Las cuotas varían según el tipo de recurso.  Amazon DynamoDB: 3000  Grupos de Amazon EC2 Auto Scaling: 200  Todos los demás tipos de recursos: 500
Número máximo de planes de escalado	100
Número máximo de instrucciones de escalado por plan de escalado	500
Número máximo de configuraciones de seguimiento de destino por instrucción de escalado	10

Tenga en cuenta las cuotas del servicio a medida que escale las cargas de trabajo. Por ejemplo, cuando alcance el número máximo de unidades de capacidad permitidas por un servicio, el escalado se detendrá. Si la demanda cae y la capacidad actual disminuye, AWS Auto Scaling puede realizar de nuevo el escalado. Para evitar volver a alcanzar este límite de cuota de servicio, puede solicitar un aumento. Cada servicio tiene sus propias cuotas predeterminadas para la capacidad máxima del recurso. Para obtener información sobre las cuotas predeterminadas de otros servicios de AWS, consulte [Puntos de enlace y cuotas de servicio](#) en la Referencia general de Amazon Web Services.

# Recursos de AWS Auto Scaling

Los recursos relacionados siguientes pueden serle de ayuda cuando trabaje con este servicio.

- [AWS Auto Scaling](#) – Página web principal para obtener información sobre AWS Auto Scaling.
- [Preguntas frecuentes de AWS Auto Scaling](#) – las respuestas a las preguntas de los clientes sobre AWS Auto Scaling.
- [AWS Auto Scaling: foro de debate de](#) : –obtiene ayuda de la comunidad.
- [Etiquetado de grupos e instancias de](#) : [Auto Scaling obtenga información sobre cómo etiquetar sus grupos de](#) .–Auto Scaling
- [Etiquetado de DynamoDB](#) – Obtenga información sobre cómo etiquetar las tablas de Amazon DynamoDB o los índices secundarios globales.
- [Etiquetado de recursos de Amazon RDS](#): obtenga información sobre cómo etiquetar los clústeres de base de datos de .–Aurora
- [Trabajar con Tag Editor](#): obtenga información sobre cómo usar Tag Editor, incluidos los recursos que admite Tag Editor.–
- [Target tracking scaling policies \(Políticas de escalado de seguimiento de destino\)](#) para Amazon EC2 Auto Scaling obtenga información sobre las políticas de escalado de seguimiento de destino para los grupos de .–Amazon EC2 Auto Scaling
- [Target tracking scaling policies \(Políticas de escalado de seguimiento de destino\)](#) para todos los demás recursos obtenga información sobre las políticas de escalado de seguimiento de destino para otros recursos distintos de –, como los índices y las tablas de Amazon EC2 y los servicios de DynamoDB. Amazon ECS
- [AWS Auto Scaling Guías de referencia de la API y la CLI de](#) Documentación para las llamadas a la API y los comandos de la – que puede utilizar para crear, modificar y eliminar planes de AWS CLI. Auto Scaling
- [Registro de llamadas a la API con CloudTrail](#) – obtenga información sobre el monitoreo de las llamadas realizadas a la API de su cuenta, incluidas las llamadas realizadas por la Consola de administración de AWS, las herramientas de línea de comandos y otros servicios.

Los siguientes recursos adicionales están disponibles para ayudarle a obtener más información acerca de AWS.

- [Clases y talleres](#): enlaces a cursos basados en roles y especializados, y también a laboratorios autoguiados para ayudarle a desarrollar sus conocimientos de AWS y obtener experiencia práctica.
- [Herramientas para desarrolladores de AWS](#): enlaces a herramientas para desarrolladores, SKD, conjuntos de herramientas de IDE y herramientas de línea de comandos para desarrollar y administrar aplicaciones de AWS.
- [Documentos técnicos de AWS](#): enlaces a una completa lista de documentos técnicos de AWS que cubre una gran variedad de temas técnicos, como arquitecturas, seguridad y economía de la nube, escritos por arquitectos de soluciones de AWS o expertos técnicos.
- [Centro de soporte de AWS](#): centro para crear y administrar sus casos de AWS Support. También incluye enlaces a otros recursos útiles como foros, preguntas técnicas frecuentes, estado de los servicios y AWS Trusted Advisor.
- [AWS Support](#): página web principal para obtener información sobre AWS Support, un canal de soporte individualizado y de respuesta rápida que le ayudará a crear y ejecutar aplicaciones en la nube.
- [Contacte con nosotros](#): un punto central de contacto para las consultas relacionadas con la facturación de AWS, cuentas, eventos, uso indebido y otros problemas.
- [Términos del sitio de AWS](#): información detallada sobre nuestros derechos de autor y marca comercial, su cuenta, licencia y acceso al sitio, entre otros temas.



# Historial de revisión

En la siguiente tabla, se describen los cambios importantes de la documentación de AWS Auto Scaling. Para obtener notificaciones sobre las actualizaciones de esta documentación, puede suscribirse a la fuente RSS.

update-history-change	update-history-description	update-history-date
<a href="#">Nuevo capítulo sobre seguridad (p. 38)</a>	El nuevo capítulo <a href="#">Seguridad</a> de la Guía del usuario de AWS Auto Scaling ayuda a entender cómo aplicar el <a href="#">modelo de responsabilidad compartida</a> cuando se usa AWS Auto Scaling. Durante esta actualización, el capítulo de la guía del usuario "Autenticación y control de acceso" se ha sustituido por una nueva sección más útil, <a href="#">Identity and Access Management para AWS Auto Scaling</a> .	March 12, 2020
<a href="#">Compatibilidad con los puntos de enlace de la VPC de Amazon (p. 38)</a>	Ahora puede establecer una conexión privada entre su VPC y AWS Auto Scaling. Para obtener consideraciones e instrucciones sobre la migración, consulte <a href="#">AWS Auto Scaling y puntos de enlace de la VPC de tipo interfaz</a> .	November 22, 2019
<a href="#">Compatibilidad para aumentar la capacidad máxima por encima de la capacidad de previsión, además de cambios en la guía (p. 38)</a>	Se añadido compatibilidad a la consola para permitir que el plan de escalado aumente la capacidad máxima por encima de la capacidad de previsión en un valor de búfer especificado. Para obtener más información, consulte <a href="#">Configuración del escalado predictivo</a> en la Guía del usuario de AWS Auto Scaling. Esta versión también incluye varias secciones reescritas en el tutorial <a href="#">Introducción a AWS Auto Scaling</a> .	March 9, 2019
<a href="#">Escalado predictivo y mejoras (p. 38)</a>	Ahora puede utilizar el escalado predictivo para escalar de forma proactiva los grupos de Amazon EC2 Auto Scaling. Esta versión también permite sustituir las políticas de escalado creadas fuera del plan de escalado (por ejemplo, desde otras consolas)	November 20, 2018

	y controlar la habilitación de la característica de escalado dinámico del plan. Para obtener más información, consulte <a href="#">Introducción a AWS Auto Scaling</a> .	
<a href="#">Compatibilidad con la configuración personalizada de los recursos (p. 38)</a>	Se ha añadido compatibilidad para personalizar diversos valores de configuración para cada recurso individual o varios recursos al mismo tiempo. Para obtener más información, consulte <a href="#">Introducción a AWS Auto Scaling</a> .	October 9, 2018
<a href="#">Etiquetas como un origen de aplicación (p. 38)</a>	Esta versión incorpora soporte para especificar un conjunto de etiquetas como un origen de aplicación.	April 23, 2018
<a href="#">Nuevo servicio (p. 38)</a>	Versión inicial de AWS Auto Scaling.	January 16, 2018

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.