Auto Scaling

Getting Started

Product Introduction

©2013-2018 Tencent Cloud. All rights reserved.

Copyright Notice

©2013-2018 Tencent Cloud. All rights reserved.

Copyright in this document is exclusively owned by Tencent Cloud. You must not reproduce, modify, copy or distribute in any way, in whole or in part, the contents of this document without Tencent Cloud's the prior written consent.

Trademark Notice

Tencent Cloud

All trademarks associated with Tencent Cloud and its services are owned by Tencent Cloud Computing (Beijing) Company Limited and its affiliated companies. Trademarks of third parties referred to in this document are owned by their respective proprietors.

Service Statement

This document is intended to provide users with general information about Tencent Cloud's products and services only and does not form part of Tencent Cloud's terms and conditions. Tencent Cloud's products or services are subject to change. Specific products and services and the standards applicable to them are exclusively provided for in Tencent Cloud's applicable terms and conditions.

Contents

Getting Started
- Creating a Scaling Solution
- Step 1: Create a startup configuration
- Step 2: Create a scaling group
- Step 3: Create a scaling policy

©2013-2018 Tencent Cloud. All rights reserved.

Getting Started
Creating a Scaling Solution

Last updated：2018-05-29 14:40:24

You can create a complete auto scaling solution by performing the following three steps:

> Note: The example below are performed in the console. If you want to use API, refer to API Usage Examples.

Step 1: Create a Scaling Configuration

The scaling configuration defines the configuration information of CVM instances used for auto scaling, including the image, storage, network, security group, login method of the CVM.

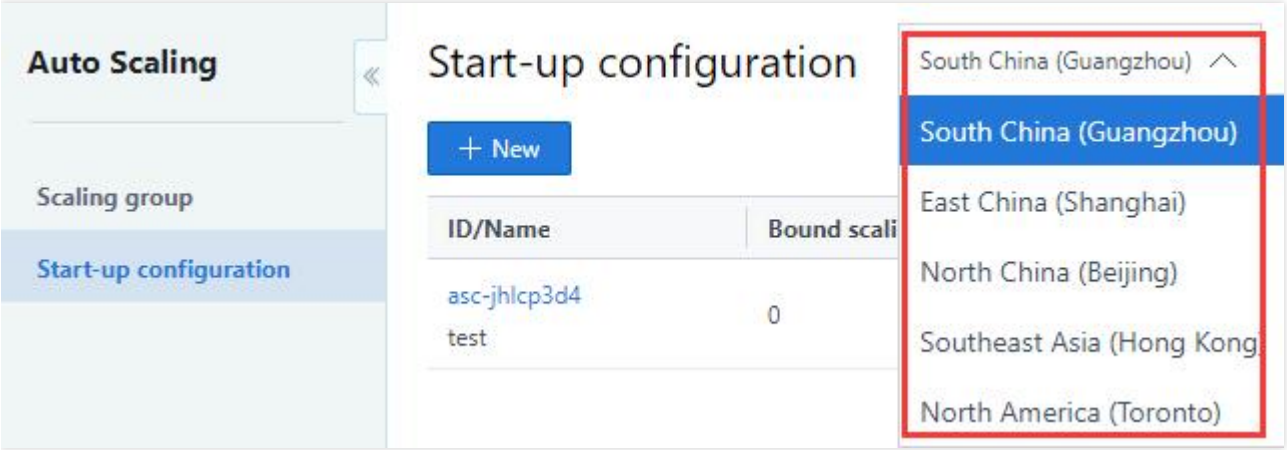
> Note: The scaling configuration is created **completely free of charge**.

Log in to Auto Scaling Console, and click **Scaling Configuration** in the navigation bar.

Select a Region

The CVMs which can be added manually and the cloud load balancers which can be bound are restricted by your selected region. For example, if you select Guangzhou as the region of the scaling configuration, the CVM of Guangzhou is automatically added to the scaling group. In a scaling group of Guangzhou region, you cannot add CVMs in other regions (Shanghai, Beijing, Hong Kong, Toronto, etc.) manually, nor bind load balancers in other regions (Shanghai, Beijing, Hong Kong, Toronto, etc.).

©2013-2018 Tencent Cloud. All rights reserved.

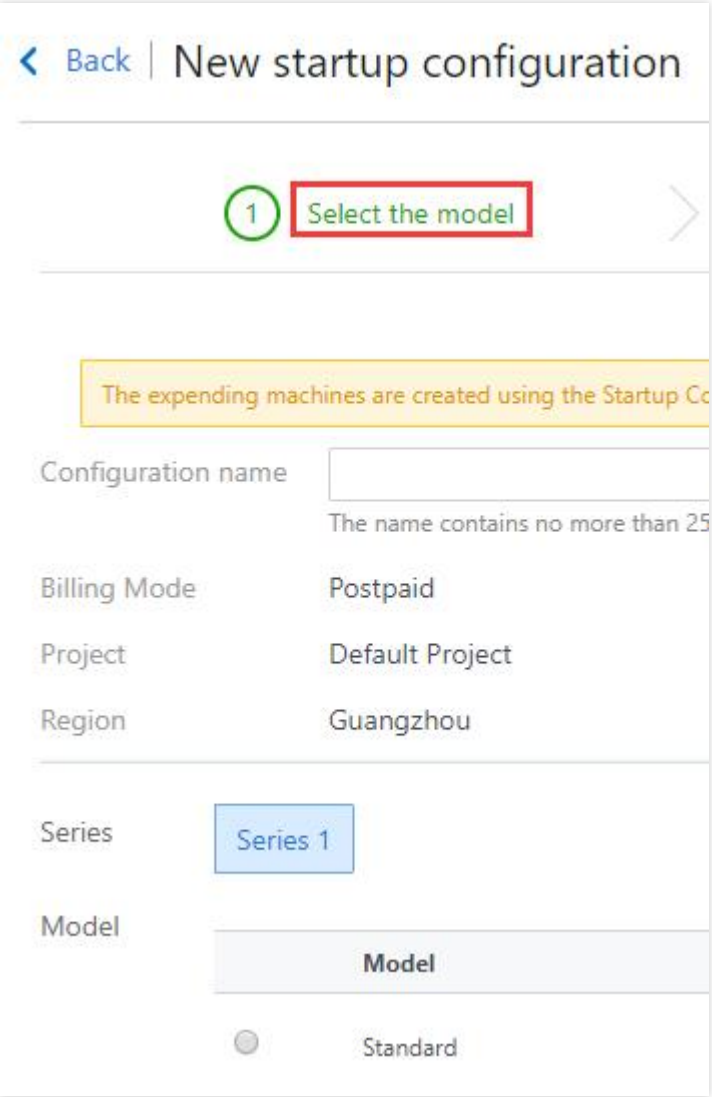


Click , and fill in the basic scaling configuration information in the pop-up page.

Select a Model

- Name the configuration.
- Select the same CVM model to which the scaling group is going to bind.

©2013-2018 Tencent Cloud. All rights reserved.



Select an Image

You can use a public image or a custom image to create the scaling configuration.

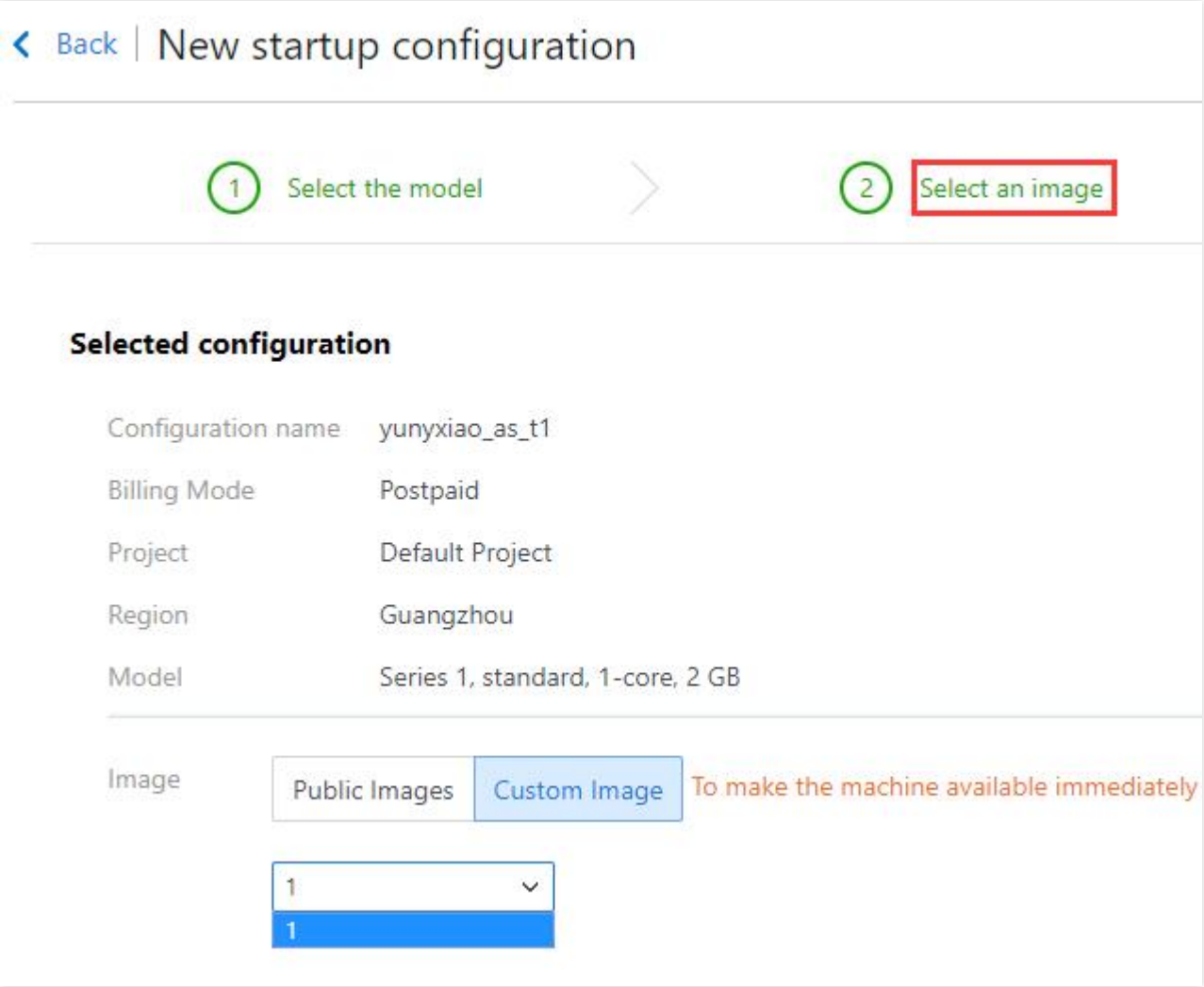
It is strongly recommended that you use the custom image of which the environment has been deployed. The reasons include the following:

- If you select a public image, the scaled CVM instance is a clear OS, and you need to manually deploy the application environment;
- If you select a custom image, by creating an image for the CVM instance of which the environment has been deployed, and using the image to create CVM instances in batches, the CVM instances will have the same software environment as the original CVM instance after created successfully, and thus the batch deployment is achieved.

Therefore, it is recommended to select a custom image here.
Bind the image of the **CVM to which the scaling group is going to bind**.

©2013-2018 Tencent Cloud. All rights reserved.

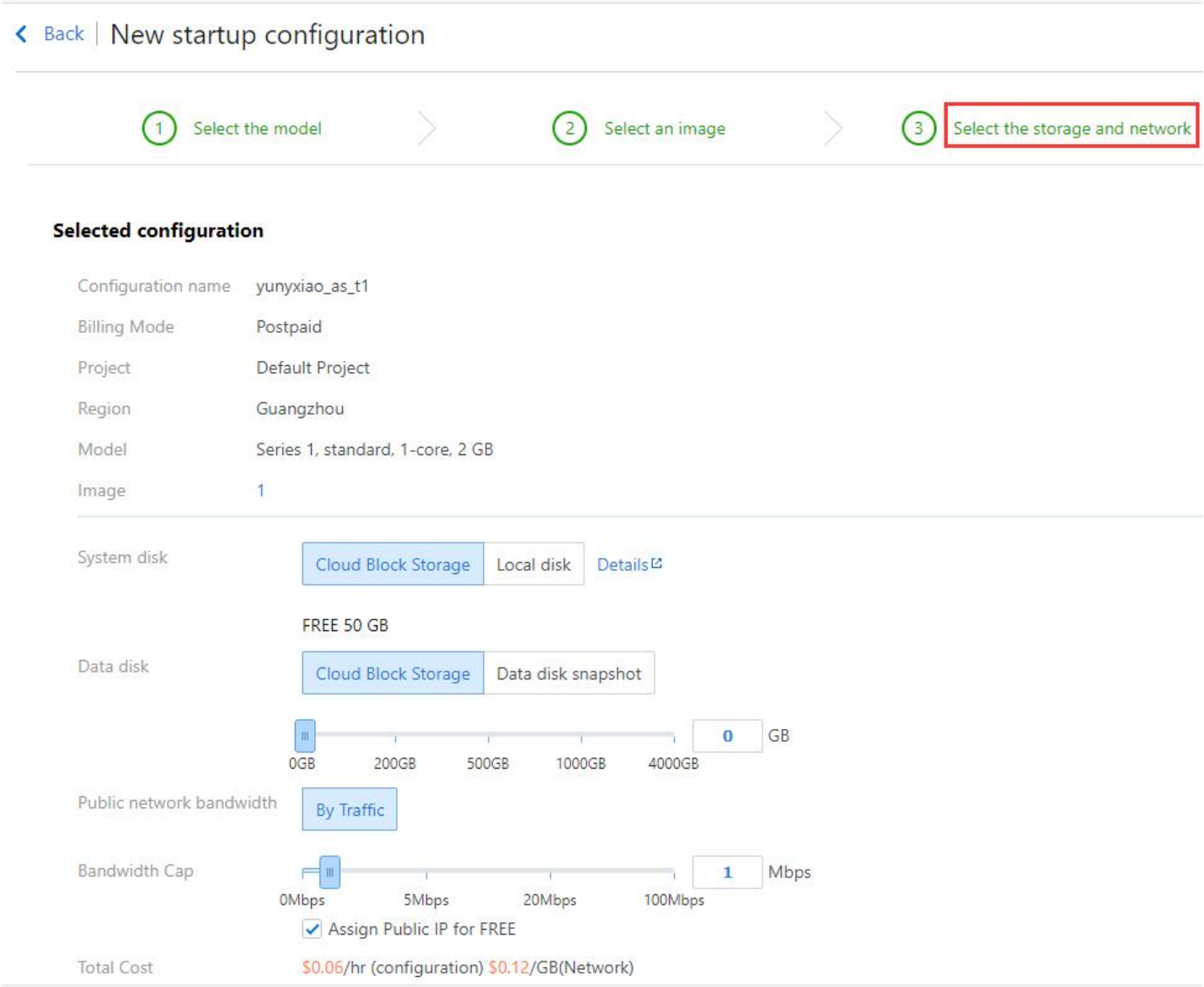
How to create the image for the "CVM to which the scaling group is going to bind"?



Select a Storage and a Network

Set storage and network in this page.

©2013-2018 Tencent Cloud. All rights reserved.



Please note that if you select the cloud disk as the system disk, then you can select the data disk snapshot for the data disk.

For users with large amount of data, they often use data disks to store data. When data disk A creates a snapshot file, users can use it to quickly clone multiple disks, achieving a fast server deployment.

When the auto scaling automatically adds a new CVM instance, if the data disk snapshot is specified in scaling configuration, Tencent Cloud's cloud disk can allow the automatic mounting of the data disk containing the set data after the CVM instance is activated, so as to meet the needs of automatic data copy.

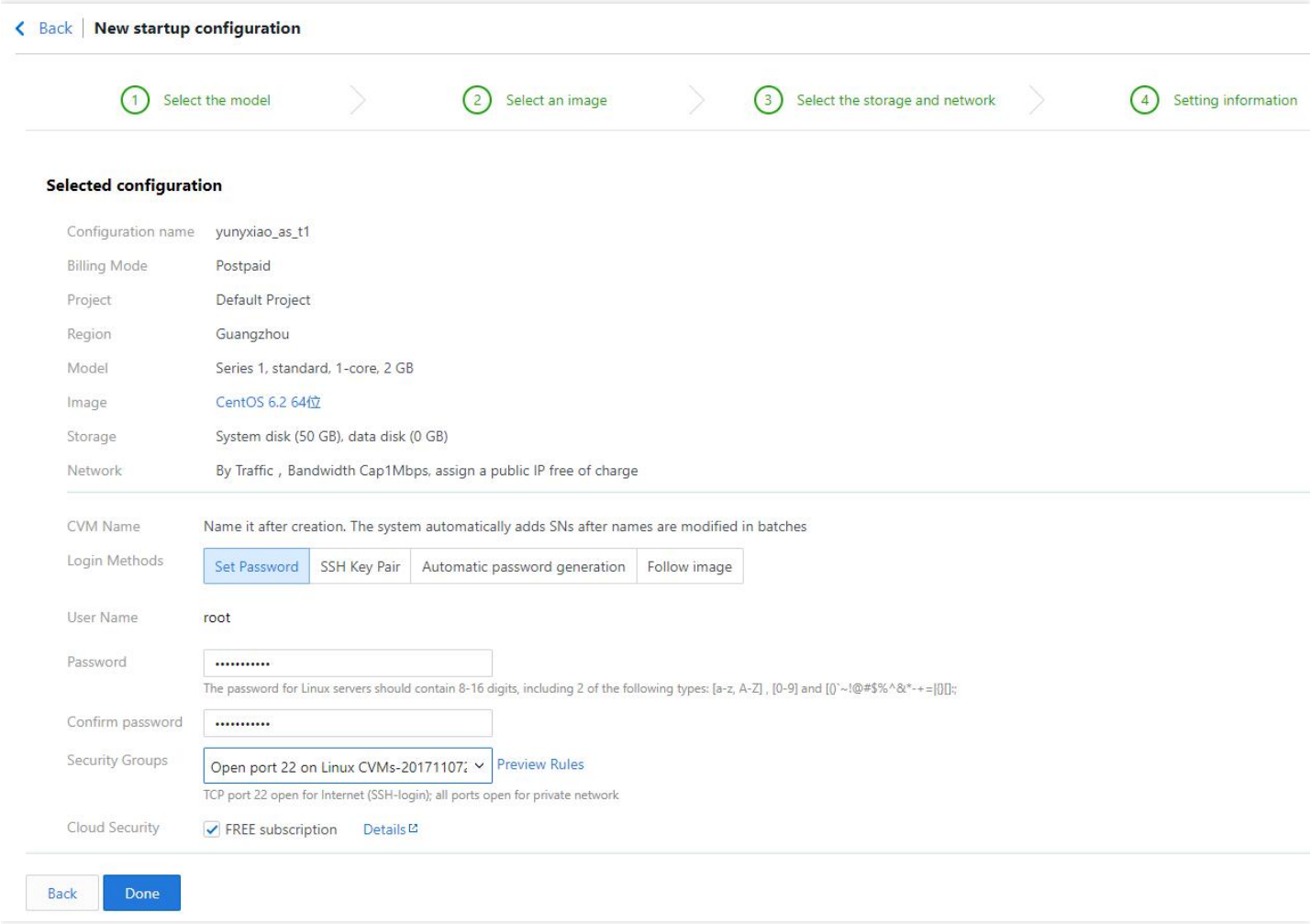
If the data disk snapshot is specified in the scaling configuration, you need to ensure that the data disk can be mounted automatically and correctly for the successful automatic scale-up of the scaling group. You need to perform some operations on the original instance of the data disk snapshot before setting

©2013-2018 Tencent Cloud. All rights reserved.

the auto scaling, so as to realize the automatic mounting of data disk when activating a new CVM instance. For instructions on how to do this, refer to How to Mount Data Disk Automatically When Activating New Instance Using Custom Image and Data Disk Snapshot

> Note:
>
> - Auto Scaling service is free of charge, and newly added servers, hard disks and networks will be charged by the traffic of CVM instances, hard disks and networks. This page will display prices based on your settings.

Set Information



Select the login method and security group in the **Set Information** page. The CVM instances added via the Auto Scaling service use Cloud Security and Cloud Monitor services for free by default.

After configuration, this entry will be displayed in the scaling configuration list, as shown in the figure below:

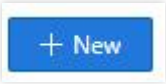
©2013-2018 Tencent Cloud. All rights reserved.

ID/Name	Bound scaling ...	CPU/MEM	Bandwidth/network billing mode	System disk/Data disk	Validity	Image	Operation
yunyxiao_as_t1	0	1-core 2GB	1Mbps Bill by traffic	Cloud Block Storage 50GB Cloud Block Storage 0GB	Valid	img-6fltx6zl	Delete

Step 2: Create a Scaling Group

A scaling group is a collection of CVM instances following the same rules and serving the same scenario. Log in to Auto Scaling Console, and click **Scaling Group** in the navigation bar.

Create a Scaling Group



Click , and fill in the basic information on scaling group in the pop-up page. Items with are required.

- The current number of CVM instances in the scaling group will be kept between the minimum and the maximum scaling group sizes.
  - The initial number of instances defines the number of CVM instances in the scaling group when created;
  - If the current number of CVM instances is less than the minimum scaling group size, the Auto Scaling service will automatically add instances to make it equal to the minimum scaling group size;
  - If the current number of CVM instances is greater than the maximum scaling group size, the Auto Scaling service will automatically decrease instances to make it equal to the maximum scaling group size;
- You can select the existing scaling configuration, or create a new scaling configuration.
- Select a network, availability zone, and remove policy.
- (Optional) You can choose to associate with an existing cloud load balance policy or create a new load balancer.

©2013-2018 Tencent Cloud. All rights reserved.

©2013-2018 Tencent Cloud. All rights reserved.

New scaling group submit a ticket ×

1 Basic Configuration > 2 Load Balancer Configuration

Name * yunyxiao_sg_t1

The name contains no more than 25 characters, and only supports letters, numbers, underscores, and separators"-", and decimal point

Minimum scaling number * 0

The current minimum instance number is 0, which may cause no instance is supporting your business in the scaling group

Number of initial instances 0

Maximum scaling number * 1

Start-up configuration * yunyxiao_as_t1 New startup configuration

Supported Networks * Basic Network

Supported Availability Zones *

- [ ] Guangzhou Zone 1
- [x] Guangzhou Zone 2
- [ ] Guangzhou Zone 3

You can select multiple availability zones. In the case of auto-expansion, CVMs will be created in these zones randomly to implement cross-region disaster recovery.

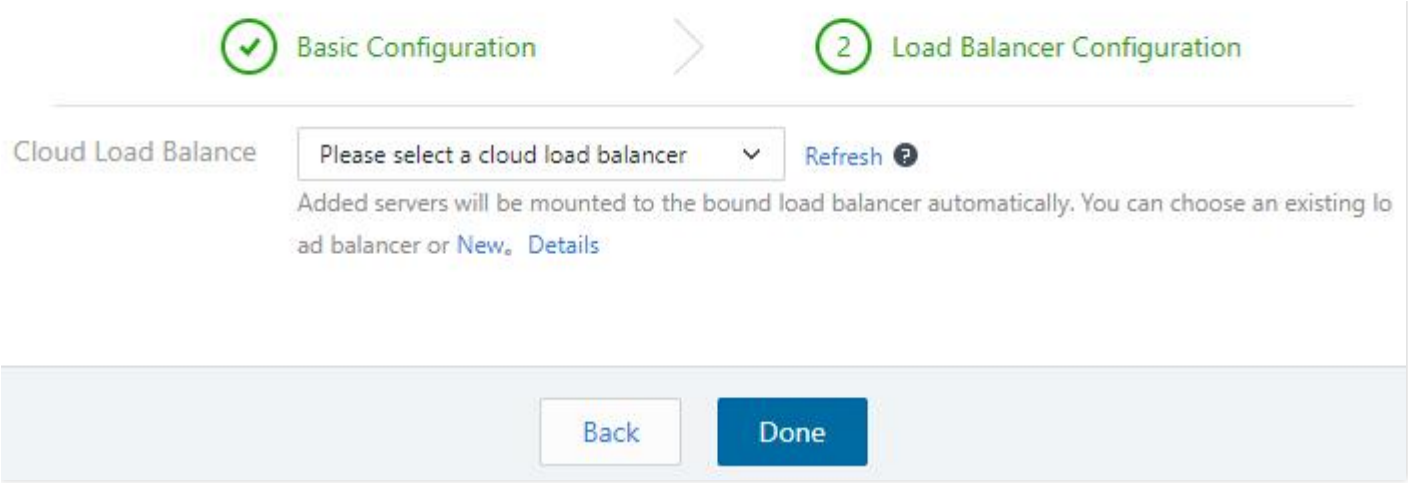
Suggested Settings

Removal policy * Remove the oldest CVM Learn more

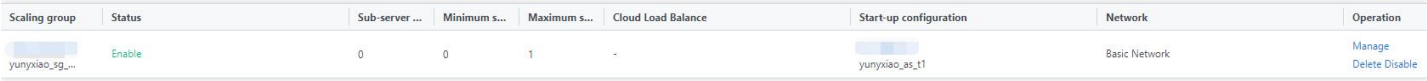
Cancel Next

New scaling group submit a ticket ×

©2013-2018 Tencent Cloud. All rights reserved.



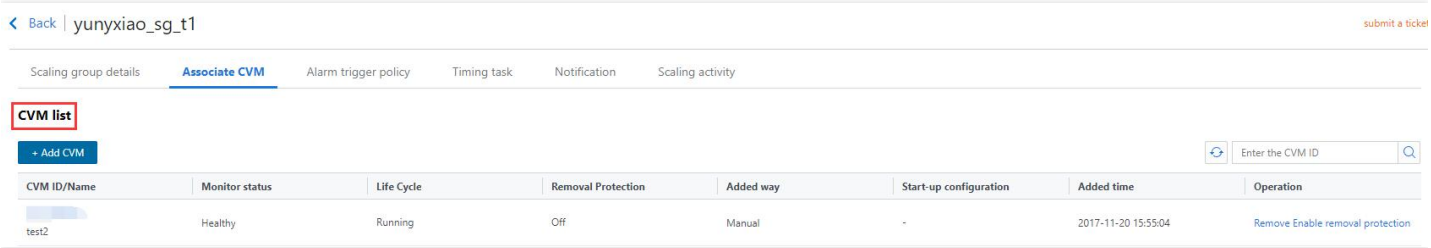
After configuration, this entry will be displayed in the scaling group list, as shown in the figure below:



Add a CVM (Optional)

Now, let's bind a CVM.

Add the CVM instance to be bound in the CVM list. After configuration, this entry will be displayed in the scaling configuration list, as shown in the figure below:



> Note: If you cannot add or remove a CVM, please check the maximum and minimum scaling group sizes you set.

Step 3: Create a Scaling Policy

The scaling group can adjust the number of CVMs based on the scaling policy:

- Create a **scheduled task** to perform scaling activities as scheduled, and you can also set to execute the scheduled task periodically;

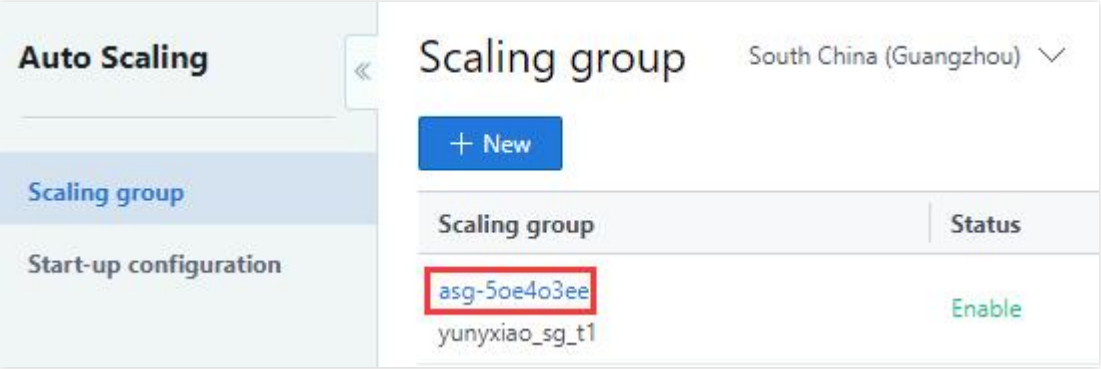
©2013-2018 Tencent Cloud. All rights reserved.

- Create an **alarm trigger policy** to perform scaling activities according to cloud monitoring indicators (such as CPU utilization and memory usage).

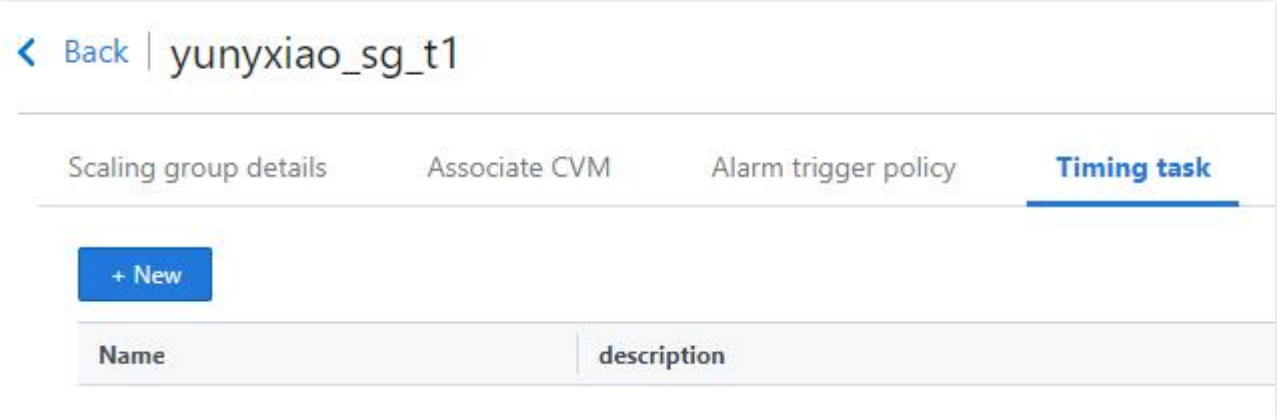
Create a Scheduled Task

If your load changes are predictable, you can set a scheduled task to plan your scaling activities. You can use this feature to automatically increase or decrease CVM instances on a scheduled and periodical basis to flexibly deal with traffic load changes and improve device utilization while saving deployment and instance costs.

In the **Scaling Group** page, click the scaling group ID to enter the scaling group management page.

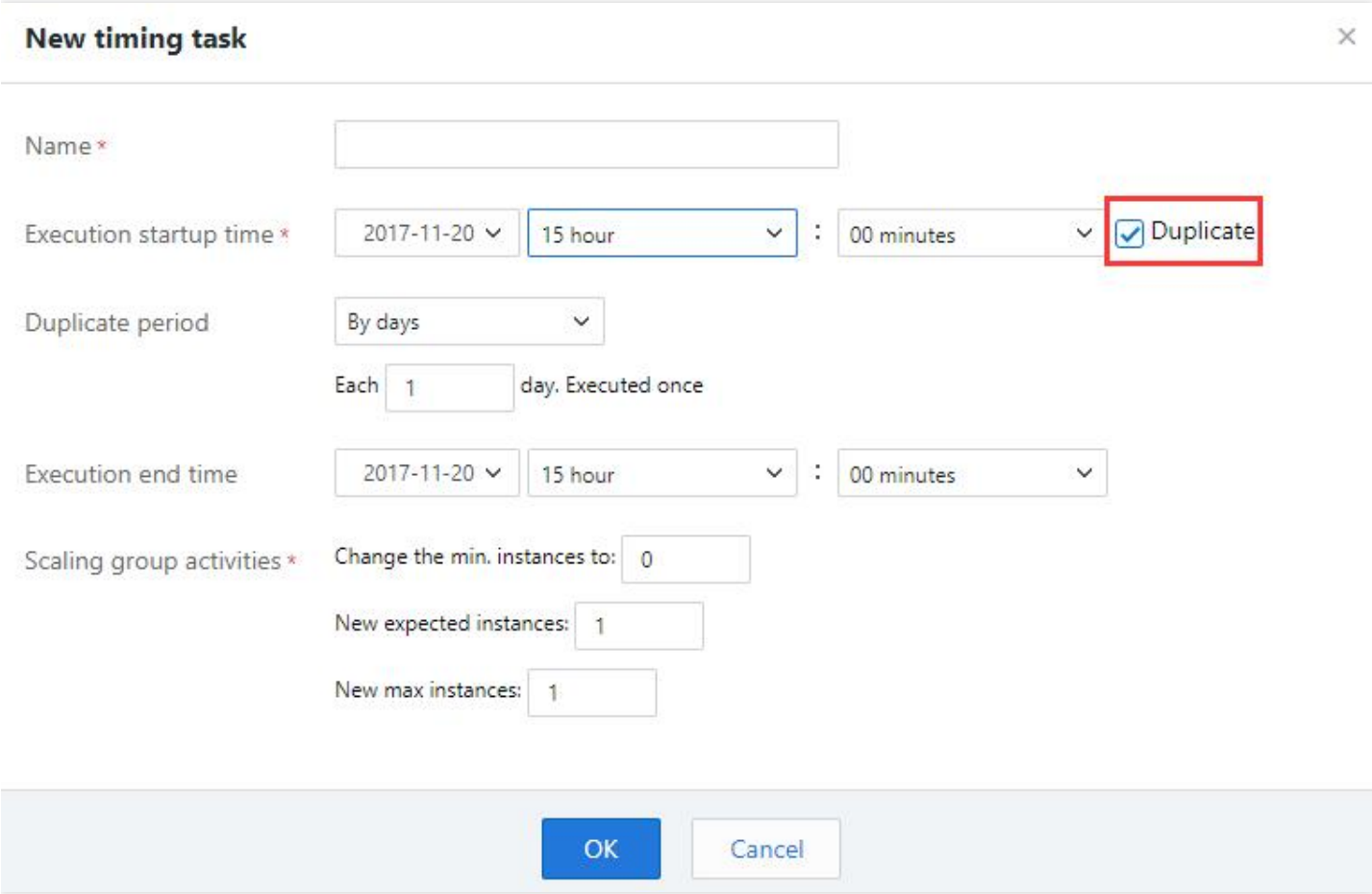


Select the **Timing Task** tab, and click **New**.

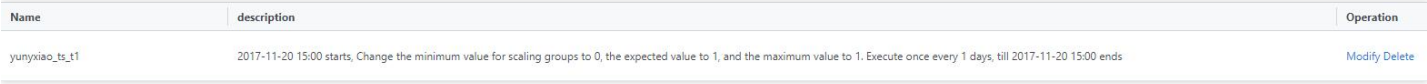


Specify information such as the scheduled task name, execution time, and activities to be executed in the New page. You can also check **Duplicate** to define the interval for the execution of a scheduled task.

©2013-2018 Tencent Cloud. All rights reserved.



After setting, the scheduled task will be displayed in the list on the page, as shown in the figure below:



Name	description	Operation
yunyxiao_ts_t1	2017-11-20 15:00 starts, Change the minimum value for scaling groups to 0, the expected value to 1, and the maximum value to 1. Execute once every 1 days, till 2017-11-20 15:00 ends	Modify Delete

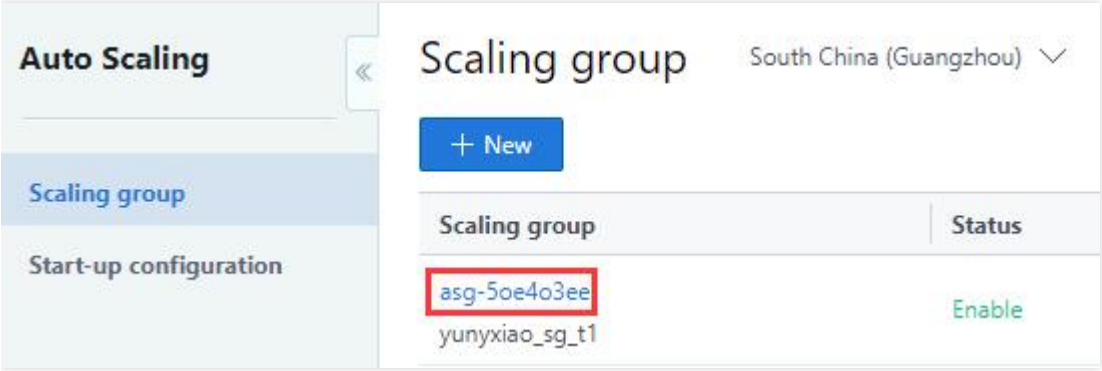
Create an Alarm Trigger Policy

If you wish to adjust business deployment based on CVM metrics, you can customize the alarm trigger policy, which will automatically increase or decrease the number of CVM instances when business load pushes the metrics to the threshold. This flexibly deals with traffic load changes, improves device utilization, and saves deployment and instance costs.

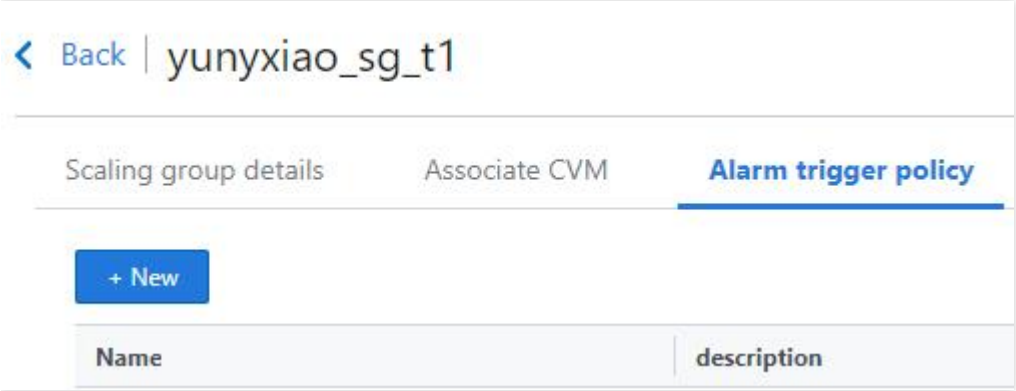
> - When a scaling group is created, a ping unreachable alarm trigger policy is created by default to replace the unhealthy machine.
> - Before using the alarm trigger policy, you need to install a new version of Cloud Monitor Agent in the CVM image. For the installation method, refer to Install Monitoring Components

©2013-2018 Tencent Cloud. All rights reserved.

In the **Scaling Group** page, click the scaling group ID to enter the scaling group management page.



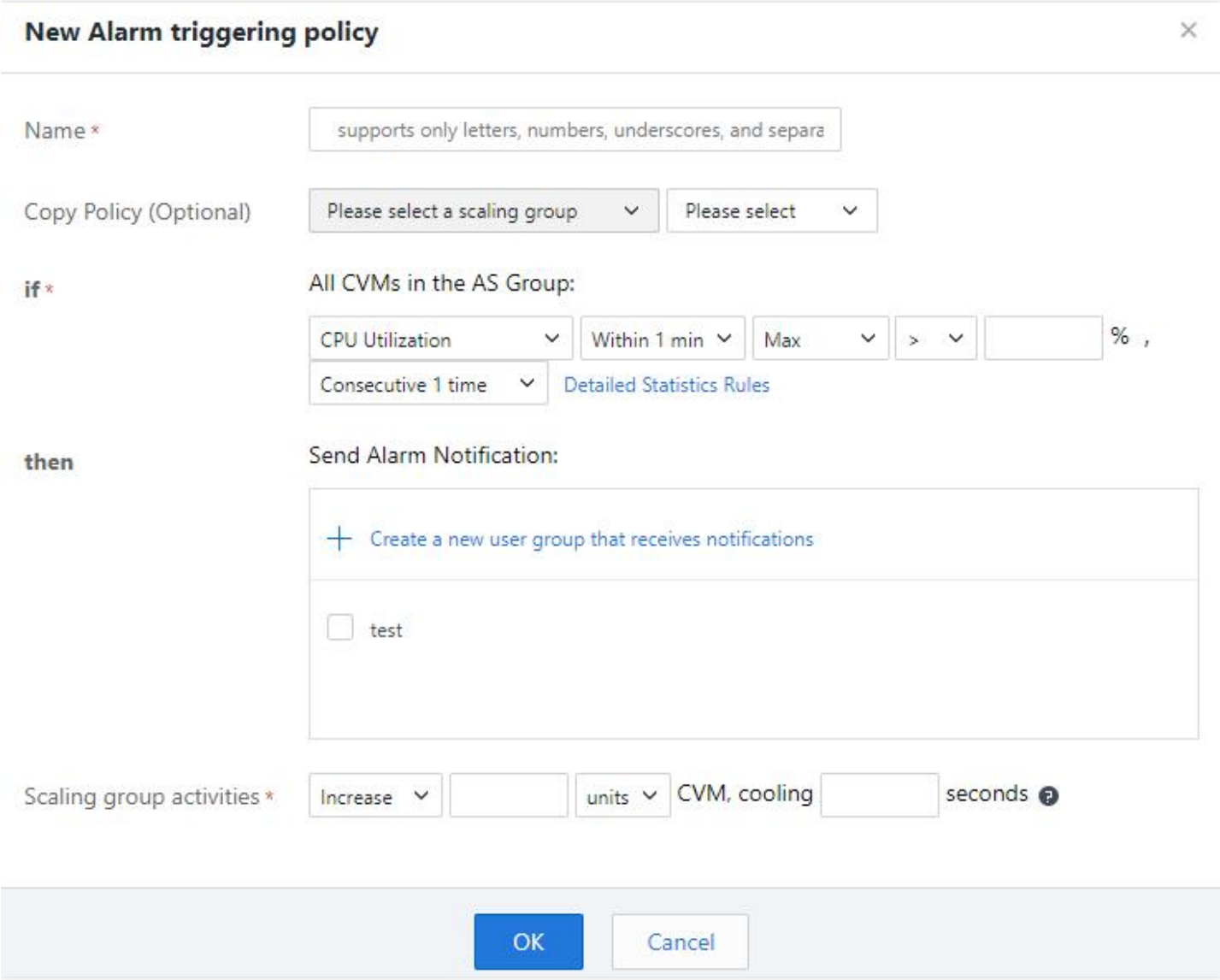
Select the **Alarm Trigger Policy** tab, and click **New**.



Set the alarm policy in the New page to automatically increase or decrease CVM instances by a specified number or percentage for the scaling group based on cloud monitoring performance metrics (such as CPU, memory, and bandwidth).

You can also copy the existing policy of an existing scaling group to the current scaling group by setting in the **Copy Policy (Optional)**.

©2013-2018 Tencent Cloud. All rights reserved.



After setting, the alarm trigger policy will be displayed in the list on the page, as shown in the figure below:

Name	description	Notification receiving group	Operation
yunyxiao_atp_t1	If CPU Utilization of CVM occurs 1 time(s) at Max greater than 70 %, within 1 minutes, the host Increase5 units. Cooled for 300 seconds.	-	Modify Delete

©2013-2018 Tencent Cloud. All rights reserved.

Step 1: Create a startup configuration

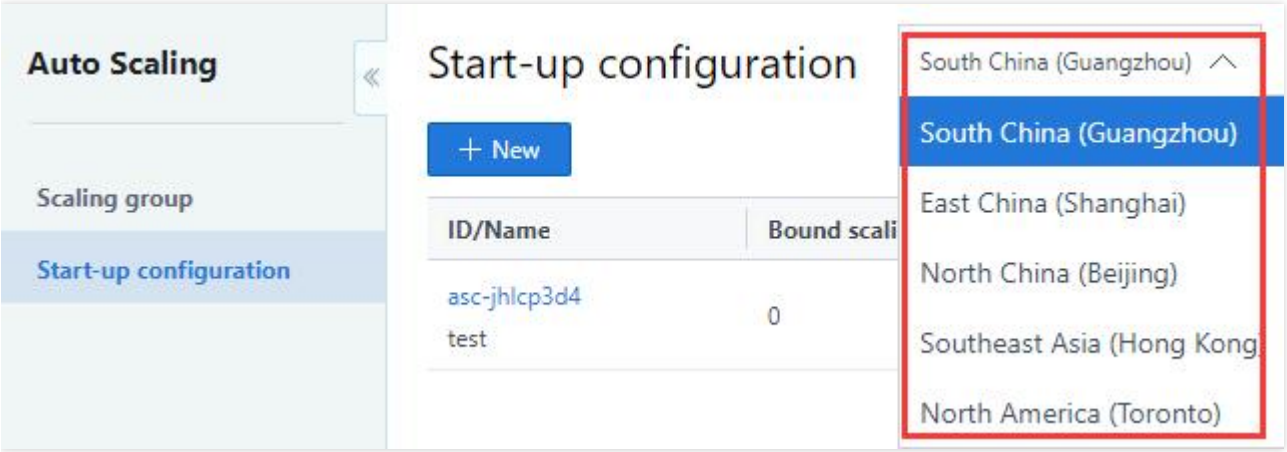
Last updated：2018-05-29 14:42:47

The scaling configuration defines the configuration information of CVM instances used for auto scaling, including the image, storage, network, security group, login method of the CVM.

> Note: The creation of a scaling configuration is free of charge.

Log in to Auto Scaling Console, and click **Scaling Configuration** in the navigation bar.

Select a Region



Please note that you must select the region where the CVM to which the scaling group needs to bind locates.

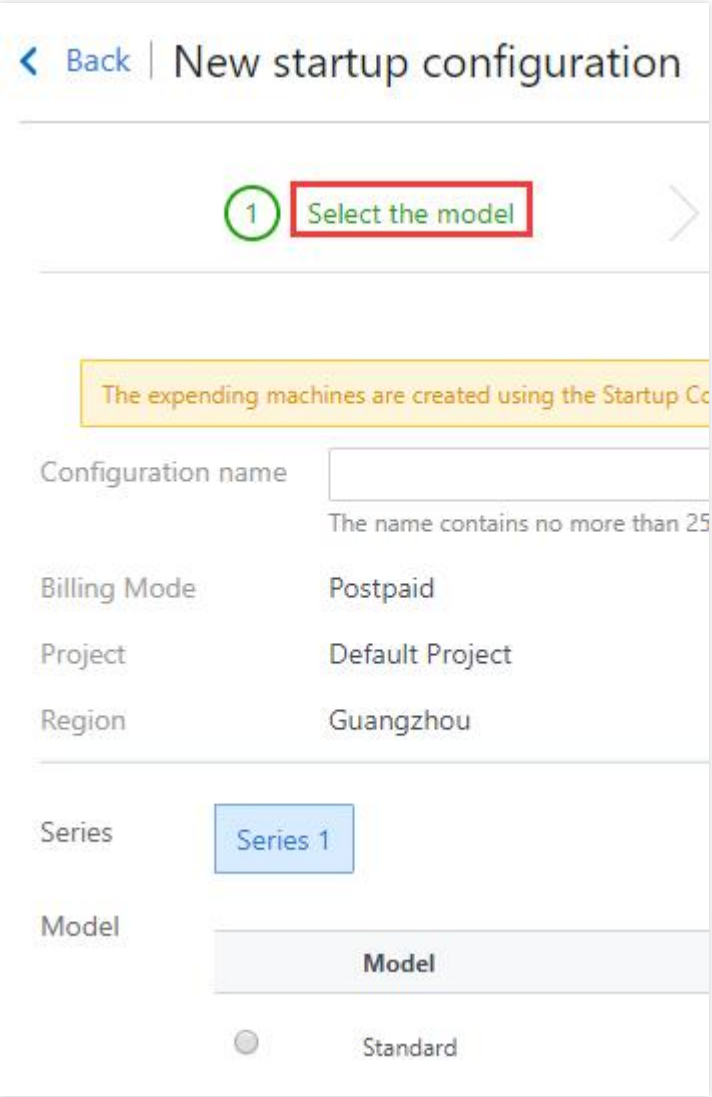
The CVMs which can be added manually and the cloud load balancers which can be bound are restricted by your selected region. For example, if you select Guangzhou as the region of the scaling configuration, the CVM of Guangzhou is automatically added to the scaling group. In a scaling group of Guangzhou region, you cannot add CVMs in other regions (Shanghai, Beijing, Hong Kong, Toronto, etc.) manually, nor bind load balancers in other regions (Shanghai, Beijing, Hong Kong, Toronto, etc.)

Select a Model

Click

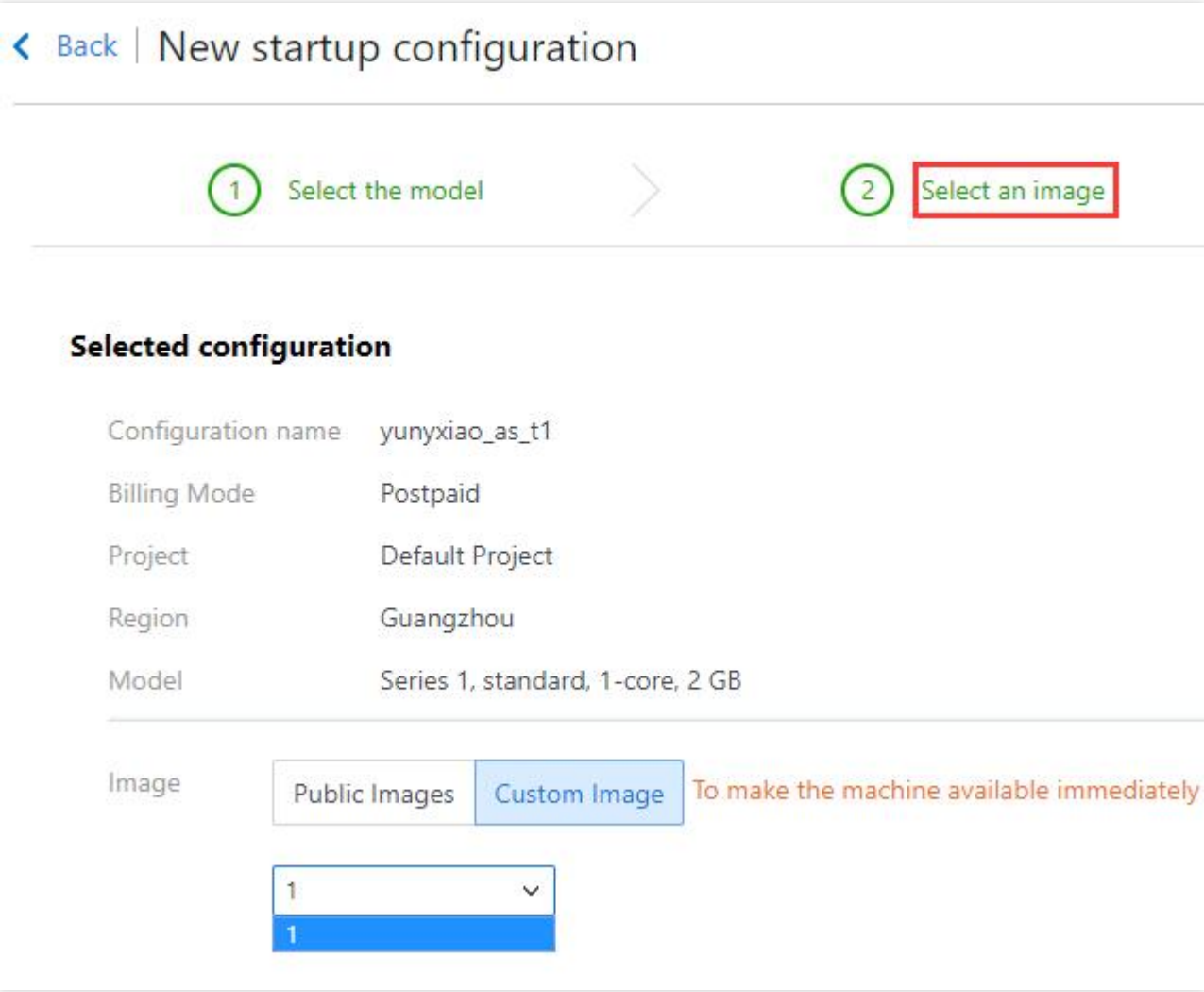
, and fill in the basic scaling configuration information in the pop-up page.

©2013-2018 Tencent Cloud. All rights reserved.



Select an Image

©2013-2018 Tencent Cloud. All rights reserved.



Public images and custom images are available.

If you select a public image, it should be consistent with the operating system of the CVM to which the scaling group is going to bind. If you select a custom image, you need to create the custom image through the CVM's image creation feature.

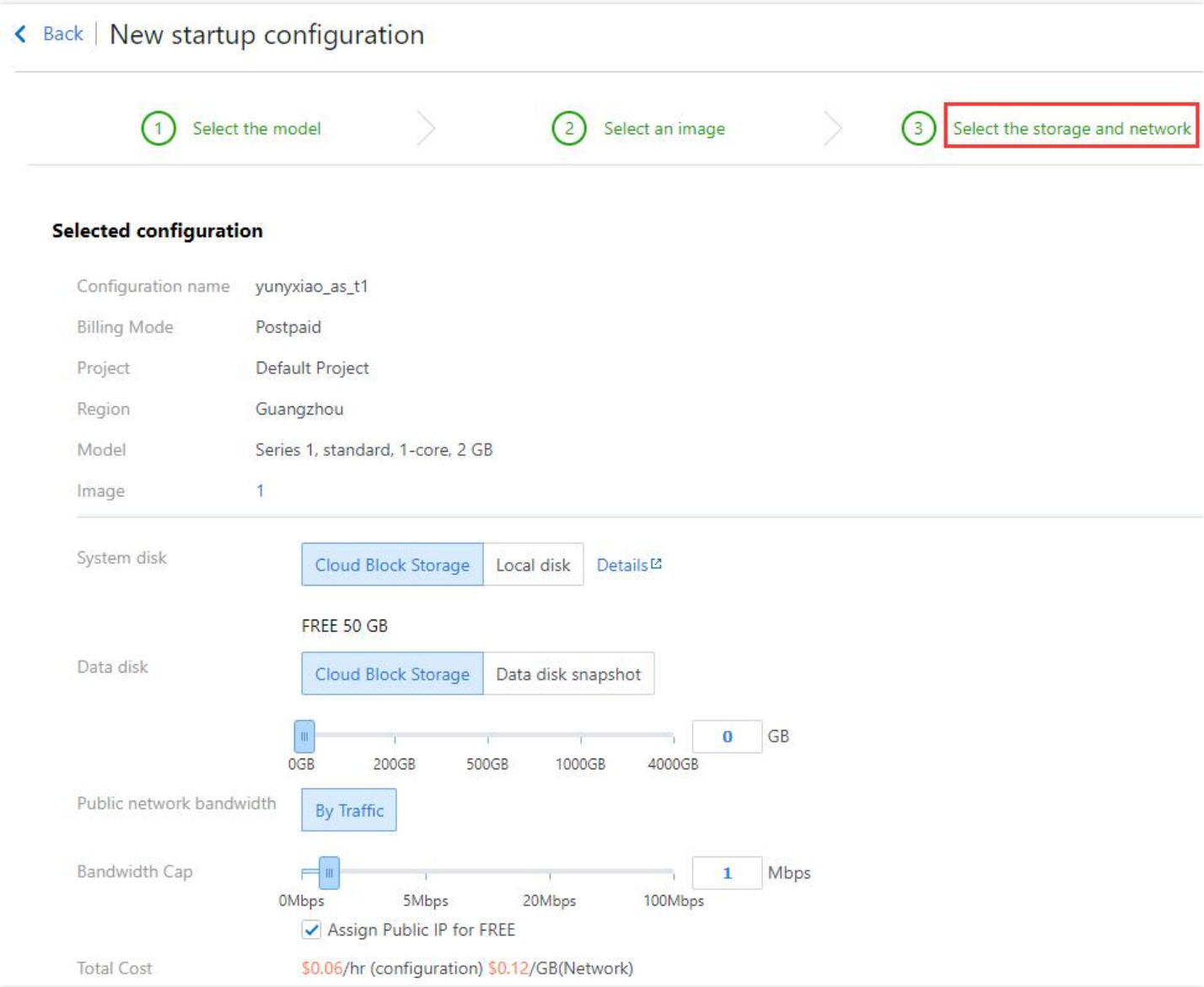
The difference is that, if you select a public image, the scaled CVM instance is not activated and cannot be used directly, and you need to manually deploy the application environment; if you select a custom image, by creating an image for the CVM instance of which the environment has been deployed, and using the image as the operating system when batch creating CVM instances, the CVM instances will have the same software environment as the original CVM instance after created successfully, and thus the batch deployment of software environments is achieved.

Therefore, it is recommended to select a custom image here. Bind the image of the "CVM to which the scaling group is going to bind".

> Note: How to create the image for the "CVM to which the scaling group is going to bind"?

©2013-2018 Tencent Cloud. All rights reserved.

Select a Storage and a Network



Set the disk and network in the **Select Storage and Network** page.

Please note that if you select the cloud disk as the system disk, then you can select the data disk snapshot for the data disk.

For users with large amount of data, they often use data disks to store data. When a data disk creates a snapshot file, users can use it to quickly clone multiple disks, achieving a fast server deployment.

When the auto scaling automatically adds a new CVM instance, if the data disk snapshot is specified in scaling configuration, Tencent Cloud's cloud disk can allow the automatic mounting of the data disk containing the set data after the CVM instance is activated, so as to meet the needs of automatic data copy.

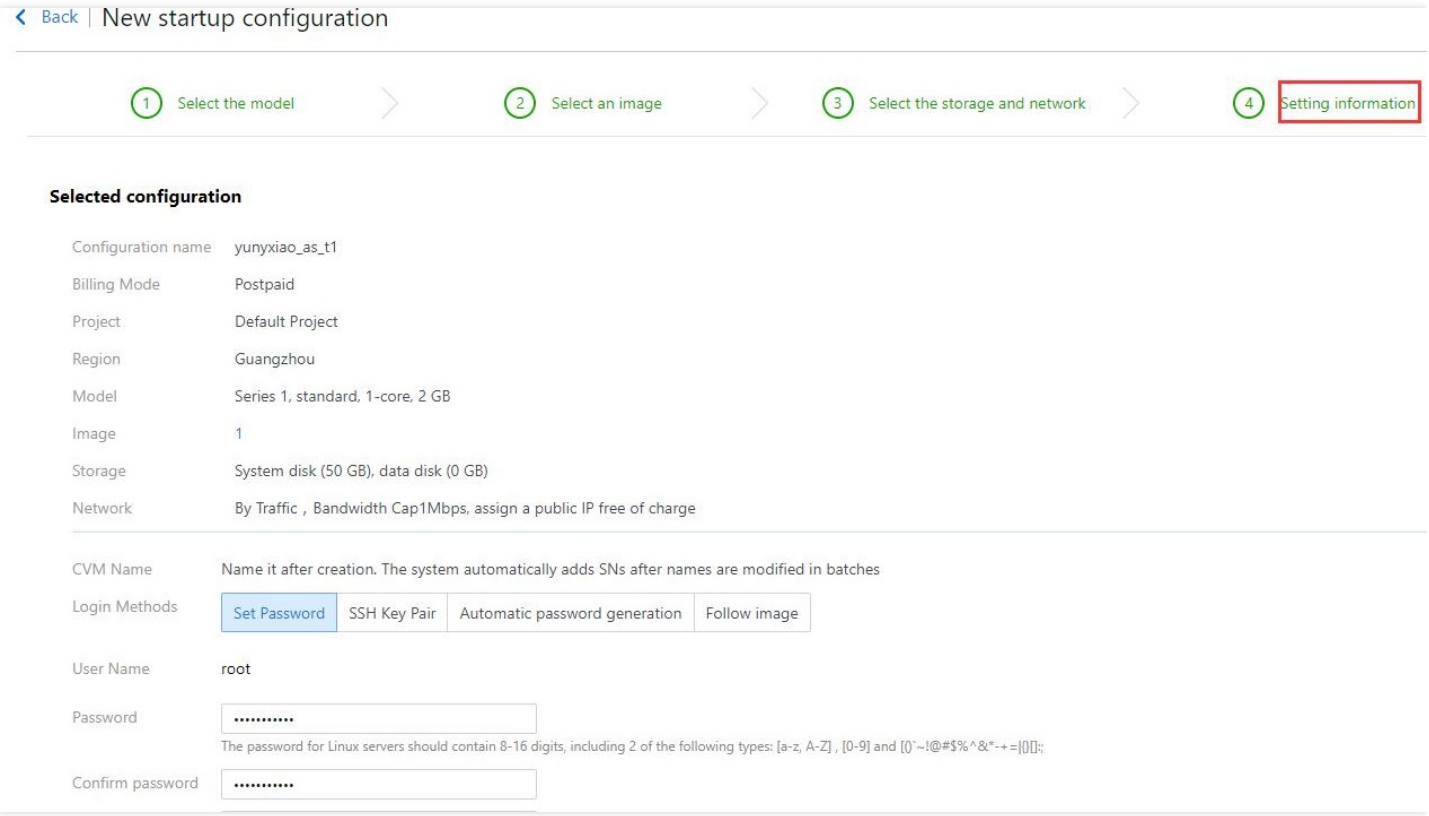
©2013-2018 Tencent Cloud. All rights reserved.

If the data disk snapshot is specified in the scaling configuration, you need to ensure that the data disk can be mounted automatically and correctly for the successful automatic scale-up of the scaling group. You need to perform some operations on the original instance of the data disk snapshot before setting the auto scaling, so as to realize the automatic mounting of data disk when activating a new CVM instance. For instructions on how to do this, refer to How to Mount Data Disk Automatically When Activating New Instance Using Custom Image and Data Disk Snapshot

> Note:
>
> - Auto Scaling service is free of charge, and newly added servers, hard disks and networks will be charged by the traffic of CVM instances, hard disks and networks. This page will display prices based on your settings.

Set Information

Select the login method and security group in the **Set Information** page.



> Note: The CVM instances added via the Auto Scaling service use Cloud Security and Cloud Monitor services for free by default.

©2013-2018 Tencent Cloud. All rights reserved.

After configuration, this entry will be displayed in the scaling configuration list, as shown in the figure below:

ID/Name	Bound scaling ...	CPU/MEM	Bandwidth/network billing mode	System disk/Data disk	Validity	Image	Operation
yunyxiao_as_t1	0	1-core 2GB	1Mbps Bill by traffic	Cloud Block Storage 50GB Cloud Block Storage 0GB	Valid	img-6fltx6zl	Delete

©2013-2018 Tencent Cloud. All rights reserved.

Step 2: Create a scaling group

Last updated：2018-05-29 14:44:03

A scaling group is a collection of CVM instances following the same rules and serving the same scenario.

Create a Scaling Group

Enter the Auto Scaling Console, and click **Scaling Group** in the navigation bar. Click

, and fill in the basic information on scaling group in the pop-up page. Items with are required.

- The current number of CVM instances in the scaling group will be kept between the minimum and the maximum scaling group sizes.
- The initial number of instances defines the number of CVM instances in the scaling group when created;
  - If the current number of CVM instances is less than the minimum scaling group size, the Auto Scaling service will automatically add instances to make it equal to the minimum scaling group size;
  - If the current number of CVM instances is greater than the maximum scaling group size, the Auto Scaling service will automatically decrease instances to make it equal to the maximum scaling group size;
- You can select the existing scaling configuration, or create a scaling configuration.
- Select a network, availability zone and Removal policy.
- **(Optional) You can choose to associate with an existing cloud load balance policy or create a load balancer.**

©2013-2018 Tencent Cloud. All rights reserved.

©2013-2018 Tencent Cloud. All rights reserved.

New scaling group submit a ticket ×

1 Basic Configuration > 2 Load Balancer Configuration

Name * yunyxiao_sg_t1

The name contains no more than 25 characters, and only supports letters, numbers, underscores, and separators"-", and decimal point

Minimum scaling number * 0

The current minimum instance number is 0, which may cause no instance is supporting your business in the scaling group

Number of initial instances 0

Maximum scaling number * 1

Start-up configuration * yunyxiao_as_t1 New startup configuration

Supported Networks * Basic Network

Supported Availability Zones *

- [ ] Guangzhou Zone 1
- [x] Guangzhou Zone 2
- [ ] Guangzhou Zone 3

You can select multiple availability zones. In the case of auto-expansion, CVMs will be created in these zones randomly to implement cross-region disaster recovery.

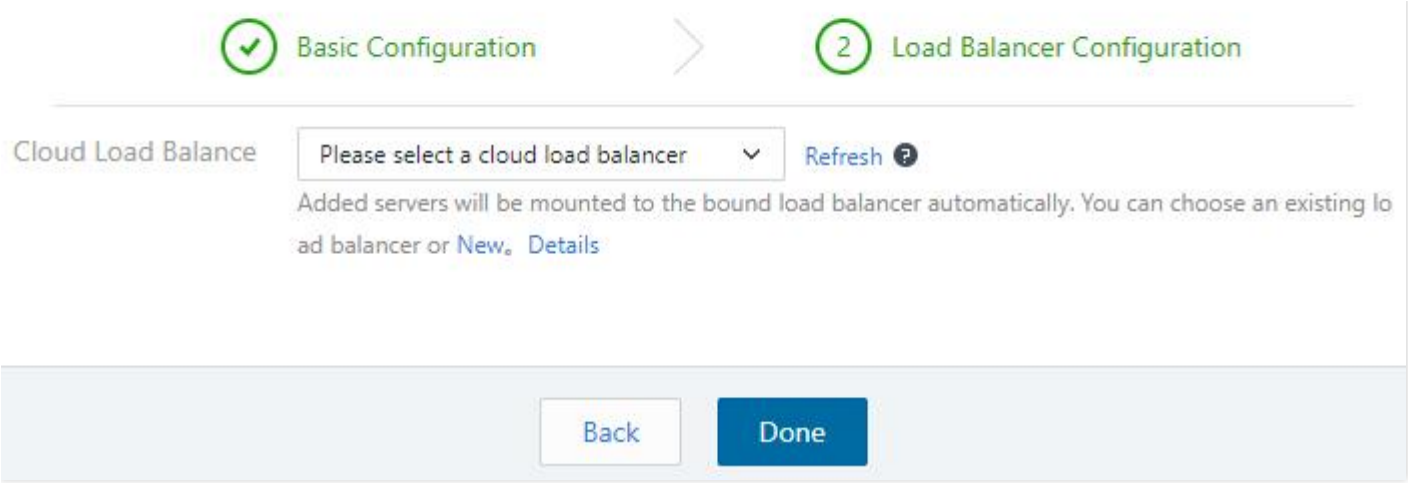
Suggested Settings

Removal policy * Remove the oldest CVM Learn more

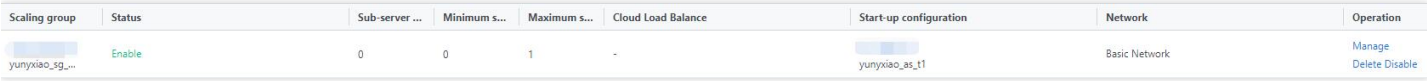
Cancel Next

New scaling group submit a ticket ×

©2013-2018 Tencent Cloud. All rights reserved.



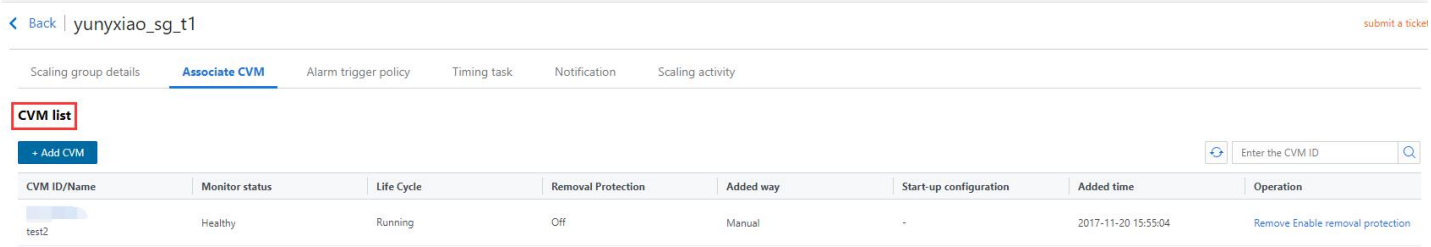
After configuration, this entry will be displayed in the scaling group list, as shown in the figure below:



Add a CVM (Optional)

Now, let's bind a CVM.

Add the CVM instance to be bound in the CVM list. After configuration, this entry will be displayed in the scaling configuration list, as shown in the figure below:



> Note: If you cannot add or remove a CVM, please check the maximum and minimum scaling group sizes you set.

©2013-2018 Tencent Cloud. All rights reserved.

Step 3: Create a scaling policy

Last updated：2018-05-29 14:53:31

The scaling group can adjust the number of CVMs based on the scaling policy:

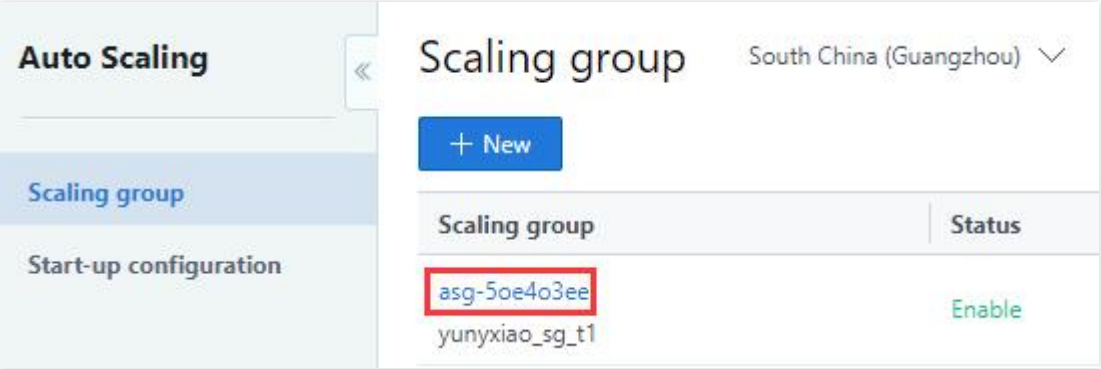
- Create a **scheduled task** to perform scaling activities as scheduled, and you can also set to execute the scheduled task periodically;
- Create an **alarm trigger policy** to perform scaling activities according to cloud monitoring indicators (such as CPU utilization and memory usage).

Create a Scheduled Task

If your load changes are predictable, you can set a scheduled task to plan your scaling activities. You can use this feature to automatically increase or decrease CVM instances on a scheduled and periodical basis to flexibly deal with traffic load changes and improve device utilization while saving deployment and instance costs.

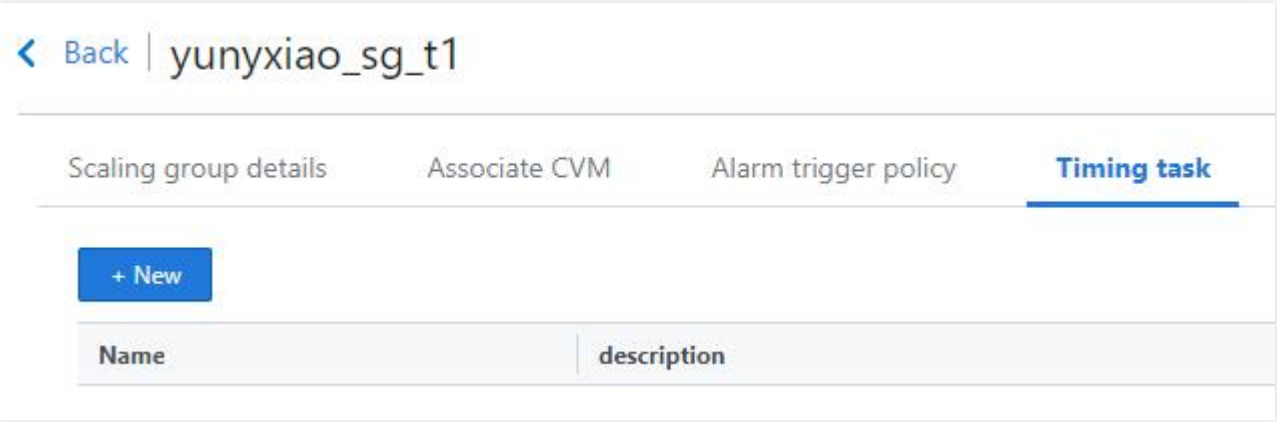
Steps:

In the **Scaling Group** page, click the scaling group ID to enter the scaling group management page.

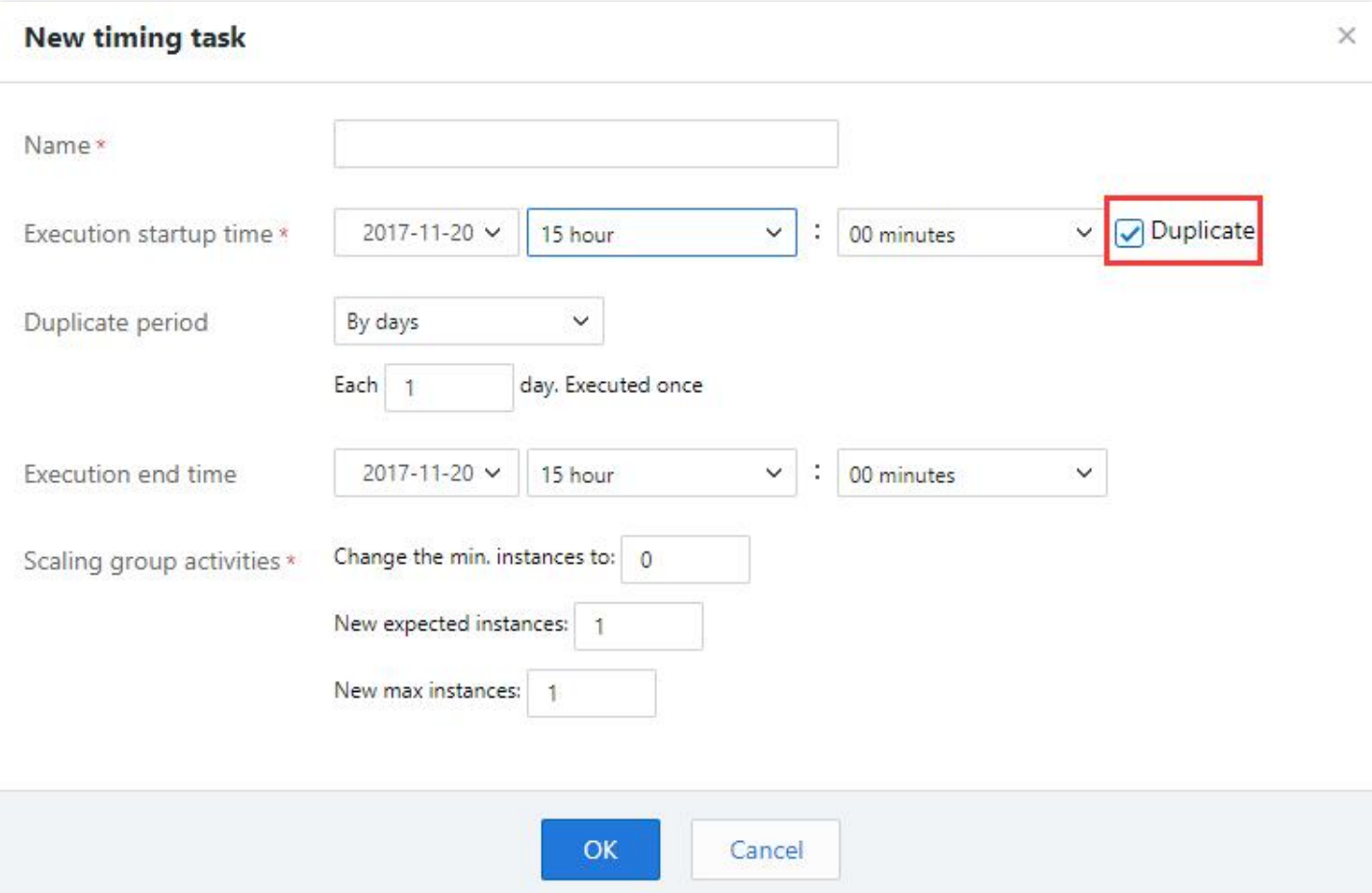


©2013-2018 Tencent Cloud. All rights reserved.

Select the **Timing Task** tab, and click **New**.



Specify information such as the scheduled task name, execution time, and activities to be executed in the New page. You can also check **Duplicate** to define the interval for the execution of a scheduled task.



©2013-2018 Tencent Cloud. All rights reserved.

After setting, the scheduled task will be displayed in the list on the page, as shown in the figure below:

Name	description	Operation
yunyxiao_ts_t1	2017-11-20 15:00 starts, Change the minimum value for scaling groups to 0, the expected value to 1, and the maximum value to 1. Execute once every 1 days, till 2017-11-20 15:00 ends	Modify Delete

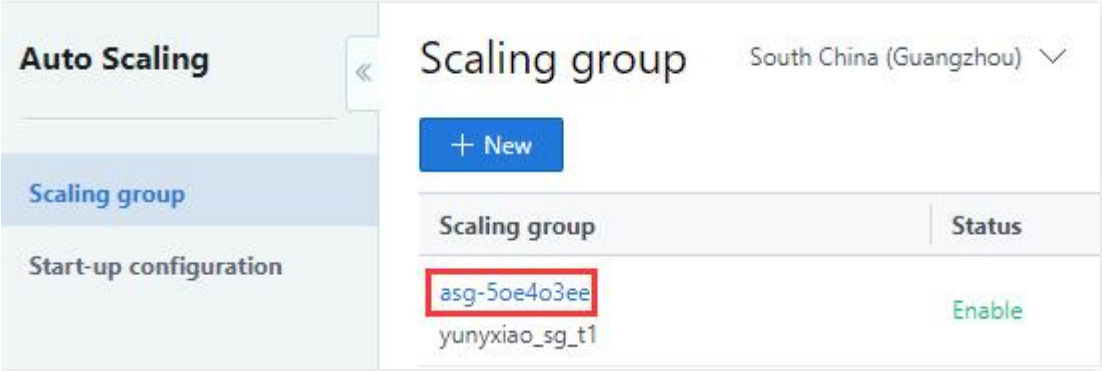
Create an Alarm Trigger Policy

If you wish to adjust business deployment based on CVM metrics, you can customize the alarm trigger policy, which will automatically increase or decrease the number of CVM instances when business load pushes the metrics to the threshold. This flexibly deals with traffic load changes, improves device utilization, and saves deployment and instance costs.

Steps:

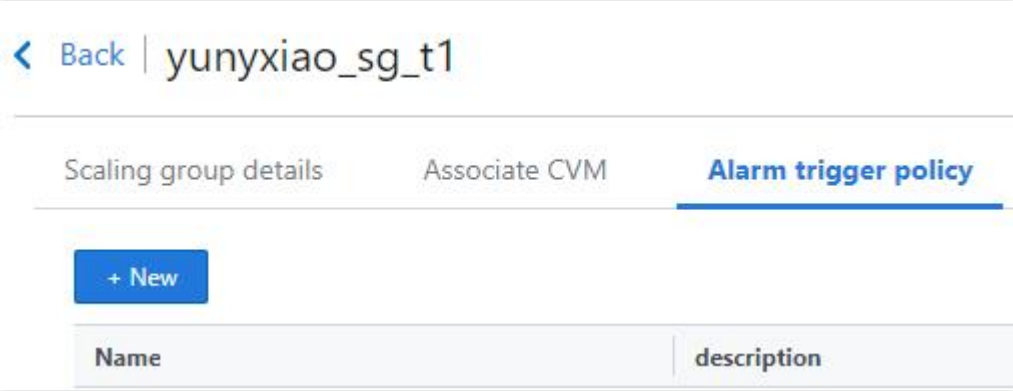
> - When a scaling group is created, a ping unreachable alarm trigger policy is created by default to replace the unhealthy sub-machine.
> - Before using the alarm trigger policy, you need to install a new version of Cloud Monitor Agent in the CVM image. For the installation method, refer to Install Monitoring Components

In the **Scaling Group** page, click the scaling group ID to enter the scaling group management page.



©2013-2018 Tencent Cloud. All rights reserved.

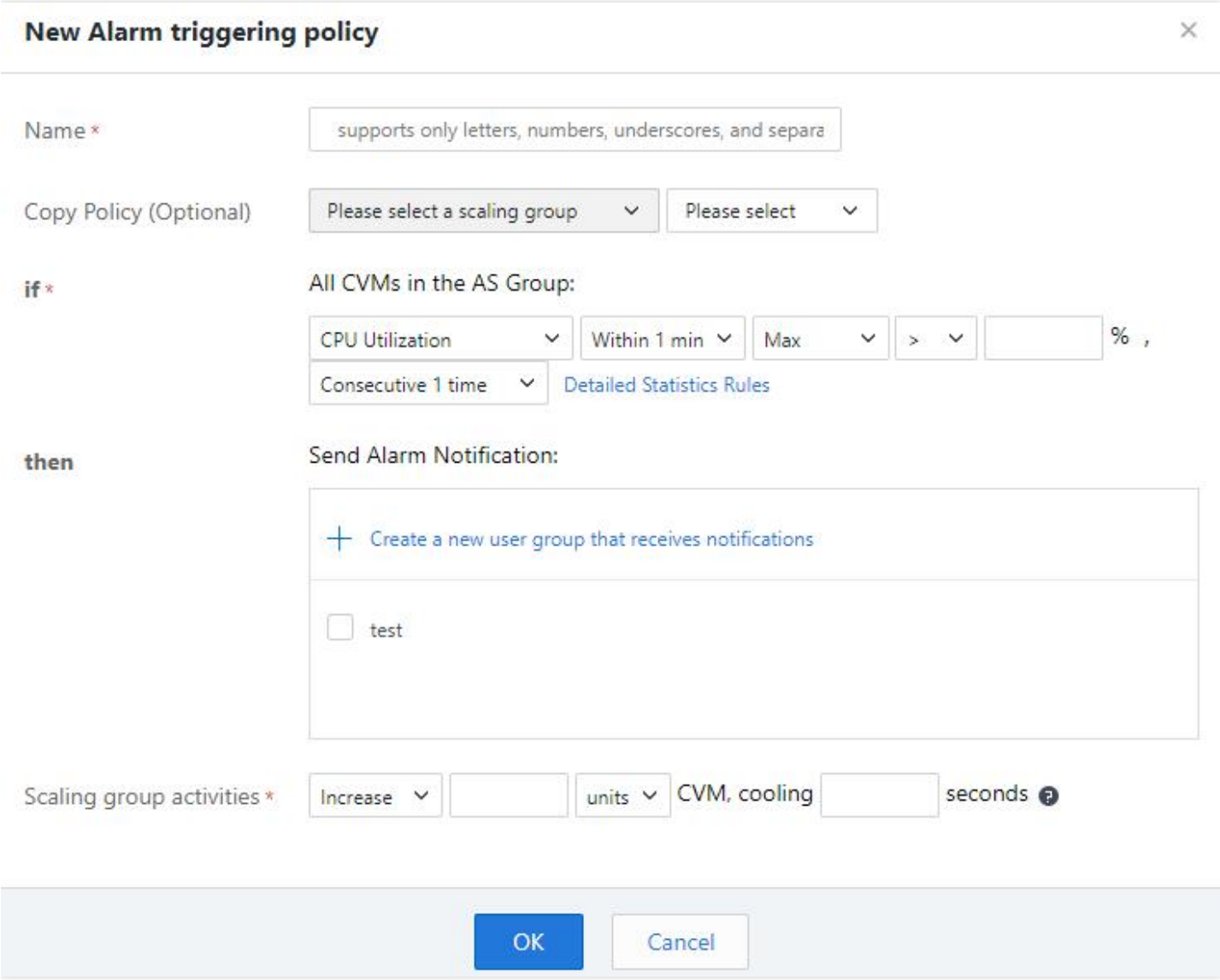
Select the **Alarm Trigger Policy** tab, and click **New**.



Set the alarm policy in the New page to automatically increase or decrease CVM instances by a specified number or percentage for the scaling group based on cloud monitoring performance metrics (such as CPU, memory, and bandwidth).

You can also copy the existing policy of an existing scaling group to the current scaling group by setting in the **Copy Policy (Optional)**.

©2013-2018 Tencent Cloud. All rights reserved.



After setting, the alarm trigger policy will be displayed in the list on the page, as shown in the figure below:

Name	description	Notification receiving group	Operation
yunyxiao_atp_t1	If CPU Utilization of CVM occurs 1 time(s) at Max greater than 70 %, within 1 minutes, the host Increase5 units. Cooled for 300 seconds.	-	Modify Delete

©2013-2018 Tencent Cloud. All rights reserved.