弹性伸缩服务

最佳实践

文档版本 02

发布日期 2024-08-20

华为云计算技术有限公司

商标声明


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云计算技术有限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，华为云计算技术有限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。

目 录

1 搭建可自动伸缩的 Discuz!论坛网站

应用场景

Discuz!论坛是全球成熟度最高、覆盖率最大的论坛软件系统之一。用户对论坛的访问可分为高峰期和平峰期，若论坛采用多服务器部署模式且满足高峰时期的负载需求，平峰期必有部分服务器处于闲置状态，增加了不必要的成本，也造成了资源浪费。

弹性伸缩可帮助您解决以上问题。当您在论坛的服务器系统中应用弹性伸缩后，弹性伸缩可以根据您设定的策略，自动地增加或减少服务器的数量，在保证您的网站正常运转的同时节约成本。本实践以搭建可自动伸缩的Discuz!论坛为例，介绍了如何使用弹性伸缩服务搭建一个可自动增加或减少弹性云服务器数量的Web服务。

方案介绍

为了实现创建可自动伸缩的Discuz!论坛，您需要按照**表1-1**中的步骤进行网站的搭建，本文重点介绍创建弹性伸缩实现云服务器自动伸缩的过程。当网站的负载增加时云服务器的CPU使用率会增大，负载降低时CPU使用率会降低。配置两条监控CPU使用率的告警策略，分别在CPU使用率高于70%时增加一台云服务器，在CPU使用率低于30%时减少一台云服务器，保证Discuz!论坛始终有合适数量的云服务器，实现自动伸缩云服务器的功能。

表 1-1 搭建 Discuz!论坛步骤

任务	分类	子任务描述	说明
搭建网站	申请服务	申请虚拟私有云	申请为云服务器提供网络服务的虚拟私有云vpc-DISCUZ。
		购买弹性公网IP	需申请使云服务器和互联网互通的弹性公网IP。
		创建安全组并添加规则	为了保证论坛的网络安全，需要设置安全组对网络访问进行控制。创建的安全组sg-DISCUZ。

任务	分类	子任务描述	说明
		购买弹性云服务器	需要购买两台弹性云服务器，云服务器discuz01用于部署论坛数据库，discuz02用于部署论坛业务。购买云服务器discuz01时绑定之前购买的弹性公网IP，discuz02暂不绑定弹性公网IP。
	配置服务器	在discuz01上搭建数据库	在discuz01上安装云数据库 RDS for MySQL，启动RDS for MySQL，设置开机自启动。
		在discuz02上部署网站代码	先将discuz01上的弹性公网IP解绑，再绑定至discuz02，在discuz02上部署Web环境和网站代码。
	配置特性	释放弹性公网IP	为了节省弹性公网IP资源，使用负载均衡服务前请先释放discuz02绑定的弹性公网IP。
		创建弹性负载均衡	为了在伸缩组中均衡访问网站的流量，需要购买增强型负载均衡监听器elb-DISCUZ。
		制作镜像	为了后续增加的云服务器可以自动搭建Web环境和部署网站代码，需要制作discuz02的镜像discuz_centos6.5(40GB)，该镜像在创建伸缩配置时作为私有镜像使用。
创建弹性伸缩	-	创建伸缩配置	伸缩配置是伸缩组内实例（弹性云服务器）的模板，定义了伸缩组内待添加的实例的规格数据。创建伸缩配置as-config-discuz。
		创建伸缩组	伸缩组是云服务器进行伸缩的基本单位，伸缩活动将会以伸缩组为单位进行。创建弹性伸缩组as-group-discuz。
		创建伸缩策略	伸缩策略能够触发伸缩活动，配置两条监控CPU使用率的告警策略，在业务负载增加时增加云服务器数量，在业务负载减少时减少云服务器数量。
		手动移入实例	为保证discuz02可以和后续移入伸缩组中的服务器共同承载论坛业务，需要将discuz02手动移入伸缩组。
		修改最小实例数	最小实例数定义了伸缩组中云服务的最少数量，修改最小实例数为1后，伸缩组至少会保证有一台云服务器。discuz02是手动移入，在实例移除策略中被移出的优先级最低，故修改最小实例数可以保证discuz02在伸缩组中不被移出。
访问网站	验证配置结果	验证网站是否可以正常访问	获取负载均衡服务的弹性公网IP地址，在浏览器中输入http://*弹性公网IP地址*/forum.php进行验证。若可以访问则说明各项配置已生效。

前期准备

请您参考《**搭建Discuz!论坛网站**》完成**表1-1**中搭建网站部分的任务。

创建伸缩配置

伸缩配置定义了移入伸缩组的云服务器的规格，为了移入伸缩组的云服务器能自动承载业务，使用镜像discuz_centos6.5(40GB)，并使伸缩配置中的参数和discuz02保持一致。

1. 登录管理控制台，选择“计算 > 弹性伸缩”。
2. 在“伸缩实例”页面，单击“创建伸缩配置”。
   参考**表1-2**进行关键参数配置，未列出的参数选择默认值即可。

图 1-1 伸缩配置参数

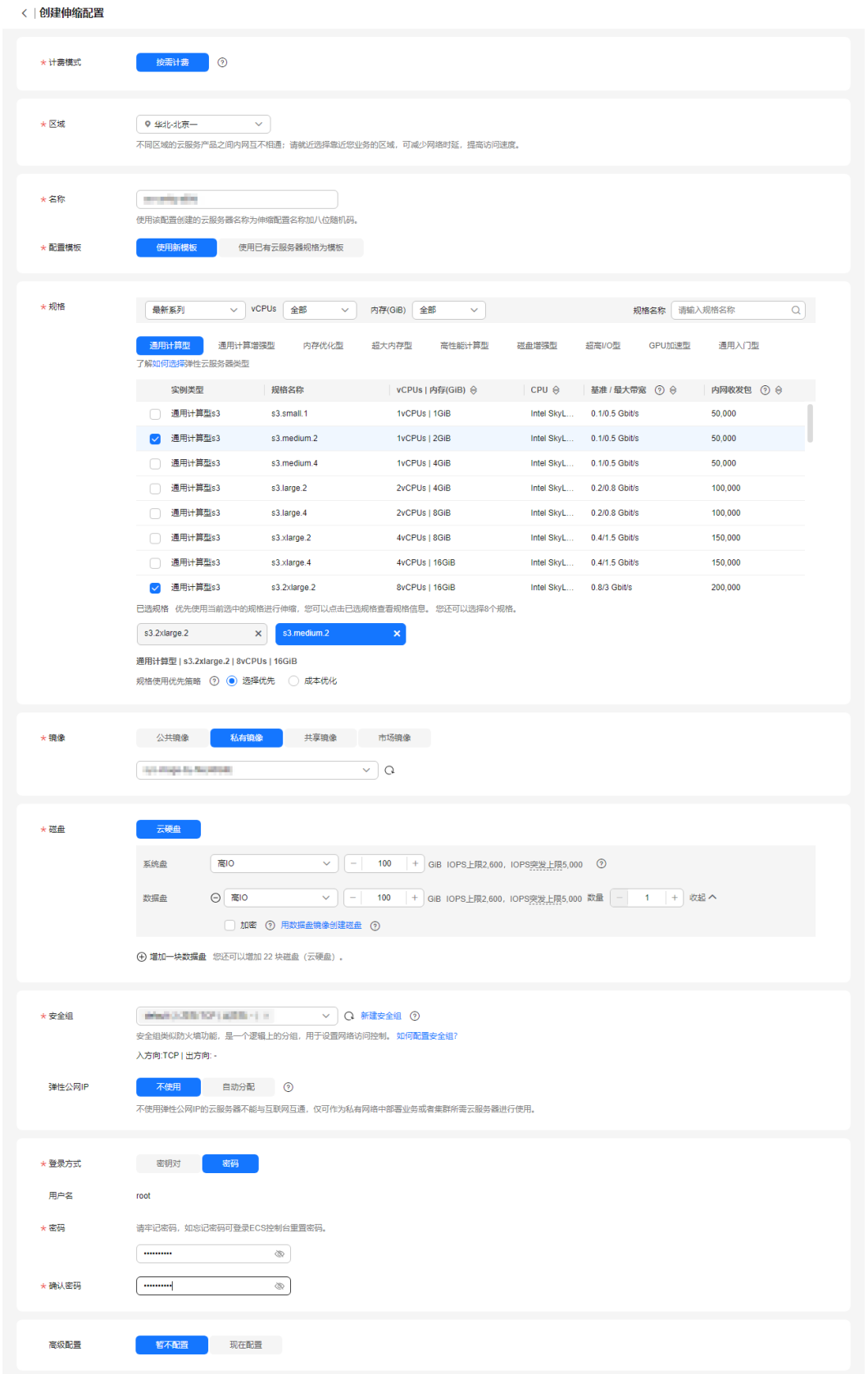


表 1-2 伸缩配置关键参数

参数	解释	取值样例
配置模板	选择“使用新模板”，重新选择云服务器类型、vCPUs、内存、镜像、磁盘等参数信息，创建新的弹性伸缩配置。	使用新模板
规格	可以选择多个规格，避免在伸缩时规格售罄的风险。规格使用优先策略包括“选择优先”和“成本优先”，请根据需要选择。	s3.medium.2 s3.large.2
镜像	为伸缩组中移入的实例提供软件和系统应用配置的模板，选择私有镜像discuz_centos6.5(40GB)。	私有镜像 discuz_centos6.5(40GB)
磁盘	为伸缩组中的移入的实例提供存储和存储管理功能。	系统盘 高IO 40GB 数据盘 高IO 100GB
安全组	安全组是一个逻辑上的分组，用来实现安全组内和组间弹性云服务器的访问控制，加强弹性云服务器的安全保护。选择安全组sg-DISCUZ。	sg-DISCUZ
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。	不使用

3. 伸缩配置参数配置完成后，单击“立即创建”。

创建伸缩组

1. 单击“创建弹性伸缩组”。
   参考**表1-3**进行关键参数配置，未列出的参数选择默认值即可。

图 1-2 设置伸缩组参数

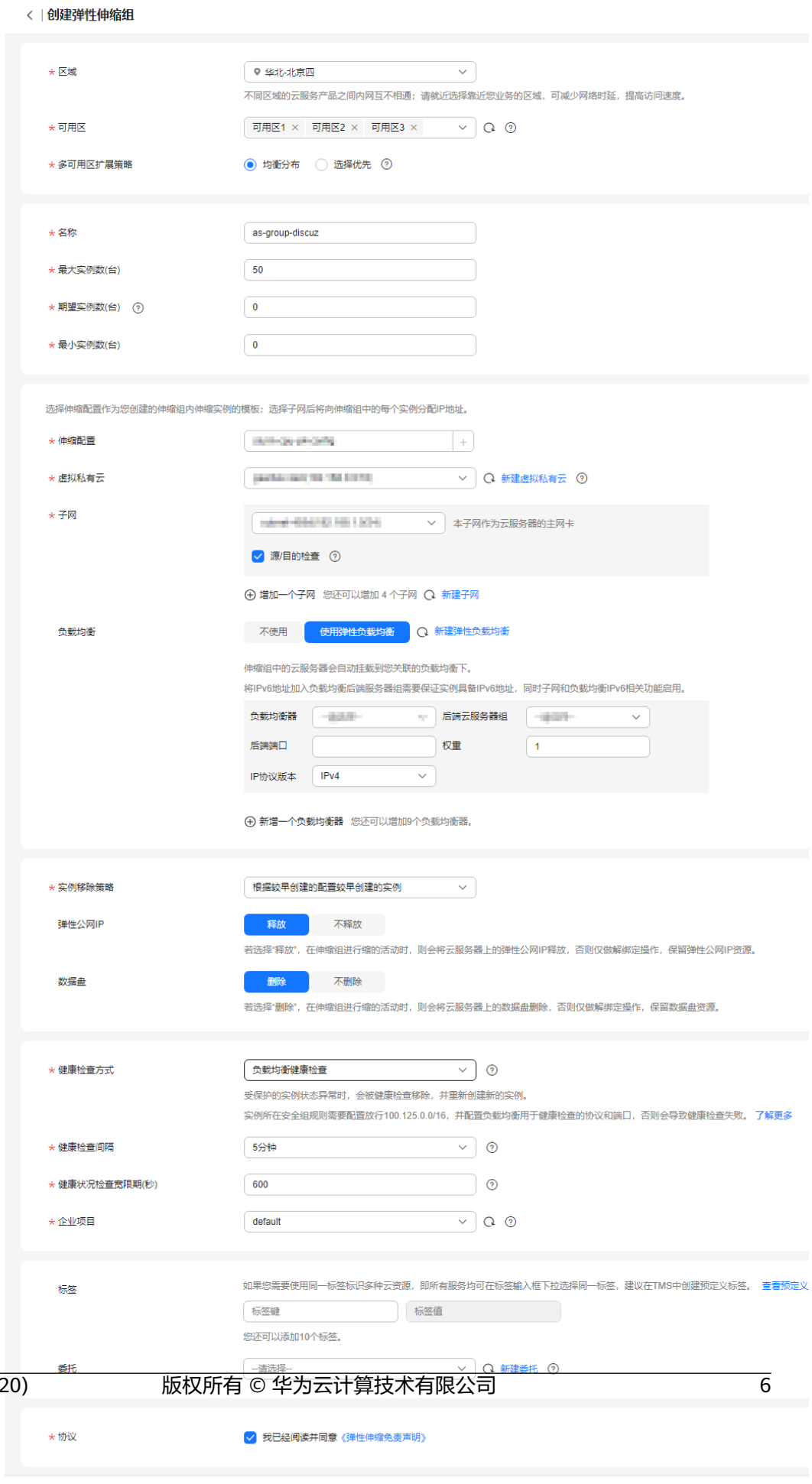


表 1-3 伸缩组关键参数

参数	解释	取值样例
最大实例数	伸缩组中弹性云服务器数量的最大值。	50
期望实例数	伸缩组中期望的云服务器数量，本实践中要将搭建Discuz!论坛的云服务器手动移入，为避免移入前发生伸缩活动，将期望实例数设置为0。	0
最小实例数	伸缩组中弹性云服务器数量的最小值。	0
虚拟私有云	为伸缩组中的实例提供所使用的网络。必须和云服务器discuz02属于同一VPC。	VPC-DISCUZ
子网	子网可以方便您管理vpc中的网络。选择中申请虚拟私有云时创建的子网。	vpc-test
负载均衡	为伸缩组中的实例均分流量，选择增强型负载均衡器elb-DISCUZ。后端端口配置为需要监听的业务端口，示例中配置为80，权重为1。	使用增强型
健康检查方式	健康检查方式选择“负载均衡健康检查”，负载均衡健康检查是通过系统向后端云服务器发起心跳检查的方式来实现的，推荐使用该方式。	负载均衡健康检查

2. 参数配置完后，单击“立即创建”。
3. 返回弹性伸缩组列表，若伸缩组为“已启用”状态，说明伸缩组创建成功。

创建伸缩策略

为了能实现云服务器的自动伸缩，配置两条监控CPU使用率的告警策略，在业务负载上升时增加云服务器数量的策略as-policy-discuz01，在业务负载降低时减少云服务器数量的策略as-policy-discuz02。

1. 在已创建的弹性伸缩组“as-group-discuz”所在行，单击操作列的“查看伸缩策略”。
2. 单击“添加伸缩策略”。

   参考**表1-4**配置伸缩策略as-policy-discuz01的参数，当系统连续3次监控到CPU使用率超过70%时，触发伸缩策略as-policy-discuz01，伸缩组会增加一台弹性云服务器。

图 1-3 伸缩策略 as-policy-discuz01 参数

表 1-4 伸缩策略 as-policy-discuz01 关键参数

参数	解释	取值样例
策略名称	创建伸缩策略的名称。	as-policy-discuz01
策略类型	选择“告警策略”。	告警策略
告警规则	可选择“现在创建”或“使用已有”。	现在创建
告警规则名称	新建告警规则的名称。	as-alarm-cpu-01
监控类型	定义监控指标的类型，是系统支持的或是自定义的。选择“系统监控”。	系统监控

参数	解释	取值样例
触发条件	选择弹性伸缩支持的监控指标并对监控指标设定告警条件。	CPU使用率最大值 > 70%
监控周期	告警规则刷新告警状态的周期。	5分钟
连续出现次数	触发告警时的采样点数目。	3
执行动作	设置伸缩活动执行动作及实例的个数或实例百分比。 执行动作包括： ● 增加 当执行伸缩活动时，向伸缩组增加实例。 ● 减少 当执行伸缩活动时，从伸缩组中减少实例。 ● 设置为 将伸缩组中的期望实例数设置为固定值。	增加1个实例
冷却时间	为了避免告警策略频繁触发，必须设置冷却时间。	900

3. 单击“确定”。
4. 再次单击“添加伸缩策略”，配置伸缩策略as-policy-discuz02的参数，当系统连续3次监控到CPU使用率低于30%时，触发伸缩策略as-policy-discuz02，伸缩组会减少一台弹性云服务器。

图 1-4 伸缩策略 as-policy-discuz02 参数

5. 单击“确定”。
6. 返回伸缩策略列表页面，若伸缩策略为“已启用”状态，说明伸缩策略创建成功。

手动移入实例

手动将云服务器discuz02移入伸缩组。

1. 单击伸缩组as-group-discuz名称进入伸缩组详情页面。
2. 切换到“伸缩实例”页签，将discuz02手动移入伸缩组中。

修改最小实例数

为保证discuz02不被伸缩活动移出伸缩组，需修改伸缩组的最小实例数。

1. 单击伸缩组as-group-discuz名称，进入伸缩组详情页面。
2. 单击页面右上角的“修改伸缩组”。修改最小实例数为1。

图 1-5 修改最小实例数

3. 修改完成后，单击“确定”。

结果验证

若论坛可以正常使用，当伸缩组中的云服务器CPU使用率持续高于70%（在伸缩组的“监控”页签可对监控指标进行观察），伸缩组会自动增加一台云服务器（在伸缩组的“活动历史”页签可对伸缩活动历史进行查看）。当伸缩组中的云服务器CPU使用率持续低于30%，且伸缩组中至少存在两台云服务器时，伸缩组会自动减少一台云服务器，则本次实践是成功的。若不然，请联系技术支持定位伸缩组不能正常进行伸缩活动的原因。

实践扩展

- 当应用场景有变化，需要在云服务器上部署新的软件时，可使用弹性伸缩的生命周期挂钩功能，在实例加入和移出伸缩组时进行自定义操作，灵活的管理加入或移出弹性伸缩组的实例。具体操作可参见**生命周期挂钩**。
- 当所需的弹性云服务器的规格变更时，可创建新的伸缩配置，操作可参考**使用新模板创建伸缩配置**。创建完成后，可参考**为伸缩组更换伸缩配置**，即可改变伸缩组新加入实例的规格。

2 结合自定义监控配置伸缩组的告警策略

弹性伸缩进行伸缩活动时，需定义如何按照不断变化的需求进行伸缩活动，即动态扩展资源。当业务需求变化频繁且无固定规律时，可通过配置告警策略实现动态扩缩资源的目的。

当您在CES（云监控服务）中没找到对应的监控指标，无法根据现有的监控指标创建告警策略时，您可以使用自定义监控将自身产品关心的业务指标上报至CES中，在CES界面根据上报的监控数据来创建对应的告警规则，然后配置伸缩组的告警策略。

您可以参考本文步骤结合自定义监控配置伸缩组的告警策略，更好的实现动态扩缩资源。

前提条件

- 已创建弹性云服务器
- 已创建弹性伸缩组

步骤一：上报监控数据

准备工作

1. **获取AK/SK**
2. **获取project_id**
3. **获取华为SDK**

上报监控数据

您可以在**API Explorer**中调试接口，上报自定义监控数据，详细参数说明请参见**添加监控数据**。API Explorer可以自动生成SDK代码示例，并提供SDK代码示例调试功能，您可以参考**华为SDK的参考示例**上报自定义监控指标数据。

例如添加自定义命名空间为MINE.APP，自定义监控维度instance_id，值为资源ID，根据实际情况填写，此处示例为0dc2f23c-6e63-4162-b59b-5f554d492655，上报云服务器CPU使用率cpu_util的数据，请求示例如下所示。

```
[{
    "metric": {
        "namespace": "MINE.APP",
        "dimensions": [
            {
                "name": "instance_id",
                "value": "0dc2f23c-6e63-4162-b59b-5f554d492655"
```

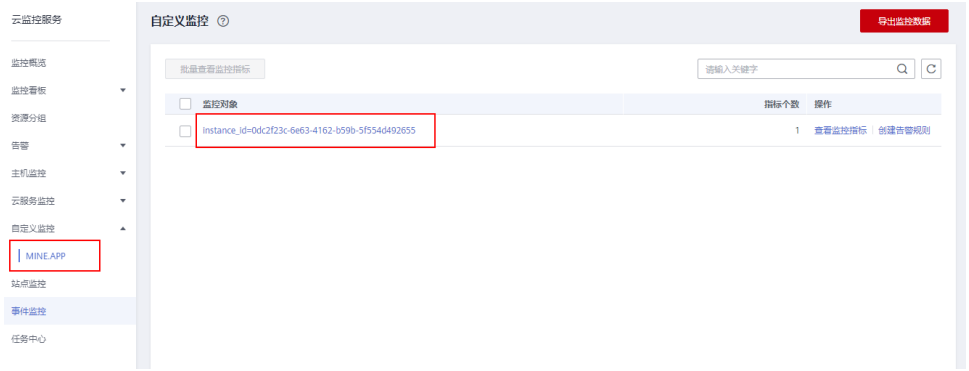


```
            }
        ],
        "metric_name": "cpu_util"
    },
    "ttl": 172800,
    "collect_time": 1695872430398,
    "type": "float",
    "value": 0.09,
    "unit": "%"
}]
```


查看监控数据

1. 登录管理控制台。
2. 单击“服务列表 > 云监控服务”。
3. 单击页面左侧的“自定义监控”。
4. 在“自定义监控”页面，可以查看当前用户通过**添加监控数据**接口上报至云监控服务的相关数据，包括自定义上报的服务，指标等。

图 2-1 自定义监控指标



5. 选择待查看的云服务资源所在行的“查看监控指标”，进入“监控指标”页面。

   在这个页面，用户可以选择页面左上方的时间范围按钮，查看该云服务资源“近1小时”、“近3小时”、“近12小时”、“近24小时”和“近7天”的监控原始数据曲线图，同时监控指标视图右上角会动态显示对应时段内监控指标的最大值与最小值。

图 2-2 查看监控指标

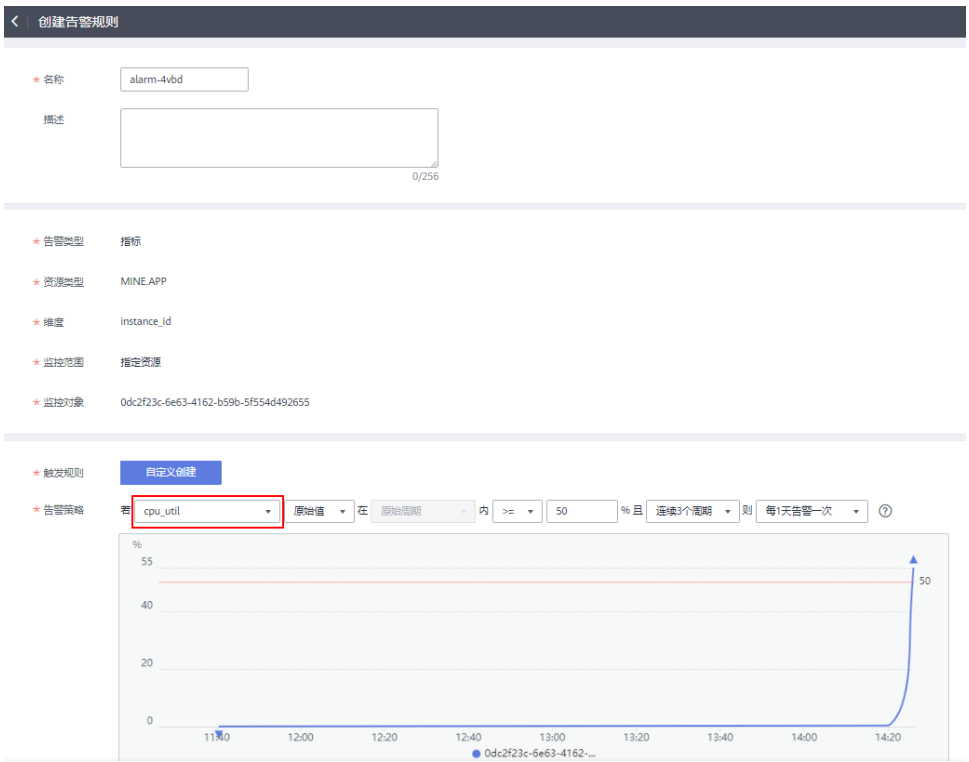
步骤二：创建告警规则

1. 登录管理控制台。
2. 单击“服务列表 > 云监控服务”。
3. 单击页面左侧的“自定义监控”。
4. 在“自定义监控”页面，单击待创建的云服务资源所在行的“创建告警规则”。

图 2-3 创建告警规则

5. 在“创建告警规则”页面，根据界面提示配置参数，具体参数说明请参见表1-表3。
例如下图中创建告警规则alarm-4vbd，告警策略选择自定义监控指标cpu_util。

图 2-4 配置告警规则



6. 单击“立即创建”，完成告警规则的创建。

图 2-5 告警规则列表

步骤三：配置伸缩策略

1. 登录管理控制台。
2. 选择“计算 > 弹性伸缩 > 伸缩实例”。
3. 在伸缩组所在行的“操作”列下，单击“查看伸缩策略”。
4. 在“伸缩策略”页签，单击“添加伸缩策略”。
5. 根据界面进行参数配置。详细参数说明请参见**表1**。

   例如下图选择使用已有的告警策略，即**步骤二：创建告警规则**中创建的告警规则alarm-4vbd。

6. 单击“确定”。

   在“伸缩策略”页签中可查看新创建的伸缩策略，新创建的伸缩策略默认的状态为“已启用”。

结果验证

若您的业务正常上报自定义监控指标cpu_util数据，当数据连续3次超过50%时，会触发告警策略执行伸缩活动，即增加一个实例至伸缩组（在伸缩组的“活动历史”页签可对伸缩活动历史进行查看），则本次实践是成功的。若不然，请联系技术支持定位伸缩组不能正常进行伸缩活动的原因。

3 AS&FunctionGraph 支持优雅关机

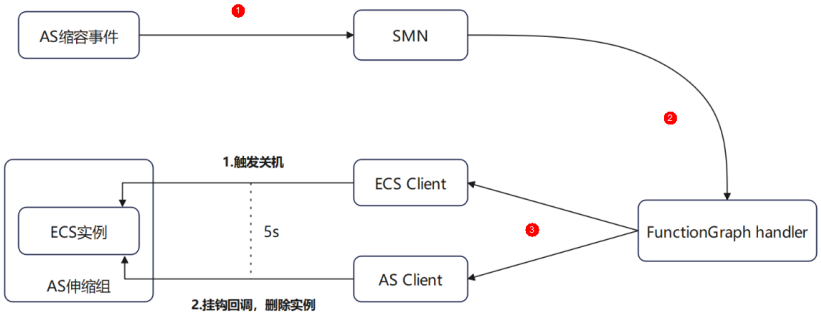
3.1 案例概述

在弹性伸缩服务中，伸缩组的实例缩容过程期间，先对实例进行关机，待客户关机后的清理工作完成后，再由弹性伸缩继续移除并删除对应实例，达到优雅关机的效果。

应用场景

1. 通过配置AS缩容事件的消息通知，转发缩容消息至SMN消息通知服务。
2. 再通过函数工作流服务，接收SMN通知转发过来的伸缩组的缩容消息，经过自定义函数获取伸缩实例等信息。
3. 调用ECS服务对缩容实例进行关机操作。关机操作结束后，等待5秒（这里可以根据业务需要进行修改），再调用AS服务继续伸缩组的缩容操作，对实例进行删除。

图 3-1 AS 优雅关机流程图



3.2 准备

1. 获取弹性伸缩服务优雅关机的程序包。
2. 创建委托ASOperation，添加“ECS FullAccess”以及“AutoScaling FullAccess”权限，具体详情请参考**创建委托**。

3. 请在工单系统中提交工单，申请在弹性伸缩服务（AS）中开通配置消息通知白名单。

> **说明**
>
> 提交工单时需要用户提供项目ID，获取项目ID方法请参见**获取项目ID**。

创建委托

1. 登录**统一身份认证服务控制台**。
2. 在统一身份认证服务的左侧导航栏中，选择“委托”页签，单击右上方的“+创建委托”。

图 3-2 创建委托

3. 开始配置委托。
   - 委托名称：ASOperation。
   - 委托类型：选择“云服务”。
   - 云服务：选择“函数工作流 FunctionGraph”。
   - 持续时间：选择“永久”。
   - 描述：填写描述信息。
4. 单击“下一步”，进入委托选择页面，在右侧搜索框中搜索“ECS FullAccess”与“AutoScaling FullAccess”权限并勾选。
5. 单击“下一步”，请根据业务需要选择权限的作用范围。
6. 单击“确定”，完成权限委托设置。

3.3 构建程序

本例提供了支持优雅关机功能的程序，使用空白模板创建函数，用户可以学习使用。

创建函数

1. 登录**函数工作流控制台**，在左侧导航栏选择“函数 > 函数列表”，进入函数列表界面。
2. 单击“创建函数”，进入创建函数流程。
3. 选择“创建空白函数”，填写函数配置信息。输入基础配置信息，完成后单击“创建函数”。
   - 函数类型：事件函数。
   - 函数名称：输入“as_graceful_shutdown”。
   - 委托名称：选择**创建委托**中创建的“ASOperation”。
   - 运行时语言：选择“Python3.6”。
4. 进入as_graceful_shutdown函数详情页，在“代码”页签，代码选择“上传自ZIP文件”，将如下程序保存成扩展名为“.py”的文件格式，压缩成zip包，并上传。然后，在“设置 > 常规设置”页签，设置如下信息，完成后单击“保存”。

```
# -*- coding:utf-8 -*-
import json
import time

from huaweicloudsdkas.v1 import AsClient
from huaweicloudsdkas.v1 import AttachCallbackInstanceLifeCycleHookRequest
from huaweicloudsdkas.v1 import CallbackLifeCycleHookOption
from huaweicloudsdkas.v1.region.as_region import AsRegion
from huaweicloudsdkecs.v2.region.ecs_region import EcsRegion
from huaweicloudsdkcore.auth.credentials import BasicCredentials
from huaweicloudsdkcore.exceptions import exceptions
from huaweicloudsdkecs.v2 import EcsClient
from huaweicloudsdkecs.v2 import ServerId
from huaweicloudsdkecs.v2 import BatchStopServersOption
from huaweicloudsdkecs.v2 import BatchStopServersRequest
from huaweicloudsdkecs.v2 import BatchStopServersRequestBody


def handler(event, context):
    # 用户配置需要在FuctionGraph前台配置处手动加入环境变量
    # Need to configure environmental variables at FuctionGraph console page
    # 客户的project_id
    # your project_id
    # region为华为云的局点名，例如cn-north-4
    # Huawei region alias，like cn-north-4 for Beijing4
    # 客户账户名与密码ak/sk
    # your ak/sk code
    project_id = context.getUserData('projectId', '').strip()
    region = context.getUserData('region', '').strip()
    ak = context.getAccessKey().strip()
    sk = context.getSecretKey().strip()

    if not project_id:
        raise Exception("'project_id' not configured")

    if not region:
        raise Exception("'region' not configured")

    if not ak or not sk:
        ak = context.getUserData('ak', '').strip()
        sk = context.getUserData('sk', '').strip()
        if not ak or not sk:
            raise Exception("ak/sk empty")

    logger = context.getLogger()
    logger.info("get incoming scaling activity, %s", event)

    credentials = BasicCredentials(ak, sk).with_project_id(project_id)

    # 从SMN传入的event中获取ECS虚拟机的实例ID与伸缩组ID信息
    # Get ECS instance ID and scaling group ID from SMN event message
    instance_id, group_id, hook_name = get_event_info(event, logger)
    if not instance_id or not group_id or not hook_name:
        logger.info("no need to perform gracefully shutdown op")
        return

    # 将获取到的实例ID进行关机操作
    # Execute soft shutdown given the instance ID
    stop_ecs_instance(credentials, instance_id, logger, region)
    logger.info("finish stop op")

    # 等待实例关机等业务后置逻辑的执行
    # Waiting for the workload after the instance shutdown
    time.sleep(5)

    # 关机完成后，调用AS的生命周期挂钩回调接口，执行继续操作，避免等待生命周期设置的超时时间
    # After everything is done, start to trigger callback interface for life cycle hook,
    # directly remove the instance rather than waiting for life cycle hook timeout
    execute_hook_callback(credentials, instance_id, group_id, hook_name, logger, region)
```

```
    logger.info("finish hook callback op")


def get_event_info(event, logger):
    record = event.get("record")[0]
    message = record.get("smn").get("message").replace("\"{", "{").replace("}\"", "}")
    parsed_message = json.loads(message)
    if parsed_message.get("lifecycle_hook_type") is None:
        # 伸缩组对接SMN后，伸缩活动、伸缩组异常，生命周期都会发送消息给生命周期都会发送消息给
FunctionGraph
        # 这里只处理由生命周期挂钩终止实例发来的事件，也可以在console上只配置生命周期-终止实例的
消息通知
        # After connecting to SMN, the scaling group activity will be sent to FunctionGraph
        # here only deal with the message sent from group_lifecycle_hook
        logger.info("current op is not triggered by life cycle")
        return None, None, None

    if parsed_message.get("lifecycle_hook_type") != "INSTANCE_TERMINATING":
        # 本示例展示实例删除前的优雅关机操作，这里的动作需识别为实例删除
        # This example is the showcase of graceful shutdown before instance remove,
        # so here need to identify as INSTANCE_TERMINATING
        logger.info("current op is not instance terminate")
        return None, None, None

    return parsed_message.get("scaling_instance").get("instance_id"),
parsed_message.get("scaling_group").get("scaling_group_id"),
parsed_message.get("lifecycle_hook_name")


def stop_ecs_instance(credentials, instance_id, logger, region):
    ecs_client = EcsClient.new_builder() \
        .with_credentials(credentials) \
        .with_region(EcsRegion.value_of(region)) \
        .build()

    try:
        logger.info(f"fulfill ECS request with instance_id:{instance_id}")
        request = BatchStopServersRequest()
        list_servers_os_stop = list()
        list_servers_os_stop.append(ServerId(id=instance_id))
        os_stop_body = BatchStopServersOption(servers=list_servers_os_stop, type="SOFT")
        request.body = BatchStopServersRequestBody(os_stop=os_stop_body)
        ecs_client.batch_stop_servers(request)
    except exceptions.ClientRequestException as e:
        logger.error(e.status_code)
        logger.error(e.request_id)
        logger.error(e.error_code)
        logger.error(e.error_msg)


def execute_hook_callback(credentials, instance_id, group_id, hook_name, logger, region):
    as_client = AsClient.new_builder() \
        .with_credentials(credentials) \
        .with_region(AsRegion.value_of(region)) \
        .build()

    try:
        logger.info(f"fulfill AS request with instance_id:{instance_id},group_id:{group_id},hook_name:
{hook_name}")
        request = AttachCallbackInstanceLifeCycleHookRequest()
        request.scaling_group_id = group_id
        request.body = CallbackLifeCycleHookOption(
            lifecycle_action_result="CONTINUE", instance_id=instance_id,
lifecycle_hook_name=hook_name)
        as_client.attach_callback_instance_life_cycle_hook(request)
    except exceptions.ClientRequestException as e:
        logger.error(e.status_code)
        logger.error(e.request_id)
```

```
    logger.error(e.error_code)
    logger.error(e.error_msg)
```

- 内存：选择“128”。
- 执行超时时间：输入“10”。
- 函数执行入口：默认“index.handler”，无需修改。
- 所属应用：默认“default”。
- 描述：输入“AS优雅关机”。

设置环境变量

进入as_graceful_shutdown函数，在“设置 > 环境变量”页签，配置环境变量，说明如表3-1所示，完成后单击“保存”。

图 3-3 配置环境变量

表 3-1 环境变量说明

环境变量	说明
region	伸缩组所在的区域
projectId	伸缩组所在的Project ID
ak	AK（Access Key ID）：访问密钥ID。与私有访问密钥关联的唯一标识符；访问密钥ID和私有访问密钥一起使用，对请求进行加密签名。如何获取AK请参考获取AK/SK。
sk	SK（Secret Access Key）：与访问密钥ID结合使用的密钥，对请求进行加密签名，可标识发送方，并防止请求被修改。如何获取SK请参考获取AK/SK。

添加依赖包

1. 制作“huaweicloudsdk_ecs_core_py3.6”与“huaweicloudsdk_as_core_py3.6”依赖包。制作依赖包详细操作请参见配置函数依赖。

说明

函数工作流服务除了支持用户自定义创建依赖包外，平台也提供了现成依赖包供用户使用，下载请参见如下：

- **huaweicloudsdk_ecs_core_py3.6**
- huaweicloudsdk_as_core_py3.6（该依赖包暂时无提供下载链接，敬请等待平台后续补充）

2. 返回函数工作流控制台，在左侧导航栏选择“函数 > 依赖包管理 > 创建依赖包”，分别创建“huaweicloudsdk_ecs_core_py3.6”与“huaweicloudsdk_as_core_py3.6”依赖包。
3. 填写依赖包信息，输入依赖包信息，完成后单击“确定”。
   - 依赖包名称：输入“huaweicloudsdk_ecs_core_py3.6或者huaweicloudsdk_as_core_py3.6”。
   - 上传方式：选择“上传ZIP文件”。
   - 文件上传：选择需要上传的zip包文件。
   - 运行时语言：选择“Python3.6”。

图 3-4 创建依赖包

4. 用户进入as_graceful_shutdown函数详情页，在“代码”页签，单击页面最底部的“添加依赖包”。
5. 添加“huaweicloudsdk_ecs_core_py3.6”与“huaweicloudsdk_as_core_py3.6”依赖包。

3.4 配置消息通知

1. 登录“消息通知服务 SMN”服务控制台，选择“主题”，单击右上方的“创建主题”。
2. 输入主题信息，完成后单击“确定”。
   - 主题名称：“AStoFunctionGraph”

- 企业项目：“default”（或者根据实际情况选择）

图 3-5 创建主题

3. 登录“弹性伸缩 AS”服务控制台，进入需要配置优雅关机的伸缩组的控制界面，选择“生命周期挂钩”，单击“添加生命周期挂钩”。

图 3-6 生命周期挂钩

4. 为伸缩组配置生命周期挂钩，完成后单击“确定”。这样缩容的实例就会被生命周期挂钩挂起，并发送消息通知至SMN主题。
   - 挂钩类型：“实例终止”。
   - 默认回调操作：“继续”。
   - 超时时间：“300”。
   - 通知主题：“AStoFunctionGraph”。

图 3-7 添加生命周期挂钩

5. 登录“函数工作流 FunctionGraph”服务控制台，进入函数“as_graceful_shutdown”详情页，在“设置 > 触发器”页签，单击“创建触发器”，弹出创建触发器界面。
6. 触发器类型选择SMN，主题名称选择SMN通知主题“AStoFunctionGraph”，这样SMN主题接收到的消息通知就会触发handler函数进行处理。

图 3-8 创建触发器

3.5 处理展示

触发伸缩活动

在伸缩组管理页面手动修改伸缩组的期望实例数，使其小于伸缩组的当前实例数，触发缩容活动。

查看伸缩实例为移出挂起状态，代表实例被生命周期成功挂起。

图 3-9 伸缩实例状态

触发函数运行

1. 在FunctionGraph的函数管理页面，选择监控-日志，观察FunctionGraph是否收到通知和函数的执行结果。

图 3-10 函数日志

2. 在伸缩组管理页面，选择伸缩实例，查看之前被挂起的伸缩实例状态为关机。

图 3-11 伸缩实例状态

3. 在伸缩组管理页面，等待伸缩组自动触发实例的移出和删除操作。

图 3-12 实例被移出并删除