Alibaba Cloud Auto Scaling

API-Reference

Issue: 20190603

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

1. **You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.**
2. **No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company, or individual in any form or by any means without the prior written consent of Alibaba Cloud.**
3. **The content of this document may be changed due to product version upgrades , adjustments, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and the updated versions of this document will be occasionally released through Alibaba Cloud-authorized channels. You shall pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.**
4. **This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides the document in the context that Alibaba Cloud products and services are provided on an "as is", "with all faults " and "as available" basis. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity , applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not bear any liability for any errors or financial losses incurred by any organizations, companies, or individuals arising from their download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, bear responsibility for any indirect, consequential, exemplary, incidental, special, or punitive damages, including lost profits arising from the use**

or trust in this document, even if Alibaba Cloud has been notified of the possibility of such a loss.

5. By law, all the content of the Alibaba Cloud website, including but not limited to works, products, images, archives, information, materials, website architecture, website graphic layout, and webpage design, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of the Alibaba Cloud website, product programs, or content shall be used, modified , reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates . The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates).
6. Please contact Alibaba Cloud directly if you discover any errors in this document.

Generic conventions

Table -1: Style conventions

Style	Description	Example
	This warning information indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	 Danger: Resetting will result in the loss of user configuration data.
	This warning information indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	 Warning: Restarting will cause business interruption. About 10 minutes are required to restore business.
	This indicates warning information, supplementary instructions, and other content that the user must understand.	 Notice: Take the necessary precautions to save exported data containing sensitive information.
	This indicates supplemental instructions, best practices, tips, and other content that is good to know for the user.	 Note: You can use Ctrl + A to select all files.
>	Multi-level menu cascade.	Settings > Network > Set network type
Bold	It is used for buttons, menus , page names, and other UI elements.	Click OK.
`Courier font`	It is used for commands.	Run the `cd / d C :/ windows` command to enter the Windows system folder.
Italics	It is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	It indicates that it is a optional value, and only one item can be selected.	`ipconfig` *[-all|-t]*

Style	Description	Example
{} or {a|b}	It indicates that it is a required value, and only one item can be selected.	`swich {stand | slave}`

Contents

1 Scaling rule

1.1 DeleteScalingRule

This topic describes how to delete a scaling rule using the API.

Description

Deletes a specified scaling rule.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface, required. The parameter value is DeleteScalingRule.
ScalingRuleId	String	Yes	ID of a scaling rule.

Response parameters

Public parameters.

Error code

For common errors, see *client errors* or *server errors*.

Error message	Error code	Description	HTTP status code
The specified scaling rule does not exist in this account.	InvalidScalingRuleId.NotFound	The specified scaling rule does not exist.	404

Request example

```
 http :// ess . aliyuncs . com /? Action = DeleteScal  ingRule
& ScalingRul  eId = eMKWG8SRNb  9dBLAjweNI  1Ik
&< Public   Request   Parameters >
```

Response example

XML format:

```
< DeleteScal  ingRuleRes  ponse >
```

```
    < RequestId > F0595173 - CA3A - 4597 - B0D3 - 97A07A042B  9C </
 RequestId >
</ DeleteScal  ingRuleRes  ponse >
```

JSON format:

```
" RequestId ": " F0595173 - CA3A - 4597 - B0D3 - 97A07A042B  9C "
```

1.2 CreateScalingRule

This topic describes how to create a scaling rule using the API.

Description

A scaling rule defines specific scaling actions, for example, adding or removing n ECS instances.

If the execution of a scaling rule results in a number of ECS instances in the scaling group that is less than the MinSize or greater than the MaxSize, Auto Scaling automatically adjusts the number of ECS instances to be added or removed by executing the “Adjust scaling group instance quantity to MinSize” or “Adjust scaling group instance quantity to MaxSize” rule.

- Example 1: If a scaling group has a MaxSize of 3, the current number of instances (Total Capacity) is 2, and a scaling rule instructs the system to “add 3 ECS instances ”, this operation actually only adds one ECS instance. (The values in the scaling rule are not changed.)
- Example 2: If a scaling group has a MinSize of 2, the current number of instances (Total Capacity) is 3, and a scaling rule instructs the system to “remove 5 ECS instances”, this operation actually only removes one ECS instance. (The values in the scaling rule are not changed.)

This operation creates a scaling rule according to input parameters.

- When AdjustmentType is TotalCapaciy, the quantity of ECS instances in the current scaling group is adjusted to the specified value, and the corresponding Adjustment Value must be no less than 0.
- When AdjustmentType is QuantityChangeInCapacity or PercentChangeInCapac ity, if the corresponding AdjustmentValue is a positive number, additional ECS instances are increased; if the corresponding AdjustmentValue is a negative number, ECS instances are decreased.

- When AdjustmentType is PercentChangeInCapacity, the Auto Scaling service uses the formula: current number of instances (Total Capacity) * AdjustmentValue/100 and follow standard rounding principles to confirm the number of ECS instances to be added or removed.
- When a cool-down time (Cooldown) is specified in a scaling rule, the specified Cooldown is applied to the scaling group after the scaling activity to perform this rule is completed. Otherwise, the DefaultCooldown is applied to the scaling group.
- A maximum of 50 scaling rules can be created in a scaling group.
- The only identifier of a scaling rule returned (ScalingRuleAri) can be primarily used by the following interface:
  - By specifying the ExecuteScalingRule’s ScalingRuleAri parameter, you can manually perform a scaling rule.
  - By specifying the CreateScheduledTask’s ScheduledAction parameter, you can schedule the performing of a scaling rule.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface, required. The parameter value is CreateScalingRule.
ScalingGroupId	String	Yes	ID of the scaling group of a scaling rule.

Name	Type	Required	Description
AdjustmentType	String	Yes	Adjustment mode of a scaling rule. Optional values: · QuantityChangeInCapacity: It is used to increase or decrease a specified number of ECS instances. · PercentChangeInCapacity: It is used to increase or decrease a specified proportion of ECS instances. · TotalCapacity: It is used to adjust the quantity of ECS instances in the current scaling group to a specified value.

Name	Type	Required	Description
AdjustmentValue	Integer	Yes	Adjusted value of a scaling rule. The number of ECS instances that are increased or decreased at a time cannot exceed 500. Value range: · QuantityChangeInCapacity: [-500, 500] · PercentChangeInCapacity: [-100, 10000] · TotalCapacity: [0, 1000]
ScalingRuleName	String	No	Name shown for the scaling group, which is a string containing 2 to 40 English or Chinese characters. It must begin with a number, a letter (case-insensitive) or a Chinese character and can contain numbers, “_”, “-“ or “.”. The account name in the same scaling group is unique in the same region. If this parameter value is not specified, the default value is ScalingRuleId.

Name	Type	Required	Description
Cooldown	Integer	No	Cool-down time of a scaling rule. Value range: [0, 86 ,400], in seconds. The default value is empty.

Response parameters

Name	Type	Description
ScalingRuleId	String	ID of a scaling rule, generated by the system and globally unique.
ScalingRuleAri	String	Unique identifier of a scaling rule.

Examples

Request example

```
 http :// ess . aliyuncs . com /? Action = CreateScal  ingRule
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& Adjustment  Type = QuantityCh  angeInCapa  city
& Adjustment  Value =- 10
&< Public   Request   Parameters >
```

Response example

`XML` format

```
< CreateScal  ingRuleRes  ponse >
    < ScalingRul  eAri > ari : acs : ess : cn - qingdao : 1344371 :
 scalingrul  e / eMKWG8SRNb  9dBLAjweNI  1Ik </ ScalingRul  eAri >
    < ScalingRul  eId > eMKWG8SRNb  9dBLAjweNI  1Ik </ ScalingRul  eId
 >
    < RequestId > 570C84F4 - A434 - 488A - AFA1 - 1E3213682B  33 </
 RequestId >
</ CreateScal  ingRuleRes  ponse >
```

`JSON` format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 ",
    " ScalingRul  eId ": " eMKWG8SRNb  9dBLAjweNI  1Ik ",
    " ScalingRul  eAri ":" ari : acs : ess : cn - qingdao : 1344371 :
 scalingrul  e / eMKWG8SRNb  9dBLAjweNI  1Ik "
```

```
}
```

Error codes

For common errors, see *client errors* or *server errors*.

Error code	Error message	HttpCode	Description
InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	404	The specified scaling group does not exist in this account.
Duplicate.	The specified value of parameter <parameter name> is duplicated.	400	The scaling rule name already exists.
QuotaExceeded.ScalingRule	Scaling rule quota exceeded in the specified scaling group.	400	Scaling rule quantity exceeds the upper limit for a user to use.

1.3 ModifyScalingRule

This topic describes how to modify a scaling rule using the API.

Description

Modifies the attributes of a scaling rule.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface, required. The parameter value is ModifyScalingRule.
ScalingRuleId	String	Yes	ID of a scaling rule.

Name	Type	Required	Description
AdjustmentType	String	No	Adjustment mode of a scaling rule. Optional values: · QuantityChangeInCapacity: It is used to increase or decrease a specified number of ECS instances. · PercentChangeInCapacity: It is used to increase or decrease a specified proportion of ECS instances. · TotalCapacity: It is used to adjust the quantity of ECS instances in the current scaling group to a specified value.

Name	Type	Required	Description
AdjustmentValue	Integer	No	Adjusted value of a scaling rule. The number of ECS instances that are increased or decreased at a time cannot exceed 500. Value range: · QuantityChangeInCapacity: [-500, 500] · PercentChangeInCapacity: [-100, 10000] · TotalCapacity: [0, 1000]
ScalingRuleName	String	No	Name shown for the scaling group, which is a string containing 2 to 40 English or Chinese characters. It must begin with a number, a letter (case-insensitive) or a Chinese character and can contain numbers, “_ “, “- “ or “.” . The account name in the same scaling group is unique in the same region. If this parameter value is not specified, the default value is ScalingRuleId.

Name	Type	Required	Description
Cooldown	Integer	No	Cool-down time of a scaling rule. Value range: [0, 86 ,400], in seconds. The default value is empty.

Response parameters

Name	Type	Description
ScalingRuleId	String	ID of a scaling rule, generated by the system and globally unique.
ScalingRuleAri	String	Unique identifier of a scaling rule.

Examples

Request example

```
 Http :// ess . aliyuncs . com  /?  Action  =  fig
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& Adjustment  Type = QuantityCh  angeInCapa  city
& Adjustment  Value =- 10
&< Common   Request   Parameters >
```

Response example

`XML` format

```
< ModifyScal  ingRuleRes  ponse >
    < ScalingRul  eAri > ari : acs : ess : cn - qingdao : 1344371 :
 scalingrul  e / eMKWG8SRNb  9dBLAjweNI  1Ik </ ScalingRul  eAri >
    < ScalingRul  eId > eMKWG8SRNb  9dBLAjweNI  1Ik </ ScalingRul  eId
 >
    < RequestId > 570C84F4 - A434 - 488A - AFA1 - 1E3213682B  33 </
 RequestId >
</ ModifyScal  ingRuleRes  ponse >
```

`JSON` format

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 ",
    " ScalingRul  eId ": " eMKWG8SRNb  9dBLAjweNI  1Ik ",
    " ScalingRul  eAri ":" ari : acs : ess : cn - qingdao : 1344371 :
 scalingrul  e / eMKWG8SRNb  9dBLAjweNI  1Ik "
```

```
}
```

Error codes

For common errors, see *client errors* or *server errors*.

Error code	Error message	HttpCode	Description
InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	404	The specified scaling group does not exist in this account.
InvalidScalingRuleName.Duplicate	The specified value of parameter < parameter name > is duplicated.	400	The scaling rule name already exists.
QuotaExceeded.ScalingRule	Scaling rule quota exceeded in the specified scaling group.	400	The user's action quota has been reached.

1.4 DescribeScalingRules

Queries information of a scaling rule.

Description

You can query all scaling rules in a scaling group by specifying the scaling group ID.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface, required. The parameter value is DescribeScalingRules.
RegionId	String	Yes	Region ID of the scaling group of a scaling rule.
ScalingGroupId	String	No	Scaling group ID.

Name	Type	Required	Description
ScalingRuleId.N	String	No	ID of a scaling rule. A maximum of 10 values can be entered. Invalid scaling rule IDs are neglected in the query result and no error is reported.
ScalingRuleName.N	String	No	Name of a scaling rule. A maximum of 10 values can be entered. Invalid scaling rule names are neglected in the query result and no error is reported.
ScalingRuleAri.N	String	No	Unique identifier of a scaling rule. A maximum of 10 values can be entered. Invalid unique identifiers of scaling rules are neglected in the query result and no error is reported.
PageNumber	Integer	No	Page number of the scaling rule list. The initial value and default value are both 1.
PageSize	Integer	No	When querying by page, this parameter indicates the number of lines per page. Maximum value: 50; default value: 10.

Response parameters

Name	Type	Description
TotalCount	Integer	Total number of scaling rules
PageNumber	Integer	Current page number
PageSize	Integer	Number of lines per page
ScalingRules	ScalingRuleSetType	Scaling rule information set

ScalingRuleSetType is a set of ScalingRuleItemTypes.

Name	Type	Description
ScalingRule	ScalingRuleItemType	Scaling rule information

The attributes of ScalingRuleItemType are listed as follows.

Name	Type	Description
ScalingRuleId	String	ID of a scaling rule
ScalingGroupId	String	Scaling group ID
ScalingRuleName	String	Name of a scaling rule
Cooldown	Integer	Cool-down time
AdjustmentType	String	Adjustment mode
AdjustmentValue	Integer	Adjustment value
ScalingRuleAri	String	Unique identifier of a scaling rule

Error code

For common errors, see *client errors* or *server errors*.

Request example

```
 http :// ess . aliyuncs . com /? Action = DescribeSc  alingRules
& RegionId = cn - qingdao
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& PageSize = 50
```

```
&< Public   Request   Parameters >
```

Response example

XML format:

```
< DescribeSc  alingRules  Response >
    < PageNumber > 1 </ PageNumber >
    < PageSize > 50 </ PageSize >
    < ScalingRul  es >
        < ScalingRul  e >
            < Adjustment  Type > QuantityCh  angeInCapa  city </
 Adjustment  Type >
            < Adjustment  Value > 1 </ Adjustment  Value >
            < Cooldown > 20 </ Cooldown >
            < ScalingGro  upId > AG6CQdPU8O  KdwLjgZcJ2  eaQ </
 ScalingGro  upId >          < ScalingRul  eAri >
 ari : acs : ess : cn - qingdao : 1344371 : scalingRul  e /
 eMKWG8SRNb  9dBLAjweNI  1Ik
</ ScalingRul  eAri >
            < ScalingRul  eId > eMKWG8SRNb  9dBLAjweNI  1Ik </
 ScalingRul  eId >            < ScalingRul  eName > eMKWG8SRNb
 9dBLAjweNI  1Ik </ ScalingRul  eName >
        </ ScalingRul  e >
    </ ScalingRul  es >
    < TotalCount > 1 </ TotalCount >
    < RequestId > 3306A40D - 3412 - 4101 - 9F19 - 5F81E3055D  AD </
 RequestId >
</ DescribeSc  alingRules  Response >
```

JSON format:

```
{
  " RequestId ": " B583BFEF - A779 - 427A - 9B74 - 262DDD2497  02 ",
  " TotalCount ":  1 ,
  " PageNumber ":  1 ,
  " PageSize ":  10 ,
  " ScalingRul  es ": {
    " ScalingRul  e ": [
      {
        " ScalingRul  eId ": " efcqrZdjlo  okc0UkE3dA  5I0a ",
        " ScalingRul  eAri ": " ari : acs : ess : cn - qingdao :
 1344371 : scalingRul  e / efcqrZdjlo  okc0UkE3dA  5I0a ",
        " Cooldown ":  500 ,
        " ScalingGro  upId ": " ccMvs9dcZl  E5c9CtrwbX  zizr ",
        " Adjustment  Type ": " TotalCapac  ity ",
        " ScalingRul  eName ": " KFJoxGKXXt ",
        " Adjustment  Value ":  5
      }
    ]
  }
}
```

2 API usage instructions

Before calling the Auto Scaling API, you must activate the Auto Scaling service on the Alibaba Cloud official website, and grant Auto Scaling the API permission in the console.

For the detailed operations, see *access resources*.

Prerequisite

Before calling the Auto Scaling API, you must activate the Auto Scaling service on the Alibaba Cloud official website, and grant Auto Scaling the API permission in the console. For the detailed operations, see *access resources*.

Errors

Before calling the Auto Scaling API, you need to activate Auto Scaling at Alibaba Cloud website, and authorize your Auto Scaling to access API on the Auto Scaling console (based on Alibaba Cloud Resource Access Management [RAM], Auto Scaling uses ECS API to replace the ECS instance resources). Errors will occur in case of any inconformity:

Error codes	Error code	Description	HTTP status code
Auto Scaling is not activated and thus the API cannot be called.	Forbidden. Unsubscribed	Do not have permission to access this API.	403
You have not fully authorized Open API to Auto Scaling.	Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	403

3 API quick start

3.1 Process introduction

This topic explains how to use APIs to create and configure overall scaling solutions, including scheduled, dynamic, customized, and fixed quantity scaling.

Follow the steps shown in the following figure. The first three steps are used to create a simple scaling solution:

1. Create a scaling group. Configure the minimum (Min Size) and maximum (Max Size) number of ECS instances for scaling, and select a Server Load Balancer instance and an RDS instance.
2. Create scaling configuration. Specify the specifications of ECS instances used for Auto Scaling, such as Image ID and Instance Type.
3. Enable the scaling group according to the scaling configuration created in Step 2.
4. Create a scaling rule. For example, adding N (number) ECS instances.
5. Create a scheduled task. For example, triggering the scaling rule at 13:00.
6. Create an alarm task (CloudMonitor API PutAlarmRule). For example, adding 1 ECS instance when the CPU usage is equal to or greater than 80%.

3.2 Create a scaling group

This topic gives you an example of how to make the CreateScalingGroup API request, and specifying MinSize and MaxSize.

```
 http :// ess . aliyuncs . com /? Action = CreateScal  ingGroup
& RegionId = cn - qingdao
& MaxSize = 20
& MinSize = 2
& LoadBalanc  erId = 147b46d767  c - cn - qingdao - cm5 - a01
& DBInstance  Id .  1 = rdszzzyyun  ybaeu
& DBInstance  Id .  2 = rdsia3u3yi  a3u3y
&< Public   Request   Parameters >
```

Response example

```
< CreateScal  ingGroupRe  sponse >
    < ScalingGro  upId > dP8VqSd9EN  XPc0ciVmbc  rBT1 </ ScalingGro
 upId >
```

```
    < RequestId > 536E9CAD - DB30 - 4647 - AC87 - AA5CC38C53  82 </
 RequestId >
</ CreateScal  ingGroupRe  sponse >
```

3.3 Create a scaling configuration

You must specify the attributes of the ECS instances to be scaled, such as the ImageID and InstanceType. You must specify the ScalingGroupId returned in step 1.

```
 http :// ess . aliyuncs . com /? Action = CreateScal  ingConfigu
 ration
& ScalingGro  upId = dP8VqSd9EN  XPc0ciVmbc  rBT1
& SecurityGr  oupId = sg - 280ih3w4b
& ImageId = centos6u5_  64_20G_ali  aegis_2014  0703 . vhd
& InstanceTy  pe = ecs . t1 . xsmall
&< common   request   parameters >
```

Response example

```
< CreateScal  ingConfigu  rationResp  onse >
    < ScalingCon  figuration  Id > eOs27Kb0oX  vQcUYjEGel  JqUy </
 ScalingCon  figuration  Id >
    < RequestId > 5CC0AD41 - 08ED - 4559 - A683 - 6F56355FE0  68 </
 RequestId >
</ CreateScal  ingConfigu  rationResp  onse >
```

3.4 Enable a scaling group

This topic gives you an example of how to make the EnableScalingGroup API request according to the scaling configuration you specified, and add an existing ECS instance.

```
 http :// ess . aliyuncs . com /? Action = EnableScal  ingGroup
& ScalingGro  upId = dP8VqSd9EN  XPc0ciVmbc  rBT1
& ActiveScal  ingConfigu  rationId = eOs27Kb0oX  vQcUYjEGel  JqUy
& InstanceId .  1 = i - 283vvyytn
&< Public   Request   Parameters >
```

Response example

```
<  EnableScal  ingGroupRe  sponse >
    < RequestId > 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE </
 RequestId >
```

```
</  EnableScal  ingGroupRe  sponse >
```

3.5 Create a scaling rule

This topic is an example of API request `CreateScal  ingRule` (create a scaling rule). You must specify the `ScalingGro  upId` .

```
 http :// ess . aliyuncs . com /? Action = CreateScal  ingRule
& ScalingGro  upId = dP8VqSd9EN  XPc0ciVmbc  rBT1
& Adjustment  Type = QuantityCh  angeInCapa  city
& Adjustment  Value = 1
&< common   request   parameters >
```

Response example

```
< CreateScal  ingRuleRes  ponse >
< ScalingRul  eAri >
 ari : acs : ess : cn - qingdao : 1344371 : scalingrul  e /
 eMKWG8SRNb  9dBLAjweNI  1Ik
</ ScalingRul  eAri >
    < ScalingRul  eId > eMKWG8SRNb  9dBLAjweNI  1Ik </ ScalingRul  eId
 >
    < RequestId > 570C84F4 - A434 - 488A - AFA1 - 1E3213682B  33 </
 RequestId >
</ CreateScal  ingRuleRes  ponse >
```

3.6 Create a scheduled task

This topic is an example of API request `CreateSche  duledTask` (create a scheduled task). You must specify the `ScalingRul  eAri` .

```
 http :// ess . aliyuncs . com /? Action = CreateSche  duledTask
& RegionId = cn - qingdao
& LaunchTime = 2014 - 08 - 17T12 : 00Z
& Recurrence  Type = Daily
& Recurrence  Value = 1
& Recurrence  EndTime = 2014 - 09 - 17T16 : 55Z
& ScheduledA  ction = ari : acs : ess : cn - qingdao : 1344371 :
 scalingrul  e / eMKWG8SRNb  9dBLAjweNI  1Ik
&< common   request   parameters >
```

Response example

```
< CreateSche  duledTaskR  esponse >
    < ScheduledT  askId > edRtShc57W  GXdt8TlPbr  jsnV </ ScheduledT
 askId >
    < RequestId > 0F02D931 - 2B12 - 44D7 - A0E9 - 39925C13D1  5E </
 RequestId >
</ CreateSche  duledTaskR  esponse >
```

4 Calling method

This topic introduces the calling method.

You can call the Auto Scaling API by sending an HTTP GET request to the Auto Scaling API server and adding the relevant request parameters to the request according to the interface instructions. Results are returned based on how the request is processed .

5 Request structure

This topic introduces the request structure.

Service address

The following table describes the access endpoints of Auto Scaling API service. To reduce network latency, we recommend that you configure endpoints based on the source from which you call the service.

Region (Location)	Endpoint
Default	ess.aliyuncs.com
China (Qingdao)	ess.aliyuncs.com
China (Beijing)	ess.aliyuncs.com
China (Zhangjiakou)	ess.cn-zhangjiakou.aliyuncs.com
China (Hohhot)	ess.cn-huhehaote.aliyuncs.com
China (Hangzhou)	ess.aliyuncs.com
China (Shanghai)	ess.aliyuncs.com
China (Shenzhen)	ess.aliyuncs.com
Hong Kong	ess.aliyuncs.com
Singapore	ess.aliyuncs.com
Australia (Sydney)	ess.ap-southeast-2.aliyuncs.com
Malaysia (Kuala Lumpur)	ess.ap-southeast-3.aliyuncs.com
Indonesia (Jakarta)	ess.ap-southeast-5.aliyuncs.com
Japan (Tokyo)	ess.ap-northeast-1.aliyuncs.com
UK (London)	ess.eu-west-1.aliyuncs.com
US (Silicon Valley)	ess.aliyuncs.com
US (Virginia)	ess.aliyuncs.com
Germany (Frankfurt)	ess.eu-central-1.aliyuncs.com
UAE (Dubai)	ess.me-east-1.aliyuncs.com
India (Mumbai)	ess.ap-south-1.aliyuncs.com

Communication protocols

The system supports request communication through the HTTP or HTTPS channel. We recommend that you send requests over HTTPS for enhanced security.

Request methods

The system allows you to send HTTP GET requests. This requires request parameters to be included in the request URL.

Request parameters

For each request, you must specify the operation you want to execute, namely the Action parameter (for example, CreateScalingGroup), and each operation must contain the public and specific request parameters of the specified operation.

Character encoding

Requests and returned results are encoded using the UTF-8 character set.

6 Common parameters

This topic describes the common parameters.

Common parameters include the *common request parameters* and the *common request parameters*.

Common request parameters

The following table describes the common parameters that comprise of a URL for a GET request over HTTP or HTTPS protocol.

Name	Type	Required	Description
Action	String	Yes	The target API.
AccessKeyId	String	Yes	Equivalent to a logon password. However, an AccessKey is used to call APIs, while logon password is used to log on to the *console*. For more information, see *Create an AccessKey*.
Signature	String	Yes	Your signature. For more information, see *signatures*.
SignatureMethod	String	Yes	Signature method. Value: HMAC-SHA1.
SignatureVersion	String	Yes	Signature algorithm version. Value: 1.0.
SignatureNonce	String	Yes	Unique random number, which is used to prevent network replay attacks. Different random numbers must be used for different requests.

Name	Type	Required	Description
Timestamp	String	Yes	Request time stamp. Use the offset from Coordinated Universal Time (UTC), the time display is based on ISO8601. Format: YYYY-MM-DDThh:mm:ssZ. For example, `2018 - 01 - 01T12 : 00 : 00Z` indicates 20:00:00, Jan 01, 2018, Beijing time (UTC+8).
Version	String	Yes	The API version to use. Value: 2014-08-28.
Format	String	No	Type of the response parameters. Optional values: json | xml. Default value: json.

Request example

```
 https :// ess . aliyuncs . com /? Action = XXXXXX
& Format = xml
& Version = 2014 - 08 - 28
& Signature = Pc5WB8gokV  n0xfeu % 2FZV % 2BiNM1dgI % 3D
& SignatureM  ethod = HMAC - SHA1
& SignatureN  once = 1521552885  2396
& SignatureV  ersion = 1 . 0
& AccessKeyI  d = key - test
& Timestamp = 2018 - 01 - 01T12 : 00 : 00Z
```

```
...
```

Common response parameter

Name	Type	Description
RequestId	String	The request ID. We return a unique RequestId for every API request, whether the request is successful or not.

7 Signatures

This topic introduces the signature.

Auto Scaling performs authentication on each access request. Therefore, each request, whether being sent via HTTP or HTTPS, must contain signature information. By using AccessKeyID and AccessKeySecret, Auto Scaling performs symmetric encryption to authenticate the request sender. The AccessKeyID and AccessKeySecret are officially issued to visitors by Alibaba Cloud (visitors can apply for and manage them at Alibaba Cloud's official website). The AccessKeyID indicates the identity of the visitor. The AccessKeySecret is the secret key used to encrypt and verify the signature string on the server. It must be kept confidential and should only be available to Alibaba Cloud and the user.

When a user calls a server, the following method is used to sign the request:

1. The Canonicalized Query String is constructed using the request parameters.
   a. The request parameters are ordered alphabetically by the parameter names (this includes the "public request parameters" and user-defined parameters for the given request interfaces described in this document, but not the Signature parameter mentioned in "public request parameters").

      Note:

For a request submitted using the GET method, these parameters constitute the parameter section of the request URI (that is, the section in the URI following "?" and connected by "&").

b. The name and value of each request parameter are encoded. The names and values must be URL encoded using the UTF-8 character set. The URL encoding rules are as follows:

- English letters A–Z and a–z, digits 0–9, and characters "-", "_", ".", and "~" are not encoded.
- Other characters are encoded in the "%XY" format, with "XY" representi ng the characters’ ASCII code in hexadecimal notation. For example, the English double quotes are encoded as "%22".
- Extended UTF-8 characters are encoded in the "%XY%ZA…" format.
- Note that a space is encoded into "%20" instead of a plus sign "+".

 Note:
Generally, libraries that support URL encoding (for example, Java’s java.net.URLEncoder) are all encoded according to the rules for the "application/x-www-form-urlencoded" MIME-type. If this encoding method is used, replace the plus signs "+" in the encoded strings with "%20", the asterisks "*" with "%2A", and change "%7E" back to the tilde "~" to conform to the encoding rules described above.

c. Connect the encoded parameter names and values with the equal sign "=".

d. Then, sort the parameter name and value pairs connected by equal signs in alphabetical order, and connect them with the "&" symbol to produce the Canonicalized Query String.

2. Follow the rules below to construct the string used for signature calculation by using the Canonicalized Query String constructed in the previous step:

```
StringToSi  gn =
 HTTPMethod  + “&” +
 percentEnc  ode (“/”) + ”&” +
```

```
percentEnc  ode ( Canonicali  zedQuerySt  ring )
```

- HTTPMethod is the HTTP method used for request submission, for example, GET.
- percentEncode("/") is the coded value for the character "/" according to the URL encoding rules described in 1.ii, that is, "%2F".
- percentEncode(CanonicalizedQueryString) is the encoded string of the Canonicalized Query String constructed in Step 1, produced by following the URL encoding rules described in 1.ii.

3. According to RFC2104 definitions, use the above signature sting to calculate the signature’s HMAC value.

Note:
The Key used for calculating the signature is the AccessKeySecret held by the user, which ends with the "&" character (ASCII:38) and is based on the SHA1 hashing.

4. According to Base64 encoding rules, encode the above HMAC value into a string. This gives you the signature value.
5. Add the obtained signature value to the request parameters as the Signature parameter. This completes the request signing process.

Note:
When the obtained signature value is submitted to the ECS server as the final request parameter value, the value is URL encoded like other parameters according to RFC3986 rules.

Take DescribeScalingGroups as an example. The request URL prior to signing is as follows:

```
http :// ess . aliyuncs . com /? TimeStamp = 2014 - 08 - 15T11
% 3A10 % 3A07Z & Format = xml & AccessKeyI  d = testid & Action
= DescribeSc  alingGroup  s & SignatureM  ethod = HMAC - SHA1 &
RegionId = cn - qingdao & SignatureN  once = 1324fd0e - e2bb - 4bb1
- 917c - bd6e437f17  10 & SignatureV  ersion = 1 . 0 & Version =
2014 - 08 - 28
```

Thus, the StringToSign is:

```
GET &% 2F & AccessKeyI  d % 3Dtestid & Action % 3DDescribe
ScalingGro  ups & Format % 3Dxml & RegionId % 3Dcn - qingdao &
SignatureM  ethod % 3DHMAC - SHA1 & SignatureN  once % 3D1324fd0e
- e2bb - 4bb1 - 917c - bd6e437f17  10 & SignatureV  ersion % 3D1 .
```

```
0 & TimeStamp % 3D2014 - 08 - 15T11 % 253A10 % 253A07Z & Version %
3D2014 - 08 - 28
```

Assume that the AccessKeyID is "testid", the AccessKeySecret is "testsecret", and the Key used for HMAC calculation is "testsecret&". The calculated signature value is "SmhZuLUnXmqxSEZ%2FGqyiwGqmf%2BM=".

The signed request URL is (added with the Signature parameter):

```
http :// ess . aliyuncs . com /? TimeStamp = 2014 - 08 - 15T11
% 3A10 % 3A07Z & Format = xml & AccessKeyI  d = testid & Action
= DescribeSc  alingGroup  s & SignatureM  ethod = HMAC - SHA1 &
RegionId = cn - qingdao & SignatureN  once = 1324fd0e - e2bb - 4bb1
- 917c - bd6e437f17  10 & SignatureV  ersion = 1 . 0 & Version =
2014 - 08 - 28 & Signature = SmhZuLUnXm  qxSEZ % 2FGqyiwGqm  f %
2BM % 3D
```

For more information about signatures and request submission, see *how to call interfaces*.

8 Response results

This topic introduces the response results.

We return the results in either XML or JSON format, but JSON is the default choice. You can switch to XML schema by specifying the request parameter `Format`. For more information, see *common parameters*.

Note:
Response examples in our API documents have line breaks and indentions to make them easy to read. The actual response results are not formatted.

Success response example

Every successful response has a request ID in the RequestId element and other API-specific response parameters. The HTTP status code for a success response is 2XX.

`XML` format:

```
<?  xml   version =" 1 . 0 "  encoding =" UTF - 8 "? > <! -- Result
 root   node -->
< ActionResp  onse > <! -- Response   request   tag ->
    < RequestId > 4C467B38 - 3910 - 447D - 87BC - AC049166F2  23 </
 RequestId > <! -- Response   result   data -->
</ ActionResp  onse >
```

`JSON` format:

```
{
    " RequestId ": " 4C467B38 - 3910 - 447D - 87BC - AC049166F2  23
 " /*  Response   result   data  */
}
```

Error response example

Every error response consists of a request ID in the RequestId element and access endpoint in the HostId element, the error code, and the error message. The HTTP status code for an error response is 4xx or 5xx.

You can fix the exception according to the API-specific error codes or *common error codes* and try the request again. Alternatively, you can *open a ticket* and provide additional inputs such as the `HostId` and `RequestId` to get technical support from us.

`XML` format:

```
<?  xml   version =" 1 . 0 "  encoding =" UTF - 8 "? ><! -- Result
 root   node -->
< Error >
    < RequestId > 540CFF28 - 407A - 40B5 - B6A5 - 74Bxxxxxxx  xx </
 RequestId > <! -- Request   ID -->
    < HostId > ess . aliyuncs . com </ HostId > <! -- Endpoint -->
    < Code > InternalEr  ror </ Code > <! -- Error   code -->
    < Message > The   request   processing   has   failed   due   to
 some   unknown   error ,  exception   or   failure . </ Message > <!
 -- Error   message -->
</ Error >
```

`JSON` format:

```
{
    " RequestId ": " 540CFF28 - 407A - 40B5 - B6A5 - 74Bxxxxxxx  xx
 ", /*  Request   ID  */
    " HostId ": " ess . aliyuncs . com ", /*  Endpoint  */
    " Code ": " InternalEr  ror ", /*  Error   code  */
    " Message ": " The   request   processing   has   failed   due
 to   some   unknown   error ,  exception   or   failure ." /*  Error
   message  */
}
```

Common error codes

Error code	Error message	HTTP status code	Description
InvalidAccessKeyId.NotFound	The Access Key ID provided does not exist in our records.	400	The specified AccessKey does not exist.
InvalidParameter	The specified value of parameter < parameter name> is not valid.	400	The specified parameter is invalid.
MissingParameter	The input parameter < parameter name> that is mandatory for processing this request is not supplied.	400	You must specify the required parameters.
NoSuchVersion	The specified version does not exist.	400	The specified API version does not exist.

Error code	Error message	HTTP status code	Description
ResourceNotAvailable	Resource you requested is not available in this region or zone.	400	ESS service is unavailable in the specified region.
Throttling	Request was denied due to request throttling.	400	You have made too many frequent requests in short time. Please try again later.
UnsupportedOperation	The specified action is not supported.	400	Unable to call the specified API.
Forbidden	Users are not authorized to operate on the specified resource.	403	You cannot perform the specified action.
Forbidden.RiskControl	This operation is forbidden by Aliyun Risk Control system.	403	The specified action is under risk control.
Forbidden.Unsubscribed	Do not have permission to access this API.	403	You must enable the ESS service before calling this API.
Forbidden.UserVerification	Your user account is not verified by Aliyun.	403	You must complete the real-name registration.
SignatureDoesNotMatch	The signature we calculated does not match the one you provided.	403	The signature calculated by us is different from the one you provide.
InternalError	The request processing has failed due to some unknown error, exception or failure.	500	Internal error.

Error code	Error message	HTTP status code	Description
ServiceUnavailable	The request has failed due to a temporary failure of the server.	503	The server cannot respond to your request. Please try again later.

9 Scaling group

9.1 CreateScalingGroup

Creates a scaling group (`CreateScalingGroup`). A scaling group is a collection of ECS instances with the same application scenarios.It defines the maximum and minimum numbers of ECS instances in the group, and their associated Server Load Balancer instances, RDS instances, and other attributes.

- The scaling group, Server Load Balancer instance, and RDS instance must be in the same region. For more information, see *regions and zones*.
- You can create up to 20 scaling groups.
- The scaling group does not take effect immediately after being created. It must be enabled (`EnableScal ingGroup`) to support scaling rule trigger and perform scaling activities.
- If a Server Load Balancer instance is specified in the scaling group, notice that:
  - The Server Load Balancer instance must be in the Active (active) status.
  - The scaling group automatically attaches its ECS instances to the Server Load Balancer instance.
  - Health check must be enabled for all listener ports configured for the Server Load Balancer instance; otherwise, creation fails.
  - The default weight of an ECS instance attached to the Server Load Balancer instance is 50.
- If an RDS instance is specified in the scaling group, notice that:
  - The specified RDS instance must be in the Running (running) status.
  - the scaling group automatically attaches the intranet IP addresses of its ECS instances to the RDS access whitelist.
  - The number of IPs in the RDS instance whitelist cannot exceed the upper limit. For more information, see *whitelist* in *RDS*.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value: CreateScalingGroup
RegionId	String	Yes	ID of the region where a scaling group is located. For more information, see *regions and zones*.
MaxSize	Integer	Yes	Maximum number of ECS instances in the scaling group. Value range: [0, 100]. If the number ECS instances in the scaling group exceeds the `MaxSize`, Auto Scaling automatically removes the ECS instances.
MinSize	Integer	Yes	Minimum number of ECS instances in the scaling group. Value range: [0, 100]. If the number ECS instances in the scaling group is less than the `MinSize`, Auto Scaling automatically creates ECS instances.

Name	Type	Required	Description
Scalinggroupname	String	No	Name shown for the scaling group, which must contain 2-40 characters (English or Chinese). The name must begin with a number, an upper/ lower-case letter or a Chinese character and may contain numbers, “_ “, “- “ or “.” . The account name is unique in the same region. If this parameter is not specified, the default value is ScalingGroupId.

Name	Type	Required	Description
DefaultCooldown	Integer	No	Default cool-down time (in seconds) of the scaling group. Value range: [0, 86400]. The default value is 300s. During the cool-down time, this scaling group cannot perform any new scaling activities. Currently, this only applies to scaling activities triggered by *CloudMonitor* alarm tasks.

Name	Type	Required	Description
RemovalPolicy.N	String	No	Policy for removing ECS instances from the scaling group. For more information, see *removal policy*. Value range of N: [1,2] Optional values: · `OldestInstance` : removes the first ECS instance attached to the scaling group. · `NewestInstance` : removes the first ECS instance attached to the scaling group. · `OldestScalingConfiguration` : removes the ECS instance with the oldest scaling configuration. Default values: `OldestScalingConfiguration` and `OldestInstance` . You can enter up to two removal policies.

Name	Type	Required	Description
LoadBalancerIds	String	No	ID list of Server Load Balancer instances. A Json Array with format: [“lb-id0” , “lb-id1” , … “lb-idz”], support up to 5 Load Balancer instance.
VServerGroup.N.LoadBalancerId	String	No	The ID of Server Load Balancer instance that a VServer Group belongs to.
VServerGroup.N.VServerGroupAttribute.M.VServerGroupId	String	No	The ID of VServer Group.
VServerGroup.N.VServerGroupAttribute.M.Port	Integer	No	The port number used to attach an ECS instance to the VServer Group. Value range: 1-65535.
VServerGroup.N.VServerGroupAttribute.M.Weight	Integer	No	The weight of an ECS instance that is attached to the VServer Group. Value range: 0-100. Default value: `50` .
DBInstanceIds	String	No	ID list of an RDS instance. A Json Array with format: [“rm-id0” , “rm-id1” , … “rm-idz”], support up to 8 RDS instance.

Name	Type	Required	Description
VSwitchId	String	No	If you create a VPC scaling group, you must specify the ID of a VSwitch.

Name	Type	Required	Description
VSwitchIds.N	String	No	Parameter VSwitchIds.N is used to create instance in multiple zones. Parameter VSwitchIds.N has a priority over parameter VSwitchId. The valid range of N is [1, 5], and you can specify at most 5 VSwitches in a VPC. If you use the `VSwitchIds.N` parameter, the parameter `VSwitchId` is ignored. · VSwitches can be from multiple zones. · The priority of VSwitches descends from 1 to 5, and 1 indicates the highest priority. When the VSwitch with the highest priority cannot create an ECS instance in its zone, the system automatically uses the VSwitch with the next highest priority to create the ECS instance.

Name	Type	Required	Description
ScalingPolicy	String	No	Specify the reclaim mode of scaling group. Optional values: · `recycle`: Shutdown and Reclaim mode · `release`: Release mode For more information about instance status, see *RemoveInstances*.

Name	Type	Required	Description
MultiAZPolicy	String	No	The ECS instance scale up/down policy for a multi-zone scaling group. Optional values: · PRIORITY : Scale up/down according to the user-defined VSwitches (VSwitchIds . N). When the VSwitch with the highest priority cannot create an ECS instance in its zone, the system automatically uses the VSwitch with the next highest priority to create the ECS instance. · COST_OPTIMIZED : Available instance type with the lowest vCPU rate is given priority. Preemptible instances are created first when preemptible instance types are specified for the scaling configuration. Pay-As-You-Go instances are created automatically when all types of preemptible

Name	Type	Required	Description
LaunchTemplateId	String	No	ID of the launch template. For the specified scaling group to obtain the startup configuration information from the template.
LaunchTemplateVersion	String	No	Version of the launch template. Optional values: · Fixed template version number · `Default` : The default template version is always used. · `Latest` : The latest template version is always used.

Response parameters

Name	Type	Description
ScalingGroupId	String	ID of a scaling group, generated by the system and globally unique.

Examples

Request example

```
 http :// ess . aliyuncs . com /? Action = CreateScal  ingGroup
& RegionId = cn - qingdao
& MaxSize = 20
& MinSize = 2
& LoadBalanc  erId = 147b46d767  c - cn - qingdao - cm5 - a01
& DBInstance  Id .  1 = rdszzzyyun  ybaeu
& DBInstance  Id .  2 = rdsia3u3yi  a3u3y
&< Public   Request   Parameters >
```

Response example

XML format:

```
< CreateScal  ingGroupRe  sponse >
    < ScalingGro  upId > dP8VqSd9EN  XPc0ciVmbc  rBT1 </ ScalingGro
 upId >
    < RequestId > 536E9CAD - DB30 - 4647 - AC87 - AA5CC38C53  82 </
 RequestId >
</ CreateScal  ingGroupRe  sponse >
```

JSON format:

```
{
 " RequestId ": " 536E9CAD - DB30 - 4647 - AC87 - AA5CC38C53  82 ",
 " ScalingGro  upId ": " dP8VqSd9EN  XPc0ciVmbc  rBT1 "
}
```

Error codes

Error codes specific to the `CreateScalingGroup` API are as follows. For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
IncorrectDBInstanceStatus	The current status of DB instance “XXX” does not support this action.	400	The specified RDS instance is not in the Running status.
IncorrectLoadBalancerHealthCheck	The current health check type of specified load balancer does not support this action.	400	Health check is not enabled for the specified Server Load Balancer instance.
IncorrectLoadBalancerStatus	The current status of the specified load balancer does not support this action.	400	The specified Server Load Balancer instance is not active.
IncorrectVSwitchStatus	The current status of virtual switch does not support this operation.	400	The VSwitch is unavailable and instances cannot be created.

Error code	Error message	HTTP status code	Description
InvalidDBInstanceId.RegionMismatch	DB instance “XXX” and the specified scaling group are not in the same Region.	400	The specified RDS instance must be in the same region as the scaling group.
InvalidLoadBalancerId.IncorrectAddressType	The current address type of specified load balancer does not support this action.	400	The Server Load Balancer instance is a private network instance, but a VSwitch has been specified. For more information, see *Server Load Balancer instances*.
InvalidLoadBalancerId.IncorrectInstanceNetworkType	The network type of the instance in specified Load Balancer does not support this action.	400	The network type of the ECS instance contained in the specified Server Load Balancer is different from the network type of the scaling group.
InvalidLoadBalancerId.RegionMismatch	The specified Load Balancer and the specified scaling group are not in the same Region.	400	The specified RDS instance and scaling group are not in the same region.
InvalidLoadBalancerId.VPCMismatch	The specified virtual switch and the instance in specified Load Balancer are not in the same VPC.	400	The ECS instance contained in the specified Server Load Balancer and VSwitchId are not in the same VPC.
InvalidParameter	The specified value of parameter "ScalingPolicy" is not valid.	400	The specified reclaim mode does not exist.

Error code	Error message	HTTP status code	Description
InvalidParameter.Conflict	The value of parameter < parameter name " and parameter < parameter name" are conflict.	400	The specified `MinSize` is greater than `MaxSize` .
InvalidScalingGroupName.Duplicate	The specified value of parameter < parameter name" is duplicated.	400	The scaling group name already exists .
QuotaExceeded.DBInstanceSecurityIP	Security IP quota exceeded in DB instance “XXX” .	400	The number of IP addresses in the access whitelist of the specified RDS instance exceeds the upper limit.
QuotaExceeded.PrivateIpAddress	Private IP address quota exceeded in the specified virtual switch.	400	The private IP address quota of the VSwitch is exceeded.
QuotaExceeded.ScalingGroup	Scaling group quota exceeded.	400	The scaling group quota is exceeded.
QuotaExceeded.VPCInstance	Instance quota exceeded in the specified VPC.	400	The instance quota of the VPC is exceeded.
InvalidDBInstanceId.NotFound	DB instance “XXX” does not exist.	404	The specified RDS instance does not exist.
InvalidLoadBalancerId.NotFound	The specified Load Balancer does not exist.	404	The specified Server Load Balancer instance does not exist.
InvalidRegionId.NotFound	The specified region does not exist.	404	The specified RegionId does not exist.
InvalidVSwitchId.NotFound	The specified virtual switch does not exist.	404	The specified VSwitch does not exist.

Error code	Error message	HTTP status code	Description
LaunchTemplateVersionSet.NotFound	The specific version of launch template is not exist.	400	The specified version of the launch template does not exist.
LaunchTemplateSet.NotFound	The specified launch template set is not found.	400	The specified launch template does not exist.
TemplateMissingParameter.ImageId	The input parameter "ImageId" that is mandatory for processing this request is not supplied.	400	The required parameter ImageId is not supplied.
TemplateMissingParameter.InstanceTypes	The input parameter "InstanceTypes" that is mandatory for processing this request is not supplied.	400	The required parameter InstanceType is not supplied.
TemplateMissingParameter.SecurityGroup	The input parameter "SecurityGroup" that is mandatory for processing this request is not supplied.	400	The required parameter SecurityGroup is not supplied.
TemplateVersion.NotNumber	The input parameter "LaunchTemplateVersion" is supposed to be a string representing the version number.	400	The fixed version number of the specified launch template must be numbers.

9.2 ModifyScalingGroup

This topic introduces how to modify a scaling group using the API.

Description

- Modifies the attributes of a scaling group. However, the following attributes cannot be modified:
  - RegionId
  - LoadBalancerId
  - DBInstanceId
- The interface can be called only when the scaling group is active or inactive.
- When the scaling configuration specified for the scaling group needs to be modified, the instance type attribute of the modified scaling configuration must be consistent with that of the active scaling configuration.

  After a new scaling configuration is added to the scaling group, the running ECS instances which are created based on the previous scaling configuration remain unchanged.
- When the number (total capacity) of ECS instances in the scaling group does not meet the modified MaxSize or MinSize specification, the Auto Scaling service automatically attaches or removes ECS instances to/from the group to make odds even.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value : ModifyScalingGroup
ScalingGroupId	String	Yes	Scaling group ID

Name	Type	Required	Description
ScalingGroupName	String	No	Name shown for the scaling group, which must contain 2-40 characters (English or Chinese). The name must begin with a number, an upper/ lower-case letter or a Chinese character and may contain numbers, “_” , “- “ or “.” . The account name is unique in the same region.
ActiveScal ingConfigurationId	String	No	ID of the active scaling configurat ion in the scaling group.
MinSize	Integer	No	Minimum number of ECS instances in the scaling group. Value range: [0, 100].
MaxSize	Integer	No	Maximum number of ECS instances in the scaling group. Value range: [0, 100].
DefaultCooldown	Integer	No	Default cool-down time (in seconds) of the scaling group . Value range: [0, 86400].

Name	Type	Required	Description
RemovalPolicy.N	String	No	Policy for removing ECS instances from the scaling group. Optional values: · OldestInstance : removes the first ECS instance attached to the scaling group. · NewestInstance : removes the first ECS instance attached to the scaling group. · OldestScalingConfiguration : removes the ECS instance with the oldest scaling configuration. You can enter up to two removal policies.
LaunchTemplateId	String	No	ID of the launch template. For the specified scaling group to obtain startup configuration information from the launch template.

Name	Type	Required	Description
LaunchTemplateVersion	String	No	Version of the launch template. Optional values: · Fixed template version number · `Default` : The default template version is always used. · `Latest` : The latest template version is always used.

Name	Type	Required	Description
VSwitchIds.N	String	No	· VSwitchIds.N takes effect only when the network type of scaling group is VPC. · VSwitchIds.N can be used for specifying VSwitchs from multiple zones, and the value range of N is [1, 5]. When modifying the scaling group, you can specify up to 5 VSwitchs in the same VPC. · The VSwitchs and scaling group must belong to the same VPC. · The priorities of VSwitchs are determined by their numbers, and 1 has the highest priority. · When an instance cannot be created in the zone where the VSwitch with higher priority belong, the VSwitch with the lower priority will be selected.

Response parameters

Public parameters.

Error codes

For common errors, see *client errors* or *server errors*.

Error message	Error code	Description	HTTP status code
The specified scaling group does not exist in this account.	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	404
The scaling group name already exists.	InvalidScalingGroupName.Duplicate	The specified value of parameter <parameter name> is duplicated.	400
The specified scaling configuration does not exist in the scaling group.	InvalidScalingConfigurationId.NotFound	The specified scaling configuration does not exist.	404
The instance types of the specified scaling configuration and the active scaling configuration do not match.	InvalidScalingConfigurationId.InstanceTypeMismatch	The specified scaling configuration and existing active scaling configuration have different instance type.	400
The specified MinSize is greater than MaxSize.	InvalidParameter.Conflict	The value of parameter <parameter name> and parameter <parameter name> are confilict.	400
The specified launch template version does not exist.	LaunchTemplateVersionSet.NotFound	The specific version of launch template does not exist.	400
The specified lauch template does not exist.	LaunchTemplateSet.NotFound	The specified launch template set is not found.	400

Error message	Error code	Description	HTTP status code
The required parameter ImageId is not supplied.	TemplateMissingParameter.ImageId	The input parameter "ImageId" that is mandatory for processing this request is not supplied.	400
The required parameter InstanceTypes is not supplied.	TemplateMissingParameter.InstanceTypes	The input parameter "InstanceTypes" that is mandatory for processing this request is not supplied.	400
The required parameter SecurityGroup is not supplied.	TemplateMissingParameter.SecurityGroup	The input parameter "SecurityGroup" that is mandatory for processing this request is not supplied.	400
The input parameter LaunchTemplateVersion must be a number.	TemplateVersion.NotNumber	The input parameter "LaunchTemplateVersion" is supposed to be a string representing the version number.	400

Request example

```
 http :// ess . aliyuncs . com /? Action = ModifyScal  ingGroup
& ScalingGro  upId = cqS5Qbbhmv  GLcJbWoDbW  Lj2V
& ScalingGro  upName = ScalingGro  up
&< Public   Request   Parameters >
```

Response example

XML format:

```
<  ModifyScal  ingGroupRe  sponse >
    < RequestId > 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE </
 RequestId >
```

```
</  ModifyScal  ingGroupRe  sponse >
```

JSON format:

```
" RequestId ": " 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE "
```

9.3 EnableScalingGroup

This topic introduces how to enable a scaling group using the API.

Description

- Enables the specified scaling group.
  - After the scaling group is successfully enabled (the group is active), the ECS instances specified by the interface are attached to the group.
  - If the current number of ECS instances in the scaling group is still smaller than MinSize after the ECS instances specified by the interface are attached, the Auto Scaling service automatically creates ECS instances in Pay-As-You-Go mode to make odds even. For example, a scaling group is created with MinSize = 5. Two existing ECS instances are specified by the InstanceId.N parameter when the scaling group is enabled. Three additional ECS instances are automatically created after the two ECS instances are attached by the Auto Scaling service to the scaling group.
- The interface can be called only when the scaling group is inactive.
- If the scaling group has no active scaling configurations, you need to input scaling configurations when enabling the scaling group.
  - A single scaling group can have only one active scaling configuration at a time.
  - If an active scaling configuration has been created before the scaling group is enabled, input of a new active scaling configuration through the interface makes the previous scaling configuration inactive.

- Restrictions on attaching ECS instances:
  - The attached ECS instance and the scaling group must be in the same region.
  - The attached ECS instance and the instance with active scaling configurations must be of the same type.
  - The attached ECS instance must in the Running state.
  - The attached ECS instance has not been attached to other scaling groups.
  - The attached ECS instance supports Subscription and Pay-As-You-Go payment methods.
  - If the VswitchID is specified for a scaling group, you cannot attach Classic ECS instances or ECS instances on other VPCs to the scaling group.
  - If the VswitchID is not specified for the scaling group, ECS instances of the VPC type cannot be attached to the scaling group.
- The call fails if the number (total capacity) of instances specified by the interface plus instances in the scaling group is greater than MaxSize.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter; value : EnableScalingGroup
ScalingGroupId	String	Yes	Scaling group ID.
ActiveScalingConfigurationId	String	No	ID of the scaling configuration to be activated in a scaling group.
InstanceId.N	String	No	ID of the ECS instance to be attached to the scaling group after it is enabled. You can input up to 20 IDs.

Name	Type	Required	Description
LaunchTemplateId	String	No	ID of the launch template. For the specified scaling group to obtain the startup configuration information from the template.
LaunchTemplateVersion	String	No	Version of the launch template. Optional values: · Fixed template version number · `Default` : The default template version is always used. · `Latest` : The latest template version is always used.

Response parameters

Public parameters.

Error codes

For common errors, see *client errors* or *server errors*.

Error message	Error code	Description	HTTP status code
The specified scaling group does not exist in this account.	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	404
API is not fully authorized to the Auto Scaling service.	Forbidden.Unauthorized	A required authorization for the specified action is not supplied.	403

Error message	Error code	Description	HTTP status code
The specified scaling group is in the deleting state.	IncorrectScalingGroupStatus	The current status of the specified scaling group does not support this action.	400
The specified scaling configuration does not exist in the scaling group.	InvalidScalingConfigurationId.NotFound	The specified scaling configuration does not exist.	404
The instance types of the specified scaling configuration and the active scaling configuration do not match.	InvalidScalingConfigurationId.InstanceTypeMismatch	The specified scaling configuration and existing active scaling configuration have different instance type.	400
No active scaling configuration is specified for the scaling group.	MissingActiveScalingConfiguration	An active scaling configuration for the specified scaling group is not supplied.	400
The specified ECS instance does not exist in this account.	InvalidInstanceId.NotFound	Instance “XXX” does not exist.	404
The specified ECS instance and the scaling group are not in the same region.	InvalidInstanceId.RegionMismatch	Instance “XXX” and the specified scaling group are not in the same Region.	400
The instance types of the specified ECS instance and the scaling configuration do not match.	InvalidInstanceId. InstanceTypeMismatch	Instance “XXX” and existing active scaling configuration have different instance type.	400
The specified ECS instance is not in the Running status.	IncorrectInstanceStatus	The current status of instance “XXX” does not support this action.	400

Error message	Error code	Description	HTTP status code
The network types of the specified ECS instance and the scaling configuration do not match.	InvalidInstanceId.NetworkTypeMismatch	The network type of instance “XXX” does not support this action.	400
The specified scaling group and the attached ECS instance are not in the same VPC.	InvalidInstanceId.VPCMismatch	Instance “XXX” and the specified scaling group are not in the same VPC.	400
The specified ECS instance has been attached to another scaling group.	InvalidInstanceId.InUse	Instance "XXX" is already attached to another scaling group.	400
The specified Server Load Balancer instance is not active.	IncorrectLoadBalancerStatus	The current status of the specified load balancer does not support this action.	400
Health check is not enabled for the specified Server Load Balancer instance.	IncorrectLoadBalancerHealthCheck	The current health check type of specified load balancer does not support this action.	400
The network type of the ECS instance contained in the specified Server Load Balancer is different from the network type of the scaling group.	InvalidLoadBalancerId.IncorrectInstanceNetworkType	The network type of the instance in specified Load Balancer does not support this action.	400
The ECS instance contained in the specified Server Load Balancer and VSwitchId are not in the same VPC.	InvalidLoadBalancerId.VPCMismatch	The specified virtual switch and the instance in specified Load Balancer are not in the same VPC.	400

Error message	Error code	Description	HTTP status code
The specified RDS instance is not running.	IncorrectDBInstanceStatus	The current status of DB instance “XXX” does not support this action.	400
Total Capacity after the ECS instance is attached is greater than MaxSize.	IncorrectCapacity.MaxSize	To attach the instances, the total capacity will be greater than the max size.	400
The specified version of the launch template does not exist.	LaunchTemplateVersionSet.NotFound	The specific version of launch template does not exist.	400
The specified launch template does not exist.	LaunchTemplateSet.NotFound	The specified launch template set is not found.	400
The required parameter ImageId is not supplied.	TemplateMissingParameter.ImageId	The input parameter "ImageId" that is mandatory for processing this request is not supplied.	400
The required parameter InstanceTypes is not supplied.	TemplateMissingParameter.InstanceTypes	The input parameter "InstanceTypes" that is mandatory for processing this request is not supplied.	400
The required parameter SecurityGroup is not supplied.	TemplateMissingParameter.SecurityGroup	The input parameter "SecurityGroup" that is mandatory for processing this request is not supplied.	400

Error message	Error code	Description	HTTP status code
The fixed version number of the specified launch template must be numbers.	TemplateVersion.NotNumber	The input parameter "LaunchTemplateVersion" is supposed to be a string representing the version number.	400

Request example

```
 http :// ess . aliyuncs . com /? Action = EnableScal  ingGroup
& ScalingGro  upId = dmIDKNcyWf  zncX9MWX1b  wFV
& InstanceId .  1 = i - 283vvyytn
&< Public   Request   Parameters >
```

Response example

XML format:

```
<  EnableScal  ingGroupRe  sponse >
    < RequestId > 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE </
 RequestId >
</  EnableScal  ingGroupRe  sponse >
```

JSON format:

```
{
    " RequestId ": " 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE "
}
```

9.4 DisableScalingGroup

You can call this operation to disable a scaling group.

Description

This operation disables a specified scaling group.

- Scaling activities that are currently being executed by the specified scaling group before this operation is called will continue until they are completed. Scaling activity requests that are made to the target scaling group after this operation has been called will be rejected.
- When you call this operation, note that the scaling group must be active.

Debugging

OpenAPI Explorer simplifies API usage. You can use OpenAPI Explorer to perform operations such as retrieve APIs, call APIs, and dynamically generate SDK example code.

Request parameters

Parameter	Type	Required	Example	Description
ScalingGroupId	String	Yes	dmIDKNcyWfzncX9MWX1****	The ID of the scaling group.
Action	String	No	DisableScalingGroup	The operation that you want to perform. Set the value to DisableScalingGroup.

Response parameters

Parameter	Type	Example	Description
RequestId	String	473469C7-AA6F-4DC5-B3DB-A3DC0DE3C83E	The ID of the request. This parameter is returned regardless of whether the operation is successful.

Examples

Sample requests

```
http :// ess . aliyuncs . com /? Action = DisableSca lingGroup
& ScalingGro upId = dmIDKNcyWf zncX9MWX1 ****
&< Common request parameters >
```

Successful response examples

`XML` format

```
< DisableSca lingGroupR esponse >
  < RequestId > 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8 3E </
 RequestId >
```

```
</ DisableSca  lingGroupR  esponse >
```

`JSON` format

```
{
 " RequestId ":" 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8  3E "
}
```

Error codes

View error codes

HTTP status code	Error code	Error message	Description
404	InvalidSca lingGroupId. NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist in the current account.

9.5 DeleteScalingGroup

You can call this operation to delete a scaling group.

Description

This operation deletes a specified scaling group.

- ForceDelete indicates whether to forcibly delete a scaling group and remove or release its associated ECS instances regardless of whether it is undergoing a scaling activity.
- If ForceDelete is set to false, the scaling group can be deleted only when the following conditions are met:
  - Condition 1: The scaling group is not currently undergoing a scaling activity.
  - Condition 2: The number of ECS instances currently in the scaling group (Total Capacity) is 0.
  - If both of these conditions are met, the scaling group can be disabled and deleted.

- If ForceDelete is set to true, we recommend that you perform the following procedure:
  - Disable the scaling group to reject new scaling activity requests. Remove all ECS instances from the scaling group after the ongoing scaling activities are completed and then delete the scaling group. Manually added ECS instances are only removed from the scaling group, whereas automatically created instances are deleted.
- Note that deleting a scaling group also deletes its scaling configurations, rules, activities, and requests.
- The following tasks and instances are not deleted when a scaling group is deleted: scheduled tasks, CloudMonitor alert tasks, SLB instances, and RDS instances.

Debugging

OpenAPI Explorer simplifies API usage. You can use OpenAPI Explorer to perform operations such as retrieve APIs, call APIs, and dynamically generate SDK example code.

Request parameters

Parameter	Type	Required	Example	Description
`ScalingGroupId`	String	Yes	dmIDKNcyWfzncX9MWX1****	The ID of the scaling group.
`Action`	String	No	DeleteScalingGroup	The operation that you want to perform. Set the value to DeleteScalingGroup.
`ForceDelete`	Boolean	No	false	Indicates whether to forcibly delete the scaling group and remove or release its associated ECS instances regardless of whether it is undergoing a scaling activity. Default value: false, indicating that the scaling group is not to be forcibly deleted.

Response parameters

Parameter	Type	Example	Description
RequestId	String	473469C7-AA6F-4DC5-B3DB-A3DC0DE3C83E	The ID of the request. This parameter is returned regardless of whether the operation is successful.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = DeleteScal  ingGroup
& ScalingGro  upId = dmIDKNcyWf  zncX9MWX1 ****
&< Common   request   parameters >
```

Successful response examples

`XML` format

```
< DeleteScal  ingGroupRe  sponse >
  < RequestId > 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8  3E </
 RequestId >
</ DeleteScal  ingGroupRe  sponse >
```

`JSON` format

```
{
 " RequestId ":" 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8  3E "
}
```

Error codes

View error codes

HTTP status code	Error code	Error message	Description
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist in the current account.

HTTP status code	Error code	Error message	Description
403	Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	The error message returned when Auto Scaling is not authorized to call the specified operation.
400	InstanceInUse	You cannot delete a scaling configuration or scaling group while there is an instance associated with it.	The error message returned when the specified scaling group still contains ECS instances and cannot be deleted.

9.6 DescribeScalingGroups

This topic introduces how to query a scaling group using the API.

Description

Queries the information of a scaling group. Scaling groups have the following life cycle states:

- Active: In this state, the scaling group can receive scaling rule execution requests and trigger scaling activities.
- Inactive: In this state, the scaling group does not receive scaling rule execution requests.
- Deleting: The scaling group is being deleted and does not receive scaling rule execution requests.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value: DescribeScalingGroups

Name	Type	Required	Description
RegionId	String	Yes	ID of the region where the scaling group is located.
ScalingGroupId.N	String	No	Scaling group ID. You can enter up to 20 IDs. Invalid scaling group IDs are not displayed in query results, and no error is reported.
ScalingGroupName.N	String	No	Scaling group name. You can enter up to 20 names. Invalid scaling group names are not displayed in query results, and no error is reported.
PageNumber	Integer	No	Page number of the scaling group list, staring from 1. Default value: 1.
PageSize	Integer	No	When querying by page, this parameter indicates the number of lines per page. Maximum value: 50; default value: 10.

Response parameters

Name	Type	Description
TotalCount	Integer	Total number of scaling groups
PageNumber	Integer	Current page number
PageSize	Integer	Number of lines per page

Name	Type	Description
ScalingGroups	ScalingGroupSetType	Scaling group information set

ScalingGroupSetType is a set of ScalingGroupItemTypes.

Name	Type	Description
ScalingGroup	ScalingGroupItemType	Scaling group information

The attributes of ScalingGroupItemType are listed as follows.

Name	Type	Description
ScalingGroupId	String	Scaling group ID.
ScalingGroupName	String	Name shown for the scaling group.
ActiveScalingConfigurationId	String	ID of the active scaling configuration in the scaling group.
LaunchTemplateId	String	ID of the launch template used by the scaling group.
LaunchTemplateVersion	String	Version of the launch template used by the scaling group.
RegionId	String	ID of the region where the scaling group is located.
MinSize	Integer	Minimum number of ECS instances in the scaling group.
MaxSize	Integer	Maximum number of ECS instances in the scaling group.
DefaultCooldown	Integer	Default cool-down time of the scaling group.
RemovalPolicies	RemovalPolicySetType	A set of policies for removing ECS instances from the scaling group.
LoadBalancerIds	List of LoadBalancerId	ID of the Server Load Balancer instance.

Name	Type	Description
DBInstanceIds	DBInstanceIdSetType	ID of the RDS instance.
VSwitchId	String	ID of the virtual switch corresponding to the scaling group.
LifecycleState	String	Status of the scaling group.
TotalCapacity	Integer	Total number of ECS instances in the scaling group.
ActiveCapacity	Integer	Number of ECS instances which have been attached to the scaling group and are running properly.
PendingCapacity	Integer	Number of ECS instances which are being attached to the scaling group with relevant configurations not completed.
RemovingCapacity	Integer	Number of ECS instances which are being removed from the scaling group .
CreationTime	String	Time when the scaling group is created.

RemovalPolicySetType is a set of String types.

Name	Type	Description
RemovalPolicy	String	Policy for removing ECS instances from the scaling group.

DBInstanceIdSetType is a set of String types.

Name	Type	Description
DBInstanceId	String	ID of the RDS instance

Error code

For common errors, see *client errors* or *server errors*.

Request example

```
 http :// ess . aliyuncs . com /? Action = DescribeSc  alingGroup  s
& RegionId = cn - qingdao
& PageSize = 50
&< Public   Request   Parameters >
```

Response example

XML format:

```
< DescribeSc  alingGroup  sResponse >
    < RequestId > 6393C3A8 - B611 - 42F2 - AFA6 - F080FC45D5  D0 </
 RequestId >
    < TotalCount > 1 </ TotalCount >
    < PageNumber > 1 </ PageNumber >
    < PageSize > 10 </ PageSize >
    < ScalingGro  ups >
        < ScalingGro  up >
            < ActiveCapa  city > 1 </ ActiveCapa  city >
            < ActiveScal  ingConfigu  rationId > dyo713cNYI
 B4ddEVlKbc  pOef </ ActiveScal  ingConfigu  rationId >
            < DBInstance  Ids >
                < DBInstance  Id > rdszzzyyun  ybaeu </ DBInstance  Id
 >
            </ DBInstance  Ids >
            < VSwitchId > vpc - 25j4god4l </ VSwitchId >
            < DefaultCoo  ldown > 20 </ DefaultCoo  ldown >
            < LifecycleS  tate > Active </ LifecycleS  tate >
            < LoadBalanc  erIds >
                < LoadBalanc  erId > 147b46d767  c - cn - qingdao -
 cm5 - a01 </ LoadBalanc  erId >
            </ LoadBalanc  erIds >
            < MaxSize > 1 </ MaxSize >
            < MinSize > 0 </ MinSize >
            < PendingCap  acity > 0 </ PendingCap  acity >
            < RegionId > cn - qingdao </ RegionId >
            < RemovingCa  pacity > 0 </ RemovingCa  pacity >
            < ScalingGro  upId > dyrSuvBOtO  1dEdIlIbpl  Qb8 </
 ScalingGro  upId >
            < ScalingGro  upName > dyrSuvBOtO  1dEdIlIbpl  Qb8 </
 ScalingGro  upName >
            < RemovalPol  icies >
                < RemovalPol  icy > OldestScal  ingConfigu  ration </
 RemovalPol  icy >
                < RemovalPol  icy > OldestInst  ance </ RemovalPol
 icy >
            </ RemovalPol  icies >
            < TotalCapac  ity > 1 </ TotalCapac  ity >
            < CreationTi  me > 2014 - 08 - 14T10 : 58Z </ CreationTi
 me >
        </ ScalingGro  up >
    </ ScalingGro  ups >
</ DescribeSc  alingGroup  sResponse >
```

JSON format:

```
{
" RequestId ": " 68386699 - 8B9E - 4D5B - BC4C - 75A28F6C2A  00 ",
" TotalCount ":  1 ,
" PageSize ":  10 ,
```

```
" PageNumber ":  1 ,
  " ScalingGro  ups ": {
    " ScalingGro  up ": [
      {
        " ScalingGro  upId ": " b8pYCVbIV5  k9cz4PWpbe  0k19 ",
        " ScalingGro  upName ": " b8pYCVbIV5  k9cz4PWpbe  0k19 ",
        " RegionId ": " cn - qingdao ",
        " RemovingCa  pacity ":  0 ,
        " DefaultCoo  ldown ":  300 ,
        " MinSize ":  1 ,
        " MaxSize ":  2 ,
        " LifecycleS  tate ": " Inactive ",
        " ActiveScal  ingConfigu  rationId ": "  dyo713cNYI
 B4ddEVlKbc  pOef ",
        " LoadBalanc  erIds ": {
          " LoadBalanc  erId ": [
            " 147b46d767  c - cn - qingdao - cm5 - a01 "
          ]
        },
        " PendingCap  acity ":  0 ,
        " TotalCapac  ity ":  0 ,
        " ActiveCapa  city ":  0 ,
        " CreationTi  me ": " 2014 - 08 - 14T10 : 58Z ",
        " DBInstance  Ids ": {
          " DBInstance  Id ": [
            " rdsia3u3yi  a3u3y ",
            " rdszzzyyun  ybaeu "
          ]
        },
        " VSwitchId ":" vpc - 25j4god4l ",
        " RemovalPol  icies ": {
          " RemovalPol  icy ": [
            " OldestScal  ingConfigu  ration ",
            " OldestInst  ance "
          ]
        }
      }
    ]
  }
 }
```


9.7 DescribeScalingInstances

This topic introduces how to query the list of ECS instances in a scaling group using the API.

Description

Queries the list of ECS instances in a scaling group. You can query by scaling group ID , scaling configuration ID, health status, lifecycle status, and creation type.

- The ECS instances in a scaling group can be created automatically or attached manually.
  - ECS instances can be automatically created by the Auto Scaling service based on your scaling configurations and rules.
  - ECS instances can also be manually attached to a scaling group.
- An ECS instance in a scaling group may be in the following states during its lifecycle:
  - Pending: The ECS instance is being attached to the scaling group, with operations such as instance creation, attaching to Server Load Balancer, and adding to the RDS access whitelist.
  - InService: The ECS instance is successfully added to the scaling group and provides services properly.
  - Removing: The ECS instance is being removed from the scaling group.
- An ECS instance in a scaling group may be in the following health states:
  - Healthy: The ECS instance is healthy.
  - Unhealthy: The ECS instance is unhealthy if it is not in the Running (`Running`) status.
- The Auto Scaling service automatically removes unhealthy ECS instances from scaling groups.
  - If the unhealthy ECS instances are automatically created, the Auto Scaling service disables and releases them.
  - If the unhealthy ECS instances are manually attached, the Auto Scaling service does not disable or release them.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value : DescribeScalingInstances.
RegionId	String	Yes	ID of the region where the scaling group is located.

Name	Type	Required	Description
ScalingGroupId	String	No	Scaling group ID.
ScalingConfigurationId	String	No	ID of the associated scaling configuration.
InstanceId.N	String	No	ECS instance ID. You can input up to 20 IDs. Invalid instance IDs are not displayed in query results, and no error is reported.
HealthStatus	String	No	Health status of an ECS instance in the scaling group. Optional values: · Healthy · Unhealthy
LifecycleState	String	No	Lifecycle status of an ECS instance in the scaling group. Optional values: · InService: the ECS instance has been added to the scaling group and runs properly. · Pending: the ECS instance is being attached to the scaling group with relevant configurations not completed. · Removing: the ECS instance is being removed from the scaling group.

Name	Type	Required	Description
CreationType	String	No	ECS instance creation type. Optional values: · AutoCreated: the ECS instance is automatically created by the Auto Scaling service in the scaling group. · Attached: the ECS instance is created outside the Auto Scaling service and manually attached to the scaling group.
PageNumber	Integer	No	Page number of the ECS instance list , staring from 1. Default value: 1
PageSize	Integer	No	When querying by page, this parameter indicates the number of lines per page. Maximum value: 50; default value: 10

Response parameters

Name	Type	Description
TotalCount	Integer	Total number of ECS instances
PageNumber	Integer	Current page number
PageSize	Integer	Number of lines per page

Name	Type	Description
ScalingInstances	ScalingInstanceSetType	A set of ECS instance information

ScalingInstanceSetType is a set of ScalingInstanceItemTypes.

Name	Type	Description
ScalingInstance	ScalingInstanceItemType	ECS instance information

The attributes of ScalingInstanceItemType are listed as follows.

Name	Type	Description
InstanceId	String	ECS instance ID.
ScalingGroupId	String	ID of the scaling group to which the ECS instance belongs.
ScalingConfigurationId	String	ID of the associated scaling configuration.
HealthStatus	String	Health status of the ECS instance in the scaling group.
LifecycleState	String	Lifecycle status of the ECS instance in the scaling group.
CreationTime	String	Time when the ECS instance is attached to the scaling group.
CreationType	String	ECS instance creation type.

Error code

For common errors, see *client errors* or *server errors*.

Request example

```
 http :// ess . aliyuncs . com /? Action = DescribeSc  alingInsta
 nces
& RegionId = cn - qingdao
& ScalingGro  upId = dBCYxE26IH  KcGq1xPcTN  BwV
```

```
&< Public   Request   Parameters >
```

Response example

XML format:

```
< DescribeSc  alingInsta  ncesRespon  se >
    < RequestId > DFF8797F - 5B73 - 4BD7 - A7D0 - 03479C458F  7A </
 RequestId >
< TotalCount > 1 </ TotalCount >
    < PageNumber > 1 </ PageNumber >
    < PageSize > 10 </ PageSize >
    < ScalingIns  tances >
        < ScalingIns  tance >
            < CreationTi  me > 2014 - 08 - 15T17 : 37Z </ CreationTi
 me >
            < CreationTy  pe > AutoCreate  d </ CreationTy  pe >
            < HealthStat  us > Healthy </ HealthStat  us >
            < InstanceId > i - 283vvyytn </ InstanceId >
            < LifecycleS  tate > InService </ LifecycleS  tate >
   < ScalingCon  figuration  Id >
 cGsGHrdMBa  3DcDrrBVcc  4k2H
</ ScalingCon  figuration  Id >
            < ScalingGro  upId > dBCYxE26IH  KcGq1xPcTN  BwV </
 ScalingGro  upId >
        </ ScalingIns  tance >
    </ ScalingIns  tances >
</ DescribeSc  alingInsta  ncesRespon  se >
```

JSON format:

```
{
  " RequestId ": " 13305F2D - A4C2 - 4E6B - B7C7 - 0F2150842E  A3 ",
" TotalCount ":  1 ,
" PageNumber ":  1 ,
" PageSize ":  50 ,
  " ScalingIns  tances ": {
    " ScalingIns  tance ": [
      {
        " ScalingCon  figuration  Id ": " bU5uZHcAgt  zwcL4IeDea  vqTS
 ",
        " CreationTy  pe ": " AutoCreate  d ",
        " InstanceId ": " i - 28sov3exk ",
        " CreationTi  me ": " 2014 - 08 - 14T10 : 59Z ",
        " HealthStat  us ": " Healthy ",
        " LifecycleS  tate ": " InService ",
        " ScalingGro  upId ": " dE9YbOdCHq  aFdFZHXVdD  jQCB "
      }
    ]
  }
```

```
}
```

9.8 DescribeScalingActivities

This topic introduces how to query a scaling activity using the API.

Description

Queries a scaling activity.

- You can specify a scaling group ID to query all scaling activities in this scaling group.
- You can filter the query results based on the scaling activity status.
- Only scaling activities during the last 30 days can be returned.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value : DescribeScalingActivities
RegionId	String	Yes	Region of the scaling activity
ScalingGroupId	String	No	Scaling group ID
ScalingActivityId.N	String	No	Scaling group ID. You can enter up to 10 IDs. Invalid scaling group IDs are neglected in the query results, and no error is reported.

Name	Type	Required	Description
StatusCode	String	No	Scaling activity status. Optional values: Successful: · Successful scaling activities. · Warning: Partially successful scaling activities. · Failed: Failed scaling activities. · InProgress: Scaling activities in progress. · Rejected: The scaling activity request is rejected.
PageNumber	Integer	No	Page number of the scaling activity list starting from 1. The default value is 1.
PageSize	Integer	No	When querying by page, this parameter indicates the number of lines per page. Maximum value: 50; default value: 10

Response parameters

Name	Type	Description
TotalCount	Integer	Total number of the scaling activities
PageNumber	Integer	Current page number
PageSize	Integer	Number of lines per page

Name	Type	Description
ScalingActivities	ScalingActivitySetType	Set of the scaling activity information

ScalingActivities is a set consisting of the ScalingActivityItemType.

Name	Type	Description
ScalingActivity	ScalingActivityItemType	Scaling activity informatio n

Attributes of the ScalingActivityItemType are introduced as follows.

Name	Type	Description
ScalingActivityId	String	Scaling activity ID
ScalingGroupId	String	Scaling group ID
Description	String	Description on the scaling activity
Cause	String	Cause that triggers the scaling activity
StartTime	String	Start time of the scaling activity
EndTime	String	End time of the scaling activity
Progress	Integer	Running speed of the scaling activity
StatusCode	String	Current status of the scaling activity
StatusMessage	String	Message about the scaling activity status

Request example

```
 http :// ess . aliyuncs . com /? Action = DescribeSc  alingActiv
 ities
& RegionId = cn - qingdao
& PageSize = 50
```

```
&< Public   Request   Parameters >
```

Response example

XML format:

```
< DescribeSc  alingActiv  itiesRespo  nse >
    < RequestId > 8FAAE99E - EB43 - 4838 - 85AD - 93F6245490  4C </
 RequestId >
    < TotalCount > 1 </ TotalCount >
    < PageNumber > 1 </ PageNumber >
    < PageSize > 10 </ PageSize >
    < ScalingAct  ivities >
        < ScalingAct  ivity >
            < Cause > A   scheduled   task   excuete   scaling   rule
 " srtest ",  changing   the   Total   Capacity   from  " 0 "  to  " 1
 ". </ Cause >
            < Descriptio  n > Add  " 1 "  ECS   instance </
 Descriptio  n >
            < EndTime > 2014 - 08 - 17T12 : 39Z </ EndTime >
            < Progress > 100 </ Progress >
            < ScalingAct  ivityId > ebta5WbUzC  8gcwUWvfch  yT4U </
 ScalingAct  ivityId >
            < ScalingGro  upId > AG6CQdPU8O  KdwLjgZcJ2  eaQ </
 ScalingGro  upId >
            < StartTime > 2014 - 08 - 17T12 : 39Z </ StartTime >
            < StatusCode > Successful </ StatusCode >
            < StatusMess  age >" 1 "  ECS   instances   is   added .
 </ StatusMess  age >
        </ ScalingAct  ivity >
    </ ScalingAct  ivities >
</ DescribeSc  alingActiv  itiesRespo  nse >
```

JSON format:

```
{
" RequestId ":" 0A016AD6 - A91F - 4210 - 9DEE - D5BBD4270F  45 ",
" TotalCount ": 2582 ,
" PageNumber ": 1 ,
" PageSize ": 1 ,
" ScalingAct  ivities ":{
" ScalingAct  ivity ":[
{
   " ScalingAct  ivityId ":" bBxOR5dhKF  occGkbTrcy  HE2g ",
   " StartTime ":" 2014 - 08 - 18T20 : 49Z ",
" EndTime ":" 2014 - 08 - 18T20 : 49Z ",
" Cause ":"  A   scheduled   task   excuete   scaling   rule  \"
 srtest \",  changing   the   Total   Capacity   from  \" 0 \"  to  \"
 1 \".",
" Descriptio  n ":"  Add  \" 1 \"  ECS   instance  ",
" Progress ": 100 ,
" ScalingGro  upId ":" c7acXJbAJm  pPcGE7G3bw  wbS9 ",
" StatusCode ":" Successful ",
" StatusMess  age ":"\" 1 \"  ECS   instances   is   added . "
}
]
},
```

```
}
```

Error code

For common errors, see *client errors* or *server errors*.

9.9 AttachLoadBalancers

You can call this operation to add one or more SLB instances.

Description

Due to *limits*, the following conditions must be met when you add an SLB instance to a scaling group:

- The SLB instance and the scaling group must be in the same account.
- The SLB instance and the scaling group must be in the same region.
- The SLB instance must be in the Running state.
- The SLB instance must have at least one listener with health check enabled.
- The SLB instance and the scaling group must be in the same VPC if the network type of both of them is VPC.
- If the network type of the scaling group is VPC, the network type of the SLB instance is classic network, and the SLB back-end server contains a VPC instance, the VPC instance and the scaling group must be in the same VPC.
- The number of added SLB instances must be less than the quota of instances in the scaling group.

Debugging

OpenAPI Explorer simplifies API usage. You can use OpenAPI Explorer to perform operations such as retrieve APIs, call APIs, and dynamically generate SDK example code.

Request parameters

Parameter	Type	Required	Example	Description
LoadBalancer . N	RepeatList	Yes	lb-2zeur05gfs****	The IDs of SLB instances. Up to five SLB instances can be added at a time.
ScalingGroupId	String	Yes	AG6CQdPU8OKdwLjgZcJ****	The ID of the scaling group.

Parameter	Type	Required	Example	Description
Action	String	No	AttachLoad Balancers	The operation that you want to perform. Set the value to AttachLoadBalancers.
ForceAttach	Boolean	No	false	Indicates whether to add all instances in the current scaling group to the SLB back-end server: · true: Add all instances to the back-end server. · false: Do not add any instance to the back-end server. Default value: false.

Response parameters

Parameter	Type	Example	Description
RequestId	String	473469C7-AA6F-4DC5-B3DB-A3DC0DE3C83E	The ID of the request. This parameter is returned regardless of whether the operation is successful.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = AttachLoad  Balancers
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ ****
& LoadBalanc  er .  1 = lb - 2zeur05gfs ****
&< Common   request   parameters >
```

Successful response examples

`XML` format

```
< AttachLoad  BalancersR  esponse >
  < RequestId > 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8  3E </
 RequestId >
```

```
</ AttachLoad  BalancersR  esponse >
```

`JSON` format

```
{
 " RequestId ":" 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8  3E "
}
```

Error codes

View error codes

HTTP status code	Error code	Error message	Description
403	Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	The error message returned when Auto Scaling is not authorized to call the specified operation.
404	InvalidScalingGroupId. NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist in the current account.
400	QuotaExceeded. LoadBalancer	LoadBalancer quota exceeded in the scaling group "%s".	The error message returned when the number of SLB instances in the scaling group exceeds the quota.
404	InvalidLoadBalancerId. NotFound	The load balancer "%s" does not exist.	The error message returned when the specified SLB instance does not exist.

HTTP status code	Error code	Error message	Description
400	InvalidLoadBalancerId.RegionMismatch	The load balancer "%s" and the specified scaling group are not in the same Region.	The error message returned when the SLB instance and the scaling group are not in the same region.
400	IncorrectLoadBalancerStatus	The current status of the load balancer "%s" does not support this action.	The error message returned when the SLB instance is in a state that does not support the current operation.
400	IncorrectLoadBalancerHealthCheck	The current health check type of the load balancer "%s" does not support this action.	The error message returned when health check is not enabled for the specified SLB instance.
400	InvalidLoadBalancerId.VPCMismatch	The specified virtual switch and the instance in the load balancer "%s" are not in the same VPC.	The error message returned when the SLB instance and the scaling group are not in the same VPC.
400	QuotaExceeded.BackendServer	Backend server quota exceeded in the load balancer "%s".	The error message returned when the number of back-end servers of the SLB instance exceeds the quota.

HTTP status code	Error code	Error message	Description
404	InvalidScalingConfigurationId.NotFound	The specified scaling configuration does not exist.	The error message returned when the specified scaling configuration that is enabled for the current scaling group does not exist.

9.10 DetachLoadBalancers

You can call this operation to remove one or more SLB instances.

Debugging

OpenAPI Explorer simplifies API usage. You can use OpenAPI Explorer to perform operations such as retrieve APIs, call APIs, and dynamically generate SDK example code.

Request parameters

Parameter	Type	Required	Example	Description
`LoadBalancer . N`	RepeatList	Yes	lb-2zeur05gfs****	The IDs of SLB instances. Up to five SLB instances can be removed at a time.
`ScalingGroupId`	String	Yes	AG6CQdPU8OKdwLjgZcJ****	The ID of the scaling group.
`Action`	String	No	DetachLoadBalancers	The operation that you want to perform. Set the value to DetachLoadBalancers.

Parameter	Type	Required	Example	Description
ForceDetach	Boolean	No	false	Indicates whether to remove instances in the current scaling group from the back-end servers of the SLB instance: · true: Remove the instances. · false: Do not remove the instances. Default value: false.

Response parameters

Parameter	Type	Example	Description
RequestId	String	473469C7-AA6F-4DC5-B3DB-A3DC0DE3C83E	The ID of the request. This parameter is returned regardless of whether the operation is successful.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = DetachLoad  Balancers
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ ****
& LoadBalanc  er .  1 = lb - 2zeur05gfs ****
&< Common   request   parameters >
```

Successful response examples

`XML` format

```
< DetachLoad  BalancersR  esponse >
  < RequestId > 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8  3E
</ RequestId >
</ DetachLoad  BalancersR  esponse >
```

`JSON` format

```
{
 " RequestId ":" 473469C7 - AA6F - 4DC5 - B3DB - A3DC0DE3C8  3E "
}
```

Error codes

View error codes

HTTP status code	Error code	Error message	Description
403	Forbidden.Unauthorized	A required authorization for the specified action is not supplied.	The error message returned when Auto Scaling is not authorized to call the specified operation.
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist in the current account.
404	InvalidLoadBalancerId.NotFound	The Load Balancer "%s" does not exist.	The error message returned when the specified SLB instance does not exist in the scaling group.

9.11 DetachDBInstances

You can call this operation to remove one or more RDS instances.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to `DetachDBInstances`.
ScalingGroupId	String	Yes	The ID of the scaling group.
DBInstance.N	String	Yes	The ID of the RDS instance. You can remove up to five RDS instances at one time.
ForceDetach	Boolean	No	Indicates whether to remove the private IP addresses of instances in the scaling group from the whitelist of RDS instances. · `true`: removes private IP addresses · `false`: does not remove private IP addresses Default value: `false`.

Response parameters

Name	Type	Description
RequestId	String	The GUID generated by Alibaba Cloud for the request.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = DetachDBIn  stances
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& DBInstance .  1 = rm - bp12cy3926  1
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< DetachDBIn  stancesRes  ponse >
    < RequestId > DD0309B7 - 2613 - 4792 - 9B86 - 2759066952  53 </
 RequestId >
</ DetachDBIn  stancesRes  ponse >
```

`JSON` format

```
{
    " RequestId ": " DD0309B7 - 2613 - 4792 - 9B86 - 2759066952  53 "
}
```

Error codes

For more information about common errors, see *Client errors* or *Server errors*.

HttpCode	Error code	Error message	Description
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist in the account.
400	InvalidDBInstanceId.NotFound	DB instance "%s" does not exist.	The error message returned when the specified RDS instance does not exist.

9.12 AttachDBInstances

You can call this operation to add one or more RDS instances.

Description

Due to restrictions, the following conditions must be met when you add RDS instances to a scaling group.

- The SLB instance and the scaling group must be in the same account.
- The status of the RDS instance must be Unlocked. For more information, see *RDS usage instructions*.
- The status of the RDS instance must be Running.
- After you add an RDS instance, the number of IP addresses included in the default group of the IP whitelist for an RDS instance cannot exceed 1,000. For more information about an IP whitelist, see *Set the whitelist*.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to `AttachDBInstances`.
ScalingGroupId	String	Yes	The ID of the scaling group.
DBInstance.N	String	Yes	The ID of the RDS instance. You can add up to five RDS instances at one time.

Name	Type	Required	Description
ForceAttach	Boolean	No	Indicates whether to add all private IP addresses of instances in the scaling group to the IP whitelist of the RDS instance. · `true`: adds all private IP addresses. · `false`: does not add all private IP addresses. Default value: `false`.

Response parameters

Name	Type	Description
RequestId	String	The GUID generated by Alibaba Cloud for the request.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = AttachDBIn  stances
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& DBInstance .  1 = rm - bp12cy3926  1
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< AttachDBIn  stancesRes  ponse >
    < RequestId > DD0309B7 - 2613 - 4792 - 9B86 - 2759066952  53 </
 RequestId >
</ AttachDBIn  stancesRes  ponse >
```

`JSON` format

```
{
```

```
    " RequestId ": " DD0309B7 - 2613 - 4792 - 9B86 - 2759066952  53 "
}
```

Error codes

For more information about common errors, see *Client errors* or *Server errors*.

HttpCode	Error code	Error message	Description
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist in your account.
400	QuotaExceeded.RDS	"RDS" quota exceeded.	The error message returned when the number of RDS instances in the scaling group exceeds the quota.
400	InvalidDBInstanceId.NotFound	The specified value of parameter "%s" is not valid.	The error message returned when the specified RDS instance does not exist.
400	IncorrectDBInstanceStatus	The current status of DB instance “XXX” does not support this action.	The specified action failed due to the current status of the instance.
400	QuotaExceeded.DBInstanceSecurityIP	Security IP quota exceeded in DB instance “XXX”.	The error message returned when the number of backend IP addresses in the whitelist of the RDS instance exceeds the quota.
400	InvalidInstanceIds.PrivateIpNotFound	Can not find all private ips of instances in specific scaling group.	The error message returned when failing to obtain the private IP address of the RDS instance in the scaling group.

9.13 AttachVServerGroups

You can call the `AttachVServerGroups` operation to add one or more VServer groups to an SLB instance.

Restrictions

Due to the *restrictions* of SLB, you must make sure that the following prerequisites are met before you add a VServer group to the scaling group:

- The SLB instance and the scaling group must be under the same account.
- The SLB instance and the scaling group must be in the same region.
- The SLB instance must be in the Running status.
- The SLB instance must be configured with at least one listener with health check enabled.
- The SLB instance and the scaling group must be in the same VPC network if the network type of both of them is VPC.
- If the network type of the scaling group is VPC, the network type of the SLB instance is classic, and the backend server group of the SLB instance contains a VPC-connected instance, then the VPC-connected instance and the scaling group must be in the same VPC network.
- The VServer group that you want to add to the scaling group must have been associated with the SLB instance.
- Currently, you can add up to five VServer groups to a scaling group.

Notice:

To add a VServer group to a scaling group, you must specify the SLB instance ID (`LoadBalanc erId`), the VServer group ID (`VServerGro upId`), and the port number of the VServer group (`Port`). If you have added the same VServer group to a scaling group using different port numbers, Auto Scaling considers that multiple VServer groups are added to the scaling group. If multiple VServer groups with the same group ID and port number are specified in the request parameters, only the first VServer group is used. The other VServer groups are ignored.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to `AttachVServerGroups`.
RegionId	String	Yes	The ID of the region where the scaling group is created, for example, `cn-hangzhou` or `cn-shanghai`. For more information, see *Regions and zones*.
ScalingGroupId	String	Yes	The ID of the scaling group.
VServerGroup.N.LoadBalancerId	String	Yes	The ID of the SLB instance that the VServer group is associated with. N represents the SLB instance number. Valid values: 1 to 5.
VServerGroup.N.VServerGroupAttribute.M.VServerGroupId	String	Yes	The ID of the VServer group. N represents the SLB instance number. Valid values: 1 to 5. M represents the VServer group number that is associated with the SLB instance. Valid values: 1 to 5.

Name	Type	Required	Description
VServerGroup.N.VServerGroupAttribute.M.Port	Integer	Yes	The port number that is used by Auto Scaling to add the ECS instances to the VServer group. Valid values: 1 to 65535. N represents the SLB instance number. Valid values: 1 to 5. M represents the VServer group number that is associated with the SLB instance. Valid values: 1 to 5.
VServerGroup.N.VServerGroupAttribute.M.Weight	Integer	No	The weight set for the ECS instances that are added to the VServer group. Value range: 0 to 100. Default value: `50` . N represents the SLB instance number. Valid values: 1 to 5. M represents the VServer group number that is associated with the SLB instance. Valid values: 1 to 5.

Name	Type	Required	Description
ForceAttach	Boolean	No	Indicates whether to add the ECS instances in the scaling group to the newly added VServer group. · true : Add the ECS instances to the VServer group. · false : Do not add the ECS instances to the VServer group. Default value: false .

Response parameters

Name	Type	Description
RequestId	String	The GUID generated by Alibaba Cloud for the request.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = AttachVSer  verGroups
& ScalingGro  upId = asg - xxxxxxxxxx  x
& RegionId = cn - hangzhou
&  Vservergro  up .  1 .  LoadBalanc  erId = lb - xxxxxxx
& VServerGro  up .  1 .  VServerGro  upAttribut  e .  1 .  VServerGro
 upId = rsp - xxxxx
& VServerGro  up .  1 .  VServerGro  upAttribut  e .  1 .  Port = 22
& VServerGro  up .  1 .  VServerGro  upAttribut  e .  1 .  Weight =
 100
&< Common   request   parameters >
```

Sample responses

XML format

```
< AttachVSer  verGroupsR  esponse >
```

```
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ AttachVSer  verGroupsR  esponse >
```

JSON format

```
{
   " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error code	Error message	HTTP status code	Description
Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	403	The error message returned when Auto Scaling is not authorized to call the specified operation.
InvalidScalingGroupId. NotFound	The specified scaling group does not exist.	404	The error message returned when the specified scaling group does not exist under your account.
InvalidLoadBalancerId. NotFound	The load balancer "%s" does not exist.	404	The error message returned when the specified SLB instance does not exist under your account.
InvalidLoadBalancerId. RegionMismatch	The load balancer "%s" and the specified scaling group are not in the same Region.	400	The error message returned when the SLB instance and the scaling group are not in the same region.
IncorrectLoadBalancerStatus	The current status of the load balancer "%s" does not support this action.	400	The error message returned when the specified operation is not supported by the SLB instance.

Error code	Error message	HTTP status code	Description
IncorrectLoadBalancerHealthCheck	The current health check type of the load balancer "%s" does not support this action.	400	The error message returned when health check is not enabled for the specified SLB instance.
InvalidLoadBalancerId.VPCMismatch	The specified virtual switch and the instance in the load balancer "%s" are not in the same VPC.	400	The error message returned when the SLB instance and the scaling group are not in the same VPC network.
InvalidVServerGroupId.ForLoadBalancer	Invalid VServerGroupId For LoadBalancer "%s".	400	The error message returned when the VServer group that is specified by the `VServerGroupId` parameter has not been associated with the specified SLB instance.
QuotaExceeded.VServerGroup	VServerGroup quota exceeded in the specified scaling group.	400	The error message returned when the number of VServer groups that you can add to the scaling group exceeds the upper limit.

9.14 DetachVServerGroups

You call the `DetachVServerGroups` operation to remove one or more VServer groups.

Description

To remove a VServer group from a scaling group, you must specify the SLB instance ID (`LoadBalancerId`), the VServer group ID (`VServerGroupId`), and the port number of the VServer group (`Port`). The VServer group is removed if the VServer group specified by the request parameters matches the VServer group in the scaling

group. Otherwise, the request is ignored and no error message will be returned by the operation.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to `DetachVServerGroups`.
RegionId	String	Yes	The ID of the region where the scaling group is created, for example, `cn-hangzhou` or `cn-shanghai`. For more information, see *Regions and zones*.
ScalingGroupId	String	Yes	The ID of the scaling group.
VServerGroup.N.LoadBalancerId	String	Yes	The ID of the SLB instance that the VServer group is associated with. N indicates the number of the SLB instance. Valid values: 1 to 5.

Name	Type	Required	Description
VServerGroup .N.VServerGro upAttribute.M. VServerGroupId	String	Yes	The ID of the VServer group. N indicates the number of the SLB instance. Valid values: 1 to 5. M indicates the number of the VServer group that is associated with the SLB instance. Valid values: 1 to 5.
VServerGroup .N.VServerGro upAttribute.M.Port	Integer	Yes	The port that is used by Auto Scaling to add the ECS instances to the VServer group. Value range: 1 to 65535. N indicates the number of the SLB instance. Valid values: 1 to 5. M indicates the number of the VServer group that is associated with the SLB instance. Valid values: 1 to 5.

Name	Type	Required	Description
Forcedetach	Boolean	No	Indicates whether to remove the ECS instances from the VServer group in the specified scaling group. · true : Remove the ECS instances. · false : Do not remove the ECS instances. Default value: false .

Response parameters

Name	Type	Description
RequestId	String	The GUID generated by Alibaba Cloud for the request.

Examples

Sample requests

```
 Http :// ess . aliyuncs . com  /?  Action  =  fig
& ScalingGro  upId = asg - xxxxx
& RegionId = cn - hangzhou
& VServerGro  up .  1 .  LoadBalanc  erId = lb - xxxxxxx
& VServerGro  up .  1 .  VServerGro  upAttribut  e .  1 .  VServerGro
 upId = rsp - xxxxx
& VServerGro  up .  1 .  Vservergro  upattribut  e .  1 .  Port = 22
& VServerGro  up .  1 .  VServerGro  upAttribut  e .  2 .  VServerGro
 upId = rsp - xxxxx
& VServerGro  up .  1   VServerGro  upAttribut  e .  2 .  Port = 80
&< Common   request   parameters >
```

Sample responses

XML format

```
< DetachVSer  verGroupsR  esponse >
  < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
```

```
</ DetachVSer  verGroupsR  esponse >
```

JSON format

```
{
    " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error code	Error message	HTTP status code	Description
Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	403	The error message returned when Auto Scaling is not authorized to call the specified operation.
InvalidScalingGroupId. NotFound	The specified scaling group does not exist.	404	The error message returned when the specified scaling group does not exist under your account.
InvalidParameter	The specified value of parameter "%s" is not valid.	400	The error message returned when the parameter is invalid.
MissingParameter	The input parameter "%s" that is mandatory for processing this request is not supplied.	400	The error message returned when the required parameters are not set.

10 Scaling configuration

10.1 DeleteScalingConfiguration

Deletes a specified scaling configuration.

An active scaling configuration in a scaling group cannot be deleted.

If any ECS instance created according to a scaling configuration is still in the scaling group, the scaling configuration cannot be deleted.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface, required. The parameter value is DeleteScalingConfiguration.
ScalingConfigurationId	String	Yes	ID of a scaling configuration.

Response parameters

Public parameters.

Error codes

For common errors, see *client errors* or *server errors*.

Error message	Error code	Description	HTTP status code
The specified scaling configuration does not exist in this account.	InvalidScalingConfigurationId.NotFound	The specified scaling configuration does not exist.	404
The specified scaling configuration is not in inactive state.	IncorrectScalingConfigurationLifecycleState	The current lifecycle state of specified scaling configuration does not support this action.	400

Error message	Error code	Description	HTTP status code
The scaling configuration has an associated ECS instance not deleted yet.	InstanceInUse	You cannot delete a scaling configuration or scaling group while there is an instance associated with it.	400

Request example

```
 http :// ess . aliyuncs . com /? Action = DeleteScal  ingConfigu
 ration
& ScalingCon  figuration  Id = eOs27Kb0oX  vQcUYjEGel  JqUy
&< Public   Request   Parameters >
```

Response example

XML format:

```
< DeleteScal  ingConfigu  rationResp  onse >
    < RequestId > 61D30272 - 7111 - 44D9 - BB45 - FCB55E4A14  10 </
 RequestId >
</ DeleteScal  ingConfigu  rationResp  onse >
```

JSON format:

```
" RequestId ": " 61D30272 - 7111 - 44D9 - BB45 - FCB55E4A14  10 "
```

10.2 CreateScalingConfiguration

This interface creates a scaling configuration.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface, required. The parameter value is CreateScalingConfiguration.
ScalingGroupId	String	Yes	ID of the scaling group of a scaling configuration.

Name	Type	Required	Description
ImageId	String	No	ID of an image file , indicating the image resource selected when an instance is enabled.
ImageName	String	No	Name of an image file. An image name is unique within a region. If ImageId is set, ImageName is ignored. You cannot specify a marketplace image by ImageName.
InstanceType	String	Yes	Resource type of an ECS instance.
SecurityGroupId	String	Yes	ID of the security group to which a newly created instance belongs . Mutual access is allowed between instances in one security group.

Name	Type	Required	Description
ScalingConfigurationName	String	No	Name shown for the scheduled task. The name must contain 2-40 English or Chinese characters, and start with a number, a letter in upper or lower case or a Chinese character. The name can contain numbers, “_ “, “- “ or “.” . The account name in the same scaling group is unique in the same region. If this parameter value is not specified, the default value is ScalingConfigurationId.

Name	Type	Required	Description
InternetChargeType	String	No	Network billing type. Values: · PayByBandwidth: The users are charged by bandwidth and the bandwidth is specified by InternetMaxBandwidthOut. · PayByTraffic: The users are charged by traffic and InternetMaxBandwidthOut only specifies the upper limit of bandwidth. The users need to pay according to the actual traffic. If this parameter value is not specified, the default value is PayByBandwidth in classic network and PayByTraffic in VPC separately.

Name	Type	Required	Description
InternetMaxBandwidthIn	Integer	No	Maximum incoming bandwidth from the public network, measured in Mbps (Mega bit per second). The value range is [1,200]. If this parameter value is not specified, AliyunAPI automatically sets the value to 200 Mbps. This is for free in any cases. The inbound traffic of your instance is not billed.
InternetMaxBandwidthOut	Integer	No	Maximum outgoing bandwidth from the public network, measured in Mbps (Mega bit per second). The value range · for PayByBandwidth is [0,100]. If this parameter value is not specified, AliyunAPI automatically sets the value to 0 Mbps. · The value range for PayByTraffic is [0,100]. If this parameter value is not specified, an error is reported.

Name	Type	Required	Description
IoOptimized	String	No	Whether the instance is I/O optimized or not. For non I/O optimized instance, IoOptimized is not an available parameter. Optional values: · optimized: The specified instance is I/O optimized. · When IoOptimized is not available, - instances of generation I instance types are non I/O optimized by default. - instances of non generation I instance types are I/O optimized by default.

Name	Type	Required	Description
SystemDisk.Category	String	No	Category of the system disk. Optional values: · cloud: Common cloud disk · cloud_efficiency: Efficient cloud disk · cloud_ssd: SSD cloud disk · ephemeral_ssd: Local SSD disk The default value is cloud for instances of generation I instance types (non I/O optimized) and cloud_efficiency for other types.
SystemDisk.Size	Integer	No	Size of system disk, in GiB. Optional values: · cloud: 40-500 · cloud_efficiency: 40-500 · cloud_ssd: 40-500 · ephemeral_ssd: 40-500 The default value is {40, ImageSize}. If SystemDisk.Size is set, the system disk size must be greater than or equal to max{40, ImageSize}.

Name	Type	Required	Description
DataDisk.n.Category	String	No	Category of data disk n. The value range of n is [1, 16]. Optional values: · cloud: Common cloud disk. The DeleteWithInstance property of a common cloud disk that is created along with the instance is true. · cloud_efficiency: Efficient cloud disk · cloud_ssd: SSD cloud disk · ephemeral_ssd: Local SSD disk The default value is cloud.

Name	Type	Required	Description
DataDisk.n.Size	Integer	No	Size of data disk n, in GiB. Optional values: · cloud: 5-2,000 · cloud_efficiency: 20-32,768 · cloud_ssd: 20-32,768 · ephemeral_ssd: 5-800 If DataDisk.n.Size is set, the data disk size must be greater than or equal to the snapshot size (use SnapshotId to specify the snapshot).
DataDisk.n. SnapshotId	String	No	Snapshot used for creating the data disk n. If this parameter is specified, the DataDisk.n.Size parameter is neglected, and the size of the created disk is the size of the snapshot. If this snapshot is created before July 15, 2013 (included), the snapshot cannot be called, and InvalidSnapshot. TooOld is returned in Response.

Name	Type	Required	Description
DataDisk.n.DeleteWithInstance	Boolean	No	Whether the data disk will be released along with the instance. Optional values: · true: Release the data disk along with the instance. · false: Retain the data disk when you release the instance. Default value: true DataDisk.n.DeleteWithInstance is valid only for independent cloud disks, for which the parameter DataDisk.n.Category is set to cloud, cloud_efficiency or cloud_ssd. Otherwise, an error returns.
LoadBalancerWeight	Integer	No	The weight of the backend server, the value range is [0, 100] and the default value is 50.
UserData	String	No	The user-defined data of the instance. The UserData of an instance must be encoded in Base64 format. The maximum size of the raw data is 16 KB.

Name	Type	Required	Description
KeyPairName	String	No	Key pair name. · If a Windows ECS instance is being created, ignore this parameter. By default, no value is set. · The user password authentication method will be disabled during the initialization of a Linux instance.
RamRoleName	String	No	The name of the instance RAM role. You can query the name of an instance RAM role by using the RAM API *ListRoles*. You can also see the API *CreateRole* for more information.

Name	Type	Required	Description
Tags	String	No	The tags of an instance. You should input the information of the tag with the format of the Key-Value, such as { “key1” :” value1” ,” key2” :” … “key5” :” value5” }. At most 5 tags can be specified. Key: · It can be up to 64 characters in length. It cannot begin with aliyun . It cannot begin with http:// or https://. It cannot be a null string. · Value: It can be up to 128 characters in length. It cannot begin with aliyun . It cannot begin with http:// or https://. It can be a null string.

Name	Type	Required	Description
InstanceTypes.N	String	No	Multiple instance types InstanceType will be ignored if InstanceTypes.N is applied. N ranges from 1 to 10, which means one scaling configuration can contain up to 10 instance types. N is the priority of each instance type in your scaling configuration and 1 means the highest. In other words, the greater the number is, the lower the priority of instance type is. If you cannot launch an instance from the instance type with the relatively higher priority, the system will automatically use the instance type a level lower.
InstanceName	String	No	The name of the instance launched from the current scaling configurat ion.

Name	Type	Required	Description
HostName	String	No	Host name of the ECS instance. It cannot start or end with a period (.) or a hyphen (-). It cannot have two or more consecutive periods (.) or hyphens (-). Naming requirements: · For Windows instances: It can be [2, 15] characters in length. It can contain uppercase or lowercase letters, digits, periods (.), and hyphens (-). It cannot be all digits. · For other instances, such as Linux instances: It can be [2, 64] characters in length. It can be segments separated by periods (.). It can contain uppercase or lowercase letters, digits, and hyphens (-).

Name	Type	Required	Description
PasswordInherit	Boolean	No	Whether to use the password pre-configured in the image you select or not. For a secure access, make sure that the selected image has password configured.
SpotStrategy	String	No	The spot price you are willing to accept for a preemptible instance. It takes effect only when parameter InstanceChargeType is PostPaid. Optional values: · NoSpot: A normal Pay-As-You-Go instance. · SpotWithPriceLimit: Sets the price threshold for a preemptible instance. · SpotAsPriceGo: A price that is based on the highest Pay-As-You-Go instance. Default value: NoSpot

Name	Type	Required	Description
SpotPriceLimit.N.InstanceType	String	No	The instance type for a preemptible instance (N ranges from 1 to 10), and it takes effect only when parameter SpotStrategy is SpotWithPriceLimit.
SpotPriceLimit.N.PriceLimit	Float	No	The hourly price threshold for a preemptible instance (N ranges from 1 to 10), and it takes effect only when parameter SpotStrategy is SpotWithPriceLimit.

Response parameters

Name	Type	Description
ScalingConfigurationId	String	ID of a scaling configuration. It is generated by the system and is globally unique.

Request example

```
 http :// ess . aliyuncs . com /? Action = CreateScal  ingConfigu
 ration
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& SecurityGr  oupId = sg - 280ih3w4b
& ImageId = centos6u5_  64_20G_ali  aegis_2014  0703 . vhd
& InstanceTy  pe = ecs . t1 . xsmall
&< Public   Request   Parameters >
```

Response example

XML format:

```
< CreateScal  ingConfigu  rationResp  onse >
    < ScalingCon  figuration  Id > eOs27Kb0oX  vQcUYjEGel  JqUy </
 ScalingCon  figuration  Id >
```

```
    < RequestId > 5CC0AD41 - 08ED - 4559 - A683 - 6F56355FE0  68 </
 RequestId >
</ CreateScal  ingConfigu  rationResp  onse >
```

JSON format:

```
{
    " RequestId ": " 5CC0AD41 - 08ED - 4559 - A683 - 6F56355FE0  68 ",
    " ScalingCon  figuration  Id ": " eOs27Kb0oX  vQcUYjEGel  JqUy ",
}
```

Error codes

For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InstanceType.Mismatch	The specified scaling configuration and existing active scaling configuration have different instance type.	400	The specified scaling configuration and the existing scaling configuration have different instance types.
InvalidDataDiskSnapshotId.NotFound	Snapshot “XXX” does not exist.	404	Specified snapshot does not exist.
InvalidDataDiskSnapshotId.SizeNotSupported	The capacity of snapshot “XXX” exceeds the size limit of the specified disk category.	400	Capacity of the specified snapshot exceeds the upper limit of the disk size.
InvalidDevice.InUse	Device “XXX” has been occupied.	403	Data disk attaching point has been occupied.
InvalidImageId.InstanceTypeMismatch	The specified image does not support the specified instance type.	400	Specified image does not support the specified instance type.
InvalidImageId.NotFound	The specified image does not exist.	404	Specified image is not in this account.

Error code	Error message	HTTP status code	Description
InvalidKeyPairName.NotFound	The specified KeyPairName does not exist in our records.	400	The specified KeyPairName does not exist.
InvalidNetworkType.ForRAMRole	RAMRole can't be used For classic instance.	400	The RamRoleName is only applicable to VPC instances.
InvalidParameter	The specified value of parameter `KeyPairNam e` is not valid.	400	The parameter `KeyPairNam e` is only valid for a Windows instance.
InvalidParameter.Conflict	The value of parameter `SystemDisk . Category` and parameter `DataDisk . N . Category` are conflict.	400	Type of the specified system disk conflicts with that of the data disk .
InvalidRamRole.NotFound	The specified RamRoleName does not exist.	400	The specified RAM role name does not exist.
InvalidScalingConfigurationName.Duplicate	The specified value of parameter `ScalingCon figuration Name` is duplicated.	400	The scaling configuration name already exists.
InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	404	The specified scaling group does not exist in this account.
InvalidSecurityGroupId.IncorrectNetworkType	The network type of specified security Group does not support this action.	400	Specified network type is inconsiste nt for the specified security group and the scaling group.
InvalidSecurityGroupId.NotFound	The specified security group does not exist.	404	Specified security group is not in this account.

Error code	Error message	HTTP status code	Description
InvalidSecurityGroupId.VPCMismatch	The specified security group and the specified virtual switch are not in the same VPC.	400	The specified security group and the virtual switch are not in the same VPC.
InvalidSnapshot.TooOld	This operation is denied because the specified snapshot is created before 2013-07-15.	403	The snapshot is created before July 15, 2013 (included), and thus cannot be called.
InvalidSystemDiskCategory.ValueUnauthorized	The system disk category is not authorized.	403	You are unauthorized to create an ephemeral system disk.
InvalidUserData.Base64FormatInvalid	The specified parameter UserData must be base64 encoded.	400	The specified UserData should be encoded in Base64 format.
InvalidUserData.SizeExceeded	The specified parameter UserData exceeds the size.	400	The size of the specified UserData exceeds 16 KB.
QuotaExceeded.EphemeralDiskSize	Ephemeral disk size quota exceeded.	403	Ephemeral disk capacity exceeds 2 TB (2,048 GB).
QuotaExceeded.ScalingConfiguration	Scaling configuration quota exceeded in the specified scaling group.	400	Scaling configuration quantity exceeds the upper limit for a user to use.
QuotaExceeded.SecurityGroupInstance	Instance quota exceeded in the specified security group.	400	The number of ECS instances attached to the specified security group exceeds the upper limit.

10.3 DescribeScalingConfigurations

This operation queries scaling configuration information.

Description

You can query all scaling configurations in a scaling group by specifying the scaling group ID.

Scaling configuration states (LifecycleState) can be set to either of the following:

- **Active: Scaling groups use scaling configurations in the active status to automatica lly create ECS instances.**
- **Inactive: The scaling configuration is inactive in a scaling group. The scaling group will not use scaling configurations in inactive state to automatically create ECS instances.**

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface, required. The parameter value is `DescribeScalingConfigurations`.
RegionId	String	Yes	Region ID of the scaling group of a scaling configuration.
ScalingGroupId	String	No	Scaling group ID.

Name	Type	Required	Description
Scalingconfigurationid. N	String	No	ID of a scaling configuration. A maximum of 10 values can be entered. IDs of active and inactive scaling configurations are displayed in the query result, and can be differentiated by `LifecycleState`.
ScalingConfigurationName.N	String	No	Name of a scaling configuration. A maximum of 10 values can be entered. Invalid scaling configuration names are neglected in the query result and no error is reported.
PageNumber	Integer	No	Page number of the scaling configuration list. The initial value and default value are both 1.
PageSize	Integer	No	When querying by page, this parameter indicates the number of lines per page. Maximum value: 50; default value: 10.

Response parameters

Name	Type	Description
TotalCount	Integer	Total number of scaling configurations
PageNumber	Integer	Current page number
PageSize	Integer	Number of lines per page
ScalingConfigurations	ScalingConfiguration SetType	Scaling configuration information set

ScalingConfigurationSetType is a set of ScalingConfigurationItemTypes.

Name	Type	Description
ScalingConfiguration	ScalingConfiguration ItemType	Scaling configuration information

Attributes of the ScalingConfigurationItemType are listed as follows.

Name	Type	Description
ScalingConfigurationId	String	ID of a scaling configurat ion
ScalingConfigurationName	String	Name shown for a scaling configuration
ScalingGroupId	String	ID of the scaling group of a scaling configuration
ImageId	String	ID of an image file
InstanceType	String	Resource rule of an instance
SecurityGroupId	String	ID of a security group
InternetChargeType	String	Network billing type
InternetMaxBandwidthIn	Integer	Maximum incoming bandwidth from the public network, measured in Mbps (Mega bit per second)

Name	Type	Description
InternetMaxBandwidthOut	Integer	Maximum outgoing bandwidth from the public network, measured in Mbps (Mega bit per second)
SystemDisk.Category	String	Category of the system disk
DataDisks	DataDiskSetType	Data disk information set
LifecycleState	String	Scaling configuration status in a scaling group
CreationTime	String	Time when a scaling configuration is created

DataDiskSetType is a set of DataDiskItemTypes.

Name	Type	Description
DataDisk	DataDiskItemType	Data disk information

Attributes of the DataDiskItemType are listed as follows.

Name	Type	Description
Size	Integer	Disk capacity
Category	String	Disk category
SnapshotId	String	ID of the snapshot used for creating the data disk
Device	String	Data disk attaching point

Request example

```
 http :// ess . aliyuncs . com /? Action = CreateScal  ingConfigu
 ration
& RegionId = cn - qingdao
& PageSize = 50
&< Public   Request   Parameters >
```

Response example

XML format:

```
< DescribeSc  alingConfi  gurationsR  esponse >
    < RequestId > 804F240A - 8D3E - 40A1 - BD68 - 6B333DEA2C  A8 </
 RequestId >
```

```
    < TotalCount > 1 </ TotalCount >
    < PageNumber > 1 </ PageNumber >
    < PageSize > 50 </ PageSize >
    < ScalingCon  figuration  s >
        < ScalingCon  figuration >
            < CreationTi  me > 2014 - 08 - 14T10 : 58Z </ CreationTi
 me >
            < ImageId > centos6u5_  64_20G_ali  aegis_2014  0703 . vhd
 </ ImageId >
            < InstanceTy  pe > ecs . t1 . xsmall </ InstanceTy  pe >
            < InternetCh  argeType > PayByBandw  idth </ InternetCh
 argeType >
            < InternetMa  xBandwidth  In > 200 </ InternetMa
 xBandwidth  In >
            < InternetMa  xBandwidth  Out > 0 </ InternetMa
 xBandwidth  Out >
            < LifecycleS  tate > Active </ LifecycleS  tate >
            < ScalingCon  figuration  Id > bU5uZHcAgt  zwcL4IeDea
 vqTS </ ScalingCon  figuration  Id >
            < ScalingCon  figuration  Name > c1908dd1 - 690f - 4c9b -
 ab73 - 350f1f06e8  4f </ ScalingCon  figuration  Name >
            < ScalingGro  upId > dE9YbOdCHq  aFdFZHXVdD  jQCB </
 ScalingGro  upId >
            < SecurityGr  oupId > sg - 280ih3w4b </ SecurityGr  oupId
 >
            < SystemDisk  Category > cloud </ SystemDisk  Category >
            < DataDisks >
            < DataDisk >
                < Size > 200 </ Size >
                < Category > cloud </ Category >
                < SnapshotId > s - 280s7ngih </ SnapshotId >
                < Device >/ dev / xvdb </ Device >
            </ DataDisk >
            </ DataDisks >
        </ ScalingCon  figuration >
    </ ScalingCon  figuration  s >
</ DescribeSc  alingConfi  gurationsR  esponse >
```

JSON format:

```
{
    " RequestId ": " 67E4324F - CE14 - 4C2C - 9D60 - 5422641DB7  6F ",
    " TotalCount ":  1 ,
    " PageNumber ":  1 ,
    " PageSize ":  1 ,
    " ScalingCon  figuration  s ": {
    " ScalingCon  figuration ": [
      {
        " ScalingCon  figuration  Id ": " eqkz17cfW3  clcPExOtLN  VlD
 ",
        " SecurityGr  oupId ": " sg - 28oewzxvg ",
        " CreationTi  me ": " 2014 - 08 - 18T21 : 07Z ",
        " SystemDisk  Category ": " cloud ",
        " InternetMa  xBandwidth  In ":  200 ,
        " InternetMa  xBandwidth  Out ":  0 ,
        " LifecycleS  tate ": " Inactive ",
        " InternetCh  argeType ": " PayByBandw  idth ",
        " ImageId ": " rhel5u7_64  _20G_aliae  gis_201312  31 . vhd ",
        " InstanceTy  pe ": " ecs . s2 . small ",
        " ScalingCon  figuration  Name ": " LxVdcOqPBV ",
        " ScalingGro  upId ": " dRsEAGdvbj  R5c4SVc2bq  Lubj ",
        “ DataDisks ”:{
              “ DataDisk ”:[{
```

```
                “ Size ”: 200 ,
                “ Category ”:  cloud ,
                “ SnapshotId ”:  s - 280s7ngih ,
                “ Device ”:” / dev / xvdb ”
               }]
          }
       }
    ]
  }
}
```

Error code

For common errors, see *client errors* or *server errors*.

10.4 ModifyScalingConfiguration

You can call this operation to modify a specified scaling configuration. When you modify the name of a scaling configuration, ensure that the name is different from the existing names of scaling configurations in a scaling group.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to `ModifyScalingConfiguration`.

Name	Type	Required	Description
ScalingConfigurationName	String	No	The name of the scaling configuration. The name is a combination of 2 to 40 characters including Chinese characters, numbers, uppercase letters, lowercase letters, underscores (_), hyphens (-), and periods (.). Additionally, the name must start with a number, letter, or Chinese character. The name must be unique in a scaling group that is located in a region of an account. If you do not specify the parameter, the ID of the scaling configuration is specified as the name by default.
InstanceName	String	No	The name of the ECS instance that is created based on the current scaling configuration.

Name	Type	Required	Description
HostName	String	No	The hostname of the ECS instance. The name cannot start or end with a period (.) or hyphen (-). Additionally, a period cannot be next another period and a hyphen cannot be next to another hyphen . The naming rules for various types of instances are described as follows: · For Windows instances: The length of a hostname ranges from 2 to 15 characters. The hostname is a combination of characters including uppercase and lowercase letters , numbers, and hyphens (-), excluding periods (.). Additionally, the hostname cannot consist of all numbers. · For other instances, such as Linux instances: The length of a hostname ranges

Name	Type	Required	Description
ImageId	String	No	The ID of the image file. Specifies the image resource for an instance when you start the instance.
ImageName	String	No	The name of the image file. The name must be unique in a region. If you specify the `ImageId`, the `ImageName` is not used. If you specify the `ImageName` rather than the `ImageId`, the `ImageName` is used by the current scaling configuration to identify an image file.
ScalingConfigurationId	String	Yes	The ID of the scaling configuration. Specifies the scaling configuration to be modified.

Name	Type	Required	Description
InternetChargeType	String	No	The billing method. Valid values: · PayByBandwidth: indicates that you pay by bandwidth. In this case, the InternetMaxBandwidthOut is the specified bandwidth. · PayByTraffic: indicates that you pay by traffic. In this case, the InternetMaxBandwidthOut is the upper limit of the bandwidth. You pay for the amount of data that you use. If you do not specify the parameter, the default value is PayByBandwidth in a classic network, while the default value is PayByTraffic in a VPC network.

Name	Type	Required	Description
InternetMaxBandwidthOut	Integer	No	The maximum outgoing bandwidth for the public network. Unit: Mbit/s. Valid values: · PayByBandwidth: 1 to 100. If you do not specify the parameter, the outgoing bandwidth is set to 0 Mbit/s · PayByTraffic: 1 to 100. If you do not specify the parameter, an error occurs.

Name	Type	Required	Description
SystemDisk.Category	String	No	The types of the system disk. Valid values: · cloud : basic disks. · cloud_efficiency : ultra disks. · cloud_ssd : SSDs. · ephemeral_ssd : local SSDs. If the specification of the InstanceType is series I and the instance is not I/O optimized, the default value is cloud . Otherwise, the default value is cloud_efficiency .

Name	Type	Required	Description
SystemDisk.Size	Integer	No	The size of the system disk. Unit: GB. Valid values: · cloud : 40 to 500. · cloud_efficiency : 40 to 500. · cloud_ssd : 40 to 500. · ephemeral_ssd : 40 to 500. Default value: max{40, ImageSize}. After you specify the parameter, the size of the system disk must be equal to or greater than max{40, ImageSize}.
LoadBalancerWeight	Integer	No	The weight of the backend server. Valid values: 0 to 100. Default value: 50.
UserData	String	No	The custom data of the instance. The data is encoded by Base64. The size of the initial data is 16 KB.

Name	Type	Required	Description
KeyPairName	String	No	The name of the key pair. · For Windows ECS instances, the parameter is empty by default and is not used. · For Linux ECS instances, the logon method with a password is disabled by the initialization process.
RamRoleName	String	No	The RAM role of the instance. The role is provided and maintained by RAM. You can call the *ListRoles* operation to query roles. For more information, see *CreateRole*.

Name	Type	Required	Description
InstanceTypes.N	String	No	Includes one or more types of instances specifications. If you specify the `InstanceTypes.N`, the `InstanceType` is not used. Valid values of N: 1 to 10. Namely, you can specify up to 10 instance specifications in a scaling configuration. N is a number that represents the priority of an instance specification in the current scaling configuration. The instance specification with the number 1 indicates the highest priority. The priority decreases along with the increasing number. When an instance cannot be created based on an instance specification with a high priority, Auto Scaling creates the instance using an instance specification with the next highest level of priority.

Name	Type	Required	Description
Tags	String	No	The tags of the instance. A tag is a key-value pair. You can specify up to five sets of tags. The format is: { “key1” :” value1” ,” key2” :” value2” , … “key5” :” value5” }. Keys and values are required to follow these rules: · The length of a key is up to 64 characters. A key cannot start with a string including aliyun, http://, and https://. You cannot specify an empty string for a key. · The length of a value is up to 128 characters. A value cannot start with a string including aliyun, http://, and https://. You can specify an empty string for a value.

Name	Type	Required	Description
PasswordInherit	Boolean	No	Indicates whether you can use a password pre-defined by an image. When you specify the parameter, ensure that the selected image has a pre-defined password.
IoOptimized	String	No	Indicates whether the instance is I/O optimized. Valid values: · none : not I/O optimized. · optimized : I/O optimized. When the instance specification defined by the InstanceType parameter is in out-of-sale, the default value is none . When the instance specification specified by InstanceType is not series I, the default value is optimized .

Name	Type	Required	Description
SpotStrategy	String	No	The preemptible type of postpaid instances. When the value of `InstanceChargeType` is `PostPaid`, the parameter is valid. Valid values: · `NoSpot`: normal Pay-As-You-Go instances. · `SpotWithPriceLimit`: preemptible instances that are specified with the maximum price. · `SpotAsPriceGo`: the price is automatically set based on the current market price. Default value: `NoSpot`.

Name	Type	Required	Description
SpotPriceLimit.N.InstanceType	String	No	The instance specification of the preemptible instance. Valid values of N: 1 to 10. When the value of `SpotStrategy` is `SpotWithPriceLimit`, the parameter is valid.
SpotPriceLimit.N.PriceLimit	Float	No	The corresponding price of the preemptible instance. Valid values of N: 1 to 10. When the value of `SpotStrategy` is `SpotWithPriceLimit`, the parameter is valid.

Name	Type	Required	Description
DataDisk.N.Category	String	No	The disk type of the data disk N. Valid values: · cloud : basic disks. For a basic disk that is created along with the creation of an instance, the value of the disk's DeleteWithInstance property is true. · cloud_efficiency : ultra disks. · cloud_ssd : SSDs. · ephemeral_ssd : local SSDs. Default value: cloud .

Name	Type	Required	Description
DataDisk.N.Size	Integer	No	The size of the data disk N. Unit: GB. Valid values of N: 1 to 16. Valid values: · cloud : 5 to 2000. · cloud_efficiency : 20 to 32768. · cloud_ssd : 20 to 32768. · ephemeral_ssd : 50 to 800. After you specify the parameter, the size of the disk must be equal to or greater than the size of the snapshot. The snapshot is specified by the SnapshotId parameter.

Name	Type	Required	Description
DataDisk.N.SnapshotId	String	No	The snapshot that is used to create the data disk. Valid values of N: 1 to 4. After you specify the parameter, the `DataDisk.N.Size` parameter is not used. The actual size of the created disk is equal to the size of the specified snapshot. If the snapshot was created on or before Jul 15, 2013, this call will be rejected. Up to four pieces of `InvalidSnapshot.TooOld` can be included in a response parameter.
DataDisk.N.Device	String	No	The mount point of the data disk. Valid values of N: 1 to 4. If you do not specify the parameter, ECS allocates a mount point to an automatically created ECS instance by default. The name of the mount point ranges from */dev/xvdb* to */dev/xvdz* in alphabetical order.

Name	Type	Required	Description
DataDisk.n.DeleteWithInstance	Boolean	No	Indicates whether the data disk is released along with the instance. Valid values: · true : When you release the instance, the data disk will also be released. · false : When you release the instance, the data disk will not be released. Default value: true . You can only specify the parameter for independent disks, which means that the value of the DataDisk . n . Category parameter is cloud , cloud_efficiency , or cloud_ssd . Otherwise, an error occurs.

Response parameters

Name	Type	Description
RequestId	String	The GUID generated by Alibaba Cloud for the request.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = ModifyScal  ingConfigu
 ration
& ScalingCon  figuration  Id = asc - xxxxxxxxxx  xxxx
& ScalingCon  figuration  Name = test - modify
& ImageId = centos6u5_  64_20G_ali  aegis_2014  0703 . vhd
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< ModifyScal  ingConfigu  rationResp  onse >
  < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ ModifyScal  ingConfigu  rationResp  onse >
```

`JSON` format

```
{
    " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

For more information about common errors, see *Client errors* or *Server errors*.

HttpCode	Error code	Error message	Description
403	Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	The error message returned when you have not been authorized to perform the `Action`.
404	InvalidDat aDiskSnapshotId. NotFound	Snapshot “XXX” does not exist.	The error message returned when the specified snapshot does not exist.

HttpCode	Error code	Error message	Description
400	InvalidDataDiskSnapshotId.SizeNotSupported	The capacity of snapshot “XXX” exceeds the size limit of the specified disk category.	The error message returned when the size of the snapshot exceeds the maximum size of the disk.
404	InvalidImageId.NotFound	The specified image does not exist.	The error message returned when the specified ImageId does not exist.
400t	InvalidKeyPairName.NotFound	The specified KeyPairName does not exist in our records.	The error message returned when the specified `KeyPairName` does not exist.
400	InvalidNetworkType.ForRAMRole	RAMRole can’t be used For classic instance.	The error message returned when you specify the `RamRoleName` parameter for an instance in a classic network.
400	InvalidParamter	The specified value of parameter is not valid.	The error message returned when the value of the specified parameter is invalid.
400	InvalidScalingConfigurationName.Duplicate	The specified value of parameter is not valid.	The error message returned when the name of the scaling configuration exists.
400	InvalidSecurityGroupId.IncorrectNetworkType	The network type of specified Security Group does not support this action.	The error message returned when the network type of the specified security group is different from that of the specified scaling group.

HttpCode	Error code	Error message	Description
400	InvalidSecurityGroupId.VPCMismatch	The specified security group and the specified virtual switch are not in the same VPC.	The error message returned when the specified security group and the VSwitch are not in the same VPC.
400	InvalidTags.KeyValue	The specified tags key/value cannot be empty.	The error message returned when the `Tags` parameter is empty.
400	InvalidTags.ListSize	The specified tags list size cannot be more than “5”.	The error message returned when the length of the `Tags` list exceeds the specified length.
400	InvalidUserData.Base64FormatInvalid	The specified parameter UserData must be base64 encoded.	The error message returned when the `UserData` does not conform to Base64 encoding.
400	InvalidUserData.SizeExceeded	The specified parameter UserData exceeds the size.	The error message returned when the length of the `UserData` exceeds the specified size.

11 Instance

11.1 EnterStandby

Move an ECS instance to standby (`EnterStand by`) in a scaling group.

If Server Load Balancer is configured for the scaling group, the load balancing weight of the instance is set to zero.

- When an instance is in standby status, you can manually remove it from the scaling group before releasing it independently.
- When the number of scaling groups changes or a monitoring task triggers an automatic scaling event, instances in standby status are not removed.
- When an instance is in standby status, if it is in an unhealthy status (stopped, restarting, etc.), the instance health check status is not updated and it does not trigger a scaling event to remove unhealthy instances. The health check status is only updated after the instance leaves the standby status (*ExitStandby*).

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: EnterStandby.
ScalingGroupId	String	Yes	The scaling group ID.
InstanceId.N	String	Yes	The ECS instance ID .

Response parameters

Name	Type	Description
RequestId	String	The request ID

Request example

```
 Http :// ess . aliyuncs . com  /?  Action  =  enterstand  by
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& InstanceId .  1 = i - 28wt48iaa
```

```
&< Common   Request   Parameters >
```

Response example

XML format:

```
< EnterStand  byResponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ EnterStand  byResponse >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error code	Error message	HTTP status code	Description
Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	403	The RAM user does not have permission to call this API. Contact the account owner and request this permission before trying again.
InvalidInstanceId. NotFound	Instance “XXX” does not exist.	404	The specified ECS instance does not exist.
InvalidScalingGroupId. NotFound	The specified scaling group does not exist.	404	The specified scaling group does not exist.

11.2 ExitStandby

Removes an ECS instance from standby status in a scaling group (`ExitStandb  y`).

If Server Load Balancer is configured for the scaling group, the load balancing weight of the instance is set to the value defined in the scaling configuration.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: ExitStandby.
ScalingGroupId	String	Yes	The scaling group ID.
InstanceId.N	String	Yes	The ECS instance ID. This value can be a JSON array composed of multiple instance IDs, with the format: ["i-xxxxxxxxx", "i-yyyyyyyyy", ··· "i-zzzzzzzzz"]. This parameter supports up to 20 IDs, separated by commas (,).

Response parameters

Name	Type	Description
RequestId	String	Unique request ID

Request example

```
 http :// ess . aliyuncs . com /? Action = ExitStandb  y
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& InstanceId .  1 = i - 28wt48iaa
&< Common   Request   Parameters >
```

Response example

XML format:

```
< ExitStandb  yResponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
```

```
</ ExitStandb  yResponse >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error code	Error message	HTTP status code	Description
Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	403	The RAM user does not have permission to call this API. Contact the account owner and request this permission before trying again.
InvalidInstanceId. NotFound	A required authorization for the specified action is not supplied.	404	The specified ECS instance does not exist.
InvalidSca lingGroupId. NotFound	The specified scaling group does not exist.	404	The specified scaling group does not exist.

11.3 RebalanceInstances

You can call this operation to rebalance the distribution of ECS instances in a scaling group that is specified with multiple zones (`RebalanceI  nstances`).

Description

You can achieve the rebalance of zones by creating new ECS instances to replace the previous ECS instances. Newly created ECS instances are started first before stopping previous ECS instances. The rebalance operation will not affect the performance and availability of your applications.

- You can perform the rebalance operation for ECS instances in a scaling group that is specified with multiple zones only when you set `MultiAZPol  ic = Balance` .
- You can only perform the rebalance operation when the instances in the scaling group are significantly imbalanced.

- A single rebalance operation can replace up to 20 ECS instances.
- You are temporarily allowed to exceed 10% of the total capacity of the scaling group. This occurs during a rebalance operation when the number of ECS instances is approaching or reaches the maximum number (specified in `MaxSize`) of a scaling group and the operation is still running. The minimum number is one ECS instance. Generally, the duration of a rebalance operation from the status of exceeding the maximum capacity of a scaling group to the normal status is approximately one to six minutes.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to RebalanceInstances.
ScalingGroupId	String	Yes	The ID of the scaling group.

Response parameters

Name	Type	Description
RequestId	String	The request ID.
ScalingActivityId	String	The ID of the scaling activity.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = RebalanceI  nstances
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< RebalanceI  nstancesRe  sponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
    < ScalingAct  ivityId > asa - kjgffgdfad  ahghda </ ScalingAct
 ivityId >
```

```
</ RebalanceI  nstancesRe  sponse >
```

`JSON` format

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 ",
    " ScalingAct  ivityId ": " asa - kjgffgdfad  ahghda "
}
```

Error codes

The following table lists the error codes that the `RebalanceI  nstances` operation can return. For more information, see *Client errors* and *Server errors*.

HttpCode	Error code	Error message	Description
400	IncorrectS calingGroupStatus	The current status of the specified scaling group does not support this action.	The error message returned when you have not *enabled a scaling group*.
400	OperationDenied	This operation is denied because the specified scaling group does not support this action.	The error message returned when you have not set the `MultiAZPol  ic = Balance` policy in the specified scaling group or the great imbalance issue of the distribution of ECS instances does not occur.
403	Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	The error message returned when you have not been authorized to use the `RebalanceI  nstances` operation.
404	InvalidSca lingGroupId. NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist in this account.

11.4 SetInstancesProtection

You can call this operation to enable or disable the protection for one or more ECS instances in a scaling group (`SetInstanc esProtecti on`).

Description

After the protection is enabled for an ECS instance:

- The instance keeps running in the protected state until you disable the protection.
- The changes in the number of instances in a scaling group and scale-in activities triggered by monitoring tasks will not remove protected ECS instances. You must manually *remove an ECS instance* to release the instance.
- When an ECS instance is stopped or restarted, the health check status of the ECS instance is not updated.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to SetInstanc esProtection.
ScalingGroupId	String	Yes	The ID of the scaling group.
InstanceId.N	String	Yes	The ID of the ECS instance. Valid values of `N` : 1 to 20.

Name	Type	Required	Description
ProtectedFromScaleIn	Boolean	Yes	Indicates whether the protection for an ECS instance is enabled. Active protection when scale-in activities occur prevents the ECS instance from being stopped or removed from a scaling group. Valid values: · True · False

Response parameters

Name	Type	Description
RequestId	String	The request ID.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = SetInstanc  esProtecti  on
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& InstanceId .  1 = i - 28wt48iaa
& InstanceId .  1 = i - 28wt48ibb
& ProtectedF  romScaleIn = true
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< SetInstanc  esProtecti  onResponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ SetInstanc  esProtecti  onResponse >
```

`JSON` format

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
```

```
}
```

Error codes

The following table lists the error codes that the `SetInstanc esProtecti on` operation can return. For more information, see *Client errors* and *Server errors*.

HttpCode	Error code	Error message	Description
400	IncorrectS calingGroupStatus	The current status of the specified scaling group does not support this action.	The error message returned when you have not *enabled a scaling group*.
403	Forbidden. Unauthorized	A required authorization for the specified action is not supplied.	The error message returned when you are not authorized to use the `SetInstanc esProtecti on` operation.
404	InvalidInstanceId. NotFound	Instance “XXX” does not exist.	The error message returned when the specified ECS instance does not exist.
404	InvalidSca lingGroupId. NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist .

12 Trigger task

12.1 ExecuteScalingRule

Executes a specified scaling rule.

Description

- The interface can be called only when the scaling group is active.
- The interface can be called only when no scaling activity in the scaling group is in progress.
- When no scaling activity in the scaling group is in progress, the interface can be directly executed without cooldown.
- Successfully calling this interface only means that the Auto Scaling service has accepted the call request, and the scaling activity can be executed, but does not necessarily mean that the scaling activity can be successfully executed. You can use the returned ScalingActivityId to check the status of the scaling activity.
- When the total capacity of to-be-attached ECS instances specified by this scaling rule plus the instances in the current scaling group is greater than MaxSize, the rule Total Capacity = MaxSize is applied.
- When the total capacity of instances in the current scaling group minus the number of ECS instances to be reduced according to the scaling rule is smaller than MinSize, the rule Total Capacity = MinSize is applied.
- You can only perform auto scaling for at most 1,000 ECS instances of all scaling groups in all regions. (The number only includes the number of ECS instances automatically created, but does not include the number of manually attached ECS instances.)

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value: ExecuteScalingRule.

Name	Type	Required	Description
ScalingRuleAri	String	Yes	Unique identifier of the scaling rule.
ClientToken	String	No	Guarantees the idempotence of the request. Used to ensure request idempotence. The value is generated by a client. It must be unique among all requests and can contain a maximum of 64 ASCII characters. For more information, see *how to ensure idempotence*.

Response parameters

Name	Type	Description
ScalingActivityId	String	Scaling activity ID.

Request example

```
 http :// ess . aliyuncs . com /? Action = ExecuteSca  lingRule
& ScalingRul  eAri = ari : acs : ess : cn - qingdao : 1344371 :
 scalingRul  e / cCBpdYdQuB  e2cUxOdu6p  iOk
&< Public   Request   Parameters >
```

Response example

XML format:

```
< ExecuteSca  lingRuleRe  sponse >
    < ScalingAct  ivityId > ebta5WbUzC  8gcwUWvfch  yT4U </
 ScalingAct  ivityId >
    < RequestId > 262216B9 - F9D4 - 4D16 - BE9B - BD1C39A4F4  2B </
 RequestId >
</ ExecuteSca  lingRuleRe  sponse >
```

JSON format:

```
{
    " RequestId ": " 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE ",
    " ScalingAct  ivityId ": " ebta5WbUzC  8gcwUWvfch  yT4U "
```

```
}
```

Error codes

For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidScalingRuleAri.NotFound	The specified scaling rule Ari does not exist.	404	The specified scaling rule does not exist in this account.
Forbidden.Unauthorized	A required authorization for the specified action is not supplied.	403	API is not fully authorized to the Auto Scaling service.
IncorrectScalingGroupStatus	The current status of the specified scaling group does not support this action.	400	The specified scaling rule is not active.
ScalingActivityInProgress	You cannot delete a scaling group or launch a new scaling activity while there is a scaling activity in progress for the specified scaling group.	400	The scaling group to which the scaling rule belongs has a scaling activity in progress.
InsufficientBalance	Your account does not have enough balance.	400	Your account balance is not enough.
QuotaExceed.Instance	Living instance quota exceeded.	400	Your ECS instance quota is exceeded.
IncorrectLoadBalancerStatus	The current status of the specified Server Load Balancer instance does not support this action.	400	The Server Load Balancer instance in the scaling group to which the scaling rule belongs is not active.

Error code	Error message	HTTP status code	Description
IncorrectLoadBalancerHealthCheck	The current health check type of specified load balancer does not support this action.	400	Health check is not enabled for the Server Load Balancer in the scaling group to which the specified scaling rule belongs.
InvalidLoadBalancerId.IncorrectInstanceNetworkType	The network type of the instance in specified load balancer does not support this action.	400	The network type of the ECS instance contained in the specified Server Load Balancer is different from the network type of the scaling group.
InvalidLoadBalancerId.VPCMismatch	The specified virtual switch and the instance in specified load balancer are not in the same VPC.	400	The ECS instance contained in the specified Server Load Balancer and VSwitchId are not in the same VPC.
Incorrectdbinstancestatus	The current status of DB instance “XXX” does not support this action.	400	The RDS instance in the scaling group to which the specified scaling rule belongs is not running.
QuotaExceeded.DBInstanceSecurityIP	Security IP quota exceeded in DB instance “XXX”.	400	The number of IP addresses in the white list that can access the RDS instance in the scaling group to which the specified scaling rule belongs has reached the upper limit.

Error code	Error message	HTTP status code	Description
QuotaExceeded.SecurityGroupInstance	Instance quota exceeded in the specified security group.	400	The number of ECS instances attached to the specified security group exceeds the upper limit.
IncorrectCapacity.NoChange	To execute the specified scaling rule, the total capacity will not change.	400	The number of instances in the scaling group does not change after the scaling rule is executed.
QuotaExceeded.ScalingInstance	Scaling instance quota exceeded.	400	The Auto Scaling ECS instance quota is exceeded.
QuotaExceeded.AfterpayInstance	Living afterpay instance quota exceeded.	400	The Pay-As-You-Go ECS instance quota is exceeded.
ResourceNotAvailable.ECS	The specified region or zone does not offer the specified disk or instance category.	400	The specified ECS instance type or disk type cannot be created in the specified region.

12.2 AttachInstances

Attaches an ECS instance to a specified scaling group.

Description

- Restrictions on the attached ECS instance:
  - The attached ECS instance and the scaling group must be in the same region.
  - The attached ECS instance must in the Running status.
  - The attached ECS instance has not been attached to other scaling groups.
  - The attached ECS instance supports Subscription and Pay-As-You-Go payment methods.
  - If the VswitchID is specified for a scaling group, you cannot attach Classic ECS instances or ECS instances on other VPCs to the scaling group.
  - If the VswitchID is not specified for the scaling group, ECS instances of the VPC type cannot be attached to the scaling group.
- The interface can be called only when the scaling group is active.
- The interface can be called only when the scaling group has no scaling activity in progress.
- When the scaling group has no scaling activity in progress, the interface can be directly executed without cooldown.
- Successfully calling this interface only means that the Auto Scaling service has accepted the call request, and the scaling activity can be executed, but does not necessarily mean that the scaling activity can be successfully executed. You can use the returned ScalingActivityId to check the status of the scaling activity.
- The call attempt may fail when the total capacity of instances specified by this interface plus the instances of the scaling group is greater than MaxSize.
- The manually attached ECS instance is not associated with the active scaling configurations of the scaling group.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value: AttachInstances.

Name	Type	Required	Description
ScalingGroupId	String	Yes	Scaling group ID.
InstanceId.N	String	Yes	ECS instance ID. You can input up to 20 IDs.

Response parameter

Name	Type	Description
ScalingActivityId	String	Scaling activity ID

Error codes

For common errors, see *client errors* or *server errors*.

Error message	Error code	Description	HTTP status code
The specified scaling group does not exist in this account.	Invalidscalinggroupid.NotFound	The specified scaling group does not exist.	404
The user has not granted Open API authorization to Auto Scaling	API is not fully authorized to the Auto Scaling service.	A required authorization for the specified action is not supplied.	403
The specified scaling group is not active.	IncorrectScalingGroupStatus	The current status of the specified scaling group does not support this action.	400
The specified ECS instance does not exist in this account.	InvalidInstanceId.NotFound	Instance “XXX” does not exist.	404
The specified ECS instance and the scaling group are not in the same region.	InvalidInstanceId.RegionMismatch	Instance “XXX” and the specified scaling group are not in the same Region.	400

Error message	Error code	Description	HTTP status code
The specified ECS instance and the instance with valid scaling configurations do not match.	InvalidInstanceId.InstanceTypeMismatch	Instance “XXX” and existing active scaling configurations have different instance types.	400
The specified ECS instance is not in the Running status.	IncorrectInstanceStatus	The current status of instance “XXX” does not support this action.	400
The network types of the specified ECS instance and the scaling group do not match.	InvalidInstanceId. NetworkTypeMismatch	The network type of instance “XXX” does not support this action.	400
The specified scaling group and the attached ECS instance are not in the same VPC.	InvalidInstanceId. VPCMismatch	Instance “XXX” and the specified scaling group are not in the same VPC.	400
The specified ECS instance has been attached to another scaling group.	InvalidInstanceId. InUse	Instance “XXX” is already attached to another scaling group.	400
The specified scaling group has an in-progress scaling activity.	ScalingActivityInProgress	You cannot delete a scaling group or launch a new scaling activity while there is a scaling activity in progress for the specified scaling group.	400
The load balancing instance for the specified scaling group is not active	The Server Load Balancer instance of the specified scaling group is not in the Active status.	The current status of the specified Server Load Balancer does not support this action.	400

Error message	Error code	Description	HTTP status code
Health check is not enabled for the Server Load Balancer of the specified scaling group.	IncorrectLoadBalancerHealthCheck	The current health check type of specified load balancer does not support this action.	400
The network type of the ECS instance contained in the specified Server Load Balancer is different from the network type of the scaling group.	InvalidLoadBalancerId.IncorrectInstanceNetworkType	The network type of the instance in specified Server Load Balancer does not support this action.	400
The ECS instance contained in the specified Server Load Balancer and VSwitchId are not in the same VPC.	InvalidLoadBalancerId.VPCMismatch	The specified virtual switch and the instance in specified load balancer are not in the same VPC.	400
The RDS instance in the specified scaling group is not in the Running status.	IncorrectDBInstanceStatus	The current status of DB instance “XXX” does not support this action.	400
The number of IP addresses in the whitelist that can access the RDS instance in the scaling group has exceeded the upper limit.	QuotaExceeded.DBInstanceSecurityIP	Security IP quota exceeded in DB instance “XXX”.	400
The number of ECS instances attached to the specified security group exceeds the upper limit.	QuotaExceeded.SecurityGroupInstance	Instance quota exceeded in the specified security group.	400

Error message	Error code	Description	HTTP status code
Total Capacity after the ECS instance is attached is greater than MaxSize.	IncorrectCapacity.MaxSize	To attach the instances, the total capacity will be greater than the MaxSize.	400

Request example

```
 http :// ess . aliyuncs . com /? Action = AttachInst  ances
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& InstanceId .  1 = i - 28wt48iaa
&< Public   Request   Parameters >
```

Response example

XML format:

```
< AttachInst  ancesRespo  nse >
    < ScalingAct  ivityId > bybj9OcaOT  4ucPMbFhcq  HfA3 </
 ScalingAct  ivityId >
    < RequestId > DD0309B7 - 2613 - 4792 - 9B86 - 2759066952  53 </
 RequestId >
</ AttachInst  ancesRespo  nse >
```

JSON format:

```
" RequestId ": " 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE ",
" ScalingAct  ivityId ": " ebta5WbUzC  8gcwUWvfch  yT4U "
```

12.3 RemoveInstances

Removes an ECS instance from a specified scaling group.

Description

- When the ECS instance automatically created by the Auto Scaling service is removed from the scaling group, the ECS instance is disabled and released.
- When the manually attached ECS instance is removed from the scaling group, the ECS instance is neither disabled nor released.
- The interface can be called only when the scaling group is active.
- The interface can be called only when no scaling activity in the scaling group is in progress.
- When no scaling activity in the scaling group is in progress, the interface can be directly executed without cooldown.

- Successfully calling this interface only means that the Auto Scaling service has accepted the call request, and the scaling activity can be executed, but does not necessarily mean that the scaling activity can be successfully executed. You can use the returned ScalingActivityId to check the status of the scaling activity.
- When the total capacity of instances of the scaling group minus instances specified by this interface is smaller than than MinSize, the call fails.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value: RemoveInstances.
ScalingGroupId	String	Yes	Scaling group ID.
InstanceId.N	String	Yes	ECS instance ID. You can input up to 20 IDs.

Name	Type	Required	Description
RemovePolicy	String	No	Specify the status of instances that are removed from scaling group. Optional values: · `recycle` : The instance enters No fees for stopped instances (VPC-Connected) status. Notice: Only takes effect when `ScalingPolicy` is `recycle` . · `release` : The instance is released. By default, the value is `release` .

Response parameters

Name	Type	Description
ScalingActivityId	String	Scaling activity ID

Error codes

For common errors, see *client errors* or *server errors*.

Error message	Error code	Description	HTTP status code
The specified scaling group does not exist in this account.	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	404

Error message	Error code	Description	HTTP status code
The specified ECS instance does not exist in the scaling group.	InvalidInstanceId.NotFound	Instance "XXX" does not exist.	404
The specified group does not support the specified RemovePolicy.	InvalidParameter	The scaling group does not support the reclaim mode.	400
API is not fully authorized to the Auto Scaling service.	Forbidden.Unauthorized	A required authorization for the specified action is not supplied.	403
The specified scaling group is not active.	IncorrectScalingGroupStatus	The current status of the specified scaling group does not support this action.	400
The specified scaling group has an in-progress scaling activity.	ScalingActivityInProgress	You cannot delete a scaling group or launch a new scaling activity while there is a scaling activity in progress for the specified scaling group.	400
The Server Load Balancer instance in the scaling group to which the scaling rule belongs is not active.	IncorrectLoadBalancerStatus	The current status of the specified load balancer does not support this action.	400
The RDS instance in the scaling group to which the specified scaling rule belongs is not running.	IncorrectDBInstanceStatus	The current status of DB instance "XXX" does not support this action.	400

Error message	Error code	Description	HTTP status code
After instance removal, the total capacity is lower than MinSize.	IncorrectCapacity.MinSize	To remove the instances, the total capacity will be lesser than the MinSize.	400

Request example

```
 http :// ess . aliyuncs . com /? Action = RemoveInst  ances
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& InstanceId .  1 = i - 28wt48iaa
&< Public   Request   Parameters >
```

Response example

XML format:

```
< RemoveInst  ancesRespo  nse >
    < ScalingAct  ivityId > bybj9OcaOT  4ucPMbFhcq  HfA3 </
 ScalingAct  ivityId >
    < RequestId > DD0309B7 - 2613 - 4792 - 9B86 - 2759066952  53 </
 RequestId >
</ RemoveInst  ancesRespo  nse >
```

JSON format:

```
" RequestId ": " 6469DCD0 - 13AC - 487E - 85A0 - CE4922908F  DE ",
" ScalingAct  ivityId ": " ebta5WbUzC  8gcwUWvfch  yT4U "
```

12.4 DetachInstances

You can call this operation to detach one or more ECS instances from a scaling group.

Description

After you detach an ECS instance, the ECS instance can exist independently without attaching a scaling group. You can attach the ECS instance to another scaling group (*AttachInstances*). Detaching an ECS instance will not stop or release the ECS instance. If applicable, you can manually *stop* or *release* an ECS instance.

Note:

- The status of the target scaling group must be Enable (`Enable`).
- Ensure that no in-progress scaling activity exists in the target scaling group.

- When no in-progress scaling activity exists in the target scaling group, you can immediately perform the operation without following the *cool-down time*.
- Successfully calling an operation only means the calling request of the operation is received. You can trigger a scaling activity as usual, but you cannot ensure that the scaling activity is executed. You must view the status of the scaling activity based on the returned result of `ScalingAct ivityId`.
- The remaining number of ECS instances cannot be less than the minimum number of the scaling group (`MinSize`). The remaining number of ECS instances is the total number of ECS instances of the target scaling group minus the number of detached ECS instances.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to DetachInst ances.
ScalingGroupId	String	Yes	The ID of the scaling group.
InstanceId.N	String	Yes	The ID of the ECS instance. Valid values of `N` : 1 to 20.

Response parameters

Name	Type	Description
RequestId	String	The request ID.
ScalingActivityId	String	The ID of the scaling activity.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = DetachInst  ances
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJX  XXXX
& InstanceId .  1 = i - 28wt48iaa
```

```
&< Common   request   parameters >
```

Successful response examples

`XML` format

```
< DetachInst  ancesRespo  nse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
    < ScalingAct  ivityId > asa - xxxxxxxxx </ ScalingAct  ivityId >
</ DetachInst  ancesRespo  nse >
```

`JSON` format

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 ",
    " ScalingAct  ivityId ": " asa - xxxxxxxxx "
}
```

Error codes

HttpCode	Error code	Error message	Description
400	IncorrectScalingGroupStatus	The current status of the specified scaling group does not support this action.	The status of the target scaling group must be Enable `Enable`.
400	ScalingActivityInProgress	You cannot delete a scaling group or launch a new scaling activity while there is a scaling activity in progress.	The error message returned when the specified scaling group is currently undergoing a scaling activity.
400	IncorrectLoadBalancerStatus	The current status of the specified load balancer does not support this action.	The status of SLB instances in the target scaling group must be Running `Running`.
400	IncorrectDBInstanceStatus	The current status of DB instance “XXX” does not support this action.	The status of RDS instances in the target scaling group must be Running `Running`.

HttpCode	Error code	Error message	Description
400	IncorrectCapacity.MinSize	To remove the instances, the total capacity will be lesser than the MinSize.	The remaining number of ECS instances cannot be less than the minimum number of the scaling group (`MinSize`). The remaining number of ECS instances is the total number of ECS instances of the target scaling group minus the number of detached ECS instances.
403	Forbidden.Unauthorized	A required authorization for the specified action is not supplied.	The error message returned when you have not been authorized to use the `DetachInstances` operation.
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist.
404	InvalidInstanceId.NotFound	Instance “XXX” does not exist.	The error message returned when the specified ECS instance does not exist.

13 Scheduled tasks

13.1 DescribeScheduledTasks

Queries information of a scheduled task.

You can use the ScheduledAction to query the relevant scheduled task.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value: DescribeScheduledTasks.
RegionId	String	Yes	Region of the scheduled task.
ScheduledTaskId.N	String	No	ID of the scheduled task. You can enter at most 20 IDs. Invalid scheduled task IDs are not displayed in the query results, and no error is reported.
ScheduledTaskName.N	String	No	Display name of the schedules task. You can enter at most 20 display names. Names of invalid scheduled tasks will be neglected in the query results, and no error is reported.

Name	Type	Required	Description
ScheduledAction.N	String	No	Operations performed when the scheduled task is triggered. You can enter at most 20 operations. Invalid operations are not displayed in the query results, and no error is reported.
PageNumber	Integer	No	Page number of the scheduled task list, starting from 1. The default value is 1.
PageSize	Integer	No	When querying by page, this parameter indicates the number of lines per page. Maximum value: 50; default value: 10.

Response parameters

Name	Type	Description
TotalCount	Integer	Total number of scheduled tasks
PageNumber	Integer	Current page number
PageSize	Integer	Number of lines per page
ScheduledTasks	ScheduledTaskSetType	A set of scheduled task information

ScheduledTaskSetType is a set of ScheduledTaskItemTypes.

Name	Type	Description
ScheduledTask	ScheduledTaskItemType	Information of the scheduled task

The attributes of the ScheduledTaskItemType are listed as follows.

Name	Type	Description
ScheduledTaskId	String	ID of the scheduled task
ScheduledTaskName	String	Name of the scheduled task
Description	String	Description of the scheduled task
ScheduledAction	String	Operations performed when the scheduled task is triggered
LaunchTime	String	Time point at which the scheduled task is triggered
LaunchExpirationTime	Integer	Retry interval for the failed scheduled task
RecurrenceType	String	Type of the scheduled task to be repeated
RecurrenceValue	String	Value of the scheduled task to be repeated
RecurrenceEndTime	String	End time of the scheduled task to be repeated
TaskEnabled	Bool	Whether to enable the scheduled task

Error code

For common errors, see *client errors* or *server errors*.

Request example

```
 http :// ess . aliyuncs . com /? Action = DescribeSc  heduledTas  ks
& RegionId = cn - qingdao
& ScheduledT  askId .  1 = edRtShc57W  GXdt8TlPbr  jsnV
& PageSize = 50
 Public   Request   Parameters
```

Response example

XML format:

```
< DescribeSc  heduledTas  ksResponse >
< RequestId > B9B498DA - E836 - 45FF - 83C7 - 1930492FDD  5A </
 RequestId >
    < TotalCount > 1 </ TotalCount >
```

```
    < PageNumber > 1 </ PageNumber >
    < PageSize > 50 </ PageSize >
    < ScheduledT  asks >
        < ScheduledT  ask >
            < Descriptio  n />
            < LaunchExpi  rationTime > 600 </ LaunchExpi  rationTime >
            < LaunchTime > 2014 - 08 - 18T10 : 52Z </ LaunchTime >
            < Recurrence  EndTime > 2014 - 08 - 20T16 : 55Z </
 Recurrence  EndTime >
            < Recurrence  Type > Daily </ Recurrence  Type >
            < Recurrence  Value > 1 </ Recurrence  Value >
< ScheduledA  ction >
 ari : acs : ess : cn - qingdao : 1344371 : scalingrul  e /
 cCBpdYdQuB  e2cUxOdu6p  iOk
</ ScheduledA  ction >
            < ScheduledT  askId > edRtShc57W  GXdt8TlPbr  jsnV </
 ScheduledT  askId >
            < ScheduledT  askName > edRtShc57W  GXdt8TlPbr  jsnV </
 ScheduledT  askName >
            < TaskEnable  d > true </ TaskEnable  d >
        </ ScheduledT  ask >
    </ ScheduledT  asks >
</ DescribeSc  heduledTas  ksResponse >
```

JSON format:

```
{
  " RequestId ": " 43434132 - 91C4 - 4264 - 8343 - 681130760A  5C ",
" TotalCount ":  1 ,
" PageSize ":  1 ,
  " PageNumber ":  1 ,
  " ScheduledT  asks ": {
    " ScheduledT  ask ": [
      {
        " TaskEnable  d ":  true ,
        " ScheduledT  askId ": " b27CLSc8T4  78c2iqHr6f  qbF ",
        " Descriptio  n ": " ditingshig  echunqingc  hunan ",
        " ScheduledT  askName ": " 9906a33f - 14eb - 42b8 - 8bdb -
 ee8cdf9127  06 ",
        " LaunchExpi  rationTime ":  120 ,
        " Recurrence  Type ": " Daily ",
        " Recurrence  EndTime ": " 2014 - 08 - 13T19 : 19Z ",
        " LaunchTime ": " 2014 - 08 - 12T17 : 55Z ",
        " Recurrence  Value ": " 1 ",
        " ScheduledA  ction ": " ari : acs : ess : cn - qingdao :
 1344371 : scalingrul  e / qGx9feK1gi  admp3XKer9  4cD "
      }
    ]
  }
```

```
}
```

13.2 DeleteScheduledTask

Deletes a specified scaling rule.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value : DeleteScheduledTask.
ScheduledTaskId	String	Yes	ID of the scheduled task.

Response parameters

Public parameters.

Error code

For common errors, see *client errors* or *server errors*.

Error message	Error code	Description	HTTP status codes
The specified scheduled task does not exist in your account.	InvalidScheduledTaskId.NotFound	The specified scheduled task does not exist.	404

Request example

```
http://ess.aliyuncs.com/?Action=DeleteScheduledTask
&ScheduledTaskId=edRtShc57W GXdt8TlPbr jsnV
&<Public Request Parameters>
```

Response example

XML format:

```
<DeleteScheduledTaskResponse>
    <RequestId>7683D637-CF8A-41DC-85A8-E128061E65FC</
RequestId>
```

```
</ DeleteSche  duledTaskR  esponse >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

13.3 CreateScheduledTask

This operation creates a scheduled task according to input parameters.

Description

- You can create up to 20 scheduled tasks.
- When the trigger of a scheduled task fails because a scaling activity in a scaling group is in progress or the scaling group is disabled, the scheduled task is automatically retried within the `LaunchExpi  rationTime` ; otherwise, the scheduled trigger task is abandoned.
- If multiple tasks are scheduled at similar times to execute the rule of the same group, the earliest task triggers the scaling activity first, and other tasks make attempts to execute the rule within their `LaunchExpi  rationTime` because a scaling group executes only one scaling activity at a time. If another scheduled task is still making triggering attempts within its `LaunchExpi  rationTime` after the scaling activity is finished, the scaling rule is executed and the corresponding scaling activity is triggered.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value : CreateSche duledTask.
RegionId	String	Yes	ID of the region in which the scheduled task is located.

Name	Type	Required	Description
ScheduledAction	String	Yes	Operations performed when the scheduled task is triggered. Fill in the unique identifier of the scaling rule.
LaunchTime	String	Yes	Time point at which the scheduled task is triggered. The date format follows the ISO8601 standard and uses UTC time. It is in the format of YYYY-MM-DDThh:mmZ. If RecurrenceType is specified, the time point specified by this attribute is the default time point at which the circle is executed. If RecurrenceType is not specified, the task is executed once on the designated date and time. A time point 90 days after creation or modification cannot be entered.

Name	Type	Required	Description
ScheduledTaskName	String	No	Display name of the scheduled task, which must be 2-40 characters (English or Chinese) long. It must begin with a number, an upper/lower-case letter or a Chinese character and may contain numbers, “_”, “-“ or “.”. The account name is unique in the same region. If this parameter is not specified, the default value ScheduledScalingTaskId is used.
Description	String	No	Description of the scheduled task, which is 2-200 characters (English or Chinese) long.
LaunchExpirationTime	Integer	No	Time period within which the failed scheduled task is retried. The default value is 600s. Value range: [0, 21600]

Name	Type	Required	Description
RecurrenceType	String	No	Type of the scheduled task to be repeated. Optional values: · Daily: Recurrence interval by day for a scheduled task · Weekly: Recurrence interval by week for a scheduled task · Monthly: Recurrence interval by month for a scheduled task · Cron: Execute a scheduled task according to the specified Cron expression. RecurrenceType, RecurrenceValue and RecurrenceEndTime must be specified.

Name	Type	Required	Description
RecurrenceValue	String	No	Value of the scheduled task to be repeated. · Daily: Only one value in the range [1,31] can be filled. · Weekly: Multiple values can be filled. The values of Sunday to Saturday are 0 to 6 in sequence . Multiple values shall be separated by commas (,). · Monthly: In the format of A-B. The value range of A and B is 1 to 31, and the B value must be greater than the A value. · - Cron: An UTC time comprises of minute, hour , day, month and week. And the expression can include wildcard characters including comma (,), question mark (?), hyphen (-), asterisk (*), hash (#), slash (/), L and W. RecurrenceType, RecurrenceValue and RecurrenceEndTime must be specified.

Name	Type	Required	Description
RecurrenceEndTime	String	No	End time of the scheduled task to be repeated. The date format follows the ISO8601 standard and uses UTC time. It is in the format of YYYY-MM-DDThh:mmZ. A time point 90 days after creation or modification cannot be entered. RecurrenceType, RecurrenceValue and RecurrenceEndTime must be specified.
TaskEnabled	Bool	No	Whether to enable the scheduled task. · When the parameter is set to true, the task is enabled. · When the parameter is set to false, the task is disabled. The default value is true.

Response parameters

Name	Type	Description
ScheduledTaskId	String	ID of the scheduled task , which is generated by the system and globally unique.

Request example

```
 http :// ess . aliyuncs . com /? Action = CreateSche  duledTask
& RegionId = cn - qingdao
& LaunchTime = 2014 - 08 - 17T16 : 52Z
& Recurrence  Type = Daily
& Recurrence  Value = 1
& Recurrence  EndTime = 2014 - 08 - 17T16 : 55Z
& ScheduledA  ction = ari : acs : ess : cn - qingdao : 1344371 :
 scalingRul  e / cCBpdYdQuB  e2cUxOdu6p  iOk
&< Public   Request   Parameters >
```

Response example

XML format:

```
< CreateSche  duledTaskR  esponse >
    < ScheduledT  askId > edRtShc57W  GXdt8TlPbr  jsnV </ ScheduledT
 askId >
    < RequestId > 0F02D931 - 2B12 - 44D7 - A0E9 - 39925C13D1  5E </
 RequestId >
</ CreateSche  duledTaskR  esponse >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 ",
    " ScheduledT  askId ": " edRtShc57W  GXdt8TlPbr  jsnV "
}
```

Error code

For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidRegionId. NotFound	The specified region does not exist.	404	The specified region does not exist.

Error code	Error message	HTTP status code	Description
InvalidSch eduledTaskName. Duplicate	The specified value of parameter `ScheduledT askName` is duplicated.	400	The scheduled task name already exists .
QuotaExceeded. ScheduledTask	Scheduled task quota exceeded.	400	The specified ScheduledAction and the specified scheduled task are not in the same region.
ScheduledAction. RegionMismatch	The specified scheduled task and the specified scheduled action are not in the same Region.	400	Your scheduled task quota is exceeded.

13.4 ModifyScheduledTasksk

Modifies the attributes of a scheduled task.

Request parameters

Name	Type	Required	Description
Action	String	Yes	Operation interface name, required parameter. Value : ModifySche duledTask.
ScheduledTaskId	String	Yes	ID of the scheduled task.

Name	Type	Required	Description
ScheduledTaskName	String	No	Display name of the scheduled task, which must be 2-40 characters (English or Chinese) long. It must begin with a number, an upper/lower-case letter or a Chinese character and may contain numbers, “_”, “-“ or “.”. The account name is unique in the same region. If this parameter is not specified, the default value ScheduledScalingTaskId is used.
Description	String	No	Description of the scheduled task, which is 2-200 characters (English or Chinese) long.
ScheduledAction	String	No	Operations performed when the scheduled task is triggered. Fill in the unique identifier of the scaling rule.

Name	Type	Required	Description
LaunchTime	String	No	Time point at which the scheduled task is triggered. The date format follows the ISO8601 standard and uses UTC time. It is in the format of YYYY-MM-DDThh:mmZ. If RecurrenceType is specified, the time point specified by this attribute is the default time point at which the circle is executed. If RecurrenceType is not specified, the task is executed once on the designated date and time. A time point 90 days after creation or modification cannot be entered.

Name	Type	Required	Description
LaunchExpirationTime	Integer	No	Time period within which the failed scheduled task is retried. The default value is 600s. Value range: [0, 21600]

Name	Type	Required	Description
RecurrenceType	String	No	Type of the scheduled task to be repeated. Optional values: · Daily: Recurrence interval by day for a scheduled task. · Weekly: Recurrence interval by week for a scheduled task. · Monthly: Recurrence interval by month for a scheduled task. · Cron: Execute a scheduled task according to the specified Cron expression. After modification, RecurrenceType, RecurrenceValue and RecurrenceEndTime must be set simultaneously.

Name	Type	Required	Description
RecurrenceValue	String	No	Value of the scheduled task to be repeated. · Daily: Only one value in the range [1,31] can be filled. · Weekly: Multiple values can be filled. The values of Sunday to Saturday are 0 to 6 in sequence . Multiple values shall be separated by commas (,). · Monthly: In the format of A-B. The value range of A and B is 1 to 31, and the B value must be greater than the A value. · Cron: An expression comprises of minute, hour , day, month and week. And the expression supports comma (,), hyphen (-), asterisk (*) and slash (/). After modification , RecurrenceType , RecurrenceValue and Recurrence EndTime must be set simultaneously.

Name	Type	Required	Description
RecurrenceEndTime	String	No	End time of the scheduled task to be repeated. The date format follows the ISO8601 standard and uses UTC time. It is in the format of YYYY-MM-DDThh:mmZ. A time point 90 days after creation or modification cannot be entered. After modification, RecurrenceType, RecurrenceValue and RecurrenceEndTime must be set simultaneously.
TaskEnabled	Bool	No	Whether to enable the scheduled task. · When the parameter is set to true, the task is enabled. · When the parameter is set to false, the task is disabled. The default value is true.

Response parameters

Public parameters.

Request example

```
 http :// ess . aliyuncs . com /? Action = ModifySche  duledTask
& ScheduledT  askId = edRtShc57W  GXdt8TlPbr  jsnV
& LaunchTime = 2014 - 08 - 18T10 : 52Z
& Recurrence  EndTime = 2014 - 08 - 20T16 : 55Z
&< Public   Request   Parameters >
```

Response example

XML format:

```
< ModifySche  duledTaskR  esponse >
    < RequestId > F9372E8D - C163 - 471F - BEB4 - 3A02B3CE17  6E </
 RequestId >
</ ModifySche  duledTaskR  esponse >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error code

For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidSch eduledTaskId. NotFound	The specified scheduled task does not exist.	404	The specified scheduled task does not exist.
InvalidSch eduledTaskName. Duplicate	The specified value of parameter `ScheduledT askName` is duplicated.	400	The specified value of parameter <parameter name> is duplicated.
ScheduledAction. RegionMismatch	The specified scheduled task and the specified scheduled action are not in the same Region.	400	The specified scheduled task and the specified scheduled action are not in the same Region.

14 Lifecycle Hook

14.1 CreateLifecycleHook

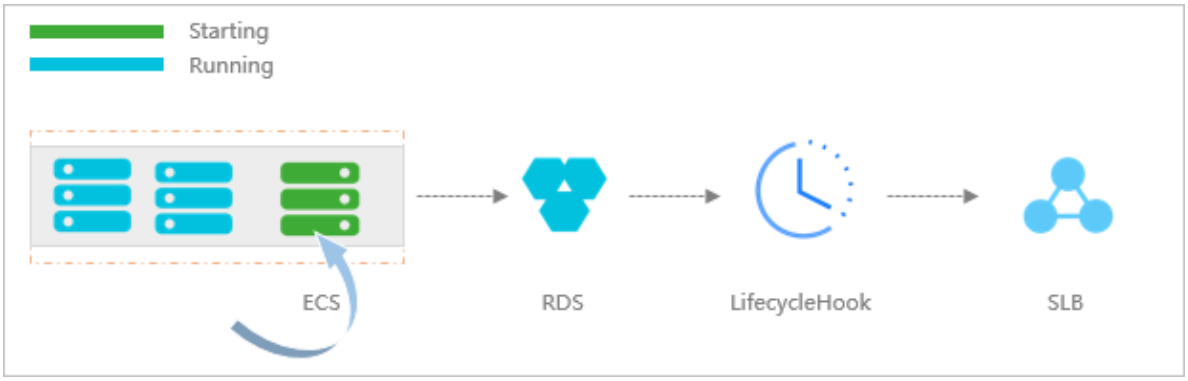
Create lifecycle hooks for a scaling group (`CreateLife cycleHook`).

Description

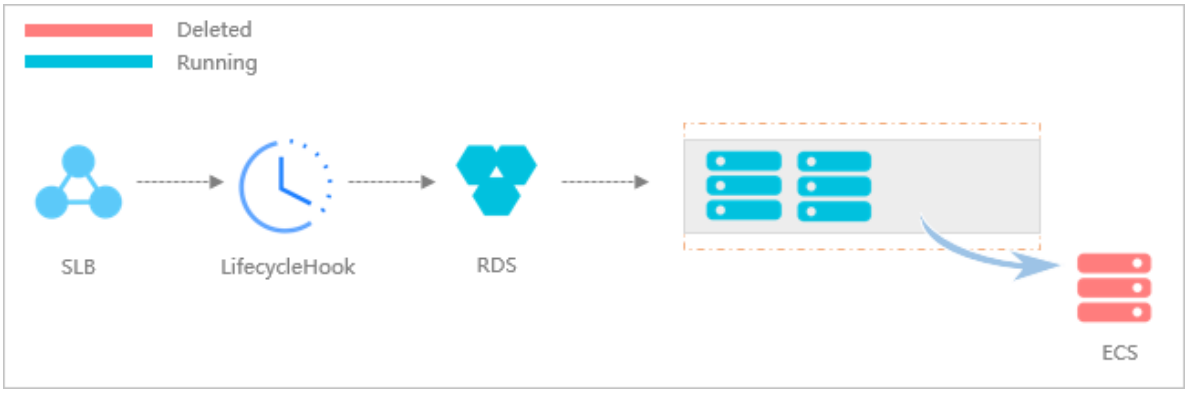
Each scaling group can have up to six lifecycle hooks. Lifecycle hooks enable scaling groups to remain in a wait state for a finite period of time during a scale-in or scale-out event. You can adjust this time by configuring the `HeartbeatT imeout` parameter.

A default period of time is specified before a scaling group scales in or out the number of ECS instances. While an ECS instance is in wait state, you can initialize the configuration of the ECS instance or retrieve data of the ECS instance. During a scale-out process, the ECS instance enters a wait state after IP addresses are added to the whitelist of a specified Relational Database Service (RDS) instance. When the wait state ends, IP addresses will be added to a specified Server Load Balancer (SLB) instance. During a scale-in process, the ECS instance enters a wait state after IP addresses are removed from the whitelist of an SLB instance. When the wait state ends, IP addresses will be removed from the whitelist of an RDS instance. The procedure is as follows.

- A scale-out event:



- A scale-in event:



We recommend that you use an Alibaba Cloud Message Service (MNS) queue or topic to receive notifications about when ECS instances are launched or terminated.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: CreateLifecycleHook
ScalingGroupId	String	Yes	The ID of the scaling group
LifecycleHookName	String	No	The name of the lifecycle hook. Each name must be unique within a scaling group. The name must be 2 to 40 characters in length and can contain letters, numbers, Chinese characters, and special characters including underscores (_), hyphens (-) and periods (.). Default value: Lifecycle Hook ID

Name	Type	Required	Description
LifecycleTransition	String	Yes	The scaling activities to which lifecycle hooks apply Valid values: · SCALE_OUT: scale-out event · SCALE_IN: scale-in event
HeartbeatTimeout	Integer	No	The time, in seconds, that can elapse before the lifecycle hook times out. If the lifecycle hook times out, the scaling group performs the default action (DefaultResult). The range is from 30 to 21,600 seconds. The default value is 600 seconds. You can prevent the lifecycle hook from timing out by calling the *RecordLifecycleActionHeartbeat* API. You can also terminate the lifecycle action by calling the *CompleteLifecycleAction* API.

Name	Type	Required	Description
DefaultResult	String	No	The action that the scaling group takes when the lifecycle hook times out. Value range: · CONTINUE: Continues the scaling activity when the current lifecycle action ends. · ABANDON: Releases the created ECS instances if the scaling activity type is scale-out. Removes the ECS instances if scaling activity type is scale-in. Default value: CONTINUE If the scaling group has multiple lifecycle hooks and one of them is terminated by the `DefaultResult = ABANDON` parameter during a scale-in event (`SCALE_IN`), the remaining lifecycle hooks under the same scaling group will also be terminated. Otherwise, the action following

Name	Type	Required	Description
NotificationArn	String	No	The Alibaba Cloud Resource Name (ARN) of the notification target that Auto Scaling will use to notify you when an instance is in the transition state for the lifecycle hook. This target can be either an MNS queue or an MNS topic. The format of the parameter value is acs:ess:{region}:{account-id}:{resource-relative-id}. · `region`: the region to which the scaling group locates · `account-id`: Alibaba Cloud ID For example: · MNS queue: acs:ess:{region}:{account-id}:queue/{queuename} · MNS topic: acs:ess:{region}:{account-id}:topic/{topicname}

Name	Type	Required	Description
Notificati onMetadata	String	No	The fixed string that you want to include when Auto Scaling sends a message about the wait state of the scaling activity to the notification target. The length of the parameter can be up to 128 characters. Auto Scaling will send the specified `Notificati onMetadata` parameter along with the notification message so that you can easily categorize your notifications The `Notificati onMetadata` parameter will only take effect after you specify the `Notificati onArn` parameter.

Response parameters

Name	Type	Required
RequestId	String	The request ID
LifecycleHookId	String	The lifecycle hook ID

Request example

```
 http :// ess . aliyuncs . com /? Action = CreateLife  cycleHook
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& LifecycleH  ookName = TESTSCALE_  OUT
& LifecycleT  ransition = SCALE_OUT
```

```
& Notificati  onArn = acs : ess : cn - hangzhou : 1111111111 : queue /
 queue1
& Notificati  onMetadata = Test
&< Public   request   parameter >
```

Response example

XML format:

```
< CreateLife  cycleHookR  esponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
    < LifecycleH  ookId > ash - xxxxxxxxxx  xxxxxxx </ LifecycleH
 ookId >
</ CreateLife  cycleHookR  esponse >
```

JSON format:

```
{
    " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 ",
    " lifecycleH  ookId ": " ash - xxxxxxxxxx  xxxxxxx "
}
```

Error codes

Error codes that are specific to this interface are as follows. For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidParamter	The specified value of parameter is not valid.	400	The specified value of the parameter is invalid.
InvalidNotificationArn	The specified parameter `notificati onArn` is invalid.	400	The specified value of the parameter `Notificati onArn` is invalid.
UnsupportedNotificationType.CurrentRegion	The `notificati onType` is not supported in the special region which scalingGroup belongs to.	400	The type of notification that is supported depends on the region to which the scaling group belongs.
QueueNotExist	The specified queue does not exist.	400	The specified MNS queue does not exist.
TopicNotExist	The specified topic does not exist.	400	The specified MNS topic does not exist.

Error code	Error message	HTTP status code	Description
InvalidLifecycleHookName.Duplicate	The specified value of parameter lifecycleHookName is duplicated.	400	The specified value of the parameter LifecycleHookName already exists.
QuotaExceeded.LifecycleHook	Lifecycle hook quota exceeded in the specified scaling group.	400	Each scaling group can have up to six lifecycle hooks.

14.2 ModifyLifecycleHook

Modify the property of a lifecycle hook (ModifyLifecycleHook).

You can modify the property of a lifecycle hook using either of the following methods:

- Specify the parameter LifecycleHookId . ScalingGroupId and LifecycleHookName are ignored.
- Specify the parameters ScalingGroupId and LifecycleHookName .

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: ModifyLifecycleHook
LifecycleHookId	String	No	The ID of the lifecycle hook
ScalingGroupId	String	No	The ID of the scaling group

Name	Type	Required	Description
LifecycleHookName	String	No	The name of the lifecycle hook. Each name must be unique within a scaling group. The name must 2 to 40 characters in length and can contain letters, numbers, Chinese characters, and special characters including underscores (_), hyphens (-), and periods (.).
LifecycleTransition	String	No	The scaling activities to which the lifecycle hooks apply Value range: · SCALE_OUT: scale-out event · SCALE_IN: scale-in event

Name	Type	Required	Description
HeartbeatTimeout	Integer	No	The time, in seconds, that can elapse before the lifecycle hook times out. If the lifecycle hook times out, the scaling group performs the default action (DefaultResult). The range is from 30 to 21,600 seconds. The default is 600 seconds. You can prevent the lifecycle hook from timing out by calling the *RecordLifecycleActionHeartbeat* API. You can also terminate the lifecycle action by calling the *CompleteLifecycleAction* API.

Name	Type	Required	Description
DefaultResult	String	No	The action that the scaling group takes when the lifecycle hook times out Value range: · CONTINUE: Continues the scaling activity when the current lifecycle action ends. · ABANDON: Releases the created ECS instances if the scaling activity type is scale-out. Removes the ECS instances if scaling activity type is scale-in. Default value: CONTINUE If the scaling group has multiple lifecycle hooks and one of them is terminated by the `DefaultResult = ABANDON` parameter during a scale-in event (`SCALE_IN`), the remaining lifecycle hooks under the same scaling group will also stop working. Otherwise, the action following

Name	Type	Required	Description
NotificationArn	String	No	The Alibaba Cloud Resource Name (ARN) of the notification target that Auto Scaling will use to notify you when an instance is in the transition state for the lifecycle hook. This target can be either an MNS queue or an MNS topic. The format of the parameter value is acs:ess:{region}:{account-id}:{resource-relative-id}. · `region`: the region to which the scaling group belongs · `account-id`: Alibaba Cloud ID For example: · MNS queue: acs:ess:{region}:{account-id}:queue/{queuename} · MNS topic: acs:ess:{region}:{account-id}:topic/{topicname}

Name	Type	Required	Description
Notificati onMetadata	String	No	The fixed string that you want to include when Auto Scaling sends a message about the wait state of the scaling activity to the notification target. The length of the parameter can be up to 128 characters. Auto Scaling will send the specified `Notificati onMetadata` parameter along with the notification message so that you can easily categorize your notifications. The `Notificati onMetadata` parameter only takes effect after you specify the `Notificati onArn` parameter.

Response parameters

Name	Type	Description
RequestId	String	The Request ID

Request example

```
 http :// ess . aliyuncs . com /? Action = ModifyLife  cycleHook
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& LifecycleH  ookName = testSCALE_  IN
& LifecycleT  ransition = SCALE_IN
& DefaultRes  ult = ABANDON
```

```
& Notificati  onArn = acs : ess : cn - hangzhou : 1111111111 : queue /
 queue2
& Notificati  onMetadata = Test
&< Public   request   parameter >
```

Response example

XML format:

```
< ModifyLife  cycleHookR  esponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ ModifyLife  cycleHookR  esponse >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error codes that are specific to this interface are as follows. For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidParamter	The specified value of parameter is not valid.	400	The specified value of the parameter is invalid.
InvalidLifecycleHookId.NotExist	The specified lifecycleHookId not exist.	400	The specified `LifecycleHookId` does not exist.
InvalidLifecycleHookName.NotExist	The specified lifecycleHookName you provided not exist.	400	The specified `LifecycleHookName` does not exist.
InvalidNotificationArn	The specified parameter `notificationArn` is invalid.	400	The specified `NotificationArn` does not exist.

Error code	Error message	HTTP status code	Description
UnsupportedNotificationType.CurrentRegion	The notificationType is not supported in the special region which scalingGroup belongs to.	400	The type of notification that is supported depends on the region to which the scaling group belongs.
LifecycleHook	The specified queue does not exist.	400	The specified MNS queue does not exist.
TopicNotExist	The specified topic does not exist.	400	The specified MNS topic does not exist.

14.3 DescribeLifecycleHooks

This interface queries one or multiple lifecycle hooks that meet the specified conditions (`DescribeLifecycleHooks`).

You can query lifecycle hooks using the following methods:

- Specify the parameter `LifecycleHookId.N` . `ScalingGroupId` and `LifecycleHookName` are ignored.
- Specify the parameter `ScalingGroupId` .
- Specify the parameters `ScalingGroupId` and `LifecycleHookName` .

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: DescribeLifecycleHooks.
LifecycleHookId.N	String	No	The list of lifecycle hook IDs. Replace `N` with the actual number to indicate a specific item lifecycle hook. The value range of N is [1, 50].

Name	Type	Required	Description
ScalingGroupId	String	No	The ID of the scaling group
LifecycleHookName	String	No	The name of the lifecycle hook
PageNumber	Integer	No	Page numbers of the instance status list. Initial value: 1; Default value: 1.
PageSize	Integer	No	Number of rows per page when performing a query. Maximum: 50. Default: 50.

Response parameters

Name	Type	Description
RequestId	String	The request ID
PageNumber	Integer	The initial page number for query
PageSize	String	The number of items returned per page when performing a query
TotalCount	String	Total number of lifecycle hooks
LifecycleHooks	*LifecycleHookModelSet*	The list of lifecycle hook data

LifecycleHookModelSet

Name	Type	Description
ScalingGroupId	String	The ID of the scaling group
LifecycleHookId	String	The ID of the lifecycle hook
LifecycleHookName	String	The name of the lifecycle hook

Name	Type	Description
DefaultResult	String	The action that the scaling group takes when the lifecycle hook times out
HeartbeatTimeout	Integer	The time, in seconds, that can elapse before the lifecycle hook times out. When the lifecycle hook times out, the scaling group performs the default action.
LifecycleTransition	String	The scaling activities to which lifecycle hooks apply
NotificationMetadata	String	The fixed string that you want to include when Auto Scaling sends a message about the wait state of the scaling activity to the notification target.
NotificationArn	String	The Alibaba Cloud Resource Name (ARN) of the notification target that Auto Scaling will use to notify you when an instance is in the transition state for the lifecycle hook .

Request example

```
 http :// ess . aliyuncs . com /? Action = DescribeLi  fecycleHoo  ks
& ScalingGro  upId = asg - xxxxx
&< Public   request   parameter >
```

Response example

XML format:

```
< DescribeLi  fecycleHoo  ksResponse >
  < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
  < PageNumber > 1 </ PageNumber >
  < PageSize > 50 </ PageSize >
  < TotalCount > 1 </ TotalCount >
  < LifecycleH  ooks >
```

```
    < LifecycleH  ook >
      < ScalingGro  upId > asg - xxx </ ScalingGro  upId >
      < LifecycleH  ookId > ash - xxx </ LifecycleH  ookId >
      < LifecycleH  ookName > Test </ LifecycleH  ookName >
      < DefaultRes  ult > CONTINUE </ DefaultRes  ult >
      < HeartbeatT  imeout > 60 </ HeartbeatT  imeout >
      < LifecycleT  ransition > SCALE_OUT </ LifecycleT  ransition >
      < Notificati  onMetadata > Test </ Notificati  onMetadata >
      < Notificati  onArn > acs : ess : cn - hangzhou : 1111111111 :
 queue / queue1 </ Notificati  onArn >
    </ LifecycleH  ook >
  </ LifecycleH  ooks >
</ DescribeLi  fecycleHoo  ksResponse >
```

JSON format:

```
{
  " lifecycleH  ooks ": [
  {
    " defaultRes  ult ": " CONTINUE ",
    " heartbeatT  imeout ":  600 ,
    " lifecycleH  ookId ": " ash - xxx ",
    " lifecycleH  ookName ": " Test ",
    " lifecycleT  ransition ": " SCALE_OUT ",
    " notificati  onArn ": " acs : ess : cn - hangzhou : 1111111111 :
 queue / queue1 ",
    " notificati  onMetadata ": " Test ",
    " scalingGro  upId ": " asg - xxx "
  }
  ],
  " pageNumber ":  1 ,
  " pageSize ":  50 ,
  " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 ",
  " totalCount ":  1
}
```

Error code

Error code specific to this interface is as follows. For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidParamter	The specified value of parameter is not valid.	400	The specified value of the parameter is invalid.

14.4 RecordLifecycleActionHeartbeat

You can use the `RecordLife cycleActio nHeartbeat` operation to extend the timeout by the length of time required to keep the instance in a wait state.

An ECS instance can be kept in the wait state for no longer than six hours. The length of time for each wait state can be extended up to 20 times.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: RecordLifecycleActionHeartbeat.
LifecycleHookId	String	Yes	The ID of the lifecycle hook
LifecycleActionToken	String	Yes	A token that indicates a specific lifecycle action associated with an instance. You can obtain this token using an MNS queue or MNS topic specified for the lifecycle hook.

Name	Type	Required	Description
HeartbeatTimeout	Integer	No	The time, in seconds, that can elapse before the lifecycle hook times out. If the lifecycle hook times out, the scaling group performs the default action (DefaultResult). The range is [30, 21600]. The default is 600 seconds. You can prevent the lifecycle hook from timing out by calling the RecordLifecycleActionHeartbea API. You can also terminate the lifecycle action by calling the *CompleteLifecycleAction* API.

Response parameters

Name	Type	Description
RequestId	String	The Request ID

Request example

```
 http :// ess . aliyuncs . com /? Action = RecordLife  cycleActio
 nHeartbeat
& LifecycleH  ookId = ash - xxxxxxxxxx  x
& LifecycleA  ctionToken = aaaa - bbbbb - cccc - ddddd
& LifecycleA  ctionResul  t = CONTINUE
```

```
&< Public   request   parameter >
```

Response example

XML format:

```
< RecordLife  cycleActio  nHeartbeat  Response >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ RecordLife  cycleActio  nHeartbeat  Response >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error codes that are specific to this interface are as follows. Error codes that are specific to this interface are as follows. For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidParamter	The specified value of parameter is not valid.	400	The specified value of the parameter is invalid.
LifecycleHookIdAndLifecycleActionToken.Invalid	The specified lifecycleActionToken and lifecycleHookId you provided does not match any in process lifecycle action.	400	The specified `LifecycleActionToken` does not match any `LifecycleHookId`.
LifecycleAction.TimeExceeded	The specified parameter heartbeatTime exceed lifecycleAction max suspend time.	400	An ECS instance can be kept in the wait state for no longer than six hours.
LifecycleAction.RecordTimesExceeded	The specified lifecycleAction exceed lifecycleAction max record times.	400	The length of time for each wait state can be extended up to 20 times.

14.5 CompleteLifecycleAction

If you finish before the timeout period ends, complete the lifecycle action (`CompleteLi fecycleAct ion`).

You can set the action following the wait state to Continue to complete this scaling event (`CONTINUE`) or to Abandon to terminate this scaling event (`ABANDON`).

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: CompleteLifecycleAction
LifecycleHookId	String	Yes	The ID of the lifecycle hook
LifecycleActionToken	String	Yes	A token that indicates a specific lifecycle action associated with an instance. You can obtain this token using an MNS queue or MNS topic specified for the lifecycle hook.

Name	Type	Required	Description
LifecycleActionResult	String	No	The action that the scaling group takes when the lifecycle hook times out Value range: · CONTINUE: Continues the scaling activity when the current lifecycle action ends. · ABANDON: Releases the created ECS instances if the scaling activity type is scale-out. Removes the ECS instances if scaling activity type is scale-in. Default value: CONTINUE If the scaling group has multiple lifecycle hooks and one of them is terminated by the `DefaultResult = ABANDON` parameter during a scale-in event (`SCALE_IN`), the remaining lifecycle hooks under the same scaling group will also be terminated. Otherwise, the action following

Response parameters

Name	Type	Description
RequestId	String	The Request ID

Request example

```
 http :// ess . aliyuncs . com /? Action = CompleteLi  fecycleAct
 ion
& LifecycleH  ookId = ash - xxxxxxxxxx  x
& LifecycleA  ctionToken = aaaa - bbbbb - cccc - ddddd
& LifecycleA  ctionResul  t = CONTINUE
&< Public   request   parameter >
```

Response example

XML format:

```
< CompleteLi  fecycleAct  ionRespons  e >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ CompleteLi  fecycleAct  ionRespons  e >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error codes that are specific to this interface are as follows. Error codes that are specific to this interface are as follows. For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidParamter	The specified value of parameter is invalid.	400	The specified value of the parameter is invalid.
LifecycleH ookIdAndLi fecycleActionToken .Invalid	The specified lifecycleA ctionToken and lifecycleHookId you provided does not match any in process lifecycle action.	400	The specified `LifecycleA ctionToken` does not match any `LifecycleH ookId`.

14.6 DeleteLifecycleHook

Delete a lifecycle hook (`DeleteLife cycleHook`).

The wait state will be terminated once the corresponding lifecycle hook is deleted.

You can delete a lifecycle hook using either of the following methods:

- Specify the parameter `LifecycleH ookId` . `ScalingGro upId` and `LifecycleH ookName` are ignored.
- Specify the parameters `ScalingGro upId` and `LifecycleH ookName` .

Request parameters

Name	Type	Required	Description
Action	String	Yes	The name of this interface. Value: DeleteLife cycleHook.
LifecycleHookId	String	No	The ID of the lifecycle hook
ScalingGroupId	String	No	The ID of the scaling group
LifecycleH ookName	String	No	The name of the lifecycle hook

Response parameters

Name	Type	Description
RequestId	String	The request ID

Request example

```
 Http :// ess . aliyuncs . com  /?  Action  =  fig
& LifecycleH  ookId = ash - xxxxxxxxxx  x
&< Public   request   parameter >
```

Response example

XML format:

```
< DeleteLife  cycleHookR  esponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
```

```
</ DeleteLife  cycleHookR  esponse >
```

JSON format:

```
{
    " RequestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

Error codes that are specific to this interface are as follows. Error codes that are specific to this interface are as follows. For common errors, see *client errors* or *server errors*.

Error code	Error message	HTTP status code	Description
InvalidParamter	The specified value of parameter is invalid.	400	The specified value of the parameter is invalid.
InvalidLifecycleHookId.NotExist	The specified lifecycleHookId does not exist.	400	The specified lifecycle hook ID does not exist.
InvalidLifecycleHookName.NotExist	The specified lifecycleHookName you provided does not exist.	400	The specified lifecycle hook name does not exist.

15 Event notification

15.1 CreateNotificationConfiguration

You can call this operation to create scaling events and resource change notifications (`CreateNotificationConfiguration`).

Description

You can receive notifications by using CloudMonitor, MNS queues, or MNS topics. When a scaling event or resource change with the specified type occurs in a scaling group, Auto Scaling will send notifications to CloudMonitor or MNS.

Currently, you cannot use MNS queues and MNS topics in several Alibaba Cloud regions. For more information, visit the *MNS console*.

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to CreateNotificationConfiguration.
ScalingGroupId	String	Yes	The ID of the scaling group.

Name	Type	Required	Description
NotificationArn	String	Yes	The identifier for the notification object. The format is: `acs : ess :{ region }:{ account - id }:{ resource - relative - id }`. · The `region` is the ID of the region where the scaling group is located. For more information, see *Regions and zones*. · The `account - id` is the account ID. · The `resource - relative - id` is the notification method. Valid values: `cloudmonitor` (system events of CloudMonitor), `queue /{ queuename }` (MNS queues), and `topic /{ topicname }` (MNS topics).

Name	Type	Required	Description
NotificationType.N	String	Yes	One or more types of scaling events and resource change notifications. Valid values of `N`: 1 to 8. Multiple values are organized in an ordered list. You can call the *DescribeNotificationTypes* operation to query parameter values.

Response parameters

Name	Type	Description
RequestId	String	The request ID.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = CreateNoti  ficationCo
 nfiguratio  n
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& Notificati  onArn = acs : ess : cn - hangzhou : 123456 : cloudmonit
 or
& Notificati  onType .  1 = AUTOSCALIN  G : SCALE_OUT_  SUCCESS
& Notificati  onType .  2 = AUTOSCALIN  G : SCALE_IN_S  UCCESS
& Notificati  onType .  3 = AUTOSCALIN  G : SCALE_OUT_  ERROR
& Notificati  onType .  4 = AUTOSCALIN  G : SCALE_IN_E  RROR
& Notificati  onType .  5 = AUTOSCALIN  G : SCALE_REJE  CT
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< CreateNoti  ficationCo  nfiguratio  nResponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
```

```
</ CreateNoti  ficationCo  nfiguratio  nResponse >
```

`JSON` format

```
{
    " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

HttpCode	Error code	Error message	Description
400	InvalidNotificationArn	The specified parameter `notificationArn` is invalid.	The error message returned when the specified `NotificationArn` is invalid.
400	InvalidNotificationTypes	The specified notificationType is invalid.	The error message returned when the specified `NotificationType . N` is invalid.
400	NotificationConfigurationExist	The specified notificationConfiguration already exist for the scalingGroup.	The error message returned when the specified event notification exists in the scaling group .
400	NotificationConfigurationQuotaExceed.ForScalingGroup	NotificationConfiguration num exceed for the specified scalingGroup.	The error message returned when the number of event notifications you specify for a scaling group exceeds the maximum number.
400	QueueNotExist	The specified queue `queuename` does not exist.	The error message returned when the specified MNS queue does not exist.

HttpCode	Error code	Error message	Description
400	TopicNotExist	The specified topic `topicname` does not exist.	The error message returned when the specified MNS topic does not exist.
400	UnsupportedNotificationType.CurrentRegion	The `notificationType` is not supported in the special region which scalingGroup belongs to.	The error message returned when the specified notification type is not supported in the region.
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist.

15.2 DescribeNotificationConfigurations

You can call this operation to query newly created scaling events and resource change notifications (`DescribeNotificationConfigurations`).

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to DescribeNotificationConfigurations.
ScalingGroupId	String	Yes	The ID of the scaling group.

Response parameters

Name	Type	Description
RequestId	String	The request ID
NotificationConfigurationModels	Array of *NotificationConfigurationModelSet*	The list of scaling events and resource change notifications.

NotificationConfigurationModelSet

Name	Type	Description
ScalingGroupId	String	The request ID
NotificationArn	String	The list of scaling events and resource change notifications.
NotificationTypes	Array of *NotificationTypeSet*	The list of scaling events and notification types of resource changes.

NotificationTypeSet

Name	Type	Description
NotificationType	String	Scaling events and notification types of resource changes. You can call the *DescribeNotificationTypes* operation to query parameter values.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = DescribeNo  tification
 Configurat  ions
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
&< Common   request   parameters >
```

Successful response examples

`XML` format

```
< DescribeNo  tification  Configurat  ionsRespon  se >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
    < Notificati  onConfigur  ationModel  s >
        < Notificati  onConfigur  ationModel >
            < Notificati  onArn > xxxxxxxxx </ Notificati  onArn >
            < ScalingGro  upId > xxxxxxxxxx  x </ ScalingGro  upId >
            < Notificati  onTypes >
                < Notificati  onType > AUTOSCALIN  G : SCALE_IN_S
 UCCESS </ Notificati  onType >
                < Notificati  onType > AUTOSCALIN  G : SCALE_IN_E
 RROR </ Notificati  onType >
            </ Notificati  onTypes >
       </ Notificati  onConfigur  ationModel >
```

```
        < Notificati  onConfigur  ationModel >
            < Notificati  onArn > xxxxxxxxx </ Notificati  onArn >
            < ScalingGro  upId > xxxxxxxxxx  x </ ScalingGro  upId >
            < Notificati  onTypes >
                < Notificati  onType > AUTOSCALIN  G : SCALE_IN_S
 UCCESS </ Notificati  onType >
                < Notificati  onType > AUTOSCALIN  G : SCALE_IN_E
 RROR </ Notificati  onType >
            </ Notificati  onTypes >
        </ Notificati  onConfigur  ationModel >
    </ Notificati  onConfigur  ationModel  s >
</ DescribeNo  tification  Configurat  ionsRespon  se >
```

`JSON` format

```
{
    " notificati  onConfigur  ationModel  s ": [
    {
      " notificati  onArn ": " xxxxxxxxxx ",
      " notificati  onTypes ": [
        " AUTOSCALIN  G : SCALE_OUT_  SUCCESS ",
        " AUTOSCALIN  G : SCALE_OUT_  ERROR ",
        " AUTOSCALIN  G : SCALE_IN_S  UCCESS ",
        " AUTOSCALIN  G : SCALE_IN_E  RROR ",
        " AUTOSCALIN  G : SCALE_REJE  CT "
      ],
      " scalingGro  upId ": " xxxxxxxxxx "
    },
    {
      " notificati  onArn ": " xxxxxxxxxx ",
      " notificati  onTypes ": [
        " AUTOSCALIN  G : SCALE_OUT_  SUCCESS ",
        " AUTOSCALIN  G : SCALE_OUT_  ERROR ",
        " AUTOSCALIN  G : SCALE_IN_S  UCCESS ",
        " AUTOSCALIN  G : SCALE_IN_E  RROR ",
        " AUTOSCALIN  G : SCALE_REJE  CT "
      ],
      " scalingGro  upId ": " xxxxxxxxxx "
    }
  ],
  " requestId ": " EEAB83FF - 318A - 41CA - 8EB6 - BB9614256B  A9 "
}
```

Error codes

HttpCode	Error code	Error message	Description
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist.

15.3 ModifyNotificationConfiguration

You can call this operation to modify scaling events and resource change notifications (`ModifyNoti ficationCo nfiguratio n`).

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to ModifyNoti ficationCo nfiguration.
ScalingGroupId	String	Yes	The ID of the scaling group.

Name	Type	Required	Description
NotificationArn	String	Yes	The identifier for the notification object. The format is: `acs : ess :{ region }:{ account - id }:{ resource - relative - id }`. · The `region` is the ID of the region where the scaling group is located. For more information, see *Regions and zones*. · The `account - id` is the account ID. · The `resource - relative - id` is the notification method. Valid values: `cloudmonitor` (system events of CloudMonitor), `queue /{ queuename }` (MNS queues), and `topic /{ topicname }` (MNS topics).

Name	Type	Required	Description
NotificationType.N	String	Yes	Includes one or more types of scaling event notifications. Valid values of `N`: 1 to 5. Multiple values are listed in ascending order. You can call the *DescribeNotificationTypes* operation to query parameter values.

Response parameters

Name	Type	Description
RequestId	String	The request ID.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = ModifyNoti  ficationCo
 nfiguratio  n
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& Notificati  onArn = acs : ess : cn - hangzhou : 123456 : cloudmonit
 or
& Notificati  onType .  1 = AUTOSCALIN  G : SCALE_OUT_  SUCCESS
& Notificati  onType .  2 = AUTOSCALIN  G : SCALE_IN_S  UCCESS
& Notificati  onType .  3 = AUTOSCALIN  G : SCALE_OUT_  ERROR
& Notificati  onType .  4 = AUTOSCALIN  G : SCALE_IN_E  RROR
& Notificati  onType .  5 = AUTOSCALIN  G : SCALE_REJE  CT
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< ModifyNoti  ficationCo  nfiguratio  nResponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ ModifyNoti  ficationCo  nfiguratio  nResponse >
```

`JSON` format

```
{
    " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
```

```
}
```

Error codes

HttpCode	Error code	Error message	Description
400	InvalidNotificationTypes	The specified notificationType is invalid.	The error message returned when the specified `NotificationType.N` is invalid.
400	NotificationConfigurationNotExist	The specified notification configuration not exist for the scalingGroup.	The error message returned when the specified scaling event and the notification type of the resource change do not exist.
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist.

15.4 DeleteNotificationConfiguration

You can call this operation to delete scaling events and resource change notifications (`DeleteNotificationConfiguration`).

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to DeleteNotificationConfiguration.
ScalingGroupId	String	Yes	The ID of the scaling group.

Name	Type	Required	Description
NotificationArn	String	Yes	The identifier for the notification object. The format is: `acs : ess :{ region }:{ account - id }:{ resource - relative - id }`. · The `region` is the ID of the region where the scaling group is located. For more information, see *Regions and zones*. · The `account - id` is the account ID. · The `resource - relative - id` is the notification method. Valid values: `cloudmonitor` (system events of CloudMonitor), `queue /{ queuename }` (MNS queues), and `topic /{ topicname }` (MNS topics).

Response parameters

Name	Type	Description
RequestId	String	The request ID

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = DeleteNoti  ficationCo
 nfiguratio  n
& ScalingGro  upId = AG6CQdPU8O  KdwLjgZcJ2  eaQ
& Notificati  onArn = acs : ess : cn - hangzhou : 123456 : cloudmonit
 or
&< Common   request   parameters >
```

Successful response examples

`XML` format

```
< DeleteNoti  ficationCo  nfiguratio  nResponse >
    < RequestId > 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 </
 RequestId >
</ DeleteNoti  ficationCo  nfiguratio  nResponse >
```

`JSON` format

```
{
    " requestId ": " 04F0F334 - 1335 - 436C - A1D7 - 6C044FE733  68 "
}
```

Error codes

HTTP status code	Error code	Error message	Description
400	NotificationConfigurationNotExist	The specified notificationConfiguration not exist for the scalingGroup.	The error message returned when the specified scaling event and resource change notification do not exist in the scaling group.
404	InvalidScalingGroupId.NotFound	The specified scaling group does not exist.	The error message returned when the specified scaling group does not exist.

15.5 DescribeNotificationTypes

You can call this operation to query scaling events and notification types of resource changes (`DescribeNo tification Types`).

Request parameters

Name	Type	Required	Description
Action	String	Yes	The operation that you want to perform. Set the value to DescribeNotificationTypes.

Response parameters

Name	Type	Description
RequestId	String	The request ID
NotificationTypes	Array of *NotificationTypeSet*	The list of scaling events and notification types of resource changes.

NotificationTypeSet

Name	Type	Description
NotificationType	String	Scaling events and notification types of resource changes. Valid values: · AUTOSCALING: SCALE_OUT_SUCCESS : indicates that a scale-out activity is successful . · AUTOSCALING: SCALE_IN_SUCCESS: indicates that a scale-in activity is successful. · AUTOSCALING: SCALE_OUT_ERROR: indicates that a scale-out activity has failed. · AUTOSCALING: SCALE_IN_ERROR: indicates that a scale-in activity has failed. · AUTOSCALING: SCALE_REJECT: indicates that a scaling activity is rejected. · AUTOSCALING: SCALE_OUT_START: indicates that a scale-out activity starts. · AUTOSCALING: SCALE_IN_START: indicates that a scale-in activity starts. · AUTOSCALING :SCHEDULE_T ASK_EXPIRING: indicates that a scheduled task expires.

Examples

Sample requests

```
 http :// ess . aliyuncs . com /? Action = DescribeNo  tification
 Types
& < Common   request   parameters >
```

Successful response examples

`XML` format

```
< DescribeNo  tification  TypesRespo  nse >
    < RequestId > 21F638C3 - 7054 - 4C1E - A143 - A74C20F507  A3 </
 RequestId >
    < Notificati  onTypes >
        < Notificati  onType > AUTOSCALIN  G : SCALE_OUT_  SUCCESS </
 Notificati  onType >
        < Notificati  onType > AUTOSCALIN  G : SCALE_IN_S  UCCESS </
 Notificati  onType >
        < Notificati  onType > AUTOSCALIN  G : SCALE_OUT_  ERROR </
 Notificati  onType >
        < Notificati  onType > AUTOSCALIN  G : SCALE_IN_E  RROR </
 Notificati  onType >
        < Notificati  onType > AUTOSCALIN  G : SCALE_REJE  CT </
 Notificati  onType >
    </ Notificati  onTypes >
</ DescribeNo  tification  TypesRespo  nse >
```

`JSON` format

```
{
    " notificati  onTypes ": [
        " AUTOSCALIN  G : SCALE_OUT_  SUCCESS ",
        " AUTOSCALIN  G : SCALE_IN_S  UCCESS ",
        " AUTOSCALIN  G : SCALE_OUT_  ERROR ",
        " AUTOSCALIN  G : SCALE_IN_E  RROR ",
        " AUTOSCALIN  G : SCALE_REJE  CT "
    ],
    " requestId ": " 21F638C3 - 7054 - 4C1E - A143 - A74C20F507  A3 "
}
```

16 Error codes

16.1 Client errors

This topic describes the client errors.

Error message	Error code	Description	HTTP status code
The parameter is missing.	MissingParameter	The input parameter < parameter name > that is mandatory for processing this request is not supplied	400
The parameter value is invalid.	InvalidParameter	The specified value of parameter < parameter name > is not valid.	400
The interface is invalid.	UnsupportedOperation	The specified action is not supported.	400
The version is invalid.	NoSuchVersion	The specified version does not exist.	400
The operation is rejected by the traffic control system.	Throttling	Request was denied due to request throttling.	400
The AccessKey is invalid.	InvalidAccessKeyId.NotFound	The AccessKey ID provided does not exist in our records.	400
The operation is forbidden.	Forbidden	Users are not authorized to operate on the specified resource.	403

Error message	Error code	Description	HTTP status code
The operation is forbidden by the risk control system.	Forbidden.RiskControl	This operation is forbidden by Aliyun Risk Control system.	403
The signature is invalid.	SignatureDoesNotMatch	The signature we calculated does not match the one you provided.	403
Auto Scaling is not activated and thus the API cannot be called.	Forbidden.Unsubscribed	Do not have permission to access this API.	403
Real-name verification is missing.	Forbidden.UserVerification	Your user account is not verified by Aliyun.	403
Auto Scaling is not supported by the specified region.	ResourceNotAvailable	Resource you requested is not available in this region or zone.	400

16.2 Server errors

This topic introduces the server errors.

Error message	Error code	Description	HTTP status code
The server fails to process the request.	InternalError	The request processing has failed due to some unknown error, exception or failure.	500
The server currently fails to process the request.	ServiceUnavailable	The request has failed due to a temporary failure of the server.	503

17 How to ensure idempotence

This topic introduces how to ensure idempotence.

If a request timeout or internal server error is encountered when the “Execute a Scaling Rule” interface is called to create or release an ECS instance, the client might try to resend the request. In this case, the client can prevent the server from creating more instances than expected by providing the optional ClientToken parameter. This parameter also ensures the idempotence of the request. ClientToken is a unique and case sensitive string which is generated by the client and cannot contain more than 64 ASCII characters.

If you use the same ClientToken value to call the CreateInstance interface, the server returns identical request results that contain the same ScalingActivityId. Therefore, when you encounter an error and try again, by providing the same ClientToken value, you can ensure that only one scaling activity is created and the corresponding ScalingActivityId is obtained.

If you provide a ClientToken that has already been used, but with different request parameters, the Auto Scaling service returns the IdempotentParameterMismatch error code. However, note that you must change the SignatureNonce, TimeStamp, and Signature parameters. Because the Auto Scaling service uses SignatureNonce to prevent replay attacks and TimeStamp to mark the time of each request, the second request must provide different SignatureNonce and TimeStamp parameter values. This also produces a different Signature value.

Generally, you only need to retry the client in the case of Error 503 (ServiceUnAvailable) or no response. When Error 200 is returned, a retry generates the same results as the last time, without affecting the server status. When Error 4xx or 500 is returned, a retry usually may fail unless the message clearly indicates try it later.