
Auto Scaling

API Reference

API Version 2010-08-01



Auto Scaling: API Reference

Copyright © 2010 - 2011 Amazon Web Services LLC or its affiliates. All rights reserved.

Table of Contents

Welcome	1
Actions	2
CreateAutoScalingGroup	4
CreateLaunchConfiguration	6
DeleteAutoScalingGroup	8
DeleteLaunchConfiguration	9
DeletePolicy	10
DeleteScheduledAction	11
DescribeAdjustmentTypes	12
DescribeAutoScalingGroups	13
DescribeAutoScalingInstances	14
DescribeLaunchConfigurations	15
DescribeMetricCollectionTypes	16
DescribePolicies	17
DescribeScalingActivities	19
DescribeScalingProcessTypes	21
DescribeScheduledActions	22
DisableMetricsCollection	24
EnableMetricsCollection	25
ExecutePolicy	26
PutScalingPolicy	27
PutScheduledUpdateGroupAction	29
ResumeProcesses	30
SetDesiredCapacity	31
SetInstanceHealth	32
SuspendProcesses	33
TerminateInstanceInAutoScalingGroup	34
UpdateAutoScalingGroup	35
Data Types	37
Activity	38
AdjustmentType	39
Alarm	39
AutoScalingGroup	39
AutoScalingInstanceDetails	41
BlockDeviceMapping	42
DescribeAdjustmentTypesResult	42
DescribeAutoScalingGroupsResult	42
DescribeAutoScalingInstancesResult	43
DescribeLaunchConfigurationsResult	43
DescribeMetricCollectionTypesResult	43
DescribePoliciesResult	44
DescribeScalingActivitiesResult	44
DescribeScalingProcessTypesResult	45
DescribeScheduledActionsResult	45
Ebs	45
EnabledMetric	46
Instance	46
InstanceMonitoring	47
LaunchConfiguration	47
MetricCollectionType	48
MetricGranularityType	49
ProcessType	49
PutScalingPolicyResult	50
ScalingPolicy	51
ScheduledUpdateGroupAction	52

SuspendedProcess	52
TerminateInstanceInAutoScalingGroupResult	53
Common Query Parameters	54
Common Errors	56

Welcome

Auto Scaling is a web service designed to automatically launch or terminate EC2 instances based on user-defined policies, schedules, and health checks. Auto Scaling responds automatically to changing conditions. All you need to do is specify how it should respond to those changes.

Auto Scaling groups can work across multiple Availability Zones - distinct physical locations for the hosted Amazon EC2 instances - so that if an Availability Zone becomes unavailable, Auto Scaling will automatically redistribute applications to a different Availability Zone.

Every API call returns a response meta data object that contains a request identifier. Successful requests return an HTTP 200 status code. Unsuccessful requests return an error object and an HTTP status code of 400 or 500.

The current WSDL is available at:

<http://autoscaling.amazonaws.com/doc/2010-08-01/AutoScaling.wsdl>

Endpoints

For information about this product's regions and endpoints, go to [Regions and Endpoints](#) in the Amazon Web Services General Reference.

This document was last updated on April 29, 2011.

Actions

The actions described in this guide are called using the AWS Query protocol.

The following actions are supported:

- [CreateAutoScalingGroup](#) (p. 4)
- [CreateLaunchConfiguration](#) (p. 6)
- [DeleteAutoScalingGroup](#) (p. 8)
- [DeleteLaunchConfiguration](#) (p. 9)
- [DeletePolicy](#) (p. 10)
- [DeleteScheduledAction](#) (p. 11)
- [DescribeAdjustmentTypes](#) (p. 12)
- [DescribeAutoScalingGroups](#) (p. 13)
- [DescribeAutoScalingInstances](#) (p. 14)
- [DescribeLaunchConfigurations](#) (p. 15)
- [DescribeMetricCollectionTypes](#) (p. 16)
- [DescribePolicies](#) (p. 17)
- [DescribeScalingActivities](#) (p. 19)
- [DescribeScalingProcessTypes](#) (p. 21)
- [DescribeScheduledActions](#) (p. 22)
- [DisableMetricsCollection](#) (p. 24)
- [EnableMetricsCollection](#) (p. 25)
- [ExecutePolicy](#) (p. 26)
- [PutScalingPolicy](#) (p. 27)
- [PutScheduledUpdateGroupAction](#) (p. 29)
- [ResumeProcesses](#) (p. 30)
- [SetDesiredCapacity](#) (p. 31)
- [SetInstanceHealth](#) (p. 32)
- [SuspendProcesses](#) (p. 33)

- [TerminateInstanceInAutoScalingGroup](#) (p. 34)
- [UpdateAutoScalingGroup](#) (p. 35)

CreateAutoScalingGroup

Description

Creates a new Auto Scaling group with the specified name. Once the creation request is completed, the AutoScalingGroup is ready to be used in other calls.



Note

The Auto Scaling group name must be unique within the scope of your AWS account, and under the quota of Auto Scaling groups allowed for your account.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>AvailabilityZones.member.N</i>	A list of availability zones for the Auto Scaling group. Type: String list Length constraints: Minimum of 1 item(s) in the list.	Yes
<i>DefaultCooldown</i>	The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start. Type: Integer	No
<i>DesiredCapacity</i>	The number of EC2 instances that should be running in the group. For more information, see SetDesiredCapacity (p. 31) . Type: Integer	No
<i>HealthCheckGracePeriod</i>	Length of time in seconds after a new EC2 instance comes into service that Auto Scaling starts checking its health. Type: Integer	No
<i>HealthCheckType</i>	The service you want the health status from, Amazon EC2 or Elastic Load Balancer. Valid values are "EC2" or "ELB." Type: String Length constraints: Minimum value of 1. Maximum value of 32.	No

Auto Scaling API Reference Errors

Name	Description	Required
<i>LaunchConfigurationName</i>	The name of the launch configuration to use with the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>LoadBalancerNames.member.N</i>	A list of LoadBalancers to use. Type: String list	No
<i>MaxSize</i>	The maximum size of the Auto Scaling group. Type: Integer	Yes
<i>MinSize</i>	The minimum size of the Auto Scaling group. Type: Integer	Yes
<i>PlacementGroup</i>	Physical location of your cluster placement group created in Amazon EC2. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	No
<i>VPCZoneIdentifier</i>	The subnet identifier of the Virtual Private Cloud. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	No

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
<i>AlreadyExists</i>	The named AutoScalingGroup or launch configuration already exists.	400
<i>LimitExceeded</i>	The quota for capacity groups or launch configurations for this customer has already been reached.	400

CreateLaunchConfiguration

Description

Creates a new launch configuration. Once created, the new launch configuration is available for immediate use.



Note

The launch configuration name used must be unique, within the scope of the client's AWS account, and the maximum limit of launch configurations must not yet have been met, or else the call will fail.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>BlockDeviceMappings.member.N</i>	A list of mappings that specify how block devices are exposed to the instance. Each mapping is made up of a <i>VirtualName</i> , a <i>DeviceName</i> , and an <i>ebs</i> data structure that contains information about the associated Elastic Block Storage volume. For more information about Amazon EC2 <i>BlockDeviceMappings</i> , please go to Block Device Mapping in the <i>Amazon EC2 User Guide</i> . Type: BlockDeviceMapping (p. 42) list	No
<i>ImageId</i>	Unique ID of the <i>Amazon Machine Image</i> (AMI) which was assigned during registration. For more information about Amazon EC2 images, please go to Using AMIs in the <i>Amazon EC2 User Guide</i> . Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>InstanceMonitoring</i>	Enables detailed monitoring. Type: InstanceMonitoring (p. 47)	No
<i>InstanceType</i>	The instance type of the EC2 instance. For more information about Amazon EC2 instance types, please go to Using Instances in the <i>Amazon EC2 User Guide</i> . Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>KernelId</i>	The ID of the kernel associated with the EC2 AMI. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	No

Name	Description	Required
<i>KeyName</i>	The name of the EC2 key pair. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	No
<i>LaunchConfigurationName</i>	The name of the launch configuration to create. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>RamdiskId</i>	The ID of the RAM disk associated with the EC2 AMI. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	No
<i>SecurityGroups.member.N</i>	The names of the security groups with which to associate EC2 instances. For more information about Amazon EC2 security groups, go to Using Security Groups in the <i>Amazon EC2 User Guide</i> . Type: String list	No
<i>UserData</i>	The user data available to the launched EC2 instances. For more information about Amazon EC2 user data, please go to Using Instances in the <i>Amazon EC2 User Guide</i> . Type: String Length constraints: Minimum value of 0. Maximum value of 21847.	No

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
<i>AlreadyExists</i>	The named AutoScalingGroup or launch configuration already exists.	400
<i>LimitExceeded</i>	The quota for capacity groups or launch configurations for this customer has already been reached.	400

DeleteAutoScalingGroup

Description

Deletes the specified auto scaling group if the group has no instances and no scaling activities in progress.



Note

To remove all instances before calling [DeleteAutoScalingGroup](#) (p. 8), you can call [UpdateAutoScalingGroup](#) (p. 35) to set the minimum and maximum size of the AutoScalingGroup to zero.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes

Errors

For information about the common errors that all actions use, see [Common Errors](#) (p. 56).

Error	Description	HTTP Status Code
<i>ResourceInUse</i>	This is returned when you cannot delete a launch configuration or auto scaling group because it is being used.	400
<i>ScalingActivityInProgress</i>	You cannot delete an AutoScalingGroup while there are scaling activities in progress for that group.	400

DeleteLaunchConfiguration

Description

Deletes the specified [LaunchConfiguration](#) (p. 47).

The specified launch configuration must not be attached to an Auto Scaling group. Once this call completes, the launch configuration is no longer available for use.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>LaunchConfigurationName</i>	The name of the launch configuration. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes

Errors

For information about the common errors that all actions use, see [Common Errors](#) (p. 56).

Error	Description	HTTP Status Code
<i>ResourceInUse</i>	This is returned when you cannot delete a launch configuration or auto scaling group because it is being used.	400

DeletePolicy

Description

Deletes a policy created by [PutScalingPolicy](#) (p. 27)

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	No
<i>PolicyName</i>	The name or PolicyARN of the policy you want to delete Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes

DeleteScheduledAction

Description

Deletes a scheduled action previously created using the [PutScheduledUpdateGroupAction](#) (p. 29).

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the Auto Scaling group Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	No
<i>ScheduledActionName</i>	The name of the action you want to delete. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes

DescribeAdjustmentTypes

Description

Returns policy adjustment types for use in the [PutScalingPolicy](#) (p. 27) action.

Response Elements

The following elements come wrapped in a `DescribeAdjustmentTypesResult` structure.

Name	Description
AdjustmentTypes	A list of specific policy adjustment types. Type: AdjustmentType (p. 39) list

DescribeAutoScalingGroups

Description

Returns a full description of each Auto Scaling group in the given list. This includes all Amazon EC2 instances that are members of the group. If a list of names is not provided, the service returns the full details of all Auto Scaling groups.

This action supports pagination by returning a token if there are more pages to retrieve. To get the next page, call this action again with the returned token as the `NextToken` parameter.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>AutoScalingGroupNames.member.N</i>	A list of Auto Scaling group names. Type: String list	No
<i>MaxRecords</i>	The maximum number of records to return. Type: Integer	No
<i>NextToken</i>	A string that marks the start of the next batch of returned results. Type: String	No

Response Elements

The following elements come wrapped in a `DescribeAutoScalingGroupsResult` structure.

Name	Description
<code>AutoScalingGroups</code>	A list of Auto Scaling groups. Type: AutoScalingGroup (p. 39) list
<code>NextToken</code>	A string that marks the start of the next batch of returned results. Type: String

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
<code>InvalidNextToken</code>	The <code>NextToken</code> value is invalid.	400

DescribeAutoScalingInstances

Description

Returns a description of each Auto Scaling instance in the InstanceIds list. If a list is not provided, the service returns the full details of all instances up to a maximum of fifty.

This action supports pagination by returning a token if there are more pages to retrieve. To get the next page, call this action again with the returned token as the NextToken parameter.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>InstanceIds.member.N</i>	The list of Auto Scaling instances to describe. If this list is omitted, all auto scaling instances are described. The list of requested instances cannot contain more than 50 items. If unknown instances are requested, they are ignored with no error. Type: String list	No
<i>MaxRecords</i>	The maximum number of Auto Scaling instances to be described with each call. Type: Integer	No
<i>NextToken</i>	The token returned by a previous call to indicate that there is more data available. Type: String	No

Response Elements

The following elements come wrapped in a `DescribeAutoScalingInstancesResult` structure.

Name	Description
<code>AutoScalingInstances</code>	A list of Auto Scaling instances. Type: AutoScalingInstanceDetails (p. 41) list
<code>NextToken</code>	A string that marks the start of the next batch of returned results. Type: String

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
<code>InvalidNextToken</code>	The <code>NextToken</code> value is invalid.	400

DescribeLaunchConfigurations

Description

Returns a full description of the launch configurations given the specified names.

If no names are specified, then the full details of all launch configurations are returned.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>LaunchConfigurationNames.member.N</i>	A list of launch configuration names. Type: String list	No
<i>MaxRecords</i>	The maximum number of launch configurations. Type: Integer	No
<i>NextToken</i>	A string that marks the start of the next batch of returned results. Type: String	No

Response Elements

The following elements come wrapped in a `DescribeLaunchConfigurationsResult` structure.

Name	Description
LaunchConfigurations	A list of launch configurations. Type: LaunchConfiguration (p. 47) list
NextToken	A string that marks the start of the next batch of returned results. Type: String

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
InvalidNextToken	The NextToken value is invalid.	400

DescribeMetricCollectionTypes

Description

Returns a list of metrics and a corresponding list of granularities for each metric.

Response Elements

The following elements come wrapped in a `DescribeMetricCollectionTypesResult` structure.

Name	Description
Granularities	A list of Granularities for the listed Metrics. Type: MetricGranularityType (p. 49) list
Metrics	The list of Metrics collected. Type: MetricCollectionType (p. 48) list

DescribePolicies

Description

Returns descriptions of what each policy does. This action supports pagination. If the response includes a token, there are more records available. To get the additional records, repeat the request with the response token as the `NextToken` parameter.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	No
<i>MaxRecords</i>	The maximum number of policies that will be described with each call. Type: Integer	No
<i>NextToken</i>	A string that is used to mark the start of the next batch of returned results for pagination. Type: String	No
<i>PolicyNames.member.N</i>	A list of policy names or policy ARNs to be described. If this list is omitted, all policy names are described. If an auto scaling group name is provided, the results are limited to that group. The list of requested policy names cannot contain more than 50 items. If unknown policy names are requested, they are ignored with no error. Type: String list	No

Response Elements

The following elements come wrapped in a `DescribePoliciesResult` structure.

Name	Description
<i>NextToken</i>	A string that marks the start of the next batch of returned results. Type: String
<i>ScalingPolicies</i>	A list of scaling policies. Type: ScalingPolicy (p. 51) list

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Auto Scaling API Reference Errors

Error	Description	HTTP Status Code
InvalidNextToken	The NextToken value is invalid.	400

DescribeScalingActivities

Description

Returns the scaling activities for the specified Auto Scaling group.

If the specified *ActivityIds* list is empty, all the activities from the past six weeks are returned. Activities are sorted by completion time. Activities still in progress appear first on the list.

This action supports pagination. If the response includes a token, there are more records available. To get the additional records, repeat the request with the response token as the *NextToken* parameter.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>ActivityIds.member.N</i>	A list containing the activity IDs of the desired scaling activities. If this list is omitted, all activities are described. If an <i>AutoScalingGroupName</i> is provided, the results are limited to that group. The list of requested activities cannot contain more than 50 items. If unknown activities are requested, they are ignored with no error. Type: String list	No
<i>AutoScalingGroupName</i>	The name of the AutoScalingGroup (p. 39) . Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	No
<i>MaxRecords</i>	The maximum number of scaling activities to return. Type: Integer	No
<i>NextToken</i>	A string that marks the start of the next batch of returned results for pagination. Type: String	No

Response Elements

The following elements come wrapped in a *DescribeScalingActivitiesResult* structure.

Name	Description
<i>Activities</i>	A list of the requested scaling activities. Type: Activity (p. 38) list
<i>NextToken</i>	Acts as a paging mechanism for large result sets. Set to a non-empty string if there are additional results waiting to be returned. Pass this in to subsequent calls to return additional results. Type: String

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
InvalidNextToken	The <code>NextToken</code> value is invalid.	400

DescribeScalingProcessTypes

Description

Returns scaling process types for use in the [ResumeProcesses](#) (p. 30) and [SuspendProcesses](#) (p. 33) actions.

Response Elements

The following elements come wrapped in a `DescribeScalingProcessTypesResult` structure.

Name	Description
Processes	A list of ProcessType (p. 49) names. Type: ProcessType (p. 49) list

DescribeScheduledActions

Description

Lists all the actions scheduled for your Auto Scaling group that haven't been executed. To see a list of action already executed, see the activity record returned in [DescribeScalingActivities](#) (p. 19).

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	No
<i>EndTime</i>	The latest scheduled start time to return. If scheduled action names are provided, this field will be ignored. Type: DateTime	No
<i>MaxRecords</i>	The maximum number of scheduled actions to return. Type: Integer	No
<i>NextToken</i>	A string that marks the start of the next batch of returned results. Type: String	No
<i>ScheduledActionNames.member.N</i>	A list of scheduled actions to be described. If this list is omitted, all scheduled actions are described. The list of requested scheduled actions cannot contain more than 50 items. If an auto scaling group name is provided, the results are limited to that group. If unknown scheduled actions are requested, they are ignored with no error. Type: String list	No
<i>StartTime</i>	The earliest scheduled start time to return. If scheduled action names are provided, this field will be ignored. Type: DateTime	No

Response Elements

The following elements come wrapped in a `DescribeScheduledActionsResult` structure.

Name	Description
NextToken	A string that marks the start of the next batch of returned results. Type: String
ScheduledUpdateGroupActions	A list of scheduled actions designed to update an Auto Scaling group. Type: ScheduledUpdateGroupAction (p. 52) list

Errors

For information about the common errors that all actions use, see [Common Errors](#) (p. 56).

Error	Description	HTTP Status Code
InvalidNextToken	The NextToken value is invalid.	400

DisableMetricsCollection

Description

Disables monitoring of group metrics for the Auto Scaling group specified in `AutoScalingGroupName`. You can specify the list of affected metrics with the `Metrics` parameter.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name or ARN of the Auto Scaling Group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>Metrics.member.N</i>	The list of metrics to disable. If no metrics are specified, all metrics are disabled. The following metrics are supported: <ul style="list-style-type: none">GroupMinSizeGroupMaxSizeGroupDesiredCapacityGroupInServiceInstancesGroupPendingInstancesGroupTerminatingInstancesGroupTotalInstances Type: String list	No

EnableMetricsCollection

Description

Enables monitoring of group metrics for the Auto Scaling group specified in `AutoScalingGroupName`. You can specify the list of enabled metrics with the `Metrics` parameter.

Auto scaling metrics collection can be turned on only if the `InstanceMonitoring.Enabled` flag, in the Auto Scaling group's launch configuration, is set to `true`.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name or ARN of the Auto Scaling Group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>Granularity</i>	The granularity to associate with the metrics to collect. Currently, the only legal granularity is "1Minute". Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>Metrics.member.N</i>	The list of metrics to collect. If no metrics are specified, all metrics are enabled. The following metrics are supported: <ul style="list-style-type: none">GroupMinSizeGroupMaxSizeGroupDesiredCapacityGroupInServiceInstancesGroupPendingInstancesGroupTerminatingInstancesGroupTotalInstances Type: String list	No

ExecutePolicy

Description

Runs the policy you create for your Auto Scaling group in [PutScalingPolicy](#) (p. 27).

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name or ARN of the Auto Scaling Group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	No
<i>HonorCooldown</i>	Set to True if you want Auto Scaling to reject this request if the Auto Scaling group is in cooldown. Type: Boolean	No
<i>PolicyName</i>	The name or PolicyARN of the policy you want to run. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes

Errors

For information about the common errors that all actions use, see [Common Errors](#) (p. 56).

Error	Description	HTTP Status Code
<i>ScalingActivityInProgress</i>	You cannot delete an AutoScalingGroup while there are scaling activities in progress for that group.	400

PutScalingPolicy

Description

Creates or updates a policy for an Auto Scaling group. To update an existing policy, use the existing policy name and set the parameter(s) you want to change. Any existing parameter not changed in an update to an existing policy is not changed in this update request.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>AdjustmentType</i>	Specifies whether the <code>ScalingAdjustment</code> is an absolute number or a percentage of the current capacity. Valid values are <code>ChangeInCapacity</code> , <code>ExactCapacity</code> , and <code>PercentChangeInCapacity</code> . Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>AutoScalingGroupName</i>	The name or ARN of the Auto Scaling Group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>Cooldown</i>	The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start. Type: Integer	No
<i>PolicyName</i>	The name of the policy you want to create or update. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>ScalingAdjustment</i>	The number of instances by which to scale. <code>AdjustmentType</code> determines the interpretation of this number (e.g., as an absolute number or as a percentage of the existing Auto Scaling group size). A positive increment adds to the current capacity and a negative value removes from the current capacity. Type: Integer	Yes

Response Elements

The following elements come wrapped in a `PutScalingPolicyResult` structure.

Name	Description
<code>PolicyARN</code>	A policy's Amazon Resource Name (ARN). Type: String

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
LimitExceeded	The quota for capacity groups or launch configurations for this customer has already been reached.	400

PutScheduledUpdateGroupAction

Description

Creates a scheduled scaling action for a Auto Scaling group. If you leave a parameter unspecified, the corresponding value remains unchanged in the affected Auto Scaling group.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name or ARN of the Auto Scaling Group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>DesiredCapacity</i>	The number of EC2 instances that should be running in the group. Type: Integer	No
<i>MaxSize</i>	The maximum size for the Auto Scaling group. Type: Integer	No
<i>MinSize</i>	The minimum size for the new Auto Scaling group. Type: Integer	No
<i>ScheduledActionName</i>	The name of this scaling action. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	Yes
<i>Time</i>	The time for this action to start. Type: DateTime	Yes

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
<i>AlreadyExists</i>	The named AutoScalingGroup or launch configuration already exists.	400
<i>LimitExceeded</i>	The quota for capacity groups or launch configurations for this customer has already been reached.	400

ResumeProcesses

Description

Resumes Auto Scaling processes for an Auto Scaling group. For more information, see [SuspendProcesses](#) (p. 33) and [ProcessType](#) (p. 49).

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name or Amazon Resource Name (ARN) of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>ScalingProcesses.member.N</i>	The processes that you want to suspend or resume, which can include one or more of the following: <ul style="list-style-type: none">• Launch• Terminate• HealthCheck• ReplaceUnhealthy• AZRebalance• AlarmNotifications• ScheduledActions To suspend all process types, omit this parameter. Type: String list	No

SetDesiredCapacity

Description

Adjusts the desired size of the [AutoScalingGroup](#) (p. 39) by initiating scaling activities. When reducing the size of the group, it is not possible to define which EC2 instances will be terminated. This applies to any auto-scaling decisions that might result in terminating instances.

There are two common use cases for `SetDesiredCapacity`: one for users of the Auto Scaling triggering system, and another for developers who write their own triggering systems. Both use cases relate to the concept of cooldown.

In the first case, if you use the Auto Scaling triggering system, `SetDesiredCapacity` changes the size of your Auto Scaling group without regard to the cooldown period. This could be useful, for example, if Auto Scaling did something unexpected for some reason. If your cooldown period is 10 minutes, Auto Scaling would normally reject requests to change the size of the group for that entire 10 minute period. The `SetDesiredCapacity` command allows you to circumvent this restriction and change the size of the group before the end of the cooldown period.

In the second case, if you write your own triggering system, you can use `SetDesiredCapacity` to control the size of your Auto Scaling group. If you want the same cooldown functionality that Auto Scaling offers, you can configure `SetDesiredCapacity` to honor cooldown by setting the `HonorCooldown` parameter to `true`.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the AutoScalingGroup. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>DesiredCapacity</i>	The new capacity setting for the AutoScalingGroup. Type: Integer	Yes
<i>HonorCooldown</i>	By default, <code>SetDesiredCapacity</code> overrides any cooldown period. Set to <code>True</code> if you want Auto Scaling to reject this request if the Auto Scaling group is in cooldown. Type: Boolean	No

Errors

For information about the common errors that all actions use, see [Common Errors](#) (p. 56).

Error	Description	HTTP Status Code
<code>ScalingActivityInProgress</code>	You cannot delete an AutoScalingGroup while there are scaling activities in progress for that group.	400

SetInstanceHealth

Description

Sets the health status of an instance.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>HealthStatus</i>	The health status of the instance. "Healthy" means that the instance is healthy and should remain in service. "Unhealthy" means that the instance is unhealthy. Auto Scaling should terminate and replace it. Type: String Length constraints: Minimum value of 1. Maximum value of 32.	Yes
<i>InstanceId</i>	The identifier of the EC2 instance. Type: String Length constraints: Minimum value of 1. Maximum value of 16.	Yes
<i>ShouldRespectGracePeriod</i>	If True, this call should respect the grace period associated with the group. Type: Boolean	No

SuspendProcesses

Description

Suspends Auto Scaling processes for an Auto Scaling group. To suspend specific process types, specify them by name with the `ScalingProcesses.member.N` parameter. To suspend all process types, omit the `ScalingProcesses.member.N` parameter.



Important

Suspending either of the two primary process types, `Launch` or `Terminate`, can prevent other process types from functioning properly. For more information about processes and their dependencies, see [ProcessType](#) (p. 49).

To resume processes that have been suspended, use [ResumeProcesses](#) (p. 30).

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<code>AutoScalingGroupName</code>	The name or Amazon Resource Name (ARN) of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<code>ScalingProcesses.member.N</code>	The processes that you want to suspend or resume, which can include one or more of the following: <ul style="list-style-type: none">• <code>Launch</code>• <code>Terminate</code>• <code>HealthCheck</code>• <code>ReplaceUnhealthy</code>• <code>AZRebalance</code>• <code>AlarmNotifications</code>• <code>ScheduledActions</code> To suspend all process types, omit this parameter. Type: String list	No

TerminateInstanceInAutoScalingGroup

Description

Terminates the specified instance. Optionally, the desired group size can be adjusted.



Note

This call simply registers a termination request. The termination of the instance cannot happen immediately.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters \(p. 54\)](#).

Name	Description	Required
<i>InstanceId</i>	The ID of the EC2 instance to be terminated. Type: String Length constraints: Minimum value of 1. Maximum value of 16.	Yes
<i>ShouldDecrementDesiredCapacity</i>	Specifies whether (<i>true</i>) or not (<i>false</i>) terminating this instance should also decrement the size of the AutoScalingGroup (p. 39) . Type: Boolean	Yes

Response Elements

The following elements come wrapped in a `TerminateInstanceInAutoScalingGroupResult` structure.

Name	Description
Activity	A Scaling Activity. Type: Activity (p. 38)

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
ScalingActivityInProgress	You cannot delete an AutoScalingGroup while there are scaling activities in progress for that group.	400

UpdateAutoScalingGroup

Description

Updates the configuration for the specified [AutoScalingGroup](#) (p. 39).



Note

To update an Auto Scaling group with a launch configuration that has the `InstanceMonitoring.enabled` flag set to `false`, you must first ensure that collection of group metrics is disabled. Otherwise, calls to [UpdateAutoScalingGroup](#) (p. 35) will fail. If you have previously enabled group metrics collection, you can disable collection of all group metrics by calling [DisableMetricsCollection](#) (p. 24).

The new settings are registered upon the completion of this call. Any launch configuration settings take effect on any triggers after this call returns. Triggers that are currently in progress aren't affected.



Note

If the new values are specified for the `MinSize` or `MaxSize` parameters, then there will be an implicit call to [SetDesiredCapacity](#) (p. 31) to set the group to the new `MaxSize`. All optional parameters are left unchanged if not passed in the request.

Request Parameters

For information about the common parameters that all actions use, see [Common Query Parameters](#) (p. 54).

Name	Description	Required
<i>AutoScalingGroupName</i>	The name of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	Yes
<i>AvailabilityZones.member.N</i>	Availability zones for the group. Type: String list Length constraints: Minimum of 1 item(s) in the list.	No
<i>DefaultCooldown</i>	The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start. Type: Integer	No
<i>DesiredCapacity</i>	The desired capacity for the Auto Scaling group. Type: Integer	No
<i>HealthCheckGracePeriod</i>	The length of time that Auto Scaling waits before checking an instance's health status. The grace period begins when an instance comes into service. Type: Integer	No

Name	Description	Required
<i>HealthCheckType</i>	The service of interest for the health status check, either "EC2" for Amazon EC2 or "ELB" for Elastic Load Balancing. Type: String Length constraints: Minimum value of 1. Maximum value of 32.	No
<i>LaunchConfigurationName</i>	The name of the launch configuration. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.	No
<i>MaxSize</i>	The maximum size of the Auto Scaling group. Type: Integer	No
<i>MinSize</i>	The minimum size of the Auto Scaling group. Type: Integer	No
<i>PlacementGroup</i>	The name of the cluster placement group, if applicable. For more information, go to Using Cluster Instances in the <i>Amazon EC2 User Guide</i> . Type: String Length constraints: Minimum value of 1. Maximum value of 255.	No
<i>VPCZoneIdentifier</i>	The identifier for the VPC connection, if applicable. Type: String Length constraints: Minimum value of 1. Maximum value of 255.	No

Errors

For information about the common errors that all actions use, see [Common Errors \(p. 56\)](#).

Error	Description	HTTP Status Code
<i>ScalingActivityInProgress</i>	You cannot delete an AutoScalingGroup while there are scaling activities in progress for that group.	400

Data Types

The Auto Scaling API contains several data types that various actions use. This section describes each data type in detail.

**Note**

The order of each element in the response is not guaranteed. Applications should not assume a particular order.

The following data types are supported:

- [Activity](#) (p. 38)
- [AdjustmentType](#) (p. 39)
- [Alarm](#) (p. 39)
- [AutoScalingGroup](#) (p. 39)
- [AutoScalingInstanceDetails](#) (p. 41)
- [BlockDeviceMapping](#) (p. 42)
- [DescribeAdjustmentTypesResult](#) (p. 42)
- [DescribeAutoScalingGroupsResult](#) (p. 42)
- [DescribeAutoScalingInstancesResult](#) (p. 43)
- [DescribeLaunchConfigurationsResult](#) (p. 43)
- [DescribeMetricCollectionTypesResult](#) (p. 43)
- [DescribePoliciesResult](#) (p. 44)
- [DescribeScalingActivitiesResult](#) (p. 44)
- [DescribeScalingProcessTypesResult](#) (p. 45)
- [DescribeScheduledActionsResult](#) (p. 45)
- [Ebs](#) (p. 45)
- [EnabledMetric](#) (p. 46)
- [Instance](#) (p. 46)
- [InstanceMonitoring](#) (p. 47)
- [LaunchConfiguration](#) (p. 47)

- [MetricCollectionType](#) (p. 48)
- [MetricGranularityType](#) (p. 49)
- [ProcessType](#) (p. 49)
- [PutScalingPolicyResult](#) (p. 50)
- [ScalingPolicy](#) (p. 51)
- [ScheduledUpdateGroupAction](#) (p. 52)
- [SuspendedProcess](#) (p. 52)
- [TerminateInstanceInAutoScalingGroupResult](#) (p. 53)

Activity

Description

A Scaling Activity is a long-running process that represents a change to your AutoScalingGroup, such as changing the size of the group. It can also be a process to replace an instance, or a process to perform any other long-running operations supported by the API.

Contents

Name	Description
ActivityId	Specifies the ID of the activity. Type: String
AutoScalingGroupName	The name of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
Cause	Contains the reason the activity was begun. Type: String Length constraints: Minimum value of 1. Maximum value of 1023.
Description	Contains a friendly, more verbose description of the scaling activity. Type: String
EndTime	Provides the end time of this activity. Type: DateTime
Progress	Specifies a value between 0 and 100 that indicates the progress of the activity. Type: Integer
StartTime	Provides the start time of this activity. Type: DateTime
StatusCode	Contains the current status of the activity. Type: String

Name	Description
StatusMessage	Contains a friendly, more verbose description of the activity status. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

AdjustmentType

Description

Specifies whether the [PutScalingPolicy](#) (p. 27) `ScalingAdjustment` parameter is an absolute number or a percentage of the current capacity.

Contents

Name	Description
AdjustmentType	A policy adjustment type. Valid values are <code>ChangeInCapacity</code> , <code>ExactCapacity</code> , and <code>PercentChangeInCapacity</code> . Type: String Length constraints: Minimum value of 1. Maximum value of 255.

Alarm

Description

The Alarm data type.

Contents

Name	Description
AlarmARN	The Amazon Resource Name (ARN) of the alarm. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.
AlarmName	The name of the alarm Type: String Length constraints: Minimum value of 1. Maximum value of 255.

AutoScalingGroup

Description

The AutoScalingGroup data type.

Contents

Name	Description
AutoScalingGroupARN	The Amazon Resource Name (ARN) of the Auto Scaling group. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.
AutoScalingGroupName	Specifies the name of the group. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
AvailabilityZones	Contains a list of availability zones for the group. Type: String list Length constraints: Minimum of 1 item(s) in the list.
CreatedTime	Specifies the date and time the Auto Scaling group was created. Type: DateTime
DefaultCooldown	The number of seconds after a scaling activity completes before any further scaling activities can start. Type: Integer
DesiredCapacity	Specifies the desired capacity for the AutoScalingGroup. Type: Integer
EnabledMetrics	A list of metrics enabled for this Auto Scaling group. Type: EnabledMetric (p. 46) list
HealthCheckGracePeriod	The length of time that Auto Scaling waits before checking an instance's health status. The grace period begins when an instance comes into service. Type: Integer
HealthCheckType	The service of interest for the health status check, either "EC2" for Amazon EC2 or "ELB" for Elastic Load Balancing. Type: String Length constraints: Minimum value of 1. Maximum value of 32.
Instances	Provides a summary list of EC2 instances. Type: Instance (p. 46) list
LaunchConfigurationName	Specifies the name of the associated LaunchConfiguration (p. 47) . Type: String Length constraints: Minimum value of 1. Maximum value of 255.
LoadBalancerNames	A list of load balancers associated with this Auto Scaling group. Type: String list
MaxSize	Contains the maximum size of the AutoScalingGroup. Type: Integer
MinSize	Contains the minimum size of the AutoScalingGroup. Type: Integer

Name	Description
PlacementGroup	The name of the cluster placement group, if applicable. For more information, go to Using Cluster Instances in the <i>Amazon EC2 User Guide</i> . Type: String Length constraints: Minimum value of 1. Maximum value of 255.
SuspendedProcesses	Suspended processes associated with this Auto Scaling group. Type: SuspendedProcess (p. 52) list
VPCZoneIdentifier	The identifier for the VPC connection, if applicable. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

AutoScalingInstanceDetails

Description

The AutoScalingInstanceDetails data type.

Contents

Name	Description
AutoScalingGroupName	The name of the Auto Scaling group associated with this instance. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
AvailabilityZone	The availability zone in which this instance resides. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
HealthStatus	The health status of this instance. "Healthy" means that the instance is healthy and should remain in service. "Unhealthy" means that the instance is unhealthy. Auto Scaling should terminate and replace it. Type: String Length constraints: Minimum value of 1. Maximum value of 32.
InstanceId	The instance's Amazon EC2 instance ID. Type: String Length constraints: Minimum value of 1. Maximum value of 16.
LaunchConfigurationName	The launch configuration associated with this instance. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
LifecycleState	The life cycle state of this instance. Type: String Length constraints: Minimum value of 1. Maximum value of 32.

BlockDeviceMapping

Description

The BlockDeviceMapping data type.

Contents

Name	Description
DeviceName	The name of the device within Amazon EC2. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
Ebs	The Elastic Block Storage volume information. Type: Ebs (p. 45)
VirtualName	The virtual name associated with the device. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

DescribeAdjustmentTypesResult

Description

The output of the [DescribeAdjustmentTypes](#) (p. 12) action.

Contents

Name	Description
AdjustmentTypes	A list of specific policy adjustment types. Type: AdjustmentType (p. 39) list

DescribeAutoScalingGroupsResult

Description

The AutoScalingGroupsType data type.

Contents

Name	Description
AutoScalingGroups	A list of Auto Scaling groups. Type: AutoScalingGroup (p. 39) list

Name	Description
NextToken	A string that marks the start of the next batch of returned results. Type: String

DescribeAutoScalingInstancesResult

Description

The AutoScalingInstancesType data type.

Contents

Name	Description
AutoScalingInstances	A list of Auto Scaling instances. Type: AutoScalingInstanceDetails (p. 41) list
NextToken	A string that marks the start of the next batch of returned results. Type: String

DescribeLaunchConfigurationsResult

Description

The LaunchConfigurationsType data type.

Contents

Name	Description
LaunchConfigurations	A list of launch configurations. Type: LaunchConfiguration (p. 47) list
NextToken	A string that marks the start of the next batch of returned results. Type: String

DescribeMetricCollectionTypesResult

Description

The output of the [DescribeMetricCollectionTypes](#) (p. 16) action.

Contents

Name	Description
Granularities	A list of Granularities for the listed Metrics. Type: MetricGranularityType (p. 49) list
Metrics	The list of Metrics collected. Type: MetricCollectionType (p. 48) list

DescribePoliciesResult

Description

The PoliciesType data type.

Contents

Name	Description
NextToken	A string that marks the start of the next batch of returned results. Type: String
ScalingPolicies	A list of scaling policies. Type: ScalingPolicy (p. 51) list

DescribeScalingActivitiesResult

Description

The output for the [DescribeScalingActivities](#) (p. 19) action.

Contents

Name	Description
Activities	A list of the requested scaling activities. Type: Activity (p. 38) list
NextToken	Acts as a paging mechanism for large result sets. Set to a non-empty string if there are additional results waiting to be returned. Pass this in to subsequent calls to return additional results. Type: String

DescribeScalingProcessTypesResult

Description

The output of the [DescribeScalingProcessTypes](#) (p. 21) action.

Contents

Name	Description
Processes	A list of ProcessType (p. 49) names. Type: ProcessType (p. 49) list

DescribeScheduledActionsResult

Description

A scaling action that is scheduled for a future time and date. An action can be scheduled up to thirty days in advance.

Contents

Name	Description
NextToken	A string that marks the start of the next batch of returned results. Type: String
ScheduledUpdateGroupActions	A list of scheduled actions designed to update an Auto Scaling group. Type: ScheduledUpdateGroupAction (p. 52) list

Ebs

Description

The Ebs data type.

Contents

Name	Description
SnapshotId	The Snapshot ID. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

Name	Description
VolumeSize	The volume size, in GigaBytes. Type: Integer

EnabledMetric

Description

The EnabledMetric data type.

Contents

Name	Description
Granularity	The granularity of the enabled metric. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
Metric	The name of the enabled metric. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

Instance

Description

The Instance data type.

Contents

Name	Description
AvailabilityZone	Availability zones associated with this instance. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
HealthStatus	The instance's health status. Type: String Length constraints: Minimum value of 1. Maximum value of 32.
InstanceId	Specifies the EC2 instance ID. Type: String Length constraints: Minimum value of 1. Maximum value of 16.

Name	Description
LaunchConfigurationName	The launch configuration associated with this instance. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
LifecycleState	Contains a description of the current <i>lifecycle</i> state. Type: String Length constraints: Minimum value of 1. Maximum value of 32.

InstanceMonitoring

Description

The InstanceMonitoring Data Type.

Contents

Name	Description
Enabled	If true, instance monitoring is enabled. Type: Boolean

LaunchConfiguration

Description

The LaunchConfiguration data type.

Contents

Name	Description
BlockDeviceMappings	Specifies how block devices are exposed to the instance. Each mapping is made up of a <i>virtualName</i> and a <i>deviceName</i> . Type: BlockDeviceMapping (p. 42) list
CreatedTime	Provides the creation date and time for this launch configuration. Type: DateTime
ImageId	Provides the unique ID of the <i>Amazon Machine Image</i> (AMI) that was assigned during registration. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
InstanceMonitoring	Controls whether instances in this group are launched with detailed monitoring or not. Type: InstanceMonitoring (p. 47)

Name	Description
InstanceType	Specifies the instance type of the EC2 instance. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
KernelId	Provides the ID of the kernel associated with the EC2 AMI. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
KeyName	Provides the name of the EC2 key pair. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
LaunchConfigurationARN	The launch configuration's Amazon Resource Name (ARN). Type: String Length constraints: Minimum value of 1. Maximum value of 1600.
LaunchConfigurationName	Specifies the name of the launch configuration. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
RamdiskId	Provides ID of the RAM disk associated with the EC2 AMI. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
SecurityGroups	A description of the security groups to associate with the EC2 instances. Type: String list
UserData	The user data available to the launched EC2 instances. Type: String Length constraints: Minimum value of 0. Maximum value of 21847.

MetricCollectionType

Description

The MetricCollectionType data type.

Contents

Name	Description
Metric	Type: String Length constraints: Minimum value of 1. Maximum value of 255.

MetricGranularityType

Description

The MetricGranularityType data type.

Contents

Name	Description
Granularity	The granularity of a Metric. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

ProcessType

Description

There are two primary Auto Scaling process types--`Launch` and `Terminate`. The `Launch` process creates a new Amazon EC2 instance for an Auto Scaling group, and the `Terminate` process removes an existing Amazon EC2 instance.

The remaining Auto Scaling process types relate to specific Auto Scaling features:

- `AlarmNotifications`
- `AZRebalance`
- `HealthCheck`
- `ReplaceUnhealthy`
- `ScheduledActions`



Important

If you suspend `Launch` or `Terminate`, all other process types are affected to varying degrees. The following descriptions discuss how each process type is affected by a suspension of `Launch` or `Terminate`.

The `AlarmNotifications` process type accepts notifications from Amazon CloudWatch alarms that are associated with the Auto Scaling group. If you suspend the `AlarmNotifications` process type, Auto Scaling will not automatically execute scaling policies that would be triggered by alarms.

Although the `AlarmNotifications` process type is not directly affected by a suspension of `Launch` or `Terminate`, alarm notifications are often used to signal that a change in the size of the Auto Scaling group is warranted. If you suspend `Launch` or `Terminate`, Auto Scaling might not be able to implement the alarm's associated policy.

The `AZRebalance` process type seeks to maintain a balanced number of instances across Availability Zones within a Region. If you remove an Availability Zone from your Auto Scaling group or an Availability Zone otherwise becomes unhealthy or unavailable, Auto Scaling launches new instances in an unaffected Availability Zone before terminating the unhealthy or unavailable instances. When the unhealthy Availability

Zone returns to a healthy state, Auto Scaling automatically redistributes the application instances evenly across all of the designated Availability Zones.



Important

If you call [SuspendProcesses \(p. 33\)](#) on the `launch` process type, the `AZRebalance` process will neither launch new instances nor terminate existing instances. This is because the `AZRebalance` process terminates existing instances only after launching the replacement instances.

If you call [SuspendProcesses \(p. 33\)](#) on the `terminate` process type, the `AZRebalance` process can cause your Auto Scaling group to grow up to ten percent larger than the maximum size. This is because Auto Scaling allows groups to temporarily grow larger than the maximum size during rebalancing activities. If Auto Scaling cannot terminate instances, your Auto Scaling group could remain up to ten percent larger than the maximum size until you resume the `terminate` process type.

The `HealthCheck` process type checks the health of the instances. Auto Scaling marks an instance as unhealthy if Amazon EC2 or Elastic Load Balancing informs Auto Scaling that the instance is unhealthy. The `HealthCheck` process can override the health status of an instance that you set with [SetInstanceHealth \(p. 32\)](#).

The `ReplaceUnhealthy` process type terminates instances that are marked as unhealthy and subsequently creates new instances to replace them. This process calls both of the primary process types--first `Terminate` and then `Launch`.



Important

The `HealthCheck` process type works in conjunction with the `ReplaceUnhealthy` process type to provide health check functionality. If you suspend either `Launch` or `Terminate`, the `ReplaceUnhealthy` process type will not function properly.

The `ScheduledActions` process type performs scheduled actions that you create with [PutScheduledUpdateGroupAction \(p. 29\)](#). Scheduled actions often involve launching new instances or terminating existing instances. If you suspend either `Launch` or `Terminate`, your scheduled actions might not function as expected.

Contents

Name	Description
<code>ProcessName</code>	The name of a process. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

PutScalingPolicyResult

Description

The `PolicyARNType` data type.

Contents

Name	Description
PolicyARN	A policy's Amazon Resource Name (ARN). Type: String Length constraints: Minimum value of 1. Maximum value of 1600.

ScalingPolicy

Description

The ScalingPolicy data type.

Contents

Name	Description
AdjustmentType	Specifies whether the ScalingAdjustment is an absolute number or a percentage of the current capacity. Valid values are <code>ChangeInCapacity</code> , <code>ExactCapacity</code> , and <code>PercentOfCapacity</code> . Type: String Length constraints: Minimum value of 1. Maximum value of 255.
Alarms	A list of CloudWatch Alarms related to the policy. Type: Alarm (p. 39) list
AutoScalingGroupName	The name of the Auto Scaling group associated with this scaling policy. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
Cooldown	The amount of time, in seconds, after a scaling activity completes before any further trigger-related scaling activities can start. Type: Integer
PolicyARN	The Amazon Resource Name (ARN) of the policy. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.
PolicyName	The name of the scaling policy. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
ScalingAdjustment	The number associated with the specified AdjustmentType. A positive value adds to the current capacity and a negative value removes from the current capacity. Type: Integer

ScheduledUpdateGroupAction

Description

This data type stores information about an scheduled update to an Auto Scaling group.

Contents

Name	Description
AutoScalingGroupName	The name of the Auto Scaling group to be updated. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
DesiredCapacity	The number of instances you prefer to maintain in your Auto Scaling group. The desired capacity must be no less than ??? and no greater than ???. Type: Integer
MaxSize	The maximum size of the Auto Scaling group. Type: Integer
MinSize	The minimum size of the Auto Scaling group. Type: Integer
ScheduledActionARN	The Amazon Resource Name (ARN) of this scheduled action. Type: String Length constraints: Minimum value of 1. Maximum value of 1600.
ScheduledActionName	The name of this scheduled action. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
Time	The time that the action is scheduled to occur. This value can be up to one month in the future. Type: DateTime

SuspendedProcess

Description

An Auto Scaling process that has been suspended. For more information, see [ProcessType](#) (p. 49).

Contents

Name	Description
ProcessName	The name of the suspended process. Type: String Length constraints: Minimum value of 1. Maximum value of 255.
SuspensionReason	The reason that the process was suspended. Type: String Length constraints: Minimum value of 1. Maximum value of 255.

TerminateInstanceInAutoScalingGroupResult

Description

The output for the [TerminateInstanceInAutoScalingGroup](#) (p. 34) action.

Contents

Name	Description
Activity	A Scaling Activity. Type: Activity (p. 38)

Common Query Parameters

This section lists the request parameters that all actions use. Any action-specific parameters are listed in the topic for the action.

Parameter Name	Description	Required
<i>Action</i>	The action to perform. Default: None Type: String	Yes
<i>AuthParams</i>	The parameters required to authenticate a query request. Contains: AWSAccessKeyId SignatureVersion Timestamp Signature Default: None	Conditional
<i>AWSAccessKeyId</i>	The Access Key ID corresponding to the AWS Secret Access Key you used to sign the request. Default: None Type: String	Yes
<i>Expires</i>	The date and time at which the request signature expires, in the format YYYY-MM-DDThh:mm:ssZ, as specified in the ISO 8601 standard. Condition: Requests must include either <i>Timestamp</i> or <i>Expires</i> , but not both. Default: None Type: String	Conditional

Parameter Name	Description	Required
<i>Signature</i>	The digital signature you created for the request. Refer to the service's developer documentation for information about how to generate the signature. Default: None Type: String	Yes
<i>SignatureMethod</i>	The hash algorithm you used to create the request signature. Default: None Valid Values: HmacSHA256 HmacSHA1. Type: String	Yes
<i>SignatureVersion</i>	The signature version you use to sign the request. Set this value to 2. Default: None Valid Value: 2 Type: String	Yes
<i>Timestamp</i>	The date and time the request was signed, in the format YYYY-MM-DDThh:mm:ssZ, as specified in the ISO 8601 standard. Condition: Requests must include either <i>Timestamp</i> or <i>Expires</i> , but not both. Default: None Type: String	Conditional
<i>Version</i>	The API version to use, in the format YYYY-MM-DD. Default: None Type: String	Yes

Common Errors

This section lists the common errors that all actions return. Any action-specific errors are listed in the topic for the action.

Error	Description	HTTP Status Code
IncompleteSignature	The request signature does not conform to AWS standards.	400
InternalFailure	The request processing has failed due to some unknown error, exception or failure.	500
InvalidAction	The action or operation requested is invalid.	400
InvalidClientTokenId	The X.509 certificate or AWS Access Key ID provided does not exist in our records.	403
InvalidParameterCombination	Parameters that must not be used together were used together.	400
InvalidParameterValue	A bad or out-of-range value was supplied for the input parameter.	400
InvalidQueryParameter	AWS query string is malformed, does not adhere to AWS standards.	400
MalformedQueryString	The query string is malformed.	404
MissingAction	The request is missing an action or operation parameter.	400
MissingAuthenticationToken	Request must contain either a valid (registered) AWS Access Key ID or X.509 certificate.	403
MissingParameter	An input parameter that is mandatory for processing the request is not supplied.	400

Error	Description	HTTP Status Code
OptInRequired	The AWS Access Key ID needs a subscription for the service.	403
RequestExpired	Request is past expires date or the request date (either with 15 minute padding), or the request date occurs more than 15 minutes in the future.	400
ServiceUnavailable	The request has failed due to a temporary failure of the server.	503
Throttling	Request was denied due to request throttling.	400