



弹性伸缩服务

用户使用指南

天翼云科技有限公司

目录

产品简介.....	4
产品定义.....	4
术语解释.....	6
生命周期.....	7
应用场景.....	8
使用限制.....	10
与其他服务的关系.....	11
产品地域和可用区.....	13
计费说明.....	16
计费模式.....	16
快速入门.....	17
弹性伸缩使用流程.....	17
准备工作.....	18
快速扩缩容弹性云主机.....	19
动态扩展云主机.....	30
定时扩展云主机.....	34
用户指南.....	37
功能总览.....	37
伸缩组管理.....	40
伸缩配置管理.....	56
伸缩策略管理.....	65

伸缩活动管理.....	76
伸缩组内实例管理.....	84
伸缩组监控.....	90
最佳实践.....	97
利用弹性伸缩为突发业务创建低成本业务集群.....	97
利用弹性伸缩为周期性业务降低成本.....	104
部署高可用集群.....	110
常见问题.....	115
计费类.....	115
伸缩组操作类.....	115
伸缩配置操作类.....	118
伸缩策略操作类.....	119
伸缩活动操作类.....	121
伸缩组内实例管理类.....	123
关联负载均衡操作类.....	123

弹性伸缩服务

弹性伸缩服务（CT-AS, Auto Scaling）是根据用户的业务需求，通过策略自动调整其弹性计算资源的一种管理服务。该服务能够根据预设规则自动调整伸缩组内的云主机数量，在业务需求上升时自动增加云主机实例，业务需求下降时自动减少云主机实例，降低人为反复调整资源以应对业务变化和高峰压力的工作量，帮助用户节约资源和人力成本。

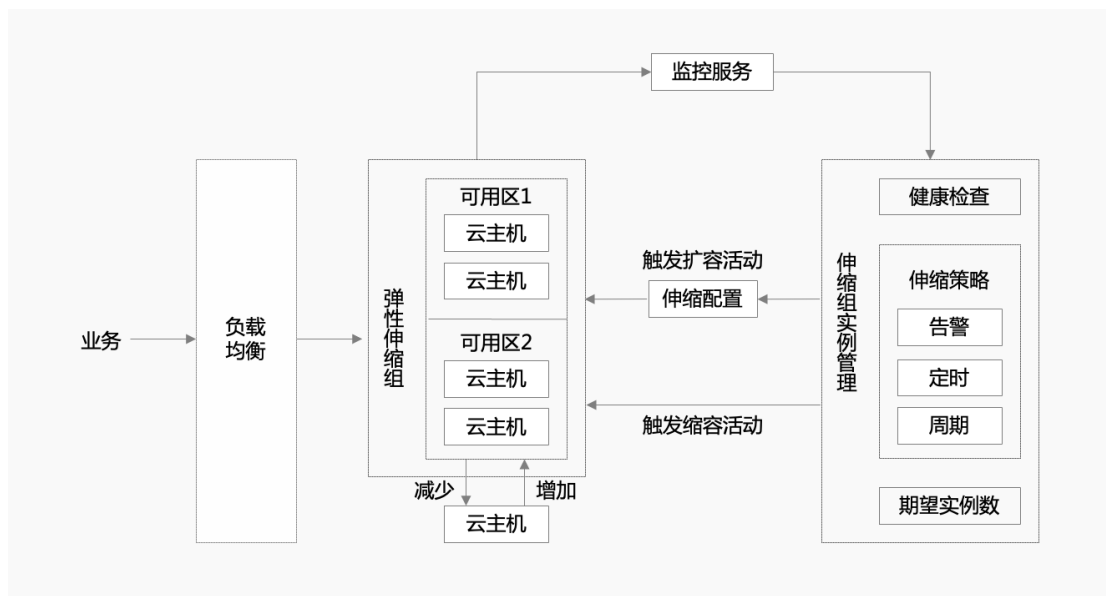
产品简介

产品定义

弹性伸缩服务（CT-AS, Auto Scaling）是根据用户的业务需求，通过策略自动调整其弹性计算资源的一种管理服务。该服务能够根据预设规则自动调整伸缩组内的云主机数量，在业务需求上升时自动增加云主机实例，业务需求下降时自动减少云主机实例，降低人为反复调整资源以应对业务变化和高峰压力的工作量，节约资源和人力成本。

工作原理

弹性伸缩服务的工作原理如下图所示：



- 业务流量通过负载均衡按照分发策略自动分发给弹性伸缩组内的弹性云主机，由多台云主机共同承载业务流量，分担计算压力。
- 用户可以通过云监控服务来监控弹性伸缩组内的云主机实例，并根据业务需求和配置方式自动调整伸缩组内的云主机实例数量。主要有三种方式：
 - 健康检查：可通过健康检查来替换伸缩组内处于“异常”状态的云主机实例。
 - 伸缩策略：可设置伸缩策略来确保伸缩组内主机实例始终满足业务计算需求，主要有三种策略，包括告警策略、定时策略、周期策略。
 - 期望实例数：可设置期望实例数将弹性伸缩组内实例数量始终维持在理想状态。

访问方式

天翼云提供如下方式进行弹性伸缩的配置和管理：

- 控制台：Web 化的服务管理平台，即利用管理控制台来配置和管理弹性伸缩。
- API：基于 API（Application programming interface）方式访问弹性伸缩，具体操作请参见[弹性伸缩 API 参考](#)。

术语解释

弹性伸缩

弹性伸缩是根据用户业务需求预先设置策略来自动调整业务中云主机资源的管理服务,当业务需求上升导致云主机负载超过阈值上限时,则自动增加弹性云主机数量以保证计算能力,当业务需求下降导致云主机负载低于阈值下限时,则自动减少弹性云主机,最大限度地节约资源与成本。

伸缩组

伸缩组是具有相同属性和应用场景的云主机实例和伸缩策略的集合。伸缩组是启停伸缩策略和进行伸缩活动的基本单位。

伸缩配置

伸缩配置是伸缩组资源扩展云主机实例的规格模板,包括云主机的类型、vCPU、内存等。

伸缩策略

伸缩策略用于定义伸缩活动的触发模式、触发条件和触发动作。伸缩策略通常有三种,告警策略、周期策略、定时策略。以下为一个定时策略示例:例如,一般情况下,凌晨 0 点到 8 点业务访问量较小,因此,为节约资源,可以添加两条策略:每天凌晨 0 点移除一台云主机;每天 8 点添加一台云主机。

伸缩活动

当执行伸缩策略、修改伸缩组参数、健康检查等操作可能会触发伸缩动作,这些动作的记录

即为伸缩活动。例如增加云主机资源或减少云主机资源的过程。

冷却时间

冷却时间是指用户在每次成功的伸缩活动后设置的一段锁定时间,当一次伸缩活动执行完成,系统开始计算冷却时间。在冷却时间内,由告警策略触发的伸缩活动将会被拒绝执行,其他类型的伸缩活动不受限制,但执行完成之后将会重新开始冷却时间计时,以避免告警策略的频繁触发。

生命周期

弹性伸缩可以根据业务需求自动调整伸缩组内的实例数量,每个实例从创建到被移出伸缩组,即为伸缩组内实例的生命周期。



- 创建实例: 伸缩组内的实例创建有两种方式,由弹性伸缩根据伸缩配置自动创建实例或由用户通过控制台、OpenAPI 创建实例。
- 运行中: 实例创建成功,处于运行中状态。
- 正在加入伸缩组: 实例正在移入伸缩组,可通过用户手动移入伸缩组或通过伸缩策略自动创建实例并移入。关联负载均衡的伸缩组,实例也会同时加入负载均衡主机组。
- 服务中: 实例移入伸缩组成功后,将进入服务中状态。
- 正在移出伸缩组: 实例正在移出伸缩组,可通过用户手动移出或伸缩策略自动移出。关联负载均衡的伸缩组,实例也会同时移出负载均衡主机组。

- 已移出：实例已移出伸缩组，不再为伸缩组服务。实例移出后，根据实例创建方式、付费方式、移出策略的差异，可能被保留或直接释放。

应用场景

弹性伸缩服务有广泛的业务场景，本文主要介绍三种典型的应用场景。

场景一：非周期波动场景

场景说明

社交平台、新闻网站会因为某个突发热点事件，访问量激增，在激增过后，访问量又下降回归正常数值，此类波动通常无法提前预测，且业务访问量变化无规律可循，无法设置定时或周期策略，用户可通过以下两种方式应对此类场景：

- 当业务访问量突增或突减时，手动对业务主机实现增减。但此方法很难做到及时且完善的增减，不推荐。
- 设置告警策略，通过观测业务主机的性能指标数据（CPU 和内存的使用率等）动态变化，来确认是否增减云主机。推荐此方法。

场景示例

以下是为此场景设置告警策略的示例：

示例：可以设置 2 个告警策略，一个告警策略设置为伸缩组内实例的 CPU 使用率平均值超过 70% 时，自动增加 n 台实例；另一个告警策略设置为伸缩组内实例的 CPU 使用率平均值低于 30% 时，自动减少 n 台实例。

产品优势

- 使用弹性伸缩设置告警策略，可以自动调整云主机实例资源，精准扩缩容，避免资

源的浪费与冗余。

- 可在确保计算能力满足需求的前提下最大程度的节省资源与人力成本。

场景二：周期波动场景

场景说明

直播平台或者在线教育网站每天 20:00-22:00 是业务高峰期,且是周期性规律的业务形态。

20:00 以后业务访问量开始突增, 22:00 以后, 业务访问量突减。

场景示例

此类场景可以通过周期策略来处理, 设置 2 个周期策略, 1 个周期策略为每天 19:55 自动增加 n 台实例, 另 1 个周期策略为每天 22:05 自动减少 n 台实例。这种方式可以很好的应对业务高峰期的业务, 并且高峰后自动释放实例, 避免资源浪费的情况。

产品优势

- 可自动调整业务主机实例数量, 确保计算能力始终跟随业务需求变化, 实现业务稳定运行。
- 可最大程度的节约资源与人力成本, 避免资源冗余浪费, 减少不必要的投入, 将精力更多投入在业务创新中。

场景三：高可用场景

场景说明

大多数企业应用都有高可用性的需求, 需要业务全天全时段健康稳定的运行, 当业务系统中出现云主机的故障, 也有可能导导致业务系统的卡顿中断, 此时需要使用弹性伸缩来替换不健康的云主机确保业务的稳定运行。

场景示例

- 此类场景可通过开启期望实例数检查来处理，伸缩组设置开启实例健康检查并设置检查时间间隔为 5 分钟。当伸缩组内某台云主机在发生故障宕机时，弹性伸缩可以通过健康检查将故障主机移出伸缩组并且迅速拉起一台新的健康主机实例。
- 在多 AZ 资源池可在伸缩组可用区配置中设置多个 AZ 进行扩容。即当某一个 AZ 主机不可用时，也可以在其他可用区迅速拉起主机实例。通过对同资源池内跨多个 AZ 进行云资源的伸缩，极大提升资源利用率，满足业务部署时对多可用区的容灾的严格要求，确保系统的高可用性。

产品优势

- 使用弹性伸缩可以更低成本的实现高可用性。
- 使用弹性伸缩产品能够自动的检测当前业务系统中存在的“异常”云主机，并联动伸缩组、负载均衡，自动替换健康的弹性云主机来承载业务运行，确保业务稳定运行。

使用限制

功能限制

- 伸缩组内云主机实例会可能根据伸缩组策略、健康检查等机制被自动释放，因此不适合用于存放业务数据、日志等信息。
- 添加至伸缩组内云主机部署的服务应是无状态且支持横向扩展。
- 若伸缩组关联负载均衡的后端云主机组被删除，则伸缩组将与该负载均衡自动解绑。

配额限制

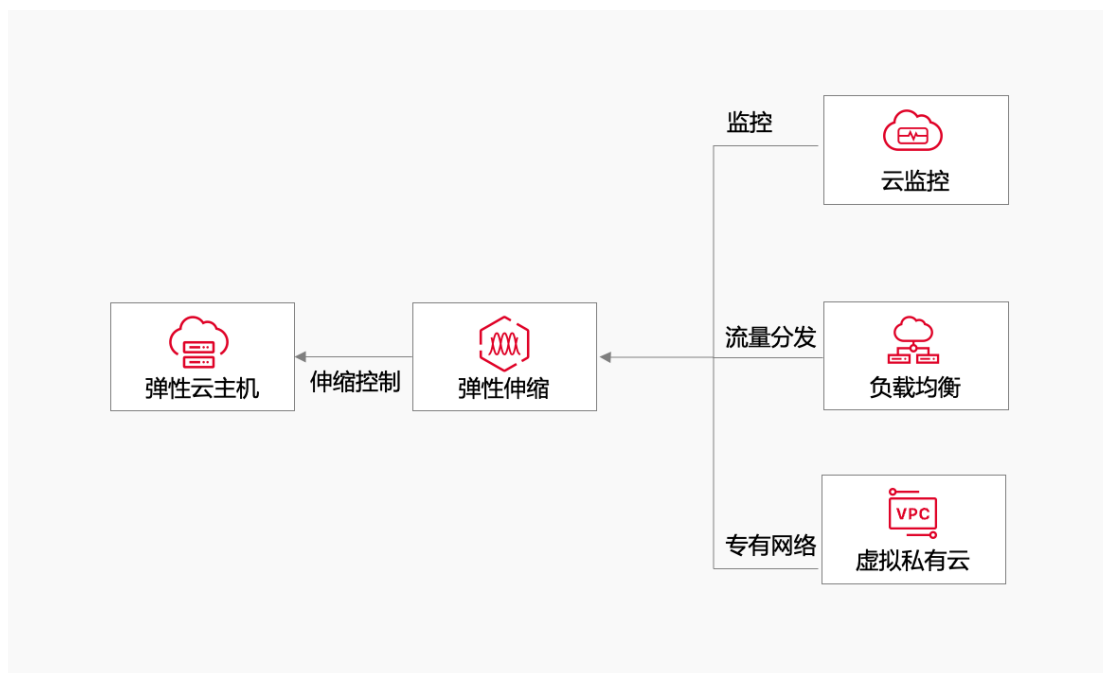
每个用户可在单个资源池内使用弹性伸缩的限制如下表所示：

资源	默认配额值	如何升级配额
单个资源池内的伸缩组总数	10	提交工单提升配额
单个资源池内伸缩配置总数	100	提交工单提升配额
单个伸缩组内的伸缩策略总数	10	提交工单提升配额
单个伸缩组内可关联的最大实例数	50	提交工单提升配额
单个伸缩组可关联的负载均衡数	10	暂不支持提升
单个伸缩组可关联的伸缩配置总数	1	暂不支持提升

与其他服务的关系

弹性伸缩依赖弹性云主机、云监控、负载均衡、虚拟私有云实现自动扩缩容功能，与其他服

务的关系如图所示：



弹性伸缩与其他服务的联系详情如下表：

相关服务	功能交互	相关链接
弹性云主机 (CT-ECS , Elastic Cloud Server)	弹性云主机是弹性伸缩中伸缩活动的主要对象，用户需要为伸缩组增加或减少弹性云主机来确保业务的稳定运行。	伸缩配置管理
弹性负载均衡 (CT-ELB , Elastic Load Balancing)	弹性伸缩的可联动负载均衡，当伸缩组中增加或减少弹性云主机时，会自动将其纳入负载均衡监听器的监听范围，并按照预设分发算法为其分发访问请求，扩展应用系统对外的服务能力，实现更高水平的应用程序容错性能。	添加负载均衡器
虚拟私有云 (CT-VPC , Virtual Private Cloud)	为不同的伸缩组提供逻辑隔离，构建一个私密、安全的虚拟私有网络环境，用户可以为每一个伸缩组进行网络配置，包括子网、安全组等。	创建伸缩组
云监控	当用户通过弹性伸缩配置了告警策略，云监	伸缩组和实例的监控

相关服务	功能交互	相关链接
	控会成为伸缩控制工具，监控预设的性能指标是否到达阈值，若达到则触发伸缩活动，若未达到则保持原样。	

产品地域和可用区

地域

地域 (Region) 是指物理的数据中心的地理区域。地域从地理位置和网络时延维度划分，同一个 Region 内共享弹性计算、块存储、对象存储、VPC 网络、弹性公网 IP、镜像等公共服务。

天翼云不同地域之间完全隔离，保证不同地域间最大程度的稳定性和容错性。为了降低访问时延、提高下载速度，建议您选择最靠近业务需求的地域。

相关特性

不同地域之间的网络完全隔离，不同地域的云产品默认不能通过内网通信。

如果不同地域之间的云产品之间有通信需求，可以通过公网 IP、VPN 等方式进行通信。

如何选择地域

在天翼云中，资源创建或购买成功后不能更换地域，因此选择地域时，您需要慎重考虑以下几个因素：

1. 地理位置：用户和资源部署地域的距离越近，网络时延越低，访问速度越快。建议您基于业务场景对时延的要求选择地域。

- a) 中国内地：一般情况下建议选择与您目标用户所在地域最为接近的数据中心，可以进一步提升用户访问速度。如果您的业务在中国内地，电信网络可以保证您在中国内地不同地域部署的业务间快速访问。
- b) 其他国家及地区：其他国家及地区提供的带宽主要面向非中国内地的用户。如果您的业务在中国内地，使用这些地域部署业务会有较长的访问延迟，不建议您使用。
2. 资源价格及资源覆盖：不同区域的资源价格可能有差异，不同地域的产品覆盖可能有差异，请根据您的需求及预算选择合适的地域。
 3. 经营性备案：如果您使用弹性云主机实例作为 Web 服务器，您需要完成经营性备案，同时需要在指定的地域购买实例。（各省（或市）通信管理局对经营性备案的审批要求不同，请以当地管理局经营性备案网站公示内容为准。）
 - 4.

可用区

可用区 (AZ, Availability Zone) 是指在同一地域内，电力和网络互相独立的物理区域。一个 AZ 是一个或多个物理数据中心的集合，具备独立的风火水电，可用区之间距离 100KM 以内，一个 Region 中的多个 AZ 间通过高速光纤相连，以满足用户跨 AZ 构建高可用性系统的需求。

相关特性

- 划分可用区的目标是能够保证可用区间故障相互隔离（大型灾害或者大型电力故障除外），不出现故障扩散，使得用户的业务持续在线服务。通过启动独立可用区内的实例，用户可以保护应用程序不受单一位置故障的影响。
- 处于相同地域不同可用区，但在同一个虚拟私有云下的云产品之间均通过内网互通，



可以直接使用内网 IP 访问。



计费说明

计费模式

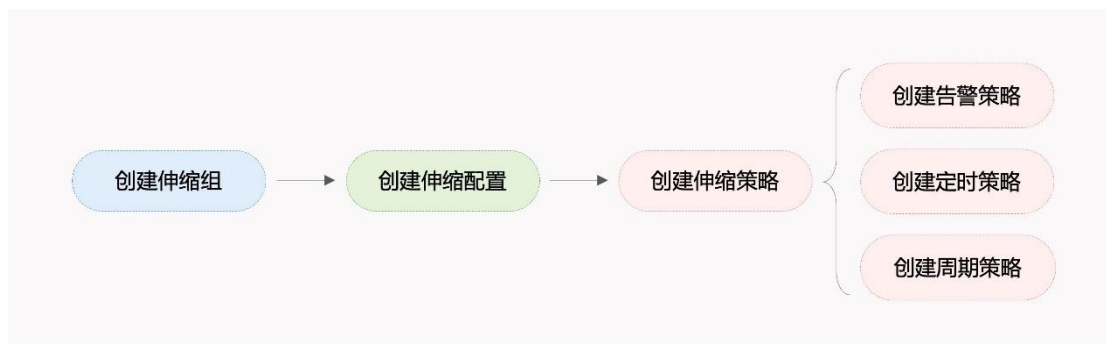
弹性伸缩服务产品免费，不收取任何费用。但弹性伸缩组内的云主机、云硬盘、弹性 IP、带宽等云产品需按照相关产品页面公示信息收费。

快速入门

弹性伸缩使用流程

弹性伸缩服务是根据您的业务需求，通过策略自动调整其弹性计算资源的一种管理服务。该服务能够根据预设规则自动调整伸缩组内的云主机数量，弹性伸缩服务的主要使用流程如下

图：



1. 准备工作：首先进行准备工作，注册天翼云，确保账户余额，具体流程参见[准备工作](#)。
2. 创建伸缩组：创建弹性伸缩首先需要创建伸缩组，伸缩组是具有相同属性和应用场景的云主机实例和伸缩策略的集合。创建伸缩组的基本信息，包括配置最大实例数、最小实例数和关联的负载均衡服务等，具体操作请参见[创建伸缩组](#)。
3. 创建伸缩配置：在创建完伸缩组后，可以创建伸缩配置，伸缩配置即伸缩活动中添加的云主机的规格。通常为一个伸缩组内添加云主机实例规格的配置模板，包括云主机的类型、vCPU、内存等。一个伸缩组支持创建一个伸缩配置。具体操作请参见[创建伸缩配置](#)。
4. 确认伸缩启用状态：在创建伸缩策略之前，需要确认伸缩组是否处于启用状态，只有伸缩组启用状态为“已启用”的伸缩组，系统才会监控该伸缩组的伸缩策略，才可能触发伸缩活动。伸缩组创建成功后，状态默认为“已启用”。仅当伸缩组状态

为“已停用”才可以启用伸缩组。具体操作步骤请参见[启用伸缩组](#)。

5. 创建伸缩策略：确认伸缩组为启用状态，用户可以创建伸缩策略。天翼云弹性伸缩服务支持创建三种策略，分别为告警策略、定时策略、周期策略，可以根据业务实际需求来创建对应的策略。具体操作请参见[创建伸缩策略](#)。

准备工作

注册天翼云账号

在使用弹性伸缩之前，您需仔细观看本文内容，完成准备工作。

如果您已经拥有一个天翼云实名账号可直接跳过本步骤，如果您还没有天翼云账号，请参考

以下流程进行创建：

1. 登录天翼云官网 www.ctyun.cn，点击右上角“免费注册”按钮。
2. 填写注册信息。填写邮箱后，设置登录密码，并通过手机验证。
3. 勾选协议，并点击“同意协议并提交”，即可完成账号注册。

实名认证

具体可参见[实名认证](#)模块对天翼云账号进行实名认证。

账户充值

您需要确保天翼云实名账户中的余额充足，具体天翼云账户充值步骤请参见[账户充值](#)。

快速扩缩容弹性云主机

操作背景


弹性伸缩服务在日常工作中可能面临以下几种场景：

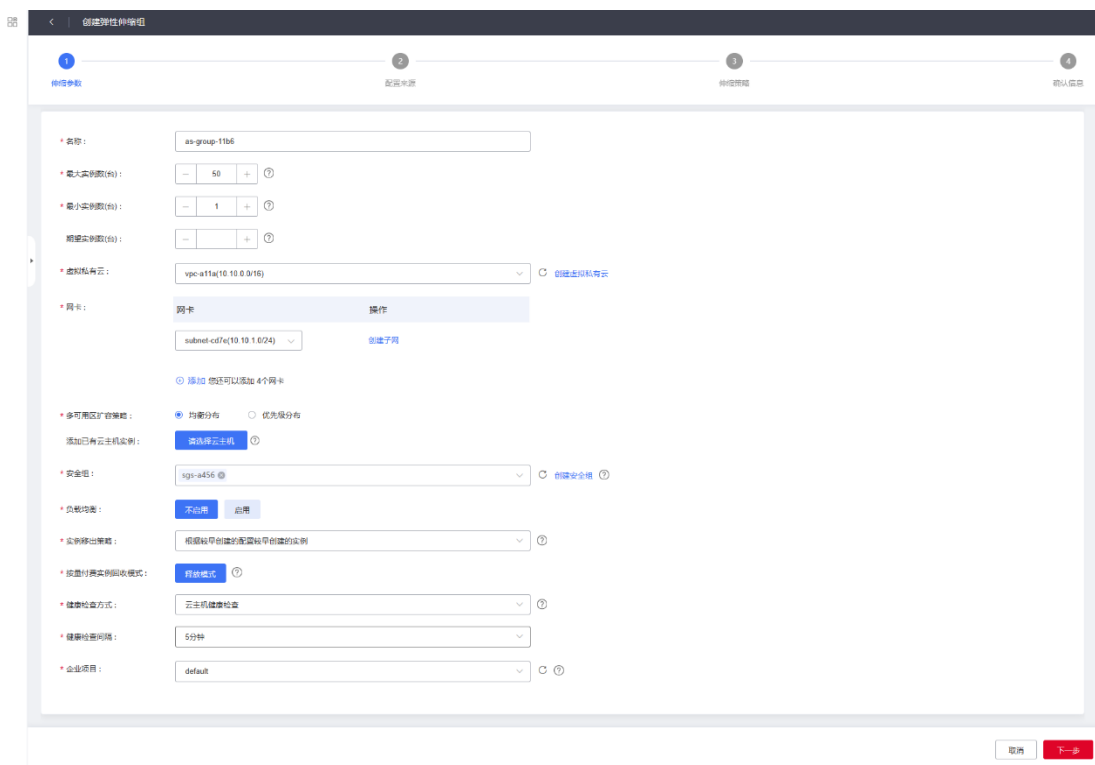
- 无明显波动的计算资源日常管理。
- 非周期波动的计算资源管理。
- 周期波动的计算资源管理。场景介绍请参见[应用场景](#)。

场景	需求说明	操作步骤
无明显波动的计算资源日常管理	创建具有相同规格与属性的云主机，放置于伸缩组内，随时监控主机运行状况。	<ol style="list-style-type: none">1. 创建伸缩组2. 创建伸缩配置（确认伸缩组为“已启用”状态）3. 绑定伸缩组与伸缩配置
非周期波动的计算资源管理	按照业务需求，为可能到来的访问波动创建告警策略。	<ol style="list-style-type: none">1. 创建伸缩组2. 创建伸缩配置（确认伸缩组为“已启用”状态）3. 创建伸缩策略（告警策略）
周期波动的计算资源管理	根据周期波动的规律预设伸缩策略，确保计算能力满足业务需求，此类业务通常设置定时/周期策略。	<ol style="list-style-type: none">1. 创建伸缩组2. 创建伸缩配置（确认伸缩组为“已启用”状态）3. 创建伸缩策略（定时/周期策略）

创建伸缩组

伸缩组是具有相同属性和应用场景的云主机实例和伸缩策略的集合。伸缩组是启停伸缩策略和进行伸缩活动的基本单位。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中，点击右上角“创建弹性伸缩组”，进入“创建弹性伸缩组>伸缩参数”配置页面。



The screenshot shows the 'Create Elastic Scaling Group' configuration page. The 'Scaling Parameters' (伸缩参数) tab is active. Key fields include:

- Name: es-group-11b6
- Maximum Instance Count: 50
- Minimum Instance Count: 1
- Desired Instance Count: (empty)
- VPC: vpc-at1a(10.10.0.0/16)
- Subnet: subnet-cd7e(10.10.1.0/24)
- Load Balancing: Enabled (启用)
- Health Check: Cloud Machine Health Check (云主机健康检查)

5. 在弹性伸缩组创建页面，填写弹性伸缩组的基本信息，例如伸缩组名称、最大实例数、最小实例数、虚拟私有云、可用区、子网、负载均衡等信息，具体参数说明见下表：

参数	是否必选	参数说明
名称	是	伸缩组的名称, 用户可自定义, 也可以选择保持系统默认分配名称。
最大实例数 (台)	是	伸缩组内包含的实例数量的最大值, 伸缩组内执行的任何伸缩活动结果不得大于最大实例数。

参数	是否必选	参数说明
		若您修改了伸缩组的最大实例数使得伸缩组的当前实例数大于最大实例数, 弹性伸缩服务会自动移出实例, 保证伸缩组的当前实例数等于伸缩最大实例数。
最小实例数 (台)	是	伸缩组内包含的实例数量的最小值, 伸缩组内执行的任何伸缩活动结果不得小于最小实例数。 若您修改了伸缩组的最小实例数使得伸缩组的当前实例数小于伸缩最小实例数时, 弹性伸缩服务会自动添加实例, 保证伸缩组的当前实例数等于伸缩最小实例数。
期望实例数 (台)	否	伸缩组内包含的实例数量的稳态值。 若您为伸缩组设置了期望实例数, 当伸缩组内实例数不等于期望实例数时, 弹性伸缩服务会自动进行扩缩容, 确保伸缩组内始终保持该数量的实例数。
虚拟私有云	是	同一伸缩组内的弹性云主机的虚拟私有云是一致的。
网卡	是	一个伸缩组可以绑定多张网卡, 默认第一张网卡为主网卡。
多可用区扩 容策略	是	多可用区资源池伸缩组需指定扩容策略。 - 均衡策略: 在伸缩组绑定的伸缩配置关联多个可用区时生效, 支持在多个可用区之间均衡分布云主机实例。 - 优先级策略: 伸缩组绑定的伸缩配置先选择的可用区优先级高。弹性伸缩优先在优先级最高的可用区尝试伸缩活动。若无法扩进行伸缩活动, 则自动在优先级次之的可用区进行。
添加已有云 主机实例	否	伸缩组创建完成后, 支持将符合条件的已有云主机实例移入伸缩组, 作为初始化实例, 最大可选择添加 10 个实例。添加已有的云主机实例数量受到最小、最大、期望实例数限制。
安全组	是	安全组类似防火墙功能, 是一个逻辑上的分组, 用于设置网络访问控制。
负载均衡	否	选择启用负载均衡后, 伸缩组中的云主机会自动挂载到您

参数	是否必选	参数说明
		<p>关联的负载均衡下。当选择使用弹性负载均衡时，需配置 4 个参数：负载均衡器、后端云主机组、后端端口和权重。访问流量将按照分发策略自动分发到伸缩组内的所有弹性云主机。</p> <p>注意：一个伸缩组可最多添加 10 组负载均衡监听器。</p>
实例移除策略	是	<p>在对伸缩组进行缩容时，会采用此策略进行移除，共有以下四种方式：</p> <ul style="list-style-type: none"> - 较早创建的配置较早创建的实例：先筛选出较早创建的配置所创建出的实例，再筛选较早创建的实例移出。 - 较晚创建的配置较晚创建的实例：先筛选出较晚创建的配置所创建出的实例，再筛选较晚创建的实例移出。 - 较早创建的实例：根据时间筛选较早创建的实例，跟是否是伸缩配置创建的无关。 - 较晚创建的实例：根据时间筛选较晚创建的实例，跟是否是伸缩配置创建的无关。
实例回收模式	是	<p>指伸缩组自动创建的实例被移出后的处理策略。</p> <p>释放模式：将移出的伸缩组自动创建的云主机释放。</p> <p>注意：对于手动移入伸缩组的云主机，如果被移出伸缩组，资源不会被释放，状态保持不变，不适用此策略。</p>
健康检查方式	是	<p>健康检查会将处于“异常”状态的云主机从伸缩组中移出，并替换同种规格的云主机来承载业务流量。有以下两种方式：</p> <ul style="list-style-type: none"> - 云主机健康检查：是指对云主机的运行状态进行检查，如关机、错误等人为或非人为状态都是云主机异常状态。 - 弹性负载均衡健康检查：是指根据负载均衡对云主机的健康检查结果进行的检查。所有监听器下检测到的云主机状态必须均为正常。 <p>注意：只有开启负载均衡时您才可以选择弹性负载均衡健康检查，建议当伸缩组开启负载均衡时，使用弹性负载均衡健康检查。</p>
健康检查间隔	是	<p>执行健康检查的周期时间。您可以根据业务实际情况设置合理的健康检查间隔（5 分钟、15 分钟、1 小时、3 小</p>

参数	是否必选	参数说明
		时), 以确保其包含主机中应用程序的预期启动时间。
企业项目	是	支持为伸缩组选择企业项目。

6. 参数配置完成之后, 请单击右下角的“下一步”, 进入“创建弹性伸缩组 > 配置来源”页面, 进行伸缩配置的创建与选择。

注意:

创建一个伸缩组时, 必须同时创建或选择一个伸缩配置。具体操作请参见[创建伸缩配置步骤](#)。

创建伸缩配置

伸缩配置即伸缩活动中添加的云主机的规格。通常为一个伸缩组内添加云主机实例规格的配置模板。

创建伸缩配置有两个进入窗口:

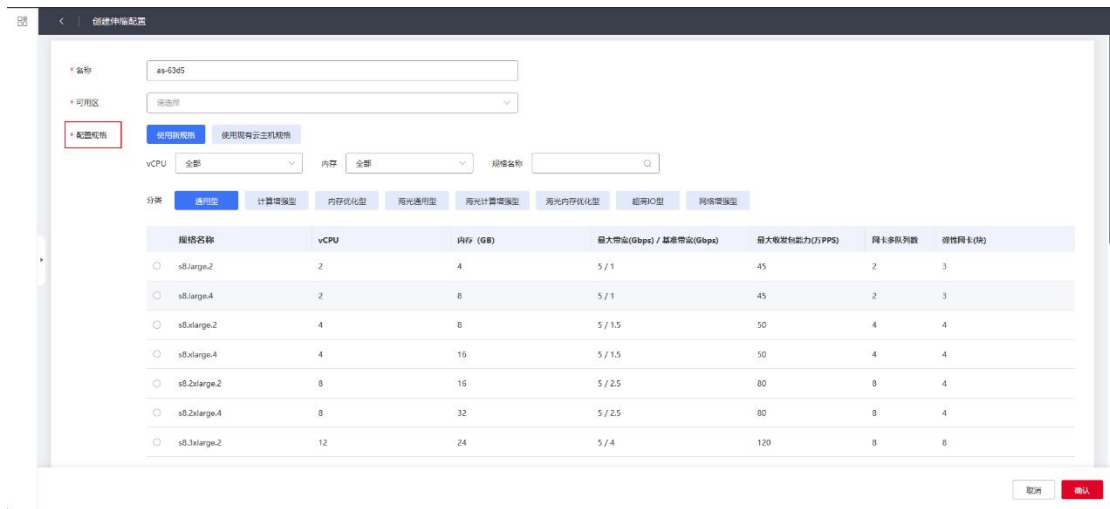
- 在弹性伸缩管理控制台, 进入“伸缩配置”页签, 点击右上角“创建伸缩配置”按钮进行创建。具体步骤请参见[“用户指南>伸缩配置管理>创建伸缩配置”](#)。
- 创建伸缩组时, 在“配置来源”页面, 即可创建新的伸缩配置。本文主要讲述此种方式。

注意:

以下操作均在[创建伸缩组](#)的步骤基础上完成。

1. 进入“创建弹性伸缩组 > 配置来源”页面, 进行伸缩配置的创建与选择。
2. 在“伸缩配置来源”选项, 您可以选择使用已有的伸缩配置或者即时创建新的伸缩

配置。此处我们选择“新建配置”，进入到伸缩配置新建页面。



3. 在伸缩配置新建页面需填写伸缩配置信息，例如配置名称、云主机镜像、云主机类型、登录方式等，具体参数说明如下表：

参数	是否必选	参数说明
名称	是	创建伸缩配置的名称。
可用区	是	多可用区资源池可选择伸缩组绑定的伸缩配置，在指定可用区扩缩容。当伸缩配置绑定的伸缩组使用优先级扩容策略时，可用区选择顺序将决定优先级。
配置规格	是	配置规格有两种选择，使用新规格与使用现有云主机规格： 选择“使用现有云主机规格 > 请选择云主机” 创建配置，云主机类型、vCPU、内存、镜像、云硬盘数据、安全组信息将默认与选择的云主机规格保持一致。 选择“使用新规格”，接下来根据实际业务需求配置主机类型、vCPU、内存、镜像、云硬盘参数、安全组信息。
弹性 IP	是	弹性 IP 可以实现虚拟私有云中的云资源通过固定的公网 IP 地址与互联网互通。您可以根据实际需求选择以下两种方式： 不使用：弹性云主机不能对外提供访问服务，仅可在虚拟私有云内部进行内网互通。 自动分配：自动为每台弹性云主机分配独享带宽的弹性 IP，带宽值可以由您设定。

参数	是否必选	参数说明
登录方式	是	天翼云提供两种登录方式供您选择，密钥对和密码。 密钥对指使用密钥作为弹性云主机的鉴权方式进行登录。选择此方式，请您在密钥对配置项中导入密钥对。说明：如果您直接从下拉列表中选择已有的密钥，请确保您已在本地获取该文件且可以登录云主机，否则，将影响您正常登录弹性云主机。 账户密码指使用设置 root 用户（Linux 操作系统）和 Administrator 用户（Windows 操作系统）的初始密码方式作为弹性云主机的鉴权方式，如果选择此方式，您可以通过用户名密码方式登录弹性云主机。您需要为您的用户设置密码，并确认密码，确保能够成功登录云主机。

4. 配置完成之后，用户可以选择单击右下角的“下一步”进入到“创建弹性伸缩组 > 伸缩策略”页面，继续进行伸缩策略的创建，也可以选择单击“立即创建”按钮，完成伸缩组与伸缩配置的创建，此处我们选择单击“下一步”继续伸缩策略的创建。

创建伸缩策略

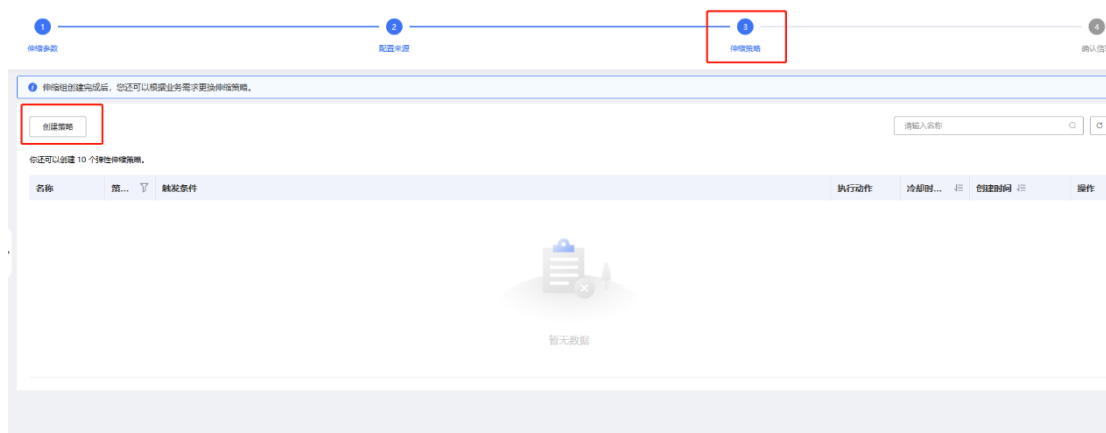
伸缩策略：伸缩策略用于定义伸缩活动的触发模式、触发条件和触发动作。伸缩策略通常有三种，告警策略、周期策略、定时策略。

创建伸缩策略有两个进入窗口：

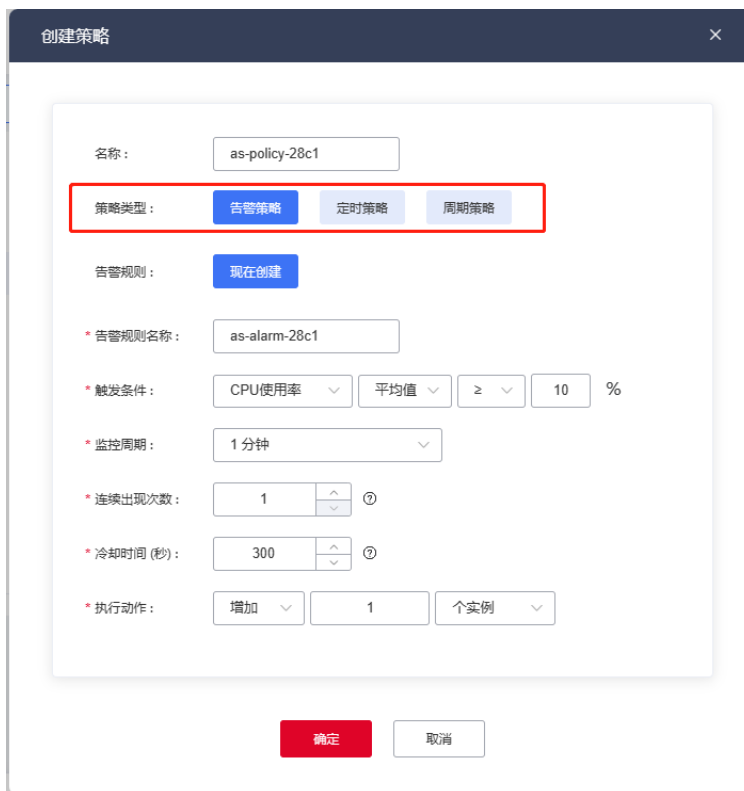
- 在待创建伸缩策略的伸缩组详情页面，点击“创建策略”，进入到伸缩策略创建页面。具体步骤请参见“[用户指南 > 伸缩策略管理 > 创建伸缩策略](#)”。
- 创建伸缩组时，完成伸缩组、伸缩配置的创建之后，点击“下一步”进入“伸缩策略”页面，即时创建新的伸缩策略。本文主要讲述此种方式。

注意：以下操作均在[创建伸缩配置](#)的步骤基础上完成。

1. 进入“创建弹性伸缩组>伸缩策略”页面，进行伸缩配置的创建与选择。
2. 点击“创建策略”按钮，在弹出的“创建策略”窗口中进行策略配置。



3. 在“创建策略”窗口中，用户可以选择三种不同的策略类型，分别为告警策略、定时策略、周期策略。



4. 三种策略拥有不同的配置参数，具体的参数说明如以下三个表：

- 告警策略参数说明：

参数	是否必选	参数说明
名称	是	创建伸缩策略的名称。
策略类型	是	此处选择“告警策略”。
告警规则	是	保持默认配置“现在创建”。当您有已经创建好的告警规则时，可以选择“使用已有”，来直接获取已创建好的告警规则模板。
告警规则名称	是	创建告警规则名称。
触发条件	是	用户可以选择弹性伸缩所支持的性能指标，并为此性能指标配置触发条件。用户可以选择的性能指标有：CPU 使用率、内存使用率、网络流入速度、网络流出速度、磁盘读速率、磁盘写速率、磁盘读请求速率、磁盘写请求速率。
监控周期	是	告警规则监控状态刷新的间隔时间。
连续出现次数	是	连续出现次数指探测结果连续几次符合您设置的规则，才会触发告警。例如设置为 3，则表示连续 3 次超过阈值后才会触发告警。
冷却时间	是	冷却时间是指用户在每次成功的伸缩活动后设置的一段锁定时间，当一次伸缩活动执行完成，系统开始计算冷却时间。在冷却时间内，由告警策略触发的伸缩活动将会被拒绝执行，其他类型的伸缩活动不受限制，但执行完成之后将会重新开始冷却时间计时，以避免告警策略的频繁触发。
执行动作	是	指当满足设置的告警策略进行伸缩活动时，弹性伸缩服务将执行的动作。用户可在此设置以下三种动作： 增加 n 台云主机实例。 减少 n 台云主机实例。 设置为 n 台云主机实例。

- 定时策略参数说明:

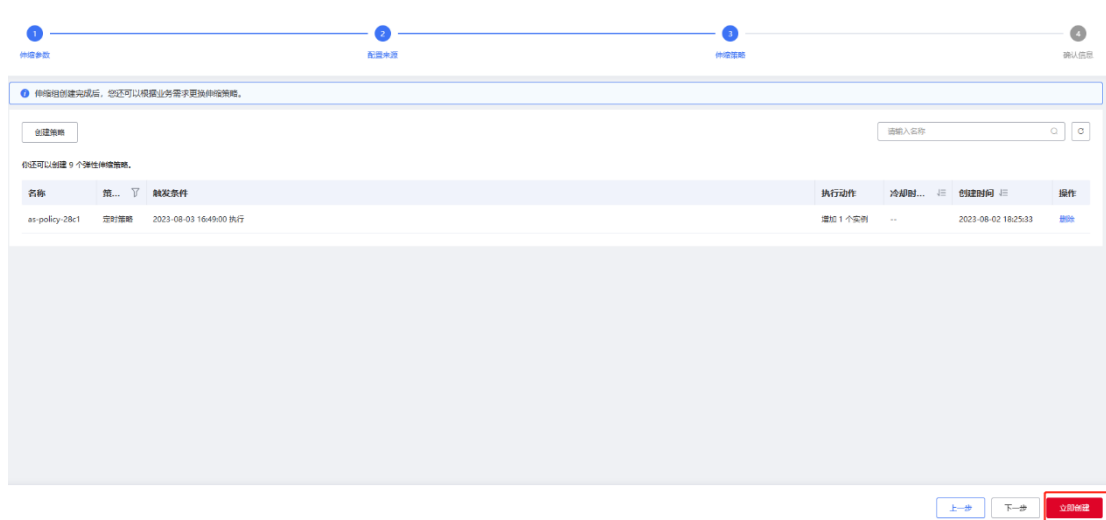
参数	是否必选	参数说明
名称	是	创建伸缩策略的名称。
策略类型	是	此处选择“定时策略”。
时区	否	为默认值: GMT+08:00, 代表北京时间。
触发时间	是	设置伸缩活动执行时间。注意: 请确保触发时间晚于默认时间, 并且晚于当前时间。
执行动作	是	指当满足设置的定时策略进行伸缩活动时, 弹性伸缩服务将执行的动作。用户可在此设置以下三种动作: 增加 n 台云主机实例。 减少 n 台云主机实例。 设置为 n 台云主机实例。

- 周期策略参数说明:

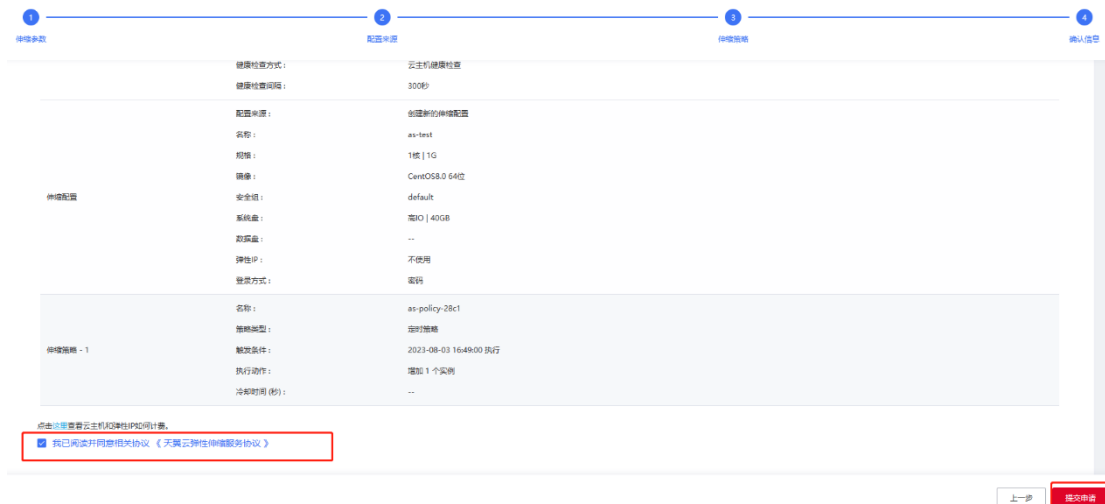
参数	是否必选	参数说明
名称	是	创建伸缩策略的名称。
策略类型	是	此处选择“周期策略”。
周期	是	执行伸缩活动的重复周期, 用户可在按天、按周、按月中选择。
时区	否	为默认值: GMT+08:00, 代表北京时间。
触发时间	是	设置伸缩活动执行时间。注意: 触发时间须在策略的生效时间段内。

参数	是否必选	参数说明
生效时间	是	伸缩策略可被触发的有效时间段。
执行动作	是	指当满足设置的周期策略进行伸缩活动时，弹性伸缩服务将执行的动作。用户可在此设置以下三种动作： 增加 n 台云主机实例。 减少 n 台云主机实例。 设置为 n 台云主机实例。

5. 用户根据上述参数说明完成伸缩策略配置，点击“创建策略”弹窗中的“确定”，回到“伸缩策略”页面，点击右下角“立即创建”，跳转至“确认信息”页面。



6. 在“确认信息”页面，确认创建的弹性伸缩组、伸缩配置、伸缩策略信息，确认无误后勾选“我已阅读并同意相关协议《天翼云弹性伸缩服务协议》”。点击右下角“提交申请”，即可完成创建。



至此，快速扩缩容弹性云主机的主要使用流程已全部操作完成，用户可回到弹性伸缩管理控制台查看详细信息。


动态扩展云主机

可以使用弹性伸缩创建低成本业务集群，根据业务流量自动扩缩容业务集群内的计算资源，减少成本浪费。方案详解如下：

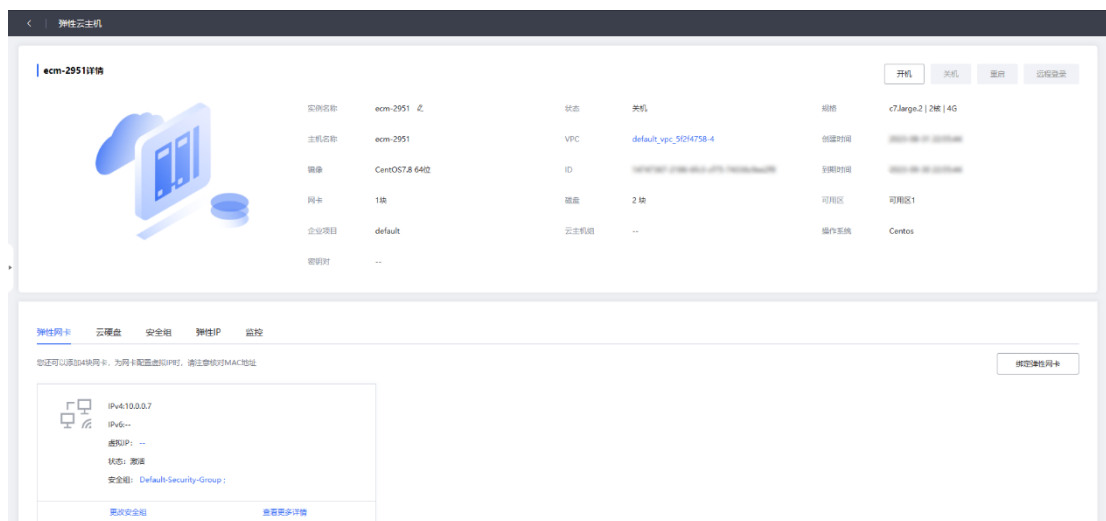
为伸缩组创建告警策略来应对突发的流量变化（突增或突减）。

步骤 1：创建包年包月弹性云主机实例

为本实践的业务集群创建能够满足日常业务需求的包年包月低成本云主机实例。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性云主机”，进入弹性云主机管理控制台。
4. 在右上角单击“创建云主机”按钮，进入到弹性云主机创建页面。
5. 在弹性云主机创建页面，完成云主机参数配置，具体参数说明可参见[创建云主机](#)。

创建成功之后如图：




注意：

付费模式请设置“包年包月”。

步骤 2：创建伸缩组并添加云主机实例

创建伸缩组，并将创建好的云主机实例添加至此伸缩组内，来满足动态扩展云服务器的需求。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择地域，此处我们选择华东 1。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中，点击右上角“创建弹性伸缩组”，进入“创建弹性伸缩组>伸缩参数”配置页面。
5. 在配置页面完成伸缩组的配置，具体参数配置可参见[创建伸缩组](#)。
6. 在完成伸缩组与伸缩配置的设置后，点击“立即创建”，即可创建伸缩组成功，创建成功如图所示：




7. 伸缩组创建成功之后，进入到伸缩组详情页面，将步骤一中创建包年包月弹性云主机添加至伸缩组内，用于日常业务需求的满足。具体操作步骤请参见将实例移入伸缩组。添加成功之后如图：

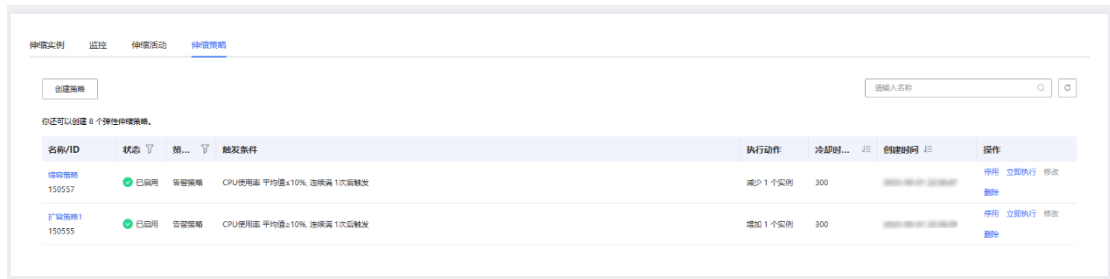


步骤 3：根据业务需求创建伸缩策略

为已创建好的伸缩组设置伸缩策略，来满足突发的业务流量变化，确保实例数可动态变化满足业务稳定运行。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择地域，此处我们选择华东 1。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击步骤 2 创建好的伸缩组名称，进入到伸缩组详情页面。

5. 在详情页面的下方,单击“伸缩策略”进入伸缩策略页签,单击“创建策略”按钮,进入到“创建伸缩策略”页面。
6. 在“创建伸缩策略”页面中完成策略的配置,在本实践中需要添加两个告警策略,一个扩容策略,一个缩容策略。具体操作步骤可参见[创建伸缩策略](#)。



7. 继续创建缩容策略。


定时扩展云主机

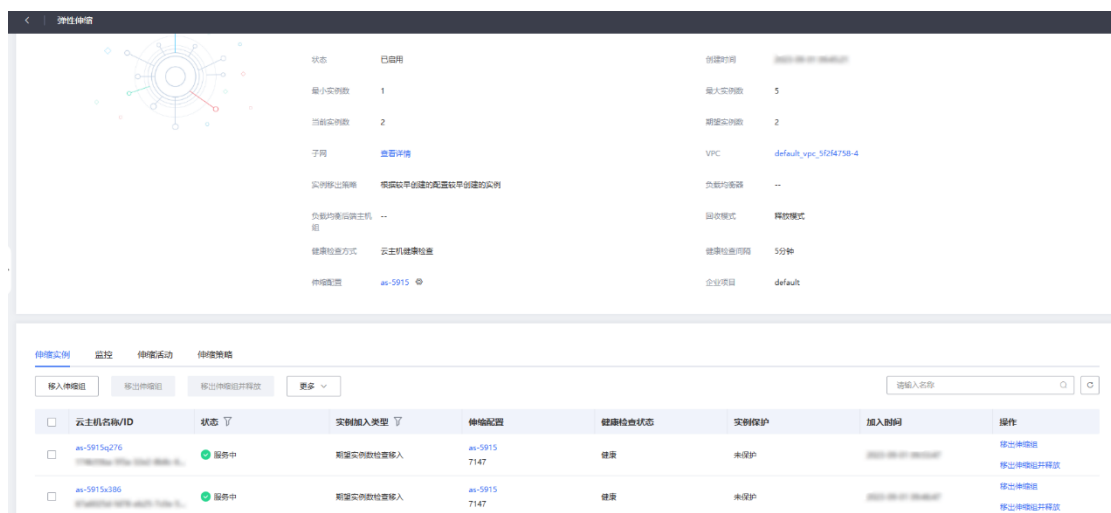
可以使用弹性伸缩创建低成本业务集群, 根据业务流量的波动周期来自动扩缩容业务集群内的计算资源, 减少成本浪费。方案详解如下:

- 针对日常业务流量, 创建伸缩组并设置最小实例数量已满足基础业务需求;
- 为伸缩组创建定时/周期策略来应对周期性的流量波动。

创建伸缩组

为周期性规律变化的业务集群创建伸缩组, 设置最小实例数为 1, 期望实例数为 2, 最大实例数为 5 来满足业务的日常访问需求。

1. 登录控制中心。
2. 单击控制中心左上角的  , 选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”, 进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中, 点击右上角“创建弹性伸缩组”, 进入“创建弹性伸缩组>伸缩参数”配置页面。
5. 在配置页面完成伸缩组的配置, 具体参数配置可参见[创建伸缩组](#)。



The screenshot displays the Elastic Scaling console interface. The top section shows the configuration for an elastic scaling group, including its status (Enabled), creation time, and various parameters like minimum, current, and maximum instances. Below this, there are tabs for 'Elastic Instances', 'Monitoring', 'Scaling Activities', and 'Scaling Policies'. The 'Elastic Instances' tab is active, showing a table of instances with columns for instance name, status, scaling type, configuration, health status, protection, and creation time.


云主机名称/ID	状态	实例加入类型	伸缩配置	健康检查状态	实例保护	加入时间	操作
ew-5915q276	服务中	期望实例数批量加入	ew-5915-7147	健康	未保护	2023-08-01 10:00:00	移出伸缩组 移出伸缩组并释放
ew-5915x386	服务中	期望实例数批量加入	ew-5915-7147	健康	未保护	2023-08-01 10:00:00	移出伸缩组 移出伸缩组并释放

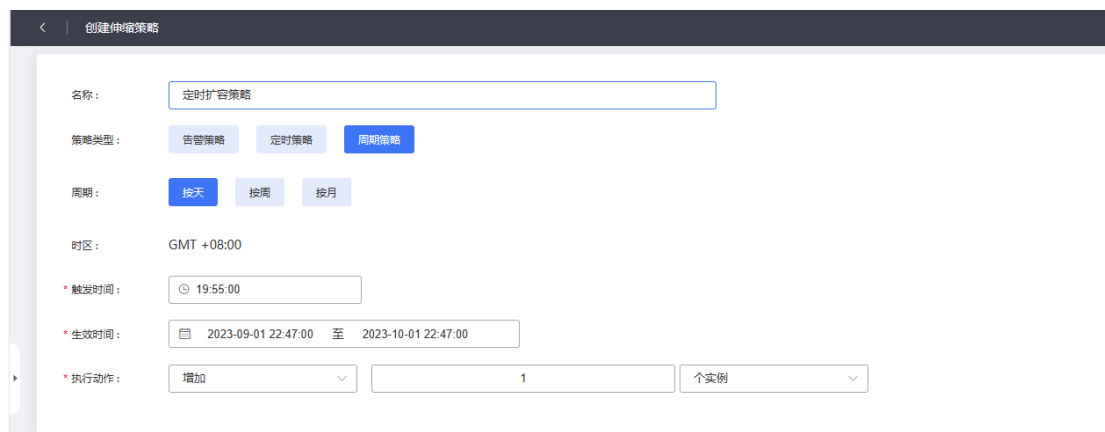
注意：

通过弹性伸缩组创建的实例仅支持按需计费。

根据业务需求创建伸缩策略

为已创建好的伸缩组设置伸缩策略，来满足突发的业务流量变化，确保业务的稳定运行。

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择地域，此处我们选择华东 1。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击通过“创建伸缩组”操作中创建好的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击“创建策略”按钮，进入到“创建伸缩策略”页面。
6. 在“创建伸缩策略”页面中完成策略的配置，在本实践中需要添加两个周期策略，一个扩容策略，一个缩容策略。具体操作步骤可参见[创建伸缩策略](#)。
 - 1) 策略类型选择周期策略。
 - 2) 时区保持默认 GMT +8:00。
 - 3) 触发时间设置为业务高峰期到来前 5 分钟，例如每天晚上 20:00 是高峰期起始时间，则周期设置为按天，9 月 1 日至 9 月 30 日期间，触发时间设置为每晚 19:55。
 - 4) 执行动作设置为增加 1 个实例。
7. 完成以上配置之后单击“确认”，即可成功创建定时扩容策略。



创建伸缩策略

名称: 定时扩容策略

策略类型: 告警策略 定时策略 周期策略

周期: 按天 按周 按月

时区: GMT +08:00

* 触发时间: 19:55:00

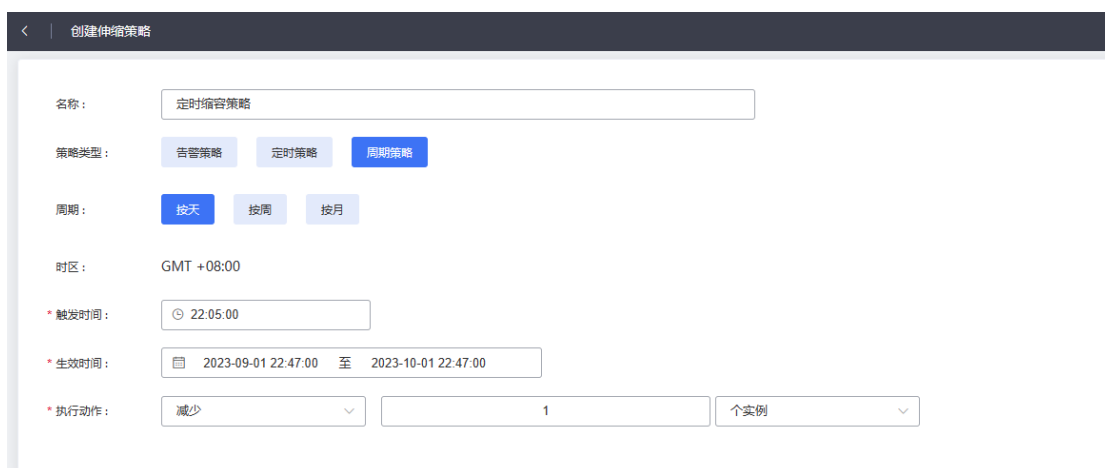
* 生效时间: 2023-09-01 22:47:00 至 2023-10-01 22:47:00

* 执行动作: 增加 1 个实例

8. 继续创建缩容策略，以下为本实践中缩容策略专有的属性值设置说明：

- 1) 策略类型选择周期策略。
- 2) 时区保持默认 GMT +8:00。
- 3) 触发时间设置为业务高峰期结束后 5 分钟，例如每天晚上 22:00 是高峰期结束时间，则周期设置为按天，8 月 1 日至 9 月 30 日期间，触发时间设置为每晚 22:05。
- 4) 执行动作设置为减少 1 个实例。

9. 完成以上配置之后单击“确认”，即可成功创建缩容策略。



创建伸缩策略

名称: 定时缩容策略

策略类型: 告警策略 定时策略 周期策略

周期: 按天 按周 按月

时区: GMT +08:00

* 触发时间: 22:05:00

* 生效时间: 2023-09-01 22:47:00 至 2023-10-01 22:47:00

* 执行动作: 减少 1 个实例

用户指南

功能总览

伸缩组

伸缩组是具有相同属性和应用场景的云主机实例和伸缩策略的集合。伸缩组是启停伸缩策略和进行伸缩活动的基本单位。后续对弹性伸缩的所有操作都是建立在伸缩组的基础之上。

功能	控制台文档	API 文档
创建伸缩组	创建伸缩组	CreateScalingGroup
启用伸缩组	启用伸缩组	EnableScalingGroup
修改伸缩组	修改伸缩组	UpdateScalingGroup
为伸缩组添加负载均衡器	添加负载均衡器	AattachLoadBalancers

伸缩配置

伸缩配置即伸缩活动中添加的云主机的规格。通常为一个伸缩组内添加云主机实例规格的配置模板，包括云主机的类型、vCPU、内存等。自动扩容时，伸缩组根据伸缩配置来创建弹性云主机实例，并将创建的云主机实例添加到伸缩组。

功能	控制台文档	API 文档
创建伸缩配置	创建伸缩配置	CreateScalingConfig
复制伸缩配置	复制伸缩配置	--
修改伸缩配置	修改伸缩配置	UpdateScalingConfig

伸缩策略

伸缩策略用于定义伸缩活动的触发模式、触发条件和触发动作。伸缩策略通常有三种：告警策略、定时策略、周期策略。

- 告警策略：通过对伸缩组内实例性能指标（CPU 使用率、内存使用率等）的监控，来确认其是否到达预设的告警条件，来自动增加或减少云主机的数量。
- 定时策略：根据业务实际情况设置一个时间点，在此时间点自动增加或减少云主机的数量。
- 周期策略：根据业务实际情况设置一段时间段，在此时间段内按照周期（按天、按周、按月）来重复执行自动增减云主机的数量。

功能	控制台文档	API 文档
创建伸缩策略	创建伸缩策略	CreateScalingRule
启用伸缩策略	启用伸缩策略	EnableScalingRule
修改伸缩策略	修改伸缩策略	UpdateScalingRule
立即执行伸缩策略	执行伸缩策略	ExecuteScalingRule

伸缩活动

当执行伸缩策略、修改伸缩组参数、健康检查等操作都会触发伸缩动作，这些动作的记录即为伸缩活动。例如增加云主机资源或减少云主机资源的过程。

功能	控制台文档	API 文档
查询伸缩活动	查询伸缩活动	QueryActivityDetail
排查伸缩活动异常	排查伸缩活动异常	--

功能	控制台文档	API 文档
按计划扩展资源	按计划扩展资源	--

伸缩组实例管理

在弹性伸缩组内，可以针对组内实例做很多操作。例如手动将实例移入伸缩组或者手动将实例移出伸缩组，手动调整实例状态，将其设置为保护中状态等。

功能	控制台文档	API 文档
将实例移入伸缩组	将实例移入伸缩组	InstanceMoveinGroup
将实例移出伸缩组	将实例移出伸缩组	InstanceMoveoutGroup
实例保护	实例保护	SetInstancesProtection

伸缩组监控

将弹性云主机实例加入伸缩组后，云监控会自动为其开启监控，可以通过云监控更好地掌握伸缩组内实例运行状况。

功能	控制台文档	API 文档
监控指标说明	监控指标说明	--
查看监控指标数据	查看监控指标数据	--
弹性伸缩健康检查	弹性伸缩健康检查	--
设置监控告警规则	设置监控告警规则	CreateAlarmRule

伸缩组管理

伸缩组概述

伸缩组：伸缩组是具有相同属性和应用场景的云主机实例和伸缩策略的集合。伸缩组是启停伸缩策略和进行伸缩活动的基本单位。后续对弹性伸缩的所有操作都是建立在伸缩组的基础之上。

弹性伸缩的伸缩组状态如下表所示：

状态	状态说明
已启用	创建的弹性伸缩组处于“已启用”状态，此时可以对弹性伸缩组做任何操作。
已停用	将处于启用状态的伸缩组停用后，此时伸缩组处于停用状态，此时伸缩组无法触发伸缩活动。


创建伸缩组

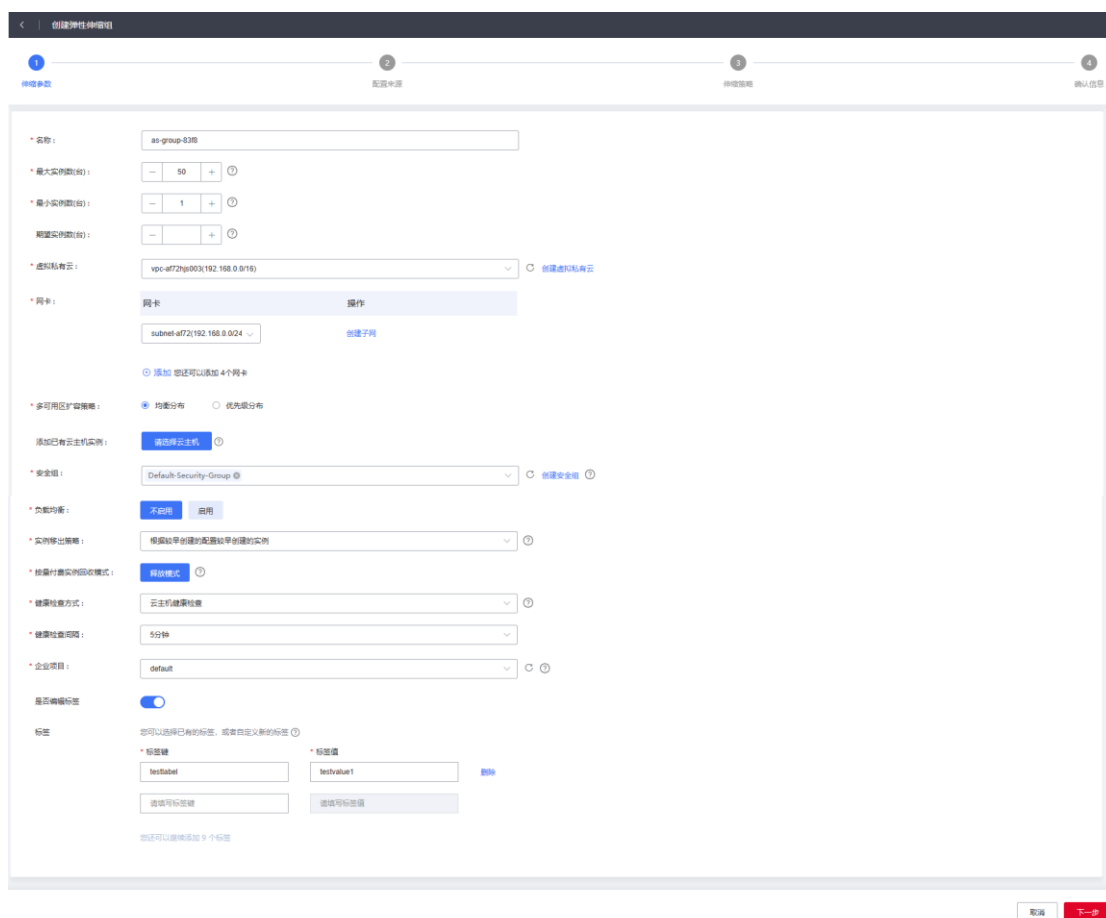
操作场景

在创建弹性伸缩组时，用户需要经历以下三个步骤：

- 配置伸缩组信息，包括最大实例数、最小实例数、负载均衡器等。
- 创建伸缩配置来源，包括实例规格、弹性 IP、登录方式等。
- 创建伸缩策略，为伸缩组创建伸缩策略，包括策略类型、告警规则等。此步骤为可选步骤，可不在创建伸缩组时进行。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中，点击右上角“创建弹性伸缩组”，进入“创建弹性伸缩组>伸缩参数”配置页面。



5. 在弹性伸缩组创建页面，填写弹性伸缩组的基本信息，例如伸缩组名称、最大实例数、最小实例数、虚拟私有云、可用区、子网、负载均衡等信息，具体参数说明见下表：

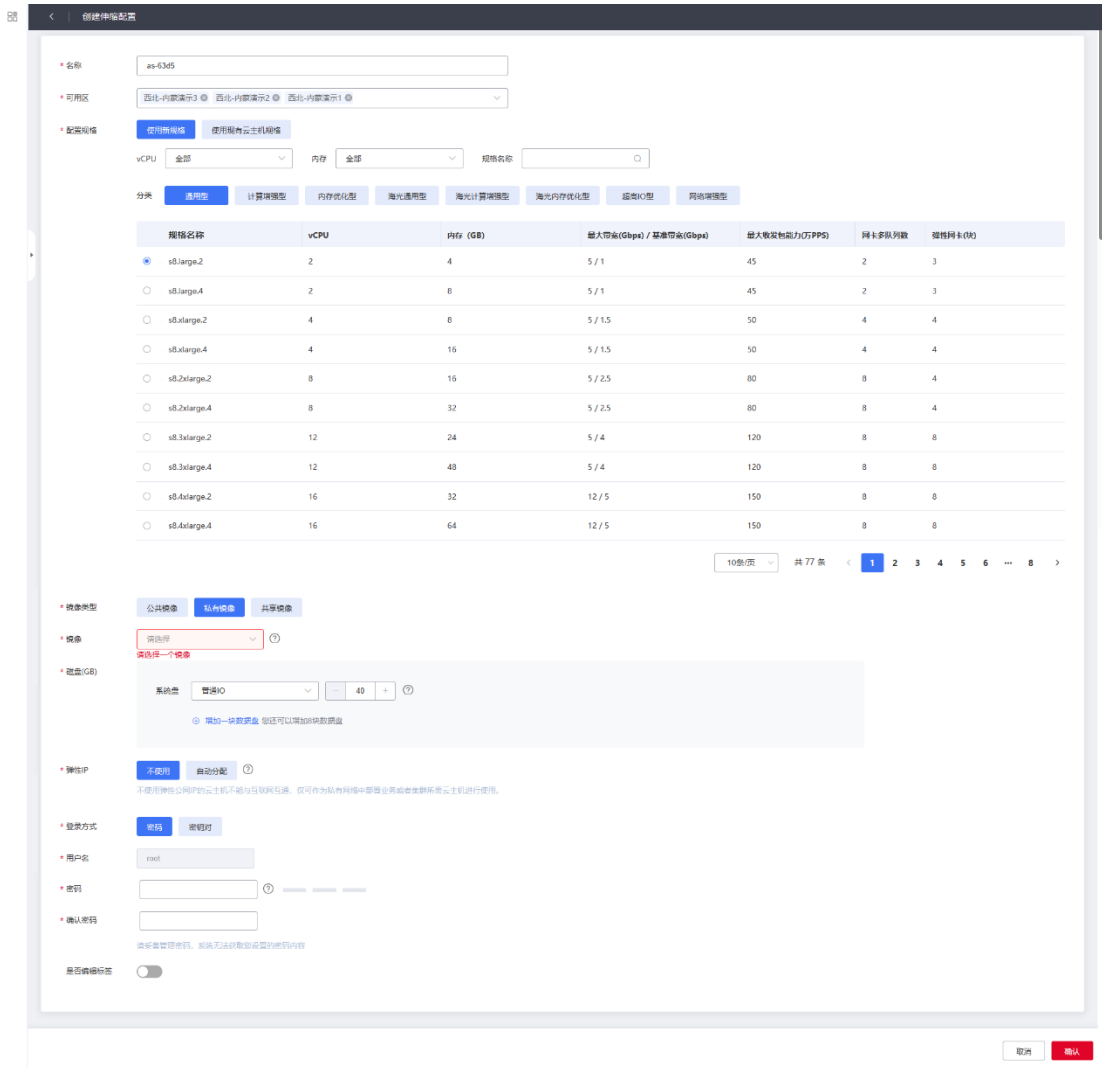
参数	是否必选	参数说明
名称	是	伸缩组的名称，用户可自定义，也可以选择保持系统默认分

参数	是否必选	参数说明
		配名称。
最大实例数 (台)	是	伸缩组中弹性云主机实例个数的最大值。当伸缩组的当前实例数大于最大实例数时，弹性伸缩服务会自动移出实例，使得伸缩组的当前实例数等于伸缩最大实例数。理论上实际实例数不允许大于最大实例数。
最小实例数 (台)	是	伸缩组中弹性云主机实例个数的最小值。当伸缩组的当前实例数小于最小实例数时，弹性伸缩服务会自动添加实例，使得伸缩组的当前实例数等于最小实例数。理论上实际实例数不允许小于最小实例数。
期望实例数 (台)	否	伸缩组期望的稳态云主机实例个数。当伸缩组内实例数不等于期望实例数时，弹性伸缩服务会自动进行扩缩容，确保伸缩组内始终保持该数量的实例数。
虚拟私有云	是	同一伸缩组内的弹性云主机的虚拟私有云是一致的。
网卡	是	一个伸缩组可以绑定多张网卡，默认第一张网卡为主网卡。
多可用区扩 容策略	是	<p>多可用区资源池伸缩组需指定扩容策略。</p> <ul style="list-style-type: none"> - 均衡策略：在伸缩组绑定的伸缩配置关联多个可用区时生效，支持在多个可用区之间均衡分布云主机实例。 - 优先级策略：伸缩组绑定的伸缩配置先选择的可用区优先级高。弹性伸缩优先在优先级最高的可用区尝试伸缩活动。若无法扩进行伸缩活动，则自动在优先级次之的可用区进行。
添加已有云 主机实例	否	伸缩组创建完成后，支持将符合条件的已有云主机实例移入伸缩组，作为初始化实例，最大可选择添加 10 个实例。添加已有的云主机实例数量受到最小、最大、期望实例数限制。
安全组	是	安全组类似防火墙功能，是一个逻辑上的分组，用于设置网络访问控制。
负载均衡	否	选择启用负载均衡后，伸缩组中的云主机会自动挂载到您关

参数	是否必选	参数说明
		联的负载均衡下。访问流量将按照分发策略自动分发到伸缩组内的所有弹性云主机。注意：一个伸缩组可最多添加 10 组负载均衡监听器。当选择使用弹性负载均衡时，需配置 4 个参数：负载均衡器、后端云主机组、后端端口和权重。
实例移除策略	是	<p>在对伸缩组进行自动缩容时，会采用此策略进行移除，共有以下四种方式：</p> <ul style="list-style-type: none"> - 较早创建的配置且较早创建的实例：先筛选出较早创建的配置所创建出的实例，再筛选较早创建的实例移出。 - 较晚创建的配置且较晚创建的实例：先筛选出较晚创建的配置所创建出的实例，再筛选较晚创建的实例移出。 - 较早创建的实例 (FIFO)：根据时间筛选较早创建的实例，跟是否是伸缩配置创建的无关。 - 较晚创建的实例 (LIFO)：根据时间筛选较晚创建的实例，跟是否是伸缩配置创建的无关。
实例回收模式	是	<p>指伸缩组自动创建的实例被移出后的处理策略。释放模式：将移出的伸缩组自动创建的云主机释放。</p> <p>注意：对于手动移入伸缩组的云主机实例，如果被移出伸缩组，资源不会被释放，状态保持不变，不适用此策略。</p>
健康检查方式	是	<p>健康检查会将处于“异常”状态的云主机从伸缩组中移出，并替换同种规格的云主机来承载业务。有以下两种方式：</p> <ul style="list-style-type: none"> - 云主机健康检查：是指对云主机的运行状态进行检查，如关机、删除都是云主机异常状态。 - 弹性负载均衡健康检查：是指根据负载均衡对云主机的健康检查结果进行的检查。所有监听器下检测到的云主机状态必须均为正常。注意：只有开启负载均衡时您才可以选择弹性负载均衡健康检查，建议当伸缩组开启负载均衡时，使用弹性负载均衡健康检查。
健康检查间隔	是	执行健康检查的周期时间。您可以根据业务实际情况设置合理的健康检查间隔 (5 分钟、15 分钟、1 小时、3 小时)。
企业项目	是	支持为伸缩组选择企业项目。

参数	是否必选	参数说明
标签	否	为伸缩组添加标签便于搜索和资源管理。一个伸缩组最多可绑定 10 组标签键和值。

- 参数配置完成之后，请单击右下角的“下一步”，进入“创建弹性伸缩组 > 配置来源”页面，进行伸缩配置的创建与选择。
- 在“伸缩配置来源”选项，您可以选择使用已有的伸缩配置或者即时创建新的伸缩配置。此处选择“新建配置”，进入到伸缩配置新建页面。



The screenshot displays the 'Create Scaling Configuration' interface. Key sections include:

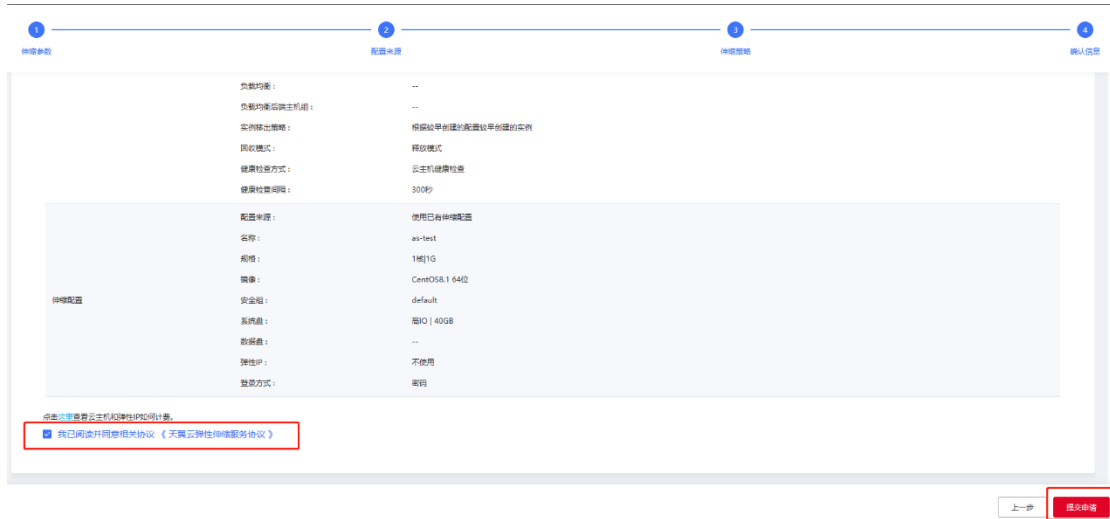
- Name:** es-4345
- Region:** 西北-内网演示3
- Configuration Source:** 使用新镜像 (Selected)
- Instance Type Selection:** A table with columns for Instance Name, vCPU, Memory (GB), Network Bandwidth, Maximum Network I/O, and Elastic IP. The selected instance is s8.large.2.
- Image Selection:** A dropdown menu with a 'Select Image' button.
- System Disk:** A section for configuring the system disk, currently set to '普通IO'.
- Instance IP:** Options for '不使用的' (Not used) or '自动分配' (Automatic allocation).
- Authentication:** Fields for Username (root), Password, and Confirm Password.
- Tags:** A toggle switch for '是否绑定标签' (Bind tags).

- 在伸缩配置新建页面需填写伸缩配置信息，例如配置名称、云主机镜像、云主机类型、登录方式等，具体参数说明如下表：

参数	是否必选	参数说明
名称	是	创建伸缩配置的名称。
可用区	是	多可用区资源池可选择伸缩组绑定的伸缩配置，在指定可用区扩缩容。当伸缩配置绑定的伸缩组使用优先级扩容策略时，可用区选择顺序将决定优先级。
配置规格	是	配置规格有两种选择，使用新规格与使用现有云主机规格： <ul style="list-style-type: none"> - 选择“使用现有云主机规格>请选择云主机”创建配置，云主机类型、vCPU、内存、镜像、云硬盘数据、安全组信息将默认与选择的云主机规格保持一致。 - 选择“使用新规格”，接下来根据实际业务需求配置云主机类型、vCPU、内存、镜像（支持公共镜像、私有镜像或共享镜像）、云硬盘参数、安全组信息。
弹性 IP	是	弹性 IP 可以实现虚拟私有云中的云资源通过固定的公网 IP 地址与互联网互通。您可以根据实际需求选择以下两种方式： <ul style="list-style-type: none"> - 不使用：弹性云主机不能对外提供访问服务，仅可在虚拟私有云内部进行内网互通。 - 自动分配：自动为每台弹性云主机分配独享带宽的弹性 IP，带宽值可以由您设定。
登录方式	是	天翼云提供两种登录方式供您选择，密钥对和密码。密钥对指使用密钥作为弹性云主机的鉴权方式进行登录。选择此方式，请您在密钥对配置项中导入密钥对。 <p>说明：</p> <ul style="list-style-type: none"> - 如果您直接从下拉列表中选择已有的密钥，请确保您已在本地获取该文件且可以登录云主机，否则，将影响您正常登录弹性云主机。 - 密码指使用设置 root 用户（Linux 操作系统）和 Administrator 用户（Windows 操作系统）的初始密码方式作为弹性云主机的鉴权方式，如果选择此方式，您可以通过用户名密码方式登录弹性云主机。您需要为您的用户设置密码，并确认密码，确保能够成功登录云主机。

9. 配置完成之后，可以单击“立即创建”按钮，完成伸缩组与伸缩配置的创建，进入

到“确认信息”页面，确认创建的弹性伸缩组、伸缩配置信息，确认无误后勾选“我已阅读并同意相关协议《天翼云弹性伸缩服务协议》”。点击右下角“提交申请”，即可完成创建。



10. 在第 9 步时，用户还可以选择单击右下角的“下一步”进入到“创建弹性伸缩组>伸缩策略”页面，继续进行伸缩策略的创建，为伸缩组绑定伸缩策略，具体操作步骤请参见“快速入门>快速扩缩容弹性云主机>创建伸缩策略”。
11. 申请提交成功后，可以回到弹性伸缩组列表，查看已创建的弹性伸缩组，当伸缩组状态为“已启用”时，说明已经创建成功。

添加负载均衡器

操作场景

弹性负载均衡（CT-ELB，Elastic Load Balancing）是一种分发和控制网络流量的服务，通过预先设定的算法将访问流量自动分发到多台云主机，扩展应用系统对外的服务能力，实现更高水平的应用系统容错性能。

弹性负载均衡可以通过流量分发扩展应用系统对外的服务能力，通过消除单点故障提升应用

系统的可用性。若您需要使用弹性负载均衡提供的功能，请参考此章节为您的伸缩组添加负载均衡器。

选择启用负载均衡后，伸缩组中的云主机会自动挂载到您关联的负载均衡下。访问流量将按照分发策略自动分发到伸缩组内的所有弹性云主机。弹性伸缩只能添加已创建的负载均衡器，且弹性伸缩组和负载均衡器必须处于同一 VPC 内。如何创建负载均衡器请参见[弹性负载均衡快速入门](#)。

注意：

一个伸缩组可最多添加 10 组负载均衡监听器。

操作步骤

创建伸缩组时，您可以在“负载均衡”配置项中选择启用，此时您需配置 4 个参数：负载均衡器、后端云主机组、后端端口和权重。




修改伸缩组信息

操作场景

修改伸缩组信息时，您可以修改的参数有：伸缩组名称、最大实例数、最小实例数、健康检查方式、健康检查间隔、实例移出策略。

当伸缩组为已停用状态、实例数为 0 且没有正在执行的伸缩活动时，才可以设置可用区和负载均衡等配置选项。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在“伸缩组”页签的弹性伸缩组列表中，单击待修改伸缩组所在行的“修改”，进入到“修改伸缩组”页面。



说明：

您也可以单击待修改信息的伸缩组名称，在伸缩组详情信息页面的右上角点击“修改”进入到修改页面进行伸缩组信息的修改。



5. 在“修改伸缩组”页面，修改配置信息，修改完成后，点击右下角的“确认”按钮。即可修改成功。

更换伸缩配置


操作场景

您可以根据实际业务需求，为伸缩组更换伸缩配置，同时也为伸缩组内的弹性云主机更换规格。

注意：

更换伸缩配置时，若伸缩组正在进行伸缩活动，当前伸缩活动中的实例依旧遵循更换前的配置，更换之后的伸缩配置将在下一次伸缩活动实例配置时生效。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在待更换伸缩配置的伸缩组所在行，单击“操作”列的“更多”，在下拉框中选择“更换伸缩配置”。



5. 在弹出的“更换伸缩配置”窗口中，勾选您需要更换的配置，并点击“确定”，即可为弹性伸缩组更换伸缩配置。



注意：

若伸缩配置列表中无合适的伸缩配置，您可以自行创建新的伸缩配置再参照本文进行更换，具体操作请参见[创建伸缩配置](#)。

启用伸缩组

操作场景

弹性伸缩的伸缩组状态如下表所示：


状态	状态说明
已启用	创建的弹性伸缩组处于“已启用”状态，此时可以对弹性伸缩组做任何操作。
已停用	将处于启用状态的伸缩组停用后，此时伸缩组处于停用状态，此时伸缩组无法触发伸缩活动。

注意：

仅当伸缩组状态为“已停用”才可以启用伸缩组。只有状态为“已启用”的伸缩组，此伸缩组才可能触发伸缩活动。启用伸缩组后，当伸缩组内的当前云主机数量小于“最小实例数”或大于“最大实例数”时，系统自动添加或减少相应数量的云主机，使其满足当前伸缩组的

创建信息。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在状态为“已停用”且待修改启用状态的伸缩组所在行，单击“操作”列的“启用”。



5. 在跳出的“启用”确认弹窗中，确认待启用的伸缩组信息，点击“确定”，即可启用此伸缩组。



说明：您也可以单击待启用的伸缩组名称，在伸缩组详情信息页面的右上角点击“启用”进行伸缩组状态的更改。



停用伸缩组


操作场景

伸缩组中的实例可能会有升级或者更换配置的需求，以防此类操作触发伸缩活动，直接将实例移除，用户可以先对伸缩组做停用操作，再对伸缩组内实例进行配置，最后重新启用伸缩组，即可稳定的实现实例管理。

约束与限制

- 伸缩组有两种状态，已停用与已启用。
- 只有当伸缩组状态为已启用时，才可以对伸缩组进行停用操作。
- 当伸缩组停用之后，将无法触发任何伸缩活动。
- 若伸缩组正在执行伸缩活动，对伸缩组执行停用操作，此时伸缩组的正在执行的伸缩活动也不会停止。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在状态为“已启用”且待修改状态的伸缩组所在行，单击“操作”列的“停用”。



弹性伸缩组 伸缩配置

你还可以创建 9 个弹性伸缩组。 [了解配置详情](#)

伸缩组名称/ID	状态	当前实...	期望实...	最小实...	最大实...	企业项目	伸缩配置	网络配置	操作
as-group-f518-5781	已启用	0	--	1	5	default	as-test-5909	VPC: vpc-72cd1192-168.0.0/16	修改 停用 更多

5. 在弹出的“停用”窗口中，确认待停用的伸缩组信息，确认无误后，单击“确定”，即可停用该弹性伸缩组。此时回到伸缩组列表看到弹性伸缩组的状态为“已停用”。

弹性伸缩组 伸缩配置

您还可以创建 9 个弹性伸缩组。 [了解配置详情](#)

请输入名称

伸缩组名称/ID	状态	当前实...	期望实...	最小实...	最大实...	企业项目	伸缩配置	网络配置	操作
as-group-f518 5781	已停用	0	...	1	5	default	as-test 5909	VPC: vpc-72cd(192.168.0.0/16)	修改 应用 更多

说明：您也可以单击待停用的伸缩组名称，在伸缩组详情信息页面的右上角点击“停用”进行伸缩组状态的更改。

as-group-f518详情



伸缩组ID	5781	伸缩组名称	as-group-f518
状态	已停用	创建时间	2023-08-08 18:07:07
最小实例数	1	最大实例数	5
当前实例数	0	期望实例数	...
子网	查看详情	VPC	vpc-72cd
实例移出策略	根据最早创建时间释放最早的实例	负载均衡	...
负载均衡器的主机	...	回收模式	释放模式
健康检查方式	云主机健康检查	健康检查间隔	5分钟
伸缩配置	as-test	企业项目	default

删除伸缩组

操作场景

当您不再需要某个伸缩组时，可以删除此伸缩组。天翼云提供两种伸缩组删除方式：

- 普通删除：伸缩组无正在进行中的伸缩活动时，才可被删除。
- 强制删除：伸缩组将被立即删除，若伸缩组有正在进行中的伸缩活动，待活动执行完后，原伸缩组内的实例将被移出或移出并释放。


若您删除了伸缩组，将自动执行如下操作：

- 伸缩组内自动创建的实例将被移出伸缩组并被释放。
- 手动移入的实例将被自动移出伸缩组但不会被释放。
- 伸缩组绑定的伸缩策略将被删除。
- 伸缩配置会被解绑但不会被删除。

注意：

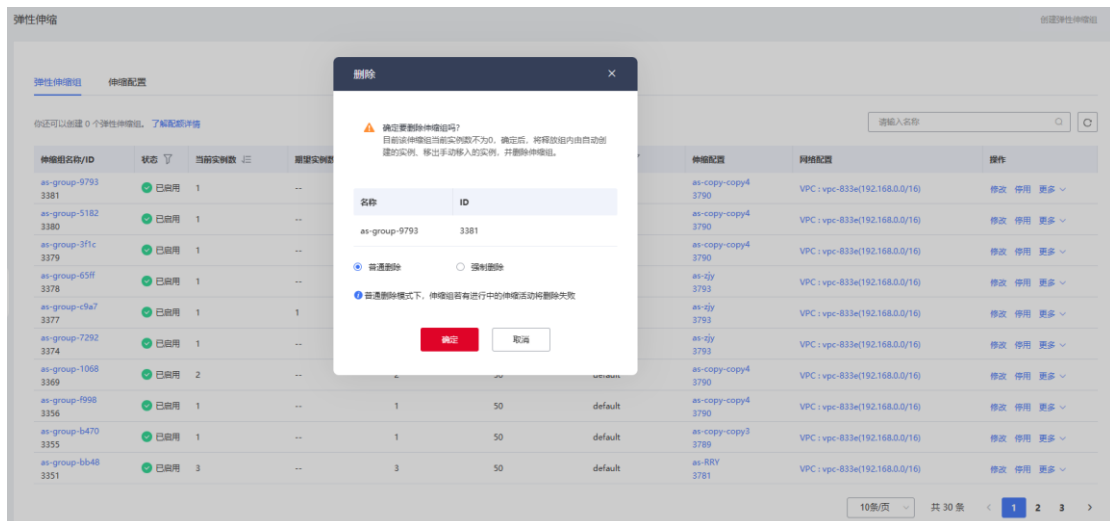
若您只是某个时间段不使用伸缩组，建议您优先选择停用伸缩组的方式，而不是直接删除。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在待删除伸缩组所在行，单击“操作”列的“更多”，下拉选择“删除”。



5. 在“删除”弹窗中，需选择伸缩组的删除方式：普通删除或强制删除。



- 普通删除：伸缩组需满足无任何正在进行中的伸缩活动的要求即可删除成功，否则将删除失败。
- 强制删除：无论伸缩组是否满足无任何正在进行中的伸缩活动的条件都将被立即删

除，删除后，伸缩组关联的伸缩活动将在执行完成后，根据实例类型执行移出或移出并释放操作。

伸缩组保护


操作场景

当您不希望伸缩组被删除时，可以为其设置保护状态，设置了保护状态的伸缩组，不可在伸缩组列表被删除。

约束与限制

- 开启保护的伸缩组无法通过控制台或 OpenAPI 删除。

设置保护状态

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在弹性伸缩组列表，选择对需要开启保护的伸缩组，单击“操作”列，选择更多菜单中的“开启伸缩组保护”。




说明：

1. 已开启保护的伸缩组，无法通过控制台、OpenAP 被 | 删除或被强制删除。

2. 伸缩组创建成功后，默认未开启伸缩组保护，需手动开启。

关闭实例保护

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在弹性伸缩组列表，选择对需要关闭保护的伸缩组，单击“操作”列，选择更多菜单中的“关闭伸缩组保护”。



伸缩配置管理

创建伸缩配置

使用现有云主机创建伸缩配置

操作场景

伸缩配置用于定义伸缩组资源扩展时云主机的规格，包括云主机的类型、vCPU、内存等。

自动扩容时，伸缩组根据伸缩配置来创建弹性云主机实例，并将创建的云主机实例添加到伸缩组。

若您需要使用与某云主机实例一样的类型、规格、镜像作为伸缩组扩容云主机的模板，可在创建伸缩配置时，可选择“使用现有云主机规格”。

约束与限制


- 单个资源池最大可创建 100 个伸缩配置。
- 单个伸缩组只能关联 1 个伸缩配置。

若单个资源池伸缩配置配额不足，可提交工单提升配额值。

操作步骤

创建伸缩配置有两个进入窗口：

- 在弹性伸缩管理控制台，进入“伸缩配置”页签，点击右上角“创建伸缩配置”按钮进行创建。本文将重点介绍此方式。
- 创建伸缩组时，在“配置来源”页面，即可创建新的伸缩配置。具体操作请参见“快速入门>快速扩缩容弹性云主机>[创建伸缩配置](#)”。

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 单击“伸缩配置”，进入“伸缩配置”页签，点击右上角“创建伸缩配置”按钮，进入“创建伸缩配置”页面。



5. 在“创建伸缩配置”页面，用户可以配置伸缩配置相关参数，具体参数配置如下表所示：

参数	是否必选	参数说明
名称	是	创建伸缩配置的名称。
配置规格	是	配置规格有两种选择，使用新规格与使用现有云主机规格： <ul style="list-style-type: none"> - 选择“使用现有云主机规格 > 请选择云主机”快速创建伸缩配置。创建配置时，云主机类型、vCPU、内存、镜像、云硬盘数据、安全组信息将默认与选择的云主机规格保持一致。 - 选择“使用新规格”，根据实际业务需求配置云主机类型、vCPU、内存、镜像、云硬盘参数、安全组信息。
弹性 IP	是	弹性 IP 可以实现虚拟私有云中的云资源通过固定的公网 IP 地址对外提供访问服务。您可以根据实际需求选择以下两种方式： <ul style="list-style-type: none"> - 不使用：弹性云主机不能与互联网互通，仅可在虚拟私有云内部使用。 - 自动分配：自动为每台弹性云主机分配独享带宽的弹性 IP，带宽大小可以根据您的业务需要设定。
登录方式	是	天翼云提供两种登录方式供您选择，密钥对和密码。 <ul style="list-style-type: none"> - 密钥对：指使用密钥作为弹性云主机的鉴权方式进行登录。选择此方式，请您在密钥对配置项中导入密钥对。如果您直接从下拉列表中选择已有的密钥，请确保您已在本地获取该文件，否则，将影响您正常登录弹性云主机。 - 密码：指使用设置 root 用户（Linux 操作系统）和 Administrator 用户（Windows 操作系统）的初始密码方式作为弹性云主机的鉴权方式，如果选择此方式，您可以通过用户名密码方式登录弹性云主机。您需要为您的用户设置密码用于登录云主机。

说明：

创建伸缩配置时选择“使用现有云主机规格”可完成伸缩配置的建设。

创建伸缩配置

* 名称: ⓘ

* 配置规格:

* 弹性IP: ⓘ

不使用弹性公网IP的云主机不能与互联网互通, 仅可作为私有网络中部署业务或者集群所需云主机进行使用。

* 登录方式:

* 用户名:

* 密码: ⓘ

* 确认密码:

请妥善保管密码, 系统无法获取您设置的密码内容

使用新模板创建伸缩配置

操作场景

伸缩配置用于定义伸缩组资源扩展时云主机的规格, 包括云主机的类型、vCPU、内存等。

自动扩容时, 伸缩组根据伸缩配置来创建弹性云主机实例, 并将创建的云主机实例添加到伸缩组。

若您需要定义全新的云主机类型、规格、镜像作为伸缩组扩容云主机模板, 可在创建伸缩配置时, 可选择“使用新规格”。

约束与限制


- 单个资源池最大可创建 100 个伸缩配置。
- 单个伸缩组只能关联 1 个伸缩配置。

若单个资源池伸缩配置配额不足, 可提交工单提升配额值。

操作步骤

创建伸缩配置有两个进入窗口:

- 在弹性伸缩管理控制台，进入“伸缩配置”页签，点击右上角“创建伸缩配置”按钮进行创建。本文将重点介绍此方式。
- 创建伸缩组时，在“配置来源”页面，即可创建新的伸缩配置。具体操作请参见“快速入门>快速扩缩容弹性云主机>[创建伸缩配置](#)”。

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 单击“伸缩配置”，进入“伸缩配置”页签，点击右上角“创建伸缩配置”按钮，进入“创建伸缩配置”页面。



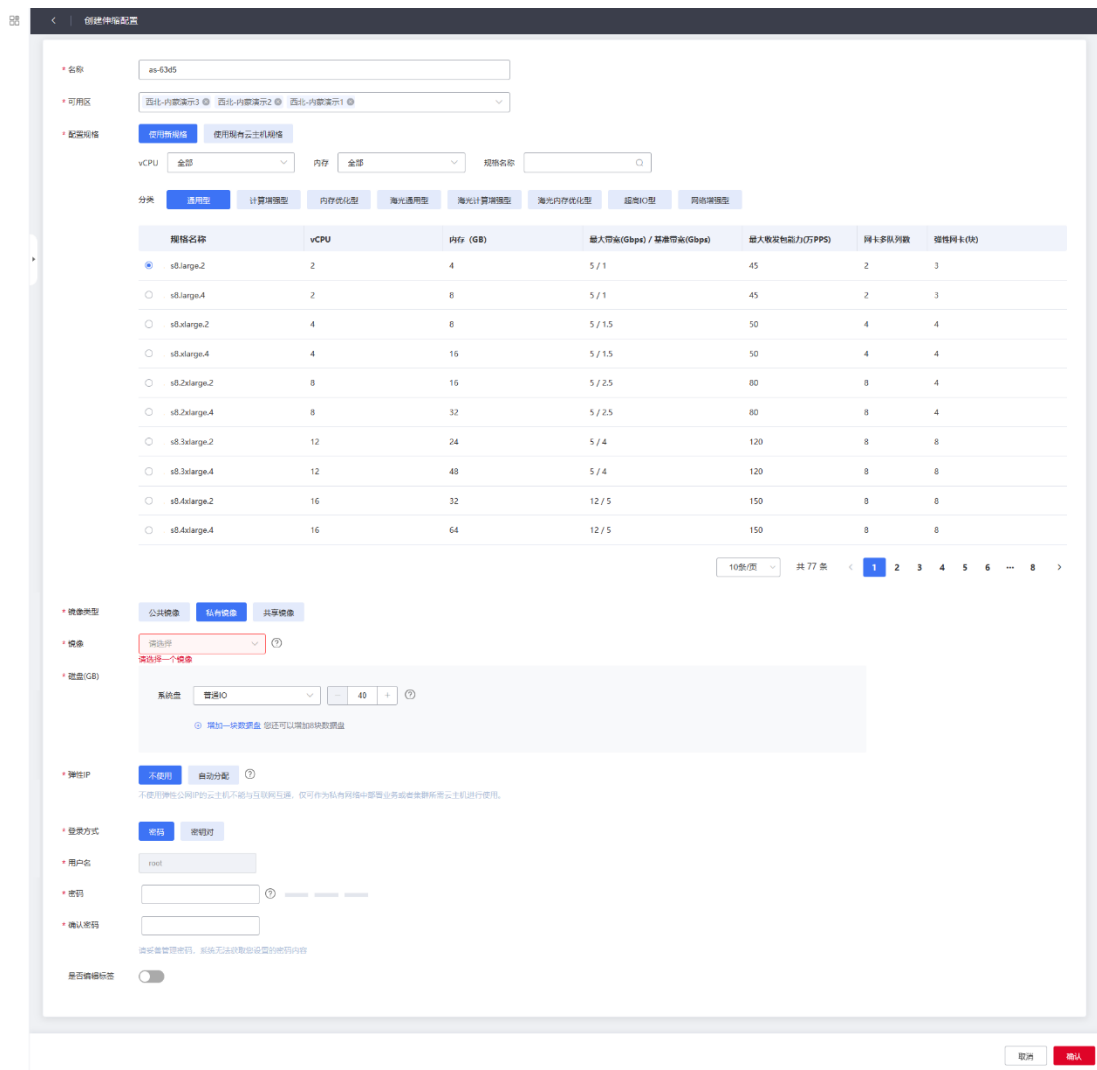
5. 在“创建伸缩配置”页面，用户可以配置伸缩配置相关参数，具体参数配置如下表所示：

参数	是否必选	参数说明
名称	是	创建伸缩配置的名称。
可用区	是	多可用区资源池可选择伸缩组绑定的伸缩配置，在指定可用区扩缩容。当伸缩配置绑定的伸缩组使用优先级扩容策略时，可用区选择顺序将决定优先级。
配置规格	是	配置规格有两种选择，使用新规格与使用现有云主机规格： - 选择“使用现有云主机规格 > 请选择云主机”快速创建伸缩配置。创建配置时，云主机类型、vCPU、内存、镜像、云硬盘数据、安全组信息将默认与选择的云主机规格保持一致。

参数	是否必选	参数说明
		<ul style="list-style-type: none">- 选择“使用新规格”，根据实际业务需求配置云主机类型、vCPU、内存、镜像镜像（支持公共镜像、私有镜像或共享镜像）、云硬盘参数、安全组信息。
弹性 IP	是	<p>弹性 IP 可以实现虚拟私有云中的云资源通过固定的公网 IP 地址对外提供访问服务。您可以根据实际需求选择以下两种方式：</p> <ul style="list-style-type: none">- 不使用：弹性云主机不能与互联网互通，仅可在虚拟私有云内部使用。- 自动分配：自动为每台弹性云主机分配独享带宽的弹性 IP，带宽大小可以根据您的业务需要设定。
登录方式	是	<p>天翼云提供两种登录方式供您选择，密钥对和密码。</p> <ul style="list-style-type: none">- 密钥对：指使用密钥作为弹性云主机的鉴权方式进行登录。选择此方式，请您在密钥对配置项中导入密钥对。如果您直接从下拉列表中选择已有的密钥，请确保您已在本地获取该文件，否则，将影响您正常登录弹性云主机。- 密码：指使用设置 root 用户（Linux 操作系统）和 Administrator 用户（Windows 操作系统）的初始密码方式作为弹性云主机的鉴权方式，如果选择此方式，您可以通过用户名密码方式登录弹性云主机。您需要为您的用户设置密码用于登录云主机。

说明：

创建伸缩配置时选择“使用新规格”可完成伸缩配置创建。



6. 配置完成之后，点击右下角的“确定”，即可创建成功。创建成功的伸缩配置可在伸缩配置列表中查看。

复制伸缩配置

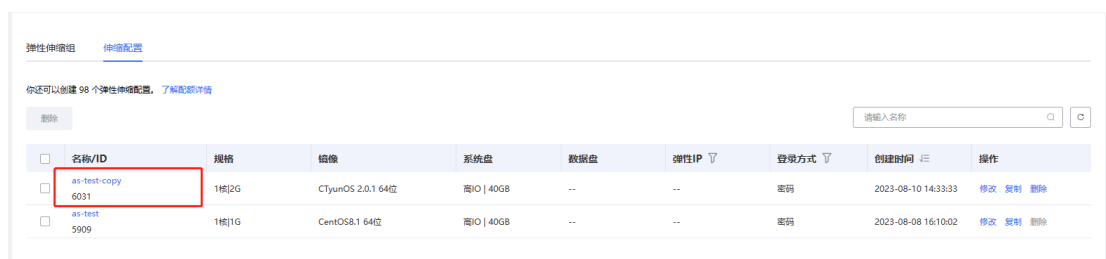
操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。

- 单击“伸缩配置”，进入“伸缩配置”页签，在待复制的伸缩配置行，单击“操作”列的“复制”。



- 在“复制伸缩配置”页面，您可以修改配置名称、云主机规格和镜像等参数，若想要同一种配置，可保持参数不变，点击右下角“确认”按钮。
- 复制伸缩配置成功后回到伸缩配置列表，用户可以看到新增的伸缩配置信息。




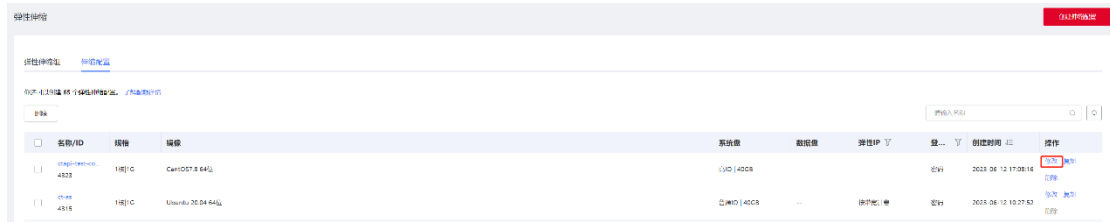
修改伸缩配置

操作场景

若当前伸缩配置不能满足现有需求，您可按需修改伸缩配置。

操作步骤

- 登录控制中心。
- 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
- 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
- 单击“伸缩配置”，进入“伸缩配置”页签，在待修改的伸缩配置行，单击“操作”列的“修改”。



5. 修改相关数据;
6. 单击【确认】, 完成修改。

删除伸缩配置


操作场景

当您不再使用某个伸缩组时, 可以删除该伸缩配置。

注意:

当删除指定伸缩配置时, 只有当该伸缩配置不被任何弹性伸缩组绑定时, 才允许被删除。您可以选择单个删除, 也可以选择批量删除。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 , 选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”, 进入弹性伸缩管理控制台。
4. 单击“伸缩配置”, 进入“伸缩配置”页签, 在待删除的伸缩配置行, 单击“操作”列的“删除”。



5. 在“删除”弹窗中，确认待删除配置，确认无误后，点击“确定”，即可删除此条伸缩配置。
6. 当您需要批量删除伸缩配置时，您可以勾选待删除伸缩配置前的复选框，单击列表上方的“删除”，即可将伸缩配置删除。

伸缩策略管理

伸缩策略概述

伸缩策略是弹性伸缩实现动态扩缩容的重要机制。通过伸缩策略，可定义自动伸缩活动的触发条件、执行的伸缩任务，从而实现自动扩缩容。

使用限制

- 一个伸缩组最多可以创建 10 个伸缩策略。
- 若伸缩策略设置了冷却时间 300s，则当策略触发后，300s 内不会再次执行同一伸缩策略。
- 当一个伸缩组拥有多个伸缩策略时，若多个伸缩策略之间互不冲突，后续实例只要满足策略条件，均会触发伸缩活动。
- 伸缩策略执行后伸缩组将会自动增加或减少云主机实例，但是伸缩组内的实例数不会超过最大实例数和最小实例数的范围，若超过最大实例数，会按照实例移除规则自动移除组内实例，若小于最小实例数将会自动增加实例。

伸缩策略

伸缩策略用于定义伸缩活动的触发模式、触发条件和触发动作。伸缩策略通常有三种：告警策略、定时策略、周期策略。

- 告警策略：通过对伸缩组内实例性能指标（CPU 使用率、内存使用率等）的监控，来确认其是否到达预设的告警条件，来自动增加或减少云主机的数量。
- 定时策略：根据业务实际情况设置一个时间点，在此时间点自动增加或减少云主机的数量。
- 周期策略：根据业务实际情况设置一个时间段，在此时间段内按照周期（按天、按周、按月）来重复执行自动增减云主机的数量。

创建伸缩策略

操作场景

伸缩策略用于定义伸缩活动的触发模式、触发条件和触发动作。伸缩策略通常有三种：告警策略、定时策略、周期策略。


- 告警策略：通过对伸缩组内实例性能指标（CPU 使用率、内存使用率等）的监控，来确认其是否到达预设的告警条件，来自动增加或减少云主机的数量。
- 定时策略：根据业务实际情况设置一个时间点，在此时间点自动增加或减少云主机的数量。
- 周期策略：根据业务实际情况设置一段时间段，在此时间段内按照周期（按天、按周、按月）来重复执行自动增减云主机的数量。

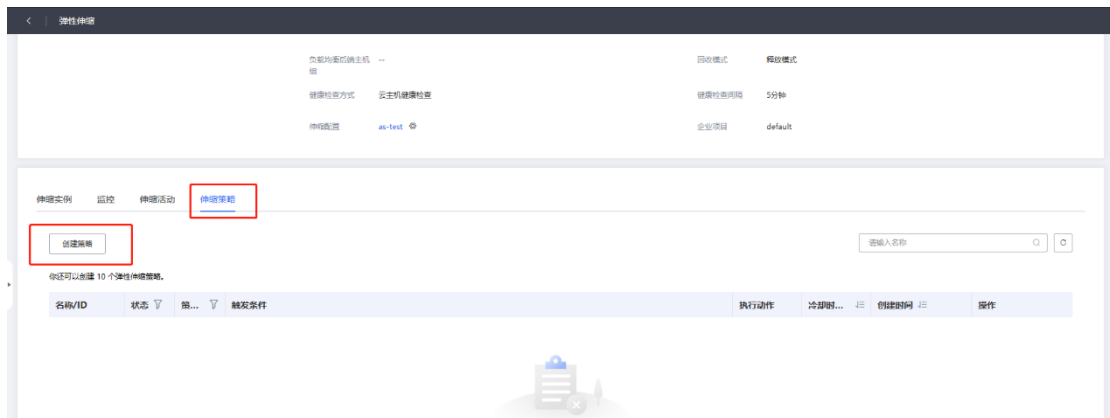
操作步骤

创建伸缩策略有两个进入窗口：

- 在待创建伸缩策略的伸缩组详情页面，点击“创建策略”，进入到伸缩策略创建页面。本文将会详细介绍此种方式。

- 创建伸缩组时，完成伸缩组、伸缩配置的创建之后，点击“下一步”进入“伸缩策略”页面，即可创建新的伸缩策略。具体操作请参见“快速入门>快速扩缩容弹性云主机>[创建伸缩策略](#)”。

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待添加伸缩策略的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击“创建策略”按钮，进入到创建伸缩策略页面。



6. 在创建伸缩策略页面，可配置具体的策略信息，用户可选择创建三种策略，分别为告警策略，定时策略，周期策略，三种策略类型的参数不同，具体参数说明如以下三个表所示：

- 告警策略参数说明：

参数	是否必选	参数说明
名称	是	创建伸缩策略的名称。
策略类型	是	此处选择“告警策略”。

参数	是否必选	参数说明
告警规则	是	保持默认配置“现在创建”。当您有已经创建好的告警规则时，可以选择“使用已有”，来直接获取已创建好的告警规则模板。
告警规则名称	是	创建告警规则名称。
触发条件	是	用户可以选择弹性伸缩所支持的性能指标，并为此性能指标配置触发条件。用户可以选择的性能指标有：CPU 使用率、内存使用率、网络流入速度、网络流出速度、磁盘读速率、磁盘写速率、磁盘读请求速率、磁盘写请求速率。
监控周期	是	告警规则监控状态刷新的间隔时间。
连续出现次数	是	连续出现次数指探测结果连续几次符合您设置的规则，才会触发告警。例如设置为 3，则表示连续 3 次超过阈值后才会触发告警。
冷却时间	是	冷却时间是指用户在每次成功的伸缩活动后设置的一段锁定时间，当一次伸缩活动执行完成，系统开始计算冷却时间。为避免告警策略的频繁触发，在冷却时间内，由告警策略触发的伸缩活动将会被拒绝执行，其他类型的伸缩活动不受限制，但执行完成之后将会重新开始冷却时间计时。 例如，将告警策略冷却时间设置为 300 秒，伸缩组 08:00 时由于告警策略触发并完成了伸缩活动，则在 08:05 前，伸缩组会拒绝新告警触发的伸缩活动，但不会拒绝在 08:00-08:05 间设置的定时或周期策略触发的伸缩活动。
执行动作	是	指当满足设置的告警策略进行伸缩活动时，弹性伸缩服务将执行的动作。用户可在此设置以下三种动作：增加 n 台云主机实例。减少 n 台云主机实例。设置为 n 台云主机实例。

- 定时策略参数说明：

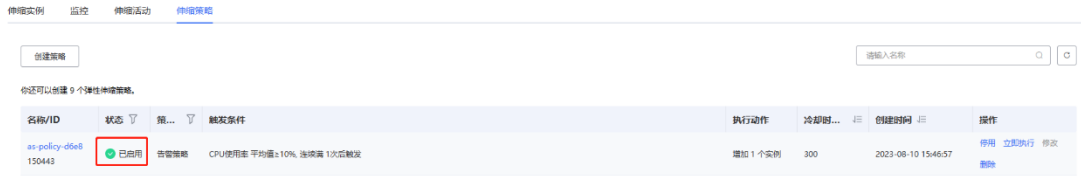
参数	是否必选	参数说明
名称	是	创建伸缩策略的名称。
策略类型	是	此处选择“定时策略”。
时区	否	为默认值：GMT+08:00，代表北京时间。
触发时间	是	设置伸缩活动执行时间。注意：请确保触发时间晚于当前时间，否则定时策略不会被触发执行。
执行动作	是	指当满足设置的定时策略进行伸缩活动时，弹性伸缩服务将执行的动作。用户可在此设置以下三种动作：增加 n 台云主机实例。减少 n 台云主机实例。设置为 n 台云主机实例。

- 周期策略参数说明：

参数	是否必选	参数说明
名称	是	创建伸缩策略的名称。
策略类型	是	此处选择“周期策略”。
周期	是	执行伸缩活动的重复周期，用户可在按天、按周、按月中选择。
时区	否	为默认值：GMT+08:00，代表北京时间。
触发时间	是	设置伸缩活动执行时间。注意：触发时间须在策略的生效时间段内。
生效时间	是	伸缩策略可被触发的有效时间段。
执行动作	是	指当满足设置的周期策略进行伸缩活动时，弹性伸缩服务将执行的动作。用户可在此设置以下三种动作：增加 n 台云主机实例。减少 n 台云主机实例。设置为 n 台云主机实例。

7. 用户根据上述参数说明完成伸缩策略配置，点击右下方“确认”，创建策略成功后，回到弹性伸缩详情页面中的“伸缩策略”页签，可以查看到已创建好的伸缩策略，

且此策略默认为“已启用”状态。




名称/ID	状态	策略...	触发条件	执行动作	冷却时...	创建时间	操作
as-policy-d6e8-150443	已启用	告警策略	CPU使用率: 平均值≥10%, 连续高 1次后触发	增加 1 个实例	300	2023-08-10 15:46:57	停用 立即执行 修改 删除

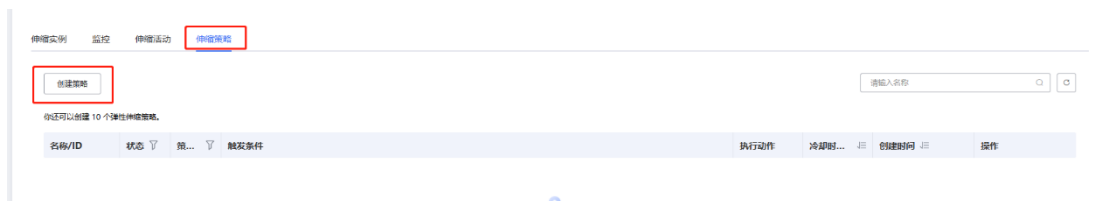
冷却时间

操作场景

冷却时间是指用户在每次成功的伸缩活动后设置的一段锁定时间,当一次伸缩活动执行完成,系统开始计算冷却时间。在冷却时间内,由告警策略触发的伸缩活动将会被拒绝执行,其他类型的伸缩活动不受限制,但执行完成之后将会重新开始冷却时间计时,以避免告警策略的频繁触发。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的  , 选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”, 进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中, 点击待添加冷却时间的伸缩策略所在伸缩组名称, 进入到伸缩组详情。
5. 在下方的“伸缩策略”页签, 点击“创建策略”, 进入到创建伸缩策略页面。



注意:

若需要修改已有策略的冷却时间，可先停用策略，停用之后，点击“操作>修改”，进入到“修改伸缩策略”页面，对冷却时间进行修改。

6. 在“创建伸缩策略”页面，当策略类型为告警策略时，用户可设置冷却时间，冷却时间默认为 300 秒。



The screenshot shows the configuration interface for a scaling policy. The 'Name' field is 'as-policy-3899'. The 'Policy Type' is 'Alert Policy'. The 'Alert Rule' is 'Create Now'. The 'Alert Rule Name' is 'as-alarm-3899'. The 'Trigger Condition' is 'CPU Usage Rate' with a 'Average Value' of '≥ 10%'. The 'Monitoring Period' is '1 minute'. The 'Continuous Occurrence Count' is '1'. The 'Cooling Time (seconds)' is '300', which is highlighted with a red box. The 'Execution Action' is 'Add' with '1' instances.

7. 配置完成，点击右下角的“确认”，即为告警策略设置冷却时间成功。

修改伸缩策略

操作场景

您可以对已经创建成功的伸缩策略执行修改操作，修改其策略名称、触发条件、冷却时间等。

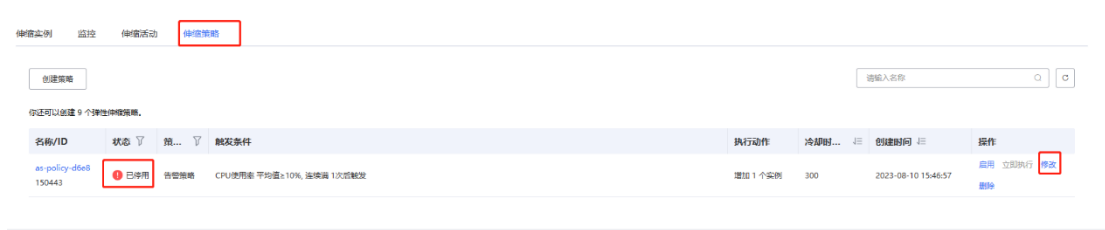
约束与限制

伸缩策略只有在“已停用”状态下才可以进行修改。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。

4. 点击待修改伸缩策略的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击待修改的伸缩策略所在行的“操作>修改”，进入到“修改伸缩策略”页面。



6. 在“修改伸缩策略”页面，您可修改策略名称、执行动作等，或者创建新告警规则，修改完成后，单击“确认”，即可成功修改伸缩策略。

启用伸缩策略

操作场景


弹性伸缩的伸缩策略状态如下表所示：

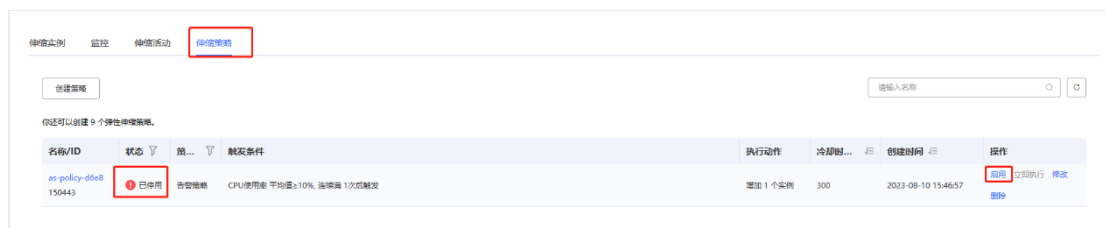
状态	状态说明
已启用	创建的弹性伸缩策略处于“已启用”状态，此时将会正常按照策略条件触发伸缩活动。新创建的伸缩策略默认为启用状态。
已停用	将处于启用状态的伸缩策略停用后，此时伸缩组将无法正常触发伸缩活动。

注意：

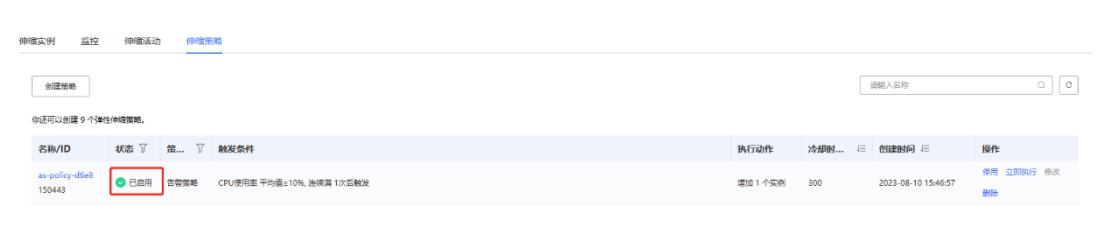
- 仅当伸缩策略状态为“已停用”时才可以启用伸缩组。
- 您可以选择单个启用，也可以选择批量启用策略，但是同一个伸缩组内的多个策略同时启用时，要确保互不冲突。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待启用伸缩策略的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击待启用的伸缩策略所在行的“操作>启用”。



6. 在弹出的“启用”确认弹窗中，确认启用信息，确认无误后，点击“确定”，即可成功启用策略。




停用伸缩策略

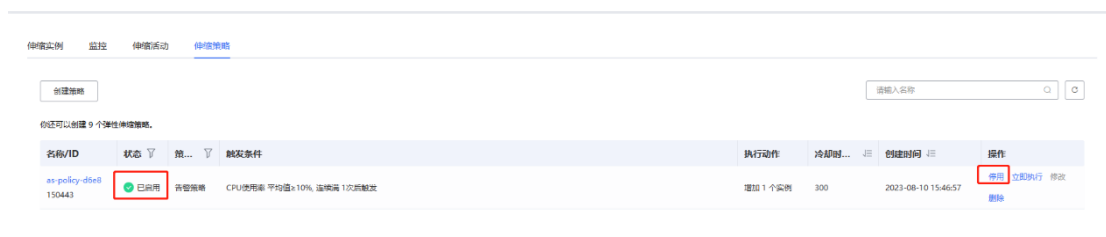
操作场景

若您某一段时间内不需要某个伸缩策略触发伸缩活动，建议您停用该策略。停用的前提条件是您的伸缩策略本身处于启用状态。

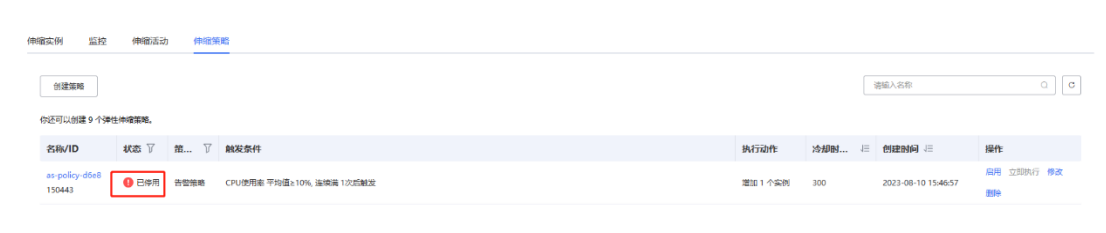
操作步骤

1. 登录控制中心。

2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待停用伸缩策略的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击待停用的伸缩策略所在行的“操作>停用”。



6. 在弹出的“停用”确认弹窗中，确认停用策略信息，确认无误后，点击“确定”，即可停用该策略。




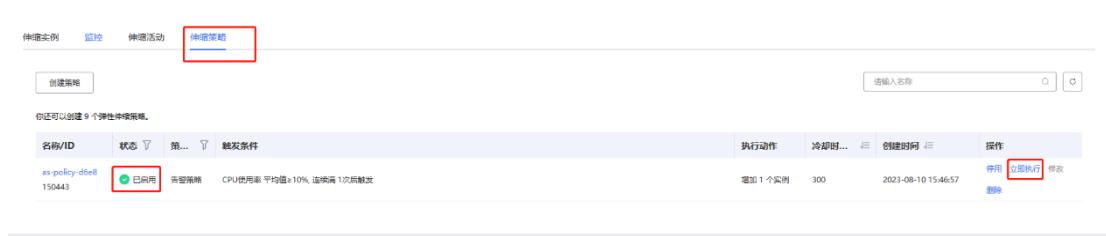
立即执行伸缩策略

操作场景

- 当伸缩组状态和伸缩策略状态都为“已启用”时，您可以立即执行伸缩策略。
- 若您设置了期望实例数，当立即执行伸缩策略触发伸缩活动后，会将伸缩组内实例数进行同步更新，使得伸缩组内实例数达到期望实例数。
- 立即执行伸缩策略时，系统不会考虑实例是否满足伸缩条件。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待立即执行伸缩策略的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击待立即执行的伸缩策略所在行的“操作>立即执行”。



6. 在“立即执行”弹窗中，确认伸缩策略信息，确认无误后，点击“确定”即可将此策略执行成功。

删除伸缩策略


操作场景

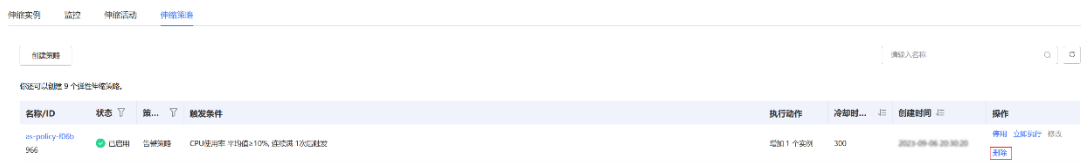
若您不再需要某个伸缩策略来触发伸缩活动，您可以选择将其删除。但您若是仅在一段时间内不需要，建议您选择停用该策略的方式，而不建议删除。

如果被删除的策略正在被触发，伸缩活动正在进行，也可以删除该策略，伸缩活动将继续执行。

操作步骤

1. 登录控制中心。

2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待删除策略的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，在待修改伸缩策略所在行的“操作”列下，单击“删除”。



6. 在“删除”确认弹窗中，确认待删除策略信息，确认无误后，单击“确定”，即可成功删除该策略。

伸缩活动管理

伸缩活动概述

当您执行伸缩策略、修改伸缩组参数、健康检查等操作都会触发伸缩动作，这些动作的记录即为伸缩活动。例如增加云主机资源或减少云主机资源的过程。

伸缩活动的状态如下表所示：

状态	状态属性	状态说明
执行中	中间状态	此伸缩组正在执行伸缩活动。
成功	稳定状态	此伸缩组中执行完成且成功的伸缩活动，例如成功添加了 1 台弹性云主机实例。
失败	稳定状态	此伸缩组中执行完成但失败的伸缩活动，伸缩失败原因可能为

状态	状态属性	状态说明
		账户余额不足或者当前可用区内资源不足等。

伸缩活动通常有以下执行类型：

- 自动伸缩：根据伸缩策略进行的自动伸缩活动。
- 手动执行策略：用户手动执行策略导致的伸缩活动。
- 手动移入：用户手动向伸缩组内移入实例。
- 手动移出：用户手动将实例从伸缩组内移出。
- 新建伸缩组满足最低数量：创建伸缩组时触发该活动。当伸缩组配置的最小实例数大于 0 且未配置期望实例数时，伸缩组创建成功即触发伸缩活动，按伸缩配置创建云主机实例，使伸缩组实例数等于最小实例数。
- 修改伸缩组满足数量：伸缩组在服务过程中，若修改最大实例数量、最小实例数量，将触发伸缩组自动检查伸缩组内实例数是否满足最大、最小实例数限制，使伸缩组内实例数量满足数量限制要求。
- 健康检查移入：为替换伸缩组内不健康云主机实例而自动移入的弹性云主机实例。
- 健康检查移出：通过云主机健康检查或负载均衡健康检查检测出伸缩组内不健康云主机实例的自动移出。
- 期望实例数检查移入：当您为伸缩组设置了期望实例数，且伸缩组内当前实例数小于期望实例数时，会自动移入云主机实例，以确保伸缩组内始终保持期望实例数。
- 期望实例数检查移出：当您为伸缩组设置了期望实例数，且伸缩组内当前实例数大于期望实例数时，会自动移出云主机实例，以确保伸缩组内始终保持期望实例数。
- 新建伸缩组满足期望实例数量：创建伸缩组时触发该活动。若创建伸缩组时，配置了期望实例数，伸缩组创建成功即触发伸缩活动，按伸缩配置创建云主机实例，使


伸缩组实例数等于期望实例数。

查询伸缩活动

操作场景

- 系统检测到伸缩策略中配置的条件满足伸缩要求。
- 手工修改伸缩组的最大/最小/期望实例数，导致当前组内实例数超过该最大/最小限制或不满足期望。
- 手动添加或移出实例。


操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待查看伸缩活动的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩活动”进入伸缩活动页签，即可查看伸缩活动列表。



The screenshot shows the '伸缩活动' (Scaling Activities) tab in the console. It displays a table with columns for Activity ID, Status, Execution Type, Before Scaling Instance Count, After Scaling Instance Count, Instance Name, Start Time, and End Time. Three activities are listed, all with a '成功' (Success) status.

活动ID	状态	执行类型	伸缩前实例总数	伸缩后实例总数	实例名称	开始时间	结束时间
5245901	成功	自动伸缩	49	50	as-config-e78e-testdx76	2023-08-14 14:49:21	2023-08-14 14:50:31
5245899	成功	自动伸缩	49	50	as-config-e78e-testqlq7	2023-08-14 14:49:21	2023-08-14 14:50:29
5245497	成功	自动伸缩	48	49	as-config-e78e-testyl3x	2023-08-14 14:44:19	2023-08-14 14:45:30

若失败，您可以点击“失败”后的  来查看失败原因。



The screenshot shows the '伸缩活动' (Scaling Activities) tab in the console. It displays a table with columns for Activity ID, Status, Execution Type, Before Scaling Instance Count, After Scaling Instance Count, Instance Name, Start Time, and End Time. Three activities are listed, all with a '失败' (Failure) status. A red box highlights the '失败' status and the information icon for the first activity.

活动ID	状态	执行类型	伸缩前实例总数	伸缩后实例总数	实例名称	开始时间	结束时间
5249419	失败	自动伸缩	51	51		2023-08-14 15:34:32	2023-08-14 15:34:32
5249043	失败	自动伸缩	51	51		2023-08-14 15:29:30	2023-08-14 15:29:31
5249041	失败	自动伸缩	51	51		2023-08-14 15:29:30	2023-08-14 15:29:30

排查伸缩活动异常

操作场景

伸缩活动有三种状态，执行中、成功、失败，具体状态说明如下：

状态	状态属性	状态说明
执行中	中间状态	此伸缩组正在执行伸缩活动。
成功	稳定状态	此伸缩组中执行完成且成功的伸缩活动，例如成功添加了 1 台弹性云主机实例。
失败	稳定状态	此伸缩组中执行完成但失败的伸缩活动，伸缩失败原因可能为账户余额不足或者当前可用区内资源不足等。

当执行失败时，用户需查证具体失败原因。

执行失败信息

失败信息	失败原因
账号余额不足	当前账号余额不足。
账号配额不足	当前账号配额不足。
实例不存在	手动移入实例在移入完成前被释放。
开通下单成功 X 台, 失败 X 台	移入实例在移入完成前被卡单、撤单、被释放。
VPC 异常	手动/自动移入伸缩组 VPC 或子网异常。
资源池余量不足	资源池或可用区资源余量不足。
退订下单成功 X 台, 失败 X 台	移出并释放的实例在释放完成前被卡单、撤单或其他异常。

失败信息	失败原因
实例已在其他伸缩组中	手动移入实例已在其他伸缩组中。
实例 VPC 与伸缩组不一致	手动移入实例 VPC 与伸缩组不一致。
伸缩组实例数量不可超过最大实例数	手动或自动移入实例至伸缩组导致组内实例总数大于最大实例数。
已有伸缩活动正在进行中	伸缩组有伸缩活动正在进行中。
伸缩组实例均在保护中，不可移出	伸缩组内实例开启保护。
伸缩组实例数量不可少于最小实例数	手动或自动移出实例至伸缩组导致组内实例总数小于最小实例数。

实例移出策略

当您执行伸缩活动时，可能会遇到需要移出伸缩组内云主机实例的情况。例如：当您执行的弹性伸缩活动导致伸缩组内当前实例数大于伸缩组最大实例数时，就需要自动移出实例。当伸缩组自动移出实例时，您可以选择实例移出策略。

实例移出共有四种方式：

移出策略	策略描述	应用场景
较早创建的配置且较早创建的实例	先筛选出较早创建的配置所创建出的实例，再筛选较早创建的实例移出。	可用于更新某个伸缩组并且逐步淘汰先前配置的实例。
较晚创建的配置且较晚创建的实例	先筛选出较晚创建的配置所创建出的实例，再筛选较晚创建的实例移出。	测试新的伸缩配置但不想在生产中保留它。

移出策略	策略描述	应用场景
较早创建的实例 (FIFO)	根据时间筛选较早创建的实例先移出, 跟是否是伸缩配置创建的无关。	将伸缩组中的实例升级为新的实例类型, 可以逐渐将较旧类型的实例替换为较新类型的实例。
较晚创建的实例 (LIFO)	根据时间筛选较晚创建的实例先移出, 跟是否是伸缩配置创建的无关。	创建时间较晚的实例优先被移出伸缩组。

当实例移出策略配置为“较早/较晚创建的实例”时, 手动移入伸缩组的云主机是不被保护的, 也有可能被移出, 如果您想保证手动移入的实例不被移出, 您可以对实例开启保护。

若未开启保护, 后续手动移入的实例在被移出时会有两种情况:

- 伸缩组自动创建的实例, 移出后将遵守您配置的“实例回收模式”规则执行, 例如直接进行释放。
- 手动移入伸缩组的实例, 移出后不删除、不更改其状态。

动态资源规划

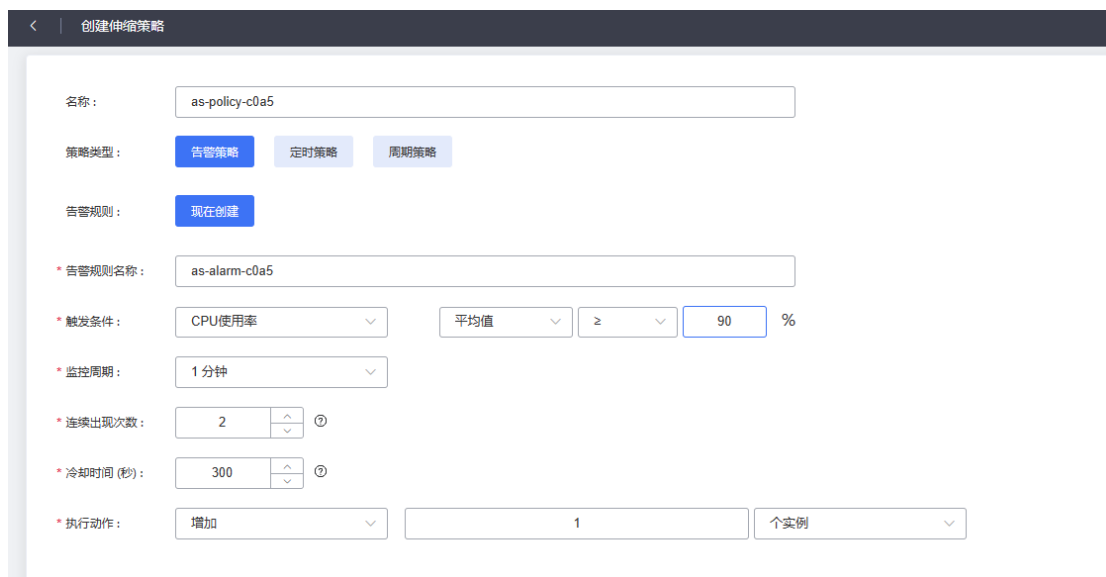
业务需求是动态变化的, 在用户实际使用弹性伸缩的过程中, 也需要依据业务变化来动态进行伸缩活动, 即动态扩展资源。

当业务访问请求是无规律且频繁变化时, 您可以选择告警策略, 业务主机的性能指标达到告警策略的阈值, 则可以开始动态伸缩。

示例: 社交平台、新闻网站会因为某个突发热点事件, 访问量激增, 此时, 若检测到该社交平台或新闻网站的后台运行主机 CPU 使用率上升到 90%, 则增加一台云主机实例, 以确保满足业务计算需求, 在 CPU 使用率下降到 30%时, 则移除一个实例, 减少资源成本的浪费。

根据以上业务场景，可以配置两条告警策略：

- 当 CPU 使用率大于 90%且连续出现 2 次时，增加 1 台云主机实例。



创建伸缩策略

名称: as-policy-c0a5

策略类型: **告警策略** 定时策略 周期策略

告警规则: **现在创建**

* 告警规则名称: as-alarm-c0a5

* 触发条件: CPU使用率 平均值 ≥ 90 %

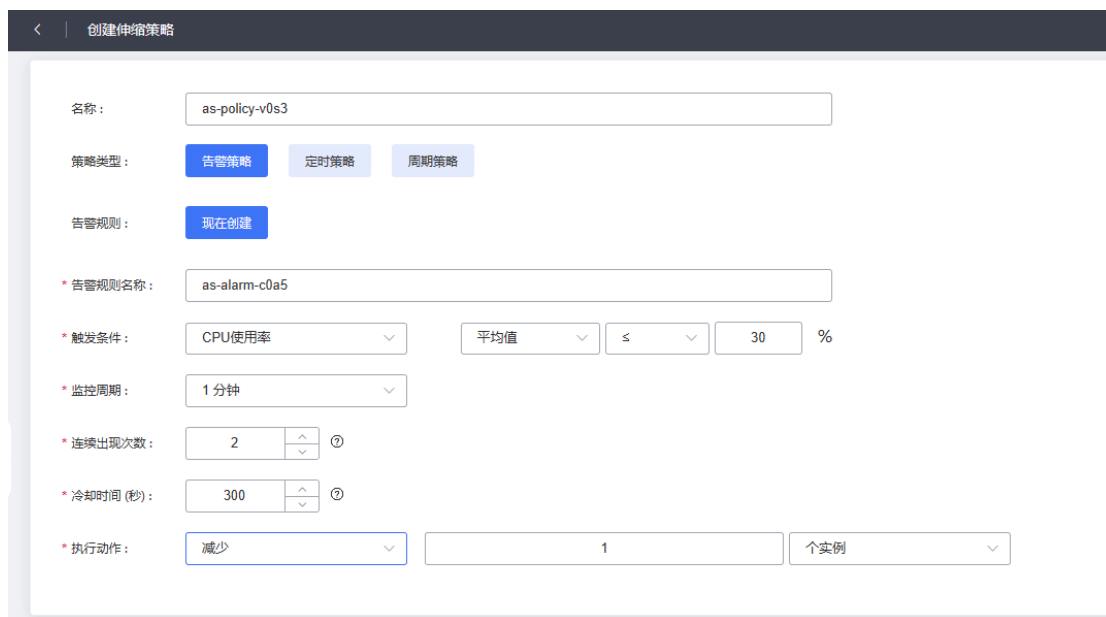
* 监控周期: 1分钟

* 连续出现次数: 2

* 冷却时间 (秒): 300

* 执行动作: 增加 1 个实例

- 当 CPU 使用率小于 30%且连续出现 2 次时，减少 1 台云主机实例。



创建伸缩策略

名称: as-policy-v0s3

策略类型: **告警策略** 定时策略 周期策略

告警规则: **现在创建**

* 告警规则名称: as-alarm-c0a5

* 触发条件: CPU使用率 平均值 ≤ 30 %

* 监控周期: 1分钟

* 连续出现次数: 2

* 冷却时间 (秒): 300

* 执行动作: 减少 1 个实例

按计划扩展资源

应对需求有规律变化的场景，可按照规律定期或者周期性的进行伸缩活动，通过设置定时策略或周期策略来动态调整资源。此类业务可以为其预设伸缩计划，设置定时策略或周期策略，

使其按计划扩展资源。

示例：直播平台或者在线教育网站每天 20:00-22:00 是业务高峰期，且是周期性规律的业务形态。20:00 以后业务访问量开始突增，22:00 以后，业务访问量突减。

可以为此示例配置两条周期策略：

- 未来一个月每天 19:55 自动增加 3 台实例。

名称：

策略类型：

周期：

时区：GMT +08:00

* 触发时间：

* 生效时间：

* 执行动作：

- 未来一个月每天 22:05 自动减少 3 台实例。

名称：

策略类型：

周期：

时区：GMT +08:00

* 触发时间：

* 生效时间：

* 执行动作：

伸缩组内实例管理

实例管理概述

在弹性伸缩组内，您可以针对组内实例做很多操作。例如手动将实例移入伸缩组或者手动将实例移出伸缩组，手动调整实例状态，将其设置为保护中状态等。

下表为实例的主要状态说明：

状态	状态属性	状态说明
服务中	稳定状态	当前伸缩组下正在承载业务，且稳定服务的云主机实例。
正在加入	中间状态	当前伸缩组下通过自动伸缩或手动操作而加入的云主机实例。
正在移出	中间状态	当前伸缩组下按照实例移除策略或手动操作而移出的云主机实例。

当实例被添加至伸缩组后，可能会有不同的状态，这些状态组成了实例的生命周期，弹性云主机加入伸缩组的方式共有两种，自动创建的实例与手动创建的实例，对应的其生命周期也会略有不同。

实例加入类型	加入方式	生命周期管理
自动创建的实例	通过为伸缩组预设的伸缩配置自动创建的弹性云主机实例，通常包括伸缩策略自动添加等。	此类实例的生命周期是完全跟随弹性伸缩活动进行的，当弹性扩容时，将会自动创建添加实例，当弹性缩容时，将会自动释放移除实例。
手动创建的实例	用户手动创建弹性云主机实例，然后再将其手动移入弹性伸缩组中。	会在弹性缩容时自动释放移除云主机实例。

在这两种生命周期外，用户还可以选择为伸缩组内实例设置保护模式，若设置保护模式，云

主机实例将不会被自动释放与移除。

将实例移入伸缩组


操作场景

您可以为伸缩组手动添加云主机实例。若伸缩组绑定了负载均衡，则新加入的云主机实例，也将加入到伸缩组内的负载均衡后端云主机组，来承载业务流量请求。



约束与限制

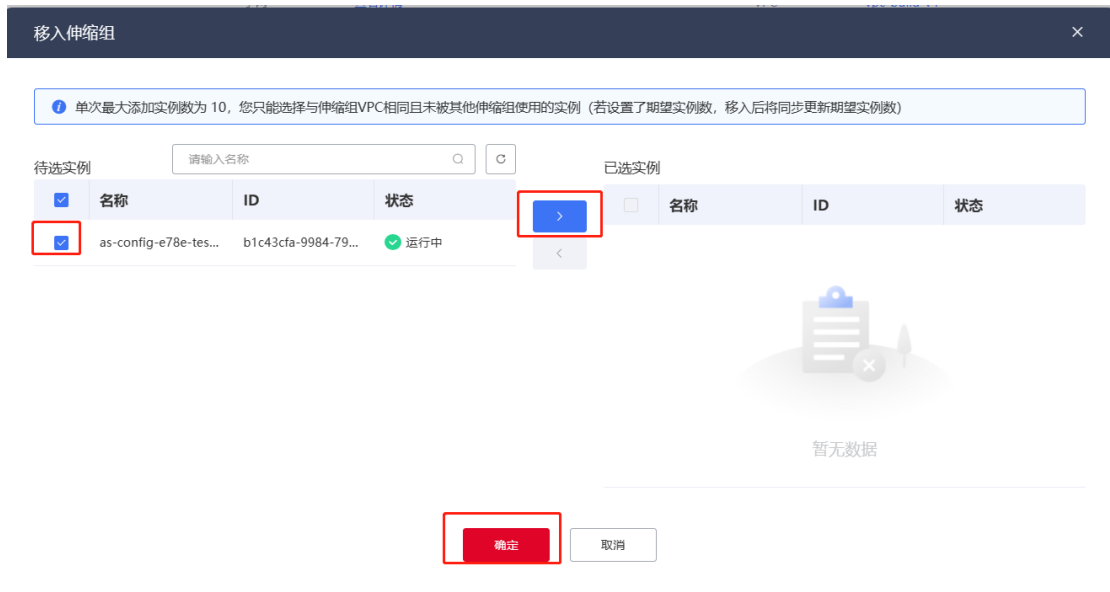
- 伸缩组处于“已启用”状态。
- 将被移入的云主机实例所在虚拟私有云必须与伸缩组所在虚拟私有云一致。
- 云主机不能已处于其他伸缩组内。
- 移入的云主机必须处于“运行中”的健康状态。
- 移入后实例数不能超过伸缩组实例上限。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待添加实例的伸缩组名称，进入到伸缩组详情页面。
5. 在伸缩组详情页面的下方，单击“伸缩实例”进入伸缩实例列表页签，单击“移入伸缩组”按钮。



- 在“移入伸缩组”弹窗中，在“待选实例”列表中勾选将要添加至此伸缩组内的实例，并单击  按钮，将此云主机实例添加至“已选实例”列表中，并单击“确定”。若您不想继续添加，可在“已选实例”列表中勾选目标实例，单击  按钮，将其弃选。



- 确定选择后，回到“伸缩实例”页签中，可查看到已移入的实例。此时，说明已成功将云主机实例移入至伸缩组中。

将实例移出伸缩组

操作场景

当您需要更新实例或者排查实例的问题时，可以选择将实例移出伸缩组。若伸缩组绑定了负

载均衡，当实例移出伸缩组后，实例将不再承担业务流量请求。


将实例移出伸缩组有两种情况，移出伸缩组与移出伸缩组并释放。

- 手动移出实例不会影响实例本身的性状，实例不会被释放、删除、关闭，您还可以用此云主机实例继续做其他操作。
- 当选择移出伸缩组并释放时，云主机实例在被移出伸缩组后还会被释放掉，但此功能对手动移入的云主机实例不生效。

约束与限制

- 伸缩实例为“已启用”状态。
- 伸缩组没有正在执行的伸缩活动。
- 移出后实例数不能低于伸缩组实例下限，即最小实例数。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待移出实例的伸缩组名称，进入到伸缩组详情页面。
5. 在伸缩组详情页面的下方，单击“伸缩实例”进入伸缩实例列表页签，在待移出实例所在行，单击“操作”列的“移出伸缩组”或“移出伸缩组并释放”。



云主机名称/ID	状态	实例加入类型	伸缩配置	健康检查状态	实例保护	加入时间	操作
openapi-bare 0b7c4017-02ca-a25a-0ec3-ef...	服务中	手动移入	---	健康	未保护	2022-11-24 15:19:42	移出伸缩组 移出伸缩组并释放
vm-ipv6-v4 bcf78349-5798-c37b-51f5-8e...	服务中	手动移入	---	健康	未保护	2022-11-24 15:19:42	移出伸缩组 移出伸缩组并释放

6. 若您想批量操作，您可以勾选待操作的云主机实例名称前的复选框，在实例列表上方单击“移出伸缩组”或“移出伸缩组并释放”按钮。



云主机名称/ID	状态	实例加入类型	伸缩配置	健康检查状态	实例保护	加入时间	操作
openapi-bare 0b7c4017-02ca-a25a-0ec3-ef...	服务中	手动移入	--	健康	未保护	2022-11-24 15:19:42	移出伸缩组 移出伸缩组并释放
vm-ipv6-v4 bcf78349-5798-c37b-51f5-8e...	服务中	手动移入	--	健康	未保护	2022-11-24 15:19:42	移出伸缩组 移出伸缩组并释放

7. 在弹出的“移出伸缩组”与“移出伸缩组并释放”弹窗中，确认移出实例信息，确认无误后，点击“确定”按钮，即可成功将实例移出伸缩组。

注意：

手动移入的云主机实例无法进行“移出伸缩组并释放”操作。

实例保护

操作场景

当您不希望伸缩组内的云主机实例被自动移出时，可以为其设置保护状态，设置了保护状态的云主机实例，在伸缩组进行缩容活动时将不会被移出。

约束与限制

在以下场景中，即使对特定实例开启了保护也无法确保其不被移出。

- 健康检查未通过的“异常”实例将会被移出，实例保护无法保护“异常”实例。
- 手动移出实例时将不会被保护状态限制。
- 保护状态需要手动调整，所有新添加的云主机实例都默认为不保护状态。

操作步骤

设置保护状态

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。

3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待设置保护状态实例所在的伸缩组名称，进入到伸缩组详情页面。
5. 在伸缩组详情页面的下方，单击“伸缩实例”进入伸缩实例列表页签，勾选待设置保护状态实例名称前的复选框，单击列表上方的“更多”按钮，在下拉列表中选择“开启实例保护”。



6. 在“开启实例保护”弹窗中，确认实例信息，确认无误后，点击“确定”，回到实例列表，可以看到“实例保护”列的属性值变为了“已保护”，此时说明已成功为实例开启了保护状态。



关闭实例保护

1. 登录控制中心。
2. 单击控制中心左上角的📍，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待设置保护状态实例所在的伸缩组名称，进入到伸缩组详情页面。
5. 在伸缩组详情页面的下方，单击“伸缩实例”进入伸缩实例列表页签，勾选待关闭保护状态实例名称前的复选框，单击列表上方的“更多”按钮，在下拉列表中选择

“关闭实例保护”。



- 在“关闭实例保护”弹窗中，确认实例信息，确认无误后，点击“确定”，回到实例列表，可以看到“实例保护”列的属性值变为了“未保护”，此时说明已成功为实例关闭了保护状态。



伸缩组监控

监控指标说明

弹性伸缩当前所支持的监控指标及其说明如下表所示：

指标	指标含义	计算公式	取值范围	测量对象	监控周期 (原始指标)
CPU 使用率	统计弹性伸缩组内实例的 CPU 使用率平均值	所有实例 CPU 使用率之和 / 实例数	$\geq 0\%$	弹性伸缩组	5min
内存使	统计弹性伸缩组内	所有实例内存使	$\geq 0\%$	弹性伸	5min

指标	指标含义	计算公式	取值范围	测量对象	监控周期 (原始指标)
用率	实例的内存使用率平均值	用率之和/实例数		缩组	
网络流入速率	统计弹性伸缩组内实例的每秒的流入网络流量	所有实例网络流入速率之和/实例数	$\geq 0\text{Kbps}$	弹性伸缩组	5min
网络流出速率	统计弹性伸缩组内实例的每秒的流出网络流量	所有实例网络流出速率之和/实例数	$\geq 0\text{Kbps}$	弹性伸缩组	5min
磁盘读速率	统计从弹性伸缩组内实例的每秒读出的数据量	所有实例磁盘读速率之和/实例数	$\geq 0\text{Kbps}$	弹性伸缩组	5min
磁盘写速率	统计从弹性伸缩组内实例的每秒写入的数据量	所有实例磁盘写速率之和/实例数	$\geq 0\text{Kbps}$	弹性伸缩组	5min
磁盘读请求速率	统计从弹性伸缩组内实例的每秒读取数据的请求次数	所有实例磁盘读请求速率之和/实例数	$\geq 0\text{IOPS}$	弹性伸缩组	5min
磁盘写请求速率	统计从弹性伸缩组内实例的每秒写入数据的请求次数	所有实例磁盘写请求速率之和/实例数	$\geq 0\text{IOPS}$	弹性伸缩组	5min

查看监控指标数据

操作场景

将弹性云主机实例加入伸缩组后，云监控会自动为其开启监控，用户可以通过云监控更好地

掌握伸缩组内实例运行状况。

前提条件


- 当伸缩组内没有实例时，即实例数为 0 时，用户是无法在监控中查看到任何监控性能变化的。
- 当伸缩组内云主机实例处在关机、故障、删除状态时，您无法查看其 CPU 使用率、内存使用率等性能指标，必须确保弹性伸缩组内的弹性云主机处于正常运行状态。

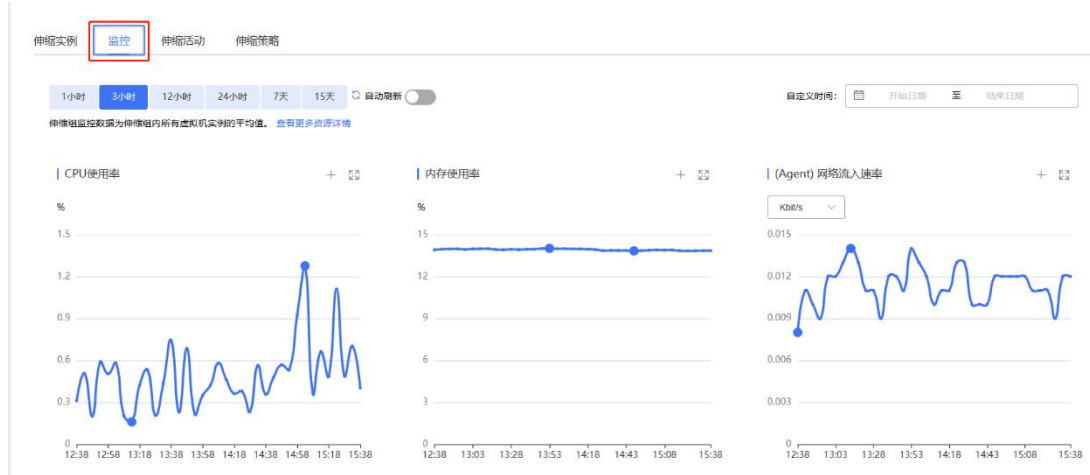
操作步骤

查看监控指标数据您可以从两个入口进入：


- 在弹性伸缩组页面查看。
- 在云监控页面查看。

在弹性伸缩组页面查看

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击待查看实例监控指标数据的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“监控”进入监控页签，即可查看伸缩组内实例的监控指标数据。

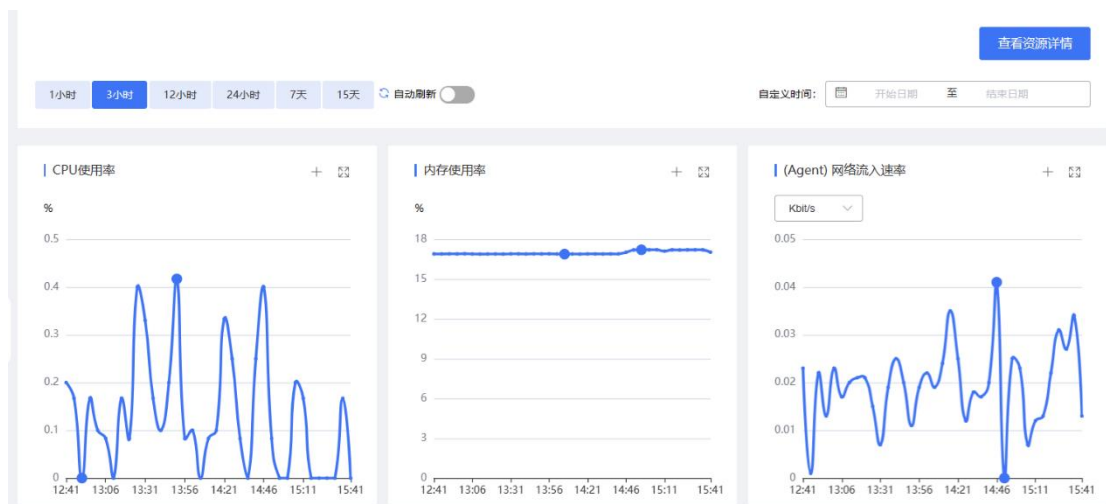



在云监控页面查看

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“管理与部署>云监控”，进入云监控管理控制台。
4. 在左侧的导航栏中单击“云服务监控>弹性伸缩监控”，进入到“弹性伸缩监控”页面。

名称	状态	ID	企业项目	操作
...	已启用	...	default	查看监控图表 创建告警规则
...	已启用	...	default	查看监控图表 创建告警规则
...	已启用	...	default	查看监控图表 创建告警规则

5. 进入到“弹性伸缩监控”页面后，可以查看到当前已创建的弹性伸缩组，在待查看监控性能数据的伸缩组所在行，单击“操作>查看监控图表”，即可进入至监控指标数据详情页面。



您可以在监控指标数据详情页面查看伸缩组内当前实例“近 1 小时”、“近 3 小时”、“近 12 小时”、“近 24 小时”、“近 7 天”“近 15 天”的监控性能数据。如果您想查看更长时间范围的监控曲线，请在监控视图中单击图标 ，进入大图模式查看。

弹性伸缩健康检查

弹性伸缩服务支持健康检查功能，保证使用运行状态健康的云主机负载业务。

弹性伸缩健康检查主要包含两种健康检查方式：云主机健康检查和弹性负载均衡健康检查，使用这两种检查方式自动替换不健康实例，避免业务中断，降低服务抖动。

- 云主机健康检查：此类检查指的是通过对云主机实例的健康检查，将处于“异常”状态的云主机移出伸缩组，例如处于“错误”等状态的云主机将会被移出，同时伸缩组将会创建新的健康实例来替代被移出伸缩组的实例，确保满足伸缩组实例数量配置要求。此方式为弹性伸缩默认检查方式。
- 弹性负载均衡健康检查：此类检查方式仅当您启用负载均衡器时才可启用。当您开启负载均衡健康检查时，您可以根据弹性负载均衡对伸缩组内云主机实例内服务的端口的状态进行检查。如果您在一个伸缩组内启用了多台负载均衡器，则只有在所

有负载均衡器均检测到云主机实例状态为健康的情况下，才会认为该云主机实例正常。只要有一个负载均衡器检测出云主机端口的服务状态是异常的，伸缩组都会将其移出。

以上两种检查方式在检测出异常云主机实例后，都会将其进行移出，但是移出之后是否删除取决于此云主机实例的加入方式与付费方式：

- 若此云主机实例是按需计费，则自动添加的实例在移出后会被释放，而手动添加进来的云主机实例在移出后将会继续保留。
- 若此云主机实例是包年包月的周期计费，则在移出后将不会被释放。

注意：

若伸缩组处于停用状态，健康检查将会继续执行，但不会触发任何伸缩活动，即不会将云主机实例进行移出。

设置监控告警规则

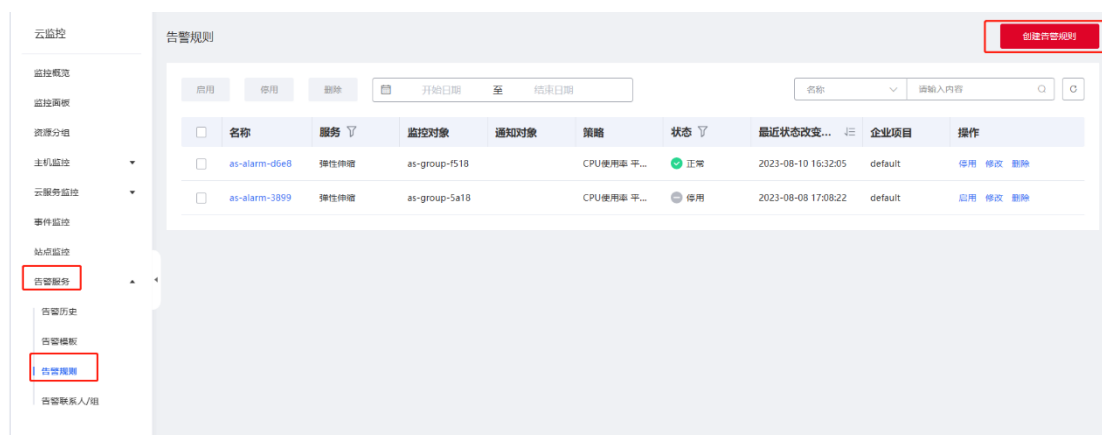
操作场景

您可以在云监控中为弹性伸缩服务的弹性伸缩组设置告警规则，通过设置告警规则来实时监控伸缩组内的实例运行状况。通过告警可以及时进行人工介入以防止出现主机崩溃等严重故障，确保业务稳定运行。

操作步骤

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“管理与部署>云监控”，进入云监控管理控制台。

4. 在左侧的导航栏中单击“告警服务>告警规则”，进入到“告警规则”页面。单击右上角的“创建告警规则”按钮，进入到“创建告警规则”页面。



5. 在“选择监控对象”部分中，“规则类型”选择“指标监控”，“云服务”选项中下拉选择“弹性伸缩”，“维度”选项中下拉选择“弹性伸缩”，“监控对象类型”选择“具体实例”，“监控对象”下拉选择已存在的弹性伸缩组。



6. 在“选择监控指标”部分中，根据业务实际需求来进行配置，具体配置可参照云监控>告警规则。
7. 在“规则信息”中，为此规则设置名称、企业项目、描述，设置完成点击“确定”，即可成功创建告警规则。

更多关于设置告警规则的信息，请参见云监控服务帮助中心。

最佳实践

利用弹性伸缩为突发业务创建低成本业务集群

场景描述

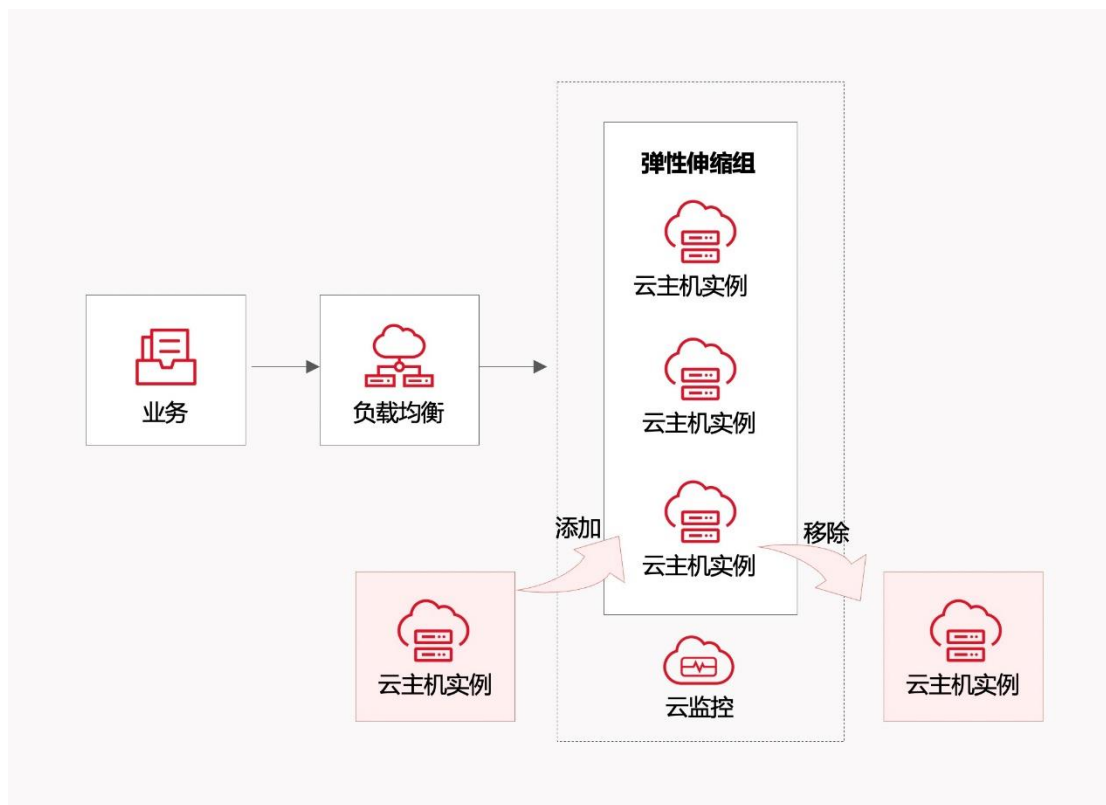
用户业务的访问量或流量波动明显，在业务高峰期业务流量较大，其他时段业务流量较低。

非高峰期的计算资源处于闲置状态，导致成本浪费，此类波动通常是无规律的，例如社交平台或者新闻媒体因为某个爆点新闻导致的突发流量。

解决方案

可以使用弹性伸缩创建低成本业务集群，根据业务流量自动扩缩容业务集群内的计算资源，减少成本浪费。方案详解如下：

- 针对日常业务流量，购买包年包月云主机实例，并将其加入弹性伸缩组做日常监控。
- 为伸缩组创建告警策略来应对突发的流量变化（突增或突减）。



前提条件

在做本实践之前，请确保您已经注册了天翼云账号，并确保您的账户中有充足的余额，具体步骤请参见[准备工作](#)。


操作步骤

为保证云主机实例提供稳定的计算服务，计划通过弹性伸缩自动调节伸缩组内云主机数量，将伸缩组内云主机 CPU 利用率维持在 50%左右。具体操作步骤如下：

- 步骤一：创建包年包月弹性云主机实例
- 步骤二：创建伸缩组并添加云主机实例
- 步骤三：根据业务需求创建伸缩策略（告警策略）

创建包年包月弹性云主机实例

为本实践的业务集群创建能够满足日常业务需求的包年包月低成本云主机实例。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性云主机”，进入弹性云主机管理控制台。
4. 在右上角点击“创建云主机”按钮，进入到弹性云主机创建页面。



5. 在弹性云主机创建页面，完成云主机参数配置，具体参数说明可参见创建云主机。

创建成功之后如图：

云主机列表

您还可以创建 9987 台云主机，使用 487 核CPU和 409585 GB内存。 [了解全部详情](#)

开机 关机 重启 更多 全部操作

筛选标签 请输入名称和ID(多个用#号分隔)

实例/主机名称	镜像	状态	标签	可...	企业项目	IPv4地址	IPv6地址	CP...	规格	付费方式/创建时间	到期时间	操作	
ecm-2cc1 ecm-2cc1	centos8...	运行中			华东1-可...	default	10.0.0.6(内)	--	X86计算	s2.small.1 1核 1G	按量付费 2023-08-07 18:46:24	--	
ecm-4b16 ecm-4b16	CentOS...	运行中			华东1-可...	default	10.0.0.4(内)	--	X86计算	s7.medium.2 1核 2G	包年包月 2023-07-27 20:22:25	2024-01-27 20:26:34	
ecm-4b16 ecm-4b16	CentOS...	运行中			华东1-可...	default	10.0.0.3(内)	--	X86计算	s7.small.1 1核 1G	包年包月 2023-07-27 20:13:17	2024-01-27 20:25:21	


10条/页 共 13 条 < 1 2 >

注意

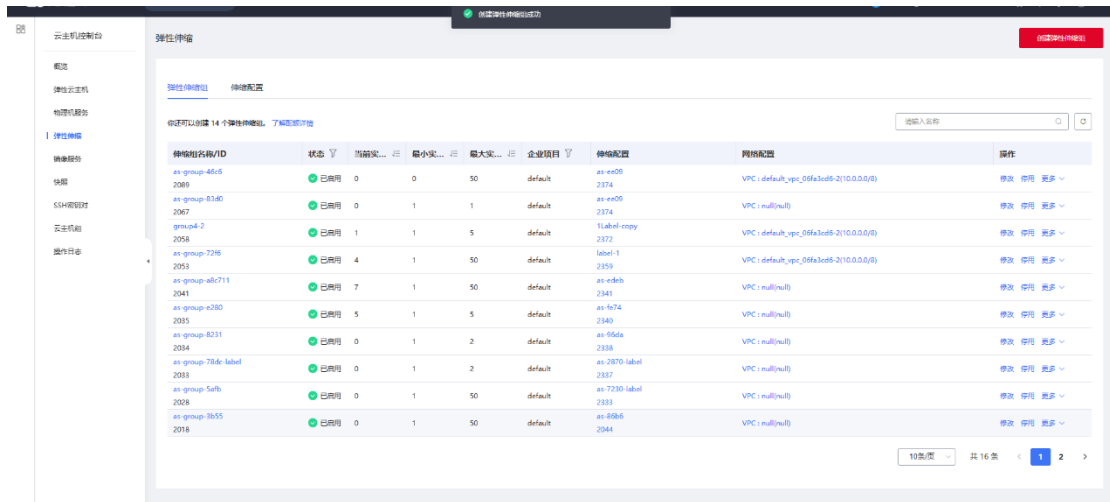
付费模式请设置“包年包月”。

创建伸缩组并添加云主机实例

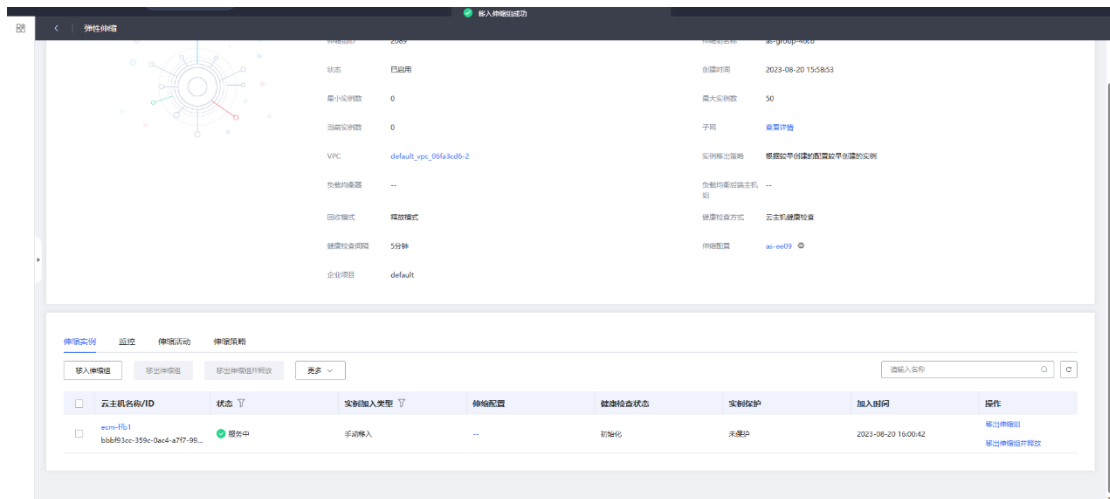
为低成本业务集群创建伸缩组，并将创建好的云主机实例添加至此伸缩组内，来满足业务的日常访问需求。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中，点击右上角“创建弹性伸缩组”，进入“创建弹性伸缩组>伸缩参数”配置页面。
5. 在配置页面完成伸缩组的配置，具体参数配置可参见[创建伸缩组](#)。以下为本实践中专有的属性值设置说明：
 - 伸缩组的虚拟私有云（VPC）需要与已创建的弹性云主机保持一致。
 - 伸缩组的镜像需要与已创建的弹性云主机保持一致。
 - 最小实例数设置为 0。
 - 根据业务需要，绑定业务模块使用的负载均衡器。
 - 实例移除策略选择较晚创建的实例。
 - 伸缩配置选择现在创建，根据业务需要选择云主机实例的规格和存储容量。

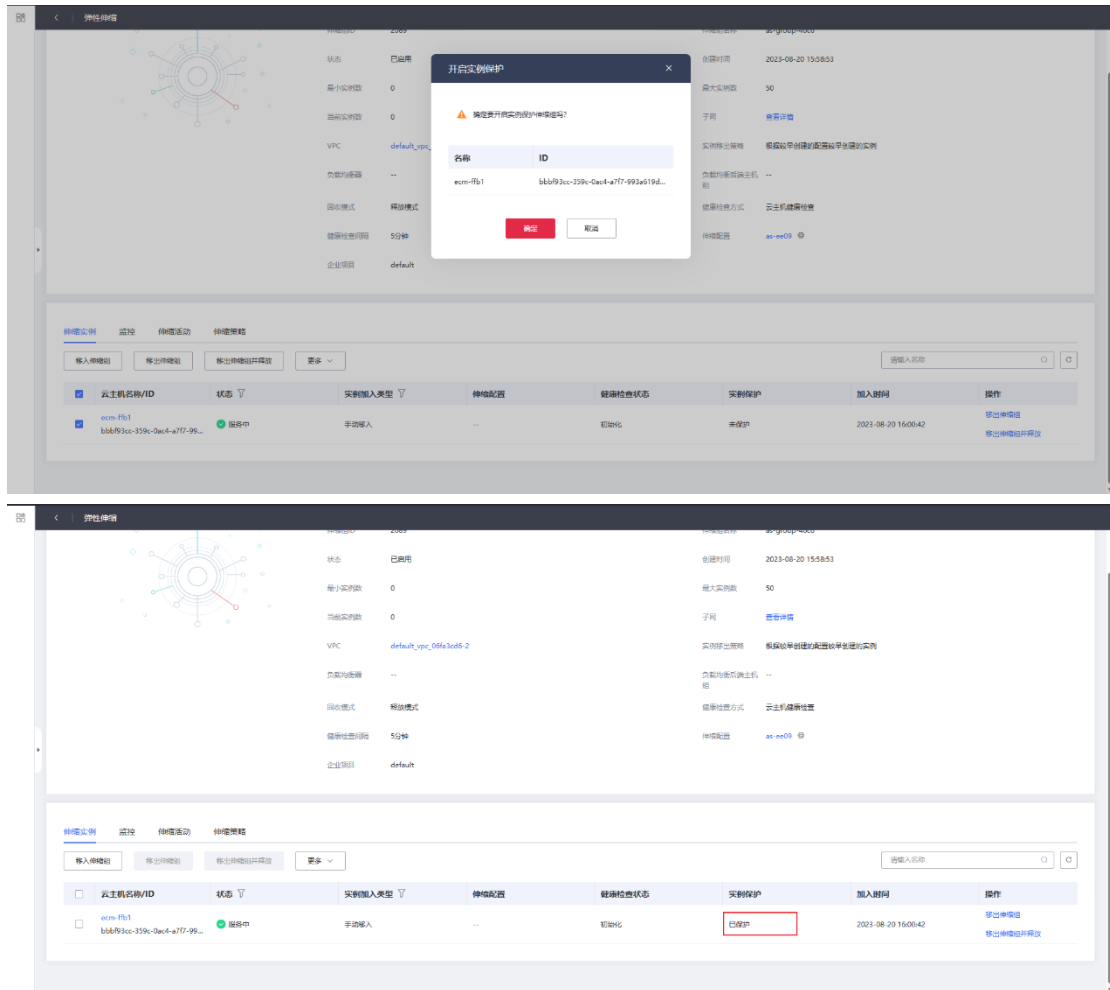
- 根据业务需要配置弹性 IP 和密码。
6. 在完成伸缩组与伸缩配置的设置后，点击“立即创建”，即可创建伸缩组成功，创建成功如图所示：



7. 伸缩组创建成功之后，进入到伸缩组详情页面，将步骤一中创建包年包月弹性云主机添加至伸缩组内，用于日常业务需求的满足。具体操作步骤请参见将实例移入伸缩组。添加成功之后如图：




8. 将实例移入伸缩组后，为防止后续实例移除时将此实例移除，为此云主机实例开启保护状态，具体操作步骤请参见实例保护。



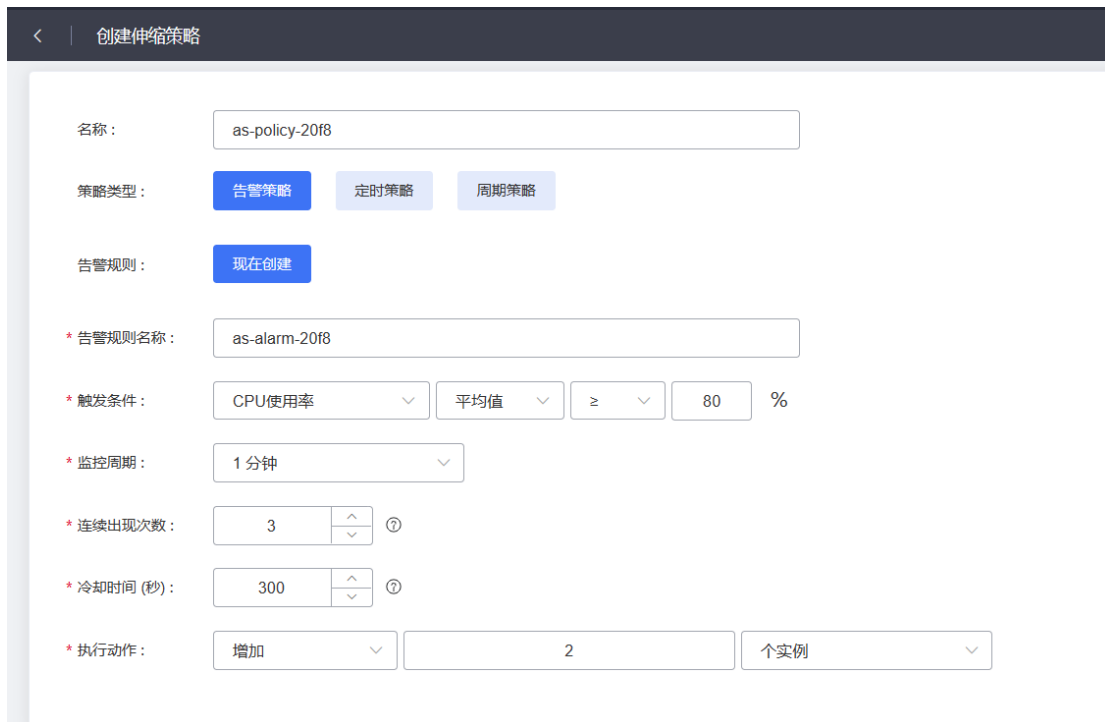
根据业务需求创建伸缩策略（告警策略）

为已创建好的伸缩组设置伸缩策略，来满足突发的业务流量变化，确保业务的稳定运行。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击步骤二创建好的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击“创建策略”按钮，进入到“创建伸缩策略”页面。
6. 在“创建伸缩策略”页面中完成策略的配置，在本实践中需要添加两个告警策略，一个扩容策略，一个缩容策略。具体操作步骤可参见[创建伸缩策略](#)。以下为本实践

中扩容策略专有的属性值设置说明：

- 策略类型选择告警策略。
- 告警规则选择现在创建。
- 触发条件设置为 CPU 使用率平均值 \geq 80%。
- 监控周期设置为 1 分钟。
- 连续出现次数设置为 3 次。
- 冷却时间设置为 300 秒。
- 执行动作设置为增加 2 个实例。



创建伸缩策略

名称：

策略类型： 告警策略 定时策略 周期策略

告警规则：

* 告警规则名称：

* 触发条件： %

* 监控周期：

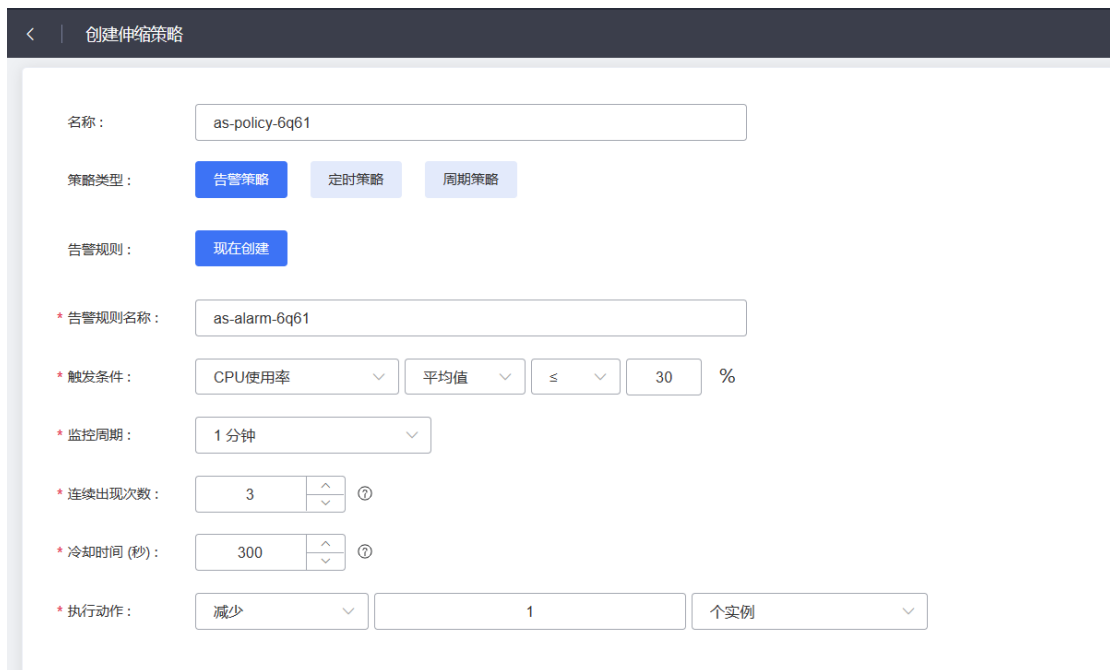
* 连续出现次数：

* 冷却时间(秒)：

* 执行动作：

7. 完成以上配置之后单击“确认”，即可成功创建扩容策略。
8. 继续创建缩容策略，以下为本实践中缩容策略专有的属性值设置说明：
 - 策略类型选择告警策略。
 - 告警规则选择现在创建。
 - 触发条件设置为 CPU 使用率平均值 \leq 30%。

- 监控周期设置为 1 分钟。
- 连续出现次数设置为 3 次。
- 冷却时间设置为 300 秒。
- 执行动作设置为减少 1 个实例。



名称：

策略类型： 告警策略 定时策略 周期策略

告警规则：

* 告警规则名称：

* 触发条件： %

* 监控周期：

* 连续出现次数：

* 冷却时间 (秒)：

* 执行动作：

9. 完成以上配置之后单击“确认”，即可成功创建缩容策略。

执行结果

- 伸缩组中手动移入的包年包月云主机实例已设置为“已保护”状态，处于已保护状态的云主机不会被伸缩策略自动移出伸缩组，保证业务集群非高峰时段的正常运行。
- 在高峰时段，即当伸缩组实例 CPU 使用率的平均值 $\geq 80\%$ 时，伸缩组将根据伸缩策略设置的扩容策略，自动增加云主机实例，确保业务在高峰时期也可以稳定运行。
- 在低谷时段，即当伸缩组实例 CPU 使用率的平均值 $\leq 30\%$ 时，伸缩组将根据伸缩策略设置的缩容策略，自动减少云主机实例，降低计算资源闲置的时间，减少成本浪费。

利用弹性伸缩为周期性业务降低成本

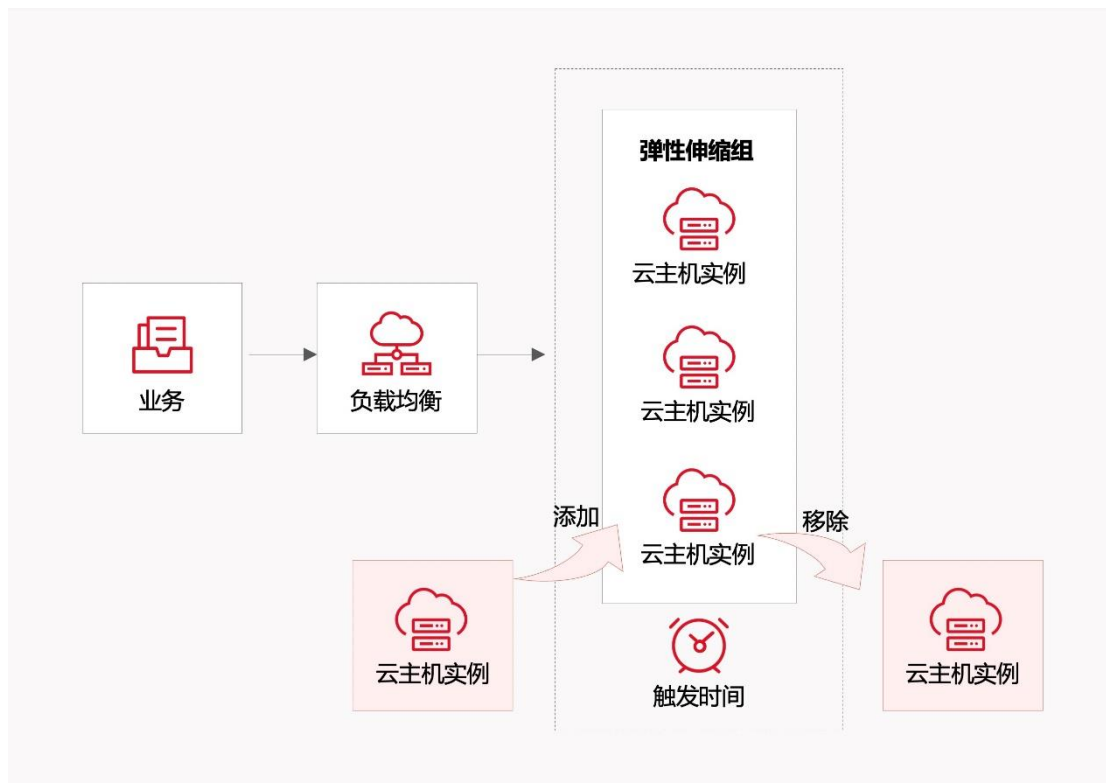
场景描述

此类场景的业务流量波动通常是有规律的周期性波动，在特定时间内业务流量较大，高峰期一过业务流量降低直到恢复正常。例如学习平台、直播平台等在 20:00 前后用户业务的访问量处于峰值，而 22:00 之后的流量会降低回到正常值。

解决方案

可以使用弹性伸缩创建低成本业务集群，根据业务流量的波动周期来自动扩缩容业务集群内的计算资源，减少成本浪费。方案详解如下：

- 针对日常业务流量，购买包年包月云主机实例，并将其加入弹性伸缩组做以日常监控，为其设置保护状态，避免被移除影响日常计算需求。
- 为伸缩组创建定时/周期策略来应对周期性的流量波动。



前提条件

在做本实践之前，请确保您已经注册了天翼云账号，并确保您的账户中有充足的余额，具体步骤请参见[准备工作](#)。


操作步骤

本实践共分为三大步：

- 步骤一：创建包年包月弹性云主机实例
- 步骤二：创建伸缩组并添加云主机实例
- 步骤三：根据业务需求创建伸缩策略（周期策略）

创建包年包月弹性云主机实例

为本实践的业务集群创建能够满足日常业务需求的包年包月低成本云主机实例。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性云主机”，进入弹性云主机管理控制台。
4. 在右上角点击“创建云主机”按钮，进入到弹性云主机创建页面。



5. 在弹性云主机创建页面，完成云主机参数配置，具体参数说明可参见[创建云主机](#)。

创建成功之后如图：

云主机列表

您还可以创建 9987 台云主机，使用 487 核vCPU和 409585 GB内存。 [了解全部详情](#)

开机 关机 重启 更多 全部操作

筛选标签 请输入名称和ID(多个ID请用逗号分隔)

实例/主机名称	镜像	状态	标签	可...	企业项目	IPv4地址	IPv6地址	CP...	规格	付费方式/创建时间	到期时间	描
ecm-2cc1 ecm-2cc1	centos8...	运行中		华东1-可...	default	10.0.0.6(内)	--	X86计算	s2.small.1 1核 1G	按量付费 2023-08-07 18:46:24	--	--
ecm-ffb1 ecm-ffb1	CentOS...	运行中		华东1-可...	default	10.0.0.4(内)	--	X86计算	s7.medium.2 1核 2G	包年包月 2023-07-27 20:22:25	2024-01-27 20:26:34	--
ecm-4b16 ecm-4b16	CentOS...	运行中		华东1-可...	default	10.0.0.3(内)	--	X86计算	s7.small.1 1核 1G	包年包月 2023-07-27 20:13:17	2024-01-27 20:25:21	--


10条/页 共 13 条 < 1 2 >

注意

付费模式请设置“包年包月”。

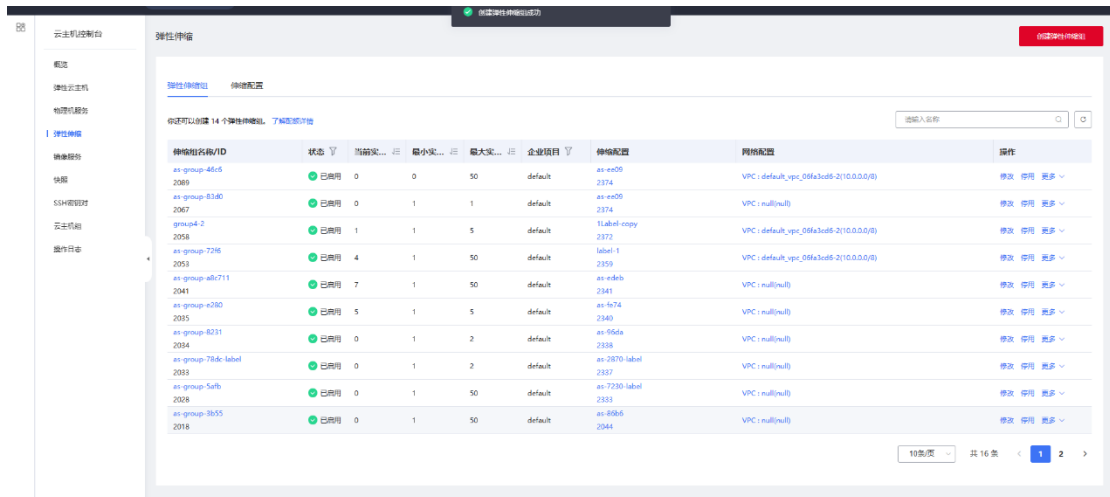
创建伸缩组并添加云主机实例

为低成本业务集群创建伸缩组，并将创建好的云主机实例添加至此伸缩组内，来满足业务的日常访问需求。

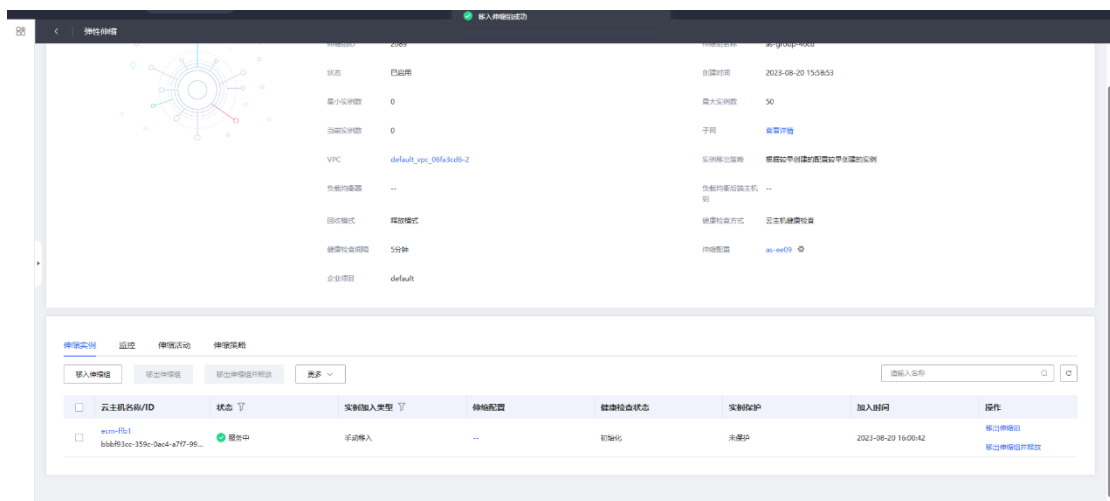
1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中，点击右上角“创建弹性伸缩组”，进入“创建弹性伸缩组>伸缩参数”配置页面。
5. 在配置页面完成伸缩组的配置，具体参数配置可参见[创建伸缩组](#)。以下为本实践中专有的属性值设置说明：
 - 伸缩组的虚拟私有云需要与已创建的弹性云主机保持一致。
 - 伸缩组的镜像需要与已创建的弹性云主机保持一致。
 - 最小实例数设置为 0。
 - 最大实例数可以设置为 10。
 - 根据业务需要，绑定业务模块使用的负载均衡器。
 - 实例移除策略选择较晚创建的实例。

- 伸缩配置选择现在创建，根据业务需要选择云主机实例的规格和存储容量。
 - 根据业务需要配置弹性 IP 和密码。
6. 在完成伸缩组与伸缩配置的设置后，点击“立即创建”，即可创建伸缩组成功，创建成功之后如图：

建成功之后如图：




7. 伸缩组创建成功之后，进入到伸缩组详情页面，将步骤一中创建包年包月弹性云主机添加至伸缩组内，用于日常业务需求的满足。具体操作步骤请参见[将实例移入伸缩组](#)。添加成功之后如图：

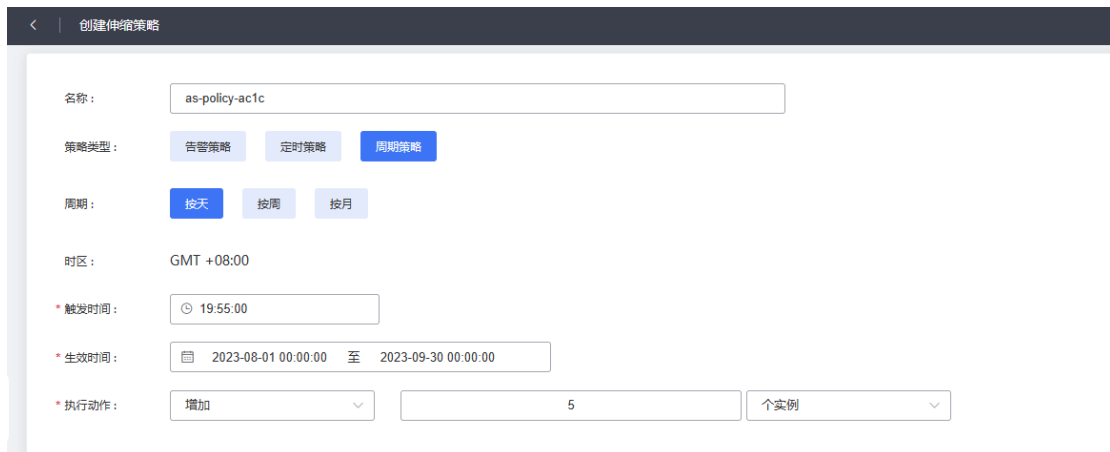


8. 将实例移入伸缩组后，为防止后续实例移除时将此实例移除，为此云主机实例开启保护状态，具体操作步骤请参见[实例保护](#)。

根据业务需求创建伸缩策略（周期策略）

为已创建好的伸缩组设置伸缩策略，来满足突发的业务流量变化，确保业务的稳定运行。

1. 登录控制中心。
2. 单击控制中心左上角的 ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 点击步骤 2 创建好的伸缩组名称，进入到伸缩组详情页面。
5. 在详情页面的下方，单击“伸缩策略”进入伸缩策略页签，单击“创建策略”按钮，进入到“创建伸缩策略”页面。
6. 在“创建伸缩策略”页面中完成策略的配置，在本实践中需要添加两个周期策略，一个扩容策略，一个缩容策略。具体操作步骤可参见[创建伸缩策略](#)。以下为本实践中扩容策略专有的属性值设置说明：
 - 策略类型选择周期策略。
 - 时区保持默认 GMT +8:00。
 - 触发时间设置为业务高峰期到来前 5 分钟，例如每天晚上 20:00 是高峰期起始时间，则周期设置为按天，8 月 1 日至 9 月 30 日期间，触发时间设置为每晚 19:55。
 - 执行动作设置为增加 5 个实例。
7. 完成以上配置之后单击“确认”，即可成功创建周期扩容策略。



创建伸缩策略

名称: as-policy-ac1c

策略类型: 告警策略 定时策略 周期策略

周期: 按天 按周 按月

时区: GMT +08:00

* 触发时间: 19:55:00

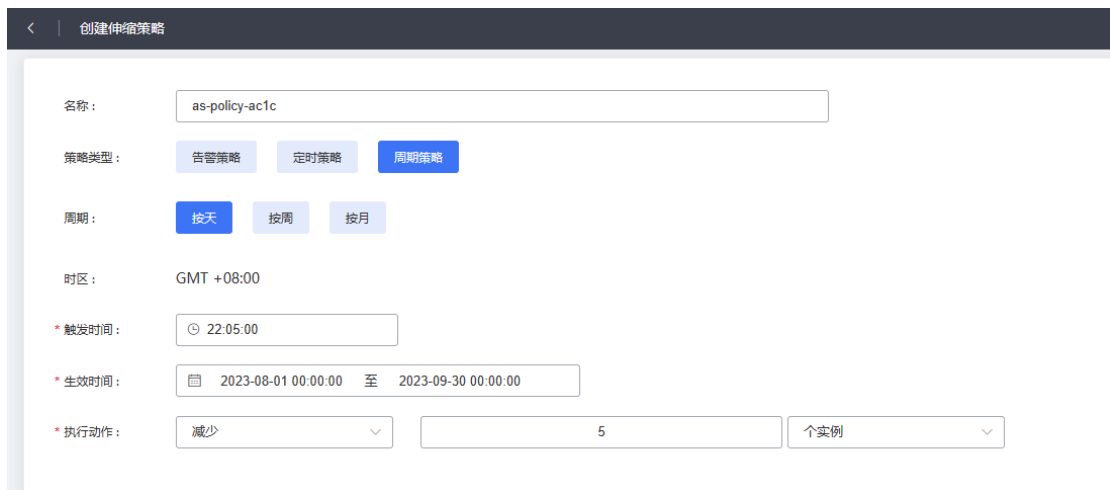
* 生效时间: 2023-08-01 00:00:00 至 2023-09-30 00:00:00

* 执行动作: 增加 5 个实例

8. 继续创建缩容策略，以下为本实践中缩容策略专有的属性值设置说明：

- 策略类型选择周期策略。
- 时区保持默认 GMT +8:00。
- 触发时间设置为业务高峰期结束后 5 分钟,例如每天晚上 22:00 是高峰期结束时间,则周期设置为按天,8 月 1 日至 9 月 30 日期间,触发时间设置为每晚 22:05。。
- 执行动作设置为减少 5 个实例。

9. 完成以上配置之后单击“确认”，即可成功创建缩容策略。



创建伸缩策略

名称: as-policy-ac1c

策略类型: 告警策略 定时策略 周期策略

周期: 按天 按周 按月

时区: GMT +08:00

* 触发时间: 22:05:00

* 生效时间: 2023-08-01 00:00:00 至 2023-09-30 00:00:00

* 执行动作: 减少 5 个实例

执行结果

- 伸缩组中手动移入的包年包月云主机实例已设置为“已保护”状态，处于已保护状态的云主机不会被伸缩策略自动移出伸缩组，保证业务集群非高峰时段的正常运行。
- 在高峰时段，即业务会在每天的 20:00 进入业务流量高访问时间段，伸缩组将根据伸缩策略设置的周期策略，自动增加云主机实例，确保业务在高峰时期也可以稳定运行。
- 在低谷时段，即业务在每天的 22:00 结束高流量访问阶段，逐渐回归正常，伸缩组将根据伸缩策略设置的周期策略，自动减少云主机实例，降低计算资源闲置的时间，减少成本浪费。

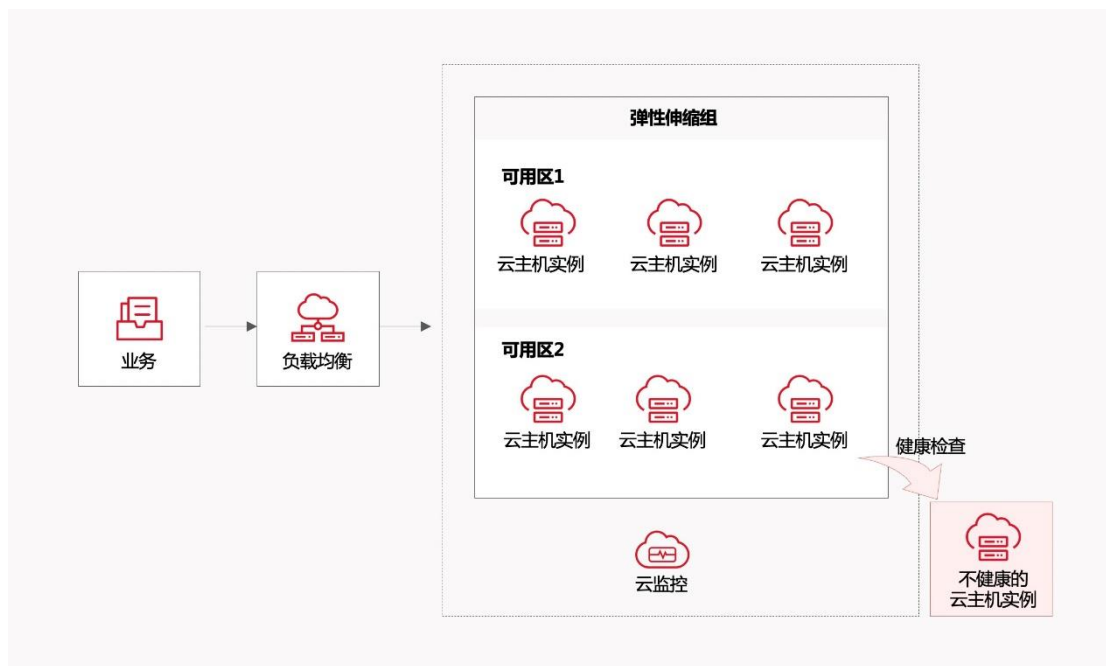
部署高可用集群

场景描述

在业务高峰期，业务系统会根据伸缩策略投放更多的计算资源，但是往往会面临资源不足、库存不足的问题，符合伸缩组伸缩配置的云主机实例是有限的，此时会导致扩容失败，严重的会导致业务中断，影响业务稳定运行。

解决方案

为业务系统设置弹性伸缩时部署高可用集群，设置伸缩组时可指定多个可用区，使得业务需求上升时，可在同一资源池下的多个可用区中伸缩，来降低电力、网络等可能出现的故障对整个系统稳定性的影响，以及库存不足导致的扩容失败等问题，来提升业务的容灾能力。




前提条件

在做本实践之前，请确保您已经注册了天翼云账号，并确保您的账户中有充足的余额，具体步骤请参见[准备工作](#)。

操作步骤

本实践主要操作重点为创建一个指定多可用区扩容的伸缩组，并为其创建伸缩配置。

1. 登录控制中心。
2. 单击控制中心左上角的  ，选择弹性伸缩组所在地域。
3. 单击“计算>弹性伸缩服务”，进入弹性伸缩管理控制台。
4. 在“弹性伸缩组”页签中，点击右上角“创建弹性伸缩组”，进入“创建弹性伸缩组>伸缩参数”配置页面。
5. 在弹性伸缩组创建页面，填写弹性伸缩组的基本信息，例如伸缩组名称、最大实例数、最小实例数、虚拟私有云、可用区、子网、负载均衡等信息，具体参数说明以及示例见下表：

参数	参数说明	本实践配置示例
名称	伸缩组的名称, 用户可自定义, 也可以选择保持系统默认分配名称。	as-group-test
最大实例数 (台)	当伸缩组的当前实例数大于最大实例数时, 弹性伸缩服务会自动移出实例, 使得伸缩组的当前实例数等于伸缩最大实例数。理论上实际实例数不允许大于最大实例数。	10
最小实例数 (台)	当伸缩组的当前实例数小于伸缩最小实例数时, 弹性伸缩服务会自动添加实例, 使得伸缩组的当前实例数等于伸缩最小实例数。理论上实际实例数不允许小于最小实例数。	0
期望实例数 (台)	当伸缩组内实例数不等于期望实例数时, 弹性伸缩服务会自动进行扩缩容, 确保伸缩组内始终保持该数量的实例数。	-
虚拟私有云	同一伸缩组内的弹性云主机的虚拟私有云是一致的。	vpc-72cd(192.168.0.0/16)
可用区	同一伸缩组内的弹性云主机可以选择多个可用区, 一个虚拟私有云子网内的弹性云服务器均	可用区 1 和可用区 2

参数	参数说明	本实践配置示例
	可以进行通信。一个伸缩组可以选择多个子网。	
安全组	安全组类似防火墙功能，是一个逻辑上的分组，用于设置网络访问控制。	Default-Security-Group
负载均衡	选择启用负载均衡后，伸缩组中的云主机会自动挂载到您关联的负载均衡下。访问流量将按照分发策略自动分发到伸缩组内的所有弹性云主机。注意：一个伸缩组可最多添加 10 组负载均衡监听器。当选择使用弹性负载均衡时，需配置 4 个参数：负载均衡器、后端云主机组、后端端口和权重。	不启用
实例移除策略	在对伸缩组进行缩容时，会采用此策略进行移除，共有以下四种方式： <ul style="list-style-type: none"> - 较早创建的配置且较早创建的实例：先筛选出较早创建的配置所创建出的实例，再筛选较早创建的实例移出。 - 较晚创建的配置且较晚创建的实例：先筛选出较晚创建的配置所创建出的实例，再筛选较晚创建的实例移出。 - 较早创建的实例（FIFO）：根据时间筛选较早创建的实例，跟是否是伸缩配置创建的无关。 - 较晚创建的实例（LIFO）：根据时间筛选较晚创建的实例，跟是否是伸缩配置创建的无关。 	较晚创建的实例
实例回收模式	指伸缩组自动创建的实例被移出后的处理策略。 释放模式：将移出的伸缩组自动创建的云主机释放。 注意：对于手动移入伸缩组的云主机，如果被移出伸缩组，资源不会被释放，状态保持不变，不适用此策略。	释放模式
健康检查方式	健康检查会将处于“异常”状态的云主机从伸缩组中移出，并替换同种规格的云主机来承载	云主机健康检查

参数	参数说明	本实践配置示例
	<p>业务流量。有以下两种方式：</p> <ul style="list-style-type: none"> - 云主机健康检查：是指对云主机的运行状态进行检查，如关机、删除都是云主机异常状态。 - 弹性负载均衡健康检查：是指根据负载均衡对云主机的健康检查结果进行的检查。所有监听器下检测到的云主机状态必须均为正常。 <p>注意：只有开启负载均衡时您才可以选择弹性负载均衡健康检查，建议当伸缩组开启负载均衡时，使用弹性负载均衡健康检查。</p>	
健康检查间隔	<p>执行健康检查的周期时间。您可以根据业务实际情况设置合理的健康检查间隔（5分钟、15分钟、1小时、3小时），以确保其包含主机中应用程序的预期启动时间。</p>	5分钟
企业项目	支持为伸缩组选择企业项目。	Default

* 名称：

* 最大实例数(台)： ①

* 最小实例数(台)： ②

期望实例数(台)：

* 虚拟私有云： [新建虚拟私有云](#)

* 可用区：

可用区	主网卡	辅助网卡	操作
<input type="text" value="可用区1"/>	<input type="text" value="subnet-72cd(192.168.0.0/24)"/>	<input type="text" value="请选择扩展网卡"/>	创建子网
<input type="text" value="可用区2"/>	<input type="text" value="subnet-72cd(192.168.0.0/24)"/>	<input type="text" value="请选择扩展网卡"/>	删除

* 安全组: [新建安全组](#) ?

* 负载均衡:

* 实例移出策略: ?

* 按量付费实例回收模式: ?

* 健康检查方式: ?

* 健康检查间隔:

* 企业项目: [新建](#) ?

6. 参数配置完成之后，请单击右下角的“下一步”，进入“创建弹性伸缩组 > 配置来源”页面，进行伸缩配置的创建与选择。

7. 在“创建弹性伸缩组 > 配置来源”页面中，“伸缩配置来源”选择“已有配置”，在已有伸缩配置列表中勾选业务需要的伸缩配置，单击“立即创建”，即可成功创建高可用伸缩组。

注意：

若需要新建伸缩配置，可参见[创建伸缩配置](#)。



伸缩组名称/ID	状态	当前实例数	期望实例数	最小实例数	最大实例数	企业项目	伸缩配置	网络配置	操作
as-group-test	已启用	0	--	1	10	default	as-test-5909	VPC : vpc-72cd(192.168.0.0/16)	修改 停用 更多

执行结果

启用伸缩组后，后续业务若遇到需要扩容的情况，当可用区 1 不可用时，将会在可用区 2 中进行扩容，不会再出现但可用区实例资源库存不足所导致的扩容失败等问题，在一定程度上保证了业务高可用性。

常见问题

计费类

弹性伸缩服务是否收取费用？

弹性伸缩服务本身不收取费用。但弹性伸缩组内的弹性云主机、云硬盘、弹性 IP 和带宽等云产品需按照相关产品计费页面公示信息收费。

注意：

如果您的账号欠费，伸缩组内所有按需计费的云主机实例将被冻结，从而导致弹性伸缩健康检查判定云主机实例状态为不健康，触发移出并释放这些云主机实例。因此为保证您的正常使用，请及时充值。

伸缩组停用后实例还收费吗？

弹性伸缩服务本身不收取费用，其中就包括创建的伸缩组，但是您需要为伸缩组内的弹性云主机实例付费。当伸缩组停用后，如果伸缩组内的弹性云主机还存在，则需要继续为弹性云主机实例、云硬盘、弹性 IP 资源付费。

伸缩组操作类

弹性伸缩服务支持数据盘的自动伸缩吗？

不支持。

弹性伸缩服务目前仅支持对弹性云主机数量的自动伸缩，没有细粒度到云硬盘级别，无法支持系统盘与数据盘的大小、数量的自动伸缩。

弹性伸缩一定要搭配弹性负载均衡、云监控才能使用吗？

不是。

弹性伸缩服务可以单独使用，也可以同弹性负载均衡、云监控一起使用。

- 云监控服务为免费服务，系统默认开通，开通之后可以监控伸缩组内实例的情况，结合伸缩策略触发伸缩活动，达到弹性伸缩的目的。
- 负载均衡服务在有需求时可以部署。例如，有分布式集群需求或者高可用需求的场景下，可以使用弹性负载均衡来分发流量给伸缩组内的云主机实例，共同承载业务请求。

单个伸缩组内可关联的最大实例数是多少？

100 台。

同时，伸缩组内可关联的实例数受到以下限制：

- 云主机配额：单个伸缩组内可关联的最大实例数为受到账户云主机配额限制，若无法满足业务需求，您可以选择提交工单提升配额。
- 负载均衡：若伸缩组关联了负载均衡器，则伸缩组可关联的实例数受到负载均衡支持关联的云主机数的限制。

已创建的实例能否被添加至弹性伸缩组内？

能。

弹性伸缩支持将已购买的云主机实例手动添加至伸缩组。但是在添加时需要满足以下条件：

- 待添加的弹性云主机实例尚未关联其他弹性伸缩组。

- 待添加的弹性云主机实例必须处于运行状态。
- 待添加的弹性云主机实例必须与伸缩组处于同一个 VPC 内。

已经创建的包年包月云主机实例可以被添加至弹性伸缩组吗？

可以。

当包年包月的云主机满足上一问题中的条件时，可以通过手动移入伸缩组的操作，将付费方式为包年包月的云主机实例添加至指定伸缩组。具体操作可参考帮助文档[将实例移入伸缩组](#)。

弹性云主机实例是否可以加入到多个伸缩组中？

不可以。

弹性云主机实例在同一时间只能被加入一个伸缩组中。针对此问题，有以下两种解决方案：

方案一：若您希望将云主机移入其他伸缩组，可现在原伸缩组中对云主机进行“移出伸缩组”操作后，再将该云主机移入新的伸缩组。

方案二：您可以再创建一台一样的云主机关联其他伸缩组，可通过云主机克隆功能创建一个新的云主机实例并移入新的伸缩组。

伸缩组如何自动移出实例？

用户可设置实例移出策略，实例移出策略共有以下四种方式：

- 较早创建的配置且较早创建的实例：先筛选出较早创建的配置所创建出的实例，再筛选较早创建的实例移出。
- 较晚创建的配置且较晚创建的实例：先筛选出较晚创建的配置所创建出的实例，再筛选较晚创建的实例移出。

- 较早创建的实例 (FIFO): 根据时间筛选较早创建的实例, 与是否是伸缩配置创建的无关。
- 较晚创建的实例 (LIFO): 根据时间筛选较晚创建的实例, 与是否是伸缩配置创建的无关。

伸缩组停用后, 伸缩组中的云主机实例会被自动释放吗?

不会。

停用伸缩组后, 伸缩组的状态会变为“已停用”。处于“已停用”状态的伸缩组, 不会触发任何弹性伸缩活动。同时, 伸缩组内的云主机实例也不会被自动释放。

伸缩组停用后, 会带来什么影响?

伸缩组处于“已停用”状态时, 伸缩组将不再通过伸缩策略触发伸缩活动, 但是正在执行中的伸缩活动会持续进行至完成。手动调整最大实例数、最小实例数或期望实例数, 不会触发伸缩活动。

伸缩配置操作类

创建伸缩配置时, 可以添加 8 核或者 16 核的云主机实例吗?

可以。

伸缩组内的云主机实例规格无限制。您可以根据您的业务需要, 选择不同架构、vCPU、内存的规格作为伸缩配置模板用于伸缩组扩容。创建伸缩配置时:

- 支持定义全新的云主机类型、规格、镜像作为伸缩组扩容云主机模板, 可在创建伸

缩配置时，选择“使用新规格”。

- 支持使用与某云主机实例一样的类型、规格、镜像作为伸缩组扩容云主机的模板，可在创建伸缩配置时，选择“使用现有云主机规格”。

创建伸缩配置时，用户可以指定数据盘的容量吗？

可以。

数据盘的类型、容量是创建伸缩配置时云主机实例主要的存储配置项，只需在创建时进行配置即可。

注意：

1. 数据盘数量、容量受配额限制，超出配额限制伸缩活动将无法正常进行。
2. 创建伸缩配置时，若您选择使用极速型 SSD 云硬盘作为数据盘，仅支持挂载至 vCPU 数量至少为 16 的通用计算增强型 c6/c7 云主机和内存优化型 m6/m7 云主机，并且一台云主机只允许挂载最多 3 块极速型 SSD 云硬盘。

伸缩策略操作类

用户是否可以设置周期性定时策略？

可以。

弹性伸缩服务支持设置三种策略：

- 告警策略：通过对伸缩组内实例性能指标（CPU 使用率、内存使用率等）的监控，来确认其是否到达预设的告警条件，来自动增加或减少云主机的数量。
- 定时策略：根据业务实际情况设置一个时间点，在此时间点自动增加或减少云主机

的数量。

- 周期策略：根据业务实际情况设置一个时间段，在此时间段内按照周期（按天、按周、按月）来重复执行自动增减云主机的数量。

具体配置操作请参见[创建伸缩策略](#)。

告警策略与定时/周期策略之间有执行的优先级吗？

没有。

告警策略与定时/周期策略之间是相互独立的策略，互不干扰。

告警策略若失败一次，并不影响其策略有效性，若继续触发告警规则，则会继续执行伸缩活动。

定时/周期策略若失败一次，您需重新检查策略有效性，重新设置触发时间与生效时间，避免策略失效无法被触发。

弹性伸缩中的告警策略可以使用的性能指标有哪些？

告警策略可以使用的性能指标主要有：CPU 使用率、内存使用率、网络流入速度、网络流出速度、磁盘读速率、磁盘写速率、磁盘读请求速率、磁盘写请求速率。您可以选择以上性能指标，并为其设置告警规则，确定增加或者减少 n 台云主机。

如何设置告警策略触发条件？

您可以在创建弹性伸缩策略时，选择策略类型为“告警策略”，根据业务实际情况选择触发条件与执行动作。您需要选择一个监控项作为指标，如 CPU 使用率、内存使用率等。您可以使用监控指标平均值、最大值或最小值作为告警策略触发条件。

- 平均值：针对伸缩组内所有云主机实例，如 CPU 使用率平均值大于 90%时触发告

警。

- 最大值：针对伸缩组内 CPU 使用率最高的云主机实例，CPU 使用率大于 90%时触发告警。
- 最小值：针对伸缩组内 CPU 使用率最低的云主机实例，CPU 使用率大于 90%时触发告警。

伸缩活动操作类

当弹性伸缩产品出现故障需要提交工单解决时，需要提交哪些信息？

您在提交工单时，若主要询问弹性伸缩产品的故障原因，可以备注弹性伸缩组所在地域 (Region)、弹性伸缩组 ID、伸缩活动 ID 及问题描述，便于快速定位问题。

弹性伸缩组执行伸缩活动添加云主机实例时，为什么会出现伸缩失败提示？

在执行伸缩活动添加云主机实例时出现的失败提示，通常为以下原因：

- 伸缩组配置异常：请检查伸缩组绑定的 VPC、负载均衡器、安全组是否存在。若不存在，请更新伸缩组配置。
- 伸缩配置异常：请检查伸缩配置使用的云主机规格、云硬盘类型、弹性 IP 是否已售罄，镜像、密钥对是否失效。若存在异常，请更换或修改伸缩配置。
- 账户异常：当您的账户余额小于 100 元或欠费时，可能无法通过弹性伸缩添加云主机，请您及时充值。
- 配额不足：当您的云主机、云硬盘、弹性 IP 配额不足时，将无法通过弹性伸缩添加云主机，请适当删除实例或提交工单提升配额。

如何避免伸缩配置中对应规格的云主机实例库存不足导致的扩容失败？

在设置伸缩组时，您可以为其指定多个可用区，以便在业务需求增加时，利用同一资源池下的多个可用区进行伸缩，从而提升业务的容灾能力，并尽量避免单一可用区云主机实例库存不足带来的扩容失败风险。

云主机实例如何确保不被移出伸缩组？

您可以对伸缩组中一个或多个正常状态的实例启用实例保护设置，当伸缩组发生自动缩容活动时，设置了实例保护的实例将不会被移除。您还可以设置伸缩组的最小实例数，配合伸缩组设置的“实例移出策略”，可以使伸缩组始终保持一定数量的云主机实例运行。

当伸缩组内的云主机实例被移出并释放后，数据还会保留吗？

不会。

当伸缩组内的云主机实例被移出后不会对云主机内已经存储的数据产生任何影响，但是当您执行的是移出并释放，云主机将会被直接释放，数据也不会被保留。您可以将重要的数据单独存储于确保不会被释放的云主机中。具体请参考帮助文档[将实例移出伸缩组](#)。

如何删除通过弹性伸缩组创建的云主机实例？

通过弹性伸缩组自动创建的实例，支持以下两种删除方式：

1. 您可以进入到弹性伸缩组中，手动移出并删除通过弹性伸缩创建的云主机实例。具体可参考帮助文档[将实例移出伸缩组](#)。
2. 当弹性伸缩执行缩容活动时，将根据实例创建方式自动移出或移出并释放云主机实例。

例。

伸缩组内实例管理类

怎样查看通过弹性伸缩创建的主机实例密码并进行登录？

系统无法获取您设置的密码内容，请您在创建伸缩配置时，妥善保管相关内容。

通过弹性伸缩创建的云主机实例，与伸缩组关联的伸缩配置中设置的登录方式和密码保持一致。若您希望更改后续通过该伸缩配置创建的云主机实例，可通过修改伸缩配置页面更新密码登录方式或密码。

自动创建的云主机实例密码是否与主机私有镜像中的密码一致？

不一致。

为了确保安全，所有自动创建的云主机实例密码都建议采用 SSH 密钥对。在自动创建云主机实例时，通常不会采用自定义镜像的密码来作为登录密码。

关联负载均衡操作类

为什么要为伸缩组关联负载均衡器？

将负载均衡器关联至伸缩组后，访问流量将自动分发到伸缩组内的所有云主机实例，扩展应用系统对外的服务能力，实现更高水平的应用程序容错性能。具体操作请参见[添加负载均衡器](#)。

怎样关联使用负载均衡器？

您可以在创建伸缩组时启用负载均衡，伸缩组中的云主机实例会自动挂载到您关联的负载均衡下。单个伸缩组内最多可关联 10 个负载均衡器。您在启用负载均衡时，需要填写负载均衡器、后端云主机组、后端端口、权重四个参数。具体操作请参见[添加负载均衡器](#)。

当伸缩组执行伸缩活动新增了云主机实例后，新增的云主机实例会进入负载均衡器的监听范围吗？

会。

当您为伸缩组关联了负载均衡后，当前伸缩组内的所有云主机实例以及后续因伸缩活动新增的云主机实例都会自动加入到负载均衡中，进入负载均衡器监听范围。

注意：

若负载均衡器关联的云主机数量超出限制，云主机加入负载均衡器监听范围可能失败，影响弹性伸缩扩容活动进行，请及时删除负载均衡关联的云主机或更换关联的负载均衡器。

能否将新创建的云主机实例添加到多台负载均衡器？

可以。

您可以为伸缩组关联多台负载均衡器，单个伸缩组内支持关联 10 个负载均衡器。结合负载均衡，负载均衡可使流量可智能分配到多个云主机上，提高处理效率和吞吐量；弹性伸缩服务可根据业务流量的变化自动调整云主机数量，保证业务处理能力的充足和稳定，提高业务的可用性和性能。

启用了负载均衡的云主机实例可以修改权重吗？

可以。

负载均衡后端云主机实例的权重分配指的是比例。比如两台云主机实例的权重为 1: 2，则代表所承载的业务比例为三分之一与三分之二。1: 1 的权重与 50: 50 的权重代表的意义相同。具体操作请参见[添加负载均衡器](#)。

当负载均衡器为外网类型时，是否要求云主机实例也配置弹性 IP 来访问外网？

不是。

理论上，弹性云主机实例可以不配公网带宽。但是为了避免不必要的管理问题，您可以在创建伸缩配置时，为伸缩组内的云主机实例都配置公网带宽。

负载均衡关联至伸缩组后，新增的云主机与负载均衡是如何联动来处理访问请求的？

当弹性伸缩组启用了负载均衡后，伸缩组中的云主机实例会自动挂载到关联的负载均衡下，被纳入进负载均衡器的监听范围内。负载均衡器在检查到您后端的云主机端口正常后，会按照预设好的转发策略来为伸缩组内的云主机实例转发访问请求，由云主机实例处理请求并返回结果。