

弹性伸缩

用户指南

用户指南

工作原理

使用流程

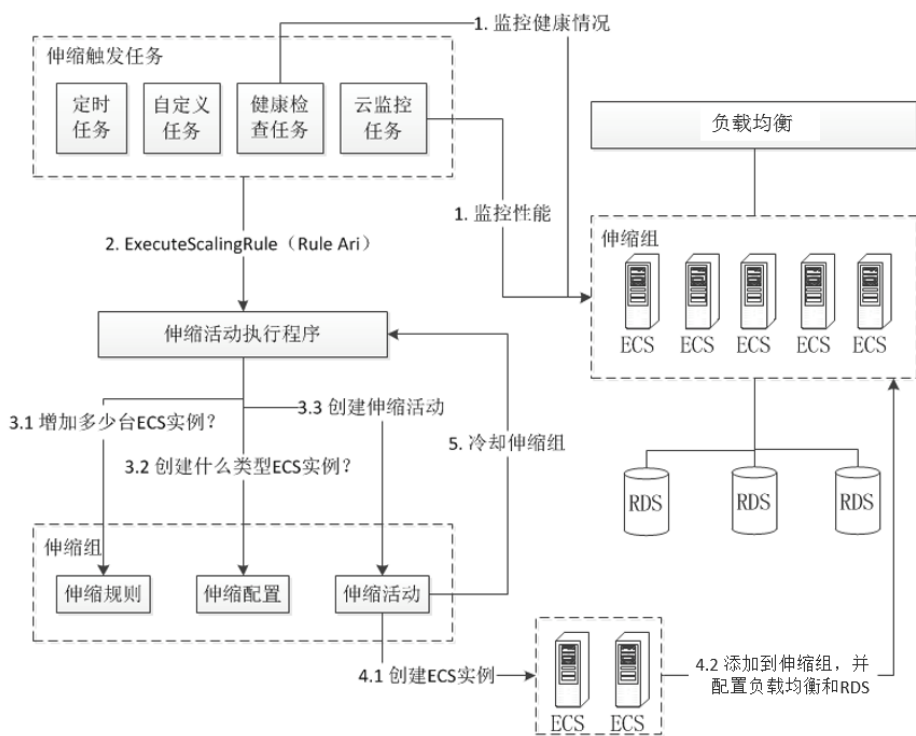
创建完整的弹性伸缩方案，您需要通过以下步骤来完成：



1. 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。
2. 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。
3. 以第二步创建的伸缩配置启用伸缩组（EnableScalingGroup）。
4. 创建伸缩规则（CreateScalingRule），如 **加 N 台 ECS 实例** 的伸缩规则。
5. 创建定时任务（CreateScheduledTask），如创建 12：00 触发第四步伸缩规则的定时任务。
6. 创建报警任务（云监控 API PutAlarmRule），如创建 CPU 平均值（也可以是最大值或最小值）大于等于 80% 则增加一台 ECS 实例的报警任务。

工作流程

下图是弹性伸缩的工作流程。



创建好伸缩组、伸缩配置、伸缩规则、伸缩触发任务后，系统会自动化执行以下流程（以增加 ECS 实例为例）：

伸缩触发任务会按照各自 **触发生效的条件** 来触发伸缩活动。

- 云监控任务会实时监控伸缩组内 ECS 实例的性能，并根据用户配置的报警规则（如伸缩组内所有 ECS 实例的 CPU 平均值大于 60%）触发执行伸缩规则请求。
- 定时任务会根据用户配置的时间来触发执行伸缩规则请求。
- 您可以根据自己的监控系统及相应的报警规则（如在线人数、作业队列）来触发执行伸缩规则请求。
- 健康检查任务会定期检查伸缩组和 ECS 实例的健康情况，如发现有不健康的 ECS 实例（如 ECS 为非running 状态）会触发执行 **移出该 ECS 实例** 的请求。

系统自动通过 ExecuteScalingRule 接口触发伸缩活动，并在该接口中指定需要执行的伸缩规则的阿里云资源唯一标识符（Ari）。

如果是用户自定义的任务，则需要用户在自己的程序中调用 ExecuteScalingRule 接口来实现。

根据步骤 2 传入的伸缩规则 Ari（Rule Ari）获取伸缩规则、伸缩组、伸缩配置的相关信息，并创建伸缩活动。

- i. 通过伸缩规则 Ari 查询伸缩规则以及相应的伸缩组信息，计算出需要增加的 ECS 实例数量，并获得需要配置的负载均衡和 RDS 信息。
- ii. 通过伸缩组查询到相应的伸缩配置信息，即获得了需要创建的ECS实例的配置信息（CPU、内存、带宽等）。
- iii. 根据需要增加的 ECS 实例数量、ECS 实例配置信息、需要配置的负载均衡实例和 RDS 实

例创建伸缩活动。

在伸缩活动中，自动创建 ECS 实例并配置负载均衡和 RDS。

- i. 按照实例配置信息创建指定数量的 ECS 实例。
- ii. 将创建好的 ECS 实例的内网 IP 添加到指定的 RDS 实例的访问白名单当中，将创建好的 ECS 实例添加到指定的负载均衡实例当中。

伸缩活动完成后，启动伸缩组的冷却功能。待冷却时间完成后，该伸缩组才能接收新的执行伸缩规则请求。

注意事项

伸缩规则

在计算和执行过程中，伸缩规则可以根据伸缩组的 MinSize、MaxSize 进行自动调整其需要增加或减少的 ECS 实例数（例：如伸缩规则中指定将伸缩组的 ECS 实例数调整至 50 台，但伸缩组 MaxSize 只有 45 台，则整个伸缩规则会按调整至 45 台来计算和执行。）

伸缩活动

同一伸缩组内、同一时刻只能有一个伸缩活动在执行。

伸缩活动不可以中断。例如，某个创建 20 台 ECS 实例的伸缩活动正在执行中，当创建到第 5 台 ECS 实例时，您无法强行终止该伸缩活动。

伸缩活动有 ECS 实例加入伸缩组失败时，需要保持 ECS 实例级事务的完整性，而非伸缩活动级事务的完整性，即只进行 ECS 实例级回滚，而不是伸缩活动级回滚。例如，当伸缩组创建了 20 台 ECS 实例，但只有 19 台 ECS 实例成功加入负载均衡时，则只对不成功的 1 台 ECS 实例进行自动释放操作。

由于弹性伸缩是借助阿里云的 RAM（Resource Access Management）服务，通过 ECS Open API 代替用户弹性伸缩 ECS 实例资源，所以回滚的 ECS 实例仍然会被扣费。

冷却时间

在冷却时间内，伸缩组只会拒绝云监控报警任务类型的伸缩活动请求，其他类型的触发任务（如用户手工执行伸缩规则、定时任务等）可以绕过冷却时间立即执行伸缩活动。

每个伸缩活动的最后一个 ECS 实例加入或移出伸缩组成功后，整个伸缩组冷却时间才开始计时。

使用须知

移出策略

有两种移出策略：默认策略和自定义策略。

默认移出策略

先按最早伸缩配置(OldestScalingConfiguration)对应的 ECS 实例进行一级筛选，再按最早创建的 ECS 实例(OldestInstance)进行二级筛选。

先在伸缩组里选最早伸缩配置(OldestScalingConfiguration)对应的 ECS 实例，然后在筛选的结果中，选最早创建的 ECS 实例(OldestInstance)。如果还有多台 ECS 实例，则随机选择一台移出伸缩组。

由于用户手工加入的既有 ECS 实例不与伸缩配置绑定，所以不会先选出该类型的 ECS 实例。

如果伸缩配置关联的 ECS 实例已经全部移出，但还需移出 ECS 实例时，则选最早用户手工移入的 ECS 实例。

自定义设置释放策略

根据用户设置的多条策略依次筛选 ECS 实例并移出伸缩组。

释放策略类型

OldestInstance：取最早创建的 ECS 实例，如果作为一级筛选条件，不区分是用户手工加入的 ECS 实例还是自动创建的 ECS 实例，直接按最早创建的实例进行移出。

NewestInstance：取最新创建的实例，如果作为一级筛选条件，不区分是用户手工加入的 ECS 实例还是自动创建的 ECS 实例，直接按最新创建的 ECS 实例进行移出。

OldestScalingConfiguration：取最早伸缩配置创建的 ECS 实例；对于用户手工移入的 ECS 实例，自动跳过此条件；如果与伸缩配置相关联的 ECS 实例已经全部移出，还要继续移出 ECS 实例时，则在用户手工加入的 ECS 实例中（没有与伸缩配置关联的 ECS 实例），随机选一台。

ECS实例生命周期的管理

加入伸缩组的 ECS 实例有两种类型：自动创建的 ECS 实例、手工添加的 ECS 实例。

自动创建的 ECS 实例

“自动创建的 ECS 实例”是指根据用户的伸缩配置和伸缩规则，由弹性伸缩服务自动创建的 ECS 实例。

弹性伸缩服务负责该类型 ECS 实例的全生命周期管理，即在伸缩组扩展活动中，负责该 ECS 实例的创建，在伸缩组的收缩活动中，负责停止和释放该 ECS 实例。

手工添加的 ECS 实例

手工添加的 ECS 实例 是指不是由弹性伸缩服务创建，但由用户手工添加到伸缩组中的 ECS 实例。

弹性伸缩服务不负责该类型 ECS 实例的全生命周期管理，即该类型的 ECS 实例是在弹性伸缩服务之外创建，只能手工加入伸缩组，在伸缩组的收缩活动中或用户手工移出该类型 ECS 实例时，弹性伸缩服务只负责将该 ECS 实例移出伸缩组，不会停止和释放该 ECS 实例。

实例状态

ECS 实例在伸缩组中的生命周期，通过以下几种状态描述：

- Pending – 表示 ECS 实例正在加入伸缩组，包括创建实例、加入负载均衡、添加 RDS 访问名单等过程。
- InService – 表示 ECS 实例已成功加入伸缩组，并正常提供服务。
- Removing – 表示 ECS 实例正在移出伸缩组。

实例的健康状态

ECS 实例在伸缩组中的健康状态为：

- Healthy(健康)
- Unhealthy(不健康)

本期仅通过 ECS 实例为非 **运行中 (Running)** 状态来判断该实例不健康。弹性伸缩会自动移出伸缩组中不健康的 ECS 实例。

- 对于 **自动创建的 ECS 实例**，弹性伸缩会停止和释放该 ECS 实例。
- 对于 **手工添加的 ECS 实例**，弹性伸缩不会停止和释放该 ECS 实例。

冷却时间

冷却时间指成功执行伸缩活动后的一段锁定时间。您可以通过以下两种方式配置冷却时间：

在伸缩组中配置默认冷却时间，默认冷却时间不能为空，请参照 [创建伸缩组](#)。

在伸缩规则中配置冷却时间，如果不配置，将使用默认冷却时间，请参照 [创建伸缩规则](#)。

说明：如果同时在伸缩组和伸缩规则中配置了冷却时间，弹性伸缩服务将优先使用伸缩规则中的冷却时间。

冷却时间规则

伸缩组成功执行伸缩活动后，弹性伸缩服务开始计算冷却时间。如果在伸缩活动中，多台 ECS 实例加入或者移出伸缩组，则从最后一台 ECS 实例加入或者移出伸缩组后开始计算冷却时间，请参照 [示例 1](#)。如果在伸缩活动中，没有 ECS 实例成功加入或者移出伸缩组，则不会开始计算冷却时间。

在冷却时间内，伸缩组会拒绝由云监控报警任务触发的伸缩活动请求。但其他类型任务（手动执行任务、定时任务）触发的伸缩活动可以立即执行，绕过冷却时间。

如果您停用伸缩组，然后再次启用伸缩组，冷却时间将会失效，请参照 [示例 2](#)。

说明：冷却时间仅锁定同一个伸缩组的伸缩活动，不会对其他伸缩组的伸缩活动产生影响。

示例场景

示例 1

一个伸缩组 asg-uf6f3xewn3dvz4bsy7r1，默认冷却时间是 10 分钟，伸缩组内存在一条伸缩规则 add3，冷

却时间为 15 分钟。

根据 add3 成功执行一次伸缩活动后，扩张了 3 台 ECS 实例，从第 3 台 ECS 实例加入伸缩组开始计算冷却时间，15 分钟内不会接受由云监控报警任务触发的伸缩活动请求。

示例 2

一个伸缩组 asg-m5efkz67re9x7a571bjh，默认冷却时间是 10 分钟，伸缩组内存在一条伸缩规则 remove1，没有设置冷却时间。

在 18:00 根据 remove1 成功执行一次伸缩活动后，收缩了 1 台 ECS 实例。正常情况下，在 18:10 前伸缩组不会接受由云监控报警任务触发的伸缩活动请求。此时停用伸缩组，然后在 18:05 再次启用伸缩组，冷却时间将失效。如果在 18:05 至 18:10 间云监控报警任务触发了伸缩活动请求，伸缩组仍会接受并执行请求。

伸缩组状态

伸缩组状态

伸缩组具有以下三种状态：Active、Inacitve 和 Deleting。

状 态	Open API 对应的状态
创建中	Inactive
已创建	Inactive
启用中	Inactive
运行中	Active
停用中	Inactive
已停止	Inactive
删除中	Deleting

伸缩活动的流程

伸缩活动的生命周期为：**判断伸缩组的健康状态和边界条件** 步骤与 **启动 cooldown** 步骤之间的所有活动。

伸缩组自动伸缩

自动扩展

判断伸缩组的健康状态和边界条件。

分配 ActivityId 和执行伸缩活动。

创建 ECS 实例。

修改 Total Capacity。

分配 IP。

添加 RDS 白名单。

启动 ECS 实例，等待启动完成。

挂载负载均衡，将权重设为创建伸缩配置时指定的 **负载均衡权重**。

伸缩活动完成，启动 cooldown。

自动收缩

判断伸缩组的健康状态和边界条件。

分配 ActivityId 和执行伸缩活动。

从负载均衡移除 ECS 实例。

停止 ECS 实例，等待停止完成。

从 RDS 白名单移除。

释放 ECS 实例。

修改 Total Capacity。

伸缩活动完成，启动 cooldown。

用户手工加入和移出既有的 ECS 实例

手工加入

判断伸缩组的健康状态、边界条件和 ECS 实例的状态、类型。

分配 ActivityId 和执行伸缩活动。

加入 ECS 实例。

修改 Total Capacity。

添加 RDS 白名单。

挂载负载均衡，将权重设为当前伸缩组中已激活的伸缩配置上指定的“负载均衡权重”。

注意：目前弹性伸缩要求手工加入实例时，实例的规格必须与伸缩组中已激活的伸缩配置上的规格相同，所以此处使用了伸缩配置上指定的“负载均衡权重”。

伸缩活动完成，启动 cooldown。

手工移出

判断伸缩组的健康状态和边界条件。

分配 ActivityId 和执行伸缩活动。

负载均衡停止向该 ECS 实例转发流量。

等待 60 秒后，从负载均衡移除 ECS 实例。

从 RDS 白名单移除。

修改 Total Capacity。

从伸缩组移出。

伸缩活动完成，启动 cooldown。

伸缩活动的状态

伸缩活动在请求阶段被拒绝，为 **Rejected** 状态。

伸缩活动执行过程中，为 **InProgress** 状态。

伸缩活动执行完成后，有 3 种状态：

成功 (Successful)：根据 MinSize、MaxSize 调整后的伸缩规则，伸缩活动成功将所有 ECS 实例加入或移出伸缩组。

说明：加入伸缩组成功表明创建 ECS 实例、加入负载均衡实例、配置 RDS 访问白名单每一步骤都成功。以上任何步骤失败都认为是该 ECS 实例加入伸缩组失败。

部分成功 (Warning)：根据 MinSize、MaxSize 调整后的伸缩规则，伸缩活动至少有一台 ECS 实例加入或移出伸缩组不成功。

全部失败 (Failed)：根据 MinSize、MaxSize 调整后的伸缩规则，伸缩活动没有一台 ECS 实例成功加入或移出伸缩组。

举例

伸缩规则：伸缩规则定义增加 5 台 ECS 实例，当前伸缩组的 Total Capacity = 3，MaxSize = 5，则执行该伸缩规则时会调整成增加 2 台 ECS 实例。伸缩活动执行完成后，3 种状态对应的情况如下：

- 成功：2 台 ECS 实例全部创建成功，并正确加入负载均衡实例和配置 RDS 访问白名单。
- 部分成功：2 台 ECS 实例全部创建成功，但只有 1 台正确加入负载均衡实例和配置 RDS 访问白名单，另外一台因无法加入负载均衡实例和配置 RDS 访问白名单而被回滚操作释放。
- 全部失败：2 台 ECS 实例全部创建失败，或者 2 台 ECS 实例全部创建成功，但这 2 台 ECS 实例都无法加入负载均衡实例和配置 RDS 访问白名单而被回滚操作释放。

伸缩活动中失败实例的回滚

伸缩活动有 ECS 实例加入伸缩组失败时，需要保持 ECS 实例级事务的完整性，而非伸缩活动级事务的完整性，即只进行 ECS 实例级回滚，而不是伸缩活动级回滚。

例如，当伸缩组创建了 20 台 ECS 实例，但只有 19 台 ECS 实例成功加入负载均衡时，则只对不成功的 1 台 ECS 实例进行自动释放操作。

由于弹性伸缩是借助阿里云的 RAM (Resource Access Management) 服务，通过 ECS Open API 代替用户弹性伸缩 ECS 实例资源，所以回滚的 ECS 实例仍然会被扣费。

移出不健康ECS实例

当 ECS 实例成功加入伸缩组后，弹性伸缩服务会定期扫描该 ECS 实例的运行状态，如果发现该 ECS 实例为非运行中 (Running) 状态，则会将该 ECS 实例移出伸缩组。

- 对于弹性服务自动创建的 ECS 实例：将立即启动移出和释放该 ECS 实例操作。
- 对于用户手工添加的 ECS 实例：将立即启动移出该 ECS 实例操作，但不会停止和释放该 ECS 实例。

移出不健康 ECS 实例，不受 MinSize 的限制，即移出不健康 ECS 实例后，有可能导致 Total Capacity 低于 MinSize，此时弹性伸缩服务会自动创建差额的 ECS 实例使得 Total Capacity 等于 MinSize。

通知

需要短信息和邮件通知的伸缩活动条件如下：

- 只有定时任务、云监控报警任务、健康检查触发的伸缩活动才通知。
- 只有创建了 ECS 实例或释放了 ECS 实例才通知。

如满足以上条件，则一个伸缩活动对应一条通知。

强制干预

弹性伸缩服务不会阻止用户从 ECS 控制台删除自动创建的 ECS 实例等强制干预行为。对于用户的强制干预

，弹性伸缩服务的处理方法如下：

资源	强制干预类型	弹性伸缩服务的处理方法
ECS	通过 ECS 控制台或 OpenAPI 强制删除伸缩组里的 ECS 实例	会按照健康检查来判断该 ECS 实例为不健康状态，并将该 ECS 实例移出伸缩组。（该 ECS 实例在 RDS 实例访问白名单中的内网 IP 无法自动删除。）如移出该 ECS 实例后，使得 Total Capacity 低于 MinSize，弹性伸缩会自动创建差额的 ECS 实例使得 Total Capacity 等于 MinSize。
ECS	客户取消将 ECS Open API 权限授予弹性伸缩	弹性伸缩将拒绝所有伸缩活动请求。
负载均衡	通过负载均衡控制台或 OpenAPI 将 ECS 实例强制移出负载均衡实例	弹性伸缩不会自动感知用户的移出行为也不会处理该类异常状态，该 ECS 实例仍存在伸缩组中。当发生收缩活动时，如果按照移出策略挑选到该 ECS 实例，仍然会把该 ECS 实例释放掉。
负载均衡	通过负载均衡控制台或 OpenAPI 强制删除负载均衡实例或关闭负载均衡实例的健康检查。	对于配置该负载均衡实例的伸缩组，不会再往伸缩组加入 ECS 实例。伸缩任务可以触发伸缩规则从伸缩组移出 ECS 实例，健康检查判断某个 ECS 实例不健康，也可将该 ECS 实例移出伸缩组。
负载均衡	系统导致负载均衡实例不可用（如欠费、故障）	除了用户手工触发的移出 ECS 实例的伸缩活动，其它所有伸缩活动执行失败。
负载均衡	客户取消将负载均衡 Open API 权限授予弹性伸缩。	对于配置了负载均衡实例的伸缩组，弹性伸缩将拒绝所有伸缩活动请求。
RDS	通过 RDS 控制台或 OpenAPI 将 ECS 实例的 ip 强制移出 RDS 访问白名单	弹性伸缩不会自动感知用户的移出行为也不会处理该类异常状态，该 ECS 实例仍存在伸缩组中。当发生收缩活动时，如果按照移出策略挑选到该 ECS 实例，仍然会把该 ECS 实例释放掉。
RDS	通过 RDS 控制台或 OpenAPI 强制删除 RDS 实例	对于配置该 RDS 实例的伸缩组，不会再往伸缩组加入 ECS 实例。伸缩任务可以触发伸缩规则从伸缩组移出 ECS 实例，健康检查判断某个 ECS 实例不健康，也可将该 ECS 实例移出伸缩组。
RDS	系统导致 RDS 实例不可用（如欠费、故障）	除了用户手工触发的移出 ECS 实例的伸缩活动，其它所有伸缩

		活动执行失败。
RDS	客户取消将 RDS Open API 权限授予弹性伸缩。	对于配置了 RDS 实例的伸缩组，弹性伸缩将拒绝所有伸缩活动请求。

数量限制

目前，弹性伸缩服务的功能项存在以下数量限制：

一个账号下最多创建 20 个伸缩组。

- 一个伸缩组内最多创建 10 个伸缩配置。
- 一个伸缩组内最多创建 50 个伸缩规则。
- 一个伸缩组内最多创建 6 个事件通知。
- 一个伸缩组内最多创建 6 个生命周期挂钩。

对于所有地域和所有伸缩组，一个账号下最多弹性伸缩 1000 台 ECS 实例。此数量只包含自动创建的 ECS 实例，不包含手动添加的 ECS 实例。

一个账号下最多创建 20 个定时任务。

伸缩配置

ECS 实例模板概述

在您的业务需求增长时，弹性伸缩可以根据伸缩规则进行弹性扩张，自动创建 ECS 实例并添加到伸缩组中。

伸缩配置是弹性伸缩用于创建 ECS 实例的模板。在创建伸缩配置时，您可以指定 ECS 实例的信息，例如实例规格、镜像类型、存储大小、登录实例用的密钥对等。您也可以修改现有伸缩配置，快速满足新的业务需求。

说明：伸缩配置依托伸缩组，因此要确保已经存在一个伸缩组，关于如何创建伸缩组，请参阅 [创建伸缩组](#)。一个伸缩组内可以创建的伸缩配置数量有限，请参阅 [使用限制](#)。

目前，您可以进行以下伸缩配置相关的操作：

- 创建伸缩配置
- 选用伸缩配置
- 修改伸缩配置
- 删除伸缩配置

创建伸缩配置

创建伸缩配置的流程和创建 ECS 实例类似，但是由于应用场景为弹性伸缩，因此不支持配置某些项目，例如地域，实际操作时请以控制台页面为准。页面上的各配置项均配有简洁说明，如果您想深入了解各配置项的含义和用法，请参阅 使用向导创建实例。

操作步骤

请参照以下步骤创建一个伸缩配置：

登录弹性伸缩控制台，在伸缩组的 **操作** 列下单击 **管理**。

伸缩组名称/ID	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
sg-yk20180724 asg-bp19mx1q86esyh8yscc9	启用	sc-yk20180814	3	0	5	300	管理 修改 停用 删除

前往 **伸缩配置** 页面，单击 **创建伸缩配置** 按钮。

基本信息 ECS实例列表	伸缩配置 创建伸缩配置									
伸缩活动	每个伸缩组中，您最多可以有10个伸缩配置。									
伸缩配置	伸缩配置	标签	实例规格组	状态	镜像	按量计费	系统盘类型	数据盘	登录凭证	操作
伸缩规则	sc-yk20180814		ecs.an1.medium (2vcpu 4GB)	生效	Windows Server 2012 R2 数据中心版 64位英文版	按量计费	高效云盘	-	未设置	查看详情 修改 删除
事件通知	sc-yk20180724		ecs.xn4.small (1vcpu 1GB)	失效	CentOS 7.4 64位	按量计费	高效云盘	-	未设置	查看详情 修改 停用 删除
生命周期结构	共有2条, 每页显示: 10 条 1									

在 **基础配置** 页面中，依次配置计费方式、实例、镜像、存储、公网带宽和安全组，然后单击 **下一步：系统配置** 按钮。



说明：弹性伸缩仅支持创建按量付费实例和抢占式实例，仅支持选择系统镜像、自定义镜像和共享镜像。

在 **系统配置** 页面中，依次配置标签（可选）、登录凭证、实例名称（可选）和高级选项（可选），然后单击 **下一步：确认配置** 按钮。



说明：高级选项包括实例 RAM 角色和实例自定义数据，选择专有网络的伸缩组方可配置高级选项。

在 **确认配置** 页面中，检查所选配置，填写伸缩配置名称，然后单击 **确认创建** 按钮。



在 **创建成功** 对话框中，单击 **启用配置** 按钮启用该伸缩配置。如果您暂时不需要使用该伸缩配置，可以直接关闭对话框。



创建成功后，伸缩配置会出现在伸缩配置列表中，您可以继续进行查看伸缩配置详情、选用伸缩配置等操作。

选用伸缩配置

在伸缩组中，伸缩配置有两种可能状态：生效 和 失效。您可以根据业务需要选用需要生效的伸缩配置，伸缩组只会根据生效的伸缩配置自动创建 ECS 实例。

伸缩配置内容较多，您可能会忘记具体的配置项。因此弹性伸缩还提供了查看伸缩配置详情的功能，供您随时了解各个伸缩配置，灵活选用 ECS 实例的模板。

操作步骤

请参照以下步骤查看一个伸缩配置的详情并选用一个伸缩配置：

登录弹性伸缩控制台，在伸缩组的 **操作** 列下单击 **管理**。

伸缩组名称/ID	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
sg-yk20180724 asg-bp19mx1q86esyh8yscc9	应用	sc-yk20180814	3	0	5	300	管理 修改 停用 删除

前往 **伸缩配置** 页面，在指定伸缩配置的 **操作** 列下单击 **查看详情**。

确定伸缩配置详情无误后，单击 **选用** 启用该伸缩配置。

基本信息

ECS实例列表

伸缩活动

伸缩配置

伸缩规则

事件通知

生命周期结构

伸缩配置

每个伸缩组中，您最多可以拥有10个伸缩配置。

伸缩配置	标签	实例规格组	状态	镜像	定价计费	系统盘类型	数据盘	登录凭证	操作
sc-yk201808141218		ecs.x7t.small (1vcpu 1GB)	失效	CentOS 7.4 64位	流量计费	高效云盘	-	密钥对 kp-yk20180814	查看详情 修改 选用 删除
sc-yk20180814		ecs.sn1.medium (2vcpu 4GB)	生效	Windows Server 2012 R2 数据中心版 64位英文版	流量计费	高效云盘	-	未设置	查看详情 修改 删除
sc-yk20180724		ecs.x7t.small (1vcpu 1GB)	失效	CentOS 7.4 64位	流量计费	高效云盘	-	未设置	查看详情 修改 选用 删除
共有3条, 每页显示: 10 条									

伸缩配置ID: asc-bp1lcv46gucka38ydc

实例规格组: ecs.x7t.small (1vcpu 1GB)

镜像ID: centos_7_04_04_20G_alibase_201701015.vhd

公网带宽: 按使用流量 5 Mbps (峰值)

登录凭证: 密钥对 kp-yk20180814

伸缩配置名称: sc-yk201808141218

镜像名称: CentOS 7.4 64位

系统盘: 高效云盘40G

负载均衡权重: 50

说明：选用一个伸缩配置后，其他伸缩配置将被置于 **失效** 状态。

选用一个伸缩配置后，在满足弹性扩张条件时，对应的伸缩组会根据该伸缩配置自动创建 ECS 实例。

修改伸缩配置

如果现有伸缩配置都无法完全满足您的业务需求，但是存在一个和期望伸缩配置大体相同的伸缩配置，您可以修改该伸缩配置，调整特定配置项，无须重新创建一个伸缩配置，提高运维效率。

操作步骤

请参照以下步骤修改一个伸缩配置：

登录弹性伸缩控制台，在伸缩组的 **操作** 列下单击 **管理**。

伸缩组名称/ID	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
sg-yk20180724 asg-bp19mx1q86esyh8yscc9	应用	sc-yk20180814	3	0	5	300	管理 修改 停用 删除

前往 **伸缩配置** 页面，在指定伸缩配置的操作列下单击 **修改**。

基本信息

ECS实例列表

伸缩活动

伸缩配置

伸缩规则

事件通知

生命周期结构

伸缩配置

每个伸缩组中，您最多可以拥有10个伸缩配置。

伸缩配置	标签	实例规格组	状态	镜像	定价计费	系统盘类型	数据盘	登录凭证	操作
sc-yk201808141405		ecs.n2.medium (2vcpu 8GB)	生效	Windows Server 2012 R2 数据中心版 64位中文版	流量计费	高效云盘	-	未设置	查看详情 修改 删除
sc-yk201808141402		ecs.xn4.small (1vcpu 1GB)	失效	CentOS 7.4 64位	流量计费	高效云盘	-	未设置	查看详情 修改 停用 删除
sc-yk201808141218		ecs.xn4.small (1vcpu 1GB)	失效	CentOS 7.4 64位	流量计费	高效云盘	-	密钥对 kp-yk20180814	查看详情 修改 停用 删除
删除									

共有3条, 每页显示: 10 条

创建伸缩配置

在 **基础配置** 页面中，修改配置项，然后单击 **下一步：系统配置** 按钮。

弹性伸缩 ESS

伸缩组名称: sg-yk20180724

返回伸缩组列表

返回伸缩配置列表

伸缩配置

系统盘

高效云盘 40 GiB

设置系统盘 / 数据盘前请确保至少已选择一个实例规格

云盘参数和性能

相同选择 SSD 云盘 / 高效云盘 / 普通云盘，请参 详细规格

数据盘

0/16

公网带宽

分配公网IP地址

系统会自动分配IP (不强制绑定)，若不需要分配公网IP，请勾选

按使用流量计费

后付费模式，按使用流量 (单位为GiB) 计费，每小时扣费，请保证余额充足

1M 25M 50M 75M 100M 5 Mbps

阿里云免费提供最高5Gbps的突发流量攻击防护，了解更多 | 提升防护能力

安全组

重新选择安全组

安全组是防火墙功能，用于设置网络访问控制，您也可以对管理控制台 新建安全组、 修改选择。

所选安全组: ecssecuritygrouptest / sg-bp17qgpothd2ic814 (已有3个实例，还可以加入397个实例)

请确保该安全组开放包含22 (Linux) 或者3389 (Windows) 端口，否则无法远程登录ECS，您可以进入ECS控制台设置。前往设置。

公网带宽: 5Mbps 按使用流量计费

配置项: 系统盘 公网带宽

下一步: 系统配置

确认配置

在 **系统配置** 页面中，修改配置项，然后单击 **下一步：确认配置** 按钮。

弹性伸缩 ESS

伸缩组名称: sg-yk20180724

返回伸缩组列表

返回伸缩配置列表

基础配置 (已填)

系统配置

确认配置 (未填)

标签

标签由区分大小写的键值对组成。键值，也可以添加一个键为“Group”且值为“Web”的标签。

标签键不可以重复，最长为64位；标签值可以为空，最长为128位。标签键和标签值都不能以“xiyun”、“acs”、“https”或“http”开头。

您最多可以设置5个标签，设置的标签将应用在本伸缩组的所有实例的云盘。

添加标签

登录凭证

密钥对

密钥对: kp-yk20180814

详细参考 | 新建密钥对

实例名称

autocscaling-yk201808141213

2-128个字符，以大小写字母或中文开头，可包含数字、“-”、“_”、“.”

高级选项 (实例 RAM 角色 & 实例自定义数据 cloud-init)

公网带宽: 5Mbps 按使用流量计费

配置项: 系统盘 公网带宽

上一步: 基础配置

下一步: 确认配置

确认配置

在 **修改成功** 对话框中，单击 **启用配置** 按钮启用该伸缩配置。如果您暂时不需要使用该伸缩配置，可以直接关闭对话框。



修改一个伸缩配置后，之前根据该伸缩配置创建的 ECS 实例仍然可以正常运行。

删除伸缩配置

如果您不再需要一个伸缩配置，可以将其删除，空出伸缩配置的配额，用于创建其他伸缩配置。

前提条件

在删除一个伸缩配置前，请确保满足以下条件，否则删除操作会失败：

- 该伸缩配置必须处于 失效 状态。
- 对应的伸缩组中不存在根据该伸缩配置自动创建出的 ECS 实例。

操作步骤

请参照以下步骤删除一个或多个伸缩配置：

登录弹性伸缩控制台，在伸缩组的 **操作** 列下单击 **管理**。

伸缩组名称/ID	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
sg-yk20180724 asg-bp19mx1q86esyh8yscq9	启用	sc-yk20180814	3	0	5	300	管理 修改 停用 删除

前往 **伸缩配置** 页面，在指定伸缩配置的操作列下单击 **删除**。



您也可以勾选伸缩配置左侧的复选框，然后单击下方的 **删除** 按钮进行批量删除。



在 **删除伸缩配置** 对话框中，单击 **确认** 按钮。

伸缩组

创建伸缩组

伸缩组（Scaling Group）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。

ECS 实例移出伸缩组的策略参考 [移出策略](#)。

限制

- 伸缩组、负载均衡实例和 RDS 实例必须在同一个地域。
- 一个用户最多只能创建 20 个伸缩组。

网络类型

伸缩组的网络类型分为经典网络和专有网络 VPC。

对于 VPC 类型伸缩组，您可以创建多可用区专有网络。即您可以为 VPC 类型伸缩组配置同一 VPC 下的多台虚拟交换机，当一个虚拟交换机所在可用区无法创建 ECS 实例时，弹性伸缩自动切换其它可用区创建 ECS 实例。

。

参数描述

根据传入参数创建伸缩组。

伸缩最大实例数（MaxSize）、伸缩最小实例数（MinSize）定义了伸缩组内 ECS 实例个数的最大值和最小值。

- 当伸缩组的当前 ECS 实例数（Total Capacity）小于 MinSize 时，弹性伸缩会自动添加 ECS 实例，使得伸缩组的当前 ECS 实例数等于 MinSize。
- 当伸缩组的当前 ECS 实例数（Total Capacity）大于 MaxSize 时，弹性伸缩会自动移出 ECS 实例，使得伸缩组的当前 ECS 实例数等于 MaxSize。

默认冷却时间（DefaultCooldown）是伸缩组的默认冷却时间。

- 一个伸缩活动（添加或移出 ECS 实例）执行完成后的一段冷却时间内，该伸缩组不执行其它的伸缩活动。
- 目前仅针对云监控报警任务触发的伸缩活动有效。

移出策略（RemovalPolicy）是指当需要从伸缩组移出 ECS 实例并且有多种选择时，则按该策略选择需要移出的 ECS 实例。

负载均衡注意事项

如果在伸缩组中指定了负载均衡实例，伸缩组会自动将加入伸缩组的 ECS 实例添加到指定的负载均衡实例当中。

- 指定的负载均衡实例必须是已启用状态。
- 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，否则创建失败。
- 加入负载均衡的 ECS 实例的权重默认为 50。

一台 SLB 实例可以挂载不同网络类型 ECS 实例：

- 若您选择的 SLB 实例是经典网络类型，该 SLB 实例可以挂载同一 VPC 下的 ECS 实例，但不限制您对经典网络 ECS 实例的挂载，只需确保同一 Region 即可。
- 若您选择的 SLB 实例是 VPC 类型，该 SLB 实例可以挂载与 SLB 相同的 VPC 下的 VPC 类型 ECS 实例，也可以挂载同一地域下的经典网络类型实例。
- 如果您创建的是 VPC 类型伸缩组，您可以指定与伸缩组相同的 VPC 下的专有网络类型 SLB 实例，也可以指定经典网络类型 SLB 实例，但是经典网络 SLB 实例不能挂载其它 VPC 下的 ECS 实例。
- 如果您创建的是经典网络类型伸缩组，您可以指定当前地域下的任意 SLB 实例。

RDS 注意事项

如果在伸缩组中指定了 RDS 实例，伸缩组会自动将加入伸缩组的 ECS 实例的内网 IP 添加到指定的 RDS 实例的访问白名单当中。

- 指定的 RDS 实例必须是运行中状态。
- 指定的 RDS 实例访问白名单的 IP 个数不能达到上限。
- 伸缩组创建成功后，伸缩组不会立即生效，只有启用伸缩组，才能接受伸缩规则的触发和执行相关的伸缩活动。

操作步骤

更多详情，请参阅 [快速入门 步骤2. 创建伸缩组](#)。

下一步

[启用伸缩组](#)

修改伸缩组

修改伸缩组的属性。

在创建伸缩组的属性中，有以下属性不可以修改：

- 地域
- 负载均衡
- RDS 数据库实例

当伸缩组为生效（Active）和失效（Inactive）状态，才可以执行此功能。

当伸缩组的 ECS 实例数（Total Capacity）不满足修改后的 MaxSize 或 MinSize，弹性伸缩服务会自动加入或移出 ECS 实例，使得伸缩组的 ECS 实例数等于 MaxSize 或 MinSize。

示例

修改伸缩组

*伸缩组名称：

classic

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）：

1

最小为0，最大为1000

*伸缩最小实例数（台）：

1

最小为0，最大为1000

*默认冷却时间（秒）：

300

最小为0，必须为整数

移出策略：

先筛选

最早伸缩配置对应的

在结果中再筛选

最早创建的实例

移出

如何保证手工添加的ECS实例不被移出伸缩组

专有网络：

vsw-uf6nx9hd8zsnp33irkwy7

多可用区扩容策略：

优先级策略

负载均衡：

-

数据库：

提交

取消

查询伸缩组

伸缩组有三种状态：

- 启用（ Enable ）：该状态下伸缩组接收伸缩规则请求并触发伸缩活动。
- 停用（ Disable ）：该状态下伸缩组不接收任何伸缩规则请求。
- 删除中（ Deleting ）：该状态下伸缩组不接收任何伸缩规则请求。

操作步骤

- 登录 弹性伸缩管理控制台。
- 在左侧导航栏中，单击 伸缩组管理。
- 选择地域。
- 查看伸缩组状态。



启用伸缩组

启用一个指定的伸缩组。

启用伸缩组成功以后（ active 状态 ），如果当前 ECS 实例数量小于 MinSize ，则弹性伸缩服务会自动创建差额的按量付费的 ECS 实例。

例如，创建伸缩组时，指定 MinSize = 5 ，启用前伸缩组已有 2 台 ECS 实例，则启动后会自动创建 3 台 ECS 实例。

当伸缩组为失效（ Inactive ）状态，才可以执行此功能。

当伸缩组有生效的伸缩配置时，才可以执行此功能。

一个伸缩组在同一时刻只能有一个 生效 的伸缩配置。

示例

伸缩组管理列表							刷新	创建伸缩组
深圳北京杭州青岛								
伸缩组	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作	
33	启用	33	0	0	0	300	管理	修改 停用 删除
22	启用	22	0	0	0	300	管理	修改 停用 删除
11	启用	11	0	0	0	300	管理	修改 停用 删除
多组	停用	111	63	0	0	300	管理	修改 启用 删除
单组3	停用	111	40	0	0	300	管理	修改 启用 删除
单组2	停用	111	40	0	0	300	管理	修改 启用 删除
单组1	停用	111	4	0	0	300	管理	修改 启用 删除
暴力	停用	伸缩配置	0	0	100	300	管理	修改 启用 删除
共有8条， 每页显示：20条							<< < 1 > >>	

停用伸缩组

停用一个指定的伸缩组。

- 只有在当前伸缩组没有任何伸缩活动进行时，才能停用伸缩组。
- 当伸缩组为生效（Active）状态，才可以执行此功能。

示例

伸缩组管理列表							刷新	创建伸缩组
深圳 北京 杭州 青岛								
伸缩组	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作	
33	启用	33	0	0	0	300	管理 修改	停用 删除
22	启用	22	0	0	0	300	管理 修改	停用 删除
11	启用	11	0	0	0	300	管理 修改	停用 删除

删除伸缩组

删除一个指定的伸缩组。

强制删除（ForceDelete）属性表示如伸缩组存在 ECS 实例或正在进行伸缩活动，是否强制删除伸缩组并移出和释放 ECS 实例。

- 只有 Open API 才会看到此属性
- 控制台删除一个伸缩组时，会默认采用强制删除的模式

如果 Forcedelete 属性为 false，必须满足以下两个条件，才能删除伸缩组：

伸缩组没有任何伸缩活动正在执行。

伸缩组当前的 ECS 实例数量（Total Capacity）为 0。

满足以上条件，会先停止伸缩组，最后再删除伸缩组。

当 Forcedelete 属性为 true 时：

先停止伸缩组，拒绝接收新的伸缩活动请求。

然后等待已有的伸缩活动完成。

最后将伸缩组内所有 ECS 实例移出伸缩组，并删除伸缩组。您手工添加的 ECS 实例会被移出伸缩组，弹性伸缩自动创建的 ECS 实例会被自动删除。

删除伸缩组，包含删除相关联的伸缩配置、伸缩规则、伸缩活动、伸缩请求的信息。

删除伸缩组，不会删除以下任务或实例：定时任务、云监控报警任务、负载均衡实例、RDS 实例。

示例

伸缩组管理列表 深圳 北京 杭州 青岛 刷新 创建伸缩组

伸缩组	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
33	启用	33	0	0	0	300	管理 修改 停用 删除
22	启用	22	0	0	0	300	管理 修改 停用 删除
11	启用	11	0	0	0	300	管理 修改 停用 删除
多组	停用	111	63	0	0	300	管理 修改 启用 删除
单组3	停用	111	40	0	0	300	管理 修改 启用 删除

伸缩规则

创建伸缩规则

什么是伸缩规则

伸缩规则指弹性伸缩服务弹性扩张或收缩 ECS 资源时所依据的规则，目前支持以下三种规则：

- 调整至 N 台：执行伸缩规则后，服务中的实例数将被调整至 N 台
- 增加 N 台：执行伸缩规则后，服务中的实例数在当前数量基础上增加 N 台
- 减少 N 台：执行伸缩规则后，服务中的实例数在当前数量基础上减少 N 台

说明：在一个伸缩组内，您最多可以创建 50 条伸缩规则。有关执行方法，请参考 执行伸缩规则。

如果执行伸缩规则后，伸缩组内服务中实例数将大于 *伸缩最大实例数* 或者小于 *伸缩最小实例数*，弹性伸缩服务会自动调整扩张或收缩的 ECS 实例数量，保证弹性伸缩结果不会超过数量限制。

示例

- 您拥有一个伸缩组 asg-bp19ik2u5w7esjcucu28，伸缩最大实例数为 3，伸缩规则 add3 为增加 3 台实例。如果当前服务中实例数为 2，执行伸缩规则 add3 时，仅会增加 1 台 ECS 实例。
- 您拥有一个伸缩组 asg-bp19ik2u5w7esjcucu28，伸缩最小实例数为 2，伸缩规则 reduce2 为减少 2 台实例。如果当前服务中实例数为 3，执行伸缩规则 reduce2 时，仅会减少 1 台 ECS 实例。

创建伸缩规则

请参照以下步骤创建伸缩规则：

- 进入 **弹性伸缩控制台**，在伸缩组的 **操作** 列下单击 **管理**。

伸缩组名称/ID	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
yk201807141052 asg-bp19ik2u5w7esjcucu28	停用		0	0	5	300	管理 修改 添加伸缩配置 启用 删除

- 前往 **伸缩规则** 页签，单击 **创建伸缩规则**。

基本信息

ECS实例列表

伸缩活动

伸缩配置

伸缩规则

事件通知

伸缩规则

每个伸缩组中，您最多可以拥有50个伸缩规则。

伸缩规则	调整类型	调整值	冷却时间	操作
remove2	减少	2台	-	查看详情 执行 修改 删除
to5	调整至	5台	-	查看详情 执行 修改 删除

- 在 **创建伸缩规则** 对话框中，填写 **规则名称**，配置 **规则**，并视需要填写 **冷却时间**，然后单击 **创建伸缩规则** 按钮。

说明： **冷却时间** 为可选配置项，如果不填写，则默认为伸缩组的冷却时间。

创建伸缩规则

*规则名称：

add3

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“-”

*规则：

增加

3

台

“增加”或“减少”选项最多只能支持100台，否则报错

“调整至N台”以及“增加或减少N%”一次只能触发100台实例的伸缩活动

冷却时间（秒）：

创建伸缩规则

取消

创建完成后，自动跳转至 伸缩规则 页签并显示刚刚创建的伸缩规则。

伸缩规则	调整类型	调整值	冷却时间	操作
add3	增加	3台	-	查看详情 执行 修改 删除
remove2	减少	2台	-	查看详情 执行 修改 删除
to5	调整至	5台	-	查看详情 执行 修改 删除

修改查询删除伸缩规则

修改、查询、删除伸缩规则的按钮如图所示。操作比较直观，不再赘述。

基本信息

ECS实例列表

伸缩活动

伸缩配置

伸缩规则

伸缩规则

创建伸缩规则

每个伸缩组中，您最多可以拥有10个伸缩规则。

伸缩规则	调整类型	调整值	冷却时间	操作
test	调整至	20台	300秒	查看详情 执行 修改 删除

共有1条，每页显示：20条

触发任务

执行伸缩规则

本文介绍如何执行一条伸缩规则。

限制条件

如果您需要执行伸缩规则，请注意以下限制条件：

- 伸缩规则所在的伸缩组的状态必须为启用。
- 伸缩规则所在的伸缩组内不能存在执行中的伸缩活动。
- 对于所有地域和所有伸缩组，一个账号最多可以弹性伸缩 1000 台 ECS 实例。此数量只包含自动创建的 ECS 实例，不包含手动添加的 ECS 实例。
- 弹性伸缩服务会自动调整扩张或收缩的 ECS 实例数量，保证弹性伸缩结果不会超过数量限制。

目前，您可以通过三种方式执行一条伸缩规则：

- 通过定时任务执行
- 通过报警任务执行
- 手动执行

通过定时任务执行

在创建定时任务时，选择一条伸缩规则。在指定的时间点，弹性伸缩服务将自动执行该伸缩规则。

创建定时任务

*任务名称：

timer07220900

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"-"

描述：

7月22日上午9点执行定时任务

最少需要2个字符

*执行时间：

2018-07-22

09

:

00

*伸缩规则：

伸缩组：

classic

伸缩规则：

add1

重试过期时间(秒)：

600

▶ 重复周期设置 (高级)

提交

取消

通过报警任务执行

在创建报警任务时，选择一条报警触发规则。在满足报警条件时，弹性伸缩服务将自动执行该伸缩规则。

创建报警任务

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。[查看帮助文档](#)

*任务名称：

alarm-cpu

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“-”

描述：

CPU使用率过高时执行报警任务

最少需要2个字符

*监控资源：

classic

监控类型

☒ 系统监控 ☐ 自定义监控

*监控项：

CPU

统计周期(分钟)：

1

*统计办法：

平均值

>=

阈值

70

%

重复几次后报警：

5次

*报警触发规则：

add1

提交

取消

手动执行

当伸缩组内不存在执行中的伸缩活动时，手动执行可以绕过冷却时间。请参照以下步骤手动执行一条伸缩规则：

进入弹性伸缩控制台，在伸缩组的 **操作** 列下单击 **管理**。

伸缩组名称/ID	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
classic asg-uf6f3xewm3dvr4bsy7r1	启用	win2016-yk	5	0	5	300	管理 修改 停用 删除

前往 **伸缩规则** 页签，在伸缩规则的 **操作** 列下单击 **执行**。

基本信息

ECS实例列表

伸缩活动

伸缩配置

伸缩规则

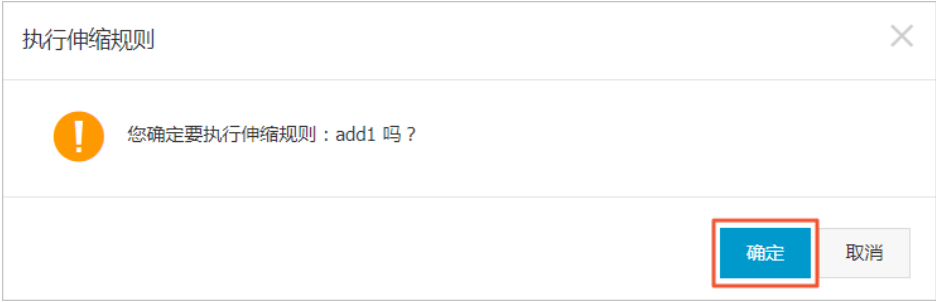
伸缩规则

每个伸缩组中，您最多可以拥有50个伸缩规则。

伸缩规则	伸缩类型	伸缩值	冷却时间	操作
add1	增加	1台	-	查看详情 执行 修改 删除

创建伸缩规则

在 **执行伸缩规则** 对话框中，单击 **确定**。



如果执行成功，页面右上角会出现成功提示。



如果执行失败，会弹出错误提示。



您可以前往 **伸缩活动** 页签，查看具体的成功或失败信息。

基本信息	伸缩活动							列表	图表
ECS实例列表									
伸缩活动	伸缩活动	变化后总实例数	开始时间	停止时间	描述	状态(全部)	操作		
伸缩配置	asa-uf66h8tstnmaajac3yye	5	2018年7月21日 13:33	2018年7月21日 13:35	Add "1" ECS ins...	成功	查看详情		
伸缩规则	asa-uf69tma1q8erhfb0wsod	4	2018年7月21日 13:33	2018年7月21日 13:33	Remove "1" ECS ...	成功	查看详情		
事件通知	asa-uf69tma1q8erhfb0wsoc	-	2018年7月21日 13:32	2018年7月21日 13:32	Remove "1" ECS ...	拒绝	查看详情		
生命周期挂钩	asa-uf69tma1q8erhfb0wsao7	-	2018年7月21日 13:31	2018年7月21日 13:31	Invalid scaling...	拒绝	查看详情		

添加 ECS 实例

本文介绍如何向伸缩组添加一个 ECS 实例，有关详细的 ECS 实例配置，请参阅 创建伸缩配置。您可以通过两种方式添加 ECS 实例：

- 执行伸缩规则自动创建一台或多台 ECS 实例。自动创建的 ECS 实例会自动满足当前伸缩配置，您无需担心产生规格限制问题。
- 手动添加一台或多台 ECS 实例。手动添加的 ECS 实例配置与当前伸缩配置没有关联。

下面将详细介绍手动添加 ECS 实例的过程。

限制条件

如果您手动添加 ECS 实例，请确保待添加的 ECS 实例满足以下条件：

- 必须与伸缩组处于同一个地域。
- 不能已加入到其它伸缩组中。
- 状态必须是 运行中。
- 可以为经典网络类型或 VPC 类型，但有以下限制：
 - 当伸缩组为经典网络类型时，只能添加经典网络类型的 ECS 实例。
 - 当伸缩组为 VPC 类型时，只能添加同一 VPC 下的 ECS 实例。

同时，请确保伸缩组满足以下条件：

- 状态必须是 启用。
- 不能存在执行中的伸缩活动。

说明：弹性伸缩服务尽力保证足额弹出待添加的 ECS 实例，但是，如果出现云服务器库存不足、待添加的 ECS 实例数超过伸缩组上限等问题，ECS 实例会无法足额弹出。这种情况下，请您检查伸缩组相关配置定位问题。如果无法解决问题，请[提交工单](#)寻求帮助。

操作步骤

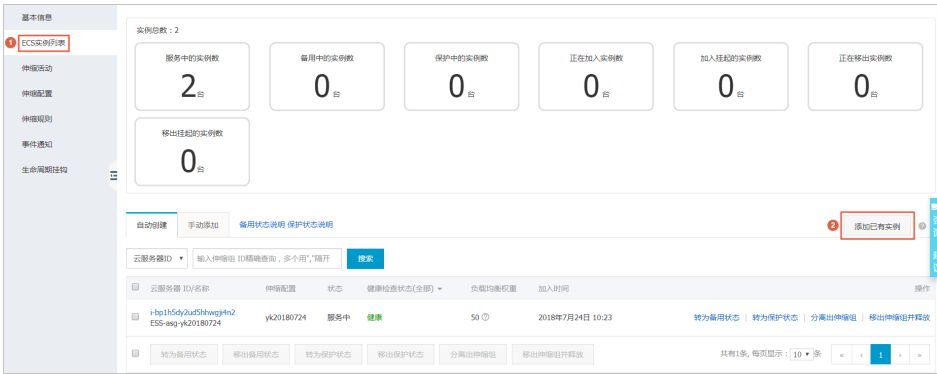
手动添加 ECS 实例可以绕过冷却时间。请参照以下步骤手动添加 ECS 实例：

登录弹性伸缩控制台。

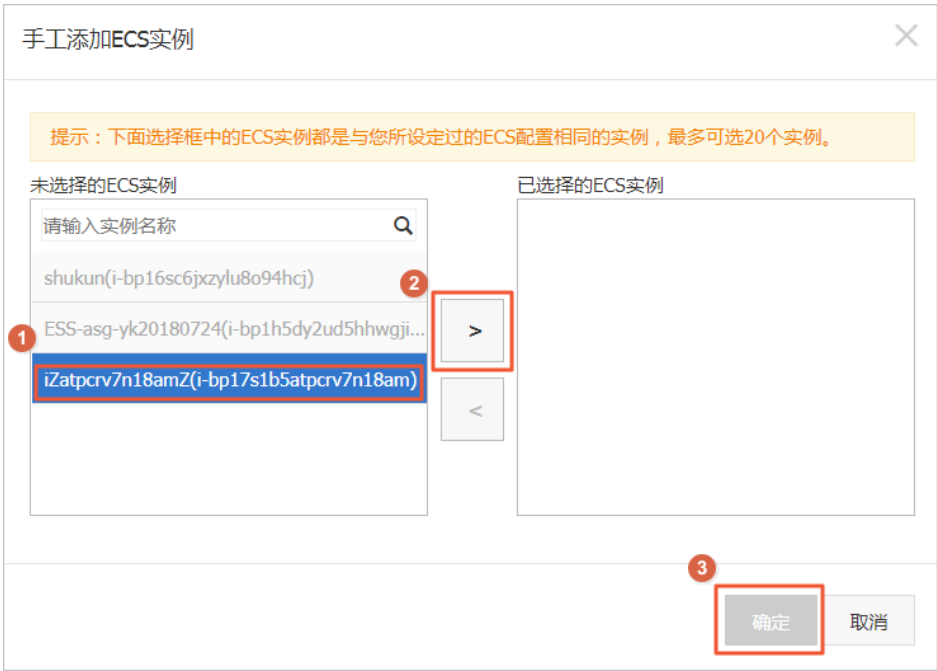
在 **伸缩组管理** 页面，单击指定伸缩组 **操作** 列下的 **管理**。



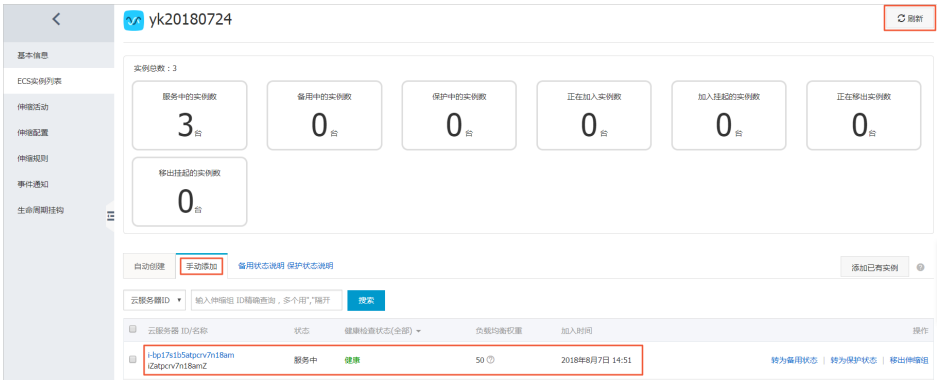
前往 **ECS 实例列表**，单击 **添加已有实例**。



从左侧列表中选择可用的 ECS 实例，单击 > 添加到伸缩组，然后单击 **确定**。



前往 **手动添加** 页签查看结果。



说明：如果页面没有自动刷新，请单击页面右上角的 **刷新** 进行手动刷新。

移出ECS实例

从指定的伸缩组里移出 ECS 实例。

- 从伸缩组移出弹性伸缩自动创建的 ECS 实例时，停止和释放该 ECS 实例。
- 从伸缩组移出用户手工加入的 ECS 实例时，不停止和释放该 ECS 实例。

执行该功能，必须同时满足以下条件：

- 伸缩组为生效（Active）状态。
- 当伸缩组没有伸缩活动正在执行。

当伸缩组没有伸缩活动正在执行时，该功能可以绕过冷却时间（Cooldown）直接执行。

执行功能返回成功，只是表示弹性伸缩服务接受了该功能的调用请求，伸缩活动可以执行，但不代表伸缩活动能够执行成功。用户需要通过返回的 ScalingActivityId 查看该伸缩活动的执行状态。

如果当前伸缩组的实例数（Total Capacity）减去该功能指定的实例数小于 MinSize 时，则执行失败。

示例

The screenshot shows the 'phpwind-cms-pa...' scaling group page. The left sidebar has tabs for '基本信息', 'ECS实例列表' (selected), '伸缩活动', '伸缩配置', and '伸缩规则'. The main area displays '实例总数: 2' and three boxes for '正在服务中的实例数' (2), '正在加入实例数' (0), and '正在移出实例数' (0). Below this is the 'ECS实例列表' table with columns: '云服务器名称', '伸缩配置', '状态', '健康检查状态(全部)', '加入时间', and '操作'. Two instances are listed, both in '服务中' (Running) and '健康' (Healthy) status. The '移出伸缩组并释放' (Remove from scaling group and release) button is highlighted with a red box. At the bottom, there is a pagination bar showing '共有2条, 每页显示: 10条' and a page number '1'.

云服务器名称	伸缩配置	状态	健康检查状态(全部)	加入时间	操作
ESS-gg-phpwind-cms-patch-ecs-i25pxu6uxw	test	服务中	健康	2014-10-07 17:11	移出伸缩组并释放
ESS-gg-phpwind-cms-patch-ecs-i25dkmmajr	test	服务中	健康	2014-10-05 17:11	移出伸缩组并释放

定时任务

创建定时任务

根据传入参数创建定时任务。每个用户最多能创建 20 个定时任务。

由于伸缩组正在发生伸缩活动或者伸缩组停用等原因，导致定时任务触发执行伸缩规则失败后，在重试过期时间（LaunchExpirationTime）内，定时任务会自动重试触发，否则放弃本次定时触发。

如果多个定时任务在相近的时间内触发执行同一个伸缩组的伸缩规则，则最早触发的定时任务会先执行伸缩活动，由于同一个伸缩组同一时刻内只能有一个伸缩活动，排在后面的定时任务会在 LaunchExpirationTime 内自动重试定时触发。如上一个伸缩活动完成之后，排在后面的定时任务还在 LaunchExpirationTime 内重试，则执行该定时任务的伸缩规则并触发相应的伸缩活动。

如果指定了同一分钟的定时任务，会选取最近创建的定时任务执行。

操作步骤

单击 **自动触发任务管理** 目录下的 **定时任务** 子目录，显示定时任务列表页。



单击 **创建定时任务**，弹出 **创建定时任务** 对话框。

填写任务名称。

填写执行时间。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。

填写重复周期。

选择定时任务需要触发的伸缩规则以及所属的伸缩组。

单击 **提交**，将在 **定时任务列表** 页展示该定时任务。

创建定时任务

*任务名称：

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“-”

描述：

最少需要2个字符

*执行时间：

2015-08-08 18 : 26

*伸缩规则：

伸缩组： -- 请选择伸缩组 --

重试过期时间(秒)：

600

重复周期设置（高级）

重复周期：按日

每 天执行一次

重复结束时间：

2015-09-06 18 : 26

提交

取消

修改、查询、删除定时任务

描述

修改定时任务的属性、查询定时任务的信息、删除一个指定的伸缩规则

弹性伸缩服务 ESS

定时任务 深圳 北京 杭州 青岛

刷新 创建定时任务

定时任务	描述	状态	伸缩组信息	执行时间	重复周期	重复停止时间	重试过期时间	操作
test	test	运行中	伸缩组：33 伸缩规则：test	2015-08-08 19:32	每周执行一次	2015-09-06 19:32	600秒	停用 修改 删除

共有1条， 每页显示：20条

<< 1 >>

报警任务

弹性伸缩报警任务

弹性伸缩（Auto Scaling）报警任务是弹性伸缩与云监控服务（CMS）深度合作，提供的一种动态管理伸缩组的方式，类似于弹性伸缩定时任务，弹性伸缩报警任务通过触发您指定的伸缩规则来执行伸缩活动，达到调整伸缩组内实例个数的目的。

定时任务可以在您指定的时间执行您指定的伸缩规则，当业务场景在时间上可预料时，能够提前做出响应，但是，在面对突发或者时间上不可预料的业务场景时，定时任务就显得捉襟见肘，此时，就需要报警任务来提供更灵活的触发伸缩规则的方式，在业务高峰期增加伸缩组内实例数量来缓解业务压力，在业务低谷时释放伸缩组内实例，减小生产成本。

报警任务通过监控特定的监控指标，对数据指标进行实时的统计，当统计值满足您指定的报警条件时，触发报警，执行您指定的伸缩规则。使用报警任务，您可以实时的根据业务的变化来不断调整伸缩组内的实例数量，保证您监控的指标维持在您期望的范围内。

弹性伸缩报警任务为您提供了一种通过监控特定监控指标来动态调整伸缩组内实例数量的方法，让您能够根据业务的变化实时的执行指定的伸缩规则，调整伸缩组内的实例数量。

弹性伸缩报警任务升级版

弹性伸缩报警任务已经全新升级，从监控范围、监控方式、监控响应速度等三个方面做出了全方位的优化，升级后的弹性伸缩报警任务将为您提供一种更全面，更可靠的利用报警任务动态管理伸缩组的方式。

升级内容主要包括：

- 增加了对系统磁盘，网卡，TCP连接数等监控指标的报警任务支持。
- 报警任务最小统计周期升级到 1分钟，提供更灵敏的监控报警。
- 增加自定义监控，为用户自有监控系统接入弹性伸缩报警任务提供标准化方式。

弹性伸缩报警任务升级版扩展了原有的监控指标，并在原有监控项的基础上，支持用户接入自定义的监控项，提供定制化的报警任务，大大增强了弹性伸缩报警任务的可用性和实用性，满足用户具体的，多样化的需求。

弹性伸缩系统监控报警任务

弹性伸缩系统监控报警任务的监控指标是云监控为用户采集的 ECS 实例的相关数据指标，比如CPU，负载等。用户在弹性伸缩中设置的系统监控的报警任务是以伸缩组作为监控粒度的，即以伸缩组内的所有实例的监控指标的统计平均值作为伸缩组的指标值，当伸缩组内实例数量发生变化时，监控指标也会同时进行更新。

支持的监控指标

弹性伸缩系统监控报警任务目前支持的监控项，其中红色字体代表升级版新增支持的监控项：

监控项	单位
CPU	%
内存	%
系统平均负载	无
内网出流量	KB/min
内网入流量	KB/min
外网出流量	KB/min
外网入流量	KB/min
系统盘写bps	Byte/s
系统盘读bps	Byte/s
系统盘写iops	个/s
系统盘读iops	个/s
外网网卡发包数（经典网络）	个/s
外网网卡收包数（经典网络）	个/s
内网网卡发包数	个/s
内网网卡收包数	个/s
TCP总连接数	个
TCP已建立连接数	个

注意事项

- 伸缩组在同一时刻只能执行一个伸缩活动，当伸缩组内存在正在执行的伸缩活动是，由报警任务触发伸缩规则产生的伸缩活动将被拒绝。
- 弹性伸缩报警任务在触发伸缩规则时受伸缩规则冷却时间影响，伸缩规则在冷却时间内时，将拒绝执行伸缩规则。伸缩组内新增加的 ECS 实例从加入伸缩组到完成系统启动配置，部署用户业务，获取到监控数据需要一定的时间（通常需要几分钟），因此您应该根据具体的业务场景，设置合适的冷却时间，防止在新增实例的监控数据缺失的这段时间中，重复触发伸缩规则。
- 弹性伸缩报警任务，默认设置了一分钟的沉默时间，即触发报警之后，一分钟内不会再次触发伸缩规则。
- 部分系统监控项（内存、负载、网卡发包数、TCP连接数）指标的采集需要为您安装云监控客户端。默认情况下，当您针对需要云监控客户端采集的监控项设置报警任务时，将为报警任务关联的伸缩组内的所有实例安装云监控客户端，同时，将为您在云监控控制台开启 新购ECS自动安装云监控，为您所有新购的 ECS 实例安装云监控客户端。

弹性伸缩自定义监控项报警任务

弹性伸缩自定义监控项报警任务的监控对象为用户自主上报到云监控中的监控指标。在一些场景下，系统监控项可能不包含您所需要的监控指标，您可能拥有自己的一套监控系统，并且关心的是与您特定业务相关的某些指标。自定义监控报警任务，为您自有的监控系统，或者与业务相关的自有监控指标提供了设置报警任务的接入点。

弹性伸缩自定义监控报警任务是针对阿里云云监控服务的自定义监控项设置报警的，用户在使用弹性伸缩自定义监控报警任务之前需要首先向云监控上报自定义监控数据，即自定义监控项。云监控自定义监控是提供给用户自由定义监控项及报警规则的一项服务，通过此服务，用户可以针对自己关心的业务指标进行监控，将采集到监控数据上报至云监控，由云监控来进行数据的处理，并可以对其设置报警规则。

上报监控数据到云监控

云监控的自定义监控服务为您提供了上报监控数据的方式，您可以将自己采集到的时序数据上报到云监控，这样的数据称作时间序列。云监控提供了OpenAPI、Java SDK 和阿里云命令行工具（CLI）三种方式上报数据，这里我们将主要关注如何使用Java SDK的方式上报监控数据。更多详细的信息您可以查看文档 [上报监控数据](#)。

使用Java sdk之前您需要首先在项目中引入相应的jar包，如果您使用maven管理项目，您只需要在项目中加入以下依赖：

```
<dependency>
<groupId>com.aliyun</groupId>
<artifactId>aliyun-java-sdk-core</artifactId>
<version>3.2.6</version>
</dependency>
<dependency>
<groupId>com.aliyun.openservices</groupId>
<artifactId>aliyun-cms</artifactId>
<version>0.2.4</version>
</dependency>
```

您可以按照如下的方式向云监控上报自定义监控项：

```
static String endPoint = "https://metrichub-cms-cn-hangzhou.aliyuncs.com";
CMSClient cmsClient = new CMSClient(endPoint, accAutoScalingKey, accAutoScalingSecret);
CustomMetricUploadRequest request = CustomMetricUploadRequest.builder()
.append(CustomMetric.builder()
.setMetricName("myCustomMetric")//自定义指标名
.setGroupId(54504L)//设置分组id
.setTime(new Date())//时间
.setType(CustomMetric.TYPE_VALUE)//类型为原始值
.appendValue(MetricAttribute.VALUE, number)//原始值，key只能为这个
.appendDimension("key1", "value1")//添加维度
```



```
.appendDimension("key2","value2")
.build()
.build();
CustomMetricUploadResponse response = cmsClient.putCustomMetric(request);//上报
```

在上述的代码片段中，我们上报了一个数据点到云监控。上报时，必须指定groupId参数，即云监控应用分组id，该分组id可以是您在云监控中已创建的应用分组，也可以是一个不存在的应用分组。您可以在 [云监控应用分组](#) 创建和查看应用应用分组信息。您上报的自定义监控项（时间序列），可以在云监控控制的 [自定义监控](#) 中查看。

我们建议您向一个云监控中已经存在的应用分组中推送自定义监控数据，云监控的应用分组是针对多种云产品的逻辑分组，向一个已存在的应用分组推送自定义监控数据将使您在需要使用云监控等相关功能时保留扩展的能力。当然，您也可以完全不必理会应用分组的概念，选择向任意一个分组id推送数据。

您上报到云监控的监控数据，云监控会自动帮您进行聚合，当然，当您需要推送的数据量太大时，您也可以选择在本地聚合之后再推送到云监控。详细信息可以参考 [上报监控数据](#)。

注意事项

云监控对用户上报监控数据设置了以下限制：

- 单云账号QPS限制为100。
- 单次最多上报100条数据，body最大为256KB。
- “metricName” 字段只支持字母、数字、下划线。需要以字母开头，非字母开头会替换为大写“A”，非法字符替换为“_”。
- “dimensions” 字段不支持“=”、“&”、“/”，非法字符会被替换为“_”。
- metricName 和dimensions的Key-value 最大均为64字节，超过64字节会被截断。
- 其他限制请关注 [计量计费](#) 说明。

创建报警任务

本文介绍如何创建一个报警任务，目前报警任务分为两种类型：系统监控项报警任务和自定义监控项报警任务。

创建系统监控项报警任务

进入 [弹性伸缩控制台](#)，分三步创建报警任务：

1. 选中报警任务侧边栏。
2. 点击创建报警任务。
3. 在弹出的创建报警任务对话框中填写定制信息。



示例

创建报警任务

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。[查看帮助文档](#)

*任务名称:

test_cpu_alarm

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“-”

描述:

cpu utilization

最少需要2个字符

*监控资源:

classic

监控类型

☒ 系统监控 ☐ 自定义监控

*监控项:

CPU

统计周期(分钟)

1

*统计办法

平均值

>=

阈值

50

%

重复几次后报警

3次

*报警触发规则

add1

提交

取消

上图中的报警任务信息含义如下：

- test_cpu_alarm 为任务名称，cpu utilization 为任务描述。
- classic 为监控资源，即报警任务监控的伸缩组。
- 系统监控 为监控类型。
- CPU（CPU使用率）为监控项。
- 1 分钟为统计周期，表示将对每1分钟的监控数据进行统计检查，判断是否满足报警条件。
- 平均值>=50% 为统计方法，重复 3次。这表示，当cpu的使用了在1分钟平均值超过阈值50%，并且连续3次满足统计办法时，将发生报警。
- 伸缩规则add1为报警触发规则，表明在发生报警时，将执行报警规则add1，为伸缩组增加1台机器。

创建自定义监控项报警任务

创建自定义监控项报警任务的过程和创建系统监控类似，唯一的不同在于，系统监控报警任务的监控项是由云监控为用户采集，而自定义监控报警任务的监控项则需要用户自己上报至云监控。

在创建自定义监控类型的报警任务时，首先需要存在已上报的自定义监控项，即时间序列。然后才能针对该时间序列设置报警规则。

创建报警任务

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。[查看帮助文档](#)

*任务名称:

test_custom_metric

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"-"

描述:

测试自定义监控

最少需要2个字符

*监控资源:

classic

监控类型

☐ 系统监控 ☒ 自定义监控

*应用分组:

54504

*监控项:

testMetric

纬度:

age=10

统计周期(分钟)

1

*统计办法

平均值

>=

阈值

0

%

重复几次后报警

3次

*报警触发规则

add1

提交

取消

在上图创建自定义报警任务之前，已经向云监控推送了一条自定义监控数据流（时间序列）。该时间序列所属的应用分组为54504，监控项名称为testMetric，维度信息为 age=10。

查看修改删除报警任务

您可以在 弹性伸缩控制台 对已创建的报警任务进行查看、修改和删除操作。

查看监控项详情

在成功创建报警任务之后，我们可以在报警任务列表看到刚刚创建的报警任务。

在系统监控列表，查看创建的系统监控报警任务。

报警任务

刷新

创建报警任务

报警任务	状态	监控资源	监控类型	统计周期	触发伸缩条件规则	触发规则数量	操作
test_cpu_alarm	正常	伸缩组: classic	系统监控	1分钟	CPU 使用率 平均值 连续 3 次 >= 50.0 %	1个	查看详情 停用 删除 更多操作

切换到自定义监控列表，查看创建的自定义报警任务。

报警任务

报警任务管理

自动创建报警任务管理

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

报警任务

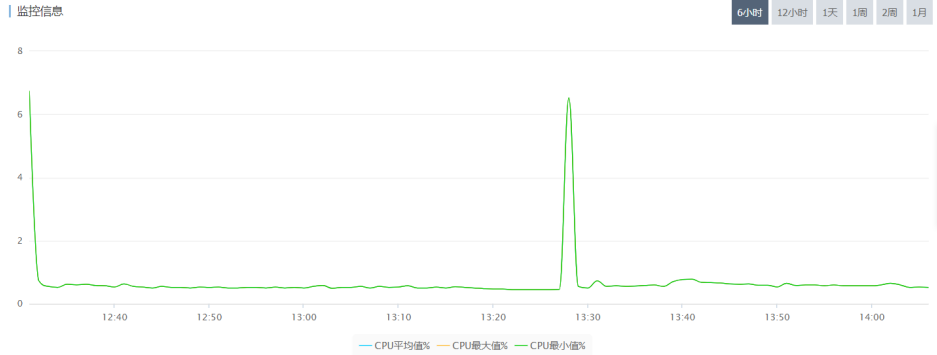
报警任务

报警任务

报警任务

报警

点击报警任务的名称，将跳转到详情页，在详情页您将看到报警任务对应监控项的历史数据。



修改报警任务

您可以在报警任务的列表页修改报警任务，也可以进入报警任务的详情页修改报警规则。

报警任务							刷新	创建报警任务
系统监控		自定义监控						
报警任务	状态	报警资源	报警类型	统计周期	触发报警条件规则	触发规则数量	操作	
test_cpu_alarm	正常	伸缩组: classic	系统监控	1分钟	CPU 使用率 平均值 连续 3 次 >= 50.0 %	1个	查看详情 停用 删除 更多操作	
共有1条, 每页显示: 10条							修改报警任务 修改报警触发规则	

报警任务详情 [返回报警任务列表](#)

基本信息

修改基本信息

报警名称：	test_cpu_alarm	是否触发规则	☒
报警状态：	正常	统计周期	1m
监控资源：	伸缩组：：classic		
监控类型	系统监控		
触发伸缩规则的条件：	CPU 使用率 平均值 连续 3 次 >= 50.0 %		
描述：	cpu utilization		

报警触发规则

您最多可以建5条触发规则，目前还可以添加4条伸缩规则

报警触发规则

规则名称	ari:acs:ess:cn-qingdao:1406926474064770:scalingrule/asr-m5e9omkfub3ypueq13
------	--

修改报警任务分为两个部分，一个是修改报警任务的基本信息，另外一个修改报警规则的触发规则。

修改基本信息包括修改任务名称、监控项、统计周期、统计办法、重复次数等，我们不建议您修改报警任务的监控项，因为，修改监控项意味着监控不同的指标，此时，新建一条针对新指标的报警任务是更合适的方法。

删除报警任务

您可以在报警任务的列表页删除报警任务。

报警任务 [刷新](#) [创建报警任务](#)

系统监控		自定义监控							
报警任务	状态	监控资源	监控类型	统计周期	触发伸缩条件规则	触发规则数量	操作		
test_cpu_alarm	正常	伸缩组：classic	系统监控	1分钟	CPU 使用率 平均值 连续 3 次 >= 50.0 %	1个	查看详情	删除	更多操作

查看伸缩活动

查询伸缩活动的信息，最多只能返回30天之内的伸缩活动。

示例

<

33

刷新

基本信息

ECS实例列表

伸缩活动

伸缩配置

伸缩规则

伸缩活动

伸缩活动

变化后总实例数

开始时间

停止时间

描述

状态(全部)

操作

dVdlgqc4dG8RcCCbPLdvpUgv	0	2015-02-19 16:59	2015-02-19 17:00	Remove "100" EC...	成功	查看详情
beMtm3cVH05xcNxtHwdJbhnQ	100	2015-02-15 17:25	2015-02-15 17:27	Add "2" ECS ins...	成功	查看详情
dSN11Rcf1sjQc1eJzNdoQpD4	98	2015-02-15 17:22	2015-02-15 17:24	Add "37" ECS in...	部分成功	查看详情
QHCoScgHwRWdx3HwgbQZu1v	63	2015-02-15 17:18	2015-02-15 17:21	Add "100" ECS i...	部分成功	查看详情

共有4条， 每页显示：20条

«

<

1

>

»

生命周期挂钩

如果您为伸缩组创建了生命周期挂钩，在伸缩组进行伸缩活动时，正在加入或正在移出伸缩组的实例将被挂钩挂起并置于等待状态。关于适用的伸缩活动类型，请参考 表1 配置参数。

功能特性

在实例保持等待状态的时间内，当前伸缩组将具有以下特性：

- 伸缩组不再运行其它的伸缩活动，包括报警任务、定时任务等触发的伸缩活动。
- 为您保留指定时长的操作时间（即挂钩的超时时间，请参考 表1 配置参数），您可以在挂起期间执行自定义操作，例如，初始化 ECS 实例配置或者获取 ECS 实例数据。
- 您可以调用相关 OpenAPI 来结束生命周期活动或者删除生命周期挂钩。
- 您可以删除生命周期挂钩来恢复执行伸缩活动。

使用限制

在一个伸缩组内，您最多可以拥有 6 个生命周期挂钩。

创建生命周期挂钩

参照以下步骤，创建一个新的生命周期挂钩：

1. 进入 弹性伸缩控制台。

前往 伸缩组管理 页签，单击要添加挂钩的伸缩组 操作 列下的 管理。

弹性伸缩

伸缩组管理

刷新

创建伸缩组

伸缩组管理

伸缩组支持事件通知

伸缩配置支持多可用区实例和跨价实例

伸缩组实例支持备用状态

伸缩组名称

输入伸缩组名称模糊查询

确定

伸缩组名称/ID	状态	伸缩配置	总实例数	最小实例数	最大实例数	默认冷却时间(秒)	操作
classic-aag-m5elyQ5ajZ24r1hc3l	启用	classic	1	1	1	300	管理 修改 停用 删除

上一步 创建弹性伸缩

前往 **生命周期挂钩** 页签，单击 **创建生命周期挂钩**。



在 **创建生命周期挂钩** 对话框中配置必要的参数，如 **表1 配置参数** 所示，然后单击 **创建生命周期挂**

创建生命周期挂钩

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"-"

*适用的伸缩活动类型：☒ 弹性收缩活动 ☐ 弹性扩张活动

超时时间（秒）：

最小为30，最大为21600，必须为整数

执行策略：☒ 继续 ☐ 拒绝 ?

通知方式：☒ MNS主题 ☐ MNS队列

*MNS主题：

通知标识 ?：

通知标识的长度不超过 128 个字符

创建生命周期挂钩 取消

钩。

表1 配置参数

配置参数	参数说明	示例
名称	生命周期挂钩名称（2-40个字符），以大小写字母、数字或中文开头，可包含"."、"_"或"-"。挂钩名字一旦确定，不能修改。	hz_yk0626
适用的伸缩活动类型	适用的伸缩活动类型包括 弹性收缩活动 和 弹性扩张活动 。	弹性收缩活动
超时时间	默认情况下，实例保持等待状态的时间。最小为 30 秒，最大为 21600 秒（6 小时），必须为整数。	600
执行策略	执行策略包括 继续 和 拒绝 。	继续

	- 继续：继续执行弹性扩张活动或者弹性收缩活动。 - 拒绝：针对弹性扩张活动，会直接释放创建出来的 ECS 实例；针对弹性收缩活动，无影响。	
通知方式	通知方式包括 MNS主题 和 MNS队列 。选择通知方式后，您还需要选择具体的MNS主题或者MNS队列，如果尚未创建MNS主题或者MNS队列，请 创建主题 或者 创建队列 。	MNS主题
通知标识	弹性伸缩每次推送消息到通知对象时，会同时发送您预先指定的通知标识，便于管理和标记不同类别的通知信息。	General information

在生命周期挂钩页签可查看到新添加的生命周期挂钩。



修改生命周期挂钩

参照以下步骤，修改一个生命周期挂钩的参数：

在 **生命周期挂钩** 页签，选择要修改的挂钩 **操作** 列下的 **修改**。



在 **修改生命周期挂钩** 对话框中，根据需要修改参数，然后单击 **修改生命周期挂钩**。修改时的参数配置请参考 [创建生命周期挂钩](#)，不能修改挂钩名称，其他参数均可修改。

修改生命周期挂钩

名称：

hz_yk0626

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"-"

*适用的伸缩活动类型：

☐ 弹性收缩活动

☒ 弹性扩张活动

超时时间（秒）：

3600

最小为30，最大为21600，必须为整数

执行策略：

☐ 继续

☒ 拒绝

?

通知方式：

☐ MNS主题

☒ MNS队列

*MNS队列：

mns-queue-cn-hangzhou

通知标识：

Running error

通知标识的长度不超过 128 个字符

修改生命周期挂钩

取消

删除生命周期挂钩

参照以下步骤，删除一个生命周期挂钩：

在 **生命周期挂钩** 页签，选择要删除的生命周期挂钩 **操作** 列下的 **删除**。

生命周期挂钩名称	适用的伸缩活动类型	通知方式	通知标识	超时时间	执行策略	操作
hz_yk0626	弹性扩张活动	MNS队列：mns-queue-cn-hangzhou	Running er...	3600秒	拒绝	修改删除

在 **删除生命周期挂钩** 对话框中，单击 **确定** 完成删除操作。

删除生命周期挂钩

!

您确定要将生命周期挂钩 hz_yk0626 删除吗？

确定

取消

如果生命周期挂钩已触发伸缩活动等待状态，删除生命周期挂钩时，则对应的等待状态会被提前结束。

设置实例备用状态

弹性伸缩支持设置ECS实例备用状态（Standby）。我们不会检查处于备用状态的ECS实例健康状态，亦不会释放ECS实例。

功能特性

设置ECS实例备用状态后：

- ECS实例保持备用状态直至您使其退出备用状态。
- 弹性伸缩不会管理ECS实例的生命周期，而是由您管理。
- 如果对应伸缩组配置了负载均衡，则备用状态ECS实例的负载均衡权重会被置零。
- 您可以 停止 实例、重启 实例或者做其他变配操作，例如 更改ECS实例规格、更换操作系统、初始化磁盘 或者 迁移经典网络到专有网络VPC 等。
- 伸缩组内发生伸缩活动时，不会移除处于备用状态的ECS实例。
- ECS实例被停止或者重启，不会更新ECS实例的健康检查状态。
- 您需要预先从伸缩组内移除该ECS实例才能将其释放。
- 如果您删除伸缩组，ECS实例会自动解除备用状态并随伸缩组一起释放。

解除ECS实例备用状态后：

- ECS实例恢复运行状态。
- 如果伸缩组设置了负载均衡实例，则将负载均衡实例中该ECS实例的权重设置为定义值。
- ECS实例被停止或者重启时，会重新更新健康检查状态。
- 弹性伸缩继续接管ECS实例的生命周期。

控制台操作

设置备用状态

登录 弹性伸缩管理控制台。

选择 地域，如华东2。

找到并单击目标伸缩组。

在左侧导航页面，单击 ECS实例列表。

勾选目标ECS实例，单击 **转为备用状态**。

解除备用状态

登录 [弹性伸缩管理控制台](#)。

选择 **地域**，如华东2。

找到并单击目标伸缩组。

在左侧导航页面，单击 **ECS实例列表**。

勾选目标ECS实例，单击 **移出备用状态**。

API操作

设置备用状态：EnterStandby

解除备用状态：ExitStandby

相关链接

- [云栖社区](#) 伸缩伸缩支持Standby
- [什么是负载均衡](#)
- [移出不健康ECS实例](#)

查询ECS实例列表

查看 ECS 实例列表

查询伸缩组内 ECS 实例列表。

加入伸缩组的 ECS 实例有两种类型：自动创建的 ECS 实例、手工添加的 ECS 实例。

- **自动创建的 ECS 实例** 是指根据用户的伸缩配置和伸缩规则，由弹性伸缩服务自动创建的 ECS 实例。
- **手工添加的 ECS 实例** 是指不是由弹性伸缩服务创建，但由用户手工添加到伸缩组中的 ECS 实例。

ECS 实例在伸缩组中的生命周期

ECS 实例在伸缩组中的生命周期，通过以下几种状态描述：

- 加入中（ Pending ） – 表示 ECS 实例正在加入伸缩组，包括创建实例、加入负载均衡、添加 RDS 访问名单等过程。
- 服务中（ InService ） – 表示 ECS 实例已成功加入伸缩组，并正常提供服务。
- 移出中（ Removing ） – 表示 ECS 实例正在移出伸缩组。

ECS 实例的健康状态

ECS 实例在伸缩组中的健康状态为：

- 健康（ Healthy ）
- 不健康（ Unhealthy ）

本期仅通过 ECS 实例为非“运行中（ Running ）”状态来判断该实例不健康，弹性伸缩会自动移出伸缩组中不健康的 ECS 实例。对于“自动创建的 ECS 实例”，弹性伸缩会停止和释放该 ECS 实例。对于“手工添加的 ECS 实例”，弹性伸缩不会停止和释放该 ECS 实例。

示例

下图是一个示例。

