# Auto Scaling

# Scaling Groups

# Product Documentation

# Contents

# Scaling Groups

# Scaling Group Overview

Last updated：2020-01-06 14:46:57

A scaling group contains a collection of CVM instances that follow the same policies and have a shared purpose. Scaling groups define attributes such as the maximum and minimum numbers of CVM instances in the group and their associated CLB instances.

# Creating a Scaling Group

Last updated：2021-01-08 16:56:12

## Overview

This document describes how to create a scaling group in the Auto Scaling console.

## Directions

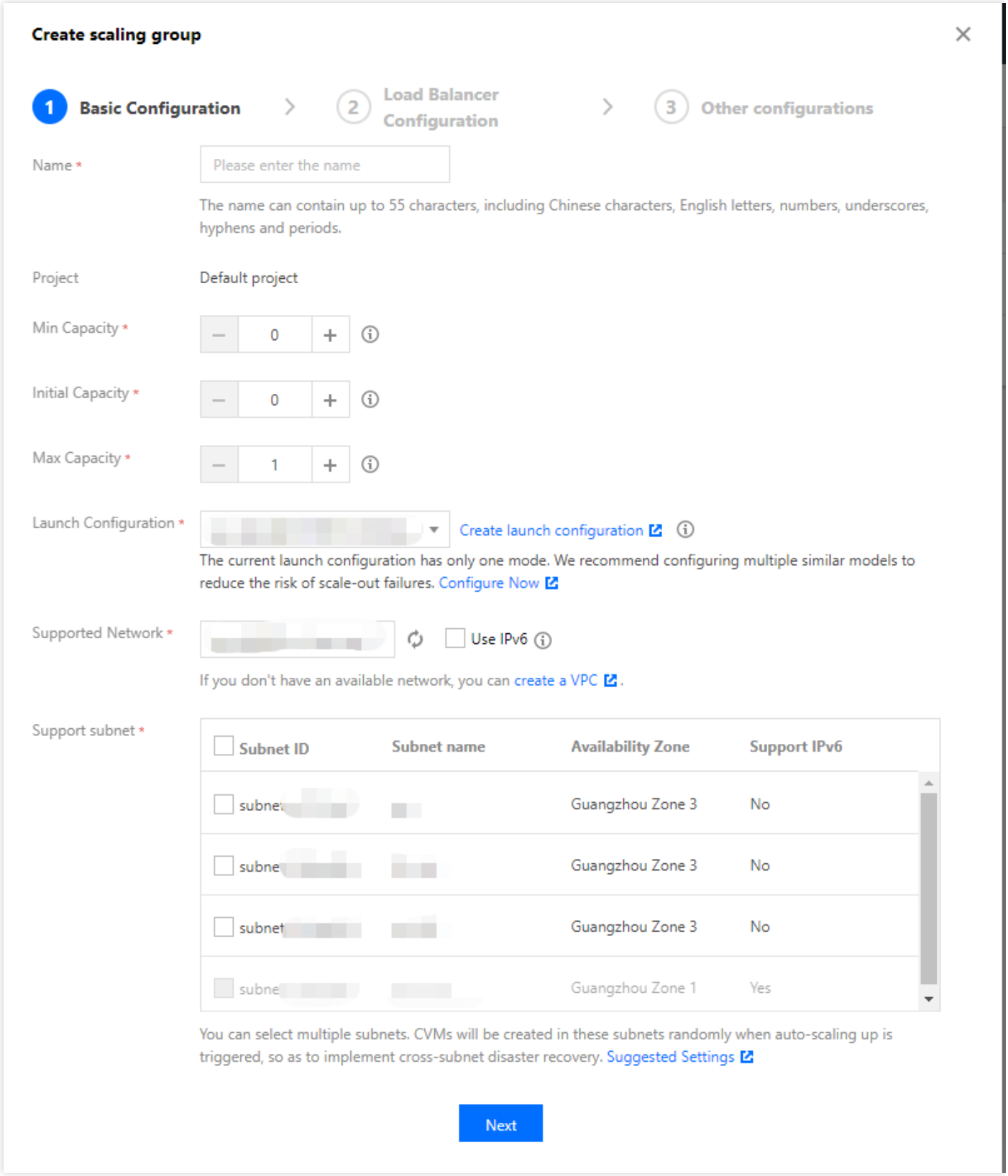
### Selecting a region

1. Log in to the Auto Scaling console and click **Scaling group** in the left sidebar.
2. At the top of the **Scaling group** page, select the region where a scaling group will be created.
   CVM instances and CLB instances must be in the same region as the one specified for launch configuration. For example, if the Guangzhou region is specified for the launch configuration, only CVM instances in Guangzhou will be automatically added to the scaling group. For a scaling group in Guangzhou, you cannot add CVM instances or bind CLB instances from other regions (such as Shanghai, Beijing, Hong Kong (China), or Toronto).

### Configuring the scaling group

1. On the **Scaling group** page, click **Create**.

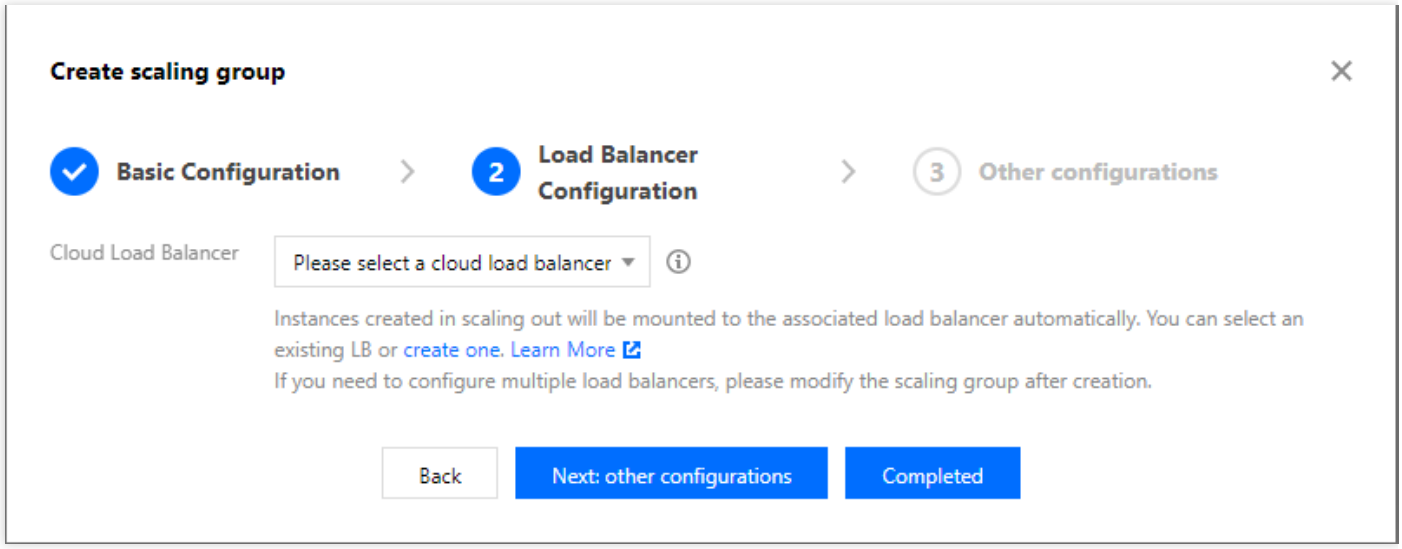
2. In the pop-up window, complete the basic configuration of the scaling group as shown below:




- **Name**: identifies the scaling group with a custom name, such as “Website logic layer”.
- **Min Capacity**: defines the minimum number of instances in the scaling group.

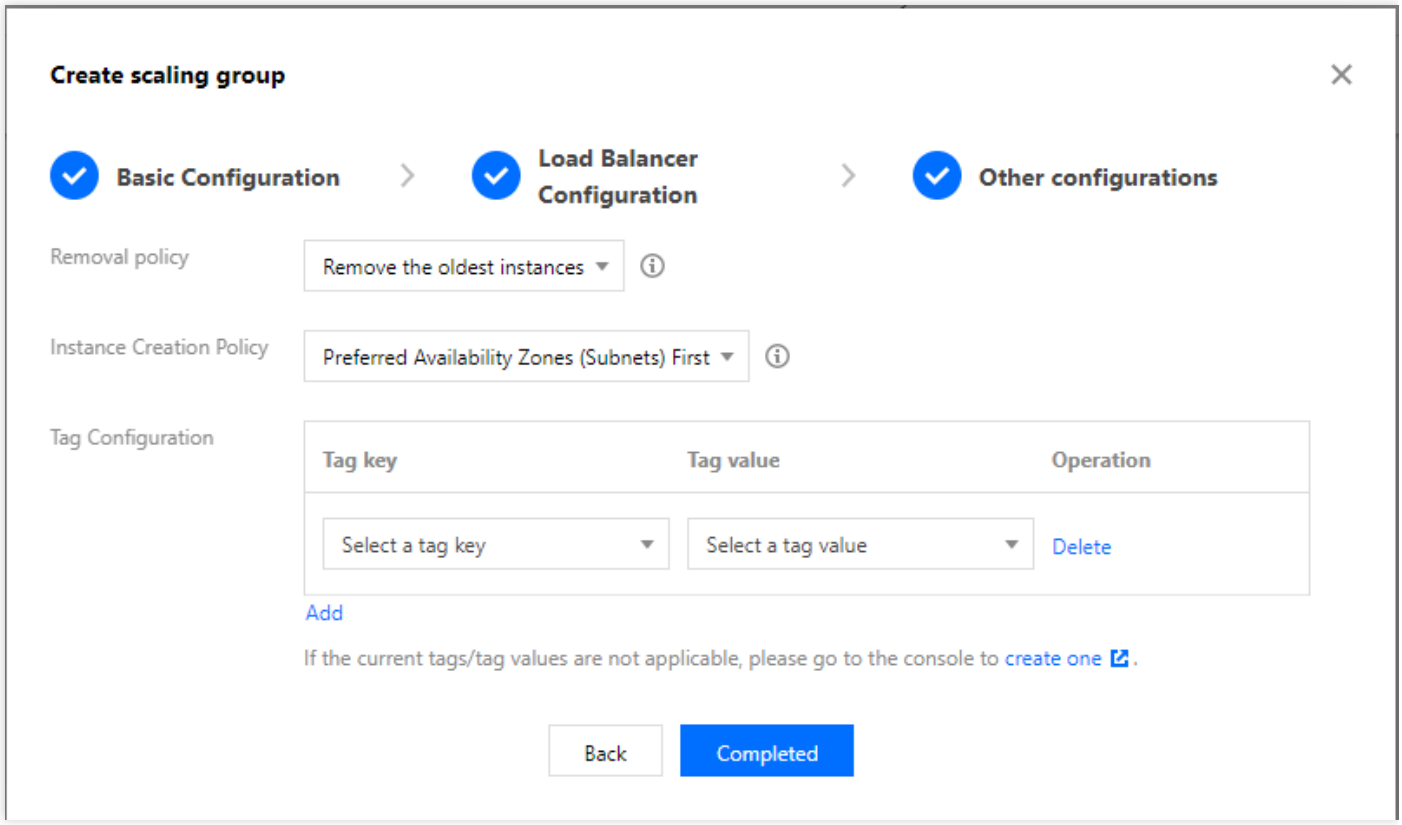
- **Initial Capacity**: defines the number of **automatically** created instances when the scaling group is created.
- **Max Capacity**: defines the maximum number of instances in the scaling group.
- **Launch Configuration**: specifies the launch configuration to scale out CVM instances.
- **Supported Network**: specifies the network attribute of the scaled-out CVMs. Select a VPC in which CVM instances will be scaled out.
- **Support subnet**: specifies the subnet in which the CVMs will be scaled out. You can select multiple subnets to locate the CVM instances automatically created for scale-out, implementing cross-subnet disaster recovery.

3. Click **Next** to complete the load balancer configuration as shown below:




You can select an existing CLB or create one. Instances created for scale-out will be mounted to the associated CLB automatically. If you need to configure multiple load balancers, please modify the scaling group after creation.

4. (Optional) Click **Next: other configurations** to complete configurations as shown below or directly click **Completed** to skip this step.




- **Removal policy**: identifies which CVMs should be removed first when AS needs to remove instances from the scaling group for scale-in:
  - **Remove the oldest instances**: removes the earliest auto-added instances in the scaling group. Instances added automatically are removed first, and then the oldest instances that are manually added. We recommend you choose this option.
  - **Remove the latest instances**: removes the latest auto-added CVMs in the scaling group. Instances added automatically are removed first, and then the latest instances that are manually added.
- **Instance Creation Policy**: specifies the policy to add instances in different availability zones of multiple subnets for scaling-out. Available options include:
  - **Preferred Availability Zones (Subnets) First**: the availability zones (subnets) will be selected sequentially from top to bottom of the configuration list till success. This mode is suitable for architectures with one primary availability zone and other secondary availability zones.
  - **Multiple Availability Zones (Subnets) Distribution**: during scale-out, the system will select availability zones (subnets) with relatively few instances in which to create new instances. This mode is suitable for architectures where instances need to be evenly distributed.
- **Tag Configuration**: categorizes and manages resources with tags. For more information, see Tag.

5. Click **Completed**.

# Subsequent Operations

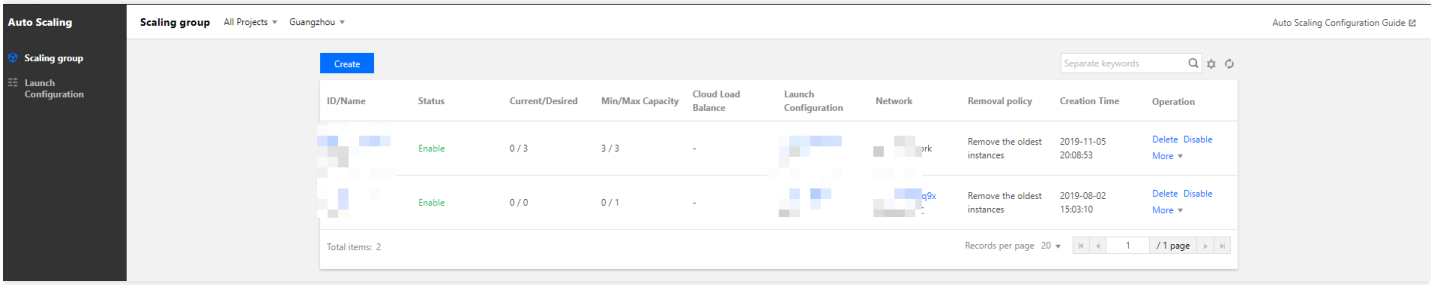
The scaling group is now created. To implement auto scaling, you need to proceed with the following 3 operations:

- Expanding Capacity Manually
- Managing an Alarm-triggered Policy
- Creating Scaling Activity Notifications

# Viewing Scaling Group List

Last updated：2020-07-16 17:02:39

Log in to the Auto Scaling Console, and click Scaling group in the left sidebar to view the list, as shown below:



# Modifying Scaling Groups

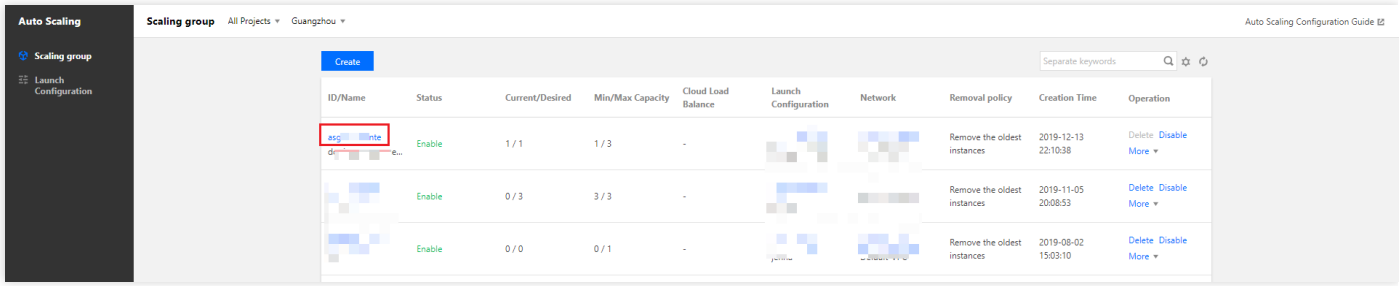
Last updated：2020-12-23 16:15:01

1. Log in to the Auto Scaling console and click **Scaling group** in the left sidebar.
2. Click the **ID**/**Name** of the scaling group to be modified to enter its details page, as shown below:
3. Select the **Scaling group details** tag, click **Edit** to modify its name, adjust the min capacity and max capacity, or modify the CVM instance removal policy, etc.
4. After modification, click **Save**.

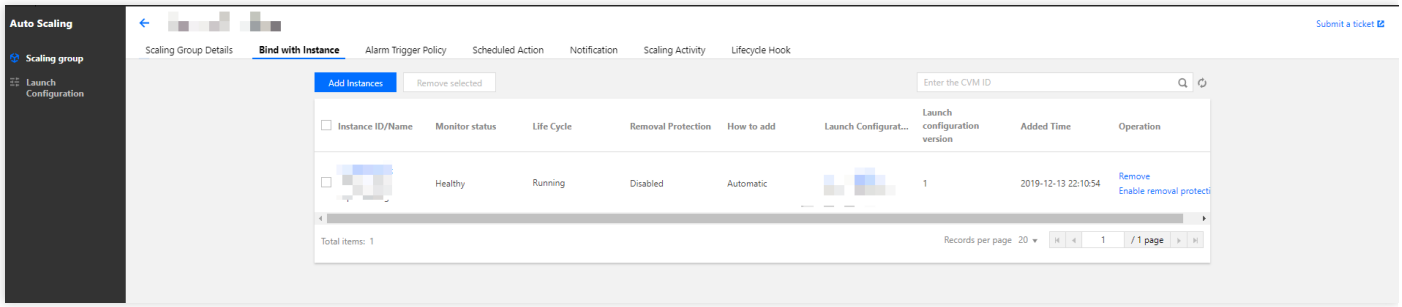
# Modifying Bound Instances

Last updated：2022-04-19 19:13:39

1. Log in to the Auto Scaling console and click **Scaling group** in the left sidebar.
2. Click the **ID/Name** of the scaling group to be modified to enter its details page, as shown in the figure below:




3. Select the **Bind with Instance** tab where you can view the list of CVMs that are bound to the scaling group, as shown in the figure below:




- To manually add CVM instances to the scaling group, click **Add Instances**, select the instance to be added (hold the **Shift** key to select multiple instances) and click **OK**.
- To unbind a specific CVM, click **Remove** under its **Operation** column.

> Note：
>
> - CVMs that are automatically created in scale-out event will be terminated when they’re removed from the scaling group.
> - CVMs that are manually added will not be terminated upon removal. They will only be removed from the scaling group, and the Cloud Load Balancer will be unbound.

# Combining Load Balancers and Scaling Groups

Last updated：2020-04-23 16:50:34

When Auto Scaling (AS) adds or deletes CVM instances, you need to ensure that app traffic is distributed across all CVM instances. If you want the scaled-out CVMs to be bound with a specific load balancer (LB) and receive the traffic forwarded by that LB without your intervention, you can specify an LB for the CVMs. In this case, the LB will work as the single point of contact for all inbound traffic towards the instances in your Auto Scaling group.

## Binding an LB to a Scaling Group

Integrate scaling groups with Cloud Load Balancer (CLB) so you can bind a CLB instance to an existing scaling group. After the CLB instance is bound, it automatically registers the instances in the scaling group and distributes inbound traffic to these instances.

1. Log in to the Auto Scaling console and click **Scaling Groups** in the left sidebar.
2. On the "Scaling Groups" page, click **Create**.
3. In the "CLB configuration" step of creating a scaling group, select the desired CLB. If no CLBs are available, click **Create** under the option to create one.

> A scaling group and its associated CLB instance (in the case of a cross-region CLB instance, its backend VPC) must be in the same network environment (the same VPC instance or the basic network in the same region).

## Unbinding a CLB from a Scaling Group

On the "Scaling Groups" page, click a scaling group ID to go to the details page for the corresponding scaling group. You can delete the corresponding CLB in the "CLB Information" section.

Once the CLB is deleted, CVMs in the scaling group will also be automatically unbound from the deleted CLB.

# Delete Scaling Groups

Last updated：2019-08-13 19:53:24

Open the Console, and select **Scaling Group** in the navigation bar.

There is a **Delete** button behind each scaling group in the scaling group list. Note: you need to delete the instances in the scaling group before your can delete the scaling group itself.