



API Reference

Amazon EC2 Auto Scaling



API Version 2011-01-01

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon EC2 Auto Scaling: API Reference

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

This documentation is a draft for private preview for regions in the AWS European Sovereign Cloud. Documentation content will continue to evolve. Published: December 30, 2025.

Table of Contents

Welcome	1
List of Actions by Function	3
Actions	6
AttachInstances	9
Request Parameters	9
Errors	10
Examples	10
See Also	10
AttachLoadBalancers	12
Request Parameters	12
Errors	13
Examples	13
See Also	14
AttachLoadBalancerTargetGroups	15
Request Parameters	15
Errors	16
Examples	17
See Also	17
AttachTrafficSources	18
Request Parameters	18
Errors	19
Examples	19
See Also	20
BatchDeleteScheduledAction	21
Request Parameters	21
Response Elements	21
Errors	22
See Also	22
BatchPutScheduledUpdateGroupAction	23
Request Parameters	23
Response Elements	23
Errors	23
See Also	24
CancelInstanceRefresh	26

Request Parameters	26
Response Elements	26
Errors	27
Examples	28
See Also	28
CompleteLifecycleAction	30
Request Parameters	30
Errors	31
Examples	32
See Also	32
CreateAutoScalingGroup	34
Request Parameters	34
Errors	45
Examples	46
See Also	46
CreateLaunchConfiguration	47
Request Parameters	47
Errors	54
Examples	55
See Also	55
CreateOrUpdateTags	56
Request Parameters	56
Errors	56
Examples	57
See Also	57
DeleteAutoScalingGroup	59
Request Parameters	59
Errors	60
Examples	60
See Also	61
DeleteLaunchConfiguration	62
Request Parameters	62
Errors	62
Examples	63
See Also	63
DeleteLifecycleHook	64

Request Parameters	64
Errors	64
Examples	65
See Also	65
DeleteNotificationConfiguration	66
Request Parameters	66
Errors	66
Examples	67
See Also	67
DeletePolicy	68
Request Parameters	68
Errors	68
Examples	69
See Also	69
DeleteScheduledAction	71
Request Parameters	71
Errors	71
Examples	72
See Also	72
DeleteTags	73
Request Parameters	73
Errors	73
Examples	73
See Also	74
DeleteWarmPool	75
Request Parameters	75
Errors	75
Examples	76
See Also	77
DescribeAccountLimits	78
Response Elements	78
Errors	78
Examples	79
See Also	79
DescribeAdjustmentTypes	81
Response Elements	81

Errors	81
Examples	81
See Also	82
DescribeAutoScalingGroups	84
Request Parameters	84
Response Elements	85
Errors	85
Examples	86
See Also	89
DescribeAutoScalingInstances	90
Request Parameters	90
Response Elements	90
Errors	91
Examples	91
See Also	92
DescribeAutoScalingNotificationTypes	94
Response Elements	94
Errors	94
Examples	94
See Also	95
DescribeInstanceRefreshes	96
Request Parameters	96
Response Elements	97
Errors	97
Examples	98
See Also	99
DescribeLaunchConfigurations	100
Request Parameters	100
Response Elements	101
Errors	101
Examples	102
See Also	103
DescribeLifecycleHooks	104
Request Parameters	104
Response Elements	104
Errors	105

Examples	105
See Also	106
DescribeLifecycleHookTypes	107
Response Elements	107
Errors	107
Examples	108
See Also	108
DescribeLoadBalancers	110
Request Parameters	110
Response Elements	111
Errors	112
Examples	112
See Also	113
DescribeLoadBalancerTargetGroups	114
Request Parameters	115
Response Elements	115
Errors	116
See Also	116
DescribeMetricCollectionTypes	118
Response Elements	118
Errors	118
Examples	118
See Also	121
DescribeNotificationConfigurations	122
Request Parameters	122
Response Elements	122
Errors	123
Examples	123
See Also	124
DescribePolicies	126
Request Parameters	126
Response Elements	127
Errors	127
Examples	128
See Also	130
DescribeScalingActivities	131

Request Parameters	131
Response Elements	132
Errors	133
Examples	133
See Also	134
DescribeScalingProcessTypes	136
Response Elements	136
Errors	136
Examples	136
See Also	138
DescribeScheduledActions	139
Request Parameters	139
Response Elements	140
Errors	141
See Also	141
DescribeTags	143
Request Parameters	143
Response Elements	144
Errors	144
Examples	145
See Also	145
DescribeTerminationPolicyTypes	147
Response Elements	147
Errors	147
Examples	147
See Also	148
DescribeTrafficSources	150
Request Parameters	150
Response Elements	151
Errors	151
Examples	152
See Also	153
DescribeWarmPool	154
Request Parameters	154
Response Elements	154
Errors	155

Examples	156
See Also	157
DetachInstances	158
Request Parameters	158
Response Elements	159
Errors	159
Examples	159
See Also	160
DetachLoadBalancers	162
Request Parameters	162
Errors	163
Examples	163
See Also	163
DetachLoadBalancerTargetGroups	165
Request Parameters	165
Errors	166
See Also	166
DetachTrafficSources	168
Request Parameters	168
Errors	168
Examples	169
See Also	169
DisableMetricsCollection	170
Request Parameters	170
Errors	171
Examples	171
See Also	172
EnableMetricsCollection	173
Request Parameters	173
Errors	175
Examples	175
See Also	175
EnterStandby	177
Request Parameters	177
Response Elements	178
Errors	178

Examples	178
See Also	179
ExecutePolicy	181
Request Parameters	181
Errors	182
See Also	183
ExitStandby	184
Request Parameters	184
Response Elements	184
Errors	185
Examples	185
See Also	186
GetPredictiveScalingForecast	187
Request Parameters	187
Response Elements	188
Errors	189
Examples	189
See Also	189
LaunchInstances	191
Request Parameters	191
Response Elements	193
Errors	194
See Also	194
PutLifecycleHook	196
Request Parameters	196
Errors	199
Examples	200
See Also	200
PutNotificationConfiguration	201
Request Parameters	201
Errors	202
Examples	202
See Also	203
PutScalingPolicy	204
Request Parameters	204
Response Elements	209

Errors	209
Examples	210
See Also	211
PutScheduledUpdateGroupAction	212
Request Parameters	212
Errors	215
Examples	215
See Also	216
PutWarmPool	218
Request Parameters	218
Errors	220
Examples	220
See Also	221
RecordLifecycleActionHeartbeat	222
Request Parameters	222
Errors	223
See Also	224
ResumeProcesses	225
Request Parameters	225
Errors	226
Examples	226
See Also	227
RollbackInstanceRefresh	228
Request Parameters	228
Response Elements	228
Errors	229
See Also	230
SetDesiredCapacity	231
Request Parameters	231
Errors	232
Examples	232
See Also	232
SetInstanceHealth	234
Request Parameters	234
Errors	235
Examples	235

See Also	235
SetInstanceProtection	237
Request Parameters	237
Errors	238
Examples	238
See Also	239
StartInstanceRefresh	240
Request Parameters	240
Response Elements	242
Errors	242
Examples	243
See Also	244
SuspendProcesses	245
Request Parameters	245
Errors	246
Examples	246
See Also	247
TerminateInstanceInAutoScalingGroup	248
Request Parameters	248
Response Elements	249
Errors	249
Examples	249
See Also	250
UpdateAutoScalingGroup	252
Request Parameters	253
Errors	261
Examples	261
See Also	262
Data Types	263
AcceleratorCountRequest	267
Contents	267
See Also	267
AcceleratorTotalMemoryMiBRequest	268
Contents	268
See Also	268
Activity	269

Contents	269
See Also	271
AdjustmentType	273
Contents	273
See Also	273
Alarm	274
Contents	274
See Also	274
AlarmSpecification	275
Contents	275
See Also	275
AutoScalingGroup	276
Contents	276
See Also	284
AutoScalingInstanceDetails	285
Contents	285
See Also	288
AvailabilityZoneDistribution	289
Contents	289
See Also	289
AvailabilityZoneImpairmentPolicy	290
Contents	290
See Also	290
BaselineEbsBandwidthMbpsRequest	291
Contents	291
See Also	291
BaselinePerformanceFactorsRequest	292
Contents	292
See Also	292
BlockDeviceMapping	293
Contents	293
See Also	294
CapacityForecast	295
Contents	295
See Also	295
CapacityReservationSpecification	296

Contents	296
See Also	296
CapacityReservationTarget	298
Contents	298
See Also	298
CpuPerformanceFactorRequest	299
Contents	299
See Also	299
CustomizedMetricSpecification	300
Contents	300
See Also	302
DesiredConfiguration	303
Contents	303
See Also	303
Ebs	305
Contents	305
See Also	307
EnabledMetric	308
Contents	308
See Also	309
FailedScheduledUpdateGroupActionRequest	310
Contents	310
See Also	310
Filter	312
Contents	312
See Also	313
Instance	314
Contents	314
See Also	317
InstanceCollection	318
Contents	318
See Also	319
InstanceLifecyclePolicy	320
Contents	320
See Also	320
InstanceMaintenancePolicy	321

Contents	321
See Also	322
InstanceMetadataOptions	323
Contents	323
See Also	324
InstanceMonitoring	325
Contents	325
See Also	325
InstanceRefresh	326
Contents	326
See Also	329
InstanceRefreshLivePoolProgress	331
Contents	331
See Also	331
InstanceRefreshProgressDetails	332
Contents	332
See Also	332
InstanceRefreshWarmPoolProgress	333
Contents	333
See Also	333
InstanceRequirements	334
Contents	334
See Also	344
InstanceReusePolicy	345
Contents	345
See Also	345
InstancesDistribution	346
Contents	346
See Also	350
LaunchConfiguration	351
Contents	351
See Also	356
LaunchInstancesError	357
Contents	357
See Also	358
LaunchTemplate	360

Contents	360
See Also	360
LaunchTemplateOverrides	361
Contents	361
See Also	364
LaunchTemplateSpecification	365
Contents	365
See Also	366
LifecycleHook	367
Contents	367
See Also	369
LifecycleHookSpecification	370
Contents	370
See Also	372
LoadBalancerState	373
Contents	373
See Also	373
LoadBalancerTargetGroupState	375
Contents	375
See Also	375
LoadForecast	377
Contents	377
See Also	377
MemoryGiBPerVCpuRequest	378
Contents	378
See Also	378
MemoryMiBRequest	379
Contents	379
See Also	379
Metric	380
Contents	380
See Also	380
MetricCollectionType	382
Contents	382
See Also	383
MetricDataQuery	384

Contents	384
See Also	386
MetricDimension	387
Contents	387
See Also	387
MetricGranularityType	388
Contents	388
See Also	388
MetricStat	389
Contents	389
See Also	390
MixedInstancesPolicy	391
Contents	391
See Also	391
NetworkBandwidthGbpsRequest	392
Contents	392
See Also	392
NetworkInterfaceCountRequest	394
Contents	394
See Also	394
NotificationConfiguration	395
Contents	395
See Also	396
PerformanceFactorReferenceRequest	397
Contents	397
See Also	398
PredefinedMetricSpecification	399
Contents	399
See Also	400
PredictiveScalingConfiguration	401
Contents	401
See Also	403
PredictiveScalingCustomizedCapacityMetric	404
Contents	404
See Also	404
PredictiveScalingCustomizedLoadMetric	405

Contents	405
See Also	405
PredictiveScalingCustomizedScalingMetric	406
Contents	406
See Also	406
PredictiveScalingMetricSpecification	407
Contents	408
See Also	409
PredictiveScalingPredefinedLoadMetric	410
Contents	410
See Also	411
PredictiveScalingPredefinedMetricPair	412
Contents	412
See Also	413
PredictiveScalingPredefinedScalingMetric	414
Contents	414
See Also	415
ProcessType	416
Contents	416
See Also	416
RefreshPreferences	418
Contents	418
See Also	422
RetentionTriggers	423
Contents	423
See Also	423
RollbackDetails	424
Contents	424
See Also	425
ScalingPolicy	426
Contents	426
See Also	429
ScheduledUpdateGroupAction	430
Contents	430
See Also	432
ScheduledUpdateGroupActionRequest	433

Contents	433
See Also	435
StepAdjustment	436
Contents	436
See Also	437
SuspendedProcess	438
Contents	438
See Also	438
Tag	439
Contents	439
See Also	440
TagDescription	441
Contents	441
See Also	442
TargetTrackingConfiguration	443
Contents	443
See Also	444
TargetTrackingMetricDataQuery	445
Contents	445
See Also	447
TargetTrackingMetricStat	448
Contents	448
See Also	449
TotalLocalStorageGBRequest	450
Contents	450
See Also	450
TrafficSourceIdentifier	451
Contents	451
See Also	452
TrafficSourceState	453
Contents	453
See Also	454
VCpuCountRequest	456
Contents	456
See Also	456
WarmPoolConfiguration	457

Contents	457
See Also	458
Common Parameters	459
Common Errors	462
Permissions	464

Welcome

This is the *Amazon EC2 Auto Scaling API Reference*. Amazon EC2 Auto Scaling is designed to automatically launch and terminate EC2 instances based on user-defined scaling policies, scheduled actions, and health checks. This guide describes the Amazon EC2 Auto Scaling operations that you can call programmatically. For more information about Amazon EC2 Auto Scaling, see the [Amazon EC2 Auto Scaling User Guide](#).

Note

To configure Amazon EC2 instances that are launched by your Auto Scaling group, you can specify a launch template or a launch configuration. We recommend that you use Amazon EC2 launch templates. Launch configurations are included as part of the same API as Auto Scaling groups, but launch templates offer more features. For more information, see [Launch templates](#) in the *Amazon EC2 Auto Scaling User Guide*.

To get started, complete the following tasks:

1. Create a launch template using [CreateLaunchTemplate](#).
2. Create an Auto Scaling group using [CreateAutoScalingGroup](#).

To make programmatic API calls, you can also use one of the AWS SDKs, which consist of libraries and sample code for various programming languages and platforms. The SDKs provide a convenient way to create programmatic access to Amazon EC2 Auto Scaling. For example, the SDKs take care of cryptographically signing requests, managing errors, and retrying requests automatically. For information about the AWS SDKs, including how to download and install them, see the [AWS SDKs](#).

API request rate

Amazon EC2 Auto Scaling uses the token bucket algorithm to implement API throttling. With this algorithm, your account has a bucket that holds a specific number of tokens. The number of tokens in the bucket represents your throttling limit at any given second. Amazon EC2 Auto Scaling throttles API requests based on a shared API bucket. For example, calls to the [DescribeAutoScalingGroups](#) and [DescribeScheduledActions](#) API operations use tokens from the same bucket. Throttling means that Amazon EC2 Auto Scaling rejects a request because the

request exceeds the service's limit for the number of requests per second. When a request is throttled, Amazon EC2 Auto Scaling returns a `RateExceeded` error.

The [DescribeAutoScalingGroups](#) API operation might be throttled when retrieving details for an Auto Scaling group that contains many instances. By default, this operation returns details for all instances in the group. To help prevent throttling, you can set the `IncludeInstances` parameter to `false` to exclude instance details from the response.

Additional information

- To specify a custom Amazon Machine Image (AMI) in your launch template, you must first create your AMI from a customized instance. For information on creating a custom AMI, see [Create an AMI](#) in the *Amazon EC2 User Guide*.
- For information about the permissions that are required for Amazon EC2 Auto Scaling, see [Granting IAM permissions for Amazon EC2 Auto Scaling actions](#).
- For information about how to provide network connectivity for instances in an Auto Scaling group, see [Provide network connectivity for your Auto Scaling instances using Amazon VPC](#) in the *Amazon EC2 Auto Scaling User Guide*.

This document was last published on December 30, 2025.

List of Actions by Function

Account Limits

- [DescribeAccountLimits](#)

Auto Scaling Groups

- [AttachLoadBalancers](#)
- [AttachLoadBalancerTargetGroups](#)
- [CreateAutoScalingGroup](#)
- [DeleteAutoScalingGroup](#)
- [DescribeAutoScalingGroups](#)
- [DescribeLoadBalancers](#)
- [DescribeLoadBalancerTargetGroups](#)
- [DetachLoadBalancerTargetGroups](#)
- [DetachLoadBalancers](#)
- [UpdateAutoScalingGroup](#)

Auto Scaling Instances

- [AttachInstances](#)
- [DescribeAutoScalingInstances](#)
- [DetachInstances](#)
- [SetInstanceHealth](#)
- [SetInstanceProtection](#)
- [TerminateInstanceInAutoScalingGroup](#)

Instance Refreshes

- [CancelInstanceRefresh](#)
- [DescribeInstanceRefreshes](#)
- [RollbackInstanceRefresh](#)

- [StartInstanceRefresh](#)

Launch Configurations

- [CreateLaunchConfiguration](#)
- [DeleteLaunchConfiguration](#)
- [DescribeLaunchConfigurations](#)

Lifecycle Hooks

- [CompleteLifecycleAction](#)
- [DeleteLifecycleHook](#)
- [DescribeLifecycleHooks](#)
- [DescribeLifecycleHookTypes](#)
- [PutLifecycleHook](#)
- [RecordLifecycleActionHeartbeat](#)

Monitoring

- [DeleteNotificationConfiguration](#)
- [DescribeAutoScalingNotificationTypes](#)
- [DescribeMetricCollectionTypes](#)
- [DescribeNotificationConfigurations](#)
- [DisableMetricsCollection](#)
- [EnableMetricsCollection](#)
- [PutNotificationConfiguration](#)

Scaling

- [DeletePolicy](#)
- [DescribeAdjustmentTypes](#)
- [DescribePolicies](#)
- [DescribeScalingActivities](#)

- [DescribeScalingProcessTypes](#)
- [DescribeTerminationPolicyTypes](#)
- [ExecutePolicy](#)
- [PutScalingPolicy](#)
- [ResumeProcesses](#)
- [SetDesiredCapacity](#)
- [SuspendProcesses](#)

Scheduled Scaling

- [BatchDeleteScheduledAction](#)
- [BatchPutScheduledUpdateGroupAction](#)
- [DeleteScheduledAction](#)
- [DescribeScheduledActions](#)
- [PutScheduledUpdateGroupAction](#)

Standby State

- [EnterStandby](#)
- [ExitStandby](#)

Tags

- [CreateOrUpdateTags](#)
- [DeleteTags](#)
- [DescribeTags](#)

Warm Pools

- [PutWarmPool](#)
- [DescribeWarmPool](#)
- [DeleteWarmPool](#)

Actions

The following actions are supported:

- [AttachInstances](#)
- [AttachLoadBalancers](#)
- [AttachLoadBalancerTargetGroups](#)
- [AttachTrafficSources](#)
- [BatchDeleteScheduledAction](#)
- [BatchPutScheduledUpdateGroupAction](#)
- [CancelInstanceRefresh](#)
- [CompleteLifecycleAction](#)
- [CreateAutoScalingGroup](#)
- [CreateLaunchConfiguration](#)
- [CreateOrUpdateTags](#)
- [DeleteAutoScalingGroup](#)
- [DeleteLaunchConfiguration](#)
- [DeleteLifecycleHook](#)
- [DeleteNotificationConfiguration](#)
- [DeletePolicy](#)
- [DeleteScheduledAction](#)
- [DeleteTags](#)
- [DeleteWarmPool](#)
- [DescribeAccountLimits](#)
- [DescribeAdjustmentTypes](#)
- [DescribeAutoScalingGroups](#)
- [DescribeAutoScalingInstances](#)
- [DescribeAutoScalingNotificationTypes](#)
- [DescribeInstanceRefreshes](#)
- [DescribeLaunchConfigurations](#)
- [DescribeLifecycleHooks](#)

- [DescribeLifecycleHookTypes](#)
- [DescribeLoadBalancers](#)
- [DescribeLoadBalancerTargetGroups](#)
- [DescribeMetricCollectionTypes](#)
- [DescribeNotificationConfigurations](#)
- [DescribePolicies](#)
- [DescribeScalingActivities](#)
- [DescribeScalingProcessTypes](#)
- [DescribeScheduledActions](#)
- [DescribeTags](#)
- [DescribeTerminationPolicyTypes](#)
- [DescribeTrafficSources](#)
- [DescribeWarmPool](#)
- [DetachInstances](#)
- [DetachLoadBalancers](#)
- [DetachLoadBalancerTargetGroups](#)
- [DetachTrafficSources](#)
- [DisableMetricsCollection](#)
- [EnableMetricsCollection](#)
- [EnterStandby](#)
- [ExecutePolicy](#)
- [ExitStandby](#)
- [GetPredictiveScalingForecast](#)
- [LaunchInstances](#)
- [PutLifecycleHook](#)
- [PutNotificationConfiguration](#)
- [PutScalingPolicy](#)
- [PutScheduledUpdateGroupAction](#)
- [PutWarmPool](#)
- [RecordLifecycleActionHeartbeat](#)

- [ResumeProcesses](#)
- [RollbackInstanceRefresh](#)
- [SetDesiredCapacity](#)
- [SetInstanceHealth](#)
- [SetInstanceProtection](#)
- [StartInstanceRefresh](#)
- [SuspendProcesses](#)
- [TerminateInstanceInAutoScalingGroup](#)
- [UpdateAutoScalingGroup](#)

AttachInstances

Attaches one or more EC2 instances to the specified Auto Scaling group.

When you attach instances, Amazon EC2 Auto Scaling increases the desired capacity of the group by the number of instances being attached. If the number of instances being attached plus the desired capacity of the group exceeds the maximum size of the group, the operation fails.

If there is a Classic Load Balancer attached to your Auto Scaling group, the instances are also registered with the load balancer. If there are target groups attached to your Auto Scaling group, the instances are also registered with the target groups.

For more information, see [Detach or attach instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceIds.member.N

The IDs of the instances. You can specify up to 20 instances.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `AttachInstances`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=AttachInstances
&AutoScalingGroupName=my-asg
&InstanceIds.member.1=i-1234567890abcdef0
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)

- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

AttachLoadBalancers

Note

This API operation is superseded by https://docs.aws.eu/autoscaling/ec2/APIReference/API_AttachTrafficSources.html, which can attach multiple traffic sources types. We recommend using `AttachTrafficSources` to simplify how you manage traffic sources. However, we continue to support `AttachLoadBalancers`. You can use both the original `AttachLoadBalancers` API operation and `AttachTrafficSources` on the same Auto Scaling group.

Attaches one or more Classic Load Balancers to the specified Auto Scaling group. Amazon EC2 Auto Scaling registers the running instances with these Classic Load Balancers.

To describe the load balancers for an Auto Scaling group, call the [DescribeLoadBalancers](#) API. To detach a load balancer from the Auto Scaling group, call the [DetachLoadBalancers](#) API.

This operation is additive and does not detach existing Classic Load Balancers or target groups from the Auto Scaling group.

For more information, see [Use Elastic Load Balancing to distribute traffic across the instances in your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

LoadBalancerNames.member.N

The names of the load balancers. You can specify up to 10 load balancers.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InstanceRefreshInProgress

The request failed because an active instance refresh already exists for the specified Auto Scaling group.

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `AttachLoadBalancers`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=AttachLoadBalancers
&AutoScalingGroupName=my-asg
&LoadBalancerNames.member.1=my-lb
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

AttachLoadBalancerTargetGroups

Note

This API operation is superseded by [AttachTrafficSources](#), which can attach multiple traffic sources types. We recommend using `AttachTrafficSources` to simplify how you manage traffic sources. However, we continue to support `AttachLoadBalancerTargetGroups`. You can use both the original `AttachLoadBalancerTargetGroups` API operation and `AttachTrafficSources` on the same Auto Scaling group.

Attaches one or more target groups to the specified Auto Scaling group.

This operation is used with the following load balancer types:

- Application Load Balancer - Operates at the application layer (layer 7) and supports HTTP and HTTPS.
- Network Load Balancer - Operates at the transport layer (layer 4) and supports TCP, TLS, and UDP.
- Gateway Load Balancer - Operates at the network layer (layer 3).

To describe the target groups for an Auto Scaling group, call the [DescribeLoadBalancerTargetGroups](#) API. To detach the target group from the Auto Scaling group, call the [DetachLoadBalancerTargetGroups](#) API.

This operation is additive and does not detach existing target groups or Classic Load Balancers from the Auto Scaling group.

For more information, see [Use Elastic Load Balancing to distribute traffic across the instances in your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

TargetGroupARNs.member.N

The Amazon Resource Names (ARNs) of the target groups. You can specify up to 10 target groups. To get the ARN of a target group, use the Elastic Load Balancing [DescribeTargetGroups](#) API operation.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InstanceRefreshInProgress

The request failed because an active instance refresh already exists for the specified Auto Scaling group.

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `AttachLoadBalancerTargetGroups`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=AttachLoadBalancerTargetGroups
&AutoScalingGroupName=my-asg
&TargetGroupARNs.member.1=arn:aws:elasticloadbalancing:us-
east-1:123456789012:targetgroup/my-targets/1234567890123456
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

AttachTrafficSources

Attaches one or more traffic sources to the specified Auto Scaling group.

You can use any of the following as traffic sources for an Auto Scaling group:

- Application Load Balancer
- Classic Load Balancer
- Gateway Load Balancer
- Network Load Balancer
- VPC Lattice

This operation is additive and does not detach existing traffic sources from the Auto Scaling group.

After the operation completes, use the [DescribeTrafficSources](#) API to return details about the state of the attachments between traffic sources and your Auto Scaling group. To detach a traffic source from the Auto Scaling group, call the [DetachTrafficSources](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

SkipZonalShiftValidation

If you enable zonal shift with cross-zone disabled load balancers, capacity could become imbalanced across Availability Zones. To skip the validation, specify `true`. For more information, see [Auto Scaling group zonal shift](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

TrafficSources.member.N

The unique identifiers of one or more traffic sources. You can specify up to 10 traffic sources.

Type: Array of [TrafficSourceIdentifier](#) objects

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InstanceRefreshInProgress

The request failed because an active instance refresh already exists for the specified Auto Scaling group.

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example attaches two VPC Lattice target groups, as specified by their ARNs, to the Auto Scaling group named my-asg.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=AttachTrafficSources
&AutoScalingGroupName=my-asg
&TrafficSources.member.1.Identifier=arn:aws:vpc-lattice:us-
west-2:123456789012:targetgroup/tg-0e2f2665eEXAMPLE
&TrafficSources.member.2.Identifier=arn:aws:vpc-lattice:us-
west-2:123456789012:targetgroup/tg-8360a9e72EXAMPLE
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchDeleteScheduledAction

Deletes one or more scheduled actions for the specified Auto Scaling group.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ScheduledActionNames.member.N

The names of the scheduled actions to delete. The maximum number allowed is 50.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Response Elements

The following element is returned by the service.

FailedScheduledActions.member.N

The names of the scheduled actions that could not be deleted, including an error message.

Type: Array of [FailedScheduledUpdateGroupActionRequest](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

BatchPutScheduledUpdateGroupAction

Creates or updates one or more scheduled scaling actions for an Auto Scaling group.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\u005C\u005C]`*

Required: Yes

ScheduledUpdateGroupActions.member.N

One or more scheduled actions. The maximum number allowed is 50.

Type: Array of [ScheduledUpdateGroupActionRequest](#) objects

Required: Yes

Response Elements

The following element is returned by the service.

FailedScheduledUpdateGroupActions.member.N

The names of the scheduled actions that could not be created or updated, including an error message.

Type: Array of [FailedScheduledUpdateGroupActionRequest](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExists

You already have an Auto Scaling group or launch configuration with this name.

message

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CancelInstanceRefresh

Cancels an instance refresh or rollback that is in progress. If an instance refresh or rollback is not in progress, an `ActiveInstanceRefreshNotFound` error occurs.

This operation is part of the [instance refresh feature](#) in Amazon EC2 Auto Scaling, which helps you update instances in your Auto Scaling group after you make configuration changes.

When you cancel an instance refresh, this does not roll back any changes that it made. Use the [RollbackInstanceRefresh](#) API to roll back instead.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

WaitForTransitioningInstances

When cancelling an instance refresh, this indicates whether to wait for in-flight launches and terminations to complete. The default is true.

When set to false, Amazon EC2 Auto Scaling cancels the instance refresh without waiting for any pending launches or terminations to complete.

Type: Boolean

Required: No

Response Elements

The following element is returned by the service.

InstanceRefreshId

The instance refresh ID associated with the request. This is the unique ID assigned to the instance refresh when it was started.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ActiveInstanceRefreshNotFound

The request failed because an active instance refresh or rollback for the specified Auto Scaling group was not found.

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `CancelInstanceRefresh`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=CancelInstanceRefresh
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<CancelInstanceRefreshResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <CancelInstanceRefreshResult>
    <InstanceRefreshId>08b91cf7-8fa6-48af-b6a6-d227f40f1b9b</InstanceRefreshId>
  </CancelInstanceRefreshResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</CancelInstanceRefreshResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LifecycleActionResult

The action for the group to take. You can specify either CONTINUE or ABANDON.

Type: String

Required: Yes

LifecycleActionToken

A universally unique identifier (UUID) that identifies a specific lifecycle action associated with an instance. Amazon EC2 Auto Scaling sends this token to the notification target you specified when you created the lifecycle hook.

Type: String

Length Constraints: Fixed length of 36.

Required: No

LifecycleHookName

The name of the lifecycle hook.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z0-9\-\_\/]+`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of CompleteLifecycleAction.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=CompleteLifecycleAction
&AutoScalingGroupName=my-asg
&LifecycleHookName=my-lifecycle-hook
&LifecycleActionResult=CONTINUE
&LifecycleActionToken=bcd2f1b8-9a78-44d3-8a7a-4dd07EXAMPLE
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateAutoScalingGroup

We strongly recommend using a launch template when calling this operation to ensure full functionality for Amazon EC2 Auto Scaling and Amazon EC2.

Creates an Auto Scaling group with the specified name and attributes.

If you exceed your maximum limit of Auto Scaling groups, the call fails. To query this limit, call the [DescribeAccountLimits](#) API. For information about updating this limit, see [Quotas for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

If you're new to Amazon EC2 Auto Scaling, see the introductory tutorials in [Get started with Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Every Auto Scaling group has three size properties (`DesiredCapacity`, `MaxSize`, and `MinSize`). Usually, you set these sizes based on a specific number of instances. However, if you configure a mixed instances policy that defines weights for the instance types, you must specify these sizes with the same units that you use for weighting instances.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group. This name must be unique per Region per account.

The name can contain any ASCII character 33 to 126 including most punctuation characters, digits, and upper and lowercased letters.

Note

You cannot use a colon (:) in the name.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

AvailabilityZoneDistribution

The instance capacity distribution across Availability Zones.

Type: [AvailabilityZoneDistribution](#) object

Required: No

AvailabilityZoneImpairmentPolicy

The policy for Availability Zone impairment.

Type: [AvailabilityZoneImpairmentPolicy](#) object

Required: No

AvailabilityZones.member.N

A list of Availability Zones where instances in the Auto Scaling group can be created. Used for launching into the default VPC subnet in each Availability Zone when not using the `VPCZoneIdentifier` property, or for attaching a network interface when an existing network interface ID is specified in a launch template.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

CapacityRebalance

Indicates whether Capacity Rebalancing is enabled. Otherwise, Capacity Rebalancing is disabled. When you turn on Capacity Rebalancing, Amazon EC2 Auto Scaling attempts to launch a Spot Instance whenever Amazon EC2 notifies that a Spot Instance is at an elevated risk of interruption. After launching a new instance, it then terminates an old instance. For more information, see [Use Capacity Rebalancing to handle Amazon EC2 Spot Interruptions](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

CapacityReservationSpecification

The capacity reservation specification for the Auto Scaling group.

Type: [CapacityReservationSpecification](#) object

Required: No

Context

Reserved.

Type: String

Required: No

DefaultCooldown

Only needed if you use simple scaling policies.

The amount of time, in seconds, between one scaling activity ending and another one starting due to simple scaling policies. For more information, see [Scaling cooldowns for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Default: 300 seconds

Type: Integer

Required: No

DefaultInstanceWarmup

The amount of time, in seconds, until a new instance is considered to have finished initializing and resource consumption to become stable after it enters the InService state.

During an instance refresh, Amazon EC2 Auto Scaling waits for the warm-up period after it replaces an instance before it moves on to replacing the next instance. Amazon EC2 Auto Scaling also waits for the warm-up period before aggregating the metrics for new instances with existing instances in the Amazon CloudWatch metrics that are used for scaling, resulting in more reliable usage data. For more information, see [Set the default instance warmup for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

⚠ Important

To manage various warm-up settings at the group level, we recommend that you set the default instance warmup, *even if it is set to 0 seconds*. To remove a value that you previously set, include the property but specify -1 for the value. However, we strongly recommend keeping the default instance warmup enabled by specifying a value of 0 or other nominal value.

Default: None

Type: Integer

Required: No

DesiredCapacity

The desired capacity is the initial capacity of the Auto Scaling group at the time of its creation and the capacity it attempts to maintain. It can scale beyond this capacity if you configure auto scaling. This number must be greater than or equal to the minimum size of the group and less than or equal to the maximum size of the group. If you do not specify a desired capacity, the default is the minimum size of the group.

Type: Integer

Required: No

DesiredCapacityType

The unit of measurement for the value specified for desired capacity. Amazon EC2 Auto Scaling supports `DesiredCapacityType` for attribute-based instance type selection only. For more information, see [Create a mixed instances group using attribute-based instance type selection](#) in the *Amazon EC2 Auto Scaling User Guide*.

By default, Amazon EC2 Auto Scaling specifies `units`, which translates into number of instances.

Valid values: `units` | `vcpu` | `memory-mib`

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

HealthCheckGracePeriod

The amount of time, in seconds, that Amazon EC2 Auto Scaling waits before checking the health status of an EC2 instance that has come into service and marking it unhealthy due to a failed health check. This is useful if your instances do not immediately pass their health checks after they enter the `InService` state. For more information, see [Set the health check grace period for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Default: 0 seconds

Type: Integer

Required: No

HealthCheckType

A comma-separated value string of one or more health check types.

The valid values are `EC2`, `EBS`, `ELB`, and `VPC_LATTICE`. `EC2` is the default health check and cannot be disabled. For more information, see [Health checks for instances in an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Only specify `EC2` if you must clear a value that was previously set.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceId

The ID of the instance used to base the launch configuration on. If specified, Amazon EC2 Auto Scaling uses the configuration values from the specified instance to create a new launch configuration. To get the instance ID, use the Amazon EC2 [DescribeInstances](#) API operation. For more information, see [Create an Auto Scaling group using parameters from an existing instance](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceLifecyclePolicy

The instance lifecycle policy for the Auto Scaling group. This policy controls instance behavior when an instance transitions through its lifecycle states. Configure retention triggers to specify when instances should move to a Retained state instead of automatic termination.

For more information, see [Control instance retention with instance lifecycle policies](#) in the *Amazon EC2 Auto Scaling User Guide*.

Note

Instances in a Retained state will continue to incur standard EC2 charges until terminated.

Type: [InstanceLifecyclePolicy](#) object

Required: No

InstanceMaintenancePolicy

An instance maintenance policy. For more information, see [Set instance maintenance policy](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [InstanceMaintenancePolicy](#) object

Required: No

LaunchConfigurationName

The name of the launch configuration to use to launch instances.

Conditional: You must specify either a launch template (LaunchTemplate or MixedInstancesPolicy) or a launch configuration (LaunchConfigurationName or InstanceId).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchTemplate

Information used to specify the launch template and version to use to launch instances.

Conditional: You must specify either a launch template (LaunchTemplate or MixedInstancesPolicy) or a launch configuration (LaunchConfigurationName or InstanceId).

Note

The launch template that is specified must be configured for use with an Auto Scaling group. For more information, see [Create a launch template for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [LaunchTemplateSpecification](#) object

Required: No

LifecycleHookSpecificationList.member.N

One or more lifecycle hooks to add to the Auto Scaling group before instances are launched.

Type: Array of [LifecycleHookSpecification](#) objects

Required: No

LoadBalancerNames.member.N

A list of Classic Load Balancers associated with this Auto Scaling group. For Application Load Balancers, Network Load Balancers, and Gateway Load Balancers, specify the TargetGroupARNs property instead.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxInstanceLifetime

The maximum amount of time, in seconds, that an instance can be in service. The default is null. If specified, the value must be either 0 or a number equal to or greater than 86,400 seconds (1 day). For more information, see [Replace Auto Scaling instances based on maximum instance lifetime](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Integer

Required: No

MaxSize

The maximum size of the group.

Note

With a mixed instances policy that uses instance weighting, Amazon EC2 Auto Scaling may need to go above `MaxSize` to meet your capacity requirements. In this event, Amazon EC2 Auto Scaling will never go above `MaxSize` by more than your largest instance weight (weights that define how many units each instance contributes to the desired capacity of the group).

Type: Integer

Required: Yes

MinSize

The minimum size of the group.

Type: Integer

Required: Yes

MixedInstancesPolicy

The mixed instances policy.

The policy includes properties that not only define the distribution of On-Demand Instances and Spot Instances and how the Auto Scaling group allocates instance types to fulfill On-Demand and Spot capacities, but also the properties that specify the instance configuration information—the launch template and instance types. The policy can also include a weight for each instance type and different launch templates for certain instance types. For more information, see [Auto Scaling groups with multiple instance types and purchase options](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [MixedInstancesPolicy](#) object

Required: No

NewInstancesProtectedFromScaleIn

Indicates whether newly launched instances are protected from termination by Amazon EC2 Auto Scaling when scaling in. For more information about preventing instances from terminating on scale in, see [Use instance scale-in protection](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

PlacementGroup

The name of the placement group into which to launch your instances. For more information, see [Placement groups](#) in the *Amazon EC2 User Guide*.

Note

A *cluster* placement group is a logical grouping of instances within a single Availability Zone. You cannot specify multiple Availability Zones and a cluster placement group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

ServiceLinkedRoleARN

The Amazon Resource Name (ARN) of the service-linked role that the Auto Scaling group uses to call other AWS service on your behalf. By default, Amazon EC2 Auto Scaling uses a service-linked role named `AWSServiceRoleForAutoScaling`, which it creates if it does not exist. For more information, see [Service-linked roles](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

SkipZonalShiftValidation

If you enable zonal shift with cross-zone disabled load balancers, capacity could become imbalanced across Availability Zones. To skip the validation, specify `true`. For more information, see [Auto Scaling group zonal shift](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

Tags.member.N

One or more tags. You can tag your Auto Scaling group and propagate the tags to the Amazon EC2 instances it launches. Tags are not propagated to Amazon EBS volumes. To add tags to Amazon EBS volumes, specify the tags in a launch template but use caution. If the launch template specifies an instance tag with a key that is also specified for the Auto Scaling group, Amazon EC2 Auto Scaling overrides the value of that instance tag with the value specified by the Auto Scaling group. For more information, see [Tag Auto Scaling groups and instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Array of [Tag](#) objects

Required: No

TargetGroupARNs.member.N

The Amazon Resource Names (ARN) of the Elastic Load Balancing target groups to associate with the Auto Scaling group. Instances are registered as targets with the target groups. The target groups receive incoming traffic and route requests to one or more registered targets. For

more information, see [Use Elastic Load Balancing to distribute traffic across the instances in your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

TerminationPolicies.member.N

A policy or a list of policies that are used to select the instance to terminate. These policies are executed in the order that you list them. For more information, see [Configure termination policies for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Valid values: Default | AllocationStrategy | ClosestToNextInstanceHour | NewestInstance | OldestInstance | OldestLaunchConfiguration | OldestLaunchTemplate | arn:aws:lambda:region:account-id:function:my-function:my-alias

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

TrafficSources.member.N

The list of traffic sources to attach to this Auto Scaling group. You can use any of the following as traffic sources for an Auto Scaling group: Classic Load Balancer, Application Load Balancer, Gateway Load Balancer, Network Load Balancer, and VPC Lattice.

Type: Array of [TrafficSourceIdentifier](#) objects

Required: No

VPCZoneIdentifier

A comma-separated list of subnet IDs for a virtual private cloud (VPC) where instances in the Auto Scaling group can be created. If you specify `VPCZoneIdentifier` with `AvailabilityZones`, the subnets that you specify must reside in those Availability Zones.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 5000.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExists

You already have an Auto Scaling group or launch configuration with this name.

message

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `CreateAutoScalingGroup`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=CreateAutoScalingGroup
&AutoScalingGroupName=my-asg
&VPCZoneIdentifier=subnet-057fa0918fEXAMPLE%2Csubnet-610acd08EXAMPLE
&MinSize=2
&MaxSize=10
&DesiredCapacity=2
&LoadBalancerNames.member.1=my-loadbalancer
&HealthCheckType=ELB
&HealthCheckGracePeriod=120
&LaunchConfigurationName=my-lc
&MaxInstanceLifetime=2592000
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateLaunchConfiguration

Creates a launch configuration.

If you exceed your maximum limit of launch configurations, the call fails. To query this limit, call the [DescribeAccountLimits](#) API. For information about updating this limit, see [Quotas for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

For more information, see [Launch configurations](#) in the *Amazon EC2 Auto Scaling User Guide*.

Note

Amazon EC2 Auto Scaling configures instances launched as part of an Auto Scaling group using either a launch template or a launch configuration. We strongly recommend that you do not use launch configurations. They do not provide full functionality for Amazon EC2 Auto Scaling or Amazon EC2. For information about using launch templates, see [Launch templates](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AssociatePublicIpAddress

Specifies whether to assign a public IPv4 address to the group's instances. If the instance is launched into a default subnet, the default is to assign a public IPv4 address, unless you disabled the option to assign a public IPv4 address on the subnet. If the instance is launched into a nondefault subnet, the default is not to assign a public IPv4 address, unless you enabled the option to assign a public IPv4 address on the subnet.

If you specify `true`, each instance in the Auto Scaling group receives a unique public IPv4 address. For more information, see [Provide network connectivity for your Auto Scaling instances using Amazon VPC](#) in the *Amazon EC2 Auto Scaling User Guide*.

If you specify this property, you must specify at least one subnet for `VPCZoneIdentifier` when you create your group.

Type: Boolean

The default value is `false`.

Type: Boolean

Required: No

IamInstanceProfile

The name or the Amazon Resource Name (ARN) of the instance profile associated with the IAM role for the instance. The instance profile contains the IAM role. For more information, see [IAM role for applications that run on Amazon EC2 instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ImageId

The ID of the Amazon Machine Image (AMI) that was assigned during registration. For more information, see [Find a Linux AMI](#) in the *Amazon EC2 User Guide*.

If you specify `InstanceId`, an `ImageId` is not required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceId

The ID of the instance to use to create the launch configuration. The new launch configuration derives attributes from the instance, except for the block device mapping.

To create a launch configuration with a block device mapping or override any other instance attributes, specify them as part of the same request.

For more information, see [Create a launch configuration](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceMonitoring

Controls whether instances in this group are launched with detailed (`true`) or basic (`false`) monitoring.

The default value is `true` (enabled).

Important

When detailed monitoring is enabled, Amazon CloudWatch generates metrics every minute and your account is charged a fee. When you disable detailed monitoring, CloudWatch generates metrics every 5 minutes. For more information, see [Configure monitoring for Auto Scaling instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [InstanceMonitoring](#) object

Required: No

InstanceType

Specifies the instance type of the EC2 instance. For information about available instance types, see [Available instance types](#) in the *Amazon EC2 User Guide*.

If you specify `InstanceId`, an `InstanceType` is not required.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

KernelId

The ID of the kernel associated with the AMI.

Note

We recommend that you use PV-GRUB instead of kernels and RAM disks. For more information, see [User provided kernels](#) in the *Amazon EC2 User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

KeyName

The name of the key pair. For more information, see [Amazon EC2 key pairs and Amazon EC2 instances](#) in the *Amazon EC2 User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchConfigurationName

The name of the launch configuration. This name must be unique per Region per account.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

MetadataOptions

The metadata options for the instances. For more information, see [Configure the instance metadata options](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [InstanceMetadataOptions](#) object

Required: No

PlacementTenancy

The tenancy of the instance, either `default` or `dedicated`. An instance with dedicated tenancy runs on isolated, single-tenant hardware and can only be launched into a VPC. To launch dedicated instances into a shared tenancy VPC (a VPC with the instance placement tenancy attribute set to `default`), you must set the value of this property to `dedicated`.

If you specify `PlacementTenancy`, you must specify at least one subnet for `VPCZoneIdentifier` when you create your group.

Valid values: `default` | `dedicated`

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

RamdiskId

The ID of the RAM disk to select.

Note

We recommend that you use PV-GRUB instead of kernels and RAM disks. For more information, see [User provided kernels](#) in the *Amazon EC2 User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SecurityGroups.member.N

A list that contains the security group IDs to assign to the instances in the Auto Scaling group. For more information, see [Control traffic to your AWS resources using security groups](#) in the *Amazon Virtual Private Cloud User Guide*.

Type: Array of strings

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SpotPrice

The maximum hourly price to be paid for any Spot Instance launched to fulfill the request. Spot Instances are launched when the price you specify exceeds the current Spot price. For more information, see [Request Spot Instances for fault-tolerant and flexible applications](#) in the *Amazon EC2 Auto Scaling User Guide*.

Valid Range: Minimum value of 0.001

Note

When you change your maximum price by creating a new launch configuration, running instances will continue to run as long as the maximum price for those running instances is higher than the current Spot price.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

UserData

The user data to make available to the launched EC2 instances. For more information, see [Instance metadata and user data](#) (Linux) and [Instance metadata and user data](#) (Windows). If you

are using a command line tool, base64-encoding is performed for you, and you can load the text from a file. Otherwise, you must provide base64-encoded text. User data is limited to 16 KB.

Type: String

Length Constraints: Maximum length of 21847.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExists

You already have an Auto Scaling group or launch configuration with this name.

message

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `CreateLaunchConfiguration`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=CreateLaunchConfiguration
&LaunchConfigurationName=my-lc
&ImageId=ami-0abcdef1234567890
&InstanceType=t2.micro
&SecurityGroups.member.1=sg-1a2b3c4d
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

CreateOrUpdateTags

Creates or updates tags for the specified Auto Scaling group.

When you specify a tag with a key that already exists, the operation overwrites the previous tag definition, and you do not get an error message.

For more information, see [Tag Auto Scaling groups and instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

Tags.member.N

One or more tags.

Type: Array of [Tag](#) objects

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExists

You already have an Auto Scaling group or launch configuration with this name.

message

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ResourceInUse

The operation can't be performed because the resource is in use.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of `CreateOrUpdateTags`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=CreateOrUpdateTags
&Tags.member.1.ResourceId=my-asg
&Tags.member.1.ResourceType=auto-scaling-group
&Tags.member.1.Key=environment
&Tags.member.1.Value=test
&Tags.member.1.PropagateAtLaunch=true
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteAutoScalingGroup

Deletes the specified Auto Scaling group.

If the group has instances or scaling activities in progress, you must specify the option to force the deletion in order for it to succeed. The force delete operation will also terminate the EC2 instances. If the group has a warm pool, the force delete option also deletes the warm pool.

To remove instances from the Auto Scaling group before deleting it, call the [DetachInstances](#) API with the list of instances and the option to decrement the desired capacity. This ensures that Amazon EC2 Auto Scaling does not launch replacement instances.

To terminate all instances before deleting the Auto Scaling group, call the [UpdateAutoScalingGroup](#) API and set the minimum size and desired capacity of the Auto Scaling group to zero.

If the group has scaling policies, deleting the group deletes the policies, the underlying alarm actions, and any alarm that no longer has an associated action.

For more information, see [Delete your Auto Scaling infrastructure](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ForceDelete

Specifies that the group is to be deleted along with all instances associated with the group, without waiting for all instances to be terminated. This action also deletes any outstanding lifecycle actions associated with the group.

Type: Boolean

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ResourceInUse

The operation can't be performed because the resource is in use.

message

HTTP Status Code: 400

ScalingActivityInProgress

The operation can't be performed because there are scaling activities in progress.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of DeleteAutoScalingGroup.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeleteAutoScalingGroup
```

```
&AutoScalingGroupName=my-asg
&ForceDelete=true
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteLaunchConfiguration

Deletes the specified launch configuration.

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

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

LaunchConfigurationName

The name of the launch configuration.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ResourceInUse

The operation can't be performed because the resource is in use.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of DeleteLaunchConfiguration.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeleteLaunchConfiguration
&LaunchConfigurationName=my-lc
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DeleteLifecycleHook.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeleteLifecycleHook
&AutoScalingGroupName=my-asg
&LifecycleHookName=my-hook
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteNotificationConfiguration

Deletes the specified notification.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

TopicARN

The Amazon Resource Name (ARN) of the Amazon SNS topic.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DeleteNotificationConfiguration.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeleteNotificationConfiguration
&AutoScalingGroupName=my-asg
&TopicARN=arn:aws:sns:us-east-1:123456789012:my-sns-topic
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeletePolicy

Deletes the specified scaling policy.

Deleting either a step scaling policy or a simple scaling policy deletes the underlying alarm action, but does not delete the alarm, even if it no longer has an associated action.

For more information, see [Delete a scaling policy](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

PolicyName

The name or Amazon Resource Name (ARN) of the policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DeletePolicy.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeletePolicy
&AutoScalingGroupName=my-asg
&PolicyName=alb1000-target-tracking-scaling-policy
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteScheduledAction

Deletes the specified scheduled action.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ScheduledActionName

The name of the action to delete.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DeleteScheduledAction.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeleteScheduledAction
&AutoScalingGroupName=my-asg
&ScheduledActionName=my-scheduled-action
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteTags

Deletes the specified tags.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

Tags.member.N

One or more tags.

Type: Array of [Tag](#) objects

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ResourceInUse

The operation can't be performed because the resource is in use.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of DeleteTags.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeleteTags
&Tags.member.1.ResourceId=my-asg
&Tags.member.1.ResourceType=auto-scaling-group
&Tags.member.1.Key=environment
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DeleteWarmPool

Deletes the warm pool for the specified Auto Scaling group.

For more information, see [Warm pools for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ForceDelete

Specifies that the warm pool is to be deleted along with all of its associated instances, without waiting for all instances to be terminated. This parameter also deletes any outstanding lifecycle actions associated with the warm pool instances.

Type: Boolean

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ResourceInUse

The operation can't be performed because the resource is in use.

message

HTTP Status Code: 400

ScalingActivityInProgress

The operation can't be performed because there are scaling activities in progress.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of DeleteWarmPool.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DeleteWarmPool
&AutoScalingGroupName=my-asg
&ForceDelete=true
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeAccountLimits

Describes the current Amazon EC2 Auto Scaling resource quotas for your account.

When you establish an AWS account, the account has initial quotas on the maximum number of Auto Scaling groups and launch configurations that you can create in a given Region. For more information, see [Quotas for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Response Elements

The following elements are returned by the service.

MaxNumberOfAutoScalingGroups

The maximum number of groups allowed for your account. The default is 200 groups per Region.

Type: Integer

MaxNumberOfLaunchConfigurations

The maximum number of launch configurations allowed for your account. The default is 200 launch configurations per Region.

Type: Integer

NumberOfAutoScalingGroups

The current number of groups for your account.

Type: Integer

NumberOfLaunchConfigurations

The current number of launch configurations for your account.

Type: Integer

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeAccountLimits.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeAccountLimits
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeAccountLimitsResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <DescribeAccountLimitsResult>
    <NumberOfLaunchConfigurations>5</NumberOfLaunchConfigurations>
    <MaxNumberOfLaunchConfigurations>200</MaxNumberOfLaunchConfigurations>
    <NumberOfAutoScalingGroups>10</NumberOfAutoScalingGroups>
    <MaxNumberOfAutoScalingGroups>200</MaxNumberOfAutoScalingGroups>
  </DescribeAccountLimitsResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeAccountLimitsResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeAdjustmentTypes

Describes the available adjustment types for step scaling and simple scaling policies.

The following adjustment types are supported:

- `ChangeInCapacity`
- `ExactCapacity`
- `PercentChangeInCapacity`

Response Elements

The following element is returned by the service.

AdjustmentTypes.member.N

The policy adjustment types.

Type: Array of [AdjustmentType](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `DescribeAdjustmentTypes`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeAdjustmentTypes
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeAdjustmentTypesResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <DescribeAdjustmentTypesResult>
    <AdjustmentTypes>
      <member>
        <AdjustmentType>ChangeInCapacity</AdjustmentType>
      </member>
      <member>
        <AdjustmentType>ExactCapacity</AdjustmentType>
      </member>
      <member>
        <AdjustmentType>PercentChangeInCapacity</AdjustmentType>
      </member>
    </AdjustmentTypes>
  </DescribeAdjustmentTypesResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeAdjustmentTypesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeAutoScalingGroups

Gets information about the Auto Scaling groups in the account and Region.

If you specify Auto Scaling group names, the output includes information for only the specified Auto Scaling groups. If you specify filters, the output includes information for only those Auto Scaling groups that meet the filter criteria. If you do not specify group names or filters, the output includes information for all Auto Scaling groups.

This operation also returns information about instances in Auto Scaling groups. To retrieve information about the instances in a warm pool, you must call the [DescribeWarmPool](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupNames.member.N

The names of the Auto Scaling groups. By default, you can only specify up to 50 names. You can optionally increase this limit using the `MaxRecords` property.

If you omit this property, all Auto Scaling groups are described.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\u005C\u005C]`*

Required: No

Filters.member.N

One or more filters to limit the results based on specific tags.

Type: Array of [Filter](#) objects

Required: No

IncludeInstances

Specifies whether to include information about Amazon EC2 instances in the response. When set to `true` (default), the response includes instance details.

Type: Boolean

Required: No

MaxRecords

The maximum number of items to return with this call. The default value is 50 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

AutoScalingGroups.member.N

The groups.

Type: Array of [AutoScalingGroup](#) objects

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example 1: Describe an Auto Scaling group

This example describes the specified Auto Scaling group.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeAutoScalingGroups
&AutoScalingGroupNames.member.1=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeAutoScalingGroupsResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <DescribeAutoScalingGroupsResult>
    <AutoScalingGroups>
      <member>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <AutoScalingGroupARN>arn:aws:autoscaling:us-
east-1:123456789012:autoScalingGroup:12345678-1234-1234-1234-123456789012:autoScalingGroupName/
my-asg</AutoScalingGroupARN>
        <LaunchConfigurationName>my-lc</LaunchConfigurationName>
        <LaunchTemplate />
```

```

    <MixedInstancesPolicy />
    <MinSize>2</MinSize>
    <MaxSize>10</MaxSize>
    <DesiredCapacity>2</DesiredCapacity>
    <DefaultCooldown>300</DefaultCooldown>
    <AvailabilityZones>
      <member>us-east-1a</member>
      <member>us-east-1b</member>
    </AvailabilityZones>
    <LoadBalancerNames />
    <TargetGroupARNs />
    <HealthCheckType>ELB</HealthCheckType>
    <HealthCheckGracePeriod>300</HealthCheckGracePeriod>
    <Instances>
      <member>
        <ProtectedFromScaleIn>false</ProtectedFromScaleIn>
        <AvailabilityZone>us-east-1c</AvailabilityZone>
        <LaunchConfigurationName>my-lc</LaunchConfigurationName>
        <LaunchTemplate />
        <InstanceId>i-1234567890abcdef0</InstanceId>
        <InstanceType>t3.micro</InstanceType>
        <HealthStatus>Healthy</HealthStatus>
        <LifecycleState>InService</LifecycleState>
        <WeightedCapacity />
      </member>
    </Instances>
    <CreatedTime>2023-05-06T17:47:15.107Z</CreatedTime>
    <SuspendedProcesses />
    <PlacementGroup />
    <VPCZoneIdentifier>subnet-057fa0918fEXAMPLE,subnet-610acd08EXAMPLE</
VPCZoneIdentifier>
    <EnabledMetrics />
    <Status />
    <Tags>
      <member>
        <ResourceId>my-asg</ResourceId>
        <PropagateAtLaunch>true</PropagateAtLaunch>
        <Value>test</Value>
        <Key>environment</Key>
        <ResourceType>auto-scaling-group</ResourceType>
      </member>
    </Tags>
    <TerminationPolicies>
      <member>Default</member>

```

```
</TerminationPolicies>
<NewInstancesProtectedFromScaleIn>false</NewInstancesProtectedFromScaleIn>
<ServiceLinkedRoleARN>arn:aws:iam::123456789012:role/aws-service-role/
autoscaling.amazonaws.com/AWSServiceRoleForAutoScaling</ServiceLinkedRoleARN>
<MaxInstanceLifetime />
<CapacityRebalance />
<WarmPoolConfiguration />
<WarmPoolSize />
<DesiredCapacityType />
<DefaultInstanceWarmup />
<TrafficSources />
</member>
</AutoScalingGroups>
</DescribeAutoScalingGroupsResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribeAutoScalingGroupsResponse>
```

Example 2: Describe Auto Scaling groups with a specific tag key and value pair

This example describes any Auto Scaling groups with the tag key and value pair of `environment=production`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeAutoScalingGroups
&Filters.member.1.Name=tag-key
&Filters.member.1.Values.member.1=environment
&Filters.member.2.Name=tag-value
&Filters.member.2.Values.member.1=production
&Version=2011-01-01
&AUTHPARAMS
```

Example 3: Describe Auto Scaling groups with tags that match multiple tag keys and values

This example combines filters to do more complex filtering. It describes any Auto Scaling groups with a specific set of tags. One tag key is `environment` and the tag value is `production` or `development`, and the other tag key is `costcenter` and the tag value is `cc123`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeAutoScalingGroups
&Filter.member.1.Name=tag:environment
&Filter.member.1.Values.member.1=production
&Filter.member.1.Values.member.2=development
&Filter.member.2.Name=tag:costcenter
&Filter.member.2.Values.member.1=cc123
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeAutoScalingInstances

Gets information about the Auto Scaling instances in the account and Region.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

InstanceIds.member.N

The IDs of the instances. If you omit this property, all Auto Scaling instances are described. If you specify an ID that does not exist, it is ignored with no error.

Array Members: Maximum number of 50 items.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxRecords

The maximum number of items to return with this call. The default value is 50 and the maximum value is 50.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

AutoScalingInstances.member.N

The instances.

Type: Array of [AutoScalingInstanceDetails](#) objects

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeAutoScalingInstances.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeAutoScalingInstances
&InstanceIds.member.1=i-1234567890abcdef0
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeAutoScalingInstancesResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <DescribeAutoScalingInstancesResult>
    <AutoScalingInstances>
      <member>
        <ProtectedFromScaleIn>false</ProtectedFromScaleIn>
        <AvailabilityZone>us-east-1b</AvailabilityZone>
        <LaunchConfigurationName>my-lc</LaunchConfigurationName>
        <LaunchTemplate />
        <InstanceId>i-1234567890abcdef0</InstanceId>
        <InstanceType>t3.micro</InstanceType>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <HealthStatus>HEALTHY</HealthStatus>
        <LifecycleState>InService</LifecycleState>
        <WeightedCapacity />
      </member>
    </AutoScalingInstances>
  </DescribeAutoScalingInstancesResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeAutoScalingInstancesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeAutoScalingNotificationTypes

Describes the notification types that are supported by Amazon EC2 Auto Scaling.

Response Elements

The following element is returned by the service.

AutoScalingNotificationTypes.member.N

The notification types.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeAutoScalingNotificationTypes.

Sample Request

```
https://autoscaling.amazonaws.com/?  
Version=2011-01-01&Action=DescribeAutoScalingNotificationTypes
```

```
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeAutoScalingNotificationTypesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeAutoScalingNotificationTypesResult>
    <AutoScalingNotificationTypes>
      <member>autoscaling:EC2_INSTANCE_LAUNCH</member>
      <member>autoscaling:EC2_INSTANCE_LAUNCH_ERROR</member>
      <member>autoscaling:EC2_INSTANCE_TERMINATE</member>
      <member>autoscaling:EC2_INSTANCE_TERMINATE_ERROR</member>
      <member>autoscaling:TEST_NOTIFICATION</member>
    </AutoScalingNotificationTypes>
  </DescribeAutoScalingNotificationTypesResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeAutoScalingNotificationTypesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeInstanceRefreshes

Gets information about the instance refreshes for the specified Auto Scaling group from the previous six weeks.

This operation is part of the [instance refresh feature](#) in Amazon EC2 Auto Scaling, which helps you update instances in your Auto Scaling group after you make configuration changes.

To help you determine the status of an instance refresh, Amazon EC2 Auto Scaling returns information about the instance refreshes you previously initiated, including their status, start time, end time, the percentage of the instance refresh that is complete, and the number of instances remaining to update before the instance refresh is complete. If a rollback is initiated while an instance refresh is in progress, Amazon EC2 Auto Scaling also returns information about the rollback of the instance refresh.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceRefreshIds.member.N

One or more instance refresh IDs.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxRecords

The maximum number of items to return with this call. The default value is 50 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

InstanceRefreshes.member.N

The instance refreshes for the specified group, sorted by creation timestamp in descending order.

Type: Array of [InstanceRefresh](#) objects

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeInstanceRefreshes.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeInstanceRefreshes
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeInstanceRefreshesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeInstanceRefreshesResult>
    <InstanceRefreshes>
      <member>
        <InstanceRefreshId>08b91cf7-8fa6-48af-b6a6-d227f40f1b9b</InstanceRefreshId>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <Status>InProgress</Status>
        <StatusReason>Waiting for instances to warm up before continuing. For example:
i-0645704820a8e83ff is warming up.</StatusReason>
```

```
<PercentageComplete>50</PercentageComplete>
<InstancesToUpdate>0</InstancesToUpdate>
<StartTime>2023-06-13T16:46:52Z</StartTime>
<Preferences>{}</Preferences>
</member>
</InstanceRefreshes>
</DescribeInstanceRefreshesResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribeInstanceRefreshesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeLaunchConfigurations

Gets information about the launch configurations in the account and Region.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

LaunchConfigurationNames.member.N

The launch configuration names. If you omit this property, all launch configurations are described.

Array Members: Maximum number of 50 items.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxRecords

The maximum number of items to return with this call. The default value is 50 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

LaunchConfigurations.member.N

The launch configurations.

Type: Array of [LaunchConfiguration](#) objects

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeLaunchConfigurations.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeLaunchConfigurations
&LaunchConfigurationNames.member.1=my-lc
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeLaunchConfigurationsResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <DescribeLaunchConfigurationsResult>
    <LaunchConfigurations>
      <member>
        <KernelId />
        <EbsOptimized>false</EbsOptimized>
        <RamdiskId />
        <UserData />
        <ImageId>ami-0abcdef1234567890</ImageId>
        <BlockDeviceMappings />
        <ClassicLinkVPCSecurityGroups />
        <InstanceType>t2.micro</InstanceType>
        <KeyName />
        <LaunchConfigurationARN>arn:aws:autoscaling:us-
east-1:123456789012:launchConfiguration:12345678-1234-1234-1234-123456789012:launchConfiguration
my-lc</LaunchConfigurationARN>
        <LaunchConfigurationName>my-lc</LaunchConfigurationName>
        <CreatedTime>2022-01-21T23:04:42.200Z</CreatedTime>
        <SecurityGroups>
          <member>sg-1a2b3c4d</member>
        </SecurityGroups>
        <InstanceMonitoring>
          <Enabled>true</Enabled>
        </InstanceMonitoring>
      </member>
    </LaunchConfigurations>
  </DescribeLaunchConfigurationsResult>
```

```
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribeLaunchConfigurationsResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeLifecycleHooks

Gets information about the lifecycle hooks for the specified Auto Scaling group.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

LifecycleHookNames.member.N

The names of one or more lifecycle hooks. If you omit this property, all lifecycle hooks are described.

Type: Array of strings

Array Members: Maximum number of 50 items.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z0-9\-\_\./]+`

Required: No

Response Elements

The following element is returned by the service.

LifecycleHooks.member.N

The lifecycle hooks for the specified group.

Type: Array of [LifecycleHook](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeLifecycleHooks.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeLifecycleHooks
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeLifecycleHooksResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <DescribeLifecycleHooksResult>
    <LifecycleHooks>
      <member>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <RoleARN>arn:aws:iam::1234567890:role/my-auto-scaling-role</RoleARN>
        <LifecycleTransition>autoscaling:EC2_INSTANCE_LAUNCHING</LifecycleTransition>
        <GlobalTimeout>172800</GlobalTimeout>
        <LifecycleHookName>my-launch-hook</LifecycleHookName>
```

```
<HeartbeatTimeout>3600</HeartbeatTimeout>
<DefaultResult>ABANDON</DefaultResult>
<NotificationTargetARN>arn:aws:sqs:us-east-1:123456789012:my-queue</
NotificationTargetARN>
</member>
</LifecycleHooks>
</DescribeLifecycleHooksResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribeLifecycleHooksResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeLifecycleHookTypes

Describes the available types of lifecycle hooks.

The following hook types are supported:

- `autoscaling:EC2_INSTANCE_LAUNCHING`
- `autoscaling:EC2_INSTANCE_TERMINATING`

Response Elements

The following element is returned by the service.

LifecycleHookTypes.member.N

The lifecycle hook types.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeLifecycleHookTypes.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeLifecycleHookTypes
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeLifecycleHookTypesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeLifecycleHookTypesResult>
    <LifecycleHookTypes>
      <member>autoscaling:EC2_INSTANCE_LAUNCHING</member>
      <member>autoscaling:EC2_INSTANCE_TERMINATING</member>
    </LifecycleHookTypes>
  </DescribeLifecycleHookTypesResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeLifecycleHookTypesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeLoadBalancers

Note

This API operation is superseded by [DescribeTrafficSources](#), which can describe multiple traffic sources types. We recommend using `DescribeTrafficSources` to simplify how you manage traffic sources. However, we continue to support `DescribeLoadBalancers`. You can use both the original `DescribeLoadBalancers` API operation and `DescribeTrafficSources` on the same Auto Scaling group.

Gets information about the load balancers for the specified Auto Scaling group.

This operation describes only Classic Load Balancers. If you have Application Load Balancers, Network Load Balancers, or Gateway Load Balancers, use the [DescribeLoadBalancerTargetGroups](#) API instead.

To determine the attachment status of the load balancer, use the `State` element in the response. When you attach a load balancer to an Auto Scaling group, the initial `State` value is `Adding`. The state transitions to `Added` after all Auto Scaling instances are registered with the load balancer. If Elastic Load Balancing health checks are enabled for the Auto Scaling group, the state transitions to `InService` after at least one Auto Scaling instance passes the health check. When the load balancer is in the `InService` state, Amazon EC2 Auto Scaling can terminate and replace any instances that are reported as unhealthy. If no registered instances pass the health checks, the load balancer doesn't enter the `InService` state.

Load balancers also have an `InService` state if you attach them in the [CreateAutoScalingGroup](#) API call. If your load balancer state is `InService`, but it is not working properly, check the scaling activities by calling [DescribeScalingActivities](#) and take any corrective actions necessary.

For help with failed health checks, see [Troubleshooting Amazon EC2 Auto Scaling: Health checks](#) in the *Amazon EC2 Auto Scaling User Guide*. For more information, see [Use Elastic Load Balancing to distribute traffic across the instances in your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

MaxRecords

The maximum number of items to return with this call. The default value is 100 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

LoadBalancers.member.N

The load balancers.

Type: Array of [LoadBalancerState](#) objects

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeLoadBalancers.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeLoadBalancers
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeLoadBalancersResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
```

```
<DescribeLoadBalancersResult>
  <LoadBalancers>
    <member>
      <LoadBalancerName>my-loadbalancer</LoadBalancerName>
      <State>Added</State>
    </member>
  </LoadBalancers>
</DescribeLoadBalancersResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribeLoadBalancersResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeLoadBalancerTargetGroups

Note

This API operation is superseded by [DescribeTrafficSources](#), which can describe multiple traffic sources types. We recommend using `DetachTrafficSources` to simplify how you manage traffic sources. However, we continue to support `DescribeLoadBalancerTargetGroups`. You can use both the original `DescribeLoadBalancerTargetGroups` API operation and `DescribeTrafficSources` on the same Auto Scaling group.

Gets information about the Elastic Load Balancing target groups for the specified Auto Scaling group.

To determine the attachment status of the target group, use the `State` element in the response. When you attach a target group to an Auto Scaling group, the initial `State` value is `Adding`. The state transitions to `Added` after all Auto Scaling instances are registered with the target group. If Elastic Load Balancing health checks are enabled for the Auto Scaling group, the state transitions to `InService` after at least one Auto Scaling instance passes the health check. When the target group is in the `InService` state, Amazon EC2 Auto Scaling can terminate and replace any instances that are reported as unhealthy. If no registered instances pass the health checks, the target group doesn't enter the `InService` state.

Target groups also have an `InService` state if you attach them in the [CreateAutoScalingGroup](#) API call. If your target group state is `InService`, but it is not working properly, check the scaling activities by calling [DescribeScalingActivities](#) and take any corrective actions necessary.

For help with failed health checks, see [Troubleshooting Amazon EC2 Auto Scaling: Health checks](#) in the *Amazon EC2 Auto Scaling User Guide*. For more information, see [Use Elastic Load Balancing to distribute traffic across the instances in your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Note

You can use this operation to describe target groups that were attached by using [AttachLoadBalancerTargetGroups](#), but not for target groups that were attached by using [AttachTrafficSources](#).

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

MaxRecords

The maximum number of items to return with this call. The default value is 100 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

LoadBalancerTargetGroups.member.N

Information about the target groups.

Type: Array of [LoadBalancerTargetGroupState](#) objects

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the `NextToken` value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The `NextToken` value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeMetricCollectionTypes

Describes the available CloudWatch metrics for Amazon EC2 Auto Scaling.

Response Elements

The following elements are returned by the service.

Granularities.member.N

The granularities for the metrics.

Type: Array of [MetricGranularityType](#) objects

Metrics.member.N

The metrics.

Type: Array of [MetricCollectionType](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeMetricCollectionTypes.

Sample Request

```
https://autoscaling.amazonaws.com/?
Version=2011-01-01&Action=DescribeMetricCollectionTypes
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeMetricCollectionTypesResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <DescribeMetricCollectionTypesResult>
    <Granularities>
      <member>
        <Granularity>1Minute</Granularity>
      </member>
    </Granularities>
    <Metrics>
      <member>
        <Metric>GroupMinSize</Metric>
      </member>
      <member>
        <Metric>GroupMaxSize</Metric>
      </member>
      <member>
        <Metric>GroupDesiredCapacity</Metric>
      </member>
      <member>
        <Metric>GroupInServiceInstances</Metric>
      </member>
      <member>
        <Metric>GroupPendingInstances</Metric>
      </member>
      <member>
        <Metric>GroupTerminatingInstances</Metric>
      </member>
      <member>
        <Metric>GroupStandbyInstances</Metric>
      </member>
      <member>
        <Metric>GroupTotalInstances</Metric>
      </member>
    </Metrics>
  </DescribeMetricCollectionTypesResult>
</DescribeMetricCollectionTypesResponse>
```

```
    <Metric>GroupInServiceCapacity</Metric>
  </member>
  <member>
    <Metric>GroupPendingCapacity</Metric>
  </member>
  <member>
    <Metric>GroupStandbyCapacity</Metric>
  </member>
  <member>
    <Metric>GroupTerminatingCapacity</Metric>
  </member>
  <member>
    <Metric>GroupTotalCapacity</Metric>
  </member>
  <member>
    <Metric>WarmPoolDesiredCapacity</Metric>
  </member>
  <member>
    <Metric>WarmPoolWarmedCapacity</Metric>
  </member>
  <member>
    <Metric>WarmPoolPendingCapacity</Metric>
  </member>
  <member>
    <Metric>WarmPoolTerminatingCapacity</Metric>
  </member>
  <member>
    <Metric>WarmPoolTotalCapacity</Metric>
  </member>
  <member>
    <Metric>GroupAndWarmPoolDesiredCapacity</Metric>
  </member>
  <member>
    <Metric>GroupAndWarmPoolTotalCapacity</Metric>
  </member>
</Metrics>
</DescribeMetricCollectionTypesResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribeMetricCollectionTypesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeNotificationConfigurations

Gets information about the Amazon SNS notifications that are configured for one or more Auto Scaling groups.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupNames.member.N

The name of the Auto Scaling group.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxRecords

The maximum number of items to return with this call. The default value is 50 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

NotificationConfigurations.member.N

The notification configurations.

Type: Array of [NotificationConfiguration](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeNotificationConfigurations.

Sample Request

```
https://autoscaling.amazonaws.com/?
Version=2011-01-01&Action=DescribeNotificationConfigurations
&AutoScalingGroupNames.member.1=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeNotificationConfigurationsResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeNotificationConfigurationsResult>
    <NotificationConfigurations>
      <member>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <NotificationType>autoscaling:EC2_INSTANCE_LAUNCH</NotificationType>
        <TopicARN>arn:aws:sns:us-east-1:123456789012:my-sns-topic</TopicARN>
      </member>
    </NotificationConfigurations>
  </DescribeNotificationConfigurationsResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeNotificationConfigurationsResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribePolicies

Gets information about the scaling policies in the account and Region.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxRecords

The maximum number of items to be returned with each call. The default value is 50 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

PolicyNames.member.N

The names of one or more policies. If you omit this property, all policies are described. If a group name is provided, the results are limited to that group. If you specify an unknown policy name, it is ignored with no error.

Array Members: Maximum number of 50 items.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

PolicyTypes.member.N

One or more policy types. The valid values are `SimpleScaling`, `StepScaling`, `TargetTrackingScaling`, and `PredictiveScaling`.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the `NextToken` value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

ScalingPolicies.member.N

The scaling policies.

Type: Array of [ScalingPolicy](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribePolicies.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribePolicies
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribePoliciesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribePoliciesResult>
    <ScalingPolicies>
      <member>
```

```

    <PolicyARN>arn:aws:autoscaling:us-
east-1:123456789012:scalingPolicy:c322761b-3172-4d56-9a21-0ed9dEXAMPLE:autoScalingGroupName/
my-asg:policyName/MyScaleInPolicy</PolicyARN>
    <AdjustmentType>ChangeInCapacity</AdjustmentType>
    <ScalingAdjustment>-1</ScalingAdjustment>
    <PolicyName>MyScaleInPolicy</PolicyName>
    <PolicyType>SimpleScaling</PolicyType>
    <AutoScalingGroupName>my-asg</AutoScalingGroupName>
    <Cooldown>60</Cooldown>
    <Alarms>
      <member>
        <AlarmName>TestQueue</AlarmName>
        <AlarmARN>arn:aws:cloudwatch:us-east-1:123456789012:alarm:TestQueue</
AlarmARN>
      </member>
    </Alarms>
  </member>
  <member>
    <PolicyARN>arn:aws:autoscaling:us-
east-1:123456789012:scalingPolicy:c55a5cdd-9be0-435b-b60b-
a8dd3EXAMPLE:autoScalingGroupName/my-asg:policyName/MyScaleOutPolicy</PolicyARN>
    <AdjustmentType>ChangeInCapacity</AdjustmentType>
    <ScalingAdjustment>1</ScalingAdjustment>
    <PolicyName>MyScaleOutPolicy</PolicyName>
    <PolicyType>SimpleScaling</PolicyType>
    <AutoScalingGroupName>my-asg</AutoScalingGroupName>
    <Cooldown>60</Cooldown>
    <Alarms>
      <member>
        <AlarmName>TestQueue</AlarmName>
        <AlarmARN>arn:aws:cloudwatch:us-east-1:123456789012:alarm:TestQueue</
AlarmARN>
      </member>
    </Alarms>
  </member>
</ScalingPolicies>
</DescribePoliciesResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribePoliciesResponse>

```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeScalingActivities

Gets information about the scaling activities in the account and Region.

When scaling events occur, you see a record of the scaling activity in the scaling activities. For more information, see [Verify a scaling activity for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

If the scaling event succeeds, the value of the `StatusCode` element in the response is `Successful`. If an attempt to launch instances failed, the `StatusCode` value is `Failed` or `Cancelled` and the `StatusMessage` element in the response indicates the cause of the failure. For help interpreting the `StatusMessage`, see [Troubleshooting Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

ActivityIds.member.N

The activity IDs of the desired scaling activities. If you omit this property, all activities for the past six weeks are described. If unknown activities are requested, they are ignored with no error. If you specify an Auto Scaling group, the results are limited to that group.

Array Members: Maximum number of 50 IDs.

Type: Array of strings

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

IncludeDeletedGroups

Indicates whether to include scaling activity from deleted Auto Scaling groups.

Type: Boolean

Required: No

MaxRecords

The maximum number of items to return with this call. The default value is 100 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

Activities.member.N

The scaling activities. Activities are sorted by start time. Activities still in progress are described first.

Type: Array of [Activity](#) objects

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeScalingActivities.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeScalingActivities
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeScalingActivitiesResponse xmlns="https://ec2.amazonaws.com/doc/2011-01-01/">
  <DescribeScalingActivitiesResult>
    <Activities>
```

```

    <member>
      <StatusCode>Failed</StatusCode>
      <Progress>0</Progress>
      <ActivityId>12345678-1234-1234-1234-123456789012</ActivityId>
      <StartTime>2019-04-12T17:32:07.882Z</StartTime>
      <AutoScalingGroupName>my-asg</AutoScalingGroupName>
      <AutoScalingGroupARN>arn:aws:autoscaling:us-east-1:123456789012:autoScalingGroup:12345678-1234-1234-1234-123456789012:autoScalingGroupName/my-asg</AutoScalingGroupARN>
      <Cause>At 2019-04-12T17:31:30Z a user request created an AutoScalingGroup changing the desired capacity from 0 to 1. At 2019-04-12T17:32:07Z an instance was started in response to a difference between desired and actual capacity, increasing the capacity from 0 to 1.</Cause>
      <Details>{}</Details>
      <Description>Launching a new EC2 instance. Status Reason: The image id 'ami-xxxxxxx' does not exist. Launching EC2 instance failed.</Description>
      <EndTime>2019-04-12T17:32:08Z</EndTime>
      <StatusMessage>The image id 'ami-xxxxxxx' does not exist. Launching EC2 instance failed.</StatusMessage>
    </member>
  </Activities>
</DescribeScalingActivitiesResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</DescribeScalingActivitiesResponse>

```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeScalingProcessTypes

Describes the scaling process types for use with the [ResumeProcesses](#) and [SuspendProcesses](#) APIs.

Response Elements

The following element is returned by the service.

Processes.member.N

The names of the process types.

Type: Array of [ProcessType](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeScalingProcessTypes.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeScalingProcessTypes
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeScalingProcessTypesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeScalingProcessTypesResult>
    <Processes>
      <member>
        <ProcessName>AZRebalance</ProcessName>
      </member>
      <member>
        <ProcessName>AddToLoadBalancer</ProcessName>
      </member>
      <member>
        <ProcessName>AlarmNotification</ProcessName>
      </member>
      <member>
        <ProcessName>HealthCheck</ProcessName>
      </member>
      <member>
        <ProcessName>InstanceRefresh</ProcessName>
      </member>
      <member>
        <ProcessName>Launch</ProcessName>
      </member>
      <member>
        <ProcessName>ReplaceUnhealthy</ProcessName>
      </member>
      <member>
        <ProcessName>ScheduledActions</ProcessName>
      </member>
      <member>
        <ProcessName>Terminate</ProcessName>
      </member>
    </Processes>
  </DescribeScalingProcessTypesResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeScalingProcessTypesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeScheduledActions

Gets information about the scheduled actions that haven't run or that have not reached their end time.

To describe the scaling activities for scheduled actions that have already run, call the [DescribeScalingActivities](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EndTime

The latest scheduled start time to return. If scheduled action names are provided, this property is ignored.

Type: Timestamp

Required: No

MaxRecords

The maximum number of items to return with this call. The default value is 50 and the maximum value is 100.

Type: Integer

Required: No

NextToken

The token for the next set of items to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ScheduledActionNames.member.N

The names of one or more scheduled actions. If you omit this property, all scheduled actions are described. If you specify an unknown scheduled action, it is ignored with no error.

Array Members: Maximum number of 50 actions.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

StartTime

The earliest scheduled start time to return. If scheduled action names are provided, this property is ignored.

Type: Timestamp

Required: No

Response Elements

The following elements are returned by the service.

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the `NextToken` value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

ScheduledUpdateGroupActions.member.N

The scheduled actions.

Type: Array of [ScheduledUpdateGroupAction](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

Required: No

Response Elements

The following elements are returned by the service.

NextToken

A string that indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Tags.member.N

One or more tags.

Type: Array of [TagDescription](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeTags.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeTags
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeTagsResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeTagsResult>
    <Tags>
      <member>
        <ResourceId>my-asg</ResourceId>
        <PropagateAtLaunch>true</PropagateAtLaunch>
        <Value>test</Value>
        <Key>environment</Key>
        <ResourceType>auto-scaling-group</ResourceType>
      </member>
    </Tags>
  </DescribeTagsResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeTagsResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)

- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeTerminationPolicyTypes

Describes the termination policies supported by Amazon EC2 Auto Scaling.

For more information, see [Configure termination policies for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Response Elements

The following element is returned by the service.

TerminationPolicyTypes.member.N

The termination policies supported by Amazon EC2 Auto Scaling: OldestInstance, OldestLaunchConfiguration, NewestInstance, ClosestToNextInstanceHour, Default, OldestLaunchTemplate, and AllocationStrategy.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeTerminationPolicyTypes.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeTerminationPolicyTypes
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeTerminationPolicyTypesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeTerminationPolicyTypesResult>
    <TerminationPolicyTypes>
      <member>AllocationStrategy</member>
      <member>ClosestToNextInstanceHour</member>
      <member>Default</member>
      <member>NewestInstance</member>
      <member>OldestInstance</member>
      <member>OldestLaunchConfiguration</member>
      <member>OldestLaunchTemplate</member>
    </TerminationPolicyTypes>
  </DescribeTerminationPolicyTypesResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeTerminationPolicyTypesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)

- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

TrafficSourceType

The traffic source type that you want to describe.

The following lists the valid values:

- `e1b` if the traffic source is a Classic Load Balancer.
- `e1bv2` if the traffic source is a Application Load Balancer, Gateway Load Balancer, or Network Load Balancer.
- `vpc-lattice` if the traffic source is VPC Lattice.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

NextToken

This string indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the `NextToken` value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

TrafficSources.member.N

Information about the traffic sources.

Type: Array of [TrafficSourceState](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeTrafficSources.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeTrafficSources
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeTrafficSourcesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeTrafficSourcesResult>
    <TrafficSources>
      <Identifier>arn:aws:vpc-lattice:us-west-2:123456789012:targetgroup/tg-0e2f2665eEXAMPLE</Identifier>
      <State>InService</State>
      <Type>vpc-lattice</Type>
    </TrafficSources>
  </DescribeTrafficSourcesResult>
```

```
<ResponseMetadata>
  <RequestId>9351b5ea-b4c0-4a88-b873-c51e4EXAMPLE</RequestId>
</ResponseMetadata>
</DescribeTrafficSourcesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DescribeWarmPool

Gets information about a warm pool and its instances.

For more information, see [Warm pools for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

MaxRecords

The maximum number of instances to return with this call. The maximum value is 50.

Type: Integer

Required: No

NextToken

The token for the next set of instances to return. (You received this token from a previous call.)

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

Instances.member.N

The instances that are currently in the warm pool.

Type: Array of [Instance](#) objects

NextToken

This string indicates that the response contains more items than can be returned in a single response. To receive additional items, specify this string for the NextToken value when requesting the next set of items. This value is null when there are no more items to return.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

WarmPoolConfiguration

The warm pool configuration details.

Type: [WarmPoolConfiguration](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InvalidNextToken

The NextToken value is not valid.

message

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DescribeWarmPool.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DescribeWarmPool
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DescribeWarmPoolResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DescribeWarmPoolResult>
    <WarmPoolConfiguration>
      <MinSize>30</MinSize>
      <MaxGroupPreparedCapacity>0</MaxGroupPreparedCapacity>
      <PoolState>Hibernated</PoolState>
    </WarmPoolConfiguration>
    <Instances />
  </DescribeWarmPoolResult>
  <ResponseMetadata>
    <RequestId>9351b5ea-b4c0-4a88-b873-c51e4EXAMPLE</RequestId>
  </ResponseMetadata>
</DescribeWarmPoolResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DetachInstances

Removes one or more instances from the specified Auto Scaling group.

After the instances are detached, you can manage them independent of the Auto Scaling group.

If you do not specify the option to decrement the desired capacity, Amazon EC2 Auto Scaling launches instances to replace the ones that are detached.

If there is a Classic Load Balancer attached to the Auto Scaling group, the instances are deregistered from the load balancer. If there are target groups attached to the Auto Scaling group, the instances are deregistered from the target groups.

For more information, see [Detach or attach instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceIds.member.N

The IDs of the instances. You can specify up to 20 instances.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ShouldDecrementDesiredCapacity

Indicates whether the Auto Scaling group decrements the desired capacity value by the number of instances detached.

Type: Boolean

Required: Yes

Response Elements

The following element is returned by the service.

Activities.member.N

The activities related to detaching the instances from the Auto Scaling group.

Type: Array of [Activity](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of DetachInstances.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DetachInstances
```

```
&AutoScalingGroupName=my-asg
&InstanceIds.member.1=i-1234567890abcdef0
&ShouldDecrementDesiredCapacity=true
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<DetachInstancesResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <DetachInstancesResult>
    <Activities>
      <member>
        <ActivityId>12345678-1234-1234-1234-123456789012</ActivityId>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <Description>Detaching EC2 instance: i-1234567890abcdef0</Description>
        <Cause>At 2022-06-14T00:07:30Z instance i-1234567890abcdef0 was detached in
response to a user request, shrinking the capacity from 4 to 3.</Cause>
        <Progress>50</Progress>
        <StartTime>2022-06-14T00:07:30.280Z</StartTime>
        <Details>{"Availability Zone":"us-
east-1a", "SubnetID":"subnet-057fa0918fEXAMPLE"}</Details>
        <StatusCode>InProgress</StatusCode>
      </member>
    </Activities>
  </DetachInstancesResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</DetachInstancesResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DetachLoadBalancers

Note

This API operation is superseded by [DetachTrafficSources](#), which can detach multiple traffic sources types. We recommend using `DetachTrafficSources` to simplify how you manage traffic sources. However, we continue to support `DetachLoadBalancers`. You can use both the original `DetachLoadBalancers` API operation and `DetachTrafficSources` on the same Auto Scaling group.

Detaches one or more Classic Load Balancers from the specified Auto Scaling group.

This operation detaches only Classic Load Balancers. If you have Application Load Balancers, Network Load Balancers, or Gateway Load Balancers, use the [DetachLoadBalancerTargetGroups](#) API instead.

When you detach a load balancer, it enters the `Removing` state while deregistering the instances in the group. When all instances are deregistered, then you can no longer describe the load balancer using the [DescribeLoadBalancers](#) API call. The instances remain running.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

LoadBalancerNames.member.N

The names of the load balancers. You can specify up to 10 load balancers.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `DetachLoadBalancers`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DetachLoadBalancers
&AutoScalingGroupName=my-asg
&LoadBalancerNames.member.1=my-lb
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DetachLoadBalancerTargetGroups

Note

This API operation is superseded by [DetachTrafficSources](#), which can detach multiple traffic sources types. We recommend using `DetachTrafficSources` to simplify how you manage traffic sources. However, we continue to support `DetachLoadBalancerTargetGroups`. You can use both the original `DetachLoadBalancerTargetGroups` API operation and `DetachTrafficSources` on the same Auto Scaling group.

Detaches one or more target groups from the specified Auto Scaling group.

When you detach a target group, it enters the `Removing` state while deregistering the instances in the group. When all instances are deregistered, then you can no longer describe the target group using the [DescribeLoadBalancerTargetGroups](#) API call. The instances remain running.

Note

You can use this operation to detach target groups that were attached by using [AttachLoadBalancerTargetGroups](#), but not for target groups that were attached by using [AttachTrafficSources](#).

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

TargetGroupARNs.member.N

The Amazon Resource Names (ARN) of the target groups. You can specify up to 10 target groups.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DetachTrafficSources

Detaches one or more traffic sources from the specified Auto Scaling group.

When you detach a traffic source, it enters the Removing state while deregistering the instances in the group. When all instances are deregistered, then you can no longer describe the traffic source using the [DescribeTrafficSources](#) API call. The instances continue to run.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

TrafficSources.member.N

The unique identifiers of one or more traffic sources. You can specify up to 10 traffic sources.

Type: Array of [TrafficSourceIdentifier](#) objects

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example detaches the Classic Load Balancer named `my-classic-load-balancer` from the Auto Scaling group named `my-asg`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DetachTrafficSources
&AutoScalingGroupName=my-asg
&TrafficSources.member.1.Identifier=my-classic-load-balancer
&TrafficSources.member.1.Type=elb
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

DisableMetricsCollection

Disables group metrics collection for the specified Auto Scaling group.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Metrics.member.N

Identifies the metrics to disable.

You can specify one or more of the following metrics:

- GroupMinSize
- GroupMaxSize
- GroupDesiredCapacity
- GroupInServiceInstances
- GroupPendingInstances
- GroupStandbyInstances
- GroupTerminatingInstances
- GroupTotalInstances
- GroupInServiceCapacity
- GroupPendingCapacity
- GroupStandbyCapacity
- GroupTerminatingCapacity
- GroupTotalCapacity

- WarmPoolDesiredCapacity
- WarmPoolWarmedCapacity
- WarmPoolPendingCapacity
- WarmPoolTerminatingCapacity
- WarmPoolTotalCapacity
- GroupAndWarmPoolDesiredCapacity
- GroupAndWarmPoolTotalCapacity

If you omit this property, all metrics are disabled.

For more information, see [Amazon CloudWatch metrics for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `DisableMetricsCollection`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=DisableMetricsCollection
&AutoScalingGroupName=my-asg
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

EnableMetricsCollection

Enables group metrics collection for the specified Auto Scaling group.

You can use these metrics to track changes in an Auto Scaling group and to set alarms on threshold values. You can view group metrics using the Amazon EC2 Auto Scaling console or the CloudWatch console. For more information, see [Monitor CloudWatch metrics for your Auto Scaling groups and instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

Granularity

The frequency at which Amazon EC2 Auto Scaling sends aggregated data to CloudWatch. The only valid value is 1Minute.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

Metrics.member.N

Identifies the metrics to enable.

You can specify one or more of the following metrics:

- GroupMinSize
- GroupMaxSize
- GroupDesiredCapacity
- GroupInServiceInstances
- GroupPendingInstances
- GroupStandbyInstances
- GroupTerminatingInstances
- GroupTotalInstances
- GroupInServiceCapacity
- GroupPendingCapacity
- GroupStandbyCapacity
- GroupTerminatingCapacity
- GroupTotalCapacity
- WarmPoolDesiredCapacity
- WarmPoolWarmedCapacity
- WarmPoolPendingCapacity
- WarmPoolTerminatingCapacity
- WarmPoolTotalCapacity
- GroupAndWarmPoolDesiredCapacity
- GroupAndWarmPoolTotalCapacity

If you specify `Granularity` and don't specify any metrics, all metrics are enabled.

For more information, see [Amazon CloudWatch metrics for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of EnableMetricsCollection.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=EnableMetricsCollection
&AutoScalingGroupName=my-asg
&Granularity=1Minute
&Metrics.member.1=GroupDesiredCapacity
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

EnterStandby

Moves the specified instances into the standby state.

If you choose to decrement the desired capacity of the Auto Scaling group, the instances can enter standby as long as the desired capacity of the Auto Scaling group after the instances are placed into standby is equal to or greater than the minimum capacity of the group.

If you choose not to decrement the desired capacity of the Auto Scaling group, the Auto Scaling group launches new instances to replace the instances on standby.

For more information, see [Temporarily removing instances from your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceIds.member.N

The IDs of the instances. You can specify up to 20 instances.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ShouldDecrementDesiredCapacity

Indicates whether to decrement the desired capacity of the Auto Scaling group by the number of instances moved to Standby mode.

Type: Boolean

Required: Yes

Response Elements

The following element is returned by the service.

Activities.member.N

The activities related to moving instances into Standby mode.

Type: Array of [Activity](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of EnterStandby.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=EnterStandby
```

```
&AutoScalingGroupName=my-asg
&InstanceIds.member.1=i-1234567890abcdef0
&ShouldDecrementDesiredCapacity=true
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<EnterStandbyResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <EnterStandbyResult>
    <Activities>
      <member>
        <ActivityId>12345678-1234-1234-1234-123456789012</ActivityId>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <Description>Moving EC2 instance to Standby: i-1234567890abcdef0</Description>
        <Progress>50</Progress>
        <Cause>At 2022-06-13T22:35:50Z instance i-1234567890abcdef0 was moved to
standby in response to a user request, shrinking the capacity from 4 to 3.</Cause>
        <StartTime>2022-06-13T22:35:50.884Z</StartTime>
        <Details>{"Availability Zone":"us-
east-1a", "SubnetID":"subnet-057fa0918fEXAMPLE"}</Details>
        <StatusCode>InProgress</StatusCode>
      </member>
    </Activities>
  </EnterStandbyResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</EnterStandbyResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ExecutePolicy

Executes the specified policy. This can be useful for testing the design of your scaling policy.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

BreachThreshold

The breach threshold for the alarm.

Required if the policy type is `StepScaling` and not supported otherwise.

Type: Double

Required: No

HonorCooldown

Indicates whether Amazon EC2 Auto Scaling waits for the cooldown period to complete before executing the policy.

Valid only if the policy type is `SimpleScaling`. For more information, see [Scaling cooldowns for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

MetricValue

The metric value to compare to `BreachThreshold`. This enables you to execute a policy of type `StepScaling` and determine which step adjustment to use. For example, if the breach

HTTP Status Code: 400

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ExitStandby

Moves the specified instances out of the standby state.

After you put the instances back in service, the desired capacity is incremented.

For more information, see [Temporarily removing instances from your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceIds.member.N

The IDs of the instances. You can specify up to 20 instances.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following element is returned by the service.

Activities.member.N

The activities related to moving instances out of Standby mode.

Type: Array of [Activity](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `ExitStandby`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=ExitStandby
&AutoScalingGroupName=my-asg
&InstanceIds.member.1=i-1234567890abcdef0
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<ExitStandbyResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <ExitStandbyResult>
    <Activities>
      <member>
        <ActivityId>12345678-1234-1234-1234-123456789012</ActivityId>
        <AutoScalingGroupName>my-asg</AutoScalingGroupName>
        <Description>Moving EC2 instance out of Standby: i-1234567890abcdef0</
Description>
        <Progress>30</Progress>
```

```
<Cause>At 2022-06-13T22:43:53Z instance i-1234567890abcdef0 was moved out of
standby in response to a user request, increasing the capacity from 3 to 4.</Cause>
<StartTime>2022-06-13T22:43:53.523Z</StartTime>
<Details>{"Availability Zone":"us-
east-1a", "SubnetID":"subnet-057fa0918fEXAMPLE"}</Details>
<StatusCode>PreInService</StatusCode>
</member>
</Activities>
</ExitStandbyResult>
<ResponseMetadata>
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
</ResponseMetadata>
</ExitStandbyResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

GetPredictiveScalingForecast

Retrieves the forecast data for a predictive scaling policy.

Load forecasts are predictions of the hourly load values using historical load data from CloudWatch and an analysis of historical trends. Capacity forecasts are represented as predicted values for the minimum capacity that is needed on an hourly basis, based on the hourly load forecast.

A minimum of 24 hours of data is required to create the initial forecasts. However, having a full 14 days of historical data results in more accurate forecasts.

For more information, see [Predictive scaling for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

EndTime

The exclusive end time of the time range for the forecast data to get. The maximum time duration between the start and end time is 30 days.

Although this parameter can accept a date and time that is more than two days in the future, the availability of forecast data has limits. Amazon EC2 Auto Scaling only issues forecasts for periods of two days in advance.

Type: Timestamp

Required: Yes

PolicyName

The name of the policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

StartTime

The inclusive start time of the time range for the forecast data to get. At most, the date and time can be one year before the current date and time.

Type: Timestamp

Required: Yes

Response Elements

The following elements are returned by the service.

CapacityForecast

The capacity forecast.

Type: [CapacityForecast](#) object

LoadForecast.member.N

The load forecast.

Type: Array of [LoadForecast](#) objects

UpdateTime

The time the forecast was made.

Type: Timestamp

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `GetPredictiveScalingForecast`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=GetPredictiveScalingForecast
&AutoScalingGroupName=my-asg
&PolicyName=cpu40-predictive-scaling-policy
&StartTime=2021-04-29T08:00:00Z
&EndTime=2021-05-29T08:00:00Z
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

LaunchInstances

Launches a specified number of instances in an Auto Scaling group. Returns instance IDs and other details if launch is successful or error details if launch is unsuccessful.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group to launch instances into.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

AvailabilityZoneIds.member.N

A list of Availability Zone IDs where instances should be launched. Must match or be included in the group's AZ configuration. You cannot specify both AvailabilityZones and AvailabilityZoneIds. Required for multi-AZ groups, optional for single-AZ groups.

Type: Array of strings

Array Members: Maximum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

AvailabilityZones.member.N

The Availability Zones for the instance launch. Must match or be included in the Auto Scaling group's Availability Zone configuration. Either AvailabilityZones or SubnetIds must be specified for groups with multiple Availability Zone configurations.

Type: Array of strings

Array Members: Maximum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ClientToken

A unique, case-sensitive identifier to ensure idempotency of the request.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[A-Za-z0-9\-\_\./]+`

Required: Yes

RequestedCapacity

The number of instances to launch. Although this value can exceed 100 for instance weights, the actual instance count is limited to 100 instances per launch.

Type: Integer

Valid Range: Minimum value of 1.

Required: Yes

RetryStrategy

Specifies whether to retry asynchronously if the synchronous launch fails. Valid values are NONE (default, no async retry) and RETRY_WITH_GROUP_CONFIGURATION (increase desired capacity and retry with group configuration).

Type: String

Valid Values: `retry-with-group-configuration` | `none`

Required: No

SubnetIds.member.N

The subnet IDs for the instance launch. Either `AvailabilityZones` or `SubnetIds` must be specified. If both are specified, the subnets must reside in the specified Availability Zones.

Type: Array of strings

Array Members: Maximum number of 1 item.

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Response Elements

The following elements are returned by the service.

AutoScalingGroupName

The name of the Auto Scaling group where the instances were launched.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

ClientToken

The idempotency token used for the request, either customer-specified or auto-generated.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[A-Za-z0-9\-\_\./]+`

Errors.member.N

A list of errors encountered during the launch attempt including details about failed instance launches with their corresponding error codes and messages.

Type: Array of [LaunchInstancesError](#) objects

Instances.member.N

A list of successfully launched instances including details such as instance type, Availability Zone, subnet, lifecycle state, and instance IDs.

Type: Array of [InstanceCollection](#) objects

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

IdempotentParameterMismatch

Indicates that the parameters in the current request do not match the parameters from a previous request with the same client token within the idempotency window.

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)

- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

PutLifecycleHook

Creates or updates a lifecycle hook for the specified Auto Scaling group.

Lifecycle hooks let you create solutions that are aware of events in the Auto Scaling instance lifecycle, and then perform a custom action on instances when the corresponding lifecycle event occurs.

This step is a part of the procedure for adding a lifecycle hook to an Auto Scaling group:

1. (Optional) Create a launch template or launch configuration with a user data script that runs while an instance is in a wait state due to a lifecycle hook.
2. (Optional) Create a Lambda function and a rule that allows Amazon EventBridge to invoke your Lambda function when an instance is put into a wait state due to a lifecycle hook.
3. (Optional) Create a notification target and an IAM role. The target can be either an Amazon SQS queue or an Amazon SNS topic. The role allows Amazon EC2 Auto Scaling to publish lifecycle notifications to the target.
4. **Create the lifecycle hook. Specify whether the hook is used when the instances launch or terminate.**
5. If you need more time, record the lifecycle action heartbeat to keep the instance in a wait state using the [RecordLifecycleActionHeartbeat](#) API call.
6. If you finish before the timeout period ends, send a callback by using the [CompleteLifecycleAction](#) API call.

For more information, see [Amazon EC2 Auto Scaling lifecycle hooks](#) in the *Amazon EC2 Auto Scaling User Guide*.

If you exceed your maximum limit of lifecycle hooks, which by default is 50 per Auto Scaling group, the call fails.

You can view the lifecycle hooks for an Auto Scaling group using the [DescribeLifecycleHooks](#) API call. If you are no longer using a lifecycle hook, you can delete it by calling the [DeleteLifecycleHook](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\u005C\u005C]`*

Required: Yes

DefaultResult

The action the Auto Scaling group takes when the lifecycle hook timeout elapses or if an unexpected failure occurs. The default value is ABANDON.

Valid values: CONTINUE | ABANDON

Type: String

Required: No

HeartbeatTimeout

The maximum time, in seconds, that can elapse before the lifecycle hook times out. The range is from 30 to 7200 seconds. The default value is 3600 seconds (1 hour).

Type: Integer

Required: No

LifecycleHookName

The name of the lifecycle hook.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z0-9\-\_\./]`+

Required: Yes

LifecycleTransition

The lifecycle transition. For Auto Scaling groups, there are two major lifecycle transitions.

- To create a lifecycle hook for scale-out events, specify `autoscaling:EC2_INSTANCE_LAUNCHING`.
- To create a lifecycle hook for scale-in events, specify `autoscaling:EC2_INSTANCE_TERMINATING`.

Required for new lifecycle hooks, but optional when updating existing hooks.

Type: String

Required: No

NotificationMetadata

Additional information that you want to include any time Amazon EC2 Auto Scaling sends a message to the notification target.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 4000.

Pattern: `[\u0009\u000A\u000D\u0020-\u007e]+`

Required: No

NotificationTargetARN

The Amazon Resource Name (ARN) of the notification target that Amazon EC2 Auto Scaling uses to notify you when an instance is in a wait state for the lifecycle hook. You can specify either an Amazon SNS topic or an Amazon SQS queue.

If you specify an empty string, this overrides the current ARN.

This operation uses the JSON format when sending notifications to an Amazon SQS queue, and an email key-value pair format when sending notifications to an Amazon SNS topic.

When you specify a notification target, Amazon EC2 Auto Scaling sends it a test message. Test messages contain the following additional key-value pair: "Event": "autoscaling:TEST_NOTIFICATION".

Type: String

Length Constraints: Minimum length of 0. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RoleARN

The ARN of the IAM role that allows the Auto Scaling group to publish to the specified notification target.

Valid only if the notification target is an Amazon SNS topic or an Amazon SQS queue. Required for new lifecycle hooks, but optional when updating existing hooks.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of PutLifecycleHook.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=PutLifecycleHook
&LifecycleHookName=my-launch-lifecycle-hook
&HeartbeatTimeout=300
&AutoScalingGroupName=my-asg
&LifecycleTransition=autoscaling:EC2_INSTANCE_LAUNCHING
&DefaultResult=CONTINUE
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

PutNotificationConfiguration

Configures an Auto Scaling group to send notifications when specified events take place. Subscribers to the specified topic can have messages delivered to an endpoint such as a web server or an email address.

This configuration overwrites any existing configuration.

For more information, see [Amazon SNS notification options for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

If you exceed your maximum limit of SNS topics, which is 10 per Auto Scaling group, the call fails.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

NotificationTypes.member.N

The type of event that causes the notification to be sent. To query the notification types supported by Amazon EC2 Auto Scaling, call the [DescribeAutoScalingNotificationTypes](#) API.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

TopicARN

The Amazon Resource Name (ARN) of the Amazon SNS topic.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of PutNotificationConfiguration.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=PutNotificationConfiguration
&AutoScalingGroupName=my-asg
&TopicARN=arn:aws:us-east-1:123456789012:my-sns-topic
&NotificationTypes.member.1=autoscaling:EC2_INSTANCE_LAUNCH
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

PutScalingPolicy

Creates or updates a scaling policy for an Auto Scaling group. Scaling policies are used to scale an Auto Scaling group based on configurable metrics. If no policies are defined, the dynamic scaling and predictive scaling features are not used.

For more information about using dynamic scaling, see [Target tracking scaling policies](#) and [Step and simple scaling policies](#) in the *Amazon EC2 Auto Scaling User Guide*.

For more information about using predictive scaling, see [Predictive scaling for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

You can view the scaling policies for an Auto Scaling group using the [DescribePolicies](#) API call. If you are no longer using a scaling policy, you can delete it by calling the [DeletePolicy](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AdjustmentType

Specifies how the scaling adjustment is interpreted (for example, an absolute number or a percentage). The valid values are `ChangeInCapacity`, `ExactCapacity`, and `PercentChangeInCapacity`.

Required if the policy type is `StepScaling` or `SimpleScaling`. For more information, see [Scaling adjustment types](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Cooldown

A cooldown period, in seconds, that applies to a specific simple scaling policy. When a cooldown period is specified here, it overrides the default cooldown.

Valid only if the policy type is `SimpleScaling`. For more information, see [Scaling cooldowns for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Default: None

Type: Integer

Required: No

Enabled

Indicates whether the scaling policy is enabled or disabled. The default is enabled. For more information, see [Disable a scaling policy for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

EstimatedInstanceWarmup

Not needed if the default instance warmup is defined for the group.

The estimated time, in seconds, until a newly launched instance can contribute to the CloudWatch metrics. This warm-up period applies to instances launched due to a specific target tracking or step scaling policy. When a warm-up period is specified here, it overrides the default instance warmup.

Valid only if the policy type is `TargetTrackingScaling` or `StepScaling`.

Note

The default is to use the value for the default instance warmup defined for the group. If default instance warmup is null, then `EstimatedInstanceWarmup` falls back to the value of default cooldown.

Type: Integer

Required: No

MetricAggregationType

The aggregation type for the CloudWatch metrics. The valid values are Minimum, Maximum, and Average. If the aggregation type is null, the value is treated as Average.

Valid only if the policy type is StepScaling.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MinAdjustmentMagnitude

The minimum value to scale by when the adjustment type is PercentChangeInCapacity. For example, suppose that you create a step scaling policy to scale out an Auto Scaling group by 25 percent and you specify a MinAdjustmentMagnitude of 2. If the group has 4 instances and the scaling policy is performed, 25 percent of 4 is 1. However, because you specified a MinAdjustmentMagnitude of 2, Amazon EC2 Auto Scaling scales out the group by 2 instances.

Valid only if the policy type is StepScaling or SimpleScaling. For more information, see [Scaling adjustment types](#) in the *Amazon EC2 Auto Scaling User Guide*.

Note

Some Auto Scaling groups use instance weights. In this case, set the MinAdjustmentMagnitude to a value that is at least as large as your largest instance weight.

Type: Integer

Required: No

MinAdjustmentStep

This parameter has been deprecated.

Available for backward compatibility. Use MinAdjustmentMagnitude instead.

Type: Integer

Required: No

PolicyName

The name of the policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

PolicyType

One of the following policy types:

- TargetTrackingScaling
- StepScaling
- SimpleScaling (default)
- PredictiveScaling

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

PredictiveScalingConfiguration

A predictive scaling policy. Provides support for predefined and custom metrics.

Predefined metrics include CPU utilization, network in/out, and the Application Load Balancer request count.

Required if the policy type is `PredictiveScaling`.

Type: [PredictiveScalingConfiguration](#) object

Required: No

ScalingAdjustment

The amount by which to scale, based on the specified adjustment type. A positive value adds to the current capacity while a negative number removes from the current capacity. For exact capacity, you must specify a non-negative value.

Required if the policy type is `SimpleScaling`. (Not used with any other policy type.)

Type: Integer

Required: No

StepAdjustments.member.N

A set of adjustments that enable you to scale based on the size of the alarm breach.

Required if the policy type is `StepScaling`. (Not used with any other policy type.)

Type: Array of [StepAdjustment](#) objects

Required: No

TargetTrackingConfiguration

A target tracking scaling policy. Provides support for predefined or custom metrics.

The following predefined metrics are available:

- `ASGAverageCPUUtilization`
- `ASGAverageNetworkIn`
- `ASGAverageNetworkOut`
- `ALBRequestCountPerTarget`

If you specify `ALBRequestCountPerTarget` for the metric, you must specify the `ResourceLabel` property with the `PredefinedMetricSpecification`.

Required if the policy type is `TargetTrackingScaling`.

Type: [TargetTrackingConfiguration](#) object

Required: No

Response Elements

The following elements are returned by the service.

Alarms.member.N

The CloudWatch alarms created for the target tracking scaling policy.

Type: Array of [Alarm](#) objects

PolicyARN

The Amazon Resource Name (ARN) of the policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of PutScalingPolicy.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=PutScalingPolicy
&AutoScalingGroupName=my-asg
&PolicyName=alb1000-target-tracking-scaling-policy
&PolicyType=TargetTrackingScaling
&TargetTrackingConfiguration.TargetValue=1000.0
&TargetTrackingConfiguration.PredefinedMetricSpecification.PredefinedMetricType=ALBRequestCountPerTarget
&TargetTrackingConfiguration.PredefinedMetricSpecification.ResourceLabel=app%2Fmy-alb
%2F778d41231b141a0f%2Ftargetgroup%2Fmy-alb-target-group%2F943f017f100becff
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<PutScalingPolicyResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <PutScalingPolicyResult>
    <PolicyARN>arn:aws:autoscaling:us-east-1:123456789012:scalingPolicy:228f02c2-
c665-4bfd-aaac-8b04080bea3c:autoScalingGroupName/my-asg:policyName/alb1000-target-
tracking-scaling-policy</PolicyARN>
    <Alarms>
      <member>
        <AlarmName>TargetTracking-my-asg-AlarmHigh-
fc0e4183-23ac-497e-9992-691c9980c38e</AlarmName>
        <AlarmARN>arn:aws:cloudwatch:us-east-1:123456789012:alarm:TargetTracking-my-
asg-AlarmHigh-fc0e4183-23ac-497e-9992-691c9980c38e</AlarmARN>
      </member>
      <member>
```

```
<AlarmName>TargetTracking-my-asg-AlarmLow-61a39305-ed0c-47af-  
bd9e-471a352ee1a2</AlarmName>  
  <AlarmARN>arn:aws:cloudwatch:us-east-1:123456789012:alarm:TargetTracking-my-  
asg-AlarmLow-61a39305-ed0c-47af-bd9e-471a352ee1a2</AlarmARN>  
  </member>  
</Alarms>  
</PutScalingPolicyResult>  
<ResponseMetadata>  
  <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>  
</ResponseMetadata>  
</PutScalingPolicyResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

PutScheduledUpdateGroupAction

Creates or updates a scheduled scaling action for an Auto Scaling group.

For more information, see [Scheduled scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

You can view the scheduled actions for an Auto Scaling group using the [DescribeScheduledActions](#) API call. If you are no longer using a scheduled action, you can delete it by calling the [DeleteScheduledAction](#) API.

If you try to schedule your action in the past, Amazon EC2 Auto Scaling returns an error message.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

DesiredCapacity

The desired capacity is the initial capacity of the Auto Scaling group after the scheduled action runs and the capacity it attempts to maintain. It can scale beyond this capacity if you add more scaling conditions.

Note

You must specify at least one of the following properties: MaxSize, MinSize, or DesiredCapacity.

Type: Integer

Required: No

EndTime

The date and time for the recurring schedule to end, in UTC. For example, "2021-06-01T00:00:00Z".

Type: Timestamp

Required: No

MaxSize

The maximum size of the Auto Scaling group.

Type: Integer

Required: No

MinSize

The minimum size of the Auto Scaling group.

Type: Integer

Required: No

Recurrence

The recurring schedule for this action. This format consists of five fields separated by white spaces: [Minute] [Hour] [Day_of_Month] [Month_of_Year] [Day_of_Week]. The value must be in quotes (for example, "30 0 1 1,6,12 *"). For more information about this format, see [Crontab](#).

When StartTime and EndTime are specified with Recurrence, they form the boundaries of when the recurring action starts and stops.

Cron expressions use Universal Coordinated Time (UTC) by default.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFF\\uD800\\uDC00-\\uDBFF\\uDFFF\\r\\n\\t]*`

Required: No

ScheduledActionName

The name of this scaling action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

StartTime

The date and time for this action to start, in YYYY-MM-DDThh:mm:ssZ format in UTC/GMT only and in quotes (for example, "2021-06-01T00:00:00Z").

If you specify Recurrence and StartTime, Amazon EC2 Auto Scaling performs the action at this time, and then performs the action based on the specified recurrence.

Type: Timestamp

Required: No

Time

This property is no longer used.

Type: Timestamp

Required: No

TimeZone

Specifies the time zone for a cron expression. If a time zone is not provided, UTC is used by default.

Valid values are the canonical names of the IANA time zones, derived from the IANA Time Zone Database (such as Etc/GMT+9 or Pacific/Tahiti). For more information, see https://en.wikipedia.org/wiki/List_of_tz_database_time_zones.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

AlreadyExists

You already have an Auto Scaling group or launch configuration with this name.

message

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example 1: Schedule based on a specific date and time

This example illustrates one usage of `PutScheduledUpdateGroupAction`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=PutScheduledUpdateGroupAction
&AutoScalingGroupName=my-asg
&ScheduledActionName=scale-out
&StartTime=2020-05-25T08:00:00Z
&DesiredCapacity=3
&Version=2011-01-01
&AUTHPARAMS
```

Example 2: Recurring Schedule

This example illustrates one usage of PutScheduledUpdateGroupAction.

Sample Request

```
https://autoscaling.amazonaws.com/?Action="PutScheduledUpdateGroupAction
&AutoScalingGroupName=my-asg
&ScheduledActionName=scale-out-schedule-year
&Recurrence="30 0 1 1,6,12 *"
&DesiredCapacity=3
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)

- [AWS SDK for Ruby V3](#)

PutWarmPool

Creates or updates a warm pool for the specified Auto Scaling group. A warm pool is a pool of pre-initialized EC2 instances that sits alongside the Auto Scaling group. Whenever your application needs to scale out, the Auto Scaling group can draw on the warm pool to meet its new desired capacity.

This operation must be called from the Region in which the Auto Scaling group was created.

You can view the instances in the warm pool using the [DescribeWarmPool](#) API call. If you are no longer using a warm pool, you can delete it by calling the [DeleteWarmPool](#) API.

For more information, see [Warm pools for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

InstanceReusePolicy

Indicates whether instances in the Auto Scaling group can be returned to the warm pool on scale in. The default is to terminate instances in the Auto Scaling group when the group scales in.

Type: [InstanceReusePolicy](#) object

Required: No

MaxGroupPreparedCapacity

Specifies the maximum number of instances that are allowed to be in the warm pool or in any state except Terminated for the Auto Scaling group. This is an optional property. Specify it

only if you do not want the warm pool size to be determined by the difference between the group's maximum capacity and its desired capacity.

⚠ Important

If a value for `MaxGroupPreparedCapacity` is not specified, Amazon EC2 Auto Scaling launches and maintains the difference between the group's maximum capacity and its desired capacity. If you specify a value for `MaxGroupPreparedCapacity`, Amazon EC2 Auto Scaling uses the difference between the `MaxGroupPreparedCapacity` and the desired capacity instead.

The size of the warm pool is dynamic. Only when `MaxGroupPreparedCapacity` and `MinSize` are set to the same value does the warm pool have an absolute size.

If the desired capacity of the Auto Scaling group is higher than the `MaxGroupPreparedCapacity`, the capacity of the warm pool is 0, unless you specify a value for `MinSize`. To remove a value that you previously set, include the property but specify -1 for the value.

Type: Integer

Valid Range: Minimum value of -1.

Required: No

MinSize

Specifies the minimum number of instances to maintain in the warm pool. This helps you to ensure that there is always a certain number of warmed instances available to handle traffic spikes. Defaults to 0 if not specified.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PoolState

Sets the instance state to transition to after the lifecycle actions are complete. Default is Stopped.

Type: String

Valid Values: Stopped | Running | Hibernated

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InstanceRefreshInProgress

The request failed because an active instance refresh already exists for the specified Auto Scaling group.

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of PutWarmPool.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=PutWarmPool
&AutoScalingGroupName=my-asg
&MinSize=30
&PoolState=Hibernated
&InstanceReusePolicy.ReuseOnScaleIn=true
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

RecordLifecycleActionHeartbeat

Records a heartbeat for the lifecycle action associated with the specified token or instance. This extends the timeout by the length of time defined using the [PutLifecycleHook](#) API call.

This step is a part of the procedure for adding a lifecycle hook to an Auto Scaling group:

1. (Optional) Create a launch template or launch configuration with a user data script that runs while an instance is in a wait state due to a lifecycle hook.
2. (Optional) Create a Lambda function and a rule that allows Amazon EventBridge to invoke your Lambda function when an instance is put into a wait state due to a lifecycle hook.
3. (Optional) Create a notification target and an IAM role. The target can be either an Amazon SQS queue or an Amazon SNS topic. The role allows Amazon EC2 Auto Scaling to publish lifecycle notifications to the target.
4. Create the lifecycle hook. Specify whether the hook is used when the instances launch or terminate.
5. **If you need more time, record the lifecycle action heartbeat to keep the instance in a wait state.**
6. If you finish before the timeout period ends, send a callback by using the [CompleteLifecycleAction](#) API call.

For more information, see [Amazon EC2 Auto Scaling lifecycle hooks](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

InstanceId

The ID of the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LifecycleActionToken

A token that uniquely identifies a specific lifecycle action associated with an instance. Amazon EC2 Auto Scaling sends this token to the notification target that you specified when you created the lifecycle hook.

Type: String

Length Constraints: Fixed length of 36.

Required: No

LifecycleHookName

The name of the lifecycle hook.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z0-9\-\_\//]+`

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

ResumeProcesses

Resumes the specified suspended auto scaling processes, or all suspended process, for the specified Auto Scaling group.

For more information, see [Suspend and resume Amazon EC2 Auto Scaling processes](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ScalingProcesses.member.N

One or more of the following processes:

- Launch
- Terminate
- AddToLoadBalancer
- AlarmNotification
- AZRebalance
- HealthCheck
- InstanceRefresh
- ReplaceUnhealthy
- ScheduledActions

If you omit this property, all processes are specified.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ResourceInUse

The operation can't be performed because the resource is in use.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of `ResumeProcesses`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=ResumeProcesses
&AutoScalingGroupName=my-asg
&ScalingProcesses.member.1=AlarmNotification
&Version=2011-01-01
```

&AUTHPARAMS

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

InstanceRefreshId

The instance refresh ID associated with the request. This is the unique ID assigned to the instance refresh when it was started.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ActiveInstanceRefreshNotFound

The request failed because an active instance refresh or rollback for the specified Auto Scaling group was not found.

HTTP Status Code: 400

IrreversibleInstanceRefresh

The request failed because a desired configuration was not found or an incompatible launch template (uses a Systems Manager parameter instead of an AMI ID) or launch template version (\$Latest or \$Default) is present on the Auto Scaling group.

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

SetDesiredCapacity

Sets the size of the specified Auto Scaling group.

If a scale-in activity occurs as a result of a new `DesiredCapacity` value that is lower than the current size of the group, the Auto Scaling group uses its termination policy to determine which instances to terminate.

For more information, see [Manual scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF\\r\\n\\t]*`

Required: Yes

DesiredCapacity

The desired capacity is the initial capacity of the Auto Scaling group after this operation completes and the capacity it attempts to maintain.

Type: Integer

Required: Yes

HonorCooldown

Indicates whether Amazon EC2 Auto Scaling waits for the cooldown period to complete before initiating a scaling activity to set your Auto Scaling group to its new capacity. By default, Amazon EC2 Auto Scaling does not honor the cooldown period during manual scaling activities.

Type: Boolean

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ScalingActivityInProgress

The operation can't be performed because there are scaling activities in progress.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of SetDesiredCapacity.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=SetDesiredCapacity
&AutoScalingGroupName=my-asg
&HonorCooldown=false
&DesiredCapacity=2
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

SetInstanceHealth

Sets the health status of the specified instance.

For more information, see [Set up a custom health check for your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

HealthStatus

The health status of the instance. Set to `Healthy` to have the instance remain in service. Set to `Unhealthy` to have the instance be out of service. Amazon EC2 Auto Scaling terminates and replaces the unhealthy instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceId

The ID of the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ShouldRespectGracePeriod

If the Auto Scaling group of the specified instance has a `HealthCheckGracePeriod` specified for the group, by default, this call respects the grace period. Set this to `False`, to have the call not respect the grace period associated with the group.

For more information about the health check grace period, see [Set the health check grace period for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of `SetInstanceHealth`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=SetInstanceHealth
&InstanceId=i-1234567890abcdef0
&HealthStatus=Unhealthy
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)

- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

SetInstanceProtection

Updates the instance protection settings of the specified instances. This operation cannot be called on instances in a warm pool.

For more information, see [Use instance scale-in protection](#) in the *Amazon EC2 Auto Scaling User Guide*.

If you exceed your maximum limit of instance IDs, which is 50 per Auto Scaling group, the call fails.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

InstanceIds.member.N

One or more instance IDs. You can specify up to 50 instances.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

ProtectedFromScaleIn

Indicates whether the instance is protected from termination by Amazon EC2 Auto Scaling when scaling in.

Type: Boolean

Required: Yes

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of SetInstanceProtection.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=SetInstanceProtection
&AutoScalingGroupName=my-asg
&InstanceIds.member.1=i-1234567890abcdef0
&ProtectedFromScaleIn=false
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

StartInstanceRefresh

Starts an instance refresh.

This operation is part of the [instance refresh feature](#) in Amazon EC2 Auto Scaling, which helps you update instances in your Auto Scaling group. This feature is helpful, for example, when you have a new AMI or a new user data script. You just need to create a new launch template that specifies the new AMI or user data script. Then start an instance refresh to immediately begin the process of updating instances in the group.

If successful, the request's response contains a unique ID that you can use to track the progress of the instance refresh. To query its status, call the [DescribeInstanceRefreshes](#) API. To describe the instance refreshes that have already run, call the [DescribeInstanceRefreshes](#) API. To cancel an instance refresh that is in progress, use the [CancelInstanceRefresh](#) API.

An instance refresh might fail for several reasons, such as EC2 launch failures, misconfigured health checks, or not ignoring or allowing the termination of instances that are in Standby state or protected from scale in. You can monitor for failed EC2 launches using the scaling activities. To find the scaling activities, call the [DescribeScalingActivities](#) API.

If you enable auto rollback, your Auto Scaling group will be rolled back automatically when the instance refresh fails. You can enable this feature before starting an instance refresh by specifying the `AutoRollback` property in the instance refresh preferences. Otherwise, to roll back an instance refresh before it finishes, use the [RollbackInstanceRefresh](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

DesiredConfiguration

The desired configuration. For example, the desired configuration can specify a new launch template or a new version of the current launch template.

Once the instance refresh succeeds, Amazon EC2 Auto Scaling updates the settings of the Auto Scaling group to reflect the new desired configuration.

Note

When you specify a new launch template or a new version of the current launch template for your desired configuration, consider enabling the `SkipMatching` property in preferences. If it's enabled, Amazon EC2 Auto Scaling skips replacing instances that already use the specified launch template and instance types. This can help you reduce the number of replacements that are required to apply updates.

Type: [DesiredConfiguration](#) object

Required: No

Preferences

Sets your preferences for the instance refresh so that it performs as expected when you start it. Includes the instance warmup time, the minimum and maximum healthy percentages, and the behaviors that you want Amazon EC2 Auto Scaling to use if instances that are in `Standby` state or protected from scale in are found. You can also choose to enable additional features, such as the following:

- Auto rollback
- Checkpoints
- CloudWatch alarms
- Skip matching
- Bake time

Type: [RefreshPreferences](#) object

Required: No

Strategy

The strategy to use for the instance refresh. The default value is `Rolling`.

Type: String

Valid Values: Rolling | ReplaceRootVolume

Required: No

Response Elements

The following element is returned by the service.

InstanceRefreshId

A unique ID for tracking the progress of the instance refresh.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

InstanceRefreshInProgress

The request failed because an active instance refresh already exists for the specified Auto Scaling group.

HTTP Status Code: 400

LimitExceeded

You have already reached a limit for your Amazon EC2 Auto Scaling resources (for example, Auto Scaling groups, launch configurations, or lifecycle hooks). For more information, see [DescribeAccountLimits](#).

message

HTTP Status Code: 400

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

Examples

Example

This example illustrates one usage of StartInstanceRefresh.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=StartInstanceRefresh
&AutoScalingGroupName=my-asg
&Preferences.AutoRollback=true
&Preferences.InstanceWarmup=200
&Preferences.MinHealthyPercentage=90
&Preferences.MaxHealthyPercentage=120
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<StartInstanceRefreshResponse xmlns="https://autoscaling.amazonaws.com/
doc/2011-01-01/">
  <StartInstanceRefreshResult>
    <InstanceRefreshId>ccfd3c2f-edb3-470d-af32-52cc57d201ca</InstanceRefreshId>
  </StartInstanceRefreshResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</StartInstanceRefreshResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

SuspendProcesses

Suspends the specified auto scaling processes, or all processes, for the specified Auto Scaling group.

If you suspend either the Launch or Terminate process types, it can prevent other process types from functioning properly. For more information, see [Suspend and resume Amazon EC2 Auto Scaling processes](#) in the *Amazon EC2 Auto Scaling User Guide*.

To resume processes that have been suspended, call the [ResumeProcesses](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ScalingProcesses.member.N

One or more of the following processes:

- Launch
- Terminate
- AddToLoadBalancer
- AlarmNotification
- AZRebalance
- HealthCheck
- InstanceRefresh
- ReplaceUnhealthy
- ScheduledActions

If you omit this property, all processes are specified.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ResourceInUse

The operation can't be performed because the resource is in use.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of `SuspendProcesses`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=SuspendProcesses
&AutoScalingGroupName=my-asg
```

```
&ScalingProcesses.member.1=AlarmNotification  
&Version=2011-01-01  
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

TerminateInstanceInAutoScalingGroup

Terminates the specified instance and optionally adjusts the desired group size. This operation cannot be called on instances in a warm pool.

This call simply makes a termination request. The instance is not terminated immediately. When an instance is terminated, the instance status changes to `terminated`. You can't connect to or start an instance after you've terminated it.

If you do not specify the option to decrement the desired capacity, Amazon EC2 Auto Scaling launches instances to replace the ones that are terminated.

By default, Amazon EC2 Auto Scaling balances instances across all Availability Zones. If you decrement the desired capacity, your Auto Scaling group can become unbalanced between Availability Zones. Amazon EC2 Auto Scaling tries to rebalance the group, and rebalancing might terminate instances in other zones. For more information, see [Manual scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

InstanceId

The ID of the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: Yes

ShouldDecrementDesiredCapacity

Indicates whether terminating the instance also decrements the size of the Auto Scaling group.

Type: Boolean

Required: Yes

Response Elements

The following element is returned by the service.

Activity

A scaling activity.

Type: [Activity](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ScalingActivityInProgress

The operation can't be performed because there are scaling activities in progress.

message

HTTP Status Code: 400

Examples

Example

This example illustrates one usage of `TerminateInstanceInAutoScalingGroup`.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=TerminateInstanceInAutoScalingGroup
```

```
&InstanceId=i-1234567890abcdef0
&ShouldDecrementDesiredCapacity=true
&Version=2011-01-01
&AUTHPARAMS
```

Sample Response

```
<TerminateInstanceInAutoScalingGroupResponse xmlns="https://autoscaling.amazonaws.com/doc/2011-01-01/">
  <TerminateInstanceInAutoScalingGroupResult>
    <Activity>
      <AutoScalingGroupName>my-asg</AutoScalingGroupName>
      <ActivityId>12345678-1234-1234-1234-123456789012</ActivityId>
      <Description>Terminating EC2 instance: i-1234567890abcdef0</Description>
      <Cause>At 2022-06-14T00:07:30Z instance i-1234567890abcdef0 was taken out of
service in response to a user request, shrinking the capacity from 4 to 3.</Cause>
      <StartTime>2022-06-14T00:07:30.280Z</StartTime>
      <Details>{"Availability Zone":"us-
east-1a","SubnetID":"subnet-057fa0918fEXAMPLE"}</Details>
      <Progress>0</Progress>
      <StatusCode>InProgress</StatusCode>
    </Activity>
  </TerminateInstanceInAutoScalingGroupResult>
  <ResponseMetadata>
    <RequestId>7c6e177f-f082-11e1-ac58-3714bEXAMPLE</RequestId>
  </ResponseMetadata>
</TerminateInstanceInAutoScalingGroupResponse>
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)

- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

UpdateAutoScalingGroup

We strongly recommend that all Auto Scaling groups use launch templates to ensure full functionality for Amazon EC2 Auto Scaling and Amazon EC2.

Updates the configuration for the specified Auto Scaling group.

To update an Auto Scaling group, specify the name of the group and the property that you want to change. Any properties that you don't specify are not changed by this update request. The new settings take effect on any scaling activities after this call returns.

If you associate a new launch configuration or template with an Auto Scaling group, all new instances will get the updated configuration. Existing instances continue to run with the configuration that they were originally launched with. When you update a group to specify a mixed instances policy instead of a launch configuration or template, existing instances may be replaced to match the new purchasing options that you specified in the policy. For example, if the group currently has 100% On-Demand capacity and the policy specifies 50% Spot capacity, this means that half of your instances will be gradually terminated and relaunched as Spot Instances. When replacing instances, Amazon EC2 Auto Scaling launches new instances before terminating the old ones, so that updating your group does not compromise the performance or availability of your application.

Note the following about changing `DesiredCapacity`, `MaxSize`, or `MinSize`:

- If a scale-in activity occurs as a result of a new `DesiredCapacity` value that is lower than the current size of the group, the Auto Scaling group uses its termination policy to determine which instances to terminate.
- If you specify a new value for `MinSize` without specifying a value for `DesiredCapacity`, and the new `MinSize` is larger than the current size of the group, this sets the group's `DesiredCapacity` to the new `MinSize` value.
- If you specify a new value for `MaxSize` without specifying a value for `DesiredCapacity`, and the new `MaxSize` is smaller than the current size of the group, this sets the group's `DesiredCapacity` to the new `MaxSize` value.

To see which properties have been set, call the [DescribeAutoScalingGroups](#) API. To view the scaling policies for an Auto Scaling group, call the [DescribePolicies](#) API. If the group has scaling policies, you can update them by calling the [PutScalingPolicy](#) API.

Request Parameters

For information about the parameters that are common to all actions, see [Common Parameters](#).

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

AvailabilityZoneDistribution

The instance capacity distribution across Availability Zones.

Type: [AvailabilityZoneDistribution](#) object

Required: No

AvailabilityZoneImpairmentPolicy

The policy for Availability Zone impairment.

Type: [AvailabilityZoneImpairmentPolicy](#) object

Required: No

AvailabilityZones.member.N

One or more Availability Zones for the group.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

CapacityRebalance

Enables or disables Capacity Rebalancing. If Capacity Rebalancing is disabled, proactive replacement of at-risk Spot Instances does not occur. For more information, see [Capacity Rebalancing in Auto Scaling to replace at-risk Spot Instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Note

To suspend rebalancing across Availability Zones, use the [SuspendProcesses](#) API.

Type: Boolean

Required: No

CapacityReservationSpecification

The capacity reservation specification for the Auto Scaling group.

Type: [CapacityReservationSpecification](#) object

Required: No

Context

Reserved.

Type: String

Required: No

DefaultCooldown

Only needed if you use simple scaling policies.

The amount of time, in seconds, between one scaling activity ending and another one starting due to simple scaling policies. For more information, see [Scaling cooldowns for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Integer

Required: No

DefaultInstanceWarmup

The amount of time, in seconds, until a new instance is considered to have finished initializing and resource consumption to become stable after it enters the InService state.

During an instance refresh, Amazon EC2 Auto Scaling waits for the warm-up period after it replaces an instance before it moves on to replacing the next instance. Amazon EC2 Auto Scaling also waits for the warm-up period before aggregating the metrics for new instances with existing instances in the Amazon CloudWatch metrics that are used for scaling, resulting in more reliable usage data. For more information, see [Set the default instance warmup for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Important

To manage various warm-up settings at the group level, we recommend that you set the default instance warmup, *even if it is set to 0 seconds*. To remove a value that you previously set, include the property but specify -1 for the value. However, we strongly recommend keeping the default instance warmup enabled by specifying a value of 0 or other nominal value.

Type: Integer

Required: No

DesiredCapacity

The desired capacity is the initial capacity of the Auto Scaling group after this operation completes and the capacity it attempts to maintain. This number must be greater than or equal to the minimum size of the group and less than or equal to the maximum size of the group.

Type: Integer

Required: No

DesiredCapacityType

The unit of measurement for the value specified for desired capacity. Amazon EC2 Auto Scaling supports DesiredCapacityType for attribute-based instance type selection only. For more information, see [Create a mixed instances group using attribute-based instance type selection](#) in the *Amazon EC2 Auto Scaling User Guide*.

By default, Amazon EC2 Auto Scaling specifies `units`, which translates into number of instances.

Valid values: `units` | `vcpu` | `memory-mib`

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

HealthCheckGracePeriod

The amount of time, in seconds, that Amazon EC2 Auto Scaling waits before checking the health status of an EC2 instance that has come into service and marking it unhealthy due to a failed health check. This is useful if your instances do not immediately pass their health checks after they enter the `InService` state. For more information, see [Set the health check grace period for an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Integer

Required: No

HealthCheckType

A comma-separated value string of one or more health check types.

The valid values are `EC2`, `EBS`, `ELB`, and `VPC_LATTICE`. `EC2` is the default health check and cannot be disabled. For more information, see [Health checks for instances in an Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Only specify `EC2` if you must clear a value that was previously set.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceLifecyclePolicy

The instance lifecycle policy for the Auto Scaling group. This policy controls instance behavior when an instance transitions through its lifecycle states. Configure retention triggers to specify when instances should move to a Retained state instead of automatic termination.

For more information, see [Control instance retention with instance lifecycle policies](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [InstanceLifecyclePolicy](#) object

Required: No

InstanceMaintenancePolicy

An instance maintenance policy. For more information, see [Set instance maintenance policy](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [InstanceMaintenancePolicy](#) object

Required: No

LaunchConfigurationName

The name of the launch configuration. If you specify LaunchConfigurationName in your update request, you can't specify LaunchTemplate or MixedInstancesPolicy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchTemplate

The launch template and version to use to specify the updates. If you specify LaunchTemplate in your update request, you can't specify LaunchConfigurationName or MixedInstancesPolicy.

Type: [LaunchTemplateSpecification](#) object

Required: No

MaxInstanceLifetime

The maximum amount of time, in seconds, that an instance can be in service. The default is null. If specified, the value must be either 0 or a number equal to or greater than 86,400 seconds (1 day). To clear a previously set value, specify a new value of 0. For more information, see [Replacing Auto Scaling instances based on maximum instance lifetime](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Integer

Required: No

MaxSize

The maximum size of the Auto Scaling group.

Note

With a mixed instances policy that uses instance weighting, Amazon EC2 Auto Scaling may need to go above MaxSize to meet your capacity requirements. In this event, Amazon EC2 Auto Scaling will never go above MaxSize by more than your largest instance weight (weights that define how many units each instance contributes to the desired capacity of the group).

Type: Integer

Required: No

MinSize

The minimum size of the Auto Scaling group.

Type: Integer

Required: No

MixedInstancesPolicy

The mixed instances policy. For more information, see [Auto Scaling groups with multiple instance types and purchase options](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [MixedInstancesPolicy](#) object

Errors

For information about the errors that are common to all actions, see [Common Errors](#).

ResourceContention

You already have a pending update to an Amazon EC2 Auto Scaling resource (for example, an Auto Scaling group, instance, or load balancer).

message

HTTP Status Code: 500

ScalingActivityInProgress

The operation can't be performed because there are scaling activities in progress.

message

HTTP Status Code: 400

ServiceLinkedRoleFailure

The service-linked role is not yet ready for use.

HTTP Status Code: 500

Examples

Example 1: Update the health check settings of an existing Auto Scaling group

This example updates the health check settings for the Auto Scaling group my-asg.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=UpdateAutoScalingGroup
&AutoScalingGroupName=my-asg
&HealthCheckType=ELB
&HealthCheckGracePeriod=300
&Version=2011-01-01
&AUTHPARAMS
```

Example 2: Update the VPC subnets of an existing Auto Scaling group

This example updates the VPC subnets for the Auto Scaling group my-asg.

Sample Request

```
https://autoscaling.amazonaws.com/?Action=UpdateAutoScalingGroup
&AutoScalingGroupName=my-asg
&VPCZoneIdentifier=subnet-057fa0918fEXAMPLE%2Csubnet-610acd08EXAMPLE
%2Csubnet-530fc83aEXAMPLE
&Version=2011-01-01
&AUTHPARAMS
```

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS Command Line Interface V2](#)
- [AWS SDK for .NET](#)
- [AWS SDK for C++](#)
- [AWS SDK for Go v2](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for JavaScript V3](#)
- [AWS SDK for Kotlin](#)
- [AWS SDK for PHP V3](#)
- [AWS SDK for Python](#)
- [AWS SDK for Ruby V3](#)

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 a data type structure is not guaranteed. Applications should not assume a particular order.

The following data types are supported:

- [AcceleratorCountRequest](#)
- [AcceleratorTotalMemoryMiBRequest](#)
- [Activity](#)
- [AdjustmentType](#)
- [Alarm](#)
- [AlarmSpecification](#)
- [AutoScalingGroup](#)
- [AutoScalingInstanceDetails](#)
- [AvailabilityZoneDistribution](#)
- [AvailabilityZoneImpairmentPolicy](#)
- [BaselineEbsBandwidthMbpsRequest](#)
- [BaselinePerformanceFactorsRequest](#)
- [BlockDeviceMapping](#)
- [CapacityForecast](#)
- [CapacityReservationSpecification](#)
- [CapacityReservationTarget](#)
- [CpuPerformanceFactorRequest](#)
- [CustomizedMetricSpecification](#)
- [DesiredConfiguration](#)
- [Ebs](#)

- [EnabledMetric](#)
- [FailedScheduledUpdateGroupActionRequest](#)
- [Filter](#)
- [Instance](#)
- [InstanceCollection](#)
- [InstanceLifecyclePolicy](#)
- [InstanceMaintenancePolicy](#)
- [InstanceMetadataOptions](#)
- [InstanceMonitoring](#)
- [InstanceRefresh](#)
- [InstanceRefreshLivePoolProgress](#)
- [InstanceRefreshProgressDetails](#)
- [InstanceRefreshWarmPoolProgress](#)
- [InstanceRequirements](#)
- [InstanceReusePolicy](#)
- [InstancesDistribution](#)
- [LaunchConfiguration](#)
- [LaunchInstancesError](#)
- [LaunchTemplate](#)
- [LaunchTemplateOverrides](#)
- [LaunchTemplateSpecification](#)
- [LifecycleHook](#)
- [LifecycleHookSpecification](#)
- [LoadBalancerState](#)
- [LoadBalancerTargetGroupState](#)
- [LoadForecast](#)
- [MemoryGiBPerVCpuRequest](#)
- [MemoryMiBRequest](#)
- [Metric](#)
- [MetricCollectionType](#)

- [MetricDataQuery](#)
- [MetricDimension](#)
- [MetricGranularityType](#)
- [MetricStat](#)
- [MixedInstancesPolicy](#)
- [NetworkBandwidthGbpsRequest](#)
- [NetworkInterfaceCountRequest](#)
- [NotificationConfiguration](#)
- [PerformanceFactorReferenceRequest](#)
- [PredefinedMetricSpecification](#)
- [PredictiveScalingConfiguration](#)
- [PredictiveScalingCustomizedCapacityMetric](#)
- [PredictiveScalingCustomizedLoadMetric](#)
- [PredictiveScalingCustomizedScalingMetric](#)
- [PredictiveScalingMetricSpecification](#)
- [PredictiveScalingPredefinedLoadMetric](#)
- [PredictiveScalingPredefinedMetricPair](#)
- [PredictiveScalingPredefinedScalingMetric](#)
- [ProcessType](#)
- [RefreshPreferences](#)
- [RetentionTriggers](#)
- [RollbackDetails](#)
- [ScalingPolicy](#)
- [ScheduledUpdateGroupAction](#)
- [ScheduledUpdateGroupActionRequest](#)
- [StepAdjustment](#)
- [SuspendedProcess](#)
- [Tag](#)
- [TagDescription](#)
- [TargetTrackingConfiguration](#)

- [TargetTrackingMetricDataQuery](#)
- [TargetTrackingMetricStat](#)
- [TotalLocalStorageGBRequest](#)
- [TrafficSourceIdentifier](#)
- [TrafficSourceState](#)
- [VCpuCountRequest](#)
- [WarmPoolConfiguration](#)

AcceleratorCountRequest

Specifies the minimum and maximum for the AcceleratorCount object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Max

The maximum value.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Min

The minimum value.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AcceleratorTotalMemoryMiBRequest

Specifies the minimum and maximum for the AcceleratorTotalMemoryMiB object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Max

The memory maximum in MiB.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Min

The memory minimum in MiB.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Activity

Describes scaling activity, which is a long-running process that represents a change to your Auto Scaling group, such as changing its size or replacing an instance.

Contents

ActivityId

The ID of the activity.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Cause

The reason the activity began.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

StartTime

The start time of the activity.

Type: Timestamp

Required: Yes

StatusCode

The current status of the activity.

Type: String

Valid Values: PendingSpotBidPlacement | WaitingForSpotInstanceRequestId | WaitingForSpotInstanceId | WaitingForInstanceId | PreInService | InProgress | WaitingForELBConnectionDraining | MidLifecycleAction | WaitingForInstanceWarmup | Successful | Failed | Cancelled | WaitingForConnectionDraining | WaitingForInPlaceUpdateToStart | WaitingForInPlaceUpdateToFinalize | InPlaceUpdateInProgress

Required: Yes

AutoScalingGroupARN

The Amazon Resource Name (ARN) of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

AutoScalingGroupState

The state of the Auto Scaling group, which is either InService or Deleted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Required: No

Description

A friendly, more verbose description of the activity.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Details

The details about the activity.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EndTime

The end time of the activity.

Type: Timestamp

Required: No

Progress

A value between 0 and 100 that indicates the progress of the activity.

Type: Integer

Required: No

StatusMessage

A friendly, more verbose description of the activity status.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AdjustmentType

Describes a policy adjustment type.

Contents

AdjustmentType

The policy adjustment type. The valid values are `ChangeInCapacity`, `ExactCapacity`, and `PercentChangeInCapacity`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Alarm

Describes an alarm.

Contents

AlarmARN

The Amazon Resource Name (ARN) of the alarm.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

AlarmName

The name of the alarm.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AlarmSpecification

Specifies the CloudWatch alarm specification to use in an instance refresh.

Contents

Alarms.member.N

The names of one or more CloudWatch alarms to monitor for the instance refresh. You can specify up to 10 alarms.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AutoScalingGroup

Describes an Auto Scaling group.

Contents

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

AvailabilityZones.member.N

One or more Availability Zones for the Auto Scaling group.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

CreatedTime

The date and time the Auto Scaling group was created.

Type: Timestamp

Required: Yes

DefaultCooldown

The duration of the default cooldown period, in seconds, for the Auto Scaling group.

Type: Integer

Required: Yes

DesiredCapacity

The desired size of the Auto Scaling group.

Type: Integer

Required: Yes

HealthCheckType

One or more comma-separated health check types for the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: [\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*

Required: Yes

MaxSize

The maximum size of the Auto Scaling group.

Type: Integer

Required: Yes

MinSize

The minimum size of the Auto Scaling group.

Type: Integer

Required: Yes

AutoScalingGroupARN

The Amazon Resource Name (ARN) of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: [\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*

Required: No

AvailabilityZoneDistribution

The EC2 instance capacity distribution across Availability Zones for the Auto Scaling group.

Type: [AvailabilityZoneDistribution](#) object

Required: No

AvailabilityZoneImpairmentPolicy

The Availability Zone impairment policy for the Auto Scaling group.

Type: [AvailabilityZoneImpairmentPolicy](#) object

Required: No

CapacityRebalance

Indicates whether Capacity Rebalancing is enabled.

Type: Boolean

Required: No

CapacityReservationSpecification

The capacity reservation specification for the Auto Scaling group.

Type: [CapacityReservationSpecification](#) object

Required: No

Context

Reserved.

Type: String

Required: No

DefaultInstanceWarmup

The duration of the default EC2 instance warmup time, in seconds, for the Auto Scaling group.

Type: Integer

Required: No

DesiredCapacityType

The unit of measurement for the value specified for desired capacity. Amazon EC2 Auto Scaling supports `DesiredCapacityType` for attribute-based instance type selection only.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EnabledMetrics.member.N

The metrics enabled for the Auto Scaling group.

Type: Array of [EnabledMetric](#) objects

Required: No

HealthCheckGracePeriod

The duration of the health check grace period, in seconds, for the Auto Scaling group.

Type: Integer

Required: No

InstanceLifecyclePolicy

The instance lifecycle policy for the Auto Scaling group.

Type: [InstanceLifecyclePolicy](#) object

Required: No

InstanceMaintenancePolicy

An instance maintenance policy.

Type: [InstanceMaintenancePolicy](#) object

Required: No

Instances.member.N

The EC2 instances associated with the Auto Scaling group.

Type: Array of [Instance](#) objects

Required: No

LaunchConfigurationName

The name of the associated launch configuration for the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchTemplate

The launch template for the Auto Scaling group.

Type: [LaunchTemplateSpecification](#) object

Required: No

LoadBalancerNames.member.N

One or more load balancers associated with the group.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MaxInstanceLifetime

The maximum amount of time, in seconds, that an EC2 instance can be in service for the Auto Scaling group.

Type: Integer

Required: No

MixedInstancesPolicy

The mixed instances policy for the group.

Type: [MixedInstancesPolicy](#) object

Required: No

NewInstancesProtectedFromScaleIn

Indicates whether newly launched EC2 instances are protected from termination when scaling in for the Auto Scaling group.

For more information about preventing instances from terminating on scale in, see [Use instance scale-in protection](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

PlacementGroup

The name of the placement group into which to launch EC2 instances for the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

PredictedCapacity

The predicted capacity of the group when it has a predictive scaling policy.

Type: Integer

Required: No

ServiceLinkedRoleARN

The Amazon Resource Name (ARN) of the service-linked role that the Auto Scaling group uses to call other Amazon Web Services on your behalf.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Status

The current state of the Auto Scaling group when the [DeleteAutoScalingGroup](#) operation is in progress.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SuspendedProcesses.member.N

The suspended processes associated with the Auto Scaling group.

Type: Array of [SuspendedProcess](#) objects

Required: No

Tags.member.N

The tags for the Auto Scaling group.

Type: Array of [TagDescription](#) objects

Required: No

TargetGroupARNs.member.N

The Amazon Resource Names (ARN) of the target groups for your load balancer.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

TerminationPolicies.member.N

The termination policies for the Auto Scaling group.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

TrafficSources.member.N

The traffic sources associated with this Auto Scaling group.

Type: Array of [TrafficSourceIdentifier](#) objects

Required: No

VPCZoneIdentifier

One or more comma-separated subnet IDs for the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 5000.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

WarmPoolConfiguration

The warm pool for the group.

Type: [WarmPoolConfiguration](#) object

Required: No

WarmPoolSize

The current size of the warm pool.

Type: Integer

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AutoScalingInstanceDetails

Describes an EC2 instance associated with an Auto Scaling group.

Contents

AutoScalingGroupName

The name of the Auto Scaling group for the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

AvailabilityZone

The Availability Zone for the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

HealthStatus

The last reported health status of this instance. `Healthy` means that the instance is healthy and should remain in service. `Unhealthy` means that the instance is unhealthy and Amazon EC2 Auto Scaling should terminate and replace it.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceId

The ID of the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

LifecycleState

The lifecycle state for the instance. The Quarantined state is not used. For more information, see [Amazon EC2 Auto Scaling instance lifecycle](#) in the *Amazon EC2 Auto Scaling User Guide*.

Valid values: Pending | Pending:Wait | Pending:Proceed | Quarantined | InService | Terminating | Terminating:Wait | Terminating:Proceed | Terminating:Retained | Terminated | Detaching | Detached | EnteringStandby | Standby | Warmed:Pending | Warmed:Pending:Wait | Warmed:Pending:Proceed | Warmed:Pending:Retained | Warmed:Terminating | Warmed:Terminating:Wait | Warmed:Terminating:Proceed | Warmed:Terminating:Retained | Warmed:Terminated | Warmed:Stopped | Warmed:Running

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

ProtectedFromScaleIn

Indicates whether the instance is protected from termination by Amazon EC2 Auto Scaling when scaling in.

Type: Boolean

Required: Yes

ImageId

The ID of the Amazon Machine Image (AMI) associated with the instance. This field shows the current AMI ID of the instance's root volume. It may differ from the original AMI used when the instance was first launched.

This field appears for:

- Instances with root volume replacements through Instance Refresh
- Instances launched with AMI overrides

This field won't appear for:

- Existing instances launched from Launch Templates without overrides
- Existing instances that didn't have their root volume replaced through Instance Refresh

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceType

The instance type of the EC2 instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchConfigurationName

The launch configuration used to launch the instance. This value is not available if you attached the instance to the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchTemplate

The launch template for the instance.

Type: [LaunchTemplateSpecification](#) object

Required: No

WeightedCapacity

The number of capacity units contributed by the instance based on its instance type.

Valid Range: Minimum value of 1. Maximum value of 999.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 3.

Pattern: `^[\u0031-\u0039][\u0030-\u0039]{0,2}$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AvailabilityZoneDistribution

Describes an Availability Zone distribution.

Contents

CapacityDistributionStrategy

If launches fail in an Availability Zone, the following strategies are available. The default is `balanced-best-effort`.

- `balanced-only` - If launches fail in an Availability Zone, Auto Scaling will continue to attempt to launch in the unhealthy zone to preserve a balanced distribution.
- `balanced-best-effort` - If launches fail in an Availability Zone, Auto Scaling will attempt to launch in another healthy Availability Zone instead.

Type: String

Valid Values: `balanced-only` | `balanced-best-effort`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

AvailabilityZoneImpairmentPolicy

Describes an Availability Zone impairment policy.

Contents

ImpairedZoneHealthCheckBehavior

Specifies the health check behavior for the impaired Availability Zone in an active zonal shift. If you select `Replace_unhealthy`, instances that appear unhealthy will be replaced in all Availability Zones. If you select `Ignore_unhealthy`, instances will not be replaced in the Availability Zone with the active zonal shift. For more information, see [Auto Scaling group zonal shift](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Valid Values: `ReplaceUnhealthy` | `IgnoreUnhealthy`

Required: No

ZonalShiftEnabled

If `true`, enable zonal shift for your Auto Scaling group.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BaselineEbsBandwidthMbpsRequest

Specifies the minimum and maximum for the `BaselineEbsBandwidthMbps` object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Max

The maximum value in Mbps.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Min

The minimum value in Mbps.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BaselinePerformanceFactorsRequest

The baseline performance to consider, using an instance family as a baseline reference. The instance family establishes the lowest acceptable level of performance. Auto Scaling uses this baseline to guide instance type selection, but there is no guarantee that the selected instance types will always exceed the baseline for every application.

Currently, this parameter only supports CPU performance as a baseline performance factor. For example, specifying `c6i` uses the CPU performance of the `c6i` family as the baseline reference.

Contents

Cpu

The CPU performance to consider, using an instance family as the baseline reference.

Type: [CpuPerformanceFactorRequest](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

BlockDeviceMapping

Describes a block device mapping.

Contents

DeviceName

The device name assigned to the volume (for example, `/dev/sdh` or `xvdh`). For more information, see [Device naming on Linux instances](#) in the *Amazon EC2 User Guide*.

Note

To define a block device mapping, set the device name and exactly one of the following properties: `Ebs`, `NoDevice`, or `VirtualName`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Ebs

Information to attach an EBS volume to an instance at launch.

Type: [Ebs](#) object

Required: No

NoDevice

Setting this value to `true` prevents a volume that is included in the block device mapping of the AMI from being mapped to the specified device name at launch.

If `NoDevice` is `true` for the root device, instances might fail the EC2 health check. In that case, Amazon EC2 Auto Scaling launches replacement instances.

Type: Boolean

Required: No

VirtualName

The name of the instance store volume (virtual device) to attach to an instance at launch. The name must be in the form ephemeralX where X is a number starting from zero (0), for example, ephemeral0.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CapacityForecast

A `GetPredictiveScalingForecast` call returns the capacity forecast for a predictive scaling policy. This structure includes the data points for that capacity forecast, along with the timestamps of those data points.

Contents

Timestamps.member.N

The timestamps for the data points, in UTC format.

Type: Array of timestamps

Required: Yes

Values.member.N

The values of the data points.

Type: Array of doubles

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CapacityReservationSpecification

Describes the Capacity Reservation preference and targeting options. If you specify `open` or `none` for `CapacityReservationPreference`, do not specify a `CapacityReservationTarget`.

Contents

CapacityReservationPreference

The capacity reservation preference. The following options are available:

- `capacity-reservations-only` - Auto Scaling will only launch instances into a Capacity Reservation or Capacity Reservation resource group. If capacity isn't available, instances will fail to launch.
- `capacity-reservations-first` - Auto Scaling will try to launch instances into a Capacity Reservation or Capacity Reservation resource group first. If capacity isn't available, instances will run in On-Demand capacity.
- `none` - Auto Scaling will not launch instances into a Capacity Reservation. Instances will run in On-Demand capacity.
- `default` - Auto Scaling uses the Capacity Reservation preference from your launch template or an open Capacity Reservation.

Type: String

Valid Values: `capacity-reservations-only` | `capacity-reservations-first` | `none` | `default`

Required: No

CapacityReservationTarget

Describes a target Capacity Reservation or Capacity Reservation resource group.

Type: [CapacityReservationTarget](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CapacityReservationTarget

The target for the Capacity Reservation. Specify Capacity Reservations IDs or Capacity Reservation resource group ARNs.

Contents

CapacityReservationIds.member.N

The Capacity Reservation IDs to launch instances into.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z0-9\-\_\/]+`

Required: No

CapacityReservationResourceGroupArns.member.N

The resource group ARNs of the Capacity Reservation to launch instances into.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CpuPerformanceFactorRequest

The CPU performance to consider, using an instance family as the baseline reference.

Contents

Reference.item.N

Specify an instance family to use as the baseline reference for CPU performance. All instance types that match your specified attributes will be compared against the CPU performance of the referenced instance family, regardless of CPU manufacturer or architecture differences.

Note

Currently only one instance family can be specified in the list.

Type: Array of [PerformanceFactorReferenceRequest](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

CustomizedMetricSpecification

Represents a CloudWatch metric of your choosing for a target tracking scaling policy to use with Amazon EC2 Auto Scaling.

To create your customized metric specification:

- Add values for each required property from CloudWatch. You can use an existing metric, or a new metric that you create. To use your own metric, you must first publish the metric to CloudWatch. For more information, see [Publish custom metrics](#) in the *Amazon CloudWatch User Guide*.
- Choose a metric that changes proportionally with capacity. The value of the metric should increase or decrease in inverse proportion to the number of capacity units. That is, the value of the metric should decrease when capacity increases.

For more information about the CloudWatch terminology below, see [Amazon CloudWatch concepts](#).

Note

Each individual service provides information about the metrics, namespace, and dimensions they use. For more information, see [AWS services that publish CloudWatch metrics](#) in the *Amazon CloudWatch User Guide*.

Contents

Dimensions.member.N

The dimensions of the metric.

Conditional: If you published your metric with dimensions, you must specify the same dimensions in your scaling policy.

Type: Array of [MetricDimension](#) objects

Required: No

MetricName

The name of the metric. To get the exact metric name, namespace, and dimensions, inspect the [Metric](#) object that is returned by a call to [ListMetrics](#).

Type: String

Required: No

Metrics.member.N

The metrics to include in the target tracking scaling policy, as a metric data query. This can include both raw metric and metric math expressions.

Type: Array of [TargetTrackingMetricDataQuery](#) objects

Required: No

Namespace

The namespace of the metric.

Type: String

Required: No

Period

The period of the metric in seconds. The default value is 60. Accepted values are 10, 30, and 60. For high resolution metric, set the value to less than 60. For more information, see [Create a target tracking policy using high-resolution metrics for faster response](#).

Type: Integer

Valid Range: Minimum value of 1.

Required: No

Statistic

The statistic of the metric.

Type: String

Valid Values: Average | Minimum | Maximum | SampleCount | Sum

Required: No

Unit

The unit of the metric. For a complete list of the units that CloudWatch supports, see the [MetricDatum](#) data type in the *Amazon CloudWatch API Reference*.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

DesiredConfiguration

Describes the desired configuration for an instance refresh.

If you specify a desired configuration, you must specify either a `LaunchTemplate` or a `MixedInstancesPolicy`.

Contents

LaunchTemplate

Describes the launch template and the version of the launch template that Amazon EC2 Auto Scaling uses to launch Amazon EC2 instances. For more information about launch templates, see [Launch templates](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [LaunchTemplateSpecification](#) object

Required: No

MixedInstancesPolicy

Use this structure to launch multiple instance types and On-Demand Instances and Spot Instances within a single Auto Scaling group.

A mixed instances policy contains information that Amazon EC2 Auto Scaling can use to launch instances and help optimize your costs. For more information, see [Auto Scaling groups with multiple instance types and purchase options](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [MixedInstancesPolicy](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Ebs

Describes information used to set up an Amazon EBS volume specified in a block device mapping.

Contents

DeleteOnTermination

Indicates whether the volume is deleted on instance termination. For Amazon EC2 Auto Scaling, the default value is `true`.

Type: Boolean

Required: No

Encrypted

Specifies whether the volume should be encrypted. Encrypted EBS volumes can only be attached to instances that support Amazon EBS encryption. For more information, see [Requirements for Amazon EBS encryption](#) in the *Amazon EBS User Guide*. If your AMI uses encrypted volumes, you can also only launch it on supported instance types.

Note

If you are creating a volume from a snapshot, you cannot create an unencrypted volume from an encrypted snapshot. Also, you cannot specify a KMS key ID when using a launch configuration.

If you enable encryption by default, the EBS volumes that you create are always encrypted, either using the AWS managed KMS key or a customer-managed KMS key, regardless of whether the snapshot was encrypted.

For more information, see [Use AWS KMS keys to encrypt Amazon EBS volumes](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

Iops

The number of input/output (I/O) operations per second (IOPS) to provision for the volume. For gp3 and io1 volumes, this represents the number of IOPS that are provisioned for the volume.

For gp2 volumes, this represents the baseline performance of the volume and the rate at which the volume accumulates I/O credits for bursting.

The following are the supported values for each volume type:

- gp3: 3,000-16,000 IOPS
- io1: 100-64,000 IOPS

For io1 volumes, we guarantee 64,000 IOPS only for [Instances built on the AWS Nitro System](#). Other instance families guarantee performance up to 32,000 IOPS.

Iops is supported when the volume type is gp3 or io1 and required only when the volume type is io1. (Not used with standard, gp2, st1, or sc1 volumes.)

Type: Integer

Valid Range: Minimum value of 100. Maximum value of 64000.

Required: No

SnapshotId

The snapshot ID of the volume to use.

You must specify either a VolumeSize or a SnapshotId.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Throughput

The throughput (MiBps) to provision for a gp3 volume.

Type: Integer

Valid Range: Minimum value of 125. Maximum value of 1000.

Required: No

VolumeSize

The volume size, in GiBs. The following are the supported volumes sizes for each volume type:

- gp2 and gp3: 1-16,384
- io1: 4-16,384
- st1 and sc1: 125-16,384
- standard: 1-1,024

You must specify either a `SnapshotId` or a `VolumeSize`. If you specify both `SnapshotId` and `VolumeSize`, the volume size must be equal or greater than the size of the snapshot.

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 16384.

Required: No

VolumeType

The volume type. For more information, see [Amazon EBS volume types](#) in the *Amazon EBS User Guide*.

Valid values: standard | io1 | gp2 | st1 | sc1 | gp3

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

EnabledMetric

Describes an enabled Auto Scaling group metric.

Contents

Granularity

The granularity of the metric. The only valid value is 1Minute.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Metric

One of the following metrics:

- GroupMinSize
- GroupMaxSize
- GroupDesiredCapacity
- GroupInServiceInstances
- GroupPendingInstances
- GroupStandbyInstances
- GroupTerminatingInstances
- GroupTotalInstances
- GroupInServiceCapacity
- GroupPendingCapacity
- GroupStandbyCapacity
- GroupTerminatingCapacity
- GroupTotalCapacity
- WarmPoolDesiredCapacity
- WarmPoolWarmedCapacity

- WarmPoolPendingCapacity
- WarmPoolTerminatingCapacity
- WarmPoolTotalCapacity
- GroupAndWarmPoolDesiredCapacity
- GroupAndWarmPoolTotalCapacity

For more information, see [Amazon CloudWatch metrics for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

FailedScheduledUpdateGroupActionRequest

Describes a scheduled action that could not be created, updated, or deleted.

Contents

ScheduledActionName

The name of the scheduled action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF\\r\\n\\t]*`

Required: Yes

ErrorCode

The error code.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ErrorMessage

The error message accompanying the error code.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Filter

Describes a filter that is used to return a more specific list of results from a describe operation.

If you specify multiple filters, the filters are automatically logically joined with an AND, and the request returns only the results that match all of the specified filters.

For more information, see [Tag Auto Scaling groups and instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

Name

The name of the filter.

The valid values for Name depend on which API operation you're using with the filter ([DescribeAutoScalingGroups](#) or [DescribeTags](#)).

DescribeAutoScalingGroups

Valid values for Name include the following:

- `tag-key` - Accepts tag keys. The results only include information about the Auto Scaling groups associated with these tag keys.
- `tag-value` - Accepts tag values. The results only include information about the Auto Scaling groups associated with these tag values.
- `tag:<key>` - Accepts the key/value combination of the tag. Use the tag key in the filter name and the tag value as the filter value. The results only include information about the Auto Scaling groups associated with the specified key/value combination.

DescribeTags

Valid values for Name include the following:

- `auto-scaling-group` - Accepts the names of Auto Scaling groups. The results only include information about the tags associated with these Auto Scaling groups.
- `key` - Accepts tag keys. The results only include information about the tags associated with these tag keys.
- `value` - Accepts tag values. The results only include information about the tags associated with these tag values.

- `propagate-at-launch` - Accepts a Boolean value, which specifies whether tags propagate to instances at launch. The results only include information about the tags associated with the specified Boolean value.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Values.member.N

One or more filter values. Filter values are case-sensitive.

If you specify multiple values for a filter, the values are automatically logically joined with an OR, and the request returns all results that match any of the specified values. For example, specify "tag:environment" for the filter name and "production,development" for the filter values to find Auto Scaling groups with the tag "environment=production" or "environment=development".

Type: Array of strings

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Instance

Describes an EC2 instance.

Contents

AvailabilityZone

The Availability Zone in which the instance is running.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

HealthStatus

The last reported health status of the instance. `Healthy` means that the instance is healthy and should remain in service. `Unhealthy` means that the instance is unhealthy and that Amazon EC2 Auto Scaling should terminate and replace it.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceId

The ID of the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

LifecycleState

A description of the current lifecycle state. The Quarantined state is not used. For more information, see [Amazon EC2 Auto Scaling instance lifecycle](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Valid Values: Pending | Pending:Wait | Pending:Proceed | Quarantined | InService | Terminating | Terminating:Wait | Terminating:Proceed | Terminated | Detaching | Detached | EnteringStandby | Standby | Warmed:Pending | Warmed:Pending:Wait | Warmed:Pending:Proceed | Warmed:Terminating | Warmed:Terminating:Wait | Warmed:Terminating:Proceed | Warmed:Terminated | Warmed:Stopped | Warmed:Running | Warmed:Hibernated

Required: Yes

ProtectedFromScaleIn

Indicates whether the instance is protected from termination by Amazon EC2 Auto Scaling when scaling in.

Type: Boolean

Required: Yes

ImageId

The ID of the Amazon Machine Image (AMI) used for the instance's current root volume. This value reflects the most recent AMI applied to the instance, including updates made through root volume replacement operations.

This field appears for:

- Instances with root volume replacements through Instance Refresh
- Instances launched with AMI overrides

This field won't appear for:

- Existing instances launched from Launch Templates without overrides
- Existing instances that didn't have their root volume replaced through Instance Refresh

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceType

The instance type of the EC2 instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchConfigurationName

The launch configuration associated with the instance.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchTemplate

The launch template for the instance.

Type: [LaunchTemplateSpecification](#) object

Required: No

WeightedCapacity

The number of capacity units contributed by the instance based on its instance type.

Valid Range: Minimum value of 1. Maximum value of 999.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 3.

Pattern: `^[\\u0031-\\u0039][\\u0030-\\u0039]{0,2}$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceCollection

Contains details about a collection of instances launched in the Auto Scaling group.

Contents

AvailabilityZone

The Availability Zone where the instances were launched.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

AvailabilityZoneId

The Availability Zone ID where the instances in this collection were launched.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceIds.member.N

A list of instance IDs for the successfully launched instances.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 19.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceType

The instance type of the launched instances.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MarketType

The market type for the instances (On-Demand or Spot).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SubnetId

The ID of the subnet where the instances were launched.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceLifecyclePolicy

The instance lifecycle policy for the Auto Scaling group. This policy controls instance behavior when an instance transitions through its lifecycle states. Configure retention triggers to specify when instances should move to a Retained state instead of automatic termination.

For more information, see [Control instance retention with instance lifecycle policies](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

RetentionTriggers

Specifies the conditions that trigger instance retention behavior. These triggers determine when instances should move to a Retained state instead of automatic termination. This allows you to maintain control over instance management when lifecycles transition and operations fail.

Type: [RetentionTriggers](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceMaintenancePolicy

Describes an instance maintenance policy.

For more information, see [Set instance maintenance policy](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

MaxHealthyPercentage

Specifies the upper threshold as a percentage of the desired capacity of the Auto Scaling group. It represents the maximum percentage of the group that can be in service and healthy, or pending, to support your workload when replacing instances. Value range is 100 to 200. To clear a previously set value, specify a value of -1.

Both MinHealthyPercentage and MaxHealthyPercentage must be specified, and the difference between them cannot be greater than 100. A large range increases the number of instances that can be replaced at the same time.

Type: Integer

Valid Range: Minimum value of -1. Maximum value of 200.

Required: No

MinHealthyPercentage

Specifies the lower threshold as a percentage of the desired capacity of the Auto Scaling group. It represents the minimum percentage of the group to keep in service, healthy, and ready to use to support your workload when replacing instances. Value range is 0 to 100. To clear a previously set value, specify a value of -1.

Type: Integer

Valid Range: Minimum value of -1. Maximum value of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceMetadataOptions

The metadata options for the instances. For more information, see [Configure the instance metadata options](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

HttpEndpoint

This parameter enables or disables the HTTP metadata endpoint on your instances. If the parameter is not specified, the default state is enabled.

Note

If you specify a value of `disabled`, you will not be able to access your instance metadata.

Type: String

Valid Values: `disabled` | `enabled`

Required: No

HttpPutResponseHopLimit

The desired HTTP PUT response hop limit for instance metadata requests. The larger the number, the further instance metadata requests can travel.

Default: 1

Type: Integer

Valid Range: Minimum value of 1. Maximum value of 64.

Required: No

HttpTokens

The state of token usage for your instance metadata requests. If the parameter is not specified in the request, the default state is `optional`.

If the state is `optional`, you can choose to retrieve instance metadata with or without a signed token header on your request. If you retrieve the IAM role credentials without a token, the version 1.0 role credentials are returned. If you retrieve the IAM role credentials using a valid signed token, the version 2.0 role credentials are returned.

If the state is `required`, you must send a signed token header with any instance metadata retrieval requests. In this state, retrieving the IAM role credentials always returns the version 2.0 credentials; the version 1.0 credentials are not available.

Type: String

Valid Values: `optional` | `required`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceMonitoring

Describes whether detailed monitoring is enabled for the Auto Scaling instances.

Contents

Enabled

If `true`, detailed monitoring is enabled. Otherwise, basic monitoring is enabled.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceRefresh

Describes an instance refresh for an Auto Scaling group.

Contents

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

DesiredConfiguration

Describes the desired configuration for the instance refresh.

Type: [DesiredConfiguration](#) object

Required: No

EndTime

The date and time at which the instance refresh ended.

Type: Timestamp

Required: No

InstanceId

The instance refresh ID.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstancesToUpdate

The number of instances remaining to update before the instance refresh is complete.

Note

If you roll back the instance refresh, `InstancesToUpdate` shows you the number of instances that were not yet updated by the instance refresh. Therefore, these instances don't need to be replaced as part of the rollback.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PercentageComplete

The percentage of the instance refresh that is complete. For each instance replacement, Amazon EC2 Auto Scaling tracks the instance's health status and warm-up time. When the instance's health status changes to healthy and the specified warm-up time passes, the instance is considered updated and is added to the percentage complete.

Note

`PercentageComplete` does not include instances that are replaced during a rollback. This value gradually goes back down to zero during a rollback.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

Preferences

The preferences for an instance refresh.

Type: [RefreshPreferences](#) object

Required: No

ProgressDetails

Additional progress details for an Auto Scaling group that has a warm pool.

Type: [InstanceRefreshProgressDetails](#) object

Required: No

RollbackDetails

The rollback details.

Type: [RollbackDetails](#) object

Required: No

StartTime

The date and time at which the instance refresh began.

Type: Timestamp

Required: No

Status

The current status for the instance refresh operation:

- **Pending** - The request was created, but the instance refresh has not started.
- **InProgress** - An instance refresh is in progress.
- **Successful** - An instance refresh completed successfully.
- **Failed** - An instance refresh failed to complete. You can troubleshoot using the status reason and the scaling activities.
- **Cancelling** - An ongoing instance refresh is being cancelled.
- **Cancelled** - The instance refresh is cancelled.
- **RollbackInProgress** - An instance refresh is being rolled back.
- **RollbackFailed** - The rollback failed to complete. You can troubleshoot using the status reason and the scaling activities.
- **RollbackSuccessful** - The rollback completed successfully.
- **Baking** - Waiting the specified bake time after an instance refresh has finished updating instances.

Type: String

Valid Values: Pending | InProgress | Successful | Failed | Cancelling | Cancelled | RollbackInProgress | RollbackFailed | RollbackSuccessful | Baking

Required: No

StatusReason

The explanation for the specific status assigned to this operation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

Strategy

The strategy to use for the instance refresh. This determines how instances in the Auto Scaling group are updated. Default is Rolling.

- `Rolling` – Terminates instances and launches replacements in batches
- `ReplaceRootVolume` – Updates instances by replacing only the root volume without terminating the instance

Type: String

Valid Values: `Rolling` | `ReplaceRootVolume`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

InstanceRefreshLivePoolProgress

Reports progress on replacing instances that are in the Auto Scaling group.

Contents

InstancesToUpdate

The number of instances remaining to update.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PercentageComplete

The percentage of instances in the Auto Scaling group that have been replaced. For each instance replacement, Amazon EC2 Auto Scaling tracks the instance's health status and warm-up time. When the instance's health status changes to healthy and the specified warm-up time passes, the instance is considered updated and is added to the percentage complete.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceRefreshProgressDetails

Reports progress on replacing instances in an Auto Scaling group that has a warm pool. This includes separate details for instances in the warm pool and instances in the Auto Scaling group (the live pool).

Contents

LivePoolProgress

Reports progress on replacing instances that are in the Auto Scaling group.

Type: [InstanceRefreshLivePoolProgress](#) object

Required: No

WarmPoolProgress

Reports progress on replacing instances that are in the warm pool.

Type: [InstanceRefreshWarmPoolProgress](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceRefreshWarmPoolProgress

Reports progress on replacing instances that are in the warm pool.

Contents

InstancesToUpdate

The number of instances remaining to update.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PercentageComplete

The percentage of instances in the warm pool that have been replaced. For each instance replacement, Amazon EC2 Auto Scaling tracks the instance's health status and warm-up time. When the instance's health status changes to healthy and the specified warm-up time passes, the instance is considered updated and is added to the percentage complete.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceRequirements

The attributes for the instance types for a mixed instances policy. Amazon EC2 Auto Scaling uses your specified requirements to identify instance types. Then, it uses your On-Demand and Spot allocation strategies to launch instances from these instance types.

When you specify multiple attributes, you get instance types that satisfy all of the specified attributes. If you specify multiple values for an attribute, you get instance types that satisfy any of the specified values.

To limit the list of instance types from which Amazon EC2 Auto Scaling can identify matching instance types, you can use one of the following parameters, but not both in the same request:

- `AllowedInstanceTypes` - The instance types to include in the list. All other instance types are ignored, even if they match your specified attributes.
- `ExcludedInstanceTypes` - The instance types to exclude from the list, even if they match your specified attributes.

Note

You must specify `VCpuCount` and `MemoryMiB`. All other attributes are optional. Any unspecified optional attribute is set to its default.

For more information, see [Create a mixed instances group using attribute-based instance type selection](#) in the *Amazon EC2 Auto Scaling User Guide*. For help determining which instance types match your attributes before you apply them to your Auto Scaling group, see [Preview instance types with specified attributes](#) in the *Amazon EC2 User Guide*.

Contents

MemoryMiB

The minimum and maximum instance memory size for an instance type, in MiB.

Type: [MemoryMiBRequest](#) object

Required: Yes

VCpuCount

The minimum and maximum number of vCPUs for an instance type.

Type: [VCpuCountRequest](#) object

Required: Yes

AcceleratorCount

The minimum and maximum number of accelerators (GPUs, FPGAs, or AWS Inferentia chips) for an instance type.

To exclude accelerator-enabled instance types, set Max to 0.

Default: No minimum or maximum limits

Type: [AcceleratorCountRequest](#) object

Required: No

AcceleratorManufacturers.member.N

Indicates whether instance types must have accelerators by specific manufacturers.

- For instance types with NVIDIA devices, specify `nvidia`.
- For instance types with AMD devices, specify `amd`.
- For instance types with AWS devices, specify `amazon-web-services`.
- For instance types with Xilinx devices, specify `xilinx`.

Default: Any manufacturer

Type: Array of strings

Valid Values: `nvidia` | `amd` | `amazon-web-services` | `xilinx`

Required: No

AcceleratorNames.member.N

Lists the accelerators that must be on an instance type.

- For instance types with NVIDIA A100 GPUs, specify `a100`.
- For instance types with NVIDIA V100 GPUs, specify `v100`.
- For instance types with NVIDIA K80 GPUs, specify `k80`.

- For instance types with NVIDIA T4 GPUs, specify `t4`.
- For instance types with NVIDIA M60 GPUs, specify `m60`.
- For instance types with AMD Radeon Pro V520 GPUs, specify `radeon-pro-v520`.
- For instance types with Xilinx VU9P FPGAs, specify `vu9p`.

Default: Any accelerator

Type: Array of strings

Valid Values: `a100` | `v100` | `k80` | `t4` | `m60` | `radeon-pro-v520` | `vu9p`

Required: No

AcceleratorTotalMemoryMiB

The minimum and maximum total memory size for the accelerators on an instance type, in MiB.

Default: No minimum or maximum limits

Type: [AcceleratorTotalMemoryMiBRequest](#) object

Required: No

AcceleratorTypes.member.N

Lists the accelerator types that must be on an instance type.

- For instance types with GPU accelerators, specify `gpu`.
- For instance types with FPGA accelerators, specify `fpga`.
- For instance types with inference accelerators, specify `inference`.

Default: Any accelerator type

Type: Array of strings

Valid Values: `gpu` | `fpga` | `inference`

Required: No

AllowedInstanceTypes.member.N

The instance types to apply your specified attributes against. All other instance types are ignored, even if they match your specified attributes.

You can use strings with one or more wild cards, represented by an asterisk (*), to allow an instance type, size, or generation. The following are examples: `m5.8xlarge`, `c5*.*`, `m5a.*`, `r*`, `*3*`.

For example, if you specify `c5*`, Amazon EC2 Auto Scaling will allow the entire C5 instance family, which includes all C5a and C5n instance types. If you specify `m5a.*`, Amazon EC2 Auto Scaling will allow all the M5a instance types, but not the M5n instance types.

 Note

If you specify `AllowedInstanceTypes`, you can't specify `ExcludedInstanceTypes`.

Default: All instance types

Type: Array of strings

Array Members: Maximum number of 400 items.

Length Constraints: Minimum length of 1. Maximum length of 30.

Pattern: `[a-zA-Z0-9\.\*\-]+`

Required: No

BareMetal

Indicates whether bare metal instance types are included, excluded, or required.

Default: `excluded`

Type: String

Valid Values: `included` | `excluded` | `required`

Required: No

BaselineEbsBandwidthMbps

The minimum and maximum baseline bandwidth performance for an instance type, in Mbps. For more information, see [Amazon EBS-optimized instances](#) in the *Amazon EC2 User Guide*.

Default: No minimum or maximum limits

Type: [BaselineEbsBandwidthMbpsRequest](#) object

Required: No

BaselinePerformanceFactors

The baseline performance factors for the instance requirements.

Type: [BaselinePerformanceFactorsRequest](#) object

Required: No

BurstablePerformance

Indicates whether burstable performance instance types are included, excluded, or required. For more information, see [Burstable performance instances](#) in the *Amazon EC2 User Guide*.

Default: excluded

Type: String

Valid Values: included | excluded | required

Required: No

CpuManufacturers.member.N

Lists which specific CPU manufacturers to include.

- For instance types with Intel CPUs, specify `intel`.
- For instance types with AMD CPUs, specify `amd`.
- For instance types with AWS CPUs, specify `amazon-web-services`.
- For instance types with Apple CPUs, specify `apple`.

Note

Don't confuse the CPU hardware manufacturer with the CPU hardware architecture. Instances will be launched with a compatible CPU architecture based on the Amazon Machine Image (AMI) that you specify in your launch template.

Default: Any manufacturer

Type: Array of strings

Valid Values: `intel` | `amd` | `amazon-web-services` | `apple`

Required: No

ExcludedInstanceTypes.member.N

The instance types to exclude. You can use strings with one or more wild cards, represented by an asterisk (*), to exclude an instance family, type, size, or generation. The following are examples: `m5.8xlarge`, `c5*.*`, `m5a.*`, `r*`, `*3*`.

For example, if you specify `c5*`, you are excluding the entire C5 instance family, which includes all C5a and C5n instance types. If you specify `m5a.*`, Amazon EC2 Auto Scaling will exclude all the M5a instance types, but not the M5n instance types.

Note

If you specify `ExcludedInstanceTypes`, you can't specify `AllowedInstanceTypes`.

Default: No excluded instance types

Type: Array of strings

Array Members: Maximum number of 400 items.

Length Constraints: Minimum length of 1. Maximum length of 30.

Pattern: `[a-zA-Z0-9\.\*\-]+`

Required: No

InstanceGenerations.member.N

Indicates whether current or previous generation instance types are included.

- For current generation instance types, specify `current`. The current generation includes EC2 instance types currently recommended for use. This typically includes the latest two to three generations in each instance family. For more information, see [Instance types](#) in the *Amazon EC2 User Guide*.
- For previous generation instance types, specify `previous`.

Default: Any current or previous generation

Type: Array of strings

Valid Values: current | previous

Required: No

LocalStorage

Indicates whether instance types with instance store volumes are included, excluded, or required. For more information, see [Amazon EC2 instance store](#) in the *Amazon EC2 User Guide*.

Default: included

Type: String

Valid Values: included | excluded | required

Required: No

LocalStorageTypes.member.N

Indicates the type of local storage that is required.

- For instance types with hard disk drive (HDD) storage, specify hdd.
- For instance types with solid state drive (SSD) storage, specify ssd.

Default: Any local storage type

Type: Array of strings

Valid Values: hdd | ssd

Required: No

MaxSpotPriceAsPercentageOfOptimalOnDemandPrice

[Price protection] The price protection threshold for Spot Instances, as a percentage of an identified On-Demand price. The identified On-Demand price is the price of the lowest priced current generation C, M, or R instance type with your specified attributes. If no current generation C, M, or R instance type matches your attributes, then the identified price is from either the lowest priced current generation instance types or, failing that, the lowest priced previous generation instance types that match your attributes. When Amazon EC2 Auto Scaling

selects instance types with your attributes, we will exclude instance types whose price exceeds your specified threshold.

The parameter accepts an integer, which Amazon EC2 Auto Scaling interprets as a percentage.

If you set `DesiredCapacityType` to `vcpu` or `memory-mib`, the price protection threshold is based on the per-vCPU or per-memory price instead of the per instance price.

 Note

Only one of `SpotMaxPricePercentageOverLowestPrice` or `MaxSpotPriceAsPercentageOfOptimalOnDemandPrice` can be specified. If you don't specify either, Amazon EC2 Auto Scaling will automatically apply optimal price protection to consistently select from a wide range of instance types. To indicate no price protection threshold for Spot Instances, meaning you want to consider all instance types that match your attributes, include one of these parameters and specify a high value, such as 999999.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

MemoryGiBPerVCpu

The minimum and maximum amount of memory per vCPU for an instance type, in GiB.

Default: No minimum or maximum limits

Type: [MemoryGiBPerVCpuRequest](#) object

Required: No

NetworkBandwidthGbps

The minimum and maximum amount of network bandwidth, in gigabits per second (Gbps).

Default: No minimum or maximum limits

Type: [NetworkBandwidthGbpsRequest](#) object

Required: No

NetworkInterfaceCount

The minimum and maximum number of network interfaces for an instance type.

Default: No minimum or maximum limits

Type: [NetworkInterfaceCountRequest](#) object

Required: No

OnDemandMaxPricePercentageOverLowestPrice

[Price protection] The price protection threshold for On-Demand Instances, as a percentage higher than an identified On-Demand price. The identified On-Demand price is the price of the lowest priced current generation C, M, or R instance type with your specified attributes. If no current generation C, M, or R instance type matches your attributes, then the identified price is from either the lowest priced current generation instance types or, failing that, the lowest priced previous generation instance types that match your attributes. When Amazon EC2 Auto Scaling selects instance types with your attributes, we will exclude instance types whose price exceeds your specified threshold.

The parameter accepts an integer, which Amazon EC2 Auto Scaling interprets as a percentage.

To turn off price protection, specify a high value, such as 999999.

If you set `DesiredCapacityType` to `vcpu` or `memory-mib`, the price protection threshold is applied based on the per-vCPU or per-memory price instead of the per instance price.

Default: 20

Type: Integer

Valid Range: Minimum value of 0.

Required: No

RequireHibernateSupport

Indicates whether instance types must provide On-Demand Instance hibernation support.

Default: false

Type: Boolean

Required: No

SpotMaxPricePercentageOverLowestPrice

[Price protection] The price protection threshold for Spot Instances, as a percentage higher than an identified Spot price. The identified Spot price is the price of the lowest priced current generation C, M, or R instance type with your specified attributes. If no current generation C, M, or R instance type matches your attributes, then the identified price is from either the lowest priced current generation instance types or, failing that, the lowest priced previous generation instance types that match your attributes. When Amazon EC2 Auto Scaling selects instance types with your attributes, we will exclude instance types whose price exceeds your specified threshold.

The parameter accepts an integer, which Amazon EC2 Auto Scaling interprets as a percentage.

If you set `DesiredCapacityType` to `vcpu` or `memory-mib`, the price protection threshold is based on the per-vCPU or per-memory price instead of the per instance price.

Note

Only one of `SpotMaxPricePercentageOverLowestPrice` or `MaxSpotPriceAsPercentageOfOptimalOnDemandPrice` can be specified. If you don't specify either, Amazon EC2 Auto Scaling will automatically apply optimal price protection to consistently select from a wide range of instance types. To indicate no price protection threshold for Spot Instances, meaning you want to consider all instance types that match your attributes, include one of these parameters and specify a high value, such as 999999.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

TotalLocalStorageGB

The minimum and maximum total local storage size for an instance type, in GB.

Default: No minimum or maximum limits

Type: [TotalLocalStorageGBRequest](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstanceReusePolicy

Describes an instance reuse policy for a warm pool.

For more information, see [Warm pools for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

ReuseOnScaleIn

Specifies whether instances in the Auto Scaling group can be returned to the warm pool on scale in.

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

InstancesDistribution

Use this structure to specify the distribution of On-Demand Instances and Spot Instances and the allocation strategies used to fulfill On-Demand and Spot capacities for a mixed instances policy.

When you define a mixed instances policy, this structure uses default values for unspecified parameters, except for `SpotMaxPrice`. However, we recommend that you review and customize the other parameters as needed and explicitly specify the allocation strategies. For help choosing allocation strategies, see [Allocation strategies](#) in the *Amazon EC2 Auto Scaling User Guide*.

When you modify `SpotAllocationStrategy`, `SpotInstancePools`, or `SpotMaxPrice` in the [UpdateAutoScalingGroup](#) API call, this update action does not deploy any changes across the running EC2 instances in the group. Your existing Spot Instances continue to run as long as the maximum price for those instances is higher than the current Spot price. When scaling out occurs, Amazon EC2 Auto Scaling launches instances based on the new settings. When scaling in occurs, Amazon EC2 Auto Scaling terminates instances according to the group's termination policies.

Contents

OnDemandAllocationStrategy

The allocation strategy to apply to your On-Demand Instances when they are launched. Possible instance types are determined by the launch template overrides that you specify.

The following lists the valid values:

lowest-price

Uses price to determine which instance types are the highest priority, launching the lowest priced instance types within an Availability Zone first. This is the default value for Auto Scaling groups that specify [InstanceRequirements](#).

prioritized

You set the order of instance types for the launch template overrides from highest to lowest priority (from first to last in the list). Amazon EC2 Auto Scaling launches your highest priority instance types first. If all your On-Demand capacity cannot be fulfilled using your highest priority instance type, then Amazon EC2 Auto Scaling launches the remaining capacity using the second priority instance type, and so on. This is the default value for Auto Scaling groups that don't specify [InstanceRequirements](#) and cannot be used for groups that do.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

OnDemandBaseCapacity

The minimum amount of the Auto Scaling group's capacity that must be fulfilled by On-Demand Instances. This base portion is launched first as your group scales.

This number has the same unit of measurement as the group's desired capacity. If you change the default unit of measurement (number of instances) by specifying weighted capacity values in your launch template overrides list, or by changing the default desired capacity type setting of the group, you must specify this number using the same unit of measurement.

Default: 0

Note

An update to this setting means a gradual replacement of instances to adjust the current On-Demand Instance levels. When replacing instances, Amazon EC2 Auto Scaling launches new instances before terminating the previous ones.

Type: Integer

Required: No

OnDemandPercentageAboveBaseCapacity

Controls the percentages of On-Demand Instances and Spot Instances for your additional capacity beyond `OnDemandBaseCapacity`. Expressed as a number (for example, 20 specifies 20% On-Demand Instances, 80% Spot Instances). If set to 100, only On-Demand Instances are used.

Default: 100

Note

An update to this setting means a gradual replacement of instances to adjust the current On-Demand and Spot Instance levels for your additional capacity higher than

the base capacity. When replacing instances, Amazon EC2 Auto Scaling launches new instances before terminating the previous ones.

Type: Integer

Required: No

SpotAllocationStrategy

The allocation strategy to apply to your Spot Instances when they are launched. Possible instance types are determined by the launch template overrides that you specify.

The following lists the valid values:

capacity-optimized

Requests Spot Instances using pools that are optimally chosen based on the available Spot capacity. This strategy has the lowest risk of interruption. To give certain instance types a higher chance of launching first, use `capacity-optimized-prioritized`.

capacity-optimized-prioritized

You set the order of instance types for the launch template overrides from highest to lowest priority (from first to last in the list). Amazon EC2 Auto Scaling honors the instance type priorities on a best effort basis but optimizes for capacity first. Note that if the On-Demand allocation strategy is set to `prioritized`, the same priority is applied when fulfilling On-Demand capacity. This is not a valid value for Auto Scaling groups that specify [InstanceRequirements](#).

lowest-price

Requests Spot Instances using the lowest priced pools within an Availability Zone, across the number of Spot pools that you specify for the `SpotInstancePools` property. To ensure that your desired capacity is met, you might receive Spot Instances from several pools. This is the default value, but it might lead to high interruption rates because this strategy only considers instance price and not available capacity.

price-capacity-optimized (recommended)

The price and capacity optimized allocation strategy looks at both price and capacity to select the Spot Instance pools that are the least likely to be interrupted and have the lowest possible price.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\x\n\t]*`

Required: No

SpotInstancePools

The number of Spot Instance pools across which to allocate your Spot Instances. The Spot pools are determined from the different instance types in the overrides. Valid only when the `SpotAllocationStrategy` is `lowest-price`. Value must be in the range of 1–20.

Default: 2

Type: Integer

Required: No

SpotMaxPrice

The maximum price per unit hour that you are willing to pay for a Spot Instance. If your maximum price is lower than the Spot price for the instance types that you selected, your Spot Instances are not launched. We do not recommend specifying a maximum price because it can lead to increased interruptions. When Spot Instances launch, you pay the current Spot price. To remove a maximum price that you previously set, include the property but specify an empty string ("") for the value.

Important

If you specify a maximum price, your instances will be interrupted more frequently than if you do not specify one.

Valid Range: Minimum value of 0.001

Type: String

Length Constraints: Minimum length of 0. Maximum length of 255.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LaunchConfiguration

Describes a launch configuration.

Contents

CreatedTime

The creation date and time for the launch configuration.

Type: Timestamp

Required: Yes

ImageId

The ID of the Amazon Machine Image (AMI) to use to launch your EC2 instances. For more information, see [Find a Linux AMI](#) in the *Amazon EC2 User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

InstanceType

The instance type for the instances. For information about available instance types, see [Available instance types](#) in the *Amazon EC2 User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

LaunchConfigurationName

The name of the launch configuration.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

AssociatePublicIpAddress

Specifies whether to assign a public IPv4 address to the group's instances. If the instance is launched into a default subnet, the default is to assign a public IPv4 address, unless you disabled the option to assign a public IPv4 address on the subnet. If the instance is launched into a nondefault subnet, the default is not to assign a public IPv4 address, unless you enabled the option to assign a public IPv4 address on the subnet. For more information, see [Provide network connectivity for your Auto Scaling instances using Amazon VPC](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

BlockDeviceMappings.member.N

The block device mapping entries that define the block devices to attach to the instances at launch. By default, the block devices specified in the block device mapping for the AMI are used. For more information, see [Block device mappings](#) in the *Amazon EC2 User Guide*.

Type: Array of [BlockDeviceMapping](#) objects

Required: No

ClassicLinkVPCId

Available for backward compatibility.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ClassicLinkVPCSecurityGroups.member.N

Available for backward compatibility.

Type: Array of strings

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

EbsOptimized

Specifies whether the launch configuration is optimized for EBS I/O (`true`) or not (`false`). For more information, see [Amazon EBS-optimized instances](#) in the *Amazon EC2 User Guide*.

Type: Boolean

Required: No

IamInstanceProfile

The name or the Amazon Resource Name (ARN) of the instance profile associated with the IAM role for the instance. The instance profile contains the IAM role. For more information, see [IAM role for applications that run on Amazon EC2 instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceMonitoring

Controls whether instances in this group are launched with detailed (`true`) or basic (`false`) monitoring.

For more information, see [Configure monitoring for Auto Scaling instances](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [InstanceMonitoring](#) object

Required: No

KernelId

The ID of the kernel associated with the AMI.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

KeyName

The name of the key pair.

For more information, see [Amazon EC2 key pairs and Amazon EC2 instances](#) in the *Amazon EC2 User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchConfigurationARN

The Amazon Resource Name (ARN) of the launch configuration.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MetadataOptions

The metadata options for the instances. For more information, see [Configure the instance metadata options](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: [InstanceMetadataOptions](#) object

Required: No

PlacementTenancy

The tenancy of the instance, either default or dedicated. An instance with dedicated tenancy runs on isolated, single-tenant hardware and can only be launched into a VPC.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RamdiskId

The ID of the RAM disk associated with the AMI.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SecurityGroups.member.N

A list that contains the security groups to assign to the instances in the Auto Scaling group. For more information, see [Control traffic to your AWS resources using security groups](#) in the *Amazon Virtual Private Cloud User Guide*.

Type: Array of strings

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SpotPrice

The maximum hourly price to be paid for any Spot Instance launched to fulfill the request. Spot Instances are launched when the price you specify exceeds the current Spot price. For more

information, see [Requesting Spot Instances for fault-tolerant and flexible applications](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Required: No

UserData

The user data to make available to the launched EC2 instances. For more information, see [Instance metadata and user data](#) in the *Amazon EC2 User Guide*. If you are using a command line tool, base64-encoding is performed for you, and you can load the text from a file. Otherwise, you must provide base64-encoded text. User data is limited to 16 KB.

Type: String

Length Constraints: Maximum length of 21847.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LaunchInstancesError

Contains details about errors encountered during instance launch attempts.

Contents

AvailabilityZone

The Availability Zone where the instance launch was attempted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

AvailabilityZoneId

The Availability Zone ID where the launch error occurred.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ErrorCode

The error code representing the type of error encountered (e.g., `InsufficientInstanceCapacity`).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ErrorMessage

A descriptive message providing details about the error encountered during the launch attempt.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

InstanceType

The instance type that failed to launch.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MarketType

The market type (On-Demand or Spot) that encountered the launch error.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SubnetId

The subnet ID where the instance launch was attempted.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LaunchTemplate

Use this structure to specify the launch templates and instance types (overrides) for a mixed instances policy.

When you update the launch template or overrides in the [UpdateAutoScalingGroup](#) API call, existing Amazon EC2 instances continue to run. When scaling out occurs, Amazon EC2 Auto Scaling launches instances to match the new settings. When scaling in occurs, Amazon EC2 Auto Scaling terminates instances according to the group's termination policies.

Contents

LaunchTemplateSpecification

The launch template.

Type: [LaunchTemplateSpecification](#) object

Required: No

Overrides.member.N

Any properties that you specify override the same properties in the launch template.

Type: Array of [LaunchTemplateOverrides](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LaunchTemplateOverrides

Use this structure to let Amazon EC2 Auto Scaling do the following when the Auto Scaling group has a mixed instances policy:

- Override the instance type that is specified in the launch template.
- Use multiple instance types.

Specify the instance types that you want, or define your instance requirements instead and let Amazon EC2 Auto Scaling provision the available instance types that meet your requirements. This can provide Amazon EC2 Auto Scaling with a larger selection of instance types to choose from when fulfilling Spot and On-Demand capacities. You can view which instance types are matched before you apply the instance requirements to your Auto Scaling group.

After you define your instance requirements, you don't have to keep updating these settings to get new EC2 instance types automatically. Amazon EC2 Auto Scaling uses the instance requirements of the Auto Scaling group to determine whether a new EC2 instance type can be used.

Contents

ImageId

The ID of the Amazon Machine Image (AMI) to use for instances launched with this override. When using Instance Refresh with ReplaceRootVolume strategy, this specifies the AMI for root volume replacement operations.

For ReplaceRootVolume operations:

- All overrides in the MixedInstancesPolicy must specify an ImageId
- The AMI must contain only a single root volume
- Root volume replacement doesn't support multi-volume AMIs

Type: String

Length Constraints: Minimum length of 5. Maximum length of 21.

Pattern: `^ami-[a-z0-9]{1,17}$`

Required: No

InstanceRequirements

The instance requirements. Amazon EC2 Auto Scaling uses your specified requirements to identify instance types. Then, it uses your On-Demand and Spot allocation strategies to launch instances from these instance types.

You can specify up to four separate sets of instance requirements per Auto Scaling group. This is useful for provisioning instances from different Amazon Machine Images (AMIs) in the same Auto Scaling group. To do this, create the AMIs and create a new launch template for each AMI. Then, create a compatible set of instance requirements for each launch template.

Note

If you specify `InstanceRequirements`, you can't specify `InstanceType`.

Type: [InstanceRequirements](#) object

Required: No

InstanceType

The instance type, such as `m3.xlarge`. You must specify an instance type that is supported in your requested Region and Availability Zones. For more information, see [Instance types](#) in the *Amazon EC2 User Guide*.

You can specify up to 40 instance types per Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchTemplateSpecification

Provides a launch template for the specified instance type or set of instance requirements. For example, some instance types might require a launch template with a different AMI. If not provided, Amazon EC2 Auto Scaling uses the launch template that's specified in the

LaunchTemplate definition. For more information, see [Specifying a different launch template for an instance type](#) in the *Amazon EC2 Auto Scaling User Guide*.

You can specify up to 20 launch templates per Auto Scaling group. The launch templates specified in the overrides and in the LaunchTemplate definition count towards this limit.

Type: [LaunchTemplateSpecification](#) object

Required: No

WeightedCapacity

If you provide a list of instance types to use, you can specify the number of capacity units provided by each instance type in terms of virtual CPUs, memory, storage, throughput, or other relative performance characteristic. When a Spot or On-Demand Instance is launched, the capacity units count toward the desired capacity. Amazon EC2 Auto Scaling launches instances until the desired capacity is totally fulfilled, even if this results in an overage. For example, if there are two units remaining to fulfill capacity, and Amazon EC2 Auto Scaling can only launch an instance with a WeightedCapacity of five units, the instance is launched, and the desired capacity is exceeded by three units. For more information, see [Configure an Auto Scaling group to use instance weights](#) in the *Amazon EC2 Auto Scaling User Guide*. Value must be in the range of 1–999.

If you specify a value for WeightedCapacity for one instance type, you must specify a value for WeightedCapacity for all of them.

Important

Every Auto Scaling group has three size parameters (DesiredCapacity, MaxSize, and MinSize). Usually, you set these sizes based on a specific number of instances. However, if you configure a mixed instances policy that defines weights for the instance types, you must specify these sizes with the same units that you use for weighting instances.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 3.

Pattern: `^[\u0031-\u0039][\u0030-\u0039]{0,2}$`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LaunchTemplateSpecification

Describes the launch template and the version of the launch template that Amazon EC2 Auto Scaling uses to launch Amazon EC2 instances. For more information about launch templates, see [Launch templates](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

LaunchTemplateId

The ID of the launch template. To get the template ID, use the Amazon EC2 [DescribeLaunchTemplates](#) API operation. New launch templates can be created using the Amazon EC2 [CreateLaunchTemplate](#) API.

Conditional: You must specify either a LaunchTemplateId or a LaunchTemplateName.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

LaunchTemplateName

The name of the launch template. To get the template name, use the Amazon EC2 [DescribeLaunchTemplates](#) API operation. New launch templates can be created using the Amazon EC2 [CreateLaunchTemplate](#) API.

Conditional: You must specify either a LaunchTemplateId or a LaunchTemplateName.

Type: String

Length Constraints: Minimum length of 3. Maximum length of 128.

Pattern: `[a-zA-Z0-9\(\)\.\-/_]+`

Required: No

Version

The version number, `$Latest`, or `$Default`. To get the version number, use the Amazon EC2 [DescribeLaunchTemplateVersions](#) API operation. New launch template versions can be created

using the Amazon EC2 [CreateLaunchTemplateVersion](#) API. If the value is `$Latest`, Amazon EC2 Auto Scaling selects the latest version of the launch template when launching instances. If the value is `$Default`, Amazon EC2 Auto Scaling selects the default version of the launch template when launching instances. The default value is `$Default`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LifecycleHook

Describes a lifecycle hook. A lifecycle hook lets you create solutions that are aware of events in the Auto Scaling instance lifecycle, and then perform a custom action on instances when the corresponding lifecycle event occurs.

Contents

AutoScalingGroupName

The name of the Auto Scaling group for the lifecycle hook.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

DefaultResult

The action the Auto Scaling group takes when the lifecycle hook timeout elapses or if an unexpected failure occurs.

Valid values: CONTINUE | ABANDON

Type: String

Required: No

GlobalTimeout

The maximum time, in seconds, that an instance can remain in a wait state. The maximum is 172800 seconds (48 hours) or 100 times HeartbeatTimeout, whichever is smaller.

Type: Integer

Required: No

HeartbeatTimeout

The maximum time, in seconds, that can elapse before the lifecycle hook times out. If the lifecycle hook times out, Amazon EC2 Auto Scaling performs the action that you specified in the DefaultResult property.

Type: Integer

Required: No

LifecycleHookName

The name of the lifecycle hook.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: [A-Za-z0-9\-_\/]+

Required: No

LifecycleTransition

The lifecycle transition.

Valid values: `autoscaling:EC2_INSTANCE_LAUNCHING` |
`autoscaling:EC2_INSTANCE_TERMINATING`

Type: String

Required: No

NotificationMetadata

Additional information that is included any time Amazon EC2 Auto Scaling sends a message to the notification target.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 4000.

Pattern: [\u0009\u000A\u000D\u0020-\u007e]+

Required: No

NotificationTargetARN

The ARN of the target that Amazon EC2 Auto Scaling sends notifications to when an instance is in a wait state for the lifecycle hook.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RoleARN

The ARN of the IAM role that allows the Auto Scaling group to publish to the specified notification target (an Amazon SNS topic or an Amazon SQS queue).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LifecycleHookSpecification

Describes information used to specify a lifecycle hook for an Auto Scaling group.

Lifecycle hooks let you create solutions that are aware of events in the Auto Scaling instance lifecycle, and then perform a custom action on instances when the corresponding lifecycle event occurs.

This step is a part of the procedure for creating a lifecycle hook for an Auto Scaling group:

1. (Optional) Create a launch template or launch configuration with a user data script that runs while an instance is in a wait state due to a lifecycle hook.
2. (Optional) Create a Lambda function and a rule that allows Amazon EventBridge to invoke your Lambda function when an instance is put into a wait state due to a lifecycle hook.
3. (Optional) Create a notification target and an IAM role. The target can be either an Amazon SQS queue or an Amazon SNS topic. The role allows Amazon EC2 Auto Scaling to publish lifecycle notifications to the target.
4. **Create the lifecycle hook. Specify whether the hook is used when the instances launch or terminate.**
5. If you need more time, record the lifecycle action heartbeat to keep the instance in a wait state by using the [RecordLifecycleActionHeartbeat](#) API call.
6. If you finish before the timeout period ends, send a callback by using the [CompleteLifecycleAction](#) API call.

For more information, see [Amazon EC2 Auto Scaling lifecycle hooks](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

LifecycleHookName

The name of the lifecycle hook.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[A-Za-z0-9\-\_\//]+`

Required: Yes

LifecycleTransition

The lifecycle transition. For Auto Scaling groups, there are two major lifecycle transitions.

- To create a lifecycle hook for scale-out events, specify `autoscaling:EC2_INSTANCE_LAUNCHING`.
- To create a lifecycle hook for scale-in events, specify `autoscaling:EC2_INSTANCE_TERMINATING`.

Type: String

Required: Yes

DefaultResult

The action the Auto Scaling group takes when the lifecycle hook timeout elapses or if an unexpected failure occurs. The default value is ABANDON.

Valid values: CONTINUE | ABANDON

Type: String

Required: No

HeartbeatTimeout

The maximum time, in seconds, that can elapse before the lifecycle hook times out. The range is from 30 to 7200 seconds. The default value is 3600 seconds (1 hour).

Type: Integer

Required: No

NotificationMetadata

Additional information that you want to include any time Amazon EC2 Auto Scaling sends a message to the notification target.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 4000.

Pattern: `[\u0009\u000A\u000D\u0020-\u007e]+`

Required: No

NotificationTargetARN

The Amazon Resource Name (ARN) of the notification target that Amazon EC2 Auto Scaling sends notifications to when an instance is in a wait state for the lifecycle hook. You can specify an Amazon SNS topic or an Amazon SQS queue.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RoleARN

The ARN of the IAM role that allows the Auto Scaling group to publish to the specified notification target. For information about creating this role, see [Prepare to add a lifecycle hook to your Auto Scaling group](#) in the *Amazon EC2 Auto Scaling User Guide*.

Valid only if the notification target is an Amazon SNS topic or an Amazon SQS queue.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LoadBalancerState

Describes the state of a Classic Load Balancer.

Contents

LoadBalancerName

The name of the load balancer.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF\\r\\n\\t]*`

Required: No

State

One of the following load balancer states:

- **Adding** - The Auto Scaling instances are being registered with the load balancer.
- **Added** - All Auto Scaling instances are registered with the load balancer.
- **InService** - At least one Auto Scaling instance passed an ELB health check.
- **Removing** - The Auto Scaling instances are being deregistered from the load balancer. If connection draining is enabled, Elastic Load Balancing waits for in-flight requests to complete before deregistering the instances.
- **Removed** - All Auto Scaling instances are deregistered from the load balancer.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LoadBalancerTargetGroupState

Describes the state of a target group.

Contents

LoadBalancerTargetGroupARN

The Amazon Resource Name (ARN) of the target group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\\u0020-\\uD7FF\\uE000-\\uFFFD\\uD800\\uDC00-\\uDBFF\\uDFFF\\r\\n\\t]*`

Required: No

State

The state of the target group.

- **Adding** - The Auto Scaling instances are being registered with the target group.
- **Added** - All Auto Scaling instances are registered with the target group.
- **InService** - At least one Auto Scaling instance passed an ELB health check.
- **Removing** - The Auto Scaling instances are being deregistered from the target group. If connection draining is enabled, Elastic Load Balancing waits for in-flight requests to complete before deregistering the instances.
- **Removed** - All Auto Scaling instances are deregistered from the target group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

LoadForecast

A `GetPredictiveScalingForecast` call returns the load forecast for a predictive scaling policy. This structure includes the data points for that load forecast, along with the timestamps of those data points and the metric specification.

Contents

MetricSpecification

The metric specification for the load forecast.

Type: [PredictiveScalingMetricSpecification](#) object

Required: Yes

Timestamps.member.N

The timestamps for the data points, in UTC format.

Type: Array of timestamps

Required: Yes

Values.member.N

The values of the data points.

Type: Array of doubles

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryGiBPerVCpuRequest

Specifies the minimum and maximum for the MemoryGiBPerVCpu object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Max

The memory maximum in GiB.

Type: Double

Valid Range: Minimum value of 0.

Required: No

Min

The memory minimum in GiB.

Type: Double

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MemoryMiBRequest

Specifies the minimum and maximum for the MemoryMiB object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Min

The memory minimum in MiB.

Type: Integer

Valid Range: Minimum value of 0.

Required: Yes

Max

The memory maximum in MiB.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Metric

Represents a specific metric.

Metric is a property of the [MetricStat](#) object.

Contents

MetricName

The name of the metric.

Type: String

Required: Yes

Namespace

The namespace of the metric. For more information, see the table in [AWS services that publish CloudWatch metrics](#) in the *Amazon CloudWatch User Guide*.

Type: String

Required: Yes

Dimensions.member.N

The dimensions for the metric. For the list of available dimensions, see the AWS documentation available from the table in [AWS services that publish CloudWatch metrics](#) in the *Amazon CloudWatch User Guide*.

Conditional: If you published your metric with dimensions, you must specify the same dimensions in your scaling policy.

Type: Array of [MetricDimension](#) objects

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetricCollectionType

Describes a metric.

Contents

Metric

One of the following metrics:

- GroupMinSize
- GroupMaxSize
- GroupDesiredCapacity
- GroupInServiceInstances
- GroupPendingInstances
- GroupStandbyInstances
- GroupTerminatingInstances
- GroupTotalInstances
- GroupInServiceCapacity
- GroupPendingCapacity
- GroupStandbyCapacity
- GroupTerminatingCapacity
- GroupTotalCapacity
- WarmPoolDesiredCapacity
- WarmPoolWarmedCapacity
- WarmPoolPendingCapacity
- WarmPoolTerminatingCapacity
- WarmPoolTotalCapacity
- GroupAndWarmPoolDesiredCapacity
- GroupAndWarmPoolTotalCapacity

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Conditional: Within each `MetricDataQuery` object, you must specify either `Expression` or `MetricStat`, but not both.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Label

A human-readable label for this metric or expression. This is especially useful if this is a math expression, so that you know what the value represents.

Type: String

Length Constraints: Maximum length of 2047.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MetricStat

Information about the metric data to return.

Conditional: Within each `MetricDataQuery` object, you must specify either `Expression` or `MetricStat`, but not both.

Type: [MetricStat](#) object

Required: No

ReturnData

Indicates whether to return the timestamps and raw data values of this metric.

If you use any math expressions, specify `true` for this value for only the final math expression that the metric specification is based on. You must specify `false` for `ReturnData` for all the other metrics and expressions used in the metric specification.

If you are only retrieving metrics and not performing any math expressions, do not specify anything for `ReturnData`. This sets it to its default (`true`).

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetricDimension

Describes the dimension of a metric.

Contents

Name

The name of the dimension.

Type: String

Required: Yes

Value

The value of the dimension.

Type: String

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetricGranularityType

Describes a granularity of a metric.

Contents

Granularity

The granularity. The only valid value is 1Minute.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MetricStat

This structure defines the CloudWatch metric to return, along with the statistic and unit.

MetricStat is a property of the [MetricDataQuery](#) object.

For more information about the CloudWatch terminology below, see [Amazon CloudWatch concepts](#) in the *Amazon CloudWatch User Guide*.

Contents

Metric

The CloudWatch metric to return, including the metric name, namespace, and dimensions. To get the exact metric name, namespace, and dimensions, inspect the [Metric](#) object that is returned by a call to [ListMetrics](#).

Type: [Metric](#) object

Required: Yes

Stat

The statistic to return. It can include any CloudWatch statistic or extended statistic. For a list of valid values, see the table in [Statistics](#) in the *Amazon CloudWatch User Guide*.

The most commonly used metrics for predictive scaling are Average and Sum.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 100.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

Unit

The unit to use for the returned data points. For a complete list of the units that CloudWatch supports, see the [MetricDatum](#) data type in the *Amazon CloudWatch API Reference*.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

MixedInstancesPolicy

Use this structure to launch multiple instance types and On-Demand Instances and Spot Instances within a single Auto Scaling group.

A mixed instances policy contains information that Amazon EC2 Auto Scaling can use to launch instances and help optimize your costs. For more information, see [Auto Scaling groups with multiple instance types and purchase options](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

InstancesDistribution

The instances distribution.

Type: [InstancesDistribution](#) object

Required: No

LaunchTemplate

One or more launch templates and the instance types (overrides) that are used to launch EC2 instances to fulfill On-Demand and Spot capacities.

Type: [LaunchTemplate](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

NetworkBandwidthGbpsRequest

Specifies the minimum and maximum for the `NetworkBandwidthGbps` object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Note

Setting the minimum bandwidth does not guarantee that your instance will achieve the minimum bandwidth. Amazon EC2 will identify instance types that support the specified minimum bandwidth, but the actual bandwidth of your instance might go below the specified minimum at times. For more information, see [Available instance bandwidth](#) in the *Amazon EC2 User Guide*.

Contents

Max

The maximum amount of network bandwidth, in gigabits per second (Gbps).

Type: Double

Valid Range: Minimum value of 0.

Required: No

Min

The minimum amount of network bandwidth, in gigabits per second (Gbps).

Type: Double

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

NetworkInterfaceCountRequest

Specifies the minimum and maximum for the `NetworkInterfaceCount` object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Max

The maximum number of network interfaces.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

Min

The minimum number of network interfaces.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

NotificationConfiguration

Describes a notification.

Contents

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

NotificationType

One of the following event notification types:

- `autoscaling:EC2_INSTANCE_LAUNCH`
- `autoscaling:EC2_INSTANCE_LAUNCH_ERROR`
- `autoscaling:EC2_INSTANCE_TERMINATE`
- `autoscaling:EC2_INSTANCE_TERMINATE_ERROR`
- `autoscaling:TEST_NOTIFICATION`

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

TopicARN

The Amazon Resource Name (ARN) of the Amazon SNS topic.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PerformanceFactorReferenceRequest

Specify an instance family to use as the baseline reference for CPU performance. All instance types that All instance types that match your specified attributes will be compared against the CPU performance of the referenced instance family, regardless of CPU manufacturer or architecture differences.

Note

Currently only one instance family can be specified in the list.

Contents

InstanceFamily

The instance family to use as a baseline reference.

Note

Make sure that you specify the correct value for the instance family. The instance family is everything before the period (.) in the instance type name. For example, in the instance `c6i.large`, the instance family is `c6i`, not `c6`. For more information, see [Amazon EC2 instance type naming conventions](#) in *Amazon EC2 Instance Types*.

The following instance types are *not supported* for performance protection.

- `c1`
- `g3` | `g3s`
- `hpc7g`
- `m1` | `m2`
- `mac1` | `mac2` | `mac2-m1ultra` | `mac2-m2` | `mac2-m2pro`
- `p3dn` | `p4d` | `p5`
- `t1`
- `u-12tb1` | `u-18tb1` | `u-24tb1` | `u-3tb1` | `u-6tb1` | `u-9tb1` | `u7i-12tb` | `u7in-16tb` | `u7in-24tb` | `u7in-32tb`

If you performance protection by specifying a supported instance family, the returned instance types will exclude the preceding unsupported instance families.

If you specify an unsupported instance family as a value for baseline performance, the API returns an empty response.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredefinedMetricSpecification

Represents a predefined metric for a target tracking scaling policy to use with Amazon EC2 Auto Scaling.

Contents

PredefinedMetricType

The metric type. The following predefined metrics are available:

- `ASGAverageCPUUtilization` - Average CPU utilization of the Auto Scaling group.
- `ASGAverageNetworkIn` - Average number of bytes received on all network interfaces by the Auto Scaling group.
- `ASGAverageNetworkOut` - Average number of bytes sent out on all network interfaces by the Auto Scaling group.
- `ALBRequestCountPerTarget` - Average Application Load Balancer request count per target for your Auto Scaling group.

Type: String

Valid Values: `ASGAverageCPUUtilization` | `ASGAverageNetworkIn` | `ASGAverageNetworkOut` | `ALBRequestCountPerTarget`

Required: Yes

ResourceLabel

A label that uniquely identifies a specific Application Load Balancer target group from which to determine the average request count served by your Auto Scaling group. You can't specify a resource label unless the target group is attached to the Auto Scaling group.

You create the resource label by appending the final portion of the load balancer ARN and the final portion of the target group ARN into a single value, separated by a forward slash (/). The format of the resource label is:

```
app/my-alb/778d41231b141a0f/targetgroup/my-alb-target-group/943f017f100becff.
```

Where:

- `app/<load-balancer-name>/<load-balancer-id>` is the final portion of the load balancer ARN
- `targetgroup/<target-group-name>/<target-group-id>` is the final portion of the target group ARN.

To find the ARN for an Application Load Balancer, use the [DescribeLoadBalancers](#) API operation. To find the ARN for the target group, use the [DescribeTargetGroups](#) API operation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingConfiguration

Represents a predictive scaling policy configuration to use with Amazon EC2 Auto Scaling.

Contents

MetricSpecifications.member.N

This structure includes the metrics and target utilization to use for predictive scaling.

This is an array, but we currently only support a single metric specification. That is, you can specify a target value and a single metric pair, or a target value and one scaling metric and one load metric.

Type: Array of [PredictiveScalingMetricSpecification](#) objects

Required: Yes

MaxCapacityBreachBehavior

Defines the behavior that should be applied if the forecast capacity approaches or exceeds the maximum capacity of the Auto Scaling group. Defaults to `HonorMaxCapacity` if not specified.

The following are possible values:

- `HonorMaxCapacity` - Amazon EC2 Auto Scaling can't increase the maximum capacity of the group when the forecast capacity is close to or exceeds the maximum capacity.
- `IncreaseMaxCapacity` - Amazon EC2 Auto Scaling can increase the maximum capacity of the group when the forecast capacity is close to or exceeds the maximum capacity. The upper limit is determined by the forecasted capacity and the value for `MaxCapacityBuffer`.

Important

Use caution when allowing the maximum capacity to be automatically increased. This can lead to more instances being launched than intended if the increased maximum capacity is not monitored and managed. The increased maximum capacity then becomes the new normal maximum capacity for the Auto Scaling group until you manually update it. The maximum capacity does not automatically decrease back to the original maximum.

Type: String

Valid Values: HonorMaxCapacity | IncreaseMaxCapacity

Required: No

MaxCapacityBuffer

The size of the capacity buffer to use when the forecast capacity is close to or exceeds the maximum capacity. The value is specified as a percentage relative to the forecast capacity. For example, if the buffer is 10, this means a 10 percent buffer, such that if the forecast capacity is 50, and the maximum capacity is 40, then the effective maximum capacity is 55.

If set to 0, Amazon EC2 Auto Scaling may scale capacity higher than the maximum capacity to equal but not exceed forecast capacity.

Required if the MaxCapacityBreachBehavior property is set to IncreaseMaxCapacity, and cannot be used otherwise.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

Mode

The predictive scaling mode. Defaults to ForecastOnly if not specified.

Type: String

Valid Values: ForecastAndScale | ForecastOnly

Required: No

SchedulingBufferTime

The amount of time, in seconds, by which the instance launch time can be advanced. For example, the forecast says to add capacity at 10:00 AM, and you choose to pre-launch instances by 5 minutes. In that case, the instances will be launched at 9:55 AM. The intention is to give resources time to be provisioned. It can take a few minutes to launch an EC2 instance. The actual amount of time required depends on several factors, such as the size of the instance and whether there are startup scripts to complete.

The value must be less than the forecast interval duration of 3600 seconds (60 minutes). Defaults to 300 seconds if not specified.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingCustomizedCapacityMetric

Describes a customized capacity metric for a predictive scaling policy.

Contents

MetricDataQueries.member.N

One or more metric data queries to provide the data points for a capacity metric. Use multiple metric data queries only if you are performing a math expression on returned data.

Type: Array of [MetricDataQuery](#) objects

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingCustomizedLoadMetric

Describes a custom load metric for a predictive scaling policy.

Contents

MetricDataQueries.member.N

One or more metric data queries to provide the data points for a load metric. Use multiple metric data queries only if you are performing a math expression on returned data.

Type: Array of [MetricDataQuery](#) objects

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingCustomizedScalingMetric

Describes a custom scaling metric for a predictive scaling policy.

Contents

MetricDataQueries.member.N

One or more metric data queries to provide the data points for a scaling metric. Use multiple metric data queries only if you are performing a math expression on returned data.

Type: Array of [MetricDataQuery](#) objects

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingMetricSpecification

This structure specifies the metrics and target utilization settings for a predictive scaling policy.

You must specify either a metric pair, or a load metric and a scaling metric individually. Specifying a metric pair instead of individual metrics provides a simpler way to configure metrics for a scaling policy. You choose the metric pair, and the policy automatically knows the correct sum and average statistics to use for the load metric and the scaling metric.

Example

- You create a predictive scaling policy and specify `ALBRequestCount` as the value for the metric pair and `1000.0` as the target value. For this type of metric, you must provide the metric dimension for the corresponding target group, so you also provide a resource label for the Application Load Balancer target group that is attached to your Auto Scaling group.
- The number of requests the target group receives per minute provides the load metric, and the request count averaged between the members of the target group provides the scaling metric. In CloudWatch, this refers to the `RequestCount` and `RequestCountPerTarget` metrics, respectively.
- For optimal use of predictive scaling, you adhere to the best practice of using a dynamic scaling policy to automatically scale between the minimum capacity and maximum capacity in response to real-time changes in resource utilization.
- Amazon EC2 Auto Scaling consumes data points for the load metric over the last 14 days and creates an hourly load forecast for predictive scaling. (A minimum of 24 hours of data is required.)
- After creating the load forecast, Amazon EC2 Auto Scaling determines when to reduce or increase the capacity of your Auto Scaling group in each hour of the forecast period so that the average number of requests received by each instance is as close to 1000 requests per minute as possible at all times.

For information about using custom metrics with predictive scaling, see [Advanced predictive scaling policy configurations using custom metrics](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

TargetValue

Specifies the target utilization.

Note

Some metrics are based on a count instead of a percentage, such as the request count for an Application Load Balancer or the number of messages in an SQS queue. If the scaling policy specifies one of these metrics, specify the target utilization as the optimal average request or message count per instance during any one-minute interval.

Type: Double

Required: Yes

CustomizedCapacityMetricSpecification

The customized capacity metric specification.

Type: [PredictiveScalingCustomizedCapacityMetric](#) object

Required: No

CustomizedLoadMetricSpecification

The customized load metric specification.

Type: [PredictiveScalingCustomizedLoadMetric](#) object

Required: No

CustomizedScalingMetricSpecification

The customized scaling metric specification.

Type: [PredictiveScalingCustomizedScalingMetric](#) object

Required: No

PredefinedLoadMetricSpecification

The predefined load metric specification.

Type: [PredictiveScalingPredefinedLoadMetric](#) object

Required: No

PredefinedMetricPairSpecification

The predefined metric pair specification from which Amazon EC2 Auto Scaling determines the appropriate scaling metric and load metric to use.

Type: [PredictiveScalingPredefinedMetricPair](#) object

Required: No

PredefinedScalingMetricSpecification

The predefined scaling metric specification.

Type: [PredictiveScalingPredefinedScalingMetric](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingPredefinedLoadMetric

Describes a load metric for a predictive scaling policy.

When returned in the output of `DescribePolicies`, it indicates that a predictive scaling policy uses individually specified load and scaling metrics instead of a metric pair.

Contents

PredefinedMetricType

The metric type.

Type: String

Valid Values: `ASGTotalCPUUtilization` | `ASGTotalNetworkIn` | `ASGTotalNetworkOut` | `ALBTargetGroupRequestCount`

Required: Yes

ResourceLabel

A label that uniquely identifies a specific Application Load Balancer target group from which to determine the request count served by your Auto Scaling group. You can't specify a resource label unless the target group is attached to the Auto Scaling group.

You create the resource label by appending the final portion of the load balancer ARN and the final portion of the target group ARN into a single value, separated by a forward slash (/). The format of the resource label is:

```
app/my-alb/778d41231b141a0f/targetgroup/my-alb-target-group/943f017f100becff.
```

Where:

- `app/<load-balancer-name>/<load-balancer-id>` is the final portion of the load balancer ARN
- `targetgroup/<target-group-name>/<target-group-id>` is the final portion of the target group ARN.

To find the ARN for an Application Load Balancer, use the [DescribeLoadBalancers](#) API operation. To find the ARN for the target group, use the [DescribeTargetGroups](#) API operation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingPredefinedMetricPair

Represents a metric pair for a predictive scaling policy.

Contents

PredefinedMetricType

Indicates which metrics to use. There are two different types of metrics for each metric type: one is a load metric and one is a scaling metric. For example, if the metric type is `ASGCPUtilization`, the Auto Scaling group's total CPU metric is used as the load metric, and the average CPU metric is used for the scaling metric.

Type: String

Valid Values: `ASGCPUtilization` | `ASGNetworkIn` | `ASGNetworkOut` | `ALBRequestCount`

Required: Yes

ResourceLabel

A label that uniquely identifies a specific Application Load Balancer target group from which to determine the total and average request count served by your Auto Scaling group. You can't specify a resource label unless the target group is attached to the Auto Scaling group.

You create the resource label by appending the final portion of the load balancer ARN and the final portion of the target group ARN into a single value, separated by a forward slash (/). The format of the resource label is:

```
app/my-alb/778d41231b141a0f/targetgroup/my-alb-target-group/943f017f100becff.
```

Where:

- `app/<load-balancer-name>/<load-balancer-id>` is the final portion of the load balancer ARN
- `targetgroup/<target-group-name>/<target-group-id>` is the final portion of the target group ARN.

To find the ARN for an Application Load Balancer, use the [DescribeLoadBalancers](#) API operation. To find the ARN for the target group, use the [DescribeTargetGroups](#) API operation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

PredictiveScalingPredefinedScalingMetric

Describes a scaling metric for a predictive scaling policy.

When returned in the output of `DescribePolicies`, it indicates that a predictive scaling policy uses individually specified load and scaling metrics instead of a metric pair.

Contents

PredefinedMetricType

The metric type.

Type: String

Valid Values: `ASGAverageCPUUtilization` | `ASGAverageNetworkIn` | `ASGAverageNetworkOut` | `ALBRequestCountPerTarget`

Required: Yes

ResourceLabel

A label that uniquely identifies a specific Application Load Balancer target group from which to determine the average request count served by your Auto Scaling group. You can't specify a resource label unless the target group is attached to the Auto Scaling group.

You create the resource label by appending the final portion of the load balancer ARN and the final portion of the target group ARN into a single value, separated by a forward slash (/). The format of the resource label is:

```
app/my-alb/778d41231b141a0f/targetgroup/my-alb-target-group/943f017f100becff.
```

Where:

- `app/<load-balancer-name>/<load-balancer-id>` is the final portion of the load balancer ARN
- `targetgroup/<target-group-name>/<target-group-id>` is the final portion of the target group ARN.

To find the ARN for an Application Load Balancer, use the [DescribeLoadBalancers](#) API operation. To find the ARN for the target group, use the [DescribeTargetGroups](#) API operation.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ProcessType

Describes a process type.

For more information, see [Types of processes](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

ProcessName

One of the following processes:

- Launch
- Terminate
- AddToLoadBalancer
- AlarmNotification
- AZRebalance
- HealthCheck
- InstanceRefresh
- ReplaceUnhealthy
- ScheduledActions

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RefreshPreferences

Describes the preferences for an instance refresh.

Contents

AlarmSpecification

(Optional) The CloudWatch alarm specification. CloudWatch alarms can be used to identify any issues and fail the operation if an alarm threshold is met.

Type: [AlarmSpecification](#) object

Required: No

AutoRollback

(Optional) Indicates whether to roll back the Auto Scaling group to its previous configuration if the instance refresh fails or a CloudWatch alarm threshold is met. The default is `false`.

A rollback is not supported in the following situations:

- There is no desired configuration specified for the instance refresh.
- The Auto Scaling group has a launch template that uses an AWS Systems Manager parameter instead of an AMI ID for the `ImageId` property.
- The Auto Scaling group uses the launch template's `$Latest` or `$Default` version.

For more information, see [Undo changes with a rollback](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

BakeTime

The amount of time, in seconds, to wait at the end of an instance refresh before the instance refresh is considered complete.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 172800.

Required: No

CheckpointDelay

(Optional) The amount of time, in seconds, to wait after a checkpoint before continuing. This property is optional, but if you specify a value for it, you must also specify a value for CheckpointPercentages. If you specify a value for CheckpointPercentages and not for CheckpointDelay, the CheckpointDelay defaults to 3600 (1 hour).

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 172800.

Required: No

CheckpointPercentages.member.N

(Optional) Threshold values for each checkpoint in ascending order. Each number must be unique. To replace all instances in the Auto Scaling group, the last number in the array must be 100.

For usage examples, see [Add checkpoints to an instance refresh](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Array of integers

Valid Range: Minimum value of 1. Maximum value of 100.

Required: No

InstanceWarmup

A time period, in seconds, during which an instance refresh waits before moving on to replacing the next instance after a new instance enters the InService state.

This property is not required for normal usage. Instead, use the DefaultInstanceWarmup property of the Auto Scaling group. The InstanceWarmup and DefaultInstanceWarmup properties work the same way. Only specify this property if you must override the DefaultInstanceWarmup property.

If you do not specify this property, the instance warmup by default is the value of the DefaultInstanceWarmup property, if defined (which is recommended in all cases), or the HealthCheckGracePeriod property otherwise.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

MaxHealthyPercentage

Specifies the maximum percentage of the group that can be in service and healthy, or pending, to support your workload when replacing instances. The value is expressed as a percentage of the desired capacity of the Auto Scaling group. Value range is 100 to 200.

If you specify `MaxHealthyPercentage`, you must also specify `MinHealthyPercentage`, and the difference between them cannot be greater than 100. A larger range increases the number of instances that can be replaced at the same time.

If you do not specify this property, the default is 100 percent, or the percentage set in the instance maintenance policy for the Auto Scaling group, if defined.

Type: Integer

Valid Range: Minimum value of 100. Maximum value of 200.

Required: No

MinHealthyPercentage

Specifies the minimum percentage of the group to keep in service, healthy, and ready to use to support your workload to allow the operation to continue. The value is expressed as a percentage of the desired capacity of the Auto Scaling group. Value range is 0 to 100.

If you do not specify this property, the default is 90 percent, or the percentage set in the instance maintenance policy for the Auto Scaling group, if defined.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

ScaleInProtectedInstances

Choose the behavior that you want Amazon EC2 Auto Scaling to use if instances protected from scale in are found.

The following lists the valid values:

Refresh

Amazon EC2 Auto Scaling replaces instances that are protected from scale in.

Ignore

Amazon EC2 Auto Scaling ignores instances that are protected from scale in and continues to replace instances that are not protected.

Wait (default)

Amazon EC2 Auto Scaling waits one hour for you to remove scale-in protection. Otherwise, the instance refresh will fail.

Type: String

Valid Values: Refresh | Ignore | Wait

Required: No

SkipMatching

(Optional) Indicates whether skip matching is enabled. If enabled (`true`), then Amazon EC2 Auto Scaling skips replacing instances that match the desired configuration. If no desired configuration is specified, then it skips replacing instances that have the same launch template and instance types that the Auto Scaling group was using before the start of the instance refresh. The default is `false`.

For more information, see [Use an instance refresh with skip matching](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: Boolean

Required: No

StandbyInstances

Choose the behavior that you want Amazon EC2 Auto Scaling to use if instances in Standby state are found.

The following lists the valid values:

Terminate

Amazon EC2 Auto Scaling terminates instances that are in Standby.

Ignore

Amazon EC2 Auto Scaling ignores instances that are in `Standby` and continues to replace instances that are in the `InService` state.

Wait (default)

Amazon EC2 Auto Scaling waits one hour for you to return the instances to service. Otherwise, the instance refresh will fail.

Type: String

Valid Values: `Terminate` | `Ignore` | `Wait`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RetentionTriggers

Defines the specific triggers that cause instances to be retained in a Retained state rather than terminated. Each trigger corresponds to a different failure scenario during the instance lifecycle. This allows fine-grained control over when to preserve instances for manual intervention.

Contents

TerminateHookAbandon

Specifies the action when a termination lifecycle hook is abandoned due to failure, timeout, or explicit abandonment (calling `CompleteLifecycleAction`).

Set to `Retain` to move instances to a Retained state. Set to `Terminate` for default termination behavior.

Retained instances don't count toward desired capacity and remain until you call `TerminateInstanceInAutoScalingGroup`.

Type: String

Valid Values: `retain` | `terminate`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

RollbackDetails

Details about an instance refresh rollback.

Contents

InstancesToUpdateOnRollback

Indicates the value of `InstancesToUpdate` at the time the rollback started.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PercentageCompleteOnRollback

Indicates the value of `PercentageComplete` at the time the rollback started.

Type: Integer

Valid Range: Minimum value of 0. Maximum value of 100.

Required: No

ProgressDetailsOnRollback

Reports progress on replacing instances in an Auto Scaling group that has a warm pool. This includes separate details for instances in the warm pool and instances in the Auto Scaling group (the live pool).

Type: [InstanceRefreshProgressDetails](#) object

Required: No

RollbackReason

The reason for this instance refresh rollback (for example, whether a manual or automatic rollback was initiated).

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1023.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

RollbackStartTime

The date and time at which the rollback began.

Type: Timestamp

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ScalingPolicy

Describes a scaling policy.

Contents

AdjustmentType

Specifies how the scaling adjustment is interpreted (for example, an absolute number or a percentage). The valid values are `ChangeInCapacity`, `ExactCapacity`, and `PercentChangeInCapacity`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Alarms.member.N

The CloudWatch alarms related to the policy.

Type: Array of [Alarm](#) objects

Required: No

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\u007F\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Cooldown

The duration of the policy's cooldown period, in seconds.

Type: Integer

Required: No

Enabled

Indicates whether the policy is enabled (`true`) or disabled (`false`).

Type: Boolean

Required: No

EstimatedInstanceWarmup

The estimated time, in seconds, until a newly launched instance can contribute to the CloudWatch metrics.

Type: Integer

Required: No

MetricAggregationType

The aggregation type for the CloudWatch metrics. The valid values are Minimum, Maximum, and Average.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 32.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MinAdjustmentMagnitude

The minimum value to scale by when the adjustment type is `PercentChangeInCapacity`.

Type: Integer

Required: No

MinAdjustmentStep

This member has been deprecated.

Available for backward compatibility. Use `MinAdjustmentMagnitude` instead.

Type: Integer

Required: No

PolicyARN

The Amazon Resource Name (ARN) of the policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

PolicyName

The name of the scaling policy.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

PolicyType

One of the following policy types:

- TargetTrackingScaling
- StepScaling
- SimpleScaling (default)
- PredictiveScaling

For more information, see [Target tracking scaling policies](#) and [Step and simple scaling policies](#) in the *Amazon EC2 Auto Scaling User Guide*.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 64.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

PredictiveScalingConfiguration

A predictive scaling policy.

Type: [PredictiveScalingConfiguration](#) object

Required: No

ScalingAdjustment

The amount by which to scale, based on the specified adjustment type. A positive value adds to the current capacity while a negative number removes from the current capacity.

Type: Integer

Required: No

StepAdjustments.member.N

A set of adjustments that enable you to scale based on the size of the alarm breach.

Type: Array of [StepAdjustment](#) objects

Required: No

TargetTrackingConfiguration

A target tracking scaling policy.

Type: [TargetTrackingConfiguration](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ScheduledUpdateGroupAction

Describes a scheduled scaling action.

Contents

AutoScalingGroupName

The name of the Auto Scaling group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

DesiredCapacity

The desired capacity is the initial capacity of the Auto Scaling group after the scheduled action runs and the capacity it attempts to maintain.

Type: Integer

Required: No

EndTime

The date and time in UTC for the recurring schedule to end. For example, "2019-06-01T00:00:00Z".

Type: Timestamp

Required: No

MaxSize

The maximum size of the Auto Scaling group.

Type: Integer

Required: No

MinSize

The minimum size of the Auto Scaling group.

Type: Integer

Required: No

Recurrence

The recurring schedule for the action, in Unix cron syntax format.

When `StartTime` and `EndTime` are specified with `Recurrence`, they form the boundaries of when the recurring action starts and stops.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ScheduledActionARN

The Amazon Resource Name (ARN) of the scheduled action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 1600.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ScheduledActionName

The name of the scheduled action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

StartTime

The date and time in UTC for this action to start. For example, "2019-06-01T00:00:00Z".

Type: Timestamp

Required: No

Time

This property is no longer used.

Type: Timestamp

Required: No

TimeZone

The time zone for the cron expression.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

ScheduledUpdateGroupActionRequest

Describes information used for one or more scheduled scaling action updates in a [BatchPutScheduledUpdateGroupAction](#) operation.

Contents

ScheduledActionName

The name of the scaling action.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

DesiredCapacity

The desired capacity is the initial capacity of the Auto Scaling group after the scheduled action runs and the capacity it attempts to maintain.

Type: Integer

Required: No

EndTime

The date and time for the recurring schedule to end, in UTC.

Type: Timestamp

Required: No

MaxSize

The maximum size of the Auto Scaling group.

Type: Integer

Required: No

Valid values are the canonical names of the IANA time zones, derived from the IANA Time Zone Database (such as Etc/GMT+9 or Pacific/Tahiti). For more information, see https://en.wikipedia.org/wiki/List_of_tz_database_time_zones.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

StepAdjustment

Describes information used to create a step adjustment for a step scaling policy.

For the following examples, suppose that you have an alarm with a breach threshold of 50:

- To trigger the adjustment when the metric is greater than or equal to 50 and less than 60, specify a lower bound of 0 and an upper bound of 10.
- To trigger the adjustment when the metric is greater than 40 and less than or equal to 50, specify a lower bound of -10 and an upper bound of 0.

There are a few rules for the step adjustments for your step policy:

- The ranges of your step adjustments can't overlap or have a gap.
- At most, one step adjustment can have a null lower bound. If one step adjustment has a negative lower bound, then there must be a step adjustment with a null lower bound.
- At most, one step adjustment can have a null upper bound. If one step adjustment has a positive upper bound, then there must be a step adjustment with a null upper bound.
- The upper and lower bound can't be null in the same step adjustment.

For more information, see [Step adjustments](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

ScalingAdjustment

The amount by which to scale, based on the specified adjustment type. A positive value adds to the current capacity while a negative number removes from the current capacity. For exact capacity, you must specify a non-negative value.

Type: Integer

Required: Yes

MetricIntervalLowerBound

The lower bound for the difference between the alarm threshold and the CloudWatch metric. If the metric value is above the breach threshold, the lower bound is inclusive (the metric must be greater than or equal to the threshold plus the lower bound). Otherwise, it is exclusive

(the metric must be greater than the threshold plus the lower bound). A null value indicates negative infinity.

Type: Double

Required: No

MetricIntervalUpperBound

The upper bound for the difference between the alarm threshold and the CloudWatch metric. If the metric value is above the breach threshold, the upper bound is exclusive (the metric must be less than the threshold plus the upper bound). Otherwise, it is inclusive (the metric must be less than or equal to the threshold plus the upper bound). A null value indicates positive infinity.

The upper bound must be greater than the lower bound.

Type: Double

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

SuspendedProcess

Describes an auto scaling process that has been suspended.

For more information, see [Types of processes](#) in the *Amazon EC2 Auto Scaling User Guide*.

Contents

ProcessName

The name of the suspended process.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

SuspensionReason

The reason that the process was suspended.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Tag

Describes a tag for an Auto Scaling group.

Contents

Key

The tag key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: Yes

PropagateAtLaunch

Determines whether the tag is added to new instances as they are launched in the group.

Type: Boolean

Required: No

ResourceId

The name of the Auto Scaling group.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ResourceType

The type of resource. The only supported value is `auto-scaling-group`.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Value

The tag value.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TagDescription

Describes a tag for an Auto Scaling group.

Contents

Key

The tag key.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 128.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

PropagateAtLaunch

Determines whether the tag is added to new instances as they are launched in the group.

Type: Boolean

Required: No

ResourceId

The name of the group.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

ResourceType

The type of resource. The only supported value is auto-scaling-group.

Type: String

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Value

The tag value.

Type: String

Length Constraints: Minimum length of 0. Maximum length of 256.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TargetTrackingConfiguration

Represents a target tracking scaling policy configuration to use with Amazon EC2 Auto Scaling.

Contents

TargetValue

The target value for the metric.

Note

Some metrics are based on a count instead of a percentage, such as the request count for an Application Load Balancer or the number of messages in an SQS queue. If the scaling policy specifies one of these metrics, specify the target utilization as the optimal average request or message count per instance during any one-minute interval.

Type: Double

Required: Yes

CustomizedMetricSpecification

A customized metric. You must specify either a predefined metric or a customized metric.

Type: [CustomizedMetricSpecification](#) object

Required: No

DisableScaleIn

Indicates whether scaling in by the target tracking scaling policy is disabled. If scaling in is disabled, the target tracking scaling policy doesn't remove instances from the Auto Scaling group. Otherwise, the target tracking scaling policy can remove instances from the Auto Scaling group. The default is `false`.

Type: Boolean

Required: No

PredefinedMetricSpecification

A predefined metric. You must specify either a predefined metric or a customized metric.

Type: [PredefinedMetricSpecification](#) object

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Type: String

Length Constraints: Minimum length of 1. Maximum length of 2047.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Label

A human-readable label for this metric or expression. This is especially useful if this is a math expression, so that you know what the value represents.

Type: String

Length Constraints: Maximum length of 2047.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

MetricStat

Information about the metric data to return.

Conditional: Within each `TargetTrackingMetricDataQuery` object, you must specify either `Expression` or `MetricStat`, but not both.

Type: [TargetTrackingMetricStat](#) object

Required: No

Period

The period of the metric in seconds. The default value is 60. Accepted values are 10, 30, and 60. For high resolution metric, set the value to less than 60. For more information, see [Create a target tracking policy using high-resolution metrics for faster response](#).

Type: Integer

Valid Range: Minimum value of 1.

Required: No

ReturnData

Indicates whether to return the timestamps and raw data values of this metric.

If you use any math expressions, specify `true` for this value for only the final math expression that the metric specification is based on. You must specify `false` for `ReturnData` for all the other metrics and expressions used in the metric specification.

If you are only retrieving metrics and not performing any math expressions, do not specify anything for `ReturnData`. This sets it to its default (`true`).

Type: Boolean

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Required: No

Unit

The unit to use for the returned data points. For a complete list of the units that CloudWatch supports, see the [MetricDatum](#) data type in the *Amazon CloudWatch API Reference*.

Type: String

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TotalLocalStorageGBRequest

Specifies the minimum and maximum for the TotalLocalStorageGB object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Max

The storage maximum in GB.

Type: Double

Valid Range: Minimum value of 0.

Required: No

Min

The storage minimum in GB.

Type: Double

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

The following lists the valid values:

- `elb` if `Identifier` is the name of a Classic Load Balancer.
- `elbv2` if `Identifier` is the ARN of an Application Load Balancer, Gateway Load Balancer, or Network Load Balancer target group.
- `vpc-lattice` if `Identifier` is the ARN of a VPC Lattice target group.

Required if the identifier is the name of a Classic Load Balancer.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

TrafficSourceState

Describes the state of a traffic source.

Contents

Identifier

The unique identifier of the traffic source.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

State

Describes the current state of a traffic source.

The state values are as follows:

- **Adding** - The Auto Scaling instances are being registered with the load balancer or target group.
- **Added** - All Auto Scaling instances are registered with the load balancer or target group.
- **InService** - For an Elastic Load Balancing load balancer or target group, at least one Auto Scaling instance passed an ELB health check. For VPC Lattice, at least one Auto Scaling instance passed an VPC_LATTICE health check.
- **Removing** - The Auto Scaling instances are being deregistered from the load balancer or target group. If connection draining (deregistration delay) is enabled, Elastic Load Balancing or VPC Lattice waits for in-flight requests to complete before deregistering the instances.
- **Removed** - All Auto Scaling instances are deregistered from the load balancer or target group.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 255.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFF\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

TrafficSource

This member has been deprecated.

This is replaced by `Identifier`.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

Type

Provides additional context for the value of `Identifier`.

The following lists the valid values:

- `elb` if `Identifier` is the name of a Classic Load Balancer.
- `elbv2` if `Identifier` is the ARN of an Application Load Balancer, Gateway Load Balancer, or Network Load Balancer target group.
- `vpc-lattice` if `Identifier` is the ARN of a VPC Lattice target group.

Required if the identifier is the name of a Classic Load Balancer.

Type: String

Length Constraints: Minimum length of 1. Maximum length of 511.

Pattern: `[\u0020-\uD7FF\uE000-\uFFFD\uD800\uDC00-\uDBFF\uDFFF\r\n\t]*`

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)

- [AWS SDK for Ruby V3](#)

VCpuCountRequest

Specifies the minimum and maximum for the VCpuCount object when you specify [InstanceRequirements](#) for an Auto Scaling group.

Contents

Min

The minimum number of vCPUs.

Type: Integer

Valid Range: Minimum value of 0.

Required: Yes

Max

The maximum number of vCPUs.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

WarmPoolConfiguration

Describes a warm pool configuration.

Contents

InstanceReusePolicy

The instance reuse policy.

Type: [InstanceReusePolicy](#) object

Required: No

MaxGroupPreparedCapacity

The maximum number of instances that are allowed to be in the warm pool or in any state except Terminated for the Auto Scaling group.

Type: Integer

Valid Range: Minimum value of -1.

Required: No

MinSize

The minimum number of instances to maintain in the warm pool.

Type: Integer

Valid Range: Minimum value of 0.

Required: No

PoolState

The instance state to transition to after the lifecycle actions are complete.

Type: String

Valid Values: Stopped | Running | Hibernated

Required: No

Status

The status of a warm pool that is marked for deletion.

Type: String

Valid Values: PendingDelete

Required: No

See Also

For more information about using this API in one of the language-specific AWS SDKs, see the following:

- [AWS SDK for C++](#)
- [AWS SDK for Java V2](#)
- [AWS SDK for Ruby V3](#)

Common Parameters

The following list contains the parameters that all actions use for signing Signature Version 4 requests with a query string. Any action-specific parameters are listed in the topic for that action. For more information about Signature Version 4, see [Signing AWS API requests](#) in the *IAM User Guide*.

Action

The action to be performed.

Type: string

Required: Yes

Version

The API version that the request is written for, expressed in the format YYYY-MM-DD.

Type: string

Required: Yes

X-Amz-Algorithm

The hash algorithm that you used to create the request signature.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Valid Values: AWS4-HMAC-SHA256

Required: Conditional

X-Amz-Credential

The credential scope value, which is a string that includes your access key, the date, the region you are targeting, the service you are requesting, and a termination string ("aws4_request"). The value is expressed in the following format: *access_key/YYYYMMDD/region/service/aws4_request*.

For more information, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-Date

The date that is used to create the signature. The format must be ISO 8601 basic format (YYYYMMDD'T'HHMMSS'Z'). For example, the following date time is a valid X-Amz-Date value: 20120325T120000Z.

Condition: X-Amz-Date is optional for all requests; it can be used to override the date used for signing requests. If the Date header is specified in the ISO 8601 basic format, X-Amz-Date is not required. When X-Amz-Date is used, it always overrides the value of the Date header. For more information, see [Elements of an AWS API request signature](#) in the *IAM User Guide*.

Type: string

Required: Conditional

X-Amz-Security-Token

The temporary security token that was obtained through a call to AWS Security Token Service (AWS STS). For a list of services that support temporary security credentials from AWS STS, see [AWS services that work with IAM](#) in the *IAM User Guide*.

Condition: If you're using temporary security credentials from AWS STS, you must include the security token.

Type: string

Required: Conditional

X-Amz-Signature

Specifies the hex-encoded signature that was calculated from the string to sign and the derived signing key.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

X-Amz-SignedHeaders

Specifies all the HTTP headers that were included as part of the canonical request. For more information about specifying signed headers, see [Create a signed AWS API request](#) in the *IAM User Guide*.

Condition: Specify this parameter when you include authentication information in a query string instead of in the HTTP authorization header.

Type: string

Required: Conditional

Common Errors

This section lists the errors common to the API actions of all AWS services. For errors specific to an API action for this service, see the topic for that API action.

AccessDeniedException

You do not have sufficient access to perform this action.

HTTP Status Code: 400

IncompleteSignature

The request signature does not conform to AWS standards.

HTTP Status Code: 400

InternalFailure

The request processing has failed because of an unknown error, exception or failure.

HTTP Status Code: 500

InvalidAction

The action or operation requested is invalid. Verify that the action is typed correctly.

HTTP Status Code: 400

InvalidClientTokenId

The X.509 certificate or AWS access key ID provided does not exist in our records.

HTTP Status Code: 403

NotAuthorized

You do not have permission to perform this action.

HTTP Status Code: 400

OptInRequired

The AWS access key ID needs a subscription for the service.

HTTP Status Code: 403

RequestExpired

The request reached the service more than 15 minutes after the date stamp on the request or more than 15 minutes after the request expiration date (such as for pre-signed URLs), or the date stamp on the request is more than 15 minutes in the future.

HTTP Status Code: 400

ServiceUnavailable

The request has failed due to a temporary failure of the server.

HTTP Status Code: 503

ThrottlingException

The request was denied due to request throttling.

HTTP Status Code: 400

ValidationError

The input fails to satisfy the constraints specified by an AWS service.

HTTP Status Code: 400

Granting IAM permissions for Amazon EC2 Auto Scaling actions

If you receive an `AccessDeniedException` when calling an Amazon EC2 Auto Scaling API action, it means that the AWS Identity and Access Management (IAM) credentials that you are using do not have the required permissions to make that call.

By default, a brand new user in your AWS account has no permissions to do anything. An IAM administrator must create and assign IAM policies that give an IAM identity (such as a user or role) permission to perform Amazon EC2 Auto Scaling API actions. For more information, see [Identity and Access Management for Amazon EC2 Auto Scaling](#) in the *Amazon EC2 Auto Scaling User Guide*.

In general, to perform an Amazon EC2 Auto Scaling action, an IAM identity must have only the matching action included in a policy, but doesn't need to be explicitly granted permission to manage Amazon EC2 instances. However, there are some operations that require multiple actions in a policy. These additional actions are called *dependent actions*. For example, if you call `CreateAutoScalingGroup` to create an Auto Scaling group with a launch template, you must also have the Amazon EC2 API permissions necessary to complete this action. For more information, see [Amazon EC2 Auto Scaling API permissions](#) in the *Amazon EC2 Auto Scaling User Guide*.