

Auto Scaling

Guia de usuário

Edição 01
Data 30-10-2021



Copyright © Huawei Technologies Co., Ltd. 2023. Todos os direitos reservados.

Nenhuma parte deste documento pode ser reproduzida ou transmitida em qualquer forma ou por qualquer meio sem consentimento prévio por escrito da Huawei Technologies Co., Ltd.

Marcas registadas e permissões



HUAWEI e outras marcas registadas da Huawei são marcas registadas da Huawei Technologies Co., Ltd.

Todos as outras marcas registadas e os nomes registados mencionados neste documento são propriedade dos seus respectivos detentores.

Aviso

Os produtos, serviços e funcionalidades adquiridos são estipulados pelo contrato feito entre a Huawei e o cliente. Todos ou parte dos produtos, serviços e funcionalidades descritos neste documento pode não estar dentro do âmbito de aquisição ou do âmbito de uso. Salvo especificação em contrário no contrato, todas as declarações, informações e recomendações neste documento são fornecidas "TAL COMO ESTÁ" sem garantias, ou representações de qualquer tipo, seja expressa ou implícita.

As informações contidas neste documento estão sujeitas a alterações sem aviso prévio. Foram feitos todos os esforços na preparação deste documento para assegurar a exatidão do conteúdo, mas todas as declarações, informações e recomendações contidas neste documento não constituem uma garantia de qualquer tipo, expressa ou implícita.

Índice

1 Grupo de AS.....	1
1.1 Criação de um grupo de AS.....	1
1.2 (Opcional) Adição de um balanceador de carga a um grupo de AS.....	8
1.3 Alteração da configuração de AS para um grupo de AS.....	8
1.4 Habilitação de um grupo de AS.....	9
1.5 Desabilitação de um grupo de AS.....	10
1.6 Modificação de um grupo de AS.....	10
1.7 Exclusão de um grupo de AS.....	11
2 Configuração de AS.....	13
2.1 Criação de uma configuração do AS.....	13
2.2 Criação de uma configuração do AS a partir de uma instância do ECS existente.....	13
2.3 Criação de uma configuração do AS a partir de um novo modelo de especificações.....	16
2.4 Cópia de uma configuração do AS.....	21
2.5 Exclusão de uma configuração do AS.....	21
3 Política de AS.....	22
3.1 Visão geral.....	22
3.2 Criação de uma política de AS.....	23
3.3 Gerenciamento de políticas de AS.....	32
4 Ação de dimensionamento.....	34
4.1 Dimensionamento dinâmico.....	34
4.2 Dimensionamento agendado.....	36
4.3 Dimensionamento manual.....	36
4.4 Configuração de uma política de remoção de instância.....	38
4.5 Visualização de uma ação de dimensionamento.....	39
4.6 Gerenciamento de ganchos do ciclo de vida.....	39
4.7 Configuração da proteção de instância.....	46
4.8 Colocar uma instância no estado de espera.....	47
5 Dimensionamento da largura de banda.....	49
5.1 Criação de uma política de dimensionamento de largura de banda.....	49
5.2 Exibição de detalhes sobre uma política de dimensionamento de largura de banda.....	55
5.3 Gerenciamento de uma política de dimensionamento de largura de banda.....	56

6 Monitoramento de instância e grupo de AS.....	59
6.1 Verificação de integridade.....	59
6.2 Configuração de notificação para um grupo de AS.....	60
6.3 Gravação de operações de recursos de AS.....	61
6.4 Adição de tags a grupos e instâncias de AS.....	64
6.5 Monitoramento de métricas.....	66
6.6 Exibição de métricas de monitoramento.....	71
6.7 Configuração de regras de alarme de monitoramento.....	72
7 Gerenciamento de permissões.....	73
7.1 Criação de um usuário e concessão de permissões de AS.....	73
7.2 Políticas personalizadas do AS.....	74
A Histórico de mudanças.....	76

1 Grupo de AS

1.1 Criação de um grupo de AS

Cenários

Um grupo de AS consiste em uma coleção de instâncias e políticas de AS que têm atributos semelhantes e se aplicam ao mesmo cenário de aplicação. Um grupo do AS é a base para habilitar ou desabilitar políticas do AS e executar ações de dimensionamento. A política do AS pré-configurada automaticamente adiciona ou exclui instâncias de ou para um grupo do AS, ou mantém um número fixo de instâncias em um grupo do AS.

Ao criar um grupo de AS, especifique uma configuração de AS para ele. Além disso, adicione uma ou mais políticas do AS para o grupo do AS.

A criação de um grupo de AS envolve a configuração dos números máximo, mínimo e esperado de instâncias e do balanceador de carga associado.

Observações

Os tipos de ECS disponíveis em diferentes AZs podem variar. Ao criar um grupo de AS, escolha uma configuração de AS que use um tipo de ECS disponível nas AZs usadas pelo grupo de AS.

- Se o tipo de ECS especificado na configuração de AS não estiver disponível em nenhuma das AZs usadas pelo grupo de AS, ocorrerão as seguintes situações:
 - Se o grupo de AS estiver desabilitado, ele não poderá ser habilitado novamente mais tarde.
 - Se o grupo de AS estiver habilitado, seu status se tornará anormal quando instâncias forem adicionadas a ele.
- Se o tipo de ECS especificado na configuração de AS estiver disponível apenas em determinadas AZs usadas pelo grupo de AS, as instâncias de ECS adicionadas por uma ação de dimensionamento serão implementadas apenas nas AZs em que esse tipo de ECS estiver disponível. Como resultado, as instâncias no grupo de AS podem não ser distribuídas uniformemente.

Procedimento

1. Log in to the management console.
2. Under **Compute**, click **Auto Scaling**. In the navigation pane on the left, choose **Instance Scaling**.
3. Click **Create AS Group**.
4. Set parameters, such as **Name**, **Max. Instances**, **Min. Instances**, and **Expected Instances**. [Tabela 1-1](#) describes the key parameters to be configured.

Tabela 1-1 Parâmetros do grupo de AS

Parâmetro	Descrição	Exemplo de valor
Region	Uma região é onde o grupo de AS é implementado. Recursos em diferentes regiões não podem se comunicar uns com os outros através de redes internas. Para menor latência de rede e acesso mais rápido aos recursos, selecione a região mais próxima dos usuários de destino.	N/A
AZ	Uma AZ é um local físico onde os recursos usam fontes de alimentação e redes independentes. As AZs são fisicamente isoladas, mas interconectadas por meio de uma rede interna. <ul style="list-style-type: none">● Se você precisar de alta disponibilidade, compre servidores em AZs diferentes.● Se você precisar de baixa latência de rede, compre servidores na mesma AZ.	N/A
Multi-AZ Scaling Policy	Este parâmetro pode ser definido como Balanced ou Sequenced . <ul style="list-style-type: none">● Balanced: ao dimensionar um grupo de AS, o sistema preferencialmente distribui instâncias do ECS uniformemente entre as AZs usadas pelo grupo do AS. Se falhar na AZ de destino, selecionará automaticamente outra AZ com base na política sequenciada.● Sequenced: ao dimensionar um grupo do AS, o sistema distribui instâncias do ECS para a AZ selecionada de acordo com a ordem em que as AZs são especificadas. NOTA Esse parâmetro precisa ser configurado quando duas ou mais AZs forem selecionadas.	Balanced
Name	Especifica o nome do grupo de AS a ser criado. O nome pode conter até 64 caracteres. Pode incluir letras, dígitos, sublinhados (_) e hifens (-).	N/A
Max. Instances	Especifica o número máximo de instâncias do ECS em um grupo de AS.	1

Parâmetro	Descrição	Exemplo de valor
Expected Instances	Especifica o número esperado de instâncias do ECS em um grupo de AS. Depois que um grupo de AS é criado, você pode alterar esse valor, o que acionará uma ação de dimensionamento.	0
Min. Instances	Especifica o número mínimo de instâncias do ECS em um grupo de AS.	0
AS configuration	Especifica a configuração de AS necessária para o grupo de AS. Uma configuração do AS define as especificações das instâncias do ECS a serem adicionadas a um grupo de AS. As especificações incluem a imagem do ECS e o tamanho do disco do sistema. Você precisa criar a configuração do AS necessária antes de criar um grupo de AS.	N/A
VPC	Fornecer uma rede para suas instâncias do ECS. Todas as instâncias do ECS no grupo do AS são implementadas nesta VPC.	N/A
Subnet	Você pode selecionar até cinco sub-redes. O grupo de AS liga automaticamente todas as NICs às instâncias do ECS criadas. A primeira sub-rede é usada pela NIC principal de uma instância do ECS por padrão, e outras sub-redes são usadas pelas NICs de extensão da instância.	N/A

Parâmetro	Descrição	Exemplo de valor
Load Balancing	<p>Este parâmetro é opcional. Um balanceador de carga distribui automaticamente o tráfego em todas as instâncias de um grupo de AS para equilibrar sua carga de serviço. Ele melhora a tolerância a falhas de suas aplicações e expande os recursos do serviço de aplicações.</p> <p>NOTA</p> <ul style="list-style-type: none"> ● Até seis balanceadores de carga podem ser adicionados a um grupo de AS. ● Depois que vários balanceadores de carga são adicionados a um grupo de AS, vários serviços podem ser ouvidos simultaneamente, melhorando assim a escalabilidade do serviço. Se ELB health check for selecionada para Health Check Method, quando qualquer um dos ouvintes detectar que uma instância se torna insalubre, o AS a substituirá por uma nova. <p>Se você selecionar Elastic load balancer, configure os seguintes parâmetros:</p> <ul style="list-style-type: none"> ● Load Balancer ● Backend ECS Group ● Backend Port: especifica a porta na qual um ECS de back-end escuta o tráfego. ● Weight: determina a parte das solicitações que um ECS de back-end processa em comparação com outros ECSs de back-end adicionados ao mesmo ouvinte. <p>Para obter mais informações sobre balanceamento de carga, consulte Guia de usuário do Elastic Load Balance.</p>	N/A

Parâmetro	Descrição	Exemplo de valor
Instance Removal Policy	<p>Controla quais instâncias são as primeiras a serem removidas durante o dimensionamento. Se as condições especificadas forem atendidas, as ações de dimensionamento serão acionadas para remover instâncias. Você pode escolher qualquer uma das seguintes políticas de remoção de instância:</p> <ul style="list-style-type: none"> ● Instância mais antiga criada a partir da configuração de AS mais antiga: a instância mais antiga criada a partir da configuração mais antiga é removida primeiro do grupo de AS. ● Instância mais recente criada a partir da configuração de AS mais antiga: a instância mais recente criada a partir da configuração mais antiga é removida primeiro do grupo de AS. ● Instância mais antiga: a instância mais antiga é removida primeiro do grupo de AS. ● Instância mais recente: a instância mais recente é removida primeiro do grupo de AS. <p>NOTA</p> <ul style="list-style-type: none"> ● O AS sempre equilibra as instâncias em AZs primeiro, independentemente de qual política de remoção é usada. ● As instâncias do ECS adicionadas manualmente são as últimas a serem removidas. Se o AS remover uma instância adicionada manualmente, ele removerá apenas a instância do grupo de AS. Ele não exclui instância. Se várias instâncias adicionadas manualmente precisarem ser removidas, o AS preferencialmente removerá primeiro a instância adicionada mais cedo. 	Instância mais antiga criada a partir da configuração de AS mais antiga
EIP	<p>Se EIP tiver sido selecionado em uma configuração de AS para um grupo de AS, um EIP será automaticamente vinculado à instância do ECS adicionada por uma ação de dimensionamento ao grupo de AS. Se você selecionar Release, o EIP vinculado a uma instância será liberado quando a instância for removida do grupo de AS. Caso contrário, o sistema desvincula o EIP da instância, mas não o libera quando a instância é removida do grupo de AS.</p>	N/A
Data Disk	<p>Se Data Disk estiver configurado na configuração de AS usada pelo grupo de AS, um disco de dados será criado automaticamente e anexado às instâncias do ECS adicionadas durante uma ação de dimensionamento ao grupo de AS.</p> <p>Se você selecionar Release, os discos de dados anexados a uma instância serão excluídos quando a instância for removida do grupo de AS. Caso contrário, o sistema desanexa os discos de dados da instância, mas não os libera quando a instância é removida do grupo de AS.</p>	N/A

Parâmetro	Descrição	Exemplo de valor
Health Check Method	<p>Quando uma verificação de integridade detecta uma instância do ECS não íntegra, o AS a substitui por uma nova. Você pode escolher um dos seguintes métodos de verificação de integridade:</p> <ul style="list-style-type: none"> ● ECS health check: verifica o status de integridade da instância do ECS. Se uma instância for interrompida ou excluída, ela é considerada insalubre. Este método é selecionado por padrão. Usando esse método, o grupo de AS avalia periodicamente o status de execução de cada instância com base nos resultados da verificação de integridade. Se os resultados da verificação de integridade mostrarem que uma instância não é saudável, o AS removerá a instância do grupo de AS. ● ELB health check: determina o status de execução do ECS usando um ouvinte de balanceamento de carga. Quando um ouvinte de balanceamento de carga detecta que uma instância não está saudável, o AS remove a instância do grupo de AS. 	N/A
Health Check Interval	Especifica o período de tempo entre as verificações de saúde. Você pode definir um intervalo de verificação de integridade, como 5 minutos, 15 minutos, 1 hora ou 3 horas, com base nos requisitos de serviço.	5 minutos
Enterprise Project	<p>Especifica o projeto da empresa ao qual o grupo de AS pertence. Se um projeto empresarial é configurado para um grupo AS, os ECSs criados neste grupo AS também pertencem a este projeto empresarial. Se você não especificar um projeto corporativo, o projeto corporativo default será usado.</p> <p>NOTA</p> <ul style="list-style-type: none"> ● Valor default indica o projeto corporativo padrão. Os recursos que não estão alocados a nenhum projeto da empresa na sua conta são exibidos no projeto da empresa padrão. ● O projeto corporativo é uma versão atualizada do IAM. Ele aloca e gerencia recursos de diferentes projetos. 	N/A

Parâmetro	Descrição	Exemplo de valor
Tag	<p>Se você tiver muitos recursos do mesmo tipo, poderá usar tags para gerenciar seus recursos com mais facilidade. Você pode identificar recursos especificados rapidamente usando as tags alocadas a eles.</p> <p>Cada tag contém uma chave e um valor. Você pode especificar a chave e o valor para cada tag.</p> <ul style="list-style-type: none"> ● Chave <ul style="list-style-type: none"> - A chave deve ser especificada. - A chave deve ser exclusiva para o grupo de AS. - A chave pode incluir até 36 caracteres. Ela não pode conter caracteres ASCII não imprimíveis (0 - 31) ou os seguintes caracteres: =* < > \, / ● Valor <ul style="list-style-type: none"> - O valor é opcional. - Uma chave pode ter apenas um valor. - O valor pode incluir até 43 caracteres. Ela não pode conter caracteres ASCII não imprimíveis (0 - 31) ou os seguintes caracteres: =* < > \, / 	N/A

5. Clique em **Next**.
 6. Na página exibida, você pode usar uma configuração do AS existente ou criar uma configuração do AS.
 7. Clique em **Next**.
 8. (Opcional) Adicione uma política do AS a um grupo do AS.
Na página exibida, clique em **Add AS Policy**.
Configure os parâmetros necessários, como **Policy Type**, **Scaling Action** e **Cooldown Period**.
-  **NOTA**
- Se uma ação de dimensionamento for acionada por uma política de AS, o período de recarga será o que estiver configurado para essa política de AS.
 - Se uma ação de dimensionamento for acionada alterando manualmente o número esperado de instâncias ou por outras ações, o período de resfriamento será o que estiver configurado para o grupo de AS. O período de resfriamento padrão é de 300 segundos.
9. Clique em **Create Now**.
 10. Verifique o grupo de AS, a configuração de AS, e as informações de política de AS. Clique em **Submit**.
 11. Confirme o resultado da criação e volte para a página **AS Groups** conforme solicitado. Depois que o grupo de AS é criado, seu status muda para **Enabled**.

1.2 (Opcional) Adição de um balanceador de carga a um grupo de AS

Elastic Load Balance (ELB) distribui automaticamente o tráfego de entrada em vários servidores de back-end com base em políticas de encaminhamento configuradas. O ELB expande os recursos de serviço das aplicações e melhora sua disponibilidade eliminando pontos únicos de falha (SPOFs).

Se as funções do ELB forem necessárias, execute as operações fornecidas nesta seção para adicionar um balanceador de carga ao seu grupo de AS. O balanceador de carga adicionado a um grupo de AS distribui o tráfego da aplicação para todas as instâncias do grupo de AS quando uma instância é adicionada ou excluída do grupo de AS.

Somente um balanceador de carga criado pode ser vinculado a um grupo de AS, e o grupo de AS e o balanceador de carga devem estar na mesma VPC. Para obter detalhes sobre como criar um balanceador de carga, consulte *Guia de usuário do Elastic Load Balance*. Para adicionar um balanceador de carga para um grupo de AS, execute as seguintes operações:

- Ao criar um grupo de AS, configure o parâmetro **Load Balancing** para adicionar um balanceador de carga. Para mais detalhes, consulte [Criação de um grupo de AS](#).
- Se um grupo de AS não tiver nenhuma ação de dimensionamento em andamento, modifique o parâmetro **Load Balancing** para adicionar um balanceador de carga. Para mais detalhes, consulte [Modificação de um grupo de AS](#).

1.3 Alteração da configuração de AS para um grupo de AS

Cenários

Se você precisar alterar as especificações de instâncias do ECS em um grupo de AS, alterar a configuração do AS usada pelo grupo de AS é uma maneira fácil de ajudá-lo a chegar lá.

Tempo efetivo da nova configuração de AS

Depois de alterar a configuração de AS para um grupo de AS, a nova configuração de AS não será usada até que todas as ações de dimensionamento em andamento sejam concluídas.

Por exemplo, se houver uma ação de dimensionamento em andamento para um grupo de AS e você alterar a configuração de AS do grupo de AS de **as-config-A** para **as-config-B**, **as-config-A** ainda é usado para as instâncias que estão sendo adicionadas na ação de dimensionamento em andamento.

as-config-B entrará em vigor na próxima ação de dimensionamento.

Figura 1-1 Alteração da configuração do AS

Name	Status	Specifications	Image	System Disk	Data Disks	Login Mode	Created	Billing Mode	Operation
as-config-B	Unbound	kc1.large.2 2 vCPUs 4 GB	CentOS 8.0 64bit with ARM	High I/O 40 GB	0	Password	Oct 09, 2020 15:17:52 ...	Pay-per-use	Copy Delete
as-config-A	Unbound	s2.small.1 1 vCPUs 1 GB	CentOS 7.6 64bit	High I/O 40 GB	0	Password	Oct 09, 2020 14:42:33 ...	Pay-per-use	Copy Delete

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Clique no nome do grupo de AS para o qual você deseja alterar a configuração de AS. Na página **Basic Information**, clique em **Change Configuration** à direita de **Configuration Name**.
Você também pode localizar a linha que contém o grupo de destino de AS e escolher **More > Change Configuration** na coluna **Operation**.
4. Na caixa de diálogo **Change AS Configuration** exibida, selecione outra configuração de AS a ser usada pelo grupo de AS.
5. Clique em **OK**.

1.4 Habilitação de um grupo de AS

Cenários

Você pode ativar um grupo de AS para dimensionar automaticamente a capacidade de entrada ou saída.

Depois que um grupo de AS é habilitado, seu status muda para **Enabled**. AS monitora a política de AS e aciona uma ação de dimensionamento para grupos de AS somente no estado **Enabled**. Depois que um grupo de AS é ativado, o AS aciona uma ação de dimensionamento para adicionar ou remover instâncias automaticamente se o número de instâncias no grupo de AS for diferente do número esperado de instâncias.

- Somente os grupos de AS no estado **Disabled** podem ser habilitados.
- Somente os grupos de AS no estado **Abnormal** podem ser ativados à força. Você pode escolher **More > Forcibly Enable** para habilitar um grupo de AS anormal. A habilitação forçada de um grupo de AS não tem consequências adversas.
- Depois de criar um grupo de AS e adicionar uma configuração de AS a um grupo de AS, o grupo de AS é ativado automaticamente.

Habilitação de um grupo de AS

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Na lista de grupo de AS, localize a linha que contém o grupo de AS de destino e clique em **Enable** na coluna **Operation**. Você também pode clicar no nome do grupo de AS e, em seguida, em **Enable** no canto superior direito da página para habilitar o grupo de AS.
4. Na caixa de diálogo **Enable AS Group**, clique em **Yes**.

Habilitação forçada de um grupo de AS

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.

3. Na lista de grupo de AS, localize a linha que contém o grupo de AS de destino e selecione **Forcibly Enable** na lista suspensa **More** na coluna **Operation**. Você também pode clicar no nome do grupo de AS e, em seguida, em **Forcibly Enable** no canto superior direito da página para habilitar o grupo de AS.
4. Na caixa de diálogo **Forcibly Enable AS Group**, clique em **Yes**.

1.5 Desabilitação de um grupo de AS

Cenários

Se você precisar parar uma instância em um grupo de AS para configuração ou atualização, desative o grupo de AS antes de executar a operação. Isso evita que a instância seja excluída em uma verificação de integridade. Quando o status da instância for restaurado, você poderá ativar o grupo de AS novamente.

Se uma ação de dimensionamento continua falhando e sendo repetida (a causa da falha pode ser visualizada na página **Elastic Cloud Server**) para um grupo de AS, use um dos seguintes métodos para impedir que a ação seja repetida:

- Desative o grupo de AS. Em seguida, após a ação de dimensionamento falhar, ela não será repetida. Habilite o grupo de AS novamente quando o ambiente for recuperado ou depois de substituir a configuração de AS.
- Desative o grupo de AS e altere o número esperado de instâncias para o número de instâncias existentes. Em seguida, após a ação de dimensionamento falhar, a ação de dimensionamento não será repetida.

Depois que um grupo de AS é desabilitado, seu status muda para **Disabled**. O AS não aciona automaticamente nenhuma ação de dimensionamento para um grupo de AS **Disabled**. Quando um grupo de AS tem uma ação de dimensionamento em andamento, a ação de dimensionamento não pára imediatamente após o grupo de AS ser desativado.

Você pode desativar um grupo de AS quando seu status for **Enabled** ou **Abnormal**.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Na lista do grupo de AS, localize a linha que contém o grupo de AS de destino e clique em **Disable** na coluna **Operation**. Você também pode clicar no nome do grupo de AS e, em seguida, em **Disable** no canto superior direito da página para desativar o grupo de AS.
4. Na caixa de diálogo **Disable AS Group**, clique em **Yes**.

1.6 Modificação de um grupo de AS

Cenários

Você pode modificar um grupo de AS, se necessário. Os valores dos seguintes parâmetros podem ser alterados: **Name**, **Max. Instances**, **Min. Instances**, **Expected Instances**, **Health Check Method**, **Health Check Interval**, **Instance Removal Policy**.

NOTA

Alterar o valor de **Expected Instances** acionará uma ação de dimensionamento. O AS aumentará ou diminuirá automaticamente o número de instâncias para o valor de **Expected Instances**.

Se o grupo de AS não estiver habilitado, não contiver nenhuma instância e não tiver nenhuma ação de dimensionamento em andamento, você poderá modificar as configurações de **Subnet**. Se um grupo de AS não tiver nenhuma ação de dimensionamento em andamento, você poderá modificar suas configurações de **AZ** e **Load Balancing**.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Na lista de grupo de AS, localize o grupo de AS que deseja modificar e escolha **More > Modify** na coluna **Operation**.
Você também pode clicar no nome do grupo de AS para alternar para a página **Overview** e clicar em **Modify** no canto superior direito.
4. Na caixa de diálogo **Modify AS Group**, modifique os dados relacionados, por exemplo, o número esperado de instâncias.
5. Clique em **OK**.

1.7 Exclusão de um grupo de AS

Cenários

Você pode excluir um grupo de AS quando ele não for mais necessário.

- Se um grupo de AS não for necessário durante um período de tempo especificado, é aconselhável desativá-lo, mas não excluí-lo.
- Para um grupo de AS que tenha instâncias ou ações de dimensionamento em andamento, se você tentar excluir forçadamente o grupo de AS e remover e excluir as instâncias no grupo de AS, o grupo de AS entrará no estado de exclusão, rejeitará novas solicitações de dimensionamento, o aguarda até que a ação de dimensionamento em andamento seja concluída e remove todas as instâncias do grupo de AS. Em seguida, o grupo de AS é automaticamente excluído. As instâncias criadas automaticamente são removidas e excluídas, mas as instâncias adicionadas manualmente só são removidas do grupo de AS. Durante este processo, outras operações não podem ser executadas no grupo de AS.
- Quando um grupo de AS é excluído, suas políticas de AS e as regras de alarme geradas com base nessas políticas de AS serão excluídas automaticamente.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Na lista do grupo de AS, localize a linha que contém o grupo de AS de destino e escolha **More > Delete** na coluna **Operation**.

4. Na caixa de diálogo **Delete AS Group** exibida, clique em **Yes**.

2 Configuração de AS

2.1 Criação de uma configuração do AS

Uma configuração do AS define as especificações das instâncias do ECS a serem adicionadas a um grupo do AS. As especificações incluem a imagem do ECS e o tamanho do disco do sistema.

Cenários

- Ao criar um grupo do AS, crie uma nova configuração do AS ou use uma configuração do AS existente.
- Crie a configuração do AS necessária na página **Instance Scaling**.
- Altere a configuração do AS na página de detalhes do grupo do AS.

Métodos

- Crie uma configuração do AS a partir de uma instância do ECS existente.
Se você criar uma configuração do AS a partir de uma instância existente do ECS, a vCPU, a memória, a imagem, o disco e o tipo de ECS serão os mesmos da instância selecionada por padrão. Para mais detalhes, consulte [Criação de uma configuração do AS a partir de uma instância do ECS existente](#).
- Crie uma configuração do AS a partir de um novo modelo de especificações.
Se você tiver requisitos especiais nos ECSs para expansão de recursos, use um novo modelo de especificações para criar a configuração do AS. Para mais detalhes, consulte [Criação de uma configuração do AS a partir de um novo modelo de especificações](#).

2.2 Criação de uma configuração do AS a partir de uma instância do ECS existente

Cenários

Você pode usar uma instância do ECS existente para criar rapidamente uma configuração do AS. Nesse caso, as configurações de parâmetro, como vCPUs, memória, imagem, disco e tipo do ECS na configuração do AS, são as mesmas da instância selecionada por padrão.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
3. Clique em **Create AS Configuration**.
4. Defina os parâmetros para a configuração do AS. [Tabela 2-1](#) lista os parâmetros de configuração do AS.

Tabela 2-1 Parâmetros de configuração do AS

Parâmetro	Descrição	Exemplo de valor
Region	Uma região é onde uma configuração do AS reside.	N/A
Name	Especifica o nome de uma configuração do AS.	N/A
Configuration Template	Selecione Use existing ECS > Select ECS . O tipo de ECS, as vCPUs, a memória, a imagem e as informações de disco na configuração do AS são as mesmas do ECS selecionado por padrão.	Use existing ECS
EIP	Um EIP é um endereço IP público estático vinculado a um ECS em uma VPC. Usando o EIP, o ECS fornece serviços externamente. As seguintes opções são fornecidas: <ul style="list-style-type: none"> ● Do not use Um ECS sem um EIP não pode acessar a Internet. No entanto, ele ainda pode ser usado como um ECS de serviço ou implementado em um cluster em uma rede privada. ● Automatically assign Um EIP com uma largura de banda dedicada é atribuído automaticamente a cada ECS. O tamanho da largura de banda é configurável. NOTA Se você selecionar Automatically assign , especifique Type , Billed By e Bandwidth .	Automatically assign

Parâmetro	Descrição	Exemplo de valor
Login Mode	<p>Um ECS pode ser autorizado usando um par de chaves ou uma senha.</p> <ul style="list-style-type: none"> ● Par de chaves As chaves são usadas para autenticar os usuários que tentam fazer logon em instâncias do ECS. Se você selecionar esse modo, crie ou importe um par de chaves na página Key Pair. <p>NOTA Se você usar uma chave existente, certifique-se de ter salvo o arquivo de chave localmente. Sem a chave, você não poderá fazer logon na instância.</p> ● Senha A senha inicial do usuário root (para Linux) ou do usuário Administrator (para Windows) é usada para autenticação. Você pode fazer logon em uma instância do ECS usando uma combinação de nome de usuário e senha. 	Admin@123
Advanced Settings	<p>Isso permite que você configure User Data e ECS Group.</p> <p>Você pode selecionar Do not configure ou Configure now.</p>	N/A
User Data	<p>Permite que um ECS injete automaticamente os dados do usuário quando o ECS é iniciado pela primeira vez. Esta configuração é opcional. Se essa função estiver ativada, o ECS injetará automaticamente os dados do usuário durante sua primeira inicialização.</p> <p>Para obter detalhes, consulte Guia de usuário do Elastic Cloud Server.</p> <p>Os dois métodos a seguir estão disponíveis:</p> <ul style="list-style-type: none"> ● Como texto: permite inserir os dados do usuário na caixa de texto abaixo. ● Como arquivo: permite injetar um script ou outros arquivos ao criar uma instância do ECS. <p>NOTA</p> <ul style="list-style-type: none"> ● Para Linux, se você usar autenticação de senha, esta função não é suportada. ● Se a imagem selecionada não suportar a injeção de dados do usuário, essa função não será suportada. 	-
ECS Group	<p>Um grupo do ECS permite criar os ECSs em diferentes hosts para melhorar a confiabilidade do serviço.</p> <p>Para obter detalhes, consulte Gerenciamento de grupos do ECS.</p>	N/A

5. Clique em **Create Now**.
6. Se quiser usar a configuração do AS recém-criada, adicione-a ao grupo do AS. Para mais detalhes, consulte [Alteração da configuração de AS para um grupo de AS](#).

2.3 Criação de uma configuração do AS a partir de um novo modelo de especificações

Cenários

Se você tiver requisitos especiais nas instâncias do ECS para expansão de recursos, use um novo modelo de especificações para criar a configuração do AS. Nesse caso, as instâncias do ECS que têm as especificações especificadas no modelo serão adicionadas ao grupo de AS nas ações de dimensionamento.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
3. Clique em **Create AS Configuration**.
4. Defina os parâmetros para a configuração do AS. [Tabela 2-2](#) lista os parâmetros de configuração do AS.

Tabela 2-2 Parâmetros de configuração do AS

Parâmetro	Descrição	Exemplo de valor
Region	Uma região é onde uma configuração do AS reside.	N/A
Name	Especifica o nome da configuração do AS a ser criada.	N/A
Configuration Template	Selecione Create new template . Se essa opção estiver selecionada, configure parâmetros, como vCPUs, memória, imagem, disco e tipo de ECS, para criar uma nova configuração do AS.	Create new template

Parâmetro	Descrição	Exemplo de valor
CPU Architecture	<p>Ambas as arquiteturas de CPU x86 e Kunpeng estão disponíveis:</p> <ul style="list-style-type: none"> ● x86: a arquitetura de CPU baseada em x86 usa CISC (Computação de conjunto de instruções complexas). ● Kunpeng: a arquitetura de CPU baseada em Kunpeng usa RISC (Computação de conjunto de instruções reduzido). <p>NOTA Esse parâmetro é exibido somente quando os ECSs baseados em x86 e baseados em Kunpeng estão disponíveis na região atual.</p>	x86
Specifications	<p>A nuvem pública fornece vários tipos de ECS para diferentes cenários de aplicações.</p> <p>Para obter mais informações, consulte Guia de usuário do Elastic Cloud Server.</p> <p>Configure as especificações do ECS, incluindo vCPUs, memória, tipo de imagem e disco, de acordo com o tipo de ECS.</p>	Memory-optimized ECS
Image	<ul style="list-style-type: none"> ● Imagem pública Uma imagem pública é uma imagem padrão e amplamente usada. Elas contêm um SO e aplicações públicas pré-instaladas e estão disponíveis para todos os usuários. Você pode configurar as aplicações ou software na imagem pública conforme necessário. ● Imagem privada Uma imagem privada é uma imagem disponível apenas para o usuário que a criou. Ela contém um SO, aplicações públicas pré-instaladas e aplicações privadas do usuário. Usar uma imagem privada para criar ECSs libera a configuração repetida de vários ECSs. ● Imagem compartilhada Uma imagem compartilhada é uma imagem privada compartilhada por outro usuário da nuvem pública. 	Public image

Parâmetro	Descrição	Exemplo de valor
Disk	<p>Inclui discos de sistema e de dados.</p> <ul style="list-style-type: none"> ● Disco do sistema I/O comum: usa unidades SATA (Serial Advanced Technology Attachment) para armazenar dados. I/O alta: usa unidades SAS (serial attached SCSI) para armazenar dados. SSD de uso geral: usa unidades de disco de estado sólido (SSD) para armazenar dados. I/O ultra-alta: usa unidades de disco de estado sólido (SSD) para armazenar dados. <p>Se uma imagem do ECS completo for usada, o disco do sistema será restaurado usando o backup em disco. No console, você só pode alterar o tipo e o tamanho do volume. Além disso, o volume não pode ser menor que o backup em disco.</p> <ul style="list-style-type: none"> ● Disco de dados Você pode criar vários discos de dados para uma instância do ECS. Além disso, você pode especificar uma imagem de disco de dados para exportar dados. Se a imagem selecionada for uma imagem do ECS completo, você poderá alterar o tipo de volume e o tamanho e os atributos de criptografia do disco de dados restaurado usando o backup em disco. Certifique-se de que o disco seja pelo menos tão grande quanto o backup em disco. Os atributos de criptografia só podem ser modificados se o backup em disco estiver na mesma região que o disco. 	<p>I/O comum para Disco do sistema</p>
Security Group	<p>Controla o acesso do ECS dentro ou entre grupos de segurança definindo regras de acesso. Os ECSs adicionados a um grupo de segurança são protegidos pelas regras de acesso definidas por você.</p>	<p>N/A</p>

Parâmetro	Descrição	Exemplo de valor
EIP	<p>Um EIP é um endereço IP público estático vinculado a um ECS em uma VPC. Usando o EIP, o ECS fornece serviços externamente.</p> <p>As seguintes opções são fornecidas:</p> <ul style="list-style-type: none"> ● Do not use: um ECS sem um EIP não pode acessar a Internet. No entanto, ele ainda pode ser usado como um ECS de serviço ou implementado em um cluster em uma rede privada. ● Automatically assign: um EIP com uma largura de banda dedicada é atribuído automaticamente a cada ECS. Você pode definir o tamanho da largura de banda. Se você selecionar Automatically assign, precisará especificar EIP Type, Billed By e Bandwidth. <p>NOTA Se você selecionar Automatically assign, precisará especificar EIP Type, Billed By e Bandwidth.</p>	Automatically assign
Bandwidth	<p>Especifica o tamanho da largura de banda em Mbit/s.</p> <p>NOTA</p> <ul style="list-style-type: none"> ● Esse parâmetro está disponível somente quando EIP está definido como Automatically assign. ● Se você selecionar Dedicated, poderá selecionar Bandwidth ou Traffic para Billed By. ● A largura de banda compartilhada pode ser faturada apenas pela largura de banda. Você pode selecionar uma largura de banda compartilhada à qual o EIP deve ser adicionado. 	100

Parâmetro	Descrição	Exemplo de valor
Login Mode	<p>Um ECS pode ser autorizado usando um par de chaves ou uma senha.</p> <ul style="list-style-type: none"> ● Par de chaves As chaves são usadas para autenticar os usuários que tentam fazer logon em instâncias do ECS. Se você selecionar esse modo, crie ou importe um par de chaves na página Key Pair. <p>NOTA Se você usar uma chave existente, certifique-se de ter salvo o arquivo de chave localmente.</p> ● Senha A senha inicial do usuário root (para Linux) ou do usuário Administrator (para Windows) é usada para autenticação. Você pode fazer logon em uma instância do ECS usando uma combinação de nome de usuário e senha. 	Admin@123
Advanced Settings	<p>Este parâmetro permite que você configure ECS Group e User Data.</p> <p>Você pode selecionar Do not configure ou Configure now.</p>	N/A
User Data	<p>Permite que um ECS injete automaticamente os dados do usuário quando o ECS é iniciado pela primeira vez. Esta configuração é opcional. Depois que esta função é habilitada, o ECS injeta automaticamente os dados do usuário durante sua primeira inicialização.</p> <p>Para obter detalhes, consulte Guia de usuário do Elastic Cloud Server.</p> <p>Os seguintes métodos estão disponíveis:</p> <ul style="list-style-type: none"> ● Como texto: permite inserir os dados do usuário na caixa de texto abaixo. ● Como arquivo: permite injetar arquivos de script ou outros arquivos ao criar um ECS. <p>NOTA</p> <ul style="list-style-type: none"> ● Para Linux, se você usar autenticação de senha, esta função não é suportada. ● Se a imagem selecionada não suportar a injeção de dados do usuário, essa função não será suportada. 	-
ECS Group	<p>Um grupo do ECS permite criar os ECSs em diferentes hosts para melhorar a confiabilidade do serviço.</p>	N/A

5. Clique em **Create Now**. O sistema exibe uma mensagem indicando que a configuração do AS foi criada com sucesso.
6. Se quiser usar a configuração do AS recém-criada, adicione-a ao grupo do AS. Para obter detalhes, consulte [Criação de um grupo de AS](#) ou [Alteração da configuração de AS para um grupo de AS](#).

2.4 Cópia de uma configuração do AS

Cenários

Você pode copiar uma configuração do AS existente.

Ao copiar uma configuração de AS, você pode modificar as configurações de parâmetros, como o nome da configuração, as especificações do ECS e a imagem da configuração de AS existente para adicionar rapidamente uma nova configuração de AS.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
3. Clique na guia **AS Configurations**, localize a linha que contém a configuração de AS de destino e clique em **Copy** na coluna **Operation**.
4. Na página **Copy AS Configuration**, modifique as definições de parâmetros, como **Name**, **Specifications** e **Image**, e configure o modo de logon do ECS com base nos requisitos de serviço.
5. Clique em **OK**.

2.5 Exclusão de uma configuração do AS

Cenários

Quando você não precisar mais de uma configuração do AS, poderá excluí-la contanto que a configuração do AS não seja usada por um grupo do AS. Você pode excluir uma única configuração do AS ou excluí-los em lotes.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
3. Clique na página de guia **AS Configurations**, localize a linha que contém a configuração de AS de destino e clique em **Delete** na coluna **Operation** para deletar esta configuração de AS. Você também pode selecionar várias configurações de AS a serem excluídas e clicar em **Delete** na parte superior da lista de configurações de AS para excluí-las todas de uma vez.

3 Política de AS

3.1 Visão geral

As políticas de AS podem acionar ações de dimensionamento para ajustar a largura de banda ou o número de instâncias em um grupo de AS. Uma política de AS define as condições para acionar uma ação de dimensionamento e a operação que será executada. Quando a condição de disparo é atendida, uma ação de dimensionamento é acionada automaticamente.

NOTA

Se forem aplicadas várias políticas de AS a um grupo de AS, uma ação de dimensionamento é desencadeada desde que qualquer uma das políticas de AS seja invocada, desde que as políticas de AS não entrem em conflito entre si.

O AS suporta as seguintes políticas:

- Política de alarme: o AS ajusta automaticamente o número de instâncias em um grupo de AS ou define o número de instâncias para o valor configurado quando um alarme é gerado para uma métrica configurada, como uso da CPU.
- Política programada: o AS aumenta ou diminui automaticamente o número de instâncias em um grupo de AS ou define o número de instâncias para o valor configurado em um momento especificado.
- Política periódica: o AS aumenta ou diminui automaticamente o número de instâncias em um grupo de AS ou define o número de instâncias para o valor configurado em um intervalo configurado, como diário, semanal e mensal.

Modos de ajuste de recursos

- Dinâmico
O AS ajusta o número de instâncias ou largura de banda quando uma política de alarme é acionada.
Esse modo é adequado para cenários em que as cargas de trabalho são imprevisíveis. As políticas de alarme são usadas para acionar ações de dimensionamento com base em dados de monitoramento em tempo real (como o uso da CPU) para ajustar dinamicamente o número de instâncias no grupo de AS.
- Planejado

O AS ajusta o número de instâncias ou a largura de banda quando uma política periódica ou agendada é acionada.

Esse modo é adequado para cenários em que as cargas de trabalho são periódicas.

- Manual

O AS permite ajustar recursos adicionando manualmente instâncias a um grupo de AS, removendo instâncias de um grupo de AS ou alterando o número esperado de instâncias.

3.2 Criação de uma política de AS

Cenários

Você pode gerenciar instâncias em um grupo de AS por meio de políticas de AS. Esta seção descreve como criar uma política de AS.

Criar uma política de alarme

1. Acesse o console de gerenciamento.
1. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
2. Localize a linha que contém o grupo de AS de destino e clique em **View AS Policy** na coluna **Operation**.
3. Na página **AS Policies**, clique em **Add AS Policy**.
4. Defina os parâmetros listados em **Tabela 3-1**.

Tabela 3-1 Parâmetros de política de AS

Parâmetro	Descrição	Exemplo de valor
Policy Name	Especifica o nome da política de AS a ser criada.	as-policy-p6g5
Policy Type	Selecione Alarm .	Alarm

Parâmetro	Descrição	Exemplo de valor
Alarm Rule	<p>Especifica se uma nova regra de alarme deve ser criada (Create) ou se uma regra de alarme existente será usada (Use existing).</p> <p>Para obter detalhes sobre como usar uma regra de alarme existente, consulte Configuração de regras de alarme de monitoramento.</p> <p>Se você optar por criar um alarme, o monitoramento do sistema e o monitoramento personalizado são suportados.</p> <ul style="list-style-type: none"> ● O monitoramento do sistema requer os parâmetros em Tabela 3-2. ● O monitoramento personalizado requer os parâmetros em Tabela 3-3. 	N/A

Parâmetro	Descrição	Exemplo de valor
Scaling Action	<p>Especifica uma ação e o número ou porcentagem de instâncias.</p> <p>As seguintes opções de ação de dimensionamento estão disponíveis:</p> <ul style="list-style-type: none"> ● Adicionar Adiciona instâncias a um grupo de AS quando a ação de dimensionamento é executada. ● Reduzir Remove instâncias de um grupo de AS quando a ação de dimensionamento é executada. ● Configurar Define o número esperado de instâncias em um grupo de AS para um valor especificado. 	<ul style="list-style-type: none"> ● Adicionar 1 instância ● Adicionar 10% de instâncias <p>O número de instâncias a serem adicionadas é 10% do número atual de instâncias no grupo de AS. Se o produto do número atual de instâncias e a porcentagem não for um número inteiro, o AS arredondará automaticamente o valor para cima ou para baixo:</p> <ul style="list-style-type: none"> - Se o valor for maior que 1, é arredondado para baixo. Por exemplo, o valor 12,7 é arredondado para 12. - Se o valor for maior que 0 mas menor que 1, é arredondado para 1. Por exemplo, o valor 0,67 é arredondado para 1. <p>Por exemplo, há 10 instâncias em um grupo de AS e a ação de dimensionamento é Add 15% instances. Quando a política de AS é acionada, o AS calcula que o número de instâncias a serem adicionadas é 1,5 e arredonda 1,5 para 1. Após a conclusão da ação de dimensionamento, há 11 instâncias no grupo de AS.</p>

Parâmetro	Descrição	Exemplo de valor
Cooldown Period	<p>Para evitar que uma política de alarmes seja invocada repetidamente pelo mesmo evento, você pode definir um período de resfriamento.</p> <p>Um período de resfriamento especifica por quanto tempo qualquer ação de dimensionamento acionada por alarme será desativada após a conclusão de uma ação de dimensionamento anterior.</p> <p>O período de resfriamento não se aplica a ações de dimensionamento agendadas ou periódicas. No entanto, o AS relata a duração do resfriamento após a conclusão de uma ação de dimensionamento programada ou periódica.</p> <p>Por exemplo, suponha que o período de resfriamento esteja definido para 300 segundos (5 minutos) e uma política agendada seja especificada para acionar uma ação de dimensionamento às 10:32, e uma ação de dimensionamento anterior acionada por uma política de alarmes termine às 10:30. Qualquer ação de dimensionamento desencadeada por alarmes será então negada durante o período de recarga das 10:30 às 10:35, mas a ação de dimensionamento programada para as 10:32 ainda ocorrerá. Se a ação de dimensionamento agendada terminar às 10:36, um novo período de resfriamento começa às 10:36 e termina às 10:41.</p>	300

Parâmetro	Descrição	Exemplo de valor
	<p>NOTA</p> <ul style="list-style-type: none"> ● Se uma ação de dimensionamento for acionada por uma política de AS, o período de recarga será o que estiver configurado para essa política de AS. ● Se uma ação de dimensionamento for acionada alterando manualmente o número esperado de instâncias ou por outras ações, o período de resfriamento será o que estiver configurado para o grupo de AS. O período de resfriamento padrão é de 300 segundos. 	

Tabela 3-2 Parâmetros de monitoramento do sistema

Parâmetro	Descrição	Exemplo de valor
Rule Name	Especifica o nome da regra de alarme.	as-alarm-7o1u
Monitoring Type	Especifica o tipo de métricas de monitoramento, que pode ser System monitoring ou Custom monitoring . Selecione System monitoring .	System monitoring
Trigger Condition	Selecione as métricas de monitoramento suportadas pelo AS e defina as condições de alarme para as métricas.	CPU Usage Max. >70%
Monitoring Interval	Especifica o intervalo no qual o status do alarme é atualizado com base na regra de alarme.	5 minutos
Consecutive Occurrences	Especifica o número de pontos de amostragem quando um alarme é acionado. Se o valor de Consecutive Occurrences for definido como n , os pontos de amostragem da regra de alarme são os pontos de amostragem em n períodos de amostragem consecutivos. O status da regra de alarme não muda para Alarm , a menos que todos os pontos de amostragem violem o limite configurado pela regra de alarme.	3

Tabela 3-3 Parâmetros de monitoramento personalizados

Parâmetro	Descrição	Exemplo de valor
Rule Name	Especifica o nome da regra de alarme.	as-alarm-7o1u

Parâmetro	Descrição	Exemplo de valor
Monitoring Type	Selecione Custom monitoring . O monitoramento personalizado atende aos requisitos de monitoramento em vários cenários.	Custom monitoring
Resource Type	Especifica o nome do serviço para o qual a regra de alarme está configurada.	AGT.ECS
Dimension	Especifica a dimensão métrica da regra de alarme.	instance_id
Monitored Object	Especifica os recursos aos quais a regra de alarme se aplica.	N/A
Trigger Condition	Selecione as métricas de monitoramento suportadas pelo AS e defina as condições de alarme para as métricas.	CPU Usage Max. >70%
Monitoring Interval	Especifica o intervalo no qual o status do alarme é atualizado com base na regra de alarme.	5 minutos
Consecutive Occurrences	Especifica o número de pontos de amostragem quando um alarme é acionado. Se o valor de Consecutive Occurrences for definido como n , os pontos de amostragem da regra de alarme são os pontos de amostragem em n períodos de amostragem consecutivos. O status da regra de alarme não muda para Alarm , a menos que todos os pontos de amostragem violem o limite configurado pela regra de alarme.	3

5. Clique em **OK**.

A política de AS recém-adicionada é exibida na guia **AS Policy**. Além disso, a política de AS está no estado **Enabled** por padrão.

Criar uma política programada ou periódica

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
3. Localize a linha que contém o grupo de AS de destino e clique em **View AS Policy** na coluna **Operation**.
4. Na página **AS Policies**, clique em **Add AS Policy**.
5. Configure os parâmetros listados em [Tabela 3-4](#).

Tabela 3-4 Descrição do parâmetro

Parâmetro	Descrição	Exemplo de valor
Policy Name	Especifica o nome da política de AS a ser criada.	as-policy-p6g5
Policy Type	<p>Selecione Scheduled ou Periodic para expandir recursos em um horário especificado.</p> <p>Se você selecionar Periodic, será necessário configurar mais dois parâmetros:</p> <ul style="list-style-type: none"> ● Período <ul style="list-style-type: none"> - Dia - Semana - Mês ● Intervalo de tempo <p>Especifica um intervalo de tempo durante o qual a política de AS pode ser acionada.</p> 	N/A
Time Zone	<p>O valor padrão é GMT+08:00.</p> <p>GMT+08:00 é 8:00 horas à frente do horário de Greenwich.</p>	GMT+08:00
Triggered At	Especifica um horário em que a política de AS é acionada.	N/A

Parâmetro	Descrição	Exemplo de valor
Scaling Action	<p>Especifica uma ação e o número de instâncias.</p> <p>As seguintes opções de ação de dimensionamento estão disponíveis:</p> <ul style="list-style-type: none"> ● Adicionar Adiciona instâncias a um grupo de AS quando a ação de dimensionamento é executada. ● Reduzir Remove instâncias de um grupo de AS quando a ação de dimensionamento é executada. ● Configurar Define o número esperado de instâncias em um grupo de AS para um valor especificado. 	<ul style="list-style-type: none"> ● Adicionar 1 instância ● Adicionar 10% de instâncias <p>O número de instâncias a serem adicionadas é 10% do número atual de instâncias no grupo de AS. Se o produto do número atual de instâncias e da porcentagem não for um inteiro, o AS automaticamente arredondará o valor para cima ou para baixo:</p> <ul style="list-style-type: none"> - Se o valor for maior que 1, é arredondado para baixo. Por exemplo, o valor 12,7 é arredondado para 12. - Se o valor for maior que 0 mas menor que 1, é arredondado para 1. Por exemplo, o valor 0,67 é arredondado para 1. <p>Por exemplo, há 10 instâncias em um grupo de AS e a ação de dimensionamento é Add 15% instances. Quando a política de AS é invocada, o AS calcula que o número de instâncias a serem adicionadas é 1,5 e arredonda 1,5 para 1. Após a conclusão da ação de dimensionamento, há 11 instâncias no grupo de AS.</p>

Parâmetro	Descrição	Exemplo de valor
Cooldown Period	<p>Para evitar que uma política de alarmes seja invocada repetidamente pelo mesmo evento, você pode definir um período de resfriamento.</p> <p>Especifica por quanto tempo qualquer ação de dimensionamento acionada por alarme será desabilitada após uma ação de dimensionamento anterior está completa.</p> <p>O período de resfriamento não se aplica a ações de dimensionamento agendadas ou periódicas. No entanto, o AS relata a duração do resfriamento após a conclusão de uma ação de dimensionamento programada ou periódica.</p> <p>Por exemplo, suponha que o período de resfriamento esteja definido para 300 segundos (5 minutos) e uma política agendada seja especificada para acionar uma ação de dimensionamento às 10:32, e uma ação de dimensionamento anterior acionada por uma política de alarmes termine às 10:30. Qualquer ação de dimensionamento desencadeada por alarmes será então negada durante o período de recarga das 10:30 às 10:35, mas a ação de dimensionamento programada para as 10:32 ainda ocorrerá. Se a ação de dimensionamento agendada terminar às 10:36, um novo período de resfriamento começa às 10:36 e termina às 10:41.</p> <p>NOTA</p> <ul style="list-style-type: none"> ● Se uma ação de dimensionamento for acionada por uma política de AS, o período de recarga será o que estiver configurado para essa política de AS. ● Se uma ação de dimensionamento for acionada alterando manualmente o número esperado de instâncias ou por outras ações, o período de resfriamento será o que estiver configurado para o grupo de AS. O período de resfriamento padrão é de 300 segundos. 	300

6. Clique em **OK**.

A política de AS recém-adicionada é exibida na guia **AS Policy**. Além disso, a política de AS está no estado **Enabled** por padrão.

NOTA

Se você tiver criado políticas de AS agendadas ou periódicas que são chamadas ao mesmo tempo, o AS executará a que foi criada posteriormente. Esta restrição não se aplica a políticas de AS acionadas por alarme.

3.3 Gerenciamento de políticas de AS

Cenários

Uma política de AS especifica as condições para disparar uma ação de dimensionamento, bem como a operação que será executada. Se as condições forem atendidas, uma ação de dimensionamento será acionada automaticamente.

Esta seção descreve como gerenciar uma política de AS, incluindo modificar, ativar, desativar, executar e excluir uma política de AS.

Modificar uma política de AS

Se uma política de AS específica não puder atender aos requisitos de serviço, você poderá modificar as configurações de parâmetro da política.

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Localize a linha que contém o grupo de AS de destino e clique em **View AS Policy** na coluna **Operation**. Na página exibida, localize a linha que contém a política de AS de destino e escolha **More > Modify** na coluna **Operation**.
4. Na caixa de diálogo **Modify AS Policy** exibida, modifique os parâmetros e clique em **OK**.

Habilitar uma política de AS

Uma política de AS pode acionar ações de dimensionamento somente quando ela e o grupo de AS estiverem habilitados. Você pode habilitar uma ou mais políticas de AS para um grupo de AS, conforme necessário.

- Antes de ativar várias políticas de AS, certifique-se de que as políticas de AS não entrem em conflito umas com as outras.
- Uma política de AS pode ser ativada somente quando seu status é **Disabled**.

Localize a linha que contém o grupo de AS de destino e clique em **View AS Policy** na coluna **Operation**. Na página exibida, localize a linha que contém a política de AS de destino e clique em **Enable** na coluna **Operation**. Para habilitar simultaneamente várias políticas de AS, selecione essas políticas de AS e clique em **Enable** na parte superior da lista de políticas de AS.

Desativar uma política de AS

Se você não quiser que uma determinada política de AS acione ações de dimensionamento dentro de um período de tempo especificado, você pode desativá-la.

- Se todas as políticas de AS configuradas para um grupo de AS estiverem desativadas, nenhuma ação de dimensionamento será acionada para esse grupo de AS. No entanto, se você alterar manualmente o valor de **Expected Instances**, uma ação de dimensionamento ainda será acionada.
- Você pode desabilitar uma política de AS somente quando seu status estiver **Enabled**.

Localize a linha que contém o grupo de AS de destino e clique em **View AS Policy** na coluna **Operation**. Na página exibida, localize a linha que contém a política de AS de destino e clique em **Disable** na coluna **Operation**. Para desativar simultaneamente várias políticas de AS, selecione essas políticas de AS e clique em **Disable** na parte superior da lista de políticas de AS.

Executar manualmente uma política de AS

Você pode fazer com que o número de instâncias em um grupo de AS atinja o número esperado de instâncias imediatamente executando manualmente uma política de AS.

- Você pode executar manualmente uma política de AS se as condições de dimensionamento configuradas na política de AS não forem atendidas.
- Você pode executar manualmente uma política de AS somente quando o grupo de AS e a política de AS estiverem no estado **Enabled**.

Localize a linha que contém o grupo de AS de destino e clique em **View AS Policy** na coluna **Operation**. Na página exibida, localize a linha que contém a política de AS de destino e clique em **Execute Now** na coluna **Operation**.

NOTA

- Se **Policy Type** for definido como **Alarm** e **Alarm Policy Type** como **Refined scaling**, a política de dimensionamento não poderá ser executada imediatamente.

Excluir uma política de AS

Você pode excluir uma política de AS que não será usada para acionar ações de dimensionamento.

Uma política de AS pode ser excluída mesmo quando a ação de dimensionamento desencadeada pela política está em andamento. A exclusão da política de AS não afeta uma ação de dimensionamento já iniciada.

Localize a linha que contém o grupo de AS de destino e clique em **View AS Policy** na coluna **Operation**. Na página exibida, localize a linha que contém a política de AS de destino e escolha **More > Delete** na coluna **Operation**.

Para excluir simultaneamente várias políticas de AS, selecione essas políticas de AS e clique em **Delete** na parte superior da lista de políticas de AS.

4 Ação de dimensionamento

4.1 Dimensionamento dinâmico

Antes de usar AS para executar ações de dimensionamento, você deve especificar como executar as ações de dimensionamento para expandir dinamicamente os recursos.

Se as demandas mudarem com frequência, você poderá configurar políticas baseadas em alarme para dimensionar recursos. Quando as condições para chamar uma política de AS são atendidas, o AS altera automaticamente o número esperado de instâncias para acionar uma ação de dimensionamento para aumentar ou diminuir os recursos. Para obter detalhes sobre como criar uma política de alarmes, consulte [Criação de uma política de AS](#).

Considere uma aplicação de reserva de bilhetes de trem. Se o uso da CPU das instâncias que executam a aplicação for de até 90%, uma instância precisará ser adicionada para garantir que os serviços sejam executados corretamente. Se o uso da CPU cair para 30%, uma instância precisará ser excluída para evitar o desperdício de recursos. Para atender aos requisitos, você pode configurar duas políticas de alarme. Uma política é usada para adicionar uma instância se o uso máximo da CPU exceder 90%. Para mais detalhes, consulte [Figura 4-1](#). A outra política é usada para remover uma instância se o uso mínimo da CPU cair abaixo de 30%. Para obter detalhes, consulte [Figura 4-2](#).

Figura 4-1 Política de alarme 01

Add AS Policy

Policy Name

Policy Type Alarm Scheduled Periodic

Alarm Rule Create Use existing

Rule Name

Monitoring Type System monitoring Custom monitoring

Trigger Condition %

The metrics that can be monitored vary somewhat by OS. [Learn more](#)
To select a metric whose name starts with (Agent), make sure the Agent has been installed on all instances in the AS group. [Learn more](#)

Monitoring Interval

Consecutive Occurrences ?

Scaling Action

Cooldown Period (s) ?

Figura 4-2 Política de alarme 02

Add AS Policy

Policy Name:

Policy Type: Alarm Scheduled Periodic

Alarm Rule: Create Use existing

Rule Name:

Monitoring Type: System monitoring Custom monitoring

Trigger Condition: %

The metrics that can be monitored vary somewhat by OS. [Learn more](#)
To select a metric whose name starts with (Agent), make sure the Agent has been installed on all instances in the AS group.
[Learn more](#)

Monitoring Interval:

Consecutive Occurrences: ⓘ

Scaling Action:

Cooldown Period (s): ⓘ

4.2 Dimensionamento agendado

Para atender às demandas que mudam regularmente, você pode configurar uma política agendada ou periódica para dimensionar os recursos em um horário especificado ou periodicamente. Para obter detalhes sobre como criar uma política agendada ou periódica, consulte [Criação de uma política de AS](#).

Tome uma aplicação web de seleção de curso on-line como exemplo. Esta aplicação é frequentemente usado quando um semestre começa e raramente usado durante outras partes do ano. Você pode configurar duas políticas agendadas para dimensionar recursos no início de cada semestre. A primeira política é usada para adicionar uma instância quando a seleção do curso começa, e a segunda política é usada para remover uma instância quando a seleção do curso termina.

4.3 Dimensionamento manual

Cenários

Você pode adicionar ou remover instâncias manualmente de ou para um grupo de AS, ou alterar o número esperado de instâncias.

Procedimento

Adicionar instâncias a um grupo de AS

Se um grupo de AS estiver ativado e não tiver nenhuma ação de dimensionamento contínua, e o número atual de instâncias for menor que o máximo, você poderá adicionar instâncias manualmente ao grupo de AS.

Antes de adicionar instâncias a um grupo de AS, certifique-se de que as seguintes condições sejam atendidas:

- As instâncias não estão em outros grupos de AS.
- As instâncias estão na mesma VPC que o grupo de AS.
- As instâncias estão nas AZs usadas pelo grupo de AS.
- Depois que as instâncias são adicionadas, o número total de instâncias é menor ou igual ao número máximo de instâncias permitidas.
- Até 10 instâncias podem ser adicionadas por vez.

Para adicionar instâncias a um grupo de AS, execute as seguintes etapas:

1. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
2. Clique na guia **AS Groups** e, em seguida, no nome do grupo de AS de destino.
3. Na página de detalhes do grupo de AS, clique na guia **Instances** e, em seguida, em **Add**.
4. Selecione as instâncias a serem adicionadas e clique em **OK**.

Remover instâncias de um grupo de AS

Você pode remover uma instância de um grupo de AS, atualizar a instância ou corrigir uma falha de instância e adicionar a instância de volta ao grupo de AS. Uma instância removida do grupo de AS não transporta mais tráfego de aplicações.

Você pode modificar a configuração de AS para um grupo de AS a qualquer momento, mas a nova configuração não será aplicada a nenhuma instância em execução. Para aplicar a nova configuração, pare uma instância e o sistema a substituirá por uma nova que tenha as especificações especificadas na nova configuração. Você também pode remover a instância do grupo de AS, atualizar a instância e adicionar a instância de volta ao grupo de AS.

Existem algumas restrições na remoção de instâncias:

- O grupo de AS não pode ter uma ação de dimensionamento em andamento, as instâncias devem estar ativas e o número total de instâncias após a remoção não pode ser menor que o número mínimo de instâncias especificadas.
- As instâncias podem ser removidas de um grupo de AS e excluídas somente se o grupo de AS não tiver nenhuma ação de dimensionamento em andamento e as instâncias forem criadas e ativas automaticamente e não forem usadas pelo SDRS (Serviço de recuperação de desastres de armazenamento).
- Por padrão, as instâncias adicionadas automaticamente a um grupo de AS são faturadas com base em pagamento por uso. Se você escolher **Remove** para essa instância, o sistema a removerá do grupo de AS, mas não a excluirá. Para remover e excluir a instância, você precisa escolher **Remove and Delete**.
- Se você alterar o modo de cobrança de uma instância em um grupo de AS de pagamento por uso para anual/mensal e escolher **Remove and Delete** para a instância, o sistema a removerá do grupo de AS, mas não a excluirá.
- As instâncias adicionadas manualmente a um grupo de AS só podem ser removidas. Eles não podem ser removidos e excluídos.

- Um máximo de 10 instâncias podem ser removidas de cada vez.

Para remover uma instância de um grupo de AS, execute as seguintes etapas:

1. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
2. Clique na guia **AS Groups** e, em seguida, no nome do grupo de AS de destino.
3. Clique na guia **Instances**, localize a linha que contém a instância de destino e clique em **Remove** ou **Remove and Delete** na coluna **Operation**.
4. Para excluir várias instâncias de um grupo de AS, marque as caixas de seleção na frente delas e clique em **Remove** ou **Remove and Delete**.

Para excluir todas as instâncias de um grupo de AS, marque a caixa de seleção à esquerda de **Name** e clique em **Remove** ou **Remove and Delete**.

Alterar o número esperado de instâncias

Altere manualmente o número esperado de instâncias para adicionar ou reduzir o número de instâncias em um grupo de AS para expandir recursos.

Para mais detalhes, consulte [Modificação de um grupo de AS](#).

4.4 Configuração de uma política de remoção de instância

Quando as instâncias são automaticamente removidas do grupo de AS, as instâncias que não estão nas AZs usadas no momento serão removidas primeiro. Além disso, o AS verificará se as instâncias estão distribuídas uniformemente nas AZs usadas atualmente. Se a carga entre as AZs estiver desbalanceada, o AS equilibra a carga entre as AZs ao remover instâncias. Se a carga entre as AZs estiver balanceada, o AS removerá as instâncias seguindo a política de remoção de instâncias pré-configurada.

O AS suporta as seguintes políticas de remoção de instâncias:

- **Instância mais antiga:** a instância mais antiga é removida primeiro do grupo de AS. Use esta política se quiser fazer upgrade de instâncias em um grupo de AS para um novo tipo de ECS. Você pode substituir gradualmente instâncias do tipo anterior por instâncias do novo tipo.
- **Instância mais recente:** a instância mais recente é removida primeiro do grupo de AS. Use esta política se quiser testar uma nova configuração de AS, mas não quiser mantê-la em produção.
- **Instância mais antiga criada a partir da configuração de AS mais antiga:** a instância mais antiga criada a partir da configuração mais antiga é removida primeiro do grupo de AS. Use essa política se quiser atualizar um grupo de AS e eliminar gradualmente as instâncias criadas a partir de uma configuração de AS anterior.
- **Instância mais recente criada a partir da configuração de AS mais antiga:** a instância mais recente criada a partir da configuração mais antiga é removida primeiro do grupo de AS.

NOTA

As instâncias adicionadas manualmente são as últimas a serem removidas e, se o AS remover uma instância adicionada manualmente, ele remove apenas a instância. Ele não exclui a instância. Se várias instâncias adicionadas manualmente precisarem ser removidas, o AS preferencialmente removerá primeiro a instância adicionada mais cedo.

4.5 Visualização de uma ação de dimensionamento

Cenários

Esta seção descreve como verificar se uma ação de dimensionamento foi executada e como exibir detalhes da ação de dimensionamento.

Visualização de ações de dimensionamento

As etapas a seguir ilustram como exibir ações de dimensionamento de um grupo de AS.

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
3. Clique na guia **AS Groups** e, em seguida, no nome do grupo de AS de destino.
4. Clique na guia **Monitoring** e visualize as ações de dimensionamento. Na página **Monitoring**, você pode exibir alterações no número de instâncias e métricas, como uso da CPU.

Exibição de ações de dimensionamento históricas

As etapas a seguir ilustram como exibir os registros históricos de ações de dimensionamento de um grupo de AS.

1. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**.
2. Clique na guia **AS Groups** e, em seguida, no nome do grupo de AS de destino.
3. Clique na guia **Scaling Actions**. Esta página exibe ações de dimensionamento históricas de um grupo de AS, incluindo dimensionamento de instâncias e migração do balanceador de carga.

Scaling Action ID, **Status**, **Scaling Action Type**, **Description**, **Start Time** e **End Time**

das ações de dimensionamento são exibidos. Clique em  antes de ID da ação de dimensionamento para exibir o nome do recurso, o status e a causa da falha. Você também pode usar a função de filtragem no canto superior direito para exibir ações de dimensionamento em um período especificado.

4.6 Gerenciamento de ganchos do ciclo de vida

Os ganchos de ciclo de vida permitem controlar de forma flexível a adição e a remoção de instâncias do ECS em grupos de AS e gerenciar o ciclo de vida de instâncias do ECS em grupos de AS. **Figura 4-3** mostra o ciclo de vida da instância quando nenhum gancho de ciclo de vida é adicionado a um grupo de AS.

Figura 4-3 Ciclo de vida da instância quando nenhum gancho de ciclo de vida é adicionado a um grupo de AS

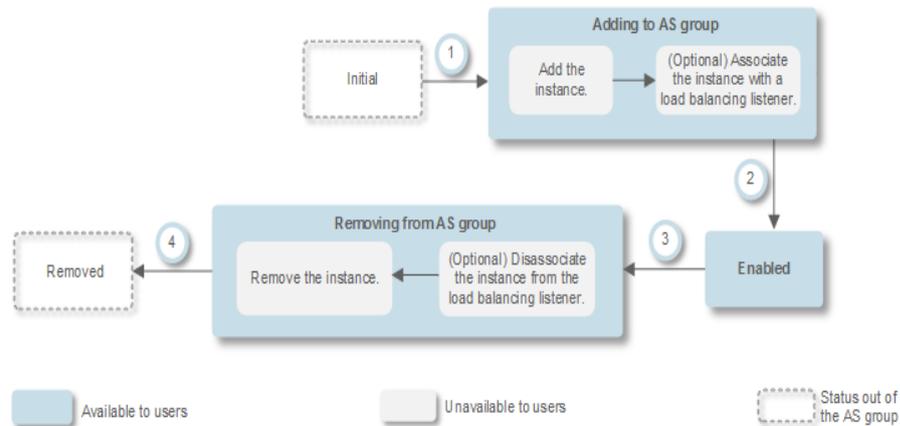
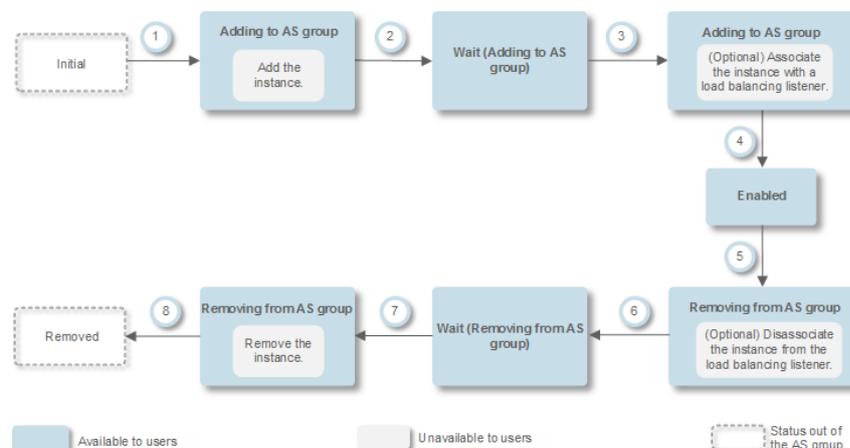


Figura 4-4 mostra o ciclo de vida da instância quando um gancho de ciclo de vida é adicionado a um grupo de AS.

Figura 4-4 Ciclo de vida da instância quando um gancho de ciclo de vida é adicionado a um grupo de AS



Quando o grupo de AS dimensiona dentro ou fora, os ganchos de ciclo de vida adicionados são acionados, a ação de dimensionamento é suspensa e a instância que está sendo adicionada ou removida é colocada em um estado de espera, conforme mostrado em 2 e 6 em **Figura 4-4**. Durante esse período, você pode executar algumas operações personalizadas na instância. Por exemplo, você pode instalar ou configurar o software em uma instância que está sendo adicionada ao grupo de AS. Uma ação de dimensionamento suspensa será retomada se ocorrer uma das seguintes situações:

- A instância está no estado de espera há mais tempo do que o período de tempo limite.
- Uma ação de retorno de chamada é executada para encerrar o estado de espera da instância.

Cenários de aplicação

- As instâncias recém-adicionadas a um grupo de AS precisam ser inicializadas antes de serem vinculadas a um ouvinte do balanceador de carga. Inicialização significa que o software está instalado e configurado e a instância está totalmente pronta para aceitar tráfego.
- Para remover uma instância de um grupo de AS, ela precisa primeiro ser desvinculada do ouvinte do balanceador de carga, interromper a aceitação de novas solicitações e finalizar o processamento de todas as solicitações aceitas.
- Antes que as instâncias sejam removidas de um grupo de AS, talvez seja necessário fazer backup de dados ou fazer o download de logs.
- Outros cenários em que operações personalizadas precisam ser executadas

Como funcionam os ganchos de ciclo de vida

Depois de adicionar ganchos de ciclo de vida a um grupo de AS, eles funcionam da seguinte maneira:

- Adicionar uma instância do ECS a um grupo de AS

Quando uma instância é inicializada e adicionada a um grupo de AS, um gancho de ciclo de vida do tipo de **Instance adding** é acionado automaticamente. A instância entra no estado **Wait (Adding to AS group)**, ou seja, a instância é suspensa pelo gancho do ciclo de vida. Se você tiver configurado um objeto de notificação, o sistema enviará uma mensagem para o objeto. Depois de receber a mensagem, você pode executar operações personalizadas, por exemplo, instalar software na instância. A instância permanece em um estado de espera até que você conclua as operações personalizadas e execute uma ação de retorno de chamada ou a duração do tempo limite termine. Depois que a instância sair de um estado de espera, a ação de retorno de chamada padrão especificada ocorrerá.

- **Continue**: a instância será adicionada ao grupo de AS.
- **Abandon**: a instância será excluída e uma nova instância será criada.

Se você configurou vários ganchos de ciclo de vida de **Instance adding**, todos eles serão acionados quando uma instância for adicionada ao grupo de AS. Se a ação de retorno de chamada padrão de qualquer gancho de ciclo de vida for **Abandon**, a instância será excluída e uma nova instância será criada. Se a ação de retorno de chamada padrão de todos os ganchos do ciclo de vida for **Continue**, a instância será adicionada ao grupo de AS após a conclusão da suspensão pelo último gancho do ciclo de vida.

- Remover uma instância de um grupo de AS

Quando uma instância é removida de um grupo de AS, a instância entra no estado **Removing from AS group**. Depois que um gancho de ciclo de vida é acionado, a instância entra no estado **Wait (Removing from AS group)**. O sistema envia mensagens para o objeto de notificação configurado. Depois de receber a mensagem, você pode realizar operações personalizadas, como desinstalar software e fazer backup de dados. A instância permanece no estado de espera até que você termine as operações personalizadas e execute a operação de retorno de chamada padrão ou a duração do tempo limite termine. Depois que a instância sair de um estado de espera, a ação de retorno de chamada padrão especificada ocorrerá.

- **Continue**: a instância é removida do grupo de AS.
- **Abandon**: a instância é removida do grupo de AS.

Se você tiver configurado vários ganchos de ciclo de vida e a ação de retorno de chamada padrão de todos os ganchos de ciclo de vida for **Continue**, a instância será

removida do grupo de AS até a suspensão pelo tempo limite dos ganchos de ciclo de vida restantes. Se a ação de retorno de chamada padrão de qualquer gancho de ciclo de vida for **Abandon**, a instância será removida diretamente do grupo de AS.

Restrições

- Você pode adicionar, modificar ou excluir um gancho de ciclo de vida quando o grupo de AS não executa uma ação de dimensionamento.
- Até cinco ganchos de ciclo de vida podem ser adicionados a um grupo de AS.

Adicionar um gancho de ciclo de vida

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Clique no nome do grupo de AS ao qual o gancho do ciclo de vida deve ser adicionado. Na página de detalhes do grupo de AS, clique na guia **Lifecycle Hooks** e, em seguida, em **Add Lifecycle Hook**.
4. Na caixa de diálogo **Add Lifecycle Hook** exibida, defina os parâmetros listados em [Tabela 4-1](#).

Tabela 4-1 Descrição do parâmetro

Parâmetro	Descrição	Exemplo de valor
Hook Name	Especifica o nome do gancho do ciclo de vida. O nome pode conter letras, dígitos, sublinhados (_) e hífens (-) e não pode exceder 32 caracteres.	we12_w
Hook Type	Especifica o tipo de gancho de ciclo de vida. O valor pode ser Instance adding ou Instance removal . Instance adding coloca uma instância que está sendo adicionada a um grupo de AS no estado Wait (Adding to AS group) . Instance removal coloca uma instância que está sendo removida de um grupo de AS no estado Wait (Removing from AS group) .	Instance adding

Parâmetro	Descrição	Exemplo de valor
Default Callback Action	<p>Especifica a ação que o sistema executa quando uma instância sai de um estado de espera.</p> <p>A ação de retorno de chamada padrão para uma Instance adding gancho de ciclo de vida pode ser Continue ou Abandon:</p> <ul style="list-style-type: none"> ● Continue: se vários ganchos de ciclo de vida estiverem configurados para o grupo de AS e a ação de retorno de chamada padrão de todos os ganchos for Continue, o sistema continuará a adicionar a instância ao grupo de AS até que todos os ganchos de ciclo de vida expirem. ● Abandon: se vários ganchos de ciclo de vida forem configurados para o grupo de AS, e a ação de retorno de chamada padrão de um gancho de ciclo de vida for Abandon, o sistema excluirá a instância e criará outra sem aguardar o tempo limite dos ganchos de ciclo de vida restantes. <p>A ação de retorno de chamada padrão para um gancho de ciclo de vida de Instance removal pode ser Continue ou Abandon:</p> <ul style="list-style-type: none"> ● Continue: se apenas um gancho de ciclo de vida for configurado para o grupo de AS, o sistema removerá a instância do grupo de AS. Se vários ganchos de ciclo de vida forem configurados para o grupo de AS e as ações de retorno de chamada padrão de todos os ganchos forem Continue, o sistema continuará a remover a instância do grupo de AS até que todos os ganchos de ciclo de vida expirem. ● Abandon: se vários ganchos de ciclo de vida estiverem configurados para o grupo de AS e a ação de retorno de chamada padrão de um gancho de ciclo de vida for Abandon, o sistema continuará a remover a instância do grupo de AS sem esperar pelo tempo limite dos ganchos de ciclo de vida restantes. 	Continue
Timeout Duration (s)	<p>Especifica a quantidade de tempo para que as instâncias permaneçam em um estado de espera. O valor varia de 300s a 86400s. O valor padrão é 3600s.</p> <p>Você pode estender a duração do tempo limite ou executar uma ação Continue ou Abandon antes que a duração do tempo limite termine. Para obter mais informações sobre ações de retorno de chamada, consulte Executar uma ação de retorno de chamada.</p>	3600

Parâmetro	Descrição	Exemplo de valor
Notificação Topic	<p>Especifica um objeto de notificação para um gancho de ciclo de vida. Para obter detalhes, consulte "Criação de um tópico" em Guia de usuário da Simple Message Notification. Quando uma instância é suspensa pelo gancho do ciclo de vida, o sistema envia uma notificação ao objeto. Essa notificação contém as informações básicas da instância, o conteúdo da notificação personalizada e o token para controlar as ações do ciclo de vida. Um exemplo de notificação é o seguinte:</p> <pre> { "service": "AutoScaling", "tenant_id": "93075aa73f6a4fc0a3209490cc57181a", "lifecycle_hook_type": "INSTANCE_LAUNCHING", "lifecycle_hook_name": "test02", "lifecycle_action_key": "4c76c562-9688-45c6- b685-7fd732df310a", "notification_metadata": "xxxxxxxxxxxxx", "scaling_instance": { "instance_id": "89b421e4-5fa6-4733- bf40-6b07a8657256", "instance_name": "as-config-kxeg_RM6OCREY", "instance_ip": "192.168.0.202" }, "scaling_group": { "scaling_group_id": "fe376277-50a6-4e36- bdb0-685da85f1a82", "scaling_group_name": "as-group-wyz01", "scaling_config_id": "16ca8027-b6cc-45fc- af2d-5a79996f685d", "scaling_config_name": "as-config-kxeg" } } </pre>	N/A
Notificação Message	Depois que um objeto de notificação é configurado, o sistema envia sua notificação personalizada para o objeto.	N/A

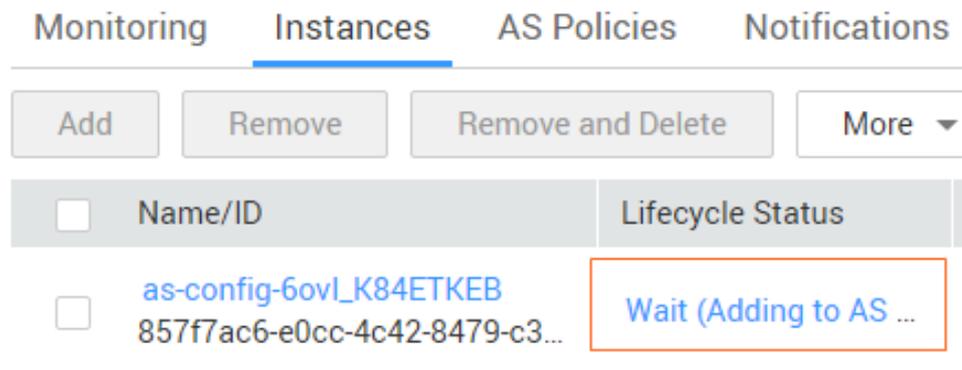
5. Clique em **OK**.

O gancho de ciclo de vida adicionado é exibido na página **Lifecycle Hooks**.

Executar uma ação de retorno de chamada

1. Na página **AS Groups**, clique no nome do grupo de AS de destino.
2. Na página exibida, clique na guia **Instances**.
3. Localize a instância que foi suspensa pelo gancho do ciclo de vida e clique em **Wait (Adding to AS group)** ou **Wait (Removing from AS group)** na coluna **Lifecycle Status**.

Figura 4-5 Executar uma ação de retorno de chamada

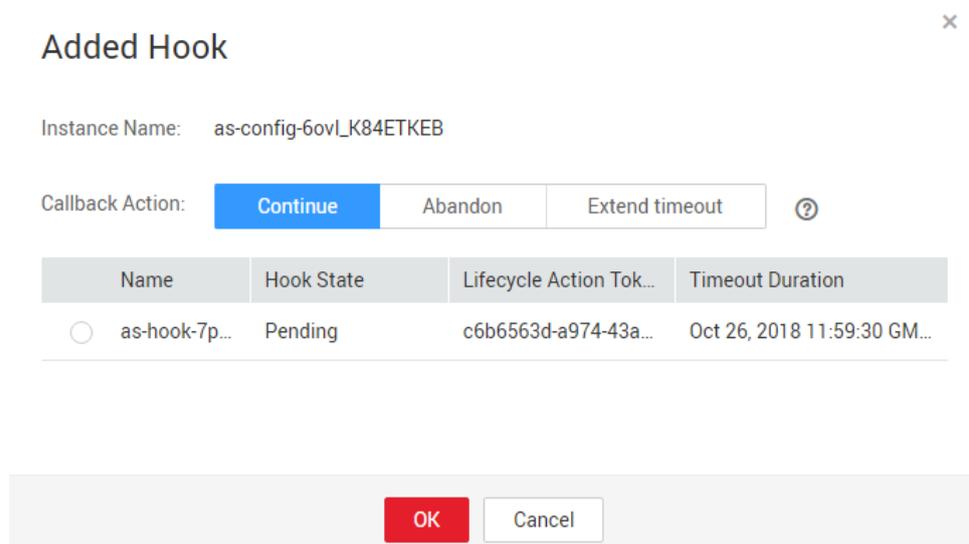


NOTA

As ações de retorno de chamada só podem ser executadas em instâncias que foram suspensas por um gancho de ciclo de vida.

- Na caixa de diálogo **Added Hook** exibida, visualize a instância suspensa e todos os ganchos do ciclo de vida e execute ações de retorno de chamada nos ganchos do ciclo de vida.

Figura 4-6 Adicionar caixa de diálogo Gancho



As ações de retorno de chamada incluem:

- **Continue**
- **Abandon**
- **Extend timeout**

Se você tiver executado operações personalizadas antes do término da duração do tempo limite, selecione **Continue** ou **Abandon** para concluir as ações do ciclo de vida. Para obter detalhes sobre **Continue** e **Abandon**, consulte [Tabela 4-1](#). Se precisar de mais tempo para operações personalizadas, selecione **Extend timeout** para estender a duração

do tempo limite. Em seguida, a duração do tempo limite será estendida em 3600 segundos a cada vez.

Modificar um gancho de ciclo de vida

Na página **Lifecycle Hooks**, localize o gancho de ciclo de vida de destino e clique em **Modify** na coluna **Operation**, consulte **Tabela 4-1** para obter parâmetros. Você pode modificar o parâmetro, exceto **Hook Name**, como **Hook Type**, **Default Callback Action** e **Timeout Duration**.

Exclusão de um gancho de ciclo de vida

Na página **Lifecycle Hooks**, localize o gancho de ciclo de vida de destino e clique em **Delete** na coluna **Operation**.

4.7 Configuração da proteção de instância

Cenários

Para controlar se uma instância pode ser removida automaticamente de um grupo de AS, use proteção de instância. Uma vez configurado, quando o AS dimensiona automaticamente no grupo de AS, a instância protegida não será removida.

Pré-requisitos

A proteção de instâncias não protege as instâncias do seguinte:

- Substituição da verificação de integridade se a instância falhar nas verificações de integridade
- Remoção manual

NOTA

- A proteção de instâncias não protege instâncias não íntegras porque essas instâncias não podem fornecer serviços.
- Por padrão, a proteção de instância não entra em vigor nos ECSs recém-criados ou adicionados a um grupo de AS.
- Se uma instância for removida de um grupo de AS, sua configuração de proteção de instância será perdida.

Ativar a proteção de instância

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Clique no nome do grupo de AS de destino.
4. Clique na guia **Instances**. Selecione uma ou mais instâncias e escolha **Enable Instance Protection** na lista suspensa **More**. Na caixa de diálogo **Enable Instance Protection** exibida, clique em **Yes**.

Desativar a proteção de instância

1. Acesse o console de gerenciamento.
1. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
2. Clique no nome do grupo de AS de destino.
3. Clique na guia **Instances**. Selecione uma ou mais instâncias e escolha **Disable Instance Protection** na lista suspensa **More**. Na caixa de diálogo **Disable Instance Protection** exibida, clique em **Yes**.

4.8 Colocar uma instância no estado de espera

Se você quiser interromper a distribuição de tráfego para algumas instâncias do grupo de AS, mas não quiser removê-las do grupo de AS, coloque as instâncias em espera. Você pode colocar uma ou mais instâncias em seu grupo de AS em espera e, em seguida, interromper ou reiniciar essas instâncias sem se preocupar com a remoção delas do grupo de AS.

Cenários de aplicação

Não é possível controlar o ciclo de vida de instâncias do ECS em um grupo de AS. O grupo de AS remove instâncias não saudáveis e não permite que você interrompa ou reinicie essas instâncias. Como resultado, algumas funções do ECS não estão disponíveis. Por exemplo, você não pode redefinir a senha ou reinstalar ou alterar o sistema operacional dessas instâncias.

Ao colocar instâncias do ECS em estado em espera, você pode controlar seu ciclo de vida e executar operações nelas conforme necessário, como interromper as instâncias. Isso facilita o gerenciamento de instâncias em seu grupo de AS e é útil em vários cenários diferentes.

- Se você quiser alterar o sistema operacional de um ECS adicionado por uma ação de dimensionamento ou interromper o ECS, poderá definir o ECS para o modo de espera. Em seguida, você pode executar todas as operações suportadas pelo serviço ECS. Depois de concluir as operações, cancele o modo em espera do ECS.

Por exemplo, você pode alterar a configuração de AS para seu grupo de AS a qualquer momento. Essa configuração será usada por qualquer instância criada no grupo de AS. No entanto, o grupo de AS não atualiza instâncias em execução. Você pode parar essas instâncias e o grupo de AS as substituirá. Como alternativa, você pode definir as instâncias para o modo em espera, atualizar o software nelas e, em seguida, cancelar o modo em espera para elas.

- Se uma instância em seu grupo de AS associada a um balanceador de carga ficar com defeito, você poderá definir a instância para o modo em espera, após o qual o balanceador de carga não distribuirá mais o tráfego de acesso à instância. Em seguida, você pode fazer logon na instância, localizar e corrigir a falha e reiniciar a instância. Depois que a instância for recuperada, cancele o modo em espera para que a instância receba tráfego novamente.

Regras de trabalho

- Defina as instâncias para o modo em espera.

Depois de definir uma instância para o modo em espera, ela será automaticamente desvinculada do balanceador de carga associado ao grupo de AS. A instância ainda está no

grupo de AS, mas nenhuma verificação de integridade será executada na instância. Nesse caso, a carga em outras instâncias aumentará. Para reduzir a carga em outras instâncias e garantir a execução adequada do serviço, você pode selecionar **Add the same number of new instances to the AS group** ao definir a instância para o modo em espera.

NOTA

- Uma instância pode ser definida para o modo em espera somente quando a instância estiver ativada e o grupo de AS não tiver nenhuma ação de dimensionamento contínua.
- As ações de dimensionamento não removerão instâncias em espera do grupo de AS.
- Você pode remover manualmente instâncias em espera do grupo de AS.
- Cancelar o modo em espera para instâncias.

Depois que você cancelar o modo em espera de uma instância, ela estará em estado de execução e receberá tráfego novamente. Se um balanceador de carga estiver associado ao grupo de AS, a instância será vinculada ao balanceador de carga. Depois que a instância começar a ser executada corretamente, a verificação de integridade será executada novamente.

NOTA

O modo em espera pode ser cancelado para uma instância somente quando a instância estiver no modo em espera e o grupo de AS não tiver nenhuma ação de dimensionamento contínua.

Configurar instâncias para o modo em espera

1. Acesse o console de gerenciamento.
2. Clique em **Service List**.
3. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
4. Clique no nome do grupo de AS de destino.
5. Clique na guia **Instances**. Selecione uma ou mais instâncias, clique em **More** e selecione **Set to Standby** na lista suspensa. Na caixa de diálogo exibida, selecione **Add the same number of new instances to the AS group** necessário e clique em **Yes**.

Cancelar o modo em espera para instâncias

1. Acesse o console de gerenciamento.
1. Clique em **Service List**.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Clique no nome do grupo de AS de destino.
4. Clique na guia **Instances**. Selecione uma ou mais instâncias, clique em **More** e selecione **Cancel Standby** na lista suspensa. Na caixa de diálogo exibida, clique em **Yes**.

5 Dimensionamento da largura de banda

5.1 Criação de uma política de dimensionamento de largura de banda

Cenários

Você pode ajustar automaticamente a largura de banda de EIP comprada e a largura de banda compartilhada usando uma política de dimensionamento de largura de banda. Esta seção descreve como criar uma política de dimensionamento de largura de banda.

Ao criar uma política de dimensionamento de largura de banda, você precisa configurar informações básicas. O sistema suporta três tipos de políticas de dimensionamento de largura de banda: baseadas em alarmes, agendadas e periódicas.

As informações básicas para criar uma política de dimensionamento de largura de banda incluem o nome da política, o tipo de recurso, o tipo de política e a condição de gatilho.

Criar uma política de dimensionamento de largura de banda baseada em alarmes

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.
3. Clique em **Create Bandwidth Scaling Policy**.
4. Defina parâmetros, como o nome da política, o tipo de política e a condição do acionador. Para mais detalhes, consulte [Tabela 5-1](#).

Tabela 5-1 Parâmetros de política de alarme

Parâmetro	Descrição	Exemplo de valor
Region	Especifica a região onde reside o grupo de AS.	N/A

Parâmetro	Descrição	Exemplo de valor
Policy Name	Especifica o nome da política de dimensionamento de largura de banda. O nome consiste apenas em letras, dígitos, sublinhados (_) e hifens (-).	N/A
Resource Type	Especifica o tipo da largura de banda a ser ajustada. Você pode selecionar EIP ou Shared bandwidth .	EIP
EIP	Especifica o endereço IP da rede pública cuja largura de banda precisa ser dimensionada. NOTA Somente as larguras de banda de EIPs de pagamento por uso podem ser dimensionadas.	N/A
Policy Type	Selecione Alarm .	Alarm
Alarm Rule	Você pode usar uma regra de alarme existente ou criar uma nova. Como alternativa, clique em Create Alarm Rule no lado direito do parâmetro Rule Name e crie uma regra de alarme na página Alarm Rules . Para mais detalhes, consulte Criação de uma regra de alarme . Para criar uma regra de alarme, configure os seguintes parâmetros: <ul style="list-style-type: none"> ● Nome da regra Especifica o nome da nova regra de alarme, por exemplo, as-alarm-7o1u. ● Condição de gatilho Selecione uma métrica de monitoramento e uma condição de gatilho com base na métrica. Tabela 5-2 lista as métricas de monitoramento suportadas. Um valor de exemplo é Tráfego de saída Avg. > 100 bit/s. ● Intervalo de monitoramento Especifica o período para a métrica, por exemplo, 5 minutos. ● Ocorrências consecutivas Especifica o número de vezes consecutivas, por exemplo, uma vez, para disparar uma ação de dimensionamento durante um período de monitoramento. 	N/A

Parâmetro	Descrição	Exemplo de valor
Scaling Action	<p>Especifica a ação de execução na política de AS.</p> <p>As seguintes opções de ação de dimensionamento estão disponíveis:</p> <ul style="list-style-type: none"> ● Adicionar Quando uma ação de dimensionamento é acionada, a largura de banda é aumentada. ● Reduzir Quando uma ação de dimensionamento é acionada, a largura de banda é diminuída. ● Configurar A largura de banda é definida para um valor fixo. <p>NOTA</p> <p>O passo (unidade mínima para ajuste de largura de banda) varia dependendo da faixa de valor de largura de banda. A largura de banda será automaticamente ajustada para o valor mais próximo de acordo com o passo real.</p> <ul style="list-style-type: none"> ● Se a largura de banda for menor ou igual a 300 Mbit/s, o passo padrão é 1 Mbit/s. ● Se a largura de banda varia de 300 Mbit/s a 1000 Mbit/s, a etapa padrão é de 50 Mbit/s. ● Se a largura de banda for maior que 1000 Mbit/s, a etapa padrão será 500 Mbit/s. 	N/A
Cooldown Period	<p>Especifica um período de tempo na unidade de segundo após a conclusão de cada ação de dimensionamento. Durante o período de resfriamento, as ações de dimensionamento acionadas por alarmes serão negadas. As ações de dimensionamento programadas e periódicas não são restritas.</p>	300s

Tabela 5-2 Métricas de monitoramento suportadas pela política de alarmes

Métrica	Descrição
Inbound Bandwidth	Indica a taxa de rede do tráfego de entrada.
Inbound Traffic	Indica o tráfego de rede que entra na plataforma de nuvem.
Outbound Bandwidth	Indica a taxa de rede do tráfego de saída.
Outbound Traffic	Indica o tráfego de rede saindo da plataforma de nuvem.
Outbound Bandwidth Usage	Indica o uso da taxa de rede do tráfego de saída Unidade: por cento.

5. Depois de definir os parâmetros, clique em **Create Now**.

A política de dimensionamento de largura de banda recém-criada é exibida na página **Bandwidth Scaling** e está no estado **Enabled** por padrão.

Criação de uma regra de alarme

Ao criar uma política de dimensionamento de largura de banda baseada em alarme, você pode clicar em **Create Alarm Rule** à direita de **Rule Name** para criar uma regra de alarme. Para fazer isso, execute as seguintes operações:

1. Clique em **Create Alarm Rule** à direita de **Rule Name** para alternar para a página **Alarm Rules** do Cloud Eye.
2. Na página **Alarm Rules**, clique em **Create Alarm Rule** no canto superior direito.
3. Definir parâmetros com base em [Figura 5-1](#) e [Tabela 5-3](#). Para obter mais informações sobre como definir regras de alarme, consulte [Guia de usuário do Cloud Eye](#).

Figura 5-1 Criação de uma regra de alarme

The screenshot displays the 'Create Alarm Rule' configuration interface. Key elements include:

- Resource Type:** Elastic IP and Bandwidth.
- Dimension:** Bandwidths.
- Monitoring Scope:** Specific resources.
- Method:** Create manually.
- Alarm Policy:** Outbound Bandwidth, Max., 5 minutes, 3 consecutive, >, 500 bit/s.
- Alarm Severity:** Major.
- Alarm Notification:** A notification channel is selected.

A line graph shows bandwidth usage over time, with a red vertical line indicating the 500 bit/s threshold. The graph shows several peaks, with one reaching approximately 48,454 bit/s.

Tabela 5-3 Parâmetros principais para criar uma regra de alarme

Parâmetro	Descrição	Exemplo de valor
Name	Especifica o nome da regra de alarme.	alarm-bandwidth
Description	(Opcional) Fornece informações suplementares sobre a regra de alarme.	-

Parâmetro	Descrição	Exemplo de valor
Enterprise Project	Especifica o projeto da empresa ao qual a regra de alarme pertence. Somente os usuários com as permissões do projeto empresarial podem exibir e gerenciar a regra de alarme.	default
Resource Type	Especifica o nome do serviço ao qual a regra de alarme se aplica. Defina este parâmetro como Elastic IP and Bandwidth .	Elastic IP and Bandwidth
Dimension	Especifica o item do serviço monitorado. O dimensionamento da largura de banda ajusta a largura de banda. Portanto, defina esse parâmetro como Bandwidths .	Bandwidths
Monitoring Scope	Especifica os recursos aos quais a regra de alarme se aplica. Defina este parâmetro como Specific resources . Pesquise recursos por nome de largura de banda ou ID, que podem ser obtidos na página que fornece detalhes sobre o EIP de destino.	Specific resources
Method	Selecione o método de criação de um alarme. Defina este parâmetro para Create manually .	Create manually
Alarm Policy	Especifica a política de alarme para acionar a regra de alarme. Defina este parâmetro conforme necessário. Para obter detalhes sobre as métricas de monitoramento, consulte Tabela 5-2 .	N/A

4. Depois de definir os parâmetros, clique em **Create**.
5. Na página **Create Bandwidth Scaling Policy**, clique em  à direita do **Rule Name** e selecione a regra de alarme criada.

Como alternativa, crie a regra de alarme desejada na página **Cloud Eye** antes de criar uma política de escala de largura de banda. Verifique se os recursos específicos selecionados durante a criação da regra de alarme são a largura de banda do EIP selecionado para a política de dimensionamento de largura de banda a ser criada. Depois que a regra de alarme é criada, você pode selecionar a regra ao criar uma política de dimensionamento de largura de banda.

Criação de uma política de dimensionamento de largura de banda agendada ou periódica

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.
3. Clique em **Create Bandwidth Scaling Policy**.
4. Defina parâmetros, como o nome da política, o tipo de política e a condição do acionador. Para mais detalhes, consulte [Tabela 5-4](#).

Tabela 5-4 Parâmetros de política programados ou periódicos

Parâmetro	Descrição	Exemplo de valor
Region	Especifica a região onde reside o grupo de AS.	N/A
Policy Name	Especifica o nome da política de dimensionamento de largura de banda. O nome consiste apenas em letras, dígitos, sublinhados (_) e hifens (-).	as-policy-p6g5
Resource Type	Especifica o tipo de largura de banda a ser gerenciada. Você pode selecionar EIP ou Shared bandwidth .	EIP
EIP	Especifica o endereço IP da rede pública cuja largura de banda precisa ser dimensionada. Este parâmetro é obrigatório quando Resource Type é definido como EIP . NOTA Somente as larguras de banda de EIPs de pagamento por uso podem ser dimensionadas.	N/A
Shared Bandwidth	Especifica a largura de banda compartilhada a ser dimensionada. Este parâmetro é obrigatório quando Resource Type está definido como Shared bandwidth .	N/A
Policy Type	Especifica o tipo de política. Você pode selecionar uma política programada ou periódica. Se você selecionar Periodic , será necessário configurar mais dois parâmetros: <ul style="list-style-type: none"> ● Intervalo de tempo Especifica um intervalo de tempo durante o qual a política de AS pode ser acionada. ● Período <ul style="list-style-type: none"> - Dia - Semana - Mês 	N/A
Triggered At	Especifica um horário em que a política de AS é acionada.	N/A

Parâmetro	Descrição	Exemplo de valor
Scaling Action	<p>Especifica a ação a ser executada.</p> <p>As seguintes opções de ação de dimensionamento estão disponíveis:</p> <ul style="list-style-type: none">● Adicionar Quando uma ação de dimensionamento é acionada, a largura de banda é aumentada.● Reduzir Quando uma ação de dimensionamento é acionada, a largura de banda é diminuída.● Configurar A largura de banda é definida para um valor fixo. <p>NOTA</p> <p>O passo (unidade mínima para ajuste de largura de banda) varia dependendo da faixa de valor de largura de banda. A largura de banda será ajustada automaticamente para o valor mais próximo de acordo com o passo real.</p> <ul style="list-style-type: none">● Se a largura de banda for menor ou igual a 300 Mbit/s, o passo padrão é 1 Mbit/s.● Se a largura de banda varia de 300 Mbit/s a 1000 Mbit/s, a etapa padrão é de 50 Mbit/s.● Se a largura de banda for maior que 1000 Mbit/s, a etapa padrão será 500 Mbit/s.	N/A
Cooldown Period	<p>Especifica um período de tempo na unidade de segundo após a conclusão de cada ação de dimensionamento.</p> <p>Durante o período de resfriamento, as ações de dimensionamento acionadas por alarmes serão negadas.</p> <p>As ações de dimensionamento programadas e periódicas não são restritas.</p>	300s

5. Depois de definir os parâmetros, clique em **Create Now**.

5.2 Exibição de detalhes sobre uma política de dimensionamento de largura de banda

Cenários

Você pode exibir detalhes sobre uma política de dimensionamento de largura de banda, incluindo suas informações básicas e logs de execução. Os logs de execução de políticas registram detalhes sobre a execução de políticas. Esta seção descreve como criar uma política de dimensionamento de largura de banda.

Procedimento

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.

3. Na página **Bandwidth Scaling**, clique no nome de uma política de dimensionamento de largura de banda para ir para a página que mostra suas informações básicas e exibir seus detalhes. Você pode exibir informações básicas sobre a política de dimensionamento, incluindo **Policy Type**, **Trigger Condition** e **Scaling Action**.

Exibir logs de execução de uma política de dimensionamento de largura de banda

Na área **Policy Execution Logs** da página de detalhes da política de dimensionamento de largura de banda, você pode exibir os logs de execução de políticas. Os logs de execução de políticas registram o status de execução, o tempo de execução, o valor original e o valor de destino de uma política de dimensionamento de largura de banda.

5.3 Gerenciamento de uma política de dimensionamento de largura de banda

Cenários

Você pode ajustar a largura de banda por meio de uma política de dimensionamento de largura de banda.

Esta seção descreve como gerenciar políticas de dimensionamento de largura de banda, incluindo ativar, desativar, modificar, excluir e executar imediatamente uma política de dimensionamento de largura de banda.

NOTA

A política de dimensionamento de largura de banda correspondente a um EIP liberado ainda ocupa a cota de política. Somente a conta e seus usuários do IAM com permissão global podem gerenciar a política de largura de banda de AS.

Ativar uma política de dimensionamento de largura de banda

Uma política de dimensionamento de largura de banda pode ser ativada somente quando seu status é **Disabled**.

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.
3. Na lista de políticas de dimensionamento de largura de banda, localize a linha que contém a política de destino e clique em **Enable** na coluna **Operation**.
4. Na caixa de diálogo **Enable Bandwidth Scaling Policy** exibida, clique em **Yes**.

Desativar uma política de dimensionamento de largura de banda

Uma política de dimensionamento de largura de banda só pode ser desativada quando o seu estado é **Enabled**.

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.

3. Na lista de políticas de dimensionamento de largura de banda, localize a linha que contém a política de destino e clique em **Disable** na coluna **Operation**.
4. Na caixa de diálogo **Disable Bandwidth Scaling Policy** exibida, clique em **Yes**.

 **NOTA**

Depois que uma política de dimensionamento de largura de banda é desabilitada, seu status muda para **Disabled**. O AS não aciona automaticamente qualquer ação de dimensionamento com base em uma política de dimensionamento de largura de banda **Disabled**.

Modificar uma política de dimensionamento de largura de banda

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.
3. Na lista de políticas de dimensionamento de largura de banda, localize a linha que contém a política de destino e clique no nome da política para alternar para a página de detalhes.

Clique em **Modify** no canto superior direito da página.

Você também pode localizar a linha que contém a política de destino, clicar em **More** na coluna **Operation** e selecionar **Modify**.

4. Modifique parâmetros. Você pode modificar os seguintes parâmetros de uma política de dimensionamento de largura de banda: **Policy Name**, **EIP**, **Policy Type**, **Scaling Action** e **Cooldown Period**.
5. Clique em **OK**.

 **NOTA**

Uma política de dimensionamento de largura de banda que está sendo executada não pode ser modificada.

Excluir uma política de dimensionamento de largura de banda

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.
3. Na lista de políticas de dimensionamento de largura de banda, localize a linha que contém a política de destino, clique em **More** na coluna **Operation** e selecione **Delete**.
4. Na caixa de diálogo **Delete Bandwidth Scaling Policy** exibida, clique em **Yes**.

Você também pode selecionar uma ou mais políticas de dimensionamento e clicar em **Delete** acima da lista para excluir uma ou mais políticas de dimensionamento.

 **NOTA**

- Você pode excluir uma política de dimensionamento de largura de banda quando não precisar mais dela. Se você não precisar dele apenas durante um período de tempo especificado, é aconselhável desativá-lo em vez de excluí-lo.
- Uma política de dimensionamento de largura de banda pode ser excluída somente quando não estiver sendo executada.

Executar uma política de dimensionamento de largura de banda

Ao executar uma política de dimensionamento de largura de banda, você pode ajustar imediatamente a largura de banda para aquela configurada na política de dimensionamento de largura de banda, em vez de ter que esperar até que a condição de gatilho seja atendida.

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Bandwidth Scaling**.
3. Na lista de política de dimensionamento de largura de banda, localize a linha que contém a política de destino e clique em **Execute Now** na coluna **Operation**.
4. Na caixa de diálogo **Execute Bandwidth Scaling Policy** exibida, clique em **Yes**.

Você também pode ir para a página de detalhes da política de dimensionamento de largura de banda e clicar em **Execute Now** no canto superior direito.

NOTA

- Uma política de dimensionamento de largura de banda só pode ser executada quando a política está ativada e nenhuma outra política de dimensionamento de largura de banda está a ser executada.
- A execução de uma política de dimensionamento de largura de banda não afeta o ajuste automático da largura de banda quando a condição de disparo da política é atendida.
- Se **Policy Type** for definido como **Alarm** e **Alarm Policy Type** como **Refined scaling**, a política de dimensionamento da largura de banda não poderá ser executada imediatamente.

6 Monitoramento de instância e grupo de AS

6.1 Verificação de integridade

Uma verificação de integridade remove instâncias anormais de um grupo de AS. Em seguida, o AS adiciona novas instâncias ao grupo de AS para que o número de instâncias seja o mesmo que o número esperado. Existem dois tipos de verificação de integridade do grupo de AS.

- **Verificação de integridade do ECS:** verifica o status de execução da instância do ECS. Se uma instância for interrompida ou excluída, ela é considerada anormal. A **verificação de integridade do ECS** é o modo de verificação de integridade padrão para um grupo de AS. O grupo de AS usa periodicamente o resultado da verificação para determinar o status de execução de cada instância no grupo de AS. Se os resultados da verificação de integridade mostrarem que uma instância não é saudável, o AS removerá a instância do grupo de AS.
- **Verificação de integridade do ELB:** determina o status de execução da instância do ECS usando um ouvinte de balanceamento de carga. Se o grupo de AS usar balanceadores de carga, o método de verificação de integridade também poderá ser a **verificação de integridade do ELB**. Se você adicionar vários balanceadores de carga a um grupo de AS, uma instância do ECS será considerada íntegro somente quando todos os balanceadores de carga detectarem que o status da instância está íntegro. Se algum balanceador de carga detectar que uma instância não está saudável, a instância será removida do grupo de AS.

Nos métodos de **verificação de integridade do ECS** e de **verificação de integridade do ELB**, o AS remove instâncias não íntegras dos grupos de AS. No entanto, as instâncias removidas são processadas de forma diferente nos dois cenários a seguir:

Para instâncias adicionadas automaticamente a um grupo de AS durante ações de dimensionamento, o AS as remove e as exclui. Para instâncias adicionadas manualmente a um grupo de AS, o AS apenas as remove do grupo de AS.

- As instâncias adicionadas automaticamente são cobradas no modo **Pagamento por uso** por padrão. Nesse caso, o AS remove e exclui tal instância se ela se tornar anormal. Para uma instância que é adicionada manualmente ao grupo de AS, o AS remove apenas a instância fora do grupo de AS.
- Se você alterar manualmente o modo de cobrança de uma instância de **Pagamento por uso** para **Anual/Mensal**, o AS removerá apenas a instância do grupo de AS.

Quando um grupo de AS é desabilitado, a verificação do status de integridade da instância continua. No entanto, o AS não removerá instâncias.

6.2 Configuração de notificação para um grupo de AS

Cenários

Depois que o serviço SMN é provisionado, você pode enviar prontamente informações do grupo de AS, como aumento de instância bem-sucedido, aumento de instância com falha, diminuição de instância bem-sucedida, diminuição de instância com falha ou exceção de grupo de AS para o usuário que usa o serviço SMN. Isso ajuda o usuário a aprender o status do grupo de AS.

Para configurar a notificação para um grupo de AS, você precisa especificar um evento e um tópico de notificação. Você pode configurar até cinco notificações para um grupo de AS. O tópico de notificação é pré-configurado no console de SMN. Quando a rede ativa está em conformidade com o cenário de notificação correspondente ao tópico de notificação, o grupo de AS envia uma notificação ao usuário.

Procedimento

1. Acesse o console de gerenciamento.
1. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
2. Clique no nome do grupo de AS de destino. Na página de detalhes do grupo de AS, clique na guia **Notifications** e, em seguida, clique em **Add Notification**.
3. Defina os parâmetros listados em **Tabela 6-1**.

Tabela 6-1 Descrição do parâmetro

Parâmetro	Descrição	Exemplo de valor
Event	Quando pelo menos uma das seguintes condições é atendida, a SMN envia uma notificação ao usuário: <ul style="list-style-type: none"> ● Criação de instância bem-sucedida ● Remoção de instância bem-sucedida ● Erros ocorrem em um grupo de AS ● Falha na criação da instância ● Falha na remoção da instância 	N/A
Topic	Selecionar um tópico criado. Para obter detalhes sobre como criar um tópico, consulte Criação de um tópico .	-

4. Clique em **OK**.

6.3 Gravação de operações de recursos de AS

Cenários

O AS pode usar o Cloud Trace Service (CTS) para registrar operações de recursos. O CTS pode registrar operações realizadas no console de gerenciamento, operações executadas chamando APIs e operações acionadas no sistema de nuvem.

Se você habilitou o CTS, quando uma chamada for feita para a API do AS, a operação será relatada ao CTS, que entregará o registro de operação a um bucket do OBS especificado para armazenamento. Com o CTS, você pode registrar operações associadas ao AS para consulta posterior, auditoria e operações de retrocesso.

Obtenção de informações de AS no CTS

Depois que você habilitar o CTS no sistema da aplicação, o sistema registrará as operações de chamada da API executadas nos recursos de AS. No console do **Cloud Trace Service**, você pode exibir os registros de operação dos últimos 7 dias. Para obter mais registros de operação, você pode ativar o Object Storage Service (OBS) e sincronizar os registros de operação com o OBS em tempo real.

A [Tabela 1](#) lista as operações do AS que podem ser registradas pelo CTS.

Tabela 6-2 Operações de AS que podem ser gravadas pelo CTS

Operação	Tipo de recurso	Nome do rastreamento
Criação de um grupo de AS	scaling_group	createScalingGroup
Modificação de um grupo de AS	scaling_group	modifyScalingGroup
Exclusão de um grupo de AS	scaling_group	deleteScalingGroup
Habilitação de um grupo de AS	scaling_group	enableScalingGroup
Desabilitação de um grupo de AS	scaling_group	disableScalingGroup
Criação de uma configuração de AS	scaling_configuration	createScalingConfiguration
Exclusão de uma configuração de AS	scaling_configuration	deleteScalingConfiguration

Operação	Tipo de recurso	Nome do rastreamento
Exclusão de configurações de AS em lotes	scaling_configuration	batchDeleteScalingConfiguration
Criação de uma política de AS	scaling_policy	createScalingPolicy
Modificação de uma política de AS	scaling_policy	modifyScalingPolicy
Exclusão de uma política de AS	scaling_policy	deleteScalingPolicy
Habilitação de uma política de AS	scaling_policy	enableScalingPolicy
Desativação de uma política de AS	scaling_policy	disableScalingPolicy
Execução de uma política de AS	scaling_policy	executeScalingPolicy
Remoção de uma instância	scaling_instance	removeInstance
Remoção de instâncias em lotes	scaling_instance	batchRemoveInstances
Adição de instâncias em lotes	scaling_instance	batchAddInstances
Ativação da proteção de instâncias em um lote	scaling_instance	batchProtectInstances
Desativação da proteção de instâncias em um lote	scaling_instance	batchUnprotectInstances
Configuração de uma notificação	scaling_notification	putScalingNotification
Exclusão de uma notificação	scaling_notification	deleteScalingNotification

Operação	Tipo de recurso	Nome do rastreamento
Criação de um gancho de ciclo de vida	scaling_lifecycle_hook	createLifecycleHook
Modificação de um gancho de ciclo de vida	scaling_lifecycle_hook	modifyLifecycleHook
Exclusão de um gancho de ciclo de vida	scaling_lifecycle_hook	deleteLifecycleHook

Exibir registros de auditoria

1. Acesse o console de gerenciamento.
2. Clique em  no canto superior esquerdo para selecionar uma região e um projeto.
3. Clique em **Service List**. Escolha **Management & Governance > Cloud Trace Service**.
4. Clique em **Trace List** no painel de navegação à esquerda.
5. Você pode usar filtros para consultar rastreamentos. Os seguintes filtros estão disponíveis:
 - **Trace Source, Resource Type e Search By**
Selecione um critério de filtro na lista suspensa.
Ao selecionar **Trace name** para **Search By**, você precisa selecionar um nome de rastreamento específico.
Ao selecionar **Resource ID** para **Search By**, você precisa selecionar ou inserir uma ID de recurso específica.
Se você selecionar **Resource name** para **Search By**, precisa selecionar ou inserir um nome de recurso específico.
 - **Operator**: selecione um operador específico (no nível do usuário em vez de no nível do locatário).
 - **Trace Status**: os valores disponíveis são **All trace statuses, normal, warning e incident**. Você só pode selecionar um deles.
 - Intervalo de tempo: no canto superior direito da página, você pode consultar rastreamentos na última hora, no último dia, na última semana ou em um período de tempo personalizado.
6. Clique em  à esquerda do rastreamento necessário para expandir seus detalhes.
7. Localize o rastreamento necessário e clique em **View Trace** na coluna **Operation**. Uma caixa de diálogo é exibida, mostrando o conteúdo do rastreamento.

Entradas de registro do CTS

Cada entrada de log consiste em um rastreamento no formato JSON. Uma entrada de registro indica uma solicitação de API do AS, incluindo a operação solicitada, a data e a hora da

operação, os parâmetros da operação e as informações sobre o usuário que envia a solicitação. As informações do usuário são obtidas do serviço Identity and Access Management (IAM).

O exemplo a seguir mostra entradas de registro do CTS para a ação `CreateScalingPolicy`:

```
{
  "time": "2016-12-15 15:27:40 GMT+08:00",
  "user": {
    "name": "xxxx",
    "id": "62ff83d2920e4d3d917e6fa5e31ddebe",
    "domain": {
      "name": "xxx",
      "id": "30274282b09749adbe7d9cabeebcbe8b"
    }
  },
  "request": {
    "scaling_policy_name": "as-policy-oonb",
    "scaling_policy_action": {
      "operation": "ADD",
      "instance_number": 1
    },
    "cool_down_time": "",
    "scheduled_policy": {
      "launch_time": "2016-12-16T07:27Z"
    },
    "scaling_policy_type": "SCHEDULED",
    "scaling_group_id": "ec4051a7-6fbd-42d2-840f-2ad8cdabee34"
  },
  "response": {
    "scaling_policy_id": "6a38d488-664b-437a-ade2-dc45f96f7f4c"
  },
  "code": 200,
  "service_type": "AS",
  "resource_type": "scaling_policy",
  "resource_name": "as-policy-oonb",
  "resource_id": "6a38d488-664b-437a-ade2-dc45f96f7f4c",
  "source_ip": "10.190.205.233",
  "trace_name": "createScalingPolicy",
  "trace_rating": "normal",
  "trace_type": "ConsoleAction",
  "api_version": "1.0",
  "record_time": "2016-12-15 15:27:40 GMT+08:00",
  "trace_id": "f627062b-c297-11e6-a606-eb2c0f48bec5"
}
```

6.4 Adição de tags a grupos e instâncias de AS

Cenários

Se você tiver muitos recursos do mesmo tipo, poderá usar uma tag para gerenciar recursos de forma flexível. Você pode identificar recursos especificados rapidamente usando as tags alocadas a eles.

Usando uma tag, você pode atribuir dados personalizados a cada grupo de AS. Você pode organizar e gerenciar grupos de AS, por exemplo, classificar recursos do grupo de AS por uso, proprietário ou ambiente.

Cada tag contém uma chave e um valor. Você pode especificar a chave e o valor para cada tag. Uma chave pode ser uma categoria associada a determinados valores, como uso, proprietário e ambiente.

Por exemplo, se você quiser distinguir o ambiente de teste e o ambiente de produção, poderá alocar uma tag com o **environment** de chave para cada grupo de AS. Para o ambiente de teste, o valor de chave é **test** e para o ambiente de produção, o valor de chave é **production**. É

recomendável usar um ou mais grupos de tags consistentes para gerenciar os recursos do grupo de AS.

Depois de alocar uma tag a um grupo de AS, o sistema adicionará automaticamente a tag às instâncias criadas automaticamente no grupo de AS. Se você adicionar uma tag a um grupo de AS ou modificar a tag, a nova tag será adicionada aos ECSs criados automaticamente no grupo de AS. Criar, excluir ou modificar a tag de um grupo de AS não terá impacto sobre os ECSs no grupo de AS.

Restrições ao uso de tags

Você deve observar as seguintes regras ao usar tags:

- Cada grupo de AS pode ter um máximo de 10 tags adicionadas a ele.
- Cada tag contém uma chave e um valor.
- Você pode definir o valor da tag para uma cadeia de caracteres vazia.
- Se você excluir um grupo de AS, todas as tags dele também serão excluídas.

Adicionar uma tag a um grupo de AS

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Clique no nome do grupo de AS. Na página de detalhes do grupo de AS, clique na guia **Tags** e, em seguida, clique em **Add Tag**.
4. Defina os parâmetros listados em **Tabela 6-3**.

Tabela 6-3 Regras de nomeação de tags

Parâmetro	Requisito	Exemplo de valor
Tag Key	<ul style="list-style-type: none"> ● O valor não pode estar vazio. ● Um grupo de AS tem uma chave única. ● Uma chave pode conter no máximo 36 caracteres, incluindo dígitos, letras, sublinhados (_), hifens (-) e caracteres Unicode de \u4e00 a \u9fff. 	Organization
Tag Value	<ul style="list-style-type: none"> ● O valor pode ser uma cadeia de caracteres vazia. ● Uma chave pode ter apenas um valor. ● Um valor de tag pode conter no máximo 43 caracteres, incluindo dígitos, letras, sublinhados (_), pontos (.), hifens (-) e caracteres Unicode de \u4e00 a \u9fff. 	Apache

5. Clique em **OK**.

Modificar ou excluir tags de um grupo de AS

1. Acesse o console de gerenciamento.
1. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
2. Clique no nome do grupo de AS. Na página **Overview**, clique na guia **Tags**.
3. Localize a linha que contém a tag e clique em **Edit** ou **Delete** na coluna **Operation**.
 Depois de clicar em **Edit**, configure os parâmetros necessários. Para mais detalhes, consulte [Tabela 6-3](#).
 Depois de clicar em **Delete**, a tag adicionada será excluída.

6.5 Monitoramento de métricas

Descrição

Esta seção descreve as métricas de monitoramento relatadas pelo AS para o Cloud Eye e define o namespace para as métricas. Você pode usar o Cloud Eye para consultar monitoramento de métricas e alarmes gerados pelo AS.

Namespace

SYS.AS

Monitoramento de métricas

[Tabela 6-4](#) lista as métricas de AS suportadas pelo Cloud Eye.

Tabela 6-4 Métricas do AS

ID da métrica	Métrica	Descrição	Intervalo de valores	Objeto Monitorado	Intervalo de monitoramento (dados brutos)
cpu_util	Uso da CPU	Uso da CPU de um grupo de AS Fórmula: uso total da CPU de todas as instâncias do ECS em um grupo de AS/ Número de instâncias do ECS no grupo de AS Unidade: por cento	≥0%	Grupo de AS	5 minutos

ID da métrica	Métrica	Descrição	Intervalo de valores	Objeto Monitorado	Intervalo de monitoramento (dados brutos)
mem_util	Uso da memória	<p>Uso de memória de um grupo de AS</p> <p>Fórmula: uso total de memória de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS</p> <p>Unidade: por cento</p> <p>NOTA Essa métrica não estará disponível se a imagem não tiver ferramentas da VM instaladas.</p>	≥0%	Grupo de AS	5 minutos
instance_num	Número de instâncias	<p>Número de instâncias do ECS disponíveis em um grupo de AS</p> <p>Fórmula: número total de instâncias do ECS no estado Enabled no grupo de AS</p>	≥0	Grupo de AS	5 minutos
network_incoming_bytes_rate_inband	Taxa de entrada em banda	<p>Número de bytes por segundo de entrada em um ECS em um grupo de AS</p> <p>Fórmula: taxas totais de entrada em banda de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS</p> <p>Unidade: byte/s</p>	≥0 Byte/s	Grupo de AS	5 minutos
network_outgoing_bytes_rate_inband	Taxa de saída em banda	<p>Número de bytes por segundo de saída em um ECS em um grupo de AS</p> <p>Fórmula: taxas totais de saída em banda de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS</p> <p>Unidade: byte/s</p>	≥0 Byte/s	Grupo de AS	5 minutos

ID da métrica	Métrica	Descrição	Intervalo de valores	Objeto Monitorado	Intervalo de monitoramento (dados brutos)
disk_read_bytes_rate	Taxa de leitura de discos	Número de bytes lidos de um grupo de AS por segundo Fórmula: taxas de leitura de discos totais de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: byte/s	≥ 0 Byte/s	Grupo de AS	5 minutos
disk_write_bytes_rate	Taxa de gravação de discos	Número de bytes gravados em um grupo de AS por segundo Fórmula: taxas de gravação totais de discos de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: byte/s	≥ 0 Byte/s	Grupo de AS	5 minutos
disk_read_requests_rate	Solicitações de leitura de discos	Número de solicitações de leitura por segundo enviadas para um disco de ECS em um grupo de AS Fórmula: taxas de leitura de discos totais de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: solicitação/s	≥ 0 solicitação/s	Grupo de AS	5 minutos

ID da métrica	Métrica	Descrição	Intervalo de valores	Objeto Monitorado	Intervalo de monitoramento (dados brutos)
disk_write_request_rate	Solicitações de gravação nos discos	Número de solicitações de gravação por segundo enviadas para um disco do ECS em um grupo de AS Fórmula: taxas de gravação totais de discos de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: solicitação/s	≥ 0 solicitação/s	Grupo de AS	5 minutos
cpu_usage	(Agente) Uso da CPU	Uso da CPU do agente de um grupo de AS Fórmula: uso total da CPU do agente de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: por cento	0-100%	Grupo de AS	1 minuto
mem_usedPercent	(Agente) Uso da memória	Uso da memória do agente de um grupo de AS Fórmula: uso total da memória do agente de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: por cento	0-100%	Grupo de AS	1 minuto
load_averagel	(Agente) Carga média de 1 minuto	Carga média da CPU de todos os ECSs em um grupo de AS no último 1 minuto Unidade: nenhum	≥ 0	Grupo de AS	1 minuto
load_averagel5	(Agente) Carga média de 5 minutos	Carga média da CPU de todos os ECSs em um grupo de AS nos últimos 5 minutos Unidade: nenhum	≥ 0	Grupo de AS	1 minuto

ID da métrica	Métrica	Descrição	Intervalo de valores	Objeto Monitorado	Intervalo de monitoramento (dados brutos)
load_average15	(Agente) Carga média de 15 minutos	Carga média da CPU de todos os ECSs em um grupo de AS nos últimos 15 minutos Unidade: nenhum	≥ 0	Grupo de AS	1 minuto
gpu_usage_gpu	(Agente) Uso da GPU	Uso da GPU do agente de um grupo de AS Fórmula: uso total da GPU do agente de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: por cento	0-100%	Grupo de AS	1 minuto
gpu_usage_mem	(Agente) Uso da memória de vídeo	Uso da memória da GPU do agente de um grupo de AS Fórmula: uso total da memória da GPU do agente de todas as instâncias do ECS em um grupo de AS/Número de instâncias do ECS no grupo de AS Unidade: por cento	0-100%	Grupo de AS	1 minuto

 **NOTA**

As métricas de monitoramento são classificadas em métricas com agente e sem agente. Para alguns sistemas operacionais, é necessário instalar o agente para obter as métricas de monitoramento correspondentes. Nesse caso, selecione as métricas de monitoramento com o agente, por exemplo, (Agente) Uso da memória.

 **NOTA**

Os SOs determinam se as métricas **Memory Usage**, **Inband Outgoing Rate** e **Inband Incoming Rate** são suportadas. Para obter detalhes, consulte [Guia de usuário do Elastic Cloud Server](#).

Antes de usar as métricas de monitoramento do agente, verifique se o agente foi instalado nas instâncias do grupo de AS. Para obter detalhes, consulte [Como fazer para instalar o plug-in do agente nas instâncias de um grupo de AS para usar as métricas de monitoramento do agente?](#)

Dimensão

Chave	Valor
AutoScalingGroup	ID de grupo do AS

6.6 Exibição de métricas de monitoramento

Cenários

A plataforma de nuvem fornece o Cloud Eye para ajudá-lo a obter o status de execução das instâncias do ECS. Esta seção descreve como exibir detalhes das métricas do grupo de AS para obter informações sobre o status das instâncias do ECS no grupo de AS.

Pré-requisitos

A instância do ECS está sendo executada corretamente.

NOTA

- As métricas de monitoramento, como **CPU Usage** e **Disks Read Rate**, estão disponíveis somente quando há pelo menos uma instância em um grupo de AS. Caso contrário, apenas a métrica **Number of Instances** está disponível.
- Os dados de monitoramento não são exibidos para uma instância interrompida, com defeito ou excluída. Depois que essa instância é reiniciada ou recuperada, os dados de monitoramento ficam disponíveis.

Exibir métricas de monitoramento no Auto Scaling

1. Acesse o console de gerenciamento.
2. Em **Compute**, clique em **Auto Scaling**. No painel de navegação à esquerda, escolha **Instance Scaling**. Em seguida, clique na guia **AS Groups**.
3. Na página **AS Groups**, localize o grupo de AS para exibir dados de monitoramento e clique em seu nome.
4. Clique na guia **Monitoring** para exibir os dados de monitoramento.

Você pode exibir dados das últimas uma, três, 12 ou 24 horas ou dos últimos 7 dias. Se você quiser exibir dados para um intervalo de tempo mais longo, clique em **View details**

para ir para a página **Cloud Eye**, passe o mouse sobre um gráfico e clique em .

Exibir métricas de monitoramento no Cloud Eye

1. Acesse o console de gerenciamento.
2. Clique em  no canto superior esquerdo para selecionar uma região e um projeto.
3. Em **Management & Governance**, selecione **Cloud Eye**.
4. No painel de navegação à esquerda, escolha **Cloud Service Monitoring > Auto Scaling**.
5. Localize a linha que contém o grupo de AS de destino e clique em **View Metric** na coluna **Operation** para exibir os dados de monitoramento.

Você pode exibir dados das últimas uma, três, 12 ou 24 horas ou dos últimos 7 dias.

Passa o mouse sobre um gráfico e clique em  para visualizar dados para um intervalo de tempo mais longo.

NOTA

Pode levar um período de tempo para obter e transferir os dados de monitoramento. Portanto, aguarde um pouco e depois verifique os dados.

6.7 Configuração de regras de alarme de monitoramento

Cenários

Definir regras de alarme permite que você personalize os objetos monitorados e as políticas de notificação e determine o status de execução das instâncias do ECS a qualquer momento.

Procedimento

1. Acesse o console de gerenciamento.
2. Clique em  no canto superior esquerdo e selecione a região e o projeto desejados.
3. Em **Management & Governance**, selecione **Cloud Eye**.
4. No painel de navegação, escolha **Alarm Management > Alarm Rules**.
5. Na página **Alarm Rules**, clique em **Create Alarm Rule** para criar uma regra de alarme para o serviço AS ou modificar uma regra de alarme existente do serviço AS.
6. Depois de definir os parâmetros, clique em **Create**.

NOTA

- Para obter mais informações sobre como definir regras de alarme, consulte o *Guia de usuário do Cloud Eye*.
- Você pode criar regras de alarme no console do Cloud Eye para expandir dinamicamente os recursos.

7 Gerenciamento de permissões

7.1 Criação de um usuário e concessão de permissões de AS

Cenários

O **IAM** pode ajudá-lo a implementar controle de permissões refinado sobre seus recursos de AS. Com o IAM, você pode

- Criar usuários do IAM para funcionários com base na estrutura organizacional da sua empresa. Cada usuário do IAM terá suas próprias credenciais de segurança para acessar os recursos do AS.
- Conceder somente as permissões necessárias para que os usuários executem uma tarefa específica.
- Usar o **IAM** para confiar uma conta da Huawei Cloud ou serviço em nuvem para executar O&M eficiente em seus recursos de AS.

Se sua conta da Huawei Cloud não exigir usuários individuais do IAM, pule esta seção.

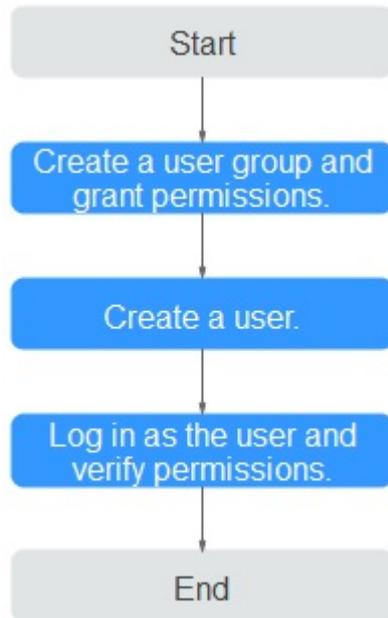
Esta seção descreve o procedimento para conceder permissões (consulte [Figura 7-1](#)).

Pré-requisitos

Saiba mais sobre as permissões (consulte [Gerenciamento de permissões](#)) suportadas pelo AS e escolha políticas ou funções de acordo com suas necessidades. Para obter as permissões de outros serviços, consulte [Permissões do sistema](#).

Fluxo do processo

Figura 7-1 Processo para conceder permissões de AS



1. **Criar um grupo de usuários e atribuir permissões a ele.**
Criar um grupo de usuários no console do IAM e anexar a política **AutoScaling ReadOnlyAccess** ao grupo.
2. **Criar um usuário do IAM e adiciona-o ao grupo de usuários.**
Criar um usuário no console do IAM e adicionar o usuário ao grupo criado em 1.
3. **Fazer logon** e verificar permissões.
Fazer logon no console de AS como o usuário criado e verifique as permissões do usuário para AS.
 - Escolher **Service List > Auto Scaling**. Em seguida, clique em **Create AS Group** no console de AS. Se aparecer uma mensagem indicando que você não tem permissões suficientes para executar a operação, a política **AutoScaling ReadOnlyAccess** já entrou em vigor.
 - Escolha qualquer outro serviço na **Service List**. Se aparecer uma mensagem indicando que você não tem permissões suficientes para acessar o serviço, a política **AutoScaling ReadOnlyAccess** já entrou em vigor.

7.2 Políticas personalizadas do AS

Cenários

As políticas personalizadas podem ser criadas para complementar as políticas definidas pelo sistema do AS. Para as ações que podem ser adicionadas às políticas personalizadas, consulte [Políticas de permissões e ações suportadas](#).

Você pode criar políticas personalizadas de uma das seguintes maneiras:

- Editor visual: selecione serviços em nuvem, ações, recursos e condições de solicitação. Isso não requer conhecimento de sintaxe de política.
- JSON: edite políticas de JSON do zero ou com base em uma política existente.

Para obter detalhes da operação, consulte [Criação de uma política personalizada](#). A seção a seguir contém exemplos de políticas personalizadas comuns de AS.

Exemplo de políticas personalizadas

- Exemplo 1: permitir que os usuários removam instâncias de um grupo de AS e criem configurações de AS

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "as:instances:delete",
        "as:configs:create"
      ]
    }
  ]
}
```

- Exemplo 2: negar a exclusão do grupo de AS

Uma política com apenas permissões "Deny" deve ser usada em conjunto com outras políticas para entrar em vigor. Se as permissões atribuídas a um usuário contiverem "Allow" e "Deny", as permissões "Deny" terão precedência sobre as permissões "Allow".

O método a seguir pode ser usado se você precisar atribuir permissões da política **AutoScaling FullAccess** a um usuário, mas quiser impedir que o usuário exclua grupos de AS. Crie uma política personalizada para negar a exclusão do grupo de AS e anexe ambas as políticas ao grupo ao qual o usuário pertence. Em seguida, o usuário pode executar todas as operações no AS, exceto excluir grupos de AS. O seguinte é um exemplo de uma política de negação:

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Action": [
        "as:groups:delete"
      ],
      "Effect": "Deny"
    }
  ]
}
```

A Histórico de mudanças

Lançado em	Descrição
30/10/2021	Esta edição é o décimo terceiro lançamento oficial. Modificação do seguinte conteúdo: adição da seção "Gerenciamento de permissões."
19/10/2020	Esta edição é o décimo segundo lançamento oficial. Modificação do seguinte conteúdo: adição da seção "Métodos de acesso".
30/08/2019	Esta edição é o décimo primeiro lançamento oficial. Modificação do seguinte conteúdo: adição de métricas de monitoramento do agente em Monitoramento de métricas .
30/01/2019	Esta edição é o décimo lançamento oficial. Modificação do seguinte conteúdo: <ul style="list-style-type: none">● Otimização de conteúdo sobre a política de AS e a política de dimensionamento de largura de banda.● Adição de Colocar uma instância no estado de espera.

Lançado em	Descrição
30/11/2018	<p>Esta edição é o nono lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● Adição das operações para exibir ações de dimensionamento históricas para Visualização de uma ação de dimensionamento. ● As perguntas frequentes organizadas por categoria. ● Adição de perguntas frequentes "Qual é o número esperado de instâncias?", "Que operação será suspensa após um grupo de AS ser desativado?", "Por que as instâncias em um grupo de AS frequentemente falham nas verificações de integridade e são excluídas e, em seguida, criadas repetidamente?", "Quais são as condições para acionar um alarme na política de AS?", e "Precisar configurar um EIP em uma configuração de AS quando um balanceador de carga foi habilitado em um grupo de AS?" ● Otimização de perguntas frequentes "O que poder fazer para permitir que minha aplicação seja implementado automaticamente em uma instância?" e "O que é um período de resfriamento? Por que é necessário?"
30/09/2018	<p>Esta edição é o oitavo lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● Adição do parâmetro ECS Group a Criação de uma configuração do AS a partir de uma instância do ECS existente e Criação de uma configuração do AS a partir de um novo modelo de especificações. ● Adição do parâmetro Bandwidth Type a Criação de uma configuração do AS a partir de uma instância do ECS existente e Criação de uma configuração do AS a partir de um novo modelo de especificações.
30/08/2018	<p>Esta edição é o sétimo lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● O período de resfriamento começa depois que uma ação de dimensionamento é concluída e o período de resfriamento padrão é de 300 segundos. ● Adição de perguntas frequentes "O que é um período de resfriamento? Como é calculado?" ● Adição dos cenários de aplicação da largura de banda de dimensionamento em Dimensionamento da largura de banda.

Lançado em	Descrição
30/07/2018	<p>Esta edição é o sexto lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● Otimização de "Criação de um grupo de AS rapidamente", "Criação de um grupo de AS" e "Expansão dinâmica de recursos". ● Adição de perguntas frequentes "Os dados de uma instância serão retidos depois que a instância for removida de um grupo de AS e excluída?" ● Adição de perguntas frequentes "O AS pode dimensionar a capacidade com base no monitoramento personalizado do Cloud Eye?" ● Adição de perguntas frequentes "O AS pode dimensionar e reduzir automaticamente vCPUs, memória e largura de banda de ECSs?" ● Adição de opção de monitoramento personalizado quando Alarm está definido para Create Alarm Rule para Dimensionamento dinâmico. ● Adição de 10s e 1 min como novas opções para o intervalo de verificação de integridade em Criação de um grupo de AS.
30/06/2018	<p>Esta edição é o quinto lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● Otimização de "Criação de um grupo de AS" e "Criação de uma configuração de AS". ● Adição de perguntas frequentes "O que poder fazer se uma política de AS falhar ao ser ativada?" ● Adição de perguntas frequentes "Poder adicionar ECSs anuais/mensais?" ● Adição de perguntas frequentes "Como fazer para impedir que os ECSs em um grupo de AS sejam removidos automaticamente?"

Lançado em	Descrição
30/05/2018	<p>Esta edição é o quarto lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● Adição do recurso de dimensionamento de largura de banda. ● Adição do parâmetro "Grupo de segurança" para Criação de uma configuração do AS a partir de um novo modelo de especificações. ● Adição de "Visualização de logs de auditoria" para Gravação de operações de recursos de AS. ● Adição de Monitoramento de métricas. ● Adição de Exibição de métricas de monitoramento. ● Adição de Configuração de regras de alarme de monitoramento. ● Adição de perguntas frequentes "O que poder fazer se o grupo de AS falhar ao ser ativado?" ● Adição de perguntas frequentes "Como lidar com instâncias não íntegras em um grupo de AS?"
30/04/2018	<p>Esta edição é o terceiro lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● adição de perguntas frequentes "Como fazer para excluir um ECS criado em uma ação de dimensionamento?"
30/03/2018	<p>Esta edição é o segundo lançamento oficial.</p> <p>Modificação do seguinte conteúdo:</p> <ul style="list-style-type: none"> ● Otimização de descrição da injeção de dados do usuário. ● Otimização de descrição da verificação de integridade da instância.
31/12/2017	<p>Esta edição é o primeiro lançamento oficial.</p>