

弹性伸缩



产品文档





【版权声明】

©2013-2024 腾讯云版权所有

本文档著作权归腾讯云单独所有,未经腾讯云事先书面许可,任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

【商标声明】



及其它腾讯云服务相关的商标均为腾讯云计算(北京)有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标,依法由权利人所有。

【服务声明】

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况,部分产品、服务的内容可能有所调整。您 所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定,除非双方另有约定,否则, 腾讯云对本文档内容不做任何明示或默示的承诺或保证。



文档目录

扩缩容

生命周期挂钩 管理定时任务

管理告警触发策略

实例健康检查

手动扩容

缩容处理

查看伸缩活动

暂停及恢复扩缩容

指定实例免于缩容

伸缩活动取消

伸缩活动失败

冷却时间说明



扩缩容 生命周期挂钩

最近更新时间:2024-07-17 15:44:52

使用场景

在伸缩组中,您可设置弹性扩容和弹性缩容活动,如果您希望在正式启动这些实例前进行自定义操作,生命周期挂 钩功能就能帮助您完成: 伸缩组扩容出实例后需要延迟一段时间挂载到 CLB,再对外提供服务。 伸缩组释放实例时执行数据备份操作。 伸缩组弹性扩容或者缩容时执行一些用户自定义操作。 **说明:** 生命周期挂钩仅在自动创建或移出实例时生效,伸缩组内其他实例不生效。 如设置了扩展模式,手动添加移出实例、开关机时也会生效。

一个伸缩组内仅可以创建10个生命周期挂钩。

工作方式

伸缩组创建生命周期挂钩后,当生命周期挂钩联动的伸缩活动类型发生时,伸缩活动被挂起,您可以在伸缩活动挂起的期间执行自定义操作。当生命周期挂钩超时时,该挂起会终止。

生命周期挂钩属性

名称	说明	示例
名称	生命周期挂钩名称, 仅支持中文、英文、数字、下划线、分隔符"-"、小数点	fehwnl_
伸缩活动 类型	弹性缩容活动/弹性扩容活动	弹性缩容 活动
扩展活动	扩展活动包括 NORMAL 和 EXTENSION,默认值为 NORMAL。 NORMAL:生命周期挂钩仅在自动创建或移出实例时生效。 EXTENSION:手动添加移出实例、开关机时生命周期挂钩也会生效。	NORMAL
超时时间	默认情况下实例保持等待状态的时间。30-7200(秒),必须为整数。	300



执行策略	执行策略包括 继续 和 拒绝 。 继续策略:挂起的伸缩活动将继续执行。 拒绝策略:针对弹性扩容活动,会直接释放已创建的 CVM 实例。针对弹性缩容活动,无影响。 当一个伸缩组设置多个生命周期挂钩时,则顺序等待操作执行。但仅执行首个拒绝 策略,后续生命周期挂钩的拒绝策略不生效。	拒绝
通知方式	通知方式包括 TDMQ 主题、TDMQ 队列和 TAT 命令。选择通知方式后,您还需要选择具体的 TDMQ 主题、TDMQ 队列或 TAT 命令。	CMQ 主 题
通知标识	弹性伸缩每次推送消息到通知对象时,会同时发送您预先指定的通知标识,便于管 理和标记不同类别的通知信息。	通知消息

创建生命周期挂钩

1. 登录 弹性伸缩控制台。

2. 选择您需要绑定生命周期挂钩的伸缩组,单击伸缩组 ID/名称,进入该伸缩组的详情页面。如下图所示:

伸缩线	且 所有项目 ▼ 广州	~						
	新建							
	ID/名称	状态	当前/期望数量	最小/最大伸缩数	负载均衡	启动配置	所属网络	移出
	asg-3hfrzggq	启用	0 / 0	0 / 0	lb-b4260ewf	asc-f04fpq1a	vpc-ck0a8h51	移出
	asg-gvov3ugi	启用	0 / 0	0 / 10		asc-mn159qw6	vpc-g7wzcv7n	移出

3. 选择**生命周期挂钩**页签,单击**新建**。

4. 在弹出的新建生命周期挂钩对话框中,填写生命周期挂钩相关信息。如下图所示:



名称 *	
	不超过128个字符,仅支持中文、英文、数字、下划线、分隔符-和小数点
伸缩活动类型 *	● 弹性缩容活动 ● 弹性扩容活动
扩展活动	包含实例开/关机、添加/移出实例
超时时间(秒)	- 300 + 范围: 30~7200
执行策略 *	● 继续
通知方式	○ TDMQ队列
	请选择 • 3
	如果没有合适的TDMQ,您可以新建 🖸

说明:

您在创建生命周期时,需要选择或新建 TDMQ。其中包括创建 TDMQ 主题 和 TDMQ 队列。

针对一个伸缩组设置多个生命周期挂钩的情况,顺序等待但只执行第一个拒绝策略,后续生命周期挂钩不生效。 如果您不填写通知方式,默认不通知您。

同一个伸缩组内生命周期挂钩名称不能重复。

修改生命周期挂钩

1. 登录 弹性伸缩控制台。

2. 选择您需要修改生命周期挂钩的伸缩组,单击伸缩组 ID/名称,进入该伸缩组的详情页面。如下图所示:



伸缩	组 所有项目 ▼ 广州	•						
	新建							
	ID/名称	状态	当前/期望数量	最小/最大伸缩数	负载均衡	启动配置	所属网络	利
	asg-3hfrzggq	启用	0 / 0	0/0	lb-b4260ewf	asc-f04fpq1a	vpc-ck0a8h51	移
	asg-gvov3ugi	启用	0/0	0 / 10		asc-mn159qw6	vpc-g7wzcv7n	移

3. 选择**生命周期挂钩**页签,在需要修改生命周期挂钩的行中,单击修改。如下图所示:

< doctest				
伸缩组详情	关联实例 告警触	发策略 定时任务	通知 伸缩活动	生命周期挂钩
当您为伸缩组创建了 以及在实例上预装载	了生命周期挂钩,在伸缩组 欠件,或移出前的数据拷贝	进行伸缩活动时,正在加入 。每个伸缩组中,最多只能	.或正在移出伸缩组的实例将被排 创建 10个生命周期挂钩。	
新建				请输入生命周期挂钩ID/名
ID/名称	伸缩活动类型	超时时间	执行策略	通知方式
ash-om07sc7p lest	弹性缩容活动 基础活动	300 秒	继续	CMQ队列

4. 在弹出的修改生命周期挂钩窗口中,根据实际需求进行修改信息即可。

删除生命周期挂钩

1. 登录 弹性伸缩控制台。

2. 选择您需要删除生命周期挂钩的伸缩组,单击伸缩组 ID/名称,进入该伸缩组的详情页面。如下图所示:



伸缩	伸缩组 所有项目 ▼ 广州 ▼									
	新建									
	ID/名称	状态	当前/期望数量	最小/最大伸缩数	负载均衡	启动配置	所属网络			
	asg-3hfrzggq	启用	0 / 0	0/0	lb-b4260ewf	asc-f04fpq1a	vpc-ck0a8h51			
	asg-gvov3ugi	启用	0 / 0	0 / 10		asc-mn159qw6	vpc-g7wzcv7n			

3. 选择**生命周期挂钩**页签,在需要删除生命周期挂钩的行中,单击**删除**。如下图所示:

← doctest				
伸缩组详情	关联实例 告警	触发策略 定时任	务通知伸缩活	5动 生命周期挂钩
当您为伸缩组创建] 以及在实例上预装载	了 生命周期挂钩,在伸缩 欠件,或移出前的数据拷	组进行伸缩活动时,正在) 坝。每个伸缩组中,最多	加入或正在移出伸缩组的实例 只能创建 10个生命周期挂钩。	将被挂起并置于等待状态, 您可在超时 的
新建				请输入生命周期挂钩ID
ID/名称	伸缩活动类型	超时时间	执行策略	通知方式
ash-om07sc7p test	弹性缩容活动 基础活动	300 秒	继续	СМQ队列

4. 在弹出的**删除生命周期挂钩**窗口中,单击确认即可。



管理定时任务

最近更新时间:2024-01-08 17:53:30

定时任务简介

定时任务,即为设定时间计划,使您的业务根据可预测的负载变化,定时扩展或缩减所使用的云服务器实例数量。 例如,您的Web应用程序的流量会在每周的星期三开始增加,并在星期四保持高流量状态,然后在星期五开始下 降。这种情况下,您可以根据Web应用程序的可预测流量模式来计划扩展活动。

要创建计划的扩展操作,请指定希望扩展操作生效的开始时间,以及用于扩展操作的新的最小大小(最小实例数)、最大大小(最大实例数)和所需大小(期望实例数)。在指定的时间,AS将依据这些设定值来更新伸缩组中的实例数量。

您可以创建仅用于一次扩展的预先计划操作,或者创建用于按经常性计划进行扩展的预先计划操作。

管理定时任务

1. 登录弹性伸缩控制台,选择左侧导航栏中的伸缩组。

2. 选择需修改的伸缩组,单击伸缩组 ID 进入伸缩组基本信息页面。如下图所示:

弹性伸缩	伸缩组 所有项目	▼ 广州 ▼						
∲ 伸缩组	新建							
弐 启动配置	ID/名称	配置建议	状态	当前/期望数量	最小/最大伸缩数	负载均衡	启动配置	所属网络
		无需调整	启用	0/0	0/1	-		
		有2项优化建议立即查看	停用	1/1	0/1	-		
	共2条							

3. 在该伸缩组详情页面,选择定时任务页签,在该页面管理与伸缩组相关联的定时任务。如下图所示:

弹性伸缩	← Manada ana						
☆ 伸缩组	伸缩组详情 关联实例 告警触发策略 定时任务 通知 伸缩活动 监控 生命周期挂钩						
弐 启动配置	新建						
	名称 描述						
	2024-03-11 21:00:00开始,更改伸缩组的最小值为0,期望值为0,最大值为1。每1天重复执行一次,到2025-03-11 21:00:00结束						
	共 1 条						

单击新建可添加新的定时任务。

选择某条定时任务,单击**修改**,可在弹出页面中修改任务名称、调整执行时间、设置是否周期执行、修改执行活动。



单击删除即可删除该条定时任务。

说明:

如果您想创建定时重复的任务,则可以指定开始时间,AS 会在该时间执行操作,然后根据重复计划执行操作。如果您指定结束时间,AS 在该时间后不再执行操作。



管理告警触发策略

最近更新时间:2024-01-08 17:53:29

简介

弹性伸缩 AS 支持根据监控的指标动态扩展伸缩组中的实例数量,您需定义告警触发策略,即触发扩展的监控指标状态以及如何按照需求变化进行扩展。

创建告警策略需指定条件和动作,如下图所示:

新建告警触发的	策略		
名称 *	Hold.		
复制策略(选填)	请选择伸缩组	▼ 请选择 ▼ 复制	
if *	伸缩组内所有实例: CPU利用率 ▼ 1分钟 详细统计规则 译	▼ 最大值 ▼ > ▼	% 连续1次
伸缩组活动 *	增加	台 实例 🔹 冷却	秒 🗊
		确定取消	

条件格式为:某个指标 + 阈值 + 周期 + 连续达到阈值的周期数。即指标在连续 N 个周期都达到了阈值。

执行动作为:发送通知+增加/减少指定数量的云服务器。

建议您为每个伸缩组各创建两个策略:一个策略用于扩展,另一个策略用于收缩。当业务量达到了告警策略指定的 条件后, AS 将执行关联的策略对伸缩组进行收缩(通过终止实例)或扩展(通过启动实例)。

场景示例



例如,您有一个电商网站应用程序,当前使用了5个实例。您做了一个运营活动,担心访问量远大于您的预估,您可以设置当前实例上的负载上升到70%时额外启动2个新的实例,然后在负载下降到40%时终止多余的实例。您可以配置伸缩组,根据这些条件自动扩展。如下图所示:



操作步骤

- 1. 登录弹性伸缩控制台,选择左侧导航栏中的伸缩组。
- 2. 选择需修改的伸缩组,单击伸缩组 ID 进入伸缩组基本信息页面。如下图所示:

弹性伸缩	伸缩组 所有项目 🔻	广州▼						
✿ 伸缩组	新建							
<u> 計</u> 启动配置	ID/名称	配置建议	状态	当前/期望数量	最小/最大伸缩数	负载均衡	启动配置	所属网络
		无需调整	启用	0/0	0/1	-		
		有2项优化建议立即查看	停用	1/1	0/1			
	共 2 条							

3. 在该伸缩组详情页面,选择**告警触发策略**页签,在该页面管理与伸缩组相关联的告警触发策略。如下图所示:



弹性伸缩	← Namalianajin		
♥ 伸缩组	伸缩组详情 关联实例	告警触发策略 定时任务 通知 伸缩活动 监控 生命周期挂钩	
王: 启动配置	新建		ជ័
	名称	描述	
	100100300	如果整个伸缩组实例的CPU利用率在1分钟内的最大值大于60%,且连续发生1次。则实例数增加1台。冷却60秒。	
华建 司沃加莱的	共 1 条 / 十:荷文 슈山 / 十· / 仁仁 町 夕		

单击新建可添加新的告警触发策略。

单击**删除**删除该条告警触发策略。

指定某台服务器不受告警伸缩策略影响

使用 auto scaling 前,也许您的系统已经有常用的服务器,您出于以下考虑,不希望机器被告警伸缩策略移出: 一机多用:集群中某台服务器除了做集群所做的事情外,还兼做其他用途。例如网站建设初期,您的某台服务器既 作为缓存服务器使用,又作为文件服务器。在缓存服务器集群放入伸缩组时,您不希望它被告警伸缩策略移出。 存放数据:该服务器是有状态的或自带其他服务器没有的数据。例如集群中其他服务器运行中产生的增量数据,都 统一保存到该服务器里。

更新镜像/快照:固定使用该服务器定期做镜像和快照

设置方法:

1. 您可以在伸缩组列表里单击服务器所在的伸缩组,进入管理页面。

2. 选择管理页面中的关联实例页签,对所要设置的实例单击设置移出保护。



实例健康检查

最近更新时间:2024-01-08 17:53:30

如果您在新建伸缩组时指定了【起始实例数】,创建启动配置和伸缩组后,伸缩组将新建与起始实例数相等的云服务器实例,同时,伸缩组会确保运行着大于【最小伸缩数】、小于【最大伸缩数】的实例。

注意:

最小伸缩数:伸缩组中允许的实例最小数量。当伸缩组的 CVM 数量小于最小伸缩数时,弹性伸缩 AS 会增加实例, 使得伸缩组当前实例数匹配最小伸缩数。

起始实例数:伸缩组刚创建时的云服务器数量。

最大伸缩数:伸缩组中允许的实例最大数量。当伸缩组的 CVM 数量大于最大伸缩数时,弹性伸缩 AS 会移出实例, 使得伸缩组当前实例数匹配最大伸缩数。

为了保持伸缩组中的实例正常运行, AS 会对伸缩组内实例的运行状况执行定期检查。如果发现实例运行状况不佳, 它将终止该实例,并启动一台新的云服务器实例。

实例健康检查

伸缩组定期检查实例运行状态来确定每个实例是否健壮,判断标准为该机器是否连续1分钟 ping 不可达。如果实例超过1分钟 ping 不可达,则 AS 会标记该实例运行状况不佳。

替换不健康实例

不健康的实例被标记为运行状况不佳之后,伸缩组将立即启动新的实例对它进行替换(设置了**移出保护**的机器除 外)。



手动扩容

最近更新时间:2024-01-08 17:53:30

弹性伸缩(Auto Scaling, AS)除支持根据业务负载自动扩缩容外,还支持您手动介入,达到快速手动扩缩容的效果。您可以通过以下两种方式达到扩容效果: 将已有的 CVM 实例加到伸缩组中 通过修改伸缩组的期望实例数,实现一键扩容

将已有的 CVM 实例添加到伸缩组中

伸缩组为您提供了添加已有实例到现有伸缩组的方式,实现与伸缩组的其他机器一起观察负载和管理的能力。

前提条件

实例处于运行状态。

实例与伸缩组位于同一地域。

实例的网络属性必须与伸缩组一样,即同属基础网络或同属于一个私有网络。

说明事项

AS 会将该组的所需容量与要添加的实例数相加。

例如您伸缩组目前的期望实例数是5,手动增加3台实例后,您伸缩组的期望实例数会变为5+3=8。如果要增加的 实例数加上所需容量超过伸缩组的最大实例数,请求将失败。

伸缩组已关联一个或多个负载均衡(CLB),手动添加的实例会自动注册到伸缩组的所有 CLB 中。

伸缩组缩容时会先移出自动创建的机器,没有自动创建的机器时,才会选择移出手动添加的机器。

伸缩组移出手动添加的实例时,只是将该实例移出伸缩组和 CLB,使实例不再通过伸缩组管理,不会销毁您的实例。

使用控制台手动添加实例

1. 登录伸缩组控制台,单击您要添加实例的伸缩组 ID。

2. 进入伸缩组详情页,选择关联实例 > 添加实例。如下图所示:



弹性伸缩	←	ر								
✿ 伸缩组	伸缩组详情	实例 告警触发策略	定时任务	通知	伸缩活动	监控	生命周期挂钩			
三: 启动配置	添加实例 批量	修出								请输入实例
	实例ID/名称	可用区	监控状态		生命周期		移出保护	加入方式	主IP地址	启动配置
	- 141		健康		运行中		未启用	手动	-	
	共 1 条									

3. 在对话框中勾选对应的实例,单击确定。如下图所示:

加入伸缩组的云服务器有什么要求? 🛽				
站择云服务器			已选择 (1)	
请输入ID/名称	Q		ID/实例名	
— ID/实例名			A 10	
				8
	l	\leftrightarrow		
✓				
转按住 shift 键进行多选				

修改期望实例数, 实现一键扩容

扩容场景



如果您的需求符合以下场景,可执行控制台进行一键扩容,并提前将 CLB 转发规则、机器配置、业务部署这类工作 做好,即使后续您的业务需要扩容,也只需一键修改伸缩组的参数,快速完成扩容。 业务的波峰波谷较难预测,但不愿把扩缩容完全交给系统决定。业务波峰波谷可预测,详情请参见管理定时任务。 您的计算需求是项目性的,且每次用的机器都类似。例如社情舆论收集、基因测序、天气预测等。

在控制台进行一键扩容

执行以下步骤设置 CVM 模板作为启动配置,并配置对应的伸缩组。

1. 创建自定义镜像,详情请参见创建自定义镜像的详细方法。

说明:

后续扩容的实例将依据此镜像部署好环境。

自定义镜像创建的推荐思路:您可选择已有的一台 CVM 或新创建一台 CVM,将您的业务部署好,并将业务设置成 随操作系统一起启动,然后导出为自定义镜像。

2. 基于该自定义镜像创建启动配置,详情请见创建启动配置。

3. 创建伸缩组。

创建时选择已创建的启动配置,最小伸缩数、最大伸缩数、起始实例数根据您需要的服务器数量的下限、上限以及 当前数量来填写。

4. 完成上述步骤后,在业务需要扩容时(例如开始基因测序任务或开通请求类机器收集数据),您可通过修改伸缩 组配置,提高最小伸缩数、最大伸缩数、期望实例数,AS将快速完成扩容。



缩容处理

最近更新时间:2024-01-08 17:53:29

对于每个伸缩组,您可以控制何时向其添加实例(即扩容)或从中删除实例(即缩容)。您可以通过添加或移出实 例,手动扩展组大小,也可以使用扩展策略让弹性伸缩自动执行该过程。

说明:

伸缩组自动缩容时,需要知道哪些实例应首先终止,选择的依据是移出策略。

在缩容时,您可以通过使用实例保护防止弹性伸缩终止特定的实例。

对于已配置负载均衡的伸缩组,在缩容、移出或删除伸缩组内实例时,实例自动与伸缩组关联的负载均衡解除挂载。

移出策略

伸缩组缩容时,会根据移出策略决定移出哪台机器。您可从以下两种移出策略中选择:

移出最旧的实例:删除最早自动增加的机器;自动增加的机器删除完后,删除最早手动增加的机器。

移出最新的实例:删除最新自动增加的机器;自动增加的机器删除完后,删除最新手动增加的机器。

注意:

不管删除最新机器还是删除最旧机器,弹性伸缩都会先删除自动创建的云服务器,然后再删除您手动加入的云服务器。

在控制台设置和修改移出策略

有两种方法设置: 创建伸缩组时,选择您需要的移出策略。 在伸缩组详情页,单击**编辑**,可修改伸缩策略。



查看伸缩活动

最近更新时间:2024-01-08 17:53:30

- 1. 登录弹性伸缩控制台,选择左侧导航栏中的伸缩组。
- 2. 选择需查看的伸缩组,单击伸缩组 ID 进入伸缩组基本信息页面。如下图所示:

弹性伸缩	伸缩组 所有项目 >	▼ 广州 ▼							
✿ 伸缩组	新建								
≕ 启动配置	ID/名称	配置建议	状态	当前/期望数量	最小V最大伸缩数	负载均衡	启动配置	所属网络	移出策略
		无需调整	启用	0/0	0/1	-			移出最旧的实例
		有2项优化建议立即查看	停用	1/1	0/1	-			移出最旧的实例

3. 在该伸缩组详情页面,选择**伸缩活动**页签,即可查看该伸缩组根据伸缩策略已执行过的伸缩活动信息。如下图所示:

弹性伸缩	← Narraillian				
✿ 伸缩组	伸缩组详情 关	联实例 告警触发策时	路 定时任务 通知 伸缩活动 监持	空 生命周期挂钩	
計 启动配置	本月 上月	近60天 近90天	选择日期 选择日期 📩		请输入伸缩
	活动ID	状态	描述	失败原因	
	۶.	成功	81-1000 cm2n	-	
	÷	成功		-	
)	成功	81-1000-010 M	-	
	• · · 1	成功	10000000 Miles	-	
	► c	成功	B000000 .	-	
		r#1 1 44			



暂停及恢复扩缩容

最近更新时间:2024-01-08 17:53:29

使用场景

如果您需要排查配置或与 Web 应用程序相关的其他问题(例如关机重置密码、升级业务等),希望在不触发自动伸缩流程的前提下对应用程序进行更改,那么您可以暂停伸缩组,完成后再恢复。

暂停伸缩组

注意事项

设置了停用伸缩组后,自动触发的活动不会进行。 自动触发的活动包括: 告警伸缩。 定时任务。 健康检查。 手动造成期望实例数不匹配。 停用伸缩组后: 手动加入实例超过最大实例数,则不允许加入。 修改伸缩组最小实例数或最大实例数,不会触发伸缩活动,但修改生效。 手动移出实例不受最小实例数限制。

操作步骤

1.登录弹性伸缩控制台,选择左侧导航栏中的伸缩组。
2.在伸缩组页面中,选择需停用伸缩组所在行右侧的停用,并在弹出窗口中进行确认。
即可查看该伸缩组已处于停用状态。如下图所示:

弹性伸缩	伸缩组 所有项目	▼ 广州 ▼							
� 伸缩组	新建								
<u></u> 启动配置	ID/名称	配置建议	状态	当前/期望数量	最小/最大伸缩数	负载均衡	启动配置	所属网络	移出策略
	×	无需调整	启用	1/1	0/1	-			移出最旧的实例
		有2项优化建议立即查看	停用	1/1	0/1	-			移出最旧的实例

恢复伸缩组



如您已完成暂停伸缩组活动期间的问题排查或操作,您可为业务恢复自动伸缩设置。

- 1. 登录弹性伸缩控制台,选择左侧导航栏中的伸缩组。
- 2. 在**伸缩组**页面中,选择需启用伸缩组所在行右侧的**启用**即可。如下图所示:

弹性伸缩	伸缩组 所有项目。	▼ ┌州 ▼							
� 伸缩组	新建								
≕ 启动配置	ID/名称	配置建议	状态	当前/期望数量	最小/最大伸缩数	负载均衡	启动配置	所属网络	移出策略
	7 2	无需调整	启用	1/1	0/1				移出最旧的实例
	2	有2项优化建议立即查看	停用	1/1	0/1	-			移出最旧的实例
	共 2 条								



指定实例免于缩容

最近更新时间:2024-01-08 17:53:30

简介

在伸缩组中,您可以指定某台子机在缩容活动时不被缩容掉。当缩容活动发生时,弹性伸缩在其他机器中选择要缩 容的子机。

您可以对一个多个伸缩组实例启用实例保护设置,可以随时更改伸缩组或实例保护设置。

如果伸缩组剩下的所有实例都受缩容保护,同时发生缩容事件,则弹性伸缩会减少所需容量,而不会移出实例。

适用场景

通常情况下,伸缩组的机器都是无状态的,所有的机器都可以随时被移走。但在实际实践中,有以下情况适用将指 定实例设置免于缩容:

一机多用:基于成本考虑,个别机器除做集群中的事情外,还兼作其他用途。例如,存储集群中产生的数据,那么 这台机器实际上是有状态的。

避免误操作:若担心策略设置错误影响业务,则可以对部分机器设置"免于缩容",这样 弹性伸缩永远不会缩容该机器,"请求-LB-子机"的通路可以保持畅通。

操作步骤

1. 登录弹性伸缩控制台,选择左侧导航栏中的伸缩组。

2. 在伸缩组页面中,选择需进行设置的伸缩组 ID,进入该伸缩组详情页面。

3. 选择关联实例页签,并单击需设置免于缩容实例所在行右侧的设置移出保护。如下图所示:

弹性伸缩	← Harrailla	-										
� 伸缩组	伸缩组详情	关联实例	告警触发策略	定时任务	通知	伸缩活动	监控	生命周期挂钩				
弐 启动配置	添加实例	批量移出									请输入实例	ID
	实例ID/名称		可用区	监控状态		生命周期		移出保护	加入方式	主IP地址	启动配置	1
			广州三区	健康		运行中		未启用	手动		-	:
	共 1 条											2

4. 在弹出提示框中单击确定即可完成设置。



伸缩活动取消

最近更新时间:2024-01-08 17:53:30

伸缩活动取消是指,定时任务时间到或者告警伸缩的条件达到,伸缩活动被触发,但是存在冲突,伸缩活动被迫取消。

冲突原因:

有进行中的伸缩活动。

伸缩组处于冷却时间中。

伸缩活动取消后是否会重试?

告警伸缩 活动如果取消,不会再重试。但是如果告警伸缩的条件继续成立,会触发下一次告警伸缩活动。

定时任务 定义的是期望实例数、最大伸缩数、最小伸缩数,所以伸缩组会一直重试,使实际存在的实例数符合期望 实例数。

注意:

伸缩组被暂停,伸缩组会直接不尝试伸缩活动,所以在"伸缩活动"记录中,不会留下取消的伸缩活动。



伸缩活动失败

最近更新时间:2024-01-08 17:53:29

伸缩活动取消 是符合预期的,伸缩活动失败 则是不符合预期的。

如何查看失败的伸缩活动?

您可查看 伸缩活动详情。 要第一时间知道伸缩活动失败,您可配置通知策略。

为什么会发生失败的伸缩活动?

我们已经将伸缩活动的原因做好归类,请查阅失败原因归类 >>



冷却时间说明

最近更新时间:2024-01-08 17:53:30

什么是冷却时间

弹性伸缩(AS)冷却时间是伸缩组的一个可配置设置,设置冷却时间,可以确保在上一扩展活动生效前 AS 不会启 动或终止其他实例。伸缩组使用简单的扩展策略动态扩展后,AS 会等待冷却时间完成,然后再继续扩展活动。 手动扩展伸缩组时,默认为不等待冷却时间,但您可以设置冷却时间覆盖默认设置。请注意,如果监测出实例运行 状况不佳,AS 会即时替换运行状况不佳的实例,而不会等待冷却时间完成。

为什么需要冷却时间

机器加入伸缩组后,需要一段时间才能将负载降下来。如果没有冷却时间,系统会在负载降下来前不断扩容,新加入的机器接管业务后,发现负载过低,然后又缩容。

在实例投入使用之前,这些实例使用配置脚本安装和配置软件,因此实例从启动到投入使用大约需要两到三分钟的 时间。(当然,实际时间取决于诸多因素,如实例大小和是否有启动脚本要完成等。)

示例场景:

业务出现流量高峰,导致告警策略的警报触发。该警报触发时,AS 会启动一个实例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,并且启动后需要时间逐渐从 CLB 接收请求。在此期间,监控警报可能会继续触发,从而导致 AS 在警报每次出现时都另外启动一个实例。

但若您设置了冷却时间,AS在启动一个实例后,将暂停所有简单扩展策略或手动扩展引起的扩展活动,直至经过了 该指定时间量(默认值为60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,所有暂停的扩展操作都会恢复。如果警报再次触发,则 AS 将启动另一个实例,而冷却时间也会再次 生效。不过,如果新增的实例足以将 CPU 使用率降为正常水平,则该组会保持其当前大小。

设置冷却时间

默认的冷却时间为60秒。

如需修改,请按以下步骤进行:

打开**伸缩组**的详情页。

单击**告警触发策略**,选择要设置的告警伸缩策略,选择**修改**,在修改框下方指定冷却时间的时长(可设置为0-999999秒)。