

Auto Scaling

FAQ

Product Introduction



Tencent
Cloud

Copyright Notice

©2013-2017 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



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

Documentation Legal Notice	2
FAQ.....	4
Elastic Scaling Issues.....	4

FAQ

Elastic Scaling Issues

1. What is cooldown period?

Cooldown period refers to a period of time when the corresponding scaling group is locked after a scaling activity (adding or removing CVM instances) is completed. During this period, no scaling activities are performed with the scaling group. The allowed value range for the cooling period is

0-999999

(seconds).

2. Is there a cooldown period for manually added CVMs?

No. Manually added CVMs need not to experience a cooldown period.

3. Do I need to pay for Auto Scaling?

No. Auto Scaling is completely free of charge.

You only need to pay for the CVM instances, no matter they're automatically created via Auto Scaling or manually added.

4. How can I increase the quota of CVMs in the scaling group?

By default, each Tencent Cloud user can have up to 30 CVMs in each availability zone. If you want more than 30 CVMs, please submit a ticket to apply for more.

5. What kind of CVMs is suitable for Auto Scaling?

The applications deployed in the CVM instances in the scaling group need to be stateless and replaceable. Because instances in the scaling group are likely to be reclaimed in scaling down, CVM instances for auto scaling cannot save status information (such as sessions) and related data (such as

databases, logs) of the applications. If you need to save the status information in your applications, you can consider saving it to a standalone CVM beyond the scaling group.

6. I failed to add a CVM to the scaling group manually.

To add a CVM to a scaling group manually, please ensure that the CVM:

- is in the same project as the scaling group;
- is in the same region as the scaling group;
- resides in the same network environment (VPC or basic network) as the scaling group;
- is in the running status.

7. I failed to bind a CLB instance with a scaling group.

To bind a CLB instance with a scaling group, please make sure that the CLB is in the same network environment (VPC or the basic network of the same region) as the scaling group.

8. Can I use Auto Scaling to automatically scale up and down CVM configuration?

No. AS can only increase or decrease the number of CVM instances according to the business requirement. However, Auto Scaling cannot change configuration of CVMs, like CPU, memory and bandwidth.

9. While scaling CVMs, are CLB and Cloud Monitor instances scaled at the same time?

It's up to you. You can set to scaling CVMs alone or scaling CVMs and bound CLBs at the same time.

10. How can I scale up a number of CVMs during a specified period of time?

You can set a pair of scheduled tasks: one defines the scale-up operation, setting the desired number of instances to the number you want to scale up to; the other defines the scale-down operation, setting the desired number of instances to the number you want to scale down to.

See [Best Practice Examples >>](#)

11. What are the specific rules for scaling group removal?

Tencent Cloud AS provides two removal policies:

- Delete the oldest: delete the oldest auto-added machine. In case no more auto-added machines exists, the earliest manual-added machine is deleted.
- Delete the latest: delete the latest auto-added machine. In case no more auto-added machines exists, the latest manual-added machine is deleted.

12. How does the alarm policy calculate the cloud monitoring information?

Take the maximum value for example. The basic rule is to take a value per minute (one value for a minute) with each passing period for the monitoring items set for each CVM. When the values taken for multiple consecutive periods (the number of periods can be customized) meet the set rules, the alarm scaling activity will be triggered.

For example, there are five CVMs in a scaling group, and the defined alarm scaling policy is "CPU utilization is more than 50% for 3 consecutive periods". The system takes a value per minute for each CVM, i.e. taking 25 CPU utilization values within one period (5 minutes). If any of the 25 values exceeds the threshold (50%), the period meets the alarm scaling rule. If the rule is met for 3 consecutive periods, the scaling activity will be triggered.

13. What is the "Expected Instance Number"?

The Expected Instance Number refers to the proper number of instances in a scaling group. It should be set between the minimum and maximum group sizes. You can manually adjust the Expected Instance Number, or configure scheduled task and alarm scaling task to trigger the adjustment. The

scaling group automatically adjusts the number of instances to match Expected Instance Number.

- Initial instance number: If an initial number of instances is set when a scaling group is created, this number is taken as the Expected Instance Number.
- Alarm scaling task adjustment: When the alarm scaling is triggered, the scaling group will adjust the current number of instances to the Expected Instance Number. For example, if the trigger action is to add two CVMs, the backend will satisfy the Expected Instance Number by adding two CVMs to the current number. If the system finds that the current number of instances is not equal to the Expected Instance Number for the scaling group, it will add two CVMs to make them equal.
- Scheduled or manual adjustment of the Expected Instance Number: When the user changes the Expected Instance Number by means of a scheduled task or a direct modification, the system, after discovering that the current number of instances is not equal to the Expected Instance Number, will trigger scaling activity until the two numbers are equal.
- System adjustment: The Expected Instance Number is between the maximum and minimum numbers of instances. The change in maximum or minimum number of instances may result in the change in the Expected Instance Number. For example, if the Expected Instance Number is 3, the minimum group size is 2, and the maximum group size is 5, in case the minimum group size is adjusted to 4, then the Expected Instance Number will be adjusted to 4 to meet the minimum group size.

14. I've set data disk snapshot in my startup configuration, but the disks are not mounted properly on added CVMs.

If a data disk snapshot is specified in the startup configuration, please follow the guide below to configure your original data disks:

See [here >>](#)

15. What happens when an AS group is disabled?

When the scaling group is disabled, the auto capacity scaling of the scaling group will not be triggered, but the restrictions on the scaling group remain in effect.

When the scaling group is disabled, the automatic activities will not performed:

- Alarm scaling;
- Scheduled task;
- Health check;
- The desired number of instances does not match due to manual operation.

But in order to ensure normal running, the basic restrictions on the scaling group still apply:

- If the instances are manually removed to be less than

min

, the removal will not be allowed;

- If the instances are manually added to greater than

max

, the addition will not be allowed;

- Manually raising

min

or

max

will not trigger scaling activity.

16. What stages are involved in the lifecycle of the CVMs that are automatically added to the scaling group?

Creating (Creating) - Indicates that a CVM instance is being created.

Running (InService) - Indicates that an instance is running.

Removing (Removing) - Indicates that an instance is being removed from the scaling group.

Binding (Attaching) - Indicates that an instance is being bound to the scaling group.

Unbinding (Detaching) - Indicates that an instance is being unbound from the scaling group.

Backuping (Backuping) - Indicates that an instance is being backed up.

UnBackuping (UnBackuping) - Indicates that the backup instance is being deleted.

Binding LB (AttachLb) - Indicates that a cloud load balance is being bound.

Unbinding LB (DetachLb) - Indicates that a cloud load balance is being unbound.

Prefetching (Preheating) - Indicates that the instance is being prefetched.

17. What happens when I remove a CVM from an AS group?

- Manually-added CVM: when a manually-added CVM is removed from an AS group, it's only removed from the AS group, and is not deleted. Its bound load balancer keeps unchanged
- CMS added by AS: when you remove an auto-added CVM from the AS group, the CVM is terminated and the load balancer is unbound