

弹性伸缩服务

# 常见问题

文档版本 01  
发布日期 2021-12-08



版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

## 商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

## 注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或暗示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

# 安全声明

## 漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

# 目录

<b>1 通用类</b>	<b>1</b>
1.1 弹性伸缩有什么限制？	1
1.2 弹性伸缩一定要搭配弹性负载均衡、云监控才能使用吗？	2
1.3 弹性伸缩是否收取费用？	2
1.4 弹性伸缩是否会因监控指标突变导致误伸缩？	2
1.5 我能创建和使用多少个伸缩策略和配置？	2
1.6 弹性伸缩是否能够自动升降云服务器的 CPU、内存和带宽？	2
1.7 弹性伸缩的配额是什么？	2
1.8 同账户下不同用户操作弹性伸缩资源时，为什么提示密钥对不存在而拦截操作？	4
1.9 伸缩带宽对弹性公网 IP 带宽和共享带宽有什么要求？	4
1.10 弹性伸缩目前支持配置的服务器类型有哪些？可以配置云耀云服务器（HECS）吗？	4
1.11 组内单台 ECS 上的程序版本会随时更新，伸缩组能否自动创建并切换为最新镜像？	4
1.12 咨询 AS 扩容实例时能否创建包周期实例？	4
1.13 弹性伸缩能根据某个场景将带宽调高吗？	5
<b>2 伸缩组类</b>	<b>11</b>
2.1 伸缩组启用失败如何处理？	11
2.2 伸缩组异常情况下如何处理？	13
2.3 停用伸缩组后，什么操作会暂停？	15
2.4 已知弹性云服务器（ECS）实例 ID，能否查看当前实例所在的伸缩组信息？以及如何查看伸缩组及其中的实例信息？	15
<b>3 伸缩策略类</b>	<b>17</b>
3.1 我能启用多少个伸缩策略？	17
3.2 告警策略支持的告警触发条件有哪些？	17
3.3 什么是冷却时间，为什么需要冷却时间？	17
3.4 弹性伸缩是否可以根据云监控中自定义监控进行动态伸缩？	18
3.5 未安装 VM Tools 对弹性伸缩组监控指标有什么影响？	18
3.6 伸缩策略启用失败如何处理？	18
3.7 如需使用 Agent 监控指标，如何为伸缩组中的实例安装 Agent 插件？	18
<b>4 实例类</b>	<b>22</b>
4.1 如何保证手动移入的 ECS 实例不被移出伸缩组？	22
4.2 多规格伸缩配置创建实例的选择的规格顺序是什么？	22
4.3 当实例被移出伸缩组并删除后，实例中的数据会保留吗？	23

4.4 我能添加已经创建的包年包月 ECS 实例吗? .....	24
4.5 按照伸缩策略增加的云服务器, 当我不用时可以自动删除吗? .....	24
4.6 什么是期望实例数? .....	24
4.7 如何删除通过弹性伸缩创建的云服务器? .....	24
4.8 包年包月的 ECS 实例出现异常后会不会被伸缩组删除? .....	25
4.9 如何处理伸缩组中状态是“异常”的实例? .....	25
4.10 当伸缩组中实例无法通过负载均衡健康检查而频繁地被删除再重新创建时应该怎么办? .....	26
4.11 如何阻止伸缩组内的云服务器被自动移除? .....	26
4.12 为什么在伸缩组内移除并删除实例后, ECS 页面还能看到实例? .....	26
4.13 手动将实例移入伸缩组时, 需要该实例已绑定弹性公网 IP 吗? .....	26
4.14 弹性伸缩 (AS) 伸缩出来的新实例必须是伸缩组已经关联的伸缩配置里的镜像, 而不能复制现有的实例状态吗? .....	26
4.15 为 ECS 自动分配的 EIP, 是自动新购, 还是从未绑定的 EIP 中绑定? .....	26
<b>5 其他.....</b>	<b>27</b>
5.1 如何自动部署应用? .....	27
5.2 支持 Cloud-Init 特性后, 对使用弹性伸缩有哪些影响? .....	27
5.3 为什么使用密钥文件无法正常登录云服务器? .....	28
5.4 伸缩组中已经添加了负载均衡, 创建伸缩配置时是否还需要配置弹性公网 IP? .....	29
5.5 如何自动初始化弹性伸缩新增的云服务器数据盘? .....	29

# 1 通用类

## 1.1 弹性伸缩有什么限制？

弹性伸缩的云服务器中运行的应用需要是无状态、可横向扩展的。因为AS会自动释放云服务器，所以弹性伸缩组内的云服务器不可以保存应用的状态信息（例如session会话）和相关数据（如数据库、日志等）。

如果应用中需要云服务器保存状态或日志信息，可以考虑把相关信息保存到独立的服务器中。

弹性伸缩对用户的资源数量或容量做的配额限制如表1-1所示。

表 1-1 配额一览表

类别	描述	默认值
弹性伸缩组	用户可以创建的最多伸缩组个数。	50
弹性伸缩配置	用户可以创建的最多伸缩配置个数。	200
弹性伸缩策略	某个弹性伸缩组下可以创建的最多伸缩策略个数。	10
弹性伸缩实例	某个弹性伸缩组下可以创建的最多实例个数。	300
伸缩带宽策略	用户最多可以创建的伸缩带宽策略个数。	10
生命周期挂钩	某个弹性伸缩组内最多可添加的生命周期挂钩个数。	5
通知	某个弹性伸缩组最多可以配置的通知个数。	5
标签	某个弹性伸缩组最多可以添加的标签个数。	10

## 1.2 弹性伸缩一定要搭配弹性负载均衡、云监控才能使用吗？

**弹性伸缩**可以单独使用，也可以同**弹性负载均衡**（ELB），**云监控**（CES）一起使用。

其中，CES服务为免费服务，系统默认开通；ELB服务在有需求时可以部署，例如，有分布式集群需求的场景下，可以使用ELB。

## 1.3 弹性伸缩是否收取费用？

弹性伸缩服务本身不收取费用，但伸缩组自动创建的按需付费实例需要支付相应的费用，实例的计费标准请参见**计费说明**。实例使用的弹性公网IP也需支付相应的费用。伸缩组进行缩容时，自动创建的实例会被移出伸缩组并删除，删除后将不再收取费用。而手动移入的实例只会被移出伸缩组，系统仍会收取该实例的使用费用。若您不再需要使用该实例，请自行在ECS页面进行退订。

例如，弹性伸缩进行扩容活动创建了两台实例，使用一个小时后，进行了缩容活动，这两台实例被移出伸缩组并删除了，则系统只收取这两台实例使用一小时产生的费用。

## 1.4 弹性伸缩是否会因监控指标突变导致误伸缩？

不会。**弹性伸缩**的监控数据基于**云监控**来获取的，监控周期可配置多个档位，如：五分钟，二十分钟，一小时等。不会因为一次指标的高峰而导致错误伸缩。

同时，弹性伸缩还支持配置冷却时间，防止由于监控的变化造成伸缩组的反复无效变化。该时间可由用户进行自定义。

## 1.5 我能创建和使用多少个伸缩策略和配置？

您默认可以创建10个弹性伸缩组，100个弹性伸缩配置。每个弹性伸缩组同一时刻支持使用1个伸缩配置，10个伸缩策略。

如果系统提供的默认配额不能满足您的需求，请联系客服进行处理。

## 1.6 弹性伸缩是否能够自动升降云服务器的 CPU、内存和带宽？

弹性伸缩目前支持自动升降带宽资源，支持自动增加或减少伸缩组内ECS实例的数量，不支持自动升降服务器的CPU、内存等配置。

## 1.7 弹性伸缩的配额是什么？

### 什么是配额？

为防止资源滥用，平台限定了各服务资源的配额，对用户的资源数量和容量做了限制。如您最多可以创建多少个伸缩组。如果有需要，您可以申请扩大配额。

本节指导您如何查询指定区域下，弹性伸缩服务各资源的使用情况，以及总配额。

## 怎样查看我的配额？


1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 在页面右上角，选择“资源 > 我的配额”。  
系统进入“服务配额”页面。

图 1-1 我的配额



4. 您可以在“服务配额”页面，查看各项资源的总配额及使用情况。  
如果当前配额不能满足业务要求，请参考后续操作，申请扩大配额。

## 如何申请扩大配额？

1. 登录管理控制台。
2. 在页面右上角，选择“资源 > 我的配额”。  
系统进入“服务配额”页面。

图 1-2 我的配额



3. 在页面右上角，单击“申请扩大配额”。



图 1-3 申请扩大配额



服务	资源类型	已用配额	总配额
弹性云服务器 ECS	实例数	0	200
	核心数	0	800
	RAM容量(MB)	0	1,000,000
弹性云服务器 IMS	块设备	0	100
	磁盘数量	1	800
函数工作流 FunctionGraph	代码存储(MB)	0	10,000
	工作流数	0	800
弹性伸缩 AS	伸缩组	0	10
	伸缩配置	0	100

4. 在“新建工单”页面，根据您的需求，填写相关参数。  
其中，“问题描述”项请填写需要调整的内容和申请原因。
5. 填写完毕后，勾选协议并单击“提交”。

## 1.8 同账户下不同用户操作弹性伸缩资源时，为什么提示密钥对不存在而拦截操作？

密钥对作为用户级资源，无法跨用户使用。如果伸缩配置中配置的是同账号下其他用户的密钥对，那么在弹性伸缩下，无法使用该伸缩配置手动触发资源下发。

如果需要用户之前互相操作伸缩配置关联资源且不受密钥对权限限制，请使用密码方式进行虚拟机鉴权。

## 1.9 伸缩带宽对弹性公网 IP 带宽和共享带宽有什么要求？

弹性伸缩仅支持对按需购买的弹性公网IP带宽和共享带宽进行调整，不支持对包年包月的带宽进行调整。

## 1.10 弹性伸缩目前支持配置的服务器类型有哪些？可以配置云耀云服务器（HECS）吗？

目前AS支持的弹性云服务器（ECS）规格，具体请以前端界面为准。ECS规格详见《[弹性云服务器用户指南](#)》。

弹性伸缩（AS）目前暂不支持云耀云服务器（HECS）。

## 1.11 组内单台 ECS 上的程序版本会随时更新，伸缩组能否自动创建并切换为最新镜像？

当前没有自动创建的功能，需要客户自行编写一个监控的插件，提交新程序版本后，自动触发调用API，先重新创建镜像，再创建弹性伸缩配置，后替换弹性伸缩组内的配置。

## 1.12 咨询 AS 扩容实例时能否创建包周期实例？

AS创建实例只能选择按需实例和竞价实例。

## 1.13 弹性伸缩能根据某个场景将带宽调高吗？

弹性伸缩可以根据某个场景将带宽调高，具体操作步骤如下：

### 操作场景

用户可以通过伸缩带宽策略对购买的弹性公网IP带宽和共享带宽进行调整。本章节介绍如何创建伸缩带宽策略。

您可以通过创建伸缩带宽策略来实现自动调整带宽。创建伸缩带宽策略时，需要配置对应的基本信息，系统支持告警策略、定时策略、周期策略三种伸缩带宽策略。

创建伸缩带宽策略的基本信息，包括配置策略名称、资源类型、策略类型、触发条件等。

### 创建伸缩带宽告警策略

1. 登录管理控制台。
2. 选择“计算 > 弹性伸缩 > 伸缩带宽”。
3. 单击“创建伸缩带宽策略”。
4. 配置策略名称、资源类型、策略类型、触发条件等参数，请参考表1-2进行配置。

表 1-2 “告警策略”参数说明

参数	解释	取值样例
区域	创建的伸缩带宽策略所在的区域。	-
策略名称	创建的伸缩带宽策略的名称。 策略名称只能由中文、英文字母、数字、下划线、和中划线组成。	-
资源类型	需要进行管理的带宽的类型，可选择“弹性公网IP”或“共享带宽”。	弹性公网IP
弹性公网IP	需要进行伸缩管理的公网IP。 <b>说明</b> 当前仅支持对按需付费模式的弹性公网IP进行伸缩，包周期的弹性公网IP不能创建弹性伸缩。	-
策略类型	选择“告警策略”。	告警策略

参数	解释	取值样例
告警规则	<p>可选择已有的告警和新建告警。也可单击右侧“新建告警规则”在云监控服务页面创建告警规则，详细操作请参考<a href="#">新建告警规则</a>。</p> <p>若选择新建告警，需配置如下参数：</p> <ul style="list-style-type: none"> <li>告警规则名称 新建告警的名称，例如as-alarm-7o1u。</li> <li>触发条件 选择弹性带宽支持的监控指标及设定该监控指标的触发条件，支持的监控指标如<a href="#">表1-3</a>所示。例如出网带宽平均值&gt;100bit/s。</li> <li>监控周期 设定对弹性带宽支持的监控指标监控的周期，例如5分钟。</li> <li>连续出现次数 在监控周期内，连续达到触发条件几次后，开始执行伸缩活动，例如1次。</li> </ul>	-
执行动作	<p>设置伸缩策略执行动作。</p> <p>执行动作包括：</p> <ul style="list-style-type: none"> <li>增加 当执行伸缩活动时，增加带宽大小。</li> <li>减少 当执行伸缩活动时，减少带宽大小。</li> <li>设置为 将带宽大小设置为固定值。</li> </ul> <p><b>说明</b> 由于带宽在不同的取值范围内步长（即可调整的最小单位）不同，最终调整后的带宽会根据实际步长自动调整为就近值。</p> <ul style="list-style-type: none"> <li>小于等于300Mbit/s：默认步长为1Mbit/s。</li> <li>300Mbit/s~1000Mbit/s：默认步长为50Mbit/s。</li> <li>大于1000Mbit/s：默认步长为500Mbit/s。</li> </ul>	-
冷却时间	<p>冷却时间是指冷却伸缩活动的时间，单位为秒。在每次伸缩活动完成之后，系统开始计算冷却时间。伸缩组在冷却时间内，会拒绝告警策略的触发，其他类型的伸缩策略（如定时策略和周期策略）及手动触发不受限制。</p>	300秒

表 1-3 告警策略监控指标说明

指标名称	解释
入网带宽	该指标用于统计测量对象入云平台的网络速度。
入网流量	该指标用于统计测量对象入云平台的网络流量。

指标名称	解释
出网带宽	该指标用于统计测量对象出云平台的网络速度。
出网流量	该指标用于统计测量对象出云平台的网络流量。
出网带宽使用率	该指标用于统计测量对象出云平台的带宽使用率，以百分比为单位。

5. 参数配置完成后，单击“立即创建”。

在“伸缩带宽”页面中可查看新创建的伸缩带宽策略，新创建的策略默认的状态为“已启用”。

## 新建告警规则

创建伸缩带宽告警策略时，您可以通过单击“告警规则名称”右侧的“新建告警规则”创建所需的告警规则。可参考如下步骤进行操作。

1. 单击“告警规则名称”右侧的“新建告警规则”，跳转到云监控的告警规则页面。
2. 单击页面右上角“创建告警规则”。
3. 参考图1-4和表1-4进行参数配置。了解更多关于创建告警规则的信息请参考《云监控用户指南》。

图 1-4 创建告警规则

\* 资源类型

\* 维度

\* 监控范围

⇨ 全选

名称

名称	ID
无记录	

⇨ 取消全选

名称

名称	ID
<input type="checkbox"/> ecs-yxx-band...	b66f6fa3-365a-4ec7...

\* 选择类型

\* 告警策略       bit/s

\* 告警级别  紧急  重要  次要  提示


告警通知

发送通知

表 1-4 创建告警规则关键参数

参数	解释	配置示例
名称	告警规则的名称。	alarm-bandwidth
描述	告警规则描述（此参数非必填项）。	-
归属企业项目	告警规则所属的企业项目。只有拥有该企业项目权限的用户才可以查看和管理该告警规则。	default
资源类型	告警规则监控的服务名称，需要选择“弹性公网IP和带宽”。	弹性公网IP和带宽
维度	指定被监控的服务的具体模块。伸缩带宽是调整带宽大小的，需要选择“带宽”。	带宽
监控范围	告警规则适用的资源范围，需要选择“指定资源”。选择资源时可通过带宽名称或ID进行搜索，带宽名称或ID请在需要调整的EIP的详情页面获取。	指定资源

参数	解释	配置示例
触发规则	<p>根据需要可选择关联模板、导入已有模板或自定义创建。</p> <p><b>说明</b> 选择关联模板后，所关联模板内容修改后，该告警规则中所包含策略也会跟随修改。</p>	自定义创建
告警策略	触发告警规则的告警策略，请按需要进行配置。监控指标的含义请参考表1-3。	-

- 参数配置完成后，单击“立即创建”。
- 返回伸缩带宽策略创建页面，单击“告警规则名称”右侧的  按钮，然后为“告警规则名称”选择刚才创建的告警规则。

您还可以在创建伸缩带宽策略前，在云监控页面创建好所需的告警规则，创建告警规则时选择的指定资源必须是创建伸缩带宽策略时的选择的EIP资源对应的带宽资源。创建完成后，在创建伸缩带宽策略时可直接选择该告警规则。

## 创建伸缩带宽定时/周期策略

- 登录管理控制台。
- 选择“计算 > 弹性伸缩 > 伸缩带宽”。
- 单击“创建伸缩带宽策略”。
- 配置策略名称、资源类型、策略类型、触发条件等参数，请参考表1-5进行配置。

表 1-5 “定时策略”或“周期策略”参数说明

参数名称	参数说明	取值样例
区域	创建的伸缩带宽策略所在的区域。	-
策略名称	<p>创建伸缩带宽策略的名称。</p> <p>策略名称只能由中文字符、英文字母、数字、下划线、和中划线组成。</p>	as-policy-p6g5
资源类型	需要进行管理的带宽的类型，可选择“弹性公网IP”或“共享带宽”。	弹性公网IP
弹性公网IP	<p>需要进行伸缩管理的公网IP。当资源类型选择“弹性公网IP”时需要配置该项。</p> <p><b>说明</b> 当前仅支持对按需付费模式的弹性公网IP进行伸缩，包周期的弹性公网IP不能创建弹性伸缩。</p>	-
共享带宽	需要进行伸缩管理的共享带宽。当资源类型选择“共享带宽”时需要配置该项。	-

参数名称	参数说明	取值样例
策略类型	<p>可选择定时策略和周期策略，在指定的时间段进行调整带宽。</p> <p>若选择周期策略除了配置表格中的参数还需配置以下两个参数：</p> <ul style="list-style-type: none"> <li>生效时间 选择伸缩策略触发的时间段。</li> <li>重复周期 <ul style="list-style-type: none"> <li>按天</li> <li>按周</li> <li>按月</li> </ul> </li> </ul>	<p>按天</p> <p>2023/03/01 00:00:00 - 2023/03/31 23:59:59</p> <p>如在3月份每天执行伸缩活动，在生效时间结束后（即2023/04/01 00:00:00），该伸缩策略失效，不再触发伸缩活动。</p>
触发时间	<p>设定伸缩策略触发时间。</p> <p><b>说明</b> 周期策略的触发时间需包含在策略的生效时间内。</p>	-
执行动作	<p>设置伸缩活动的执行动作。</p> <p>执行动作包括：</p> <ul style="list-style-type: none"> <li>增加 当执行伸缩活动时，增加带宽大小。</li> <li>减少 当执行伸缩活动时，减少带宽大小。</li> <li>设置为 将带宽大小设置为固定值。</li> </ul> <p><b>说明</b> 由于带宽在不同的取值范围内步长（即可调整的最小单位）不同，最终调整后的带宽会根据实际步长自动调整为就近值。</p> <ul style="list-style-type: none"> <li>小于等于300Mbit/s：默认步长为1Mbit/s。</li> <li>300Mbit/s~1000Mbit/s：默认步长为50Mbit/s。</li> <li>大于1000Mbit/s：默认步长为500Mbit/s。</li> </ul>	-
冷却时间	<p>冷却时间是指冷却伸缩活动的时间，单位为秒。在每次伸缩活动完成之后，系统开始计算冷却时间。伸缩组在冷却时间内，会拒绝告警策略的触发，其他类型的伸缩策略（如定时策略和周期策略）及手动触发不受限制。</p>	300秒

5. 参数配置完成后，单击“立即创建”。

# 2 伸缩组类

## 2.1 伸缩组启用失败如何处理？

伸缩组启用失败处理方法：

- 情况描述：云服务器配额不足、云硬盘配额不足、弹性公网IP配额不足等。  
可能原因：配额不足。  
处理方法：申请扩大配额或者删除不需要的资源，之后重新启用伸缩组。
- 情况描述：伸缩组使用的VPC不存在、子网不存在等。  
可能原因：虚拟私有云服务异常或者相关资源被删除。  
处理方法：等待虚拟私有云服务恢复或者修改伸缩组中VPC、子网相关参数，之后重新启用伸缩组。
- 情况描述：负载均衡监听器不存在、后端云服务器组不存在、负载均衡器不可用等。  
可能原因：负载均衡服务异常或者相关资源被删除。  
处理方法：等待负载均衡服务恢复或者修改伸缩组中负载均衡相关参数，之后重新启用伸缩组。
- 情况描述：您添加到负载均衡监听器的后端云服务器超过最大限制。  
可能原因：伸缩组使用经典型负载均衡，会自动地将加入伸缩组的实例添加到负载均衡监听器。一个监听器最多可添加300个后端云服务器。  
处理方法：您可以从监听器上移除不需要的且不在该伸缩组中的后端云服务器，之后重新启用伸缩组。
- 情况描述：伸缩配置使用的镜像不存在、规格不存在、密钥对不存在等。  
可能原因：相关资源被删除。  
处理方法：为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：生命周期挂钩使用的通知主题不存在。  
可能原因：伸缩组添加了生命周期挂钩，且伸缩活动开始前生命周期挂钩使用的通知主题被删除。如果伸缩活动开始后通知主题被删除，则下次伸缩活动时伸缩组会异常。  
处理方法：修改生命周期挂钩使用的通知主题或者删除生命周期挂钩，之后重新启用伸缩组。



- 情况描述：您选择的子网下私有IP不足。  
可能原因：伸缩组使用的子网下私有IP地址被用尽。  
处理方法：修改伸缩组中的子网信息，之后重新启用伸缩组。
- 情况描述：您选择的可用区下的该类型云服务器资源已售罄。  
可能原因：您为伸缩组选择的可用区下该类型云服务器资源售罄，或者该可用区不支持该类型云服务器资源。该类型云服务器指的是伸缩配置里选择的云服务器规格。  
处理方法：您可以为伸缩组更换新的伸缩配置，之后重新启用伸缩组。当您的伸缩组中没有实例时，还可以修改伸缩组的可用区信息，之后重新启用伸缩组。
- 情况描述：您选择的规格和磁盘不匹配。  
可能原因：伸缩配置中的云服务器类型和磁盘类型不匹配导致创建云服务器失败。  
处理方法：为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：您选择的规格和镜像不匹配。  
可能原因：伸缩配置中的云服务器类型和镜像不匹配导致创建云服务器失败。  
处理方法：为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：您选择的可用区下的该类型存储资源已售罄。  
可能原因：您为伸缩组选择的可用区下该类型存储资源售罄，或者该可用区不支持该类型存储资源。该类型存储资源指的是伸缩配置里选择的磁盘（包括系统盘和数据盘）类型。  
处理方法：您可以为伸缩组更换新的伸缩配置，之后重新启用伸缩组。当您的伸缩组中没有实例时，还可以修改伸缩组的可用区信息，之后重新启用伸缩组。
- 情况描述：您选择的伸缩配置中定义的共享带宽不存在。  
可能原因：相关资源被删除。  
处理方法：您可以使用新购买的共享带宽或者其他存在的共享带宽资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：伸缩配置中定义的共享带宽绑定的EIP个数超过最大限制。  
可能原因：单个共享带宽最多可以添加20个公网IP。  
处理方法：您可以申请扩大配额或者从共享带宽上移出不需要的公网IP，或者为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：您的伸缩配置中选择的专属主机不存在，请更换伸缩配置。  
可能原因：相关资源被删除。  
处理方法：您可以使用新购买的专属主机或者其他存在的专属主机资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：无可用的专属主机资源，请确保当前有可用的专属主机资源。  
处理方法：您可以排除专属主机的故障使其恢复可用状态或者开启专属主机自动放置属性，之后重新启用伸缩组。还可以使用新购买的专属主机资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：您的伸缩配置中选择的专属主机可用容量不足。  
处理方法：您可以删除专属主机上不需要的云服务器资源之后重新启用伸缩组，或者使用新购买的专属主机资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：您的伸缩组中选择的可用区下无可用的专属主机资源。

处理方法：您可以在该可用区下购买新的专属主机资源，重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。当您的伸缩组中没有实例时，还可以修改伸缩组的可用区信息，之后重新启用伸缩组。

- 情况描述：您的伸缩配置中选择的专属主机不支持该类型云服务器，请更换伸缩配置。

处理方法：您可以选择专属主机支持的规格重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。

- 情况描述：系统异常。

可能原因：弹性伸缩或者周边服务异常、网络异常等。

处理方法：请稍后重试，或联系技术支持。

- 情况描述：您选择的伸缩配置中定义的规格不可用。

处理方法：您可以通过错误提示信息通过重新创建伸缩配置更换新的规格，并为伸缩组更换新的伸缩配置，之后重新启用伸缩组

- 情况描述：你选择的伸缩配置无法被当前伸缩组使用。

处理方法：您可以通过错误提示信息重新创建伸缩配置，并为伸缩组更换新的伸缩配置，之后重新启用伸缩组。

- 情况描述：您的账户已欠费或余额不足。

可能原因：账户欠费或余额不足时，伸缩组无法进行扩容。

处理方法：您可以在充值之后重新启用伸缩组。

## 2.2 伸缩组异常情况下如何处理？

伸缩组异常情况及其处理方法：

- 情况描述：云服务器配额不足、云硬盘配额不足、弹性公网IP配额不足等。

可能原因：配额不足。

处理方法：申请扩大配额或者删除不需要的资源，之后重新启用伸缩组。

- 情况描述：伸缩组使用的VPC不存在、子网不存在等。

可能原因：虚拟私有云服务异常或者相关资源被删除。

处理方法：等待虚拟私有云服务恢复或者修改伸缩组中VPC、子网相关参数，之后重新启用伸缩组。

- 情况描述：负载均衡监听器不存在、后端云服务器组不存在、负载均衡器不可用等。

可能原因：负载均衡服务异常或者相关资源被删除。

处理方法：等待负载均衡服务恢复或者修改伸缩组中负载均衡相关参数，之后重新启用伸缩组。

- 情况描述：您添加到负载均衡监听器的后端云服务器超过最大限制。

可能原因：伸缩组使用经典型负载均衡，会自动地将加入伸缩组的实例添加到负载均衡监听器。一个监听器最多可添加300个后端云服务器。

处理方法：您可以从监听器上移除不需要的且不在该伸缩组中的后端云服务器，之后重新启用伸缩组。

- 情况描述：伸缩配置使用的镜像不存在、规格不存在、密钥对不存在等。

可能原因：相关资源被删除。

- 处理方法：为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
- 情况描述：生命周期挂钩使用的通知主题不存在。  
可能原因：伸缩组添加了生命周期挂钩，且伸缩活动开始前生命周期挂钩使用的通知主题被删除。如果伸缩活动开始后通知主题被删除，则下次伸缩活动时伸缩组会异常。  
处理方法：修改生命周期挂钩使用的通知主题或者删除生命周期挂钩，之后重新启用伸缩组。
  - 情况描述：您选择的子网下私有IP不足。  
可能原因：伸缩组使用的子网下私有IP地址被用尽。  
处理方法：修改伸缩组中的子网信息，之后重新启用伸缩组。
  - 情况描述：您选择的可用区下的该类型云服务器资源已售罄。  
可能原因：您为伸缩组选择的可用区下该类型云服务器资源售罄，或者该可用区不支持该类型云服务器资源。该类型云服务器指的是伸缩配置里选择的云服务器规格。  
处理方法：您可以为伸缩组更换新的伸缩配置，之后重新启用伸缩组。当您的伸缩组中没有实例时，还可以修改伸缩组的可用区信息，之后重新启用伸缩组。
  - 情况描述：您选择的规格和磁盘不匹配。  
可能原因：伸缩配置中的云服务器类型和磁盘类型不匹配导致创建云服务器失败。  
处理方法：为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
  - 情况描述：您选择的规格和镜像不匹配。  
可能原因：伸缩配置中的云服务器类型和镜像不匹配导致创建云服务器失败。  
处理方法：为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
  - 情况描述：您选择的可用区下的该类型存储资源已售罄。  
可能原因：您为伸缩组选择的可用区下该类型存储资源售罄，或者该可用区不支持该类型存储资源。该类型存储资源指的是伸缩配置里选择的磁盘（包括系统盘和数据盘）类型。  
处理方法：您可以为伸缩组更换新的伸缩配置，之后重新启用伸缩组。当您的伸缩组中没有实例时，还可以修改伸缩组的可用区信息，之后重新启用伸缩组。
  - 情况描述：您选择的伸缩配置中定义的共享带宽不存在。  
可能原因：相关资源被删除。  
处理方法：您可以使用新购买的共享带宽或者其他存在的共享带宽资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
  - 情况描述：伸缩配置中定义的共享带宽绑定的EIP个数超过最大限制。  
可能原因：单个共享带宽最多可以添加20个公网IP。  
处理方法：您可以申请扩大配额或者从共享带宽上移出不需要的公网IP，或者为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
  - 情况描述：您的伸缩配置中选择的专属主机不存在，请更换伸缩配置。  
可能原因：相关资源被删除。  
处理方法：您可以使用新购买的专属主机或者其他存在的专属主机资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。
  - 情况描述：无可用的专属主机资源，请确保当前有可用的专属主机资源。

处理方法：您可以排除专属主机的故障使其恢复可用状态或者开启专属主机自动放置属性，之后重新启用伸缩组。还可以使用新购买的专属主机资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。

- 情况描述：您的伸缩配置中选择的专属主机可用容量不足。

处理方法：您可以删除专属主机上不需要的云服务器资源之后重新启用伸缩组，或者使用新购买的专属主机资源重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。

- 情况描述：您的伸缩组中选择的可用区下无可用的专属主机资源。

处理方法：您可以在该可用区下购买新的专属主机资源，重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。当您的伸缩组中没有实例时，还可以修改伸缩组的可用区信息，之后重新启用伸缩组。

- 情况描述：您的伸缩配置中选择的专属主机不支持该类型云服务器，请更换伸缩配置。

处理方法：您可以选择专属主机支持的规格重新创建伸缩配置，为伸缩组更换新的伸缩配置，之后重新启用伸缩组。

- 情况描述：系统异常。

可能原因：弹性伸缩或者周边服务异常、网络异常等。

处理方法：请稍后重试，或联系技术支持。

- 情况描述：您选择的伸缩配置中定义的规格不可用。

处理方法：您可以通过错误提示信息通过重新创建伸缩配置更换新的规格，并为伸缩组更换新的伸缩配置，之后重新启用伸缩组。

- 情况描述：你选择的伸缩配置无法被当前伸缩组使用。

处理方法：您可以通过错误提示信息重新创建伸缩配置，并为伸缩组更换新的伸缩配置，之后重新启用伸缩组。

- 情况描述：您的账户已欠费或余额不足。

可能原因：账户欠费或余额不足时，伸缩组无法进行扩容。

处理方法：您可以在充值之后重新启用伸缩组。

## 2.3 停用伸缩组后，什么操作会暂停？

停用伸缩组之后，伸缩组将不再自动触发伸缩活动，但是已开始的伸缩活动会继续执行。伸缩策略不会自动触发伸缩活动。手动调整期望实例数后，尽管当前实例数与期望实例数不相等，但是不会触发伸缩活动。

健康检查会继续检查实例的健康状态，但不会执行移除操作。

## 2.4 已知弹性云服务器（ECS）实例 ID，能否查看当前实例所在的伸缩组信息？以及如何查看伸缩组及其中的实例信息？

不能。

查看伸缩组及其中的实例信息步骤如下：

**步骤1** 进入“伸缩实例”页面。

**步骤2** 单击“弹性伸缩组”页签下的伸缩组名称，进入伸缩组概览页，可查看伸缩组信息。

**步骤3** 单击“伸缩实例”页签，可查看该伸缩组下的实例信息。

----结束

# 3 伸缩策略类

## 3.1 我能启用多少个伸缩策略？

伸缩策略可以启用一个，也可以启用多个。

## 3.2 告警策略支持的告警触发条件有哪些？

可以针对CPU使用率、内存使用率、带内网络流入速率、带内网络流出速率、磁盘读取速率、磁盘写速率、磁盘读操作速率、磁盘写操作速率等指标进行监测告警，自动增加或减少ECS实例。

## 3.3 什么是冷却时间，为什么需要冷却时间？

冷却时间是指冷却伸缩活动的时间。每次伸缩活动完成之后，系统开始计算冷却时间。伸缩组在冷却时间内，会拒绝告警策略的触发，其他类型的伸缩策略（如定时策略和周期策略）及手动触发不受限制。

实例加入伸缩组投入使用之前，需要使用配置脚本安装和配置软件，大约需要两到三分钟（实际时间取决于诸多因素，如实例规格和是否有启动脚本等）。因此实例从启动到投入使用如果没有冷却时间，系统会在负载降下来前不断扩容，新加入的实例接管业务后，发现负载过低，然后又缩容。冷却时间避免了伸缩组重复进行不必要的伸缩活动。

冷却时间工作原理举例：

业务出现流量高峰，触发告警策略，按照配置AS会自动新增一个实例到伸缩组来帮助处理增加的需求。但是存在一个问题：该实例需要几分钟的时间才能启动，并且启动后到可以从ELB接收请求也需要一段时间。在此期间，告警可能会持续触发，从而导致告警每次触发时都会新增一个实例。若您设置了冷却时间，AS在启动一个实例后，将暂停告警策略引起的扩展活动，直至经过了该指定时间段（默认值为300秒）。这样，新启动的实例有时间开始处理应用程序流量。冷却时间过后，如果告警再次触发，AS才会启动另一个实例，而冷却时间也会再次生效。

## 3.4 弹性伸缩是否可以根据云监控中自定义监控进行动态伸缩?

弹性伸缩支持根据自定义监控进行动态伸缩。

## 3.5 未安装 VM Tools 对弹性伸缩组监控指标有什么影响?

未安装VM Tools，云监控无法监控弹性云服务器的内存使用率、带内网络流入速率和带内网络流出速率三个指标。但可以监控带外网络流入速率和带外网络流出速率指标，这样导致CPU使用率指标的精确性可能会降低。

如果弹性云服务器类型是IO优化实例，无论是否安装了vmttools，监控弹性云服务器的监控指标不包含磁盘使用率、带内网络流入速率和带内网络流出速率指标。

因此，当上述情况发生时，弹性伸缩支持的内存使用率、带内网络流入速率和带内网络流出速率三个指标会受到影响，无法获取监控数据。

## 3.6 伸缩策略启用失败如何处理?

- 情况描述：告警规则不存在。  
可能原因：告警策略中使用的告警规则被删除。  
处理方法：修改告警策略中使用的告警规则，之后重新启用策略。
- 情况描述：周期策略的触发时间不包含在策略的生效时间内。  
可能原因：周期策略的生效时间已过期。  
处理方法：修改周期策略生效的起始时间和结束时间，之后重新启用策略。
- 情况描述：定时策略触发时间必须晚于当前时间。  
可能原因：定时策略触发时间已过期。  
处理方法：修改定时策略的触发时间，之后重新启用策略。
- 情况描述：系统异常。  
处理方法：请稍后重试，或联系技术支持。

## 3.7 如需使用 Agent 监控指标，如何为伸缩组中的实例安装 Agent 插件?

### 问题背景

伸缩策略的类型为告警策略时，支持使用Agent监控指标触发伸缩活动。Agent监控即操作系统监控，相比基础监控，操作系统监控可以为用户提供服务器的系统级、主动式、细颗粒度监控服务。如需使用Agent监控指标，伸缩组中的实例必须均已安装Agent插件，本文提供了详细的操作指导。

### 操作步骤

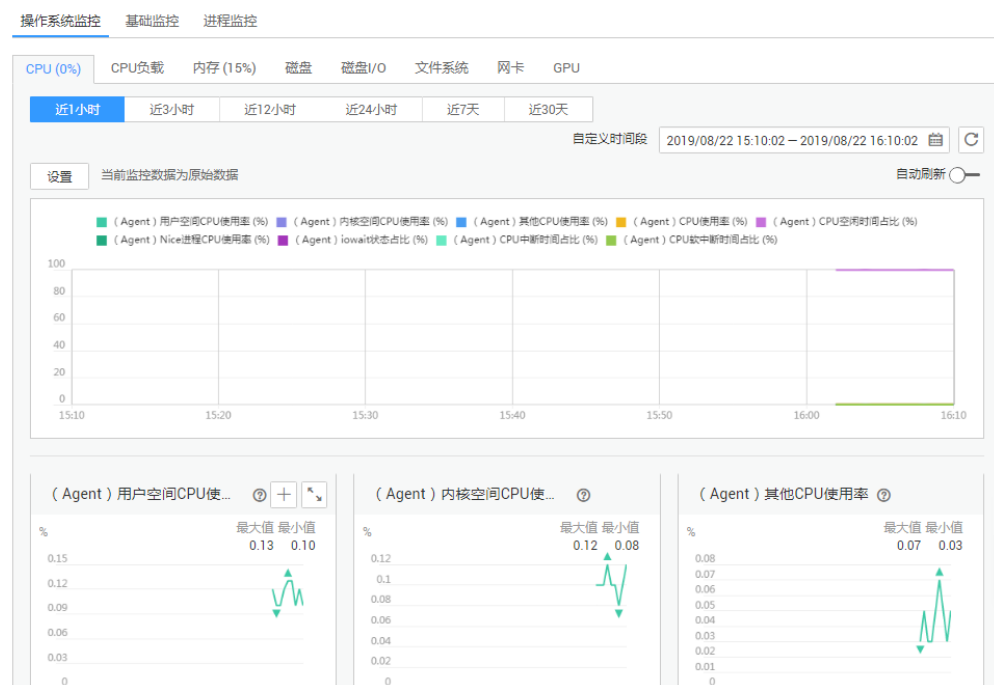
1. 登录管理控制台，选择“计算 > 弹性云服务器”。

- 进入云服务器控制台。
2. 创建一台弹性云服务器，并安装Agent插件。  
请参考“[Agent安装配置方式说明](#)”选择一种方式进行插件安装。
  3. 待Agent插件安装成功后，进入云监控控制台，选择“主机监控 > 弹性云服务器”，确保插件状态为“运行中”且能够采集Agent监控指标数据。

图 3-1 查看插件状态

<input type="checkbox"/>	名称/ID	私有IP地址	主机状态	插件状态	监控状态
<input type="checkbox"/>	ecs-ec78 fa07b856-2edd-4cd0-9aa6...	192.168....	运行中	运行中	<input checked="" type="checkbox"/>

图 3-2 查看 Agent 监控指标



4. 向弹性云服务器的conf.json配置文件中添加AccessKey/SeretKey ( AK/SK ) 信息。
  - a. 单击用户名，选择“我的凭证 > 访问密钥”，获取AK/SK。
    - 如已有访问密钥，查看创建时下载保存的credentials.csv文件，获取文件中记录的Key值。
    - 如未创建，则通过“新增访问密钥”创建新的密钥，妥善保存credentials.csv文件，并获取文件中记录的Key值。
  - b. 登录弹性云服务器，执行`cd /usr/local/telescope/bin`进入Agent安装路径。
  - c. 执行`vi conf.json`打开配置文件，输入已获取的AK/SK。



```
{
  "InstanceId": "fa07b[REDACTED]4cd0-9aa6-e5c791569e3a",
  "ProjectId": "050b1[REDACTED]572f8cc01f3740bed5",
  "AccessKey": "MK8NR3[REDACTED]7FUMJB",
  "SecretKey": "sPHiTB8[REDACTED]N4wWw3YCNwcUFqj",
  "RegionId": "cn-north-1"
}
```

如果使用“购买ECS时安装Agent”的安装方式，在注入用户数据时已经添加了AK/SK信息，此处只需检查。

- d. 按“ESC”，输入:wq保存并退出。
5. 进入镜像服务页面，将这台弹性云服务器制作作为私有镜像，详细操作请参考“[创建私有镜像](#)”。

图 3-3 创建私有镜像

\* 创建方式 **系统盘镜像** 整机镜像 数据盘镜像

\* 选择镜像源 云服务器 裸金属服务器 镜像文件

- 当前关机或开机状态的弹性云服务器才可以用来自创建私有镜像。
- 创建镜像前，请确保弹性云服务器已完成相关配置。[了解更多](#)
- 请勿在创建镜像过程中对所选择的弹性云服务器及其相关联资源进行其他操作。

所有状态 名称 搜索 刷新

名称	操作系统	运行状态	私有IP地址	创建时间
ecs-ec78	CentOS 7.4 64bit	运行中	192.168.1.158	2019/08/22 15:58:5...

当前选择: ecs-ec78|操作系统: CentOS 7.4 64bit|系统盘: 普通IO | 40 GB  
[购买弹性云服务器](#)

配置信息

加密 未加密

\* 名称 ces-agent-test

标签 如果您需要使用同一标签标识多种云资源，即所有服务均可在标签输入框下拉选择同一标签，建议在TMS中创建预定义标签。[查看预定义标签](#)

标签键 标签值

您还可以添加10个标签。

描述 0/1024

6. 进入弹性伸缩页面，使用步骤5中创建的私有镜像创建伸缩配置。

图 3-4 选择私有镜像

\* 镜像 公共镜像 **私有镜像** 共享镜像

ces-agent-test(40GB)

在“镜像”区域，单击“私有镜像”，在下拉列表中选择“ces-agent-test”，其他参数按照实际需求配置。

7. 创建伸缩组，并绑定步骤6中的伸缩配置。
8. 为伸缩组添加伸缩策略，策略类型选择“告警策略”，触发条件选择Agent相关监控指标，如：（Agent）内存使用率。

图 3-5 选择触发条件



9. 将步骤2中的云服务器手动移入伸缩组。
10. 等待并验证Agent监控指标是否生效，例如，验证以下项目：
  - 伸缩组详情页“监控”页签下显示有Agent监控指标
  - 达到告警阈值时，在伸缩组详情页“活动历史”页签查看告警策略触发成功，且发生实例的伸缩
  - 伸缩组自动扩容出来的云服务器均有Agent监控数据

# 4 实例类

## 4.1 如何保证手动移入的 ECS 实例不被移出伸缩组？

假设您向伸缩组手动移入了N台ECS实例，并且不希望这些实例被自动移出，那么您可以通过如下两种方法确保这些实例不被移出伸缩组。

### 方法一

在伸缩组同时进行如下两条配置：

- 将伸缩组的最小实例数设置为N或者大于N的值。
- 将实例移除策略配置为“根据较早创建的配置较早创建的实例”或“根据较早创建的配置较晚创建的实例”。

根据弹性伸缩的规则，手工添加的实例不会对应任何伸缩配置（因为它们不是通过伸缩配置创建的），所以弹性伸缩会先挑选通过伸缩配置自动创建的实例进行释放，只有当自动创建的实例释放完了，才会挑选手工添加的实例进行释放。由于您将最小实例数设置成N或大于N，所以手工添加的实例是不会被选中。

注意：以上是在您手动移入的实例处于正常的情况下，如果这些实例处于关机或其他异常状态，弹性伸缩会视为它们不健康，并将它们移出伸缩组，因为健康检查需要保证在伸缩组里的实例是健康的。

### 方法二

为这N台实例设置实例保护，具体操作方法可参见[实例保护](#)。

您可以同时为这N台实例设置实例保护，当伸缩组发生缩容活动时，设置了实例保护的实例将不会被移出伸缩组。注意，实例若未通过健康检查仍然会被移出伸缩组。

## 4.2 多规格伸缩配置创建实例的选择的规格顺序是什么？

当伸缩配置选择多个规格时，根据伸缩组可用区及多可用区扩展策略的配置不同，创建实例时选择的规格顺序不同，本章节将分单个可用区和多可用区情况说明。

## 单可用区

当伸缩组只选择了一个可用区时，伸缩组中的实例都会创建在该可用区中。伸缩配置选择多个规格时，可按照如下两种规格使用优先顺序创建实例：

- 选择优先：伸缩组扩容时按照您选择规格的顺序进行。例如，您依次选择了规格2、3、1。系统会按照您选择的顺序优先选择规格2创建实例，当规格2库存不足或者因为其他原因创建失败时，系统会选择规格3创建实例，当规格3也无法创建实例时才会使用规格1。
- 成本优先：伸缩组扩容时按照价格最优原则进行优先级排序。例如，您依次选择了规格1、2、3。这三个规格按成本排序为：1>3>2。系统会优先选择规格2（成本最低的规格）创建实例，当规格2无法创建实例时，选择规格3，当规格3也无法创建实例时，才会选择规格1。

## 多可用区

当伸缩组选择了两个及两个以上的可用区时，需要配置“多可用区扩展策略”（“均衡分布”或“选择优先”）。当您选择不同的多可用区扩展策略时，选择的实例规格的创建顺序也会不同。对不同的创建顺序分情况说明如下：

- 均衡分布：云服务器扩容时优先保证选择的可用区列表中各可用区下云服务器数量均衡，当无法在目标可用区下完成云服务器扩容时，按照选择优先原则选择其他可用区。创建实例选择AZ和规格的顺序举例如下：

您依次选择了可用区AZ1、AZ2、AZ3，创建伸缩配置时选择了规格1、2、3，且规格选择的优先级顺序为2、3、1。AZ1、AZ2、AZ3分别有3、2、3台实例，按照均衡分布原则AZ2的实例数较少，优先选择AZ2创建实例。先使用规格2在AZ2中创建实例，若成功则伸缩活动成功，若规格2无法在AZ2中创建实例，依次尝试使用规格3和规格1创建实例，若均失败，则AZ2中无法创建实例。根据当均衡分布无法创建实例，按照选择优先原则选择其他可用区，将在AZ1中按照规格2、3、1的顺序依次尝试创建实例。若AZ1仍无法创建云服务器，选择AZ3进行尝试，选择规格的顺序仍是2、3、1。

- 选择优先：云服务器扩容时目标可用区的选择按照选择的可用区列表的顺序进行优先级排序。创建实例选择AZ和规格的顺序举例如下：

您依次选择了可用区AZ1、AZ2、AZ3，创建伸缩配置时选择了规格1、2、3，且规格选择的优先级顺序为2、3、1。无论AZ中的实例是否均衡，系统会按照您选择AZ的顺序，即AZ1、AZ2、AZ3依次创建实例。先使用规格2在AZ1中创建实例，若无法成功，使用规格3在AZ1中创建实例，若仍无法成功，再使用规格1在AZ1中创建实例。若使用3种规格在AZ1中均无法创建实例，则尝试在AZ2中创建实例，仍按照规格顺序2、3、1进行创建。若AZ2中仍无法创建实例，选择AZ3创建实例，规格选择顺序仍为2、3、1。

### 📖 说明

该部分规格的优先级顺序是根据伸缩配置中选择的“规格使用优先策略”确定的，具体的确定方式可参考[单可用区](#)。

## 4.3 当实例被移出伸缩组并删除后，实例中的数据会保留吗？

不会。弹性伸缩会自动释放ECS实例，您需要确保伸缩组内的实例没有保存应用的状态信息或者重要数据，例如，会话记录（session）、数据库和日志等。如果您的应用需要保存状态信息，可以考虑将状态信息保存到独立的状态服务器（如ECS）、数据库（如RDS）等。

如果您需要在伸缩组中的实例被释放前执行数据备份操作或者下载日志文件，您可以为伸缩组添加“实例终止”类型的生命周期挂钩。添加生命周期挂钩后，当伸缩组进行伸缩活动时，正在移出伸缩组的实例将被挂钩挂起并置于等待状态，您能够在实例保持等待状态的时间内执行数据备份或者下载日志文件等自定义操作。

## 4.4 我能添加已经创建的包年包月 ECS 实例吗？

能。目前弹性伸缩默认自动创建按需付费的ECS实例，同时支持手动添加已经创建的包年包月和按需付费ECS实例。

## 4.5 按照伸缩策略增加的云服务器，当我不用时可以自动删除吗？

可以，但是需要您增加一条删除云服务器的伸缩策略。

## 4.6 什么是期望实例数？

期望实例数是指伸缩组中期望运行的弹性云服务器的个数，大小介于最小实例数和最大实例数之间。您可以手动调整期望实例数，也可以通过定时（周期）策略和告警策略触发调整期望实例数。

创建伸缩组时设置期望实例数：当期望实例数不为0时，伸缩组创建完成后会立即开始伸缩活动自动添加相应个数的弹性云服务器。

手动调整期望实例数：当用户通过直接修改的方式更改了期望实例数，系统发现当前实例数与期望实例数不相等，就会触发伸缩活动，使当前实例数等于期望实例数。

伸缩策略触发调整期望实例数：伸缩策略触发后，如果该策略的触发动作是增加2个实例，系统首先会将期望实例数在当前基础上+2。此时，由于伸缩组的当前实例数与期望实例数不再相等，系统就会触发伸缩活动增加2个实例，使当前实例数等于期望实例数。

## 4.7 如何删除通过弹性伸缩创建的云服务器？

### 处理方法

方法一：

1. 登录管理控制台。
2. 选择“计算 > 弹性伸缩 > 伸缩实例”。
3. 单击具体的“弹性伸缩组”名称。
4. 在“弹性伸缩组”详情页面中，选择“伸缩实例”页签。
5. 在“操作”列下选择需要删除的实例，单击“移出伸缩组并删除”。

#### 说明

如果您要删除多个实例，可以依次勾选指定实例名称左侧的方框，单击“移出伸缩组并删除”。

方法二

1. 登录管理控制台。
2. 选择“计算 > 弹性伸缩 > 伸缩实例”。
3. 单击具体的“弹性伸缩组”名称。
4. 在“弹性伸缩组”详情页面中，选择“伸缩策略”页签。
5. 在“伸缩策略”页签中增加一条删除云服务器的伸缩策略，可以按照需要减少或者调整至指定数量。

#### 方法三

1. 登录管理控制台。
1. 选择“计算 > 弹性伸缩 > 伸缩实例”。
2. 单击具体的“弹性伸缩组”名称。
3. 在“弹性伸缩组”详情页面中，单击右上方“修改”，进入“修改伸缩组”页面。
4. 手动修改伸缩组的期望实例数。

## 4.8 包年包月的 ECS 实例出现异常后会不会被伸缩组删除？

不会。伸缩组不会删除包周期的ECS实例，包周期的ECS实例出现异常后，只会将ECS实例移出伸缩组，不会进行删除操作。

## 4.9 如何处理伸缩组中状态是“异常”的实例？

正常情况下，您不要处理伸缩组中状态是“异常”的实例，弹性伸缩的健康检查功能会周期性地对伸缩组中实例的健康状态进行检查。当伸缩组为启用状态时，会将异常的实例从伸缩组中移除，然后重新创建新的实例以维持伸缩组的期望实例数和当前实例数保持一致。当伸缩组为非启用状态时，对实例的健康状态会继续进行检查，但不会执行移除操作。

值得注意的是，负载均衡健康检查是通过负载均衡系统向后端云服务器发起心跳检查的方式来实现的，而负载均衡系统和云服务器之间是通过内网进行通信的。所以，如果伸缩组使用负载均衡健康检查方式，为确保健康检查工作的正常进行，您需要确保能够通过内网访问您的云服务器，请按照以下方法排查。

1. 在监听器页面，在健康检查异常的监听器所在行，单击“健康检查”列下的“查看”。弹出健康检查配置项提示框。
  - 检查“健康检查方式”：确保后端服务器已配置相应协议并开启端口。
  - 检查“检查路径”：如果是使用HTTP协议进行健康检查，还应检查后端主机的健康检查路径是否正确。
2. 检查云服务器中防火墙等软件是否有对来自健康检查源IP的屏蔽。
3. 检查后端云主机所在安全组与网络ACL规则是否配置放行100.125.0.0/16，并配置ELB用于健康检查的协议和端口。健康检查的协议和端口在步骤1中弹出的健康检查配置项提示框中获取。
  - 若采用默认的健康检查方式：需要放行后端云服务器业务端口。
  - 若配置了不同于云服务器业务端口的健康检查端口：需要放行云服务器业务端口与健康检查端口。
4. 如果以上配置检查均正常但问题依然存在，请联系技术支持。

## 4.10 当伸缩组中实例无法通过负载均衡健康检查而频繁地被删除再重新创建时应该怎么办？

实例所在安全组规则必须放通100.125.0.0/16网段，且协议和端口号需要为ELB用于健康检查的协议和端口，否则会导致健康检查失败。

## 4.11 如何阻止伸缩组内的云服务器被自动移除？

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伸缩组。您还可以设置伸缩组的最小实例数，配合实例移除策略，可以使伸缩组始终至少保持一定数量的ECS实例运行。

值得注意的是，健康检查会将异常的实例从伸缩组中移除，并重新创建新的ECS实例。所以，请不要在ECS页面停止或者删除已经加入到伸缩组中的ECS实例，否则我们认定该ECS实例不健康，并自动移出伸缩组。当伸缩组为停用状态时，对实例的健康状态会继续进行检查，但不会执行移除操作。

## 4.12 为什么在伸缩组内移除并删除实例后，ECS 页面还能看到实例？

如果伸缩组内自动伸出的实例被锁定，那么实例移出伸缩组时，仅移除实例，不做删除操作，以确保该实例可被其他服务正常使用。

一般情况下，实例正在被其他服务使用时会被锁定，例如：实例正在被镜像服务使用，制作成私有镜像场景；实例被存储容灾服务使用场景。

## 4.13 手动将实例移入伸缩组时，需要该实例已绑定弹性公网IP 吗？

非必需。您可以将它解绑，也可以不做任何操作。

## 4.14 弹性伸缩（AS）伸缩出来的新实例必须是伸缩组已经关联的伸缩配置里的镜像，而不能复制现有的实例状态吗？

弹性伸缩伸缩出来的新实例必须是伸缩组已经关联的伸缩配置里的镜像（使用新模板创建的或使用已有云服务器规格为模板创建的），除非使用现有的实例规格为模板创建伸缩配置且该伸缩配置已关联到伸缩组，才可复制现有的实例状态。

## 4.15 为 ECS 自动分配的 EIP，是自动新购，还是从未绑定的 EIP 中绑定？

弹性伸缩过程中，为每台弹性云服务器分配的弹性公网IP，是由系统自动新购。



# 5 其他

## 5.1 如何自动部署应用？

为了使伸缩组自动加入的实例自动部署应用，您需要创建私有镜像，确保该镜像上有应用的程序软件和开机自启动设置。为伸缩组选择镜像类型为您的私有镜像的伸缩配置，新实例加入伸缩组后，就可以实现自动部署应用。详细的操作步骤指导如下：

1. 在创建私有镜像前，您需要在源云服务器中安装所需的应用程序，并设置开机自动启动。
2. 创建私有镜像，你可以根据操作系统选择对应的创建私有镜像的方法，操作指导请参考《[镜像服务用户指南](#)》。
3. 创建一个新的伸缩配置，操作可参考[使用新模板创建伸缩配置](#)，确保该伸缩配置选择的镜像为2中创建的私有镜像。
4. 切换到“弹性伸缩组”页签，单击所需的伸缩组名称，进入伸缩组详情页面。
5. 单击“配置名称”右侧的“更换配置”，在弹出的“更换伸缩配置”对话框中，选择3中创建的伸缩配置，单击“确定”。

至此全部设置已经完成，等待下次伸缩活动触发新实例加入伸缩组后，您可以查看新实例是否自动部署了所需的应用，若有问题请联系技术支持处理。

## 5.2 支持 Cloud-Init 特性后，对使用弹性伸缩有哪些影响？

Cloud-init是开源的云初始化程序，能够对新创建弹性云服务器中指定的自定义信息（主机名、密钥和用户数据等）进行初始化配置。在创建伸缩配置时，通过Cloud-Init进行对云服务器的初始化配置。

弹性伸缩组使用的伸缩配置中的私有镜像若没有安装Cloud-Init/Cloudbase-init工具，伸缩活动创建的弹性云服务器会出现如下情况：

- 如使用未安装Cloudbase-init的Windows私有镜像创建的弹性云服务器，在获取弹性云服务器密码时，系统将提示查询不到密码。您只能通过镜像本身的密码登录此云服务器。若忘记镜像本身密码，可以通过云服务器页面“重置密码”功能，自助完成云服务器的密码重置。
- 如使用未安装Cloud-Init的Linux私有镜像创建的云服务器，使用创建时输入的密码或密钥将无法登录云服务器。您只能通过镜像本身的密码或密钥登录此云服务器。若忘记镜像本身密码，或镜像本身的密钥丢失，可以通过云服务器页面“重置密码”功能，自助完成云服务器的密码重置。



- 如使用未安装Cloud-Init/Cloudbase-init的私有镜像创建云服务器时，实例自定义数据注入会失败。  
鉴于出现上述情况，在使用弹性伸缩时，请检查伸缩配置中的私有镜像是否已经安装并配置了Cloud-Init或者Cloudbase-init工具，请将使用了未安装Cloud-Init/Cloudbase-init的私有镜像的伸缩配置删除，并使用已安装Cloud-Init/Cloudbase-init的私有镜像创建新的伸缩配置。具体操作步骤：
  - a. 登录管理控制台。
  - b. 选择“计算 > 弹性伸缩 > 伸缩实例”。
  - c. 选择“伸缩配置”页签，进入伸缩配置列表页面。
  - d. 单击“创建伸缩配置”，选择已安装Cloud-Init或Cloudbase-init工具的私有镜像创建新的伸缩配置。
  - e. 在伸缩组中将伸缩配置修改为新创建的伸缩配置。

## 5.3 为什么使用密钥文件无法正常登录云服务器？

### 问题描述

用户使用密钥文件登录弹性伸缩组中的云服务器时，登录失败。

### 可能原因

该弹性伸缩组使用的伸缩配置中的镜像为用户自己制作的私有镜像，且在创建该私有镜像时用户未安装Cloud-init工具。

创建私有镜像时不安装Cloud-init工具，用户将无法对云服务器进行自定义配置，只能使用镜像原有密码或密钥登录云服务器。

### 处理方法

1. 判断是否需要继续登录该云服务器。
  - 是，请使用镜像原有密码或密钥登录云服务器。  
其中，镜像原有密码或密钥指创建私有镜像时，用户自己设置的操作系统密码或密钥。
  - 否，跳转2。
2. 更换弹性伸缩组的伸缩配置，更换方法请参见[为伸缩组添加/更换伸缩配置](#)。

#### 说明

请确保新伸缩配置中的镜像已安装了Cloud-init/Cloudbase-init工具，Cloud-init/Cloudbase-init工具的安装方法请参见《[镜像服务用户指南](#)》。

更换伸缩配置后，弹性伸缩组进行伸缩活动而新增的弹性云服务器可以直接使用密钥文件正常登录，无需再使用镜像原有密码或密钥登录云服务器。

## 5.4 伸缩组中已经添加了负载均衡，创建伸缩配置时是否还需要配置弹性公网 IP？

伸缩组中已经添加了负载均衡后，伸缩配置可以不配置弹性公网IP。系统会自动将加入伸缩组的实例添加到负载均衡上，伸缩组中的实例统一通过负载均衡绑定的弹性公网IP对外提供服务。

## 5.5 如何自动初始化弹性伸缩新增的云服务器数据盘？

### 操作场景

云服务器创建完成后，数据盘需要初始化后才能使用。当使用弹性伸缩为伸缩组增加数量较多的云服务器时，您就需要逐一手动初始化数据盘，将会占用较长时间。

本节为您介绍通过脚本自动化完成初始化磁盘的操作，包括磁盘分区和挂载指定目录。本节介绍的自动初始化脚本示例仅支持初始化一个数据盘。

本节操作以centos 6.5为例。其他操作系统配置方法略有区别，请参考对应操作系统的相关资料进行操作，文档中不对此进行详细说明。

### 操作步骤

1. 以root用户登录已有云服务器。
2. 执行以下命令，进入脚本存放目录。

```
cd /脚本目录
```

例如：

```
cd /home
```

3. 执行以下命令，创建脚本。

```
vi 脚本名称
```

例如：

```
vi fdisk_mount.sh
```

4. 按“i”，进入脚本编辑页面。

以下脚本为云服务器只有一个数据盘时的自动初始化，仅供参考，请用户根据实际情况修改。

```
#!/bin/bash
bash_scripts_name=fdisk_mount.sh
ini_path=/home/fdisk.ini
disk=
size=
mount=
partition=

function get_disk_from_ini()
{
disk=`cat $ini_path|grep disk| awk -F '=' '{print $2}'`
if [ $disk = "" ]
then
echo "disk is null in file,exit"
exit
fi
result=`fdisk -l $disk | grep $disk`
```

```
if [ $result = 1 ]
then
    echo "disk path is not exist in linux,exit"
    exit
fi
}

function get_size()
{
size=`cat $ini_path| grep size|awk -F '=' '{print $2}'`
if [ $size = "" ]
then
    echo "size is null,exit"
    exit
fi
}

function make_fs_mount()
{
mkfs.ext4 -T largefile $partition
if [ $? -ne 0 ]
then
    echo "mkfs disk failed,exit"
    exit
fi

dir=`cat $ini_path|grep mount |awk -F '=' '{print $2}'`
if [ $dir = "" ]
then
    echo "mount dir is null in file,exit"
    exit
fi

if [ ! -d $dir ]
then
    mkdir -p $dir
fi

mount $partition $dir
if [ $? -ne 0 ]
then
    echo "mount disk failed,exit"
    exit
fi

echo "$partition $dir ext3 defaults 0 0" >> /etc/fstab
}

function remove_rc()
{
cat /etc/rc.local | grep $bash_scripts_name
if [ $? ne 0 ]
then
    sed -i !/$bash_scripts_name/d /etc/rc.local
fi
}

##### start #####
##1、判断配置文件是否存在
if [ ! -f $ini_path ]
then
    echo "ini file not exist,exit"
    exit
fi

##2、获取配置文件中disk指定的设备路径
get_disk_from_ini
```

```
##3、获取配置文件中size分区大小
get_size

##4、将磁盘分区
fdisk $disk <<EOF
n
p
1
1
$size
w
EOF
partition=`fdisk -l $disk 2>/dev/null| grep "^/dev/[xsh].*d" | awk '{print $1}'`

##5、格式化分区，挂载分区到对应目录
make_fs_mount

##6、修改开机启动项，避免重试执行
remove_rc

echo 'SUCCESS'
```

5. 按“Esc”键，输入:wg，按“Enter”保存并退出编辑。
6. 执行以下命令，创建配置文件。

#### vi fdisk.ini

7. 按“i”，进入配置文件编辑页面。  
配置文件中设置了数据盘的盘符、大小和挂载目录信息，用户可参考如下所示信息，根据实际情况修改。

```
disk=/dev/xdev
size=+100G
mount=/opt/test
```

8. 按“Esc”键，输入:wq，按“Enter”保存并退出编辑。
9. 执行以下命令，打开配置文件“rc.local”。

#### vi /etc/rc.local

10. 按“i”，在“rc.local”文件中添加如下内容。

#### /home/fdisk\_mount.sh

配置“rc.local”后，云服务器首次启动时会自动执行初始化数据盘脚本。

11. 按“Esc”键，输入:wq，按“Enter”保存并退出编辑。
12. 通过已有云服务器创建私有镜像。
13. 创建伸缩配置。  
填写伸缩配置信息时，请选择上述步骤中创建的私有镜像，并选择一个数据盘。
14. 创建伸缩组。  
配置伸缩组时，请选择上述步骤中创建的伸缩配置。  
伸缩组创建成功后，通过该组的伸缩活动产生的云服务器，会按照私有镜像中的配置自动初始化数据盘。