

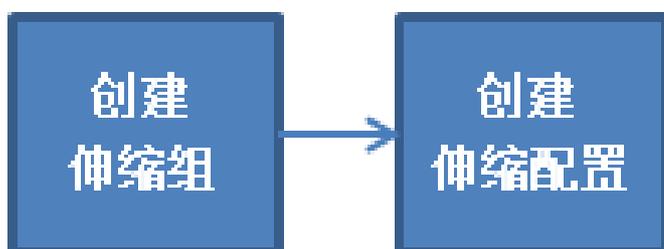
弹性伸缩

快速入门

快速入门

本快速入门介绍如何通过弹性伸缩创建一台 ECS 实例并自动添加到负载均衡，以及自动向 RDS 访问白名单添加该 ECS 实例 IP 的简易伸缩方案。

创建完整的弹性伸缩方案，您需要通过以下 2 步来完成：



登录弹性伸缩控制台，单击 **整体方案创建管理**，通过向导模式创建整体伸缩方案。



创建伸缩组

伸缩组是具有相同应用场景的 ECS 实例的集合。伸缩组定义了组内 ECS 实例数的最大值、最小值及其相关联的负载均衡实例和 RDS 实例等属性。

1. 选择伸缩组所在的 **地域** 和填写 **伸缩组名称**。
2. 将 **伸缩最大实例数** 和 **伸缩最小实例数** 设置为 **1**，则在整体伸缩方案创建完成后，将自动创建 1 台 ECS 实例。

选择 **负载均衡** 实例。指定的负载均衡实例所有配置的监听端口必须开启健康检查。

1. 组基本信息 (必填) 2. 伸缩组配置 (必填) 3. 伸缩自动触发设置 (选填) 4. 预览

1 伸缩组基本信息 伸缩组是具有相同应用场景的ECS实例的集合。伸缩组定义了组内ECS实例数的最大值、最小值及其相关联的SLB实例和RDS实例等属性。

*地域: 北京
更改地域将清空选择的负载均衡、数据库实例和实例配置。

*伸缩组名称: ESS示例
名称为2-40个字符, 以大小写字母, 数字或中文开头, 可包含"-"、"."或"_"

*伸缩最大实例数(台): 2
最大值100

*伸缩最小实例数(台): 2
请填写正整数

*默认冷却时间(秒): 300
最小为0, 必须为整数

释放策略: 执行释放 最早伸缩配置对应的实例 在组中再筛选 最早创建的实例 执行释放

负载均衡: 148a0b6a978-cn-beijing-btc-a01 SLB选定以后将不支持修改

数据库(个): 1 选择数据库 数据库选定以后将不支持修改

取消 下一步

选择 **数据库** 实例, 弹出 **选择RDS数据库** 对话框。

选择RDS数据库

注: 伸缩组最多可选择3个RDS实例。

未选择的RDS实例
请输入实例名称
暂无数据

已选择的RDS实例
rds7fziqumaifq

您共有1个可添加的实例, 已选择1个实例。

确认 取消

选择完成后, 单击 **下一步**。

创建伸缩配置

伸缩配置定义了用于弹性伸缩的 ECS 实例的配置信息。弹性伸缩为伸缩组自动增加 ECS 实例时, 会根据伸缩配置创建 ECS 实例。

1. 填写伸缩配置名称。

2. 选择需要自动化创建的实例的 CPU、内存、带宽。
3. 选择自定义镜像。需要在自定义镜像实现业务逻辑，如自动启动 Web 服务器，自动下载代码和脚本等。
4. 单击 **保存**。
5. 单击 **去预览**。

弹性伸缩服务 ESS 伸缩组管理 自动触发任务管理 整体方案创建管理

伸缩组创建流程

1. 组基本信息 (必填) 2. 伸缩组配置 (必填) 3. 伸缩自动触发设置 (选填) 4. 预览

2 伸缩配置 伸缩配置定义了用于弹性伸缩的ECS实例的配置信息。ESS为伸缩组自动增加ECS实例时，会根据伸缩配置创建ECS实例。

- 伸缩组ECS实例（按量型）配置预设

* 配置名称： 名称为2-40个字符，以大小写字母、数字或中文开头，可包含“-”、“_”或“.”

CPU： 1核 2核 4核

内存： 512MB 1GB 2GB 4GB 为了保证良好的性能体验，512MB内存不提供windows操作系统

公网带宽： 按使用流量 按固定带宽

带宽峰值： 25MB 50MB 100MB 1 Mbps

镜像类型： 公共操作系统 自定义镜像

自定义镜像：

系统盘： 云磁盘 系统盘挂载点：/dev/xvda

* 预消耗费用：配置费用：0.2元/时/台 公网流量费用：0.8元/GB

预览及完成方案创建

预览页面将统一展示前两步指定的伸缩组、伸缩配置信息，以及该方案在创建时预计花费的费用。仅收取 ECS 实例的费用，弹性伸缩免费。

弹性伸缩服务 ESS | 伸缩组管理 | 自动触发任务管理 | 整体方案创建管理

伸缩组创建流程

1. 组基本信息 (必填) | 2. 伸缩组配置 (必填) | 3. 伸缩自动触发设置 (选填) | 4. 预览

4 预览

1. 组基本信息

伸缩组名称	ESS示例	修改
伸缩组最大实例数(台)	1	
伸缩组最小实例数(台)	1	
默认冷却时间(秒)	300	
释放策略	执行释放 [最早伸缩配置对应的实例]，在结果中再筛选 [最早创建的实例] 执行释放	
负载均衡	148a0b6a978-cn-beijing-btc-a01	
数据库	rd67fzqumafq	

2. 伸缩组配置

伸缩组ECS配置			
CPU	1 核	带宽峰值	1 M
内存	512 M	镜像类型	自定义镜像
公网带宽	按使用流量计	操作系统	phowind-CHS-patch

本配置：伸缩组最小实例数为1，添加了0台已有ECS，仍需自动创建1台
预计花费：0.2 元/小时台 × 1 台

取消 上一步 完成

1. 单击 **完成**，将展示整个方案的创建结果。
2. 弹性伸缩会自动创建伸缩组、伸缩配置并启动伸缩组。
3. 弹性伸缩会根据伸缩配置自动创建一台（伸缩最小实例数）ECS 实例，并自动将该实例加入指定的负载均衡和将该实例的 IP 加入指定的 RDS 访问白名单当中。

伸缩组创建流程

✓ 伸缩组创建成功

创建详情：

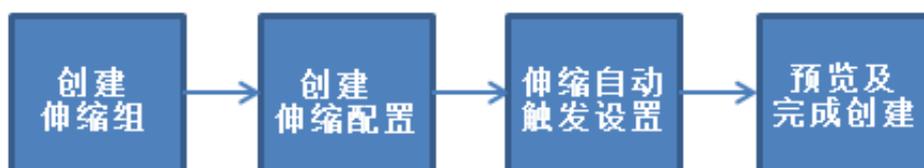
创建伸缩组	成功	创建伸缩配置	成功
启动伸缩组	成功		

您可以：

[查看所有伸缩组列表](#) [查看伸缩组详情](#) [继续创建伸缩组](#) [查看消费记录](#)

关闭

本文介绍如何创建和配置包含定时、动态、自定义、固定模式的整体伸缩方案。创建完整的弹性伸缩方案，您需要通过以下 4 步来完成：



登录弹性伸缩控制台，单击 **整体方案创建管理**，通过向导模式创建整体伸缩方案。

1. 创建伸缩组

填写伸缩组基本信息，选择伸缩组所在的 **地域**、**伸缩组名称**、**伸缩最大实例台数**、**伸缩最小实例台数** 等信息，如果您需要使用负载均衡和 RDS，需要选择 **负载均衡** 实例和选择 **数据库** 实例。

The screenshot shows the '弹性伸缩服务 ESS' console with the '伸缩组创建流程' (Scaling Group Creation Process) wizard. Step 1, '伸缩组基本信息' (Basic Information), is active. The interface includes a progress bar at the top with four steps: 1. 组基本信息 (必填) (Group Basic Information - Required), 2. 伸缩组配置 (必填) (Scaling Group Configuration - Required), 3. 伸缩自动触发设置 (选填) (Scaling Auto-trigger Settings - Optional), and 4. 预览 (Preview). Below the progress bar, a description states: '伸缩组是具有相同应用场景的ECS实例集合。它定义了组内ECS实例数的最大、最小值及其相关联的SLB/RDS实例等属性。' (A scaling group is a collection of ECS instances with the same application scenario. It defines the maximum and minimum number of ECS instances in the group and their associated SLB/RDS instances, etc.).

The '伸缩组基本信息' section contains the following fields:

- 地域:** 杭州 (Region: Hangzhou)
- 伸缩组名称:** ess-test (Scaling Group Name: ess-test)
- 伸缩最大实例数 (台):** 3 (Maximum Number of Instances: 3)
- 伸缩最小实例数 (台):** 1 (Minimum Number of Instances: 1)
- 默认冷却时间 (秒):** 300 (Default Cooling-off Time: 300)
- 移出策略:** 先选择 (最老创建的实例) 在结果中再筛选 (最新创建的实例) 移出 (Eviction Policy: Select oldest instances first, then filter in results, then evict newest instances)
- 负载均衡:** canbal-test (Load Balancing: canbal-test)
- 数据库:** rds123 (Database: rds123)

Buttons for '取消' (Cancel) and '下一步' (Next Step) are located at the bottom right.

2. 创建伸缩配置

伸缩配置定义了用于弹性伸缩的 ECS 实例的配置信息。弹性伸缩为伸缩组自动增加 ECS 实例时，会根据伸缩配置创建 ECS 实例。

The screenshot shows the '弹性伸缩服务 ESS' console with the '伸缩配置' (Scaling Configuration) wizard. Step 2, '伸缩配置' (Scaling Configuration), is active. The interface includes a progress bar at the top with four steps: 1. 组基本信息 (必填) (Group Basic Information - Required), 2. 伸缩组配置 (必填) (Scaling Group Configuration - Required), 3. 伸缩自动触发设置 (选填) (Scaling Auto-trigger Settings - Optional), and 4. 预览 (Preview). Below the progress bar, a description states: '伸缩配置定义了用于弹性伸缩的ECS实例的配置信息。ESS为伸缩组自动增加ECS实例时，会根据伸缩配置创建ECS实例。' (Scaling configuration defines the configuration information for ECS instances used for elastic scaling. ESS will create ECS instances for the scaling group based on the scaling configuration when it automatically adds ECS instances.)

The '伸缩配置' section contains the following fields:

- 配置名称:** ess-test (Configuration Name: ess-test)
- CPU:** 1核, 2核, 4核, 8核, 16核 (CPU: 1 core, 2 cores, 4 cores, 8 cores, 16 cores)
- 内存:** 2GB, 4GB, 8GB, 16GB (Memory: 2GB, 4GB, 8GB, 16GB)
- 公网带宽:** 按使用流量 (按固定带宽) (Public Network Bandwidth: Pay by usage / Pay by fixed bandwidth)
- 带宽峰值:** 25M, 50M, 100M, 1 Mbps起 (Bandwidth Peak: 25M, 50M, 100M, starting from 1 Mbps)
- 镜像类型:** 公共镜像 | 自定义镜像 (Image Type: Public Image | Custom Image)
- 公共镜像:** Ubuntu 14.04 64位 (Public Image: Ubuntu 14.04 64-bit)
- 系统盘:** 普通云盘 (系统盘挂载点: /dev/xvda) (System Disk: Standard Cloud Disk (Mount Point: /dev/xvda))
- 数据盘:** 增加一块 (您还可选配4块, 独立云磁盘支持卸载) (Data Disk: Add one (You can also select up to 4, independent cloud disks support unmounting))

At the bottom, there is a '保存' (Save) button and a note: '* 预消耗流量配置费用: 0.56元/时/台 公网流量费用: 0.8元/GB' (Pre-consumption traffic configuration fee: 0.56 yuan/hour/instance, Public network traffic fee: 0.8 yuan/GB).

Buttons for '取消' (Cancel), '上一步' (Previous Step), '下一步' (Next Step), and '去预览' (Go to Preview) are located at the bottom right.

3. 伸缩自动触发设置

伸缩自动触发有 2 种方式：

- 定时触发自动伸缩任务

- 报警触发自动伸缩任务

定时触发自动伸缩任务

如果未设置重复周期，则按指定的日期和时间执行一次。如果设置了重复周期，则此属性指定的时间点，默认为周期性任务的执行时间点。

报警触发自动伸缩任务

基于云监控性能指标（如 CPU、内存利用率），自动增加或减少伸缩组内的 ECS 实例。在使用报警任务之前，需要在 ECS 的镜像里安装新版本的云监控 Agent。

4. 预览及完成创建

预览页面将统一展示前几步指定的伸缩组、伸缩配置、添加的 ECS 实例、定时任务、报警任务信息，以及该方案在创建时预计花费的费用。仅收取 ECS 实例的费用，弹性伸缩免费。

The screenshot shows the 'Elastic Scaling Service ESS' configuration interface. It is divided into four steps: 1. Basic Information (必填), 2. Scaling Group Configuration (必填), 3. Scaling Auto-trigger Settings (选填), and 4. Complete. The current view is the 'Preview' (预览) stage.

1. 组基本信息 (必填)

伸缩组名称	ess-test
伸缩最大实例数 (台)	3
伸缩最小实例数 (台)	1
默认冷却时间 (秒)	300
移出策略	先淘汰[最早创建的实例]，在结果中再淘汰[最新创建的实例]移出
负载均衡	14fd4b08d39-cn-hangzhou-dg-a01
数据库	rds123

2. 伸缩组配置

伸缩组ECS配置

CPU	2核	带宽峰值	1 M
内存	2 G	镜像类型	公共操作系统
公网带宽	按使用流量	镜像名称	Ubuntu 14.04 64位

3. 伸缩自动触发设置

定时触发自动伸缩任务:

单击 **完成**，将展示整个方案的创建结果。

弹性伸缩会自动创建伸缩组、创建伸缩配置、启动伸缩组、创建报警任务、创建定时任务及创建对应的伸缩规则。

弹性伸缩会根据伸缩最小实例数为 1 台自动创建 1 台 ECS 实例，然后根据报警规则和定时任务规则将增加的实例加入指定的负载均衡，并将该实例的 IP 加入指定的 RDS 访问白名单当中。



API快速入门

本文介绍如何通过 Open API 创建和配置包含定时、动态、自定义、固定模式的整体伸缩方案。

创建完整的弹性伸缩方案，您需要通过以下图中所示步骤来完成（其中前三步为简易的伸缩方案）：



创建伸缩组 (CreateScalingGroup) ，配置伸缩资源的最小值 (MinSize) 、最大值 (MaxSize) 及需要关联的负载均衡实例和 RDS 实例。

创建伸缩配置 (CreateScalingConfiguration) ，指定需要弹性伸缩的 ECS 实例的相关属性，如 ImageID、InstanceType 等。

以第二步创建的伸缩配置启用伸缩组 (EnableScalingGroup) 。

创建伸缩规则 (CreateScalingRule) ，如 **加 N 台 ECS 实例** 的伸缩规则。

创建伸缩任务 (CreateScheduledTask) ，如创建 12 : 00 触发第四步伸缩规则的定时任务。

创建报警任务 (云监控 API PutAlarmRule) ，如创建 CPU 大于等于 80% 则增加一台 ECS 实例的报警任务。

创建伸缩组 (CreateScalingGroup) ，配置伸缩资源的最小值 (MinSize) 、最大值 (MaxSize) 及需要关联的负载均衡实例和 RDS 实例。

请求示例

```
http://ess.aliyuncs.com/?Action=CreateScalingGroup
&RegionId=cn-qingdao
&MaxSize=20
&MinSize=2
&LoadBalancerId=147b46d767c-cn-qingdao-cm5-a01
&DBInstanceId.1=rdszzzyyunybaeu
&DBInstanceId.2=rdsia3u3yia3u3y
&<公共请求参数>
```

返回示例

```
<CreateScalingGroupResponse>
<ScalingGroupId>dP8VqSd9ENXPc0ciVmbcrBT1</ScalingGroupId>
<RequestId>536E9CAD-DB30-4647-AC87-AA5CC38C5382</RequestId>
</CreateScalingGroupResponse>
```

创建伸缩配置 (CreateScalingConfiguration) ，指定需要弹性伸缩的 ECS 实例的相关属性，如 ImageID、InstanceType 等。请求时，需要指定第一步返回的 ScalingGroupId。

请求示例

```
http://ess.aliyuncs.com/?Action=CreateScalingConfiguration
&ScalingGroupId=dP8VqSd9ENXPc0ciVmbcBT1
&SecurityGroupId=sg-280ih3w4b
&ImageId=centos6u5_64_20G_aliaegis_20140703.vhd
&InstanceType=ecs.t1.xsmall
&<公共请求参数>
```

返回示例

```
<CreateScalingConfigurationResponse>
<ScalingConfigurationId>eOs27Kb0oXvQcUYjEGelJqUy</ScalingConfigurationId>
<RequestId>5CC0AD41-08ED-4559-A683-6F56355FE068</RequestId>
</CreateScalingConfigurationResponse>
```

以第二步创建的伸缩配置启用伸缩组（EnableScalingGroup），还可以在此步加入一台既有的 ECS 实例。

请求示例

```
http://ess.aliyuncs.com/?Action=EnableScalingGroup
&ScalingGroupId=dP8VqSd9ENXPc0ciVmbcBT1
&ActiveScalingConfigurationId=eOs27Kb0oXvQcUYjEGelJqUy
&InstanceId.1=i-283vvytn
&<公共请求参数>
```

返回示例

```
< EnableScalingGroupResponse>
<RequestId>6469DCD0-13AC-487E-85A0-CE4922908FDE</RequestId>
</ EnableScalingGroupResponse>
```

创建伸缩规则（CreateScalingRule），如加 1 台 ECS 实例的伸缩规则。在请求时，需要指定第一步返回的 ScalingGroupId。

请求示例

```
http://ess.aliyuncs.com/?Action=CreateScalingRule
```

```
&ScalingGroupId=dP8VqSd9ENXPc0ciVmbrBT1
&AdjustmentType=QuantityChangeInCapacity
&AdjustmentValue=1
&<公共请求参数>
```

返回示例

```
<CreateScalingRuleResponse>
<ScalingRuleAri>
ari:acs:ess:cn-qingdao:1344371:scalingrule/eMKWG8SRNb9dBLAjweNI1Ik
</ScalingRuleAri>
<ScalingRuleId>eMKWG8SRNb9dBLAjweNI1Ik</ScalingRuleId>
<RequestId>570C84F4-A434-488A-AFA1-1E3213682B33</RequestId>
</CreateScalingRuleResponse>
```

创建伸缩任务（CreateScheduledTask），如创建 12:00 触发第四步伸缩规则的定时任务。在请求时，需要指定第四步返回的 ScalingRuleAri。

请求示例

```
http://ess.aliyuncs.com/?Action=CreateScheduledTask
&RegionId=cn-qingdao
&LaunchTime=2014-08-17T12:00Z
&RecurrenceType=Daily
&RecurrenceValue=1
&RecurrenceEndTime=2014-09-17T16:55Z
&ScheduledAction=ari:acs:ess:cn-qingdao:1344371:scalingrule/eMKWG8SRNb9dBLAjweNI1Ik
&<公共请求参数>
```

返回示例

```
<CreateScheduledTaskResponse>
<ScheduledTaskId>edRtShc57WGXdT8TIPbrjsnV</ScheduledTaskId>
<RequestId>0F02D931-2B12-44D7-A0E9-39925C13D15E</RequestId>
</CreateScheduledTaskResponse>
```

弹性伸缩的 ECS 实例中部署的应用需要是无状态、可横向扩展的。

由于弹性伸缩会自动释放 ECS 实例，所以用于弹性伸缩的 ECS 实例不可以保存应用的状态信息（如 session）和相关数据（如数据库、日志等）。如果应用中需要保存状态信息，可以考虑把状态信息保

存到独立的状态服务器、数据库（如 RDS）及集中日志存储（如 SLS）。

弹性伸缩自动扩展出来的实例暂不支持直接自动添加到 OCS 访问白名单中,需要您自行添加。

弹性伸缩目前不支持 **纵向扩展**，即弹性伸缩暂时无法自动升降 ECS 的 CPU、内存和带宽。