

# 弹性伸缩 实践教学



腾讯云

#### 【 版权声明 】

©2013–2025 腾讯云版权所有

本文档（含所有文字、数据、图片等内容）完整的著作权归腾讯云计算（北京）有限责任公司单独所有，未经腾讯云事先明确书面许可，任何主体不得以任何形式复制、修改、使用、抄袭、传播本文档全部或部分内容。前述行为构成对腾讯云著作权的侵犯，腾讯云将依法采取措施追究法律责任。

#### 【 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。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

#### 【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。

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。

#### 【 联系我们 】

我们致力于为您提供个性化的售前购买咨询服务，及相应的技术售后服务，任何问题请联系 4009100100或95716。

## 文档目录

### 实践教程

为伸缩组设置固定出口 IP

搭建可伸缩的微信小程序

简介

步骤1: 为会话及业务服务器配置弹性伸缩策略

步骤2: 验证伸缩性和查看伸缩活动

设置服务开机启动

五分钟节省30%成本（集群版）

# 实践教程

## 为伸缩组设置固定出口 IP

最近更新时间：2024-08-02 14:29:51

本文档介绍如集群需主动外访，如何设置固定外访 IP。

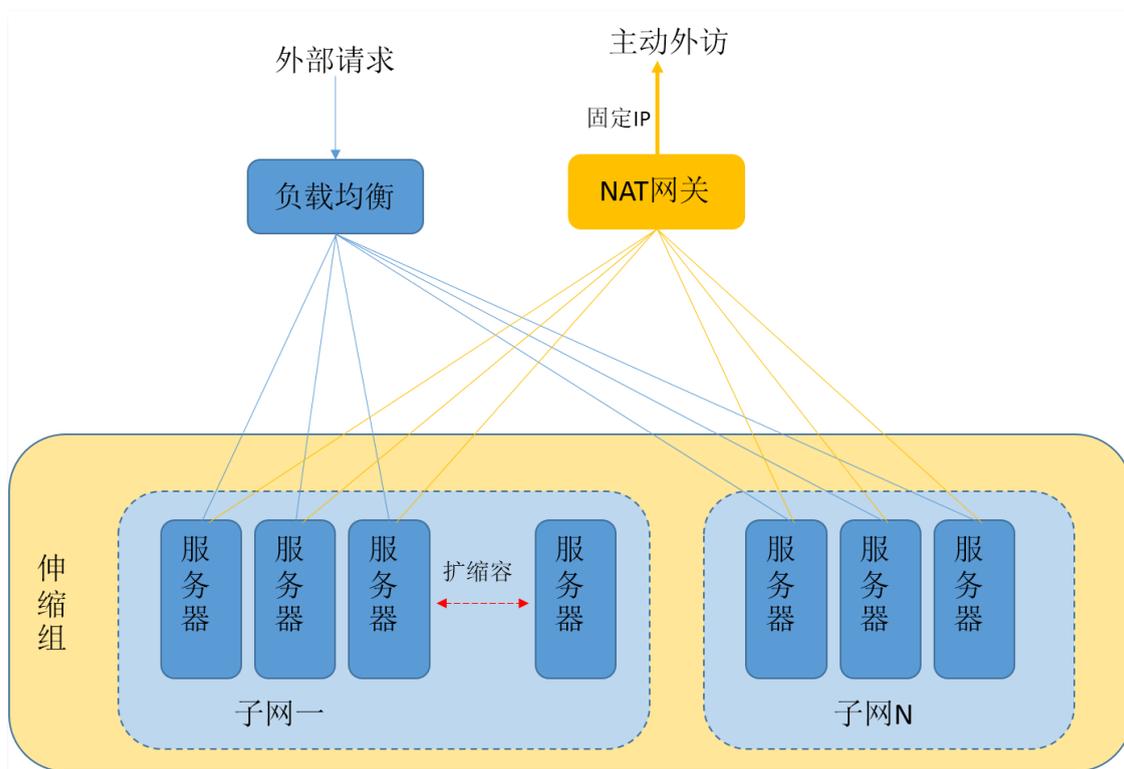
### 需求场景

如果您伸缩组中的集群，同时存在这三个需求：

- 从负载均衡 CLB 接受请求。
- 集群机器需要主动外访。
- 外访时希望用固定的外网 IP。

那么您可以按如下方案进行设置。

### 方案简述



1. 通过负载均衡 CLB 接收和响应外部请求。
2. 将机器放入私有网络 VPC 的子网中，将路由表指向 NAT 网关，主动外访请求统一经 NAT 网关的外网 IP 发出。
3. 伸缩组的网络属性设为该子网，这样扩容出来的机器都会统一用 NAT 网关主动外访。

### 设置方法

#### 步骤1: 创建 VPC 和子网

##### 创建 VPC

1. 登录私有网络控制台，选择左侧导航栏中的 **私有网络**。
2. 在私有网络页面上方，选择地域，例如，选择地域**华北地区（北京）**。

3. 单击 **+新建**，在弹出的 **新建 VPC** 窗口中，填写私有网络和子网的名称和 CIDR，并选择子网的可用区。
4. 单击**确定**即可创建 VPC。

## 创建子网

1. 在私有网络控制台中，选择左侧导航栏中的 **子网**。
2. 在 **子网** 页面上方，选择地域及 VPC。如下图所示：



3. 单击 **+新建**，在弹出的 **创建子网** 窗口中，填写子网名称、CIDR、可用区和关联路由表。
4. 单击**创建**即可，完成创建后，您就可以购买机器到这个子网中了。

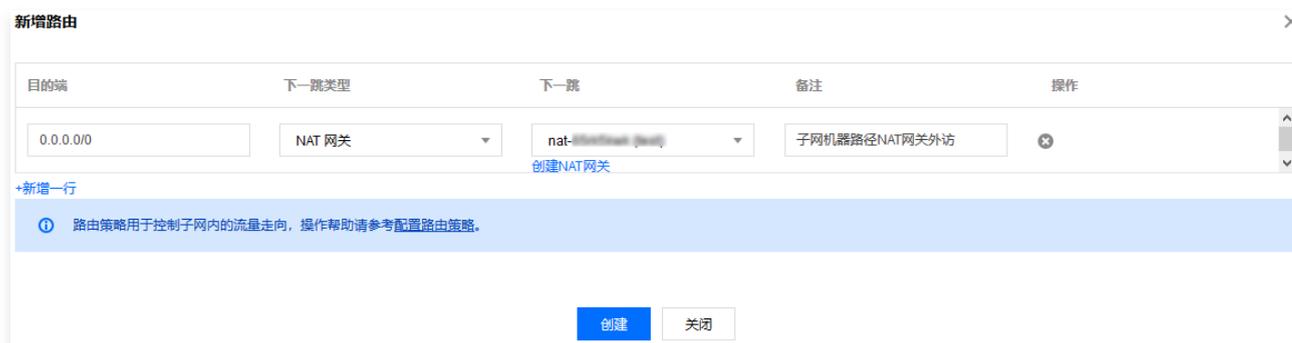
## 步骤2：创建 NAT 网关

### 新建 NAT 网关

1. 在私有网络控制台中，选择左侧导航栏中的 **NAT 网关**。
2. 在 **NAT 网关**页面中，单击 **+新建**。
3. 在弹出的 **新建NAT网关** 窗口中，依次输入或确定以下参数：
  - 网关名称
  - 网关类型（网关类型创建后可更改）
  - NAT 网关服务的私有网络（即为 **步骤1** 所创建的私有网络）
  - 为 NAT 网关分配弹性 IP（该 IP 即为您的机器外访的固定 IP）
4. 配置结束后单击**创建**，即可完成 NAT 网关的创建。  
创建完 NAT 网关后，您需要在私有网络控制台路由表页配置路由规则，以将子网流量指向 NAT 网关。

### 设置路由表（重点）

1. 在私有网络控制台中，选择左侧导航栏中的 **路由表**。
2. 在 **路由表** 页面中，选择需访问 Internet 的子网所关联的路由表 ID，进入该路由表详情页。
3. 单击 **+新增路由策略**，在弹出的“新增路由”窗口中，参考以下信息进行配置。如下图所示：



- **目的端**：此场景下可填写 `0.0.0.0/0`。
- **下一跳类型**：选择 **NAT 网关**，并选择 **步骤1** 中已创建的 NAT 网关 ID。

4. 单击**创建**即可。至此，您在这个子网中的机器即使没有公网 IP，也可以经过 NAT 网关主动外访了，对外而言还是固定的 IP。即使购买无公网 IP 且带宽为0的主机，也可以主动外访。如下图所示：

```
[root@UM_1_162_centos ~]# ping www.baidu.com
PING www.a.shifen.com (220.181.111.188) 56(84) bytes of data:
64 bytes from 220.181.111.188: icmp_seq=1 ttl=54 time=45.7 ms
64 bytes from 220.181.111.188: icmp_seq=2 ttl=54 time=45.8 ms
```

**说明：**

伸缩组需要识别这个子网，并确保机器都在这个子网上创建。

### 步骤3：设置伸缩组

此步骤的目的是将子网信息指向伸缩组，伸缩组就会把新扩容的机器放置在该子网中。扩容的机器会自动地用 NAT 网关的 IP 地址进行外访，达到固定出口 IP 的效果。

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，单击**新建**。
3. 在弹出的**新建伸缩组**页面中，填写伸缩组名称、已创建的启动配置、最大伸缩数、最小伸缩数、起始实例数等信息。其中**支持网络及支持子网**，请选择已配置好的 VPC 及子网。如下图所示：

### 新建伸缩组 ✕

1 **基本配置** > 
 2 负载均衡配置 > 
 3 竞价实例分配 > 
 4 其他配置

名称 \*

名称不超过55个字符，仅支持中文、英文、数字、下划线、分隔符和小数点

所属项目 **默认项目**

最小伸缩数 \*  -  + ⓘ

起始实例数 \*  -  + ⓘ

最大伸缩数 \*  -  + ⓘ

启动配置 \* asc (sg1实例; 规格: 1核1G) [新建启动配置](#) ⓘ

当前启动配置只有一个机型，建议在该启动配置设置多类似机型，以减少扩容失败风险，[现在去设置](#)

支持网络 \* vpc (sg1实例; 规格: 1核1G) ⓘ

如果您尚无任何支持网络，您可以[新建私有网络](#)

支持子网 \*

子网ID	子网名称	可用区	支持IPv6
<input checked="" type="checkbox"/> subnet-123456	sg1_subnet	广州东区	否
<input type="checkbox"/> subnet-123456	sg1_subnet	广州东区	否
<input type="checkbox"/> subnet-123456	sg1_subnet	广州东区	否
<input type="checkbox"/> subnet-123456	sg1_subnet	广州东区	否

您可选择多个子网，自动扩容的机器随机地从您勾选的子网创建，达到跨子网容灾的效果。[设置建议](#)

下一步

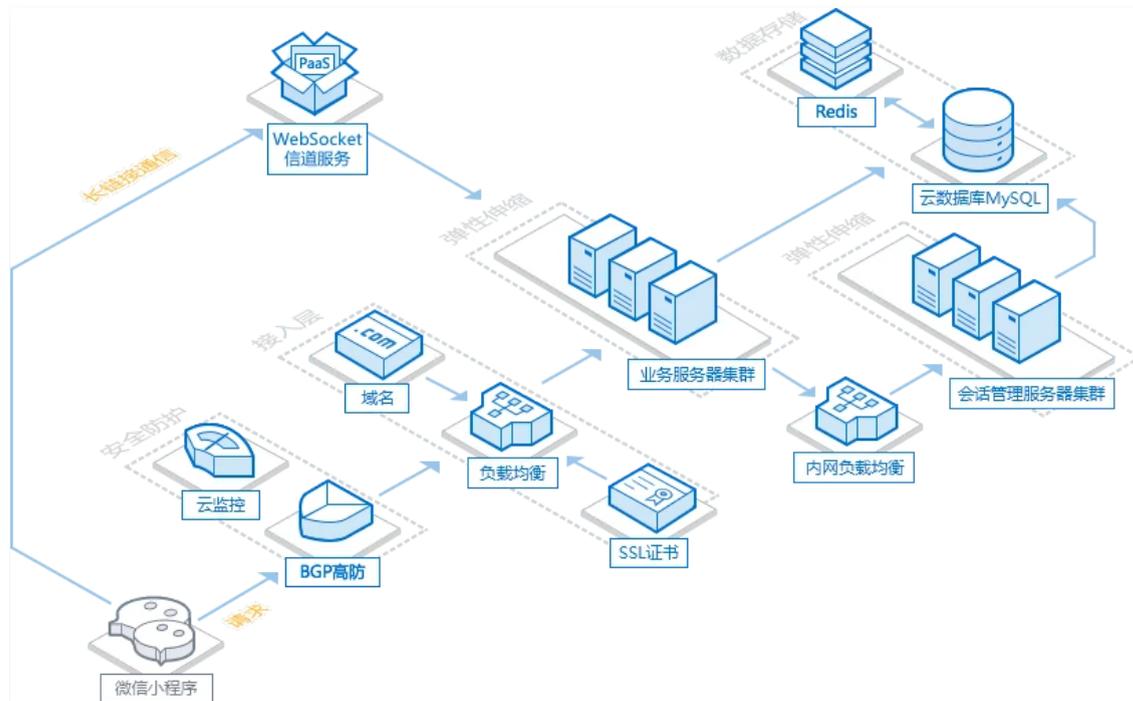
单击下一步即可完成设置。

# 搭建可伸缩的微信小程序简介

最近更新时间：2025-02-10 14:16:02

## 部署架构

我们可以按以下架构创建微信小程序，其中业务服务器集群和会话管理服务器集群需要具备弹性伸缩能力。如下图所示：



## 功能特性

弹性伸缩可以帮助您：

1. 定时给集群增加机器或减少机器。
2. 根据集群服务器的负载情况自适应地增加机器或减少机器。
3. 通过您设置的弹性伸缩策略，添加的实例会自动注册到负载均衡中，实现全自动扩容。

### ⚠ 注意：

弹性伸缩能力免费，扩容的 CVM 按秒正常计费。

## 应用场景

使用弹性伸缩，可降低成本和提高业务连续性。例如以下场景：

- 小程序访问有明显的高峰低谷：根据测算，如果业务服务器集群和会话服务器集群需要不止一台机器，且高峰短于8个小时，采用**按闲时保留固定服务器 + 高峰时期增加临时服务器**的方式，能节约30%左右成本。您可以通过 AS 设置定时扩容缩容任务，让腾讯云在忙时扩容临时服务器，闲时回收终止未充分利用的服务器。
- 小程序访问量稳定的预期下：配置基于监控告警的伸缩策略可应对意外高负载，保障服务的持续可用，给问题解决争取时间。异常高负载包括 CC 攻击以及意外流量。

### 📌 说明：

例如脸萌刚上线时远超预期的传播速度，或特定事件带来的突发访问。

# 步骤1：为会话及业务服务器配置弹性伸缩策略

最近更新时间：2025-05-22 15:47:51

## 前提条件

已搭建微信小程序，具体操作步骤请参见 [详细步骤](#)。

## 创建启动配置

### 说明：

扩容时以启动配置为模板创建机器，首先需通过启动配置指定地域、机型、镜像。

1. 登录 [弹性伸缩控制台](#)，单击左侧导航栏中的启动配置。
2. 选择小程序所在的项目和地域。如下图所示：



3. 单击新建，在弹出页面选择配置，并完成启动配置创建。

### 注意：

为了使扩容所添加的实例，无需手动配置实例环境就能够直接工作。在创建启动配置前您需要制作好镜像，确保镜像里的应用能随操作系统启动。

## 创建伸缩组

1. 登录 [弹性伸缩控制台](#)，单击新建。
2. 在新建伸缩组页面，填写集群的伸缩组管理信息，并单击下一步。如下图所示：

### 新建伸缩组

1 基本配置 > 2 负载均衡配置 > 3 竞价实例分配 > 4 其他配置

名称 \*   
名称不超过55个字符，仅支持中文、英文、数字、下划线、分隔符-和小数点

所属项目

最小伸缩数 \*    ⓘ

起始实例数 \*    ⓘ

最大伸缩数 \*    ⓘ

启动配置 \*  [新建启动配置](#) ⓘ

支持网络 \*    
如果您尚无任何支持网络，您可以[新建私有网络](#)

支持子网 \*

<input type="checkbox"/> 子网ID	子网名称	可用区	支持IPv6
<input type="checkbox"/> subnet-		广州七区	否
<input type="checkbox"/> subnet-		广州七区	否

您可选择多个子网，自动扩容的机器随机地从您勾选的子网创建，达到跨子网容灾的效果。[设置建议](#)

- **名称**：按需起一个名字，这里填“会话服务器集群”。
  - **所属项目**：可按需为伸缩组指定所属项目，未配置时默认归属“默认项目”。
  - **最小伸缩数**：集群实例数量的下限，这里填0即可。
  - **起始实例数**：伸缩组刚创建时，自动创建的实例数量。这里填0即可。
  - **最大伸缩数**：集群实例数量的上限，这里按需填写。
  - **启动配置**：选择您已创建的启动配置。
  - **支持网络**：会话服务器的网络环境，一般选“基础网络”即可。
3. 根据实际需求，在负载均衡下拉列表中选择已创建的负载均衡或新建，单击**完成**。完成创建。如下图所示：  
若无需配置负载均衡，可不选择负载均衡，单击**下一步：竞价实例分配**。



4. 在“竞价实例分配”步骤中，配置竞价实例分配策略，您也可单击**下一步：其他配置**跳过此步骤。

如需配置竞价实例，请参考以下参数说明：

- **按量基础实例数**：伸缩组内按量计费实例必须满足的最小数量。当伸缩组扩容时，首先扩容此部分的实例。
- **按量实例百分比**：除按量计费基础实例数外，按量实例所占的比例。可以指定0到100间的任意比例。
- **竞价实例创建策略**：启动配置多机型时，竞价实例创建的策略。
  - **容量优化策略**：优先选择最可用的竞价实例机型，以这种方式扩容可帮助您最好的利用竞价实例资源。
  - **成本优化策略**：优先选择单核价格最低的竞价实例机型，将从您指定的可用区中分配您的实例，以这种方式扩容可帮助您最大限度节约成本。
- **竞价实例回收监测**：开启后，弹性伸缩会尝试使用新的实例主动替换伸缩组中即将被回收的竞价实例，从而帮助您保持伸缩组内的实例数量及按量实例的比例。
- **按量实例补充竞价容量**：开启后，当您配置的机型竞价实例库存不足时为您尝试创建按量计费实例。

5. 在“其他配置”步骤中，参考以下信息设置移出策略及实例创建策略。

- **移出策略**：当伸缩组要减少实例且有多重选择时，将根据移出策略来选择移出的实例。支持“移出最旧的实例”及“移出最新的实例”。
- **实例创建策略**：
  - **首选可用区（子网）优先**：根据已配置的可用区（子网）顺序，优先选择靠前的配置项，失败后自动按顺序重试，适合以某个可用区为主，其他可用区为辅的架构。
  - **多可用区（子网）打散**：系统将根据扩容时伸缩组内实例在不同可用区（子网）的分布情况，选择相对较少的可用区（子网）创建新的实例，适合需要均匀分布实例的架构。
  - **标签配置**：可通过标签来进行资源的分类管理，详情请参见 [标签](#)。

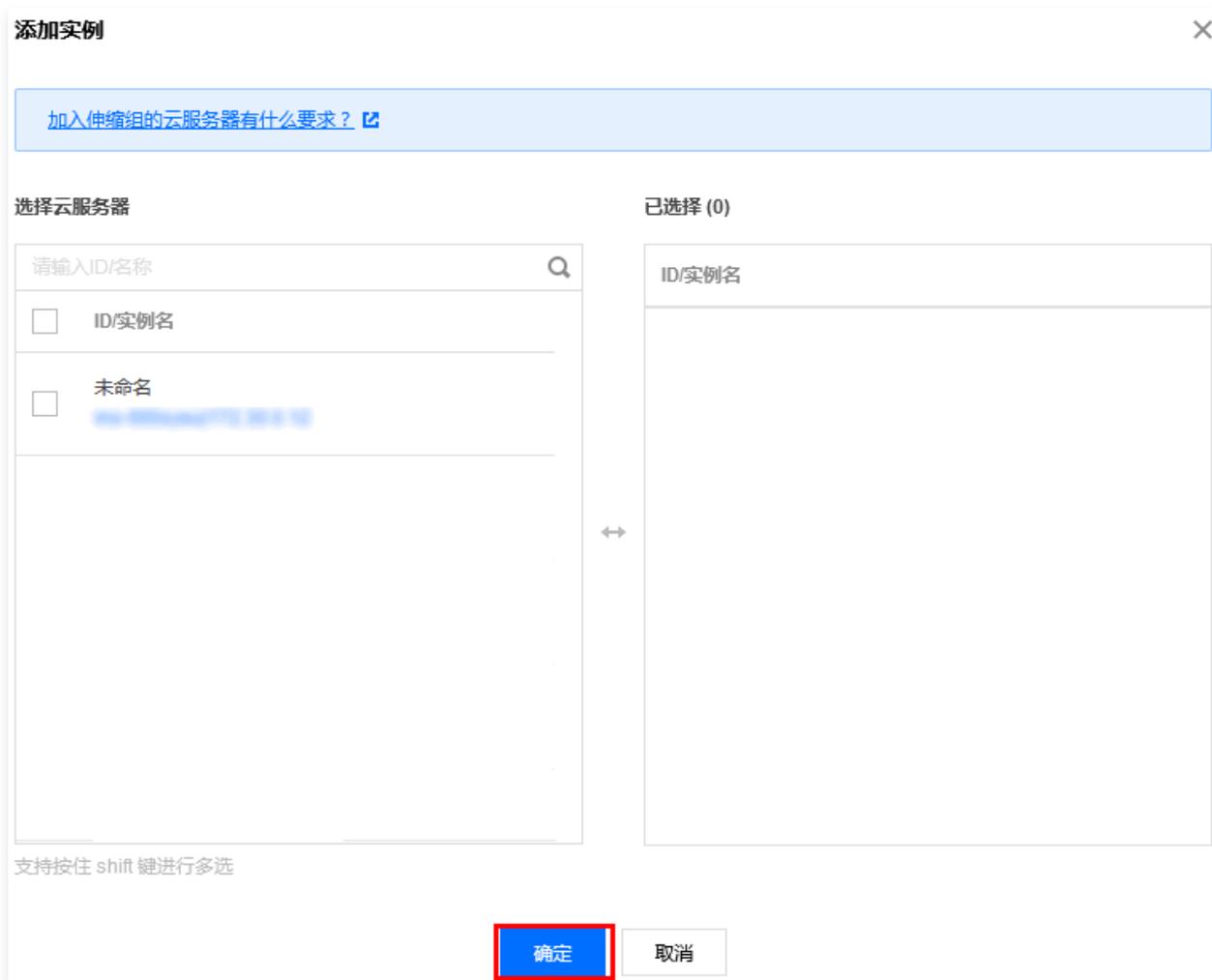
6. 单击**完成**，创建伸缩组。

## 添加现有机器进伸缩组

1. 在 [伸缩组列表页](#) 单击伸缩组 ID，进入伸缩组详情页。
2. 选择**关联实例**，单击**添加实例**。如下图所示：



3. 在添加实例页面中选择已有的会话服务器，并单击**确定**加入伸缩组。如下图所示：



添加实例成功后，进入伸缩组列表页。

4. 单击伸缩组右侧**设置移出保护**，并在弹出对话框中选择**确定**。如下图所示：



设置成功后服务器即可“免于缩容”。在缩容活动中，伸缩组不会选择这台服务器缩容。

## 设置扩缩容策略

### 说明：

通常扩容任务和缩容任务成对出现。

## 定时扩缩容

扩缩容策略可根据小程序特点设置。例如，在点餐小程序中，每天午饭时间实例的负载将会比其余时间的负载高。

您可以针对此点进行如下设置：

1. 在伸缩组详情页单击**定时任务**，并选择**新建**。如下图所示：



2. 设置定时扩容任务为**按天11:00 - 13:00扩容2台额外实例支撑负载**，并单击**确定**。如下图所示：

### 新建定时任务

名称

伸缩组活动

重复周期 **按天**  天, 执行一次

按天重复周期从每月1号开始, 例如周期为14天, 每月1号、15号、29号 (如有) 执行一次, 执行开始时间并不作为重复周期的计算起点 [详细说明](#)

执行开始时间

重复结束时间   [取消](#)

3. 设置定时缩容任务为**按天**15: 00 – 17: 00缩容2台实例减少支撑负载, 并单击**确定**如下图所示:

### 新建定时任务

名称

伸缩组活动

重复周期 **按天**  天, 执行一次

按天重复周期从每月1号开始, 例如周期为14天, 每月1号、15号、29号 (如有) 执行一次, 执行开始时间并不作为重复周期的计算起点 [详细说明](#)

执行开始时间

重复结束时间   [取消](#)

## 基于告警扩缩容

您可以进行如下设置：

设置预期不明确的扩容，应对出现意料之外的流量/攻击。

1. 在弹性伸缩组详情页单击告警触发策略，并选择新建。如下图所示：



2. 设置告警扩容策略，用于应对异常流量。如下图所示：



○ if: 此项为策略设置条件，图中设置为伸缩组内所有实例 CPU 利用率1分钟内最大值80%，连续1次。

○ 伸缩组活动: 增加2台实例，冷却0秒。

3. 设置告警缩容策略，用于清退未充分利用的服务器。如下图所示：



- **if:** 此项为策略设置条件，图中设置为伸缩组内所有实例 CPU 利用率1分钟内最大值20%，连续1次。
- **伸缩组活动:** 减少2台实例，冷却0秒。

## 为业务服务器配置弹性伸缩策略

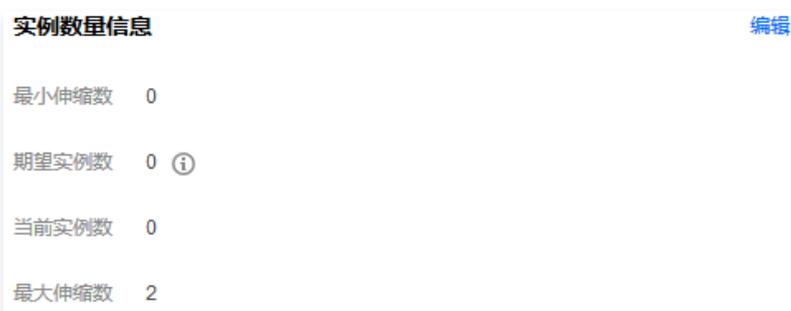
此过程与为会话服务器配置弹性伸缩策略类似。请按照 [创建启动配置](#) - [设置扩缩容策略](#) 的操作步骤，为业务服务器配置弹性伸缩策略。

## 步骤2：验证伸缩性和查看伸缩活动

最近更新时间：2024-11-22 14:56:12

### 验证伸缩性

1. 进入 [伸缩组列表](#) 页，单击伸缩组 ID，进入详情页。
2. 在伸缩组详情页，单击实例数量信息右上角编辑。如下图所示：



实例数量信息		编辑
最小伸缩数	0	
期望实例数	0 <span>ⓘ</span>	
当前实例数	0	
最大伸缩数	2	

3. 在实例数量信息编辑页，将伸缩组的期望实例数的数量设置为1，并单击保存。如下图所示：



实例数量信息	
最小伸缩数	<span>-</span> 0 <span>+</span>
期望实例数	<span>-</span> 1 <span>+</span> ⓘ
当前实例数	0
最大伸缩数	<span>-</span> 2 <span>+</span>

保存 取消

设置成功后伸缩组会自动扩容一台服务器到集群中，如果新扩容的实例能正常处理请求，说明伸缩组已正常工作。

### 查看伸缩活动

伸缩组还支持 [查询历史伸缩活动](#)，确保您完全掌控伸缩活动情况。

至此，您的小程序已经具备了智能扩缩容的能力。接下来您只需留意伸缩组通知或者不定期查看历史伸缩活动即可。

# 设置服务开机启动

最近更新时间：2025-05-22 15:47:51

## 操作场景

本文档介绍通过修改 `/etc/rc.d/rc.local` 文件，为自动扩容的机器设置服务开机启动。当使用 auto scaling 扩容时，我们希望全程无人工介入，因此建议您为自动扩容的机器设置开机后服务自动启动，例如：

- httpd 服务
- mysqld 服务
- php-fpm 服务
- tomcat 服务
- 其他服务等

## 操作步骤

### 设置服务开机自启动

#### ⚠ 注意：

本文以操作系统为 CentOS 的自动扩容机器为例。

1. 参见 [使用标准登录方式登录 Linux 实例（推荐）](#)，登录实例。
2. 执行以下命令，打开 rc.local 文件。

```
vim /etc/rc.d/rc.local
```

3. 按 `i` 进入编辑模式，并按 `↓` 到达文件末尾位置。
4. 输入以下内容，设置需自启动的服务。本文以 httpd、mysqld、php-fpm 服务为例，不同的网站所需的服务不同，请按需进行设置。

```
service httpd start
service mysqld start
service php-fpm start
```

添加完成后如下图所示：

```
#!/bin/bash
# THIS FILE IS ADDED FOR COMPATIBILITY PURPOSES
#
# It is highly advisable to create own systemd services or udev rules
# to run scripts during boot instead of using this file.
#
# In contrast to previous versions due to parallel execution during boot
# this script will NOT be run after all other services.
#
# Please note that you must run 'chmod +x /etc/rc.d/rc.local' to ensure
# that this script will be executed during boot.

touch /var/lock/subsys/local
/usr/local/qcloud/irq/net_smp_affinity.sh >/tmp/net_affinity.log 2>&1
/usr/local/qcloud/rps/set_rps.sh >/tmp/setRps.log 2>&1
/usr/local/qcloud/irq/virtio_blk_smp_affinity.sh > /tmp/virtio_blk_affinity.log 2>&1
/usr/local/qcloud/gpu/nv_gpu_conf.sh >/tmp/nv_gpu_conf.log 2>&1

service httpd start
service mysqld start
service php-fpm start
```

5. 按 Esc 键，输入 `:wq` 保存并退出。该实例再次开机后，即可自动访问网站。

## 配置验证（可选）

重启服务器（输入 `reboot` 即可重启，或者通过控制台重启）。服务器重启完成后不进入服务器，直接刷新网站的网页看是否有响应，如有即设置成功。

## 制作镜像

您可制作基于该实例的镜像，并在创建启动配置时使用该镜像。详情请参见：

- [制作自定义镜像](#)
- [制作启动配置](#)

# 五分钟节省30%成本（集群版）

最近更新时间：2025-05-22 15:47:52

## 操作场景

**说明：**

本文中提供的方案适用于集群式部署的网站或 App。

如果您的业务满足以下条件，可花5分钟配置这个方案，节省30%成本：

- 网站使用集群的方式，且集群拥有超过1台以上的云服务器（CVM）。
- 网站有较长时间的空闲。根据腾讯云的统计，90%的集群在凌晨00:00 - 早上09:00这9个小时的负载低于30%。

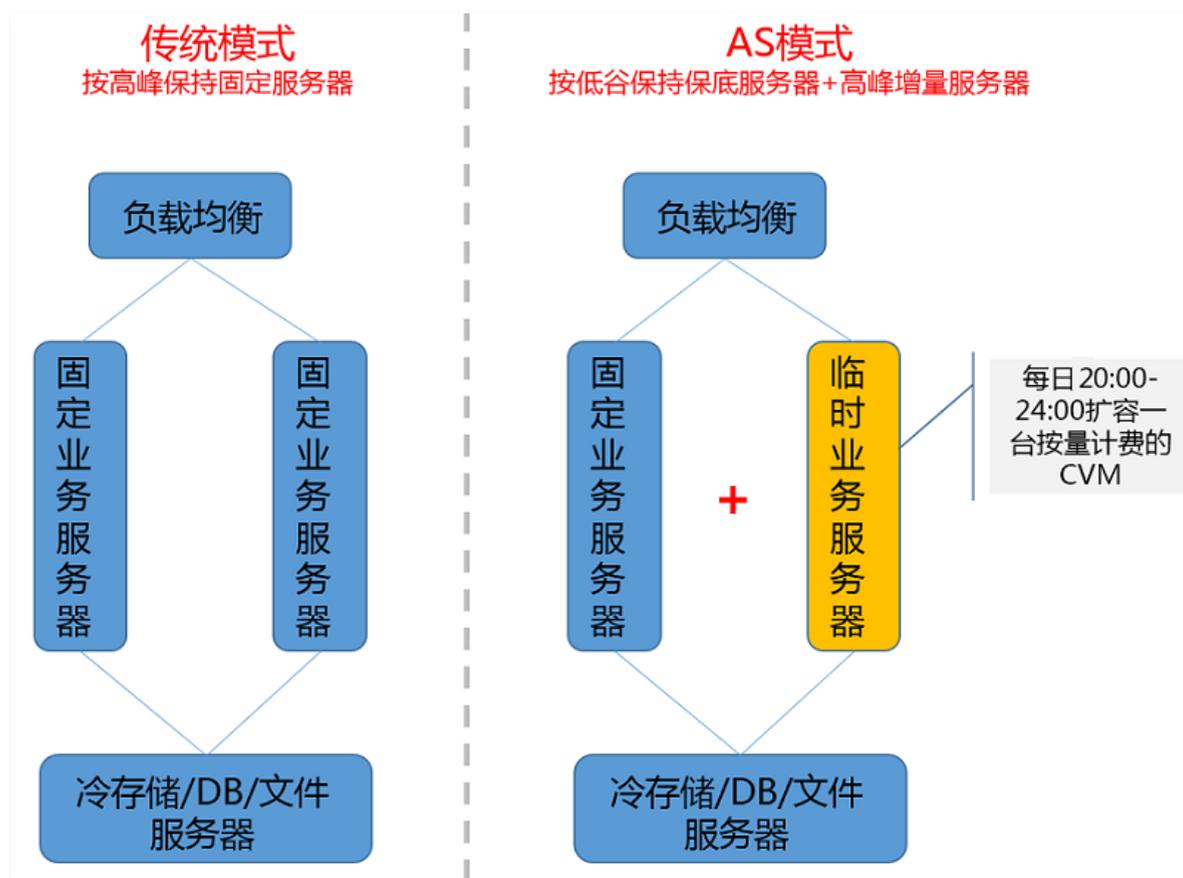
可以看出，大部分网站的高峰时间不超过8个小时，剩下的16个小时的时间，完全可以把闲置的服务器做缩容处理，这种方式能帮助您节约大量成本。

本文以某休闲类网站为例，假设该网站20:00 - 24:00是访问高峰时段。

## 方案简述

- 按非高峰时段的负载部署固定资源，可采用包年包月 CVM。
- 高峰时段的不足部分采用按量计费的 CVM，并通过定时任务在20:00扩容1台，24:00缩容回去。

新旧方案的对比：



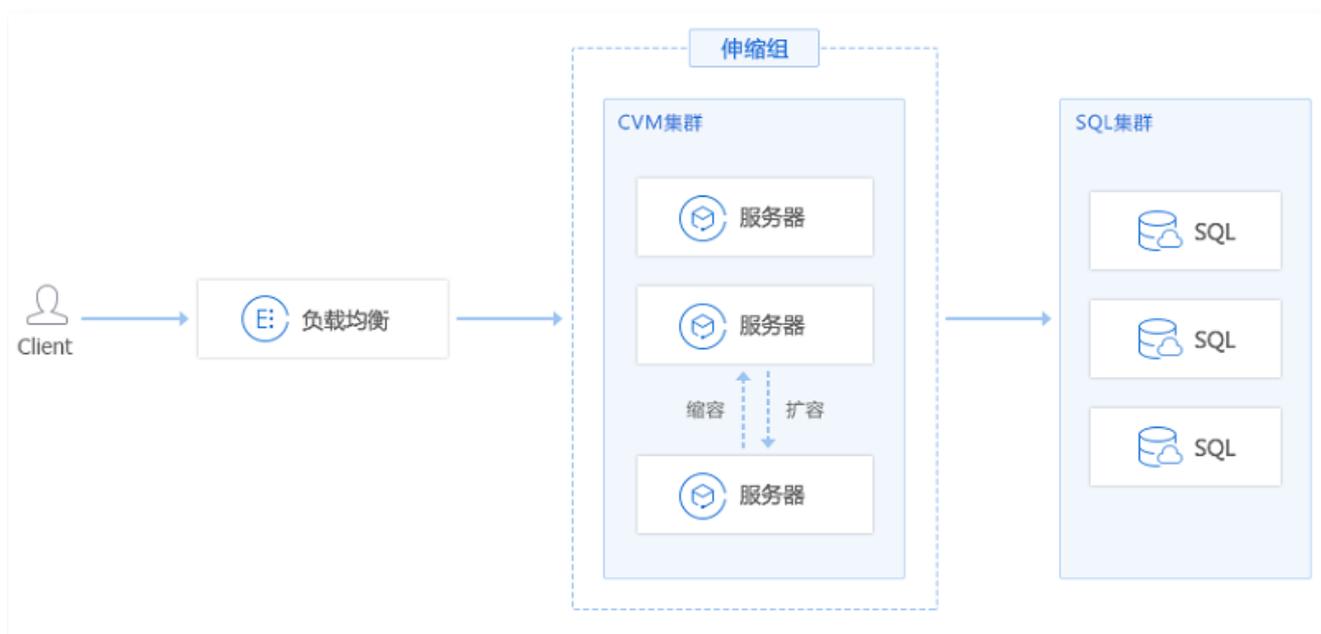
## 收益

假设原方案需要两台4核4GB的 CVM，改成一台4核4GB的 CVM + 每天4个小时临时 CVM，能节省30%左右开支。示例中的小网站每年可以节省1500元：

	传统方案			弹性伸缩方案		
	资源	数量	开销	资源	数量	开销
所需资源	4核4G的包月CVM (固定资源)	2	592	4核4G的包月CVM (固定资源)	1	296
				120小时的临时CVM (4核4G)	1	148.8
	负载均衡	1	30	负载均衡	1	30
总费用			622			474.8
节省情况	每月节省147元, 25%的幅度					

## 具体操作

示例网站的结构比较简单，只有应用服务器一个集群。如果复杂的网站，会有应用服务器集群、前端服务器集群、缓存服务器集群等，每个集群都可进行类似操作，每个集群对应一个伸缩组。如下图所示：



### 步骤1：创建集群机器的自定义镜像

请参见 [创建自定义镜像](#)，并基于一台现有的集群机器进行制作。

#### 说明：

您需要提前部署好镜像中的环境，确保镜像里的应用能随操作系统启动，这样扩容出来的机器就能直接工作，无需人工介入。

### 步骤2：创建启动配置

扩容时 AS 以启动配置为模板创建机器，因此我们需先通过创建启动配置时指定地域、机型、镜像。

1. 登录弹性伸缩控制台，选择左侧导航栏中的 [启动配置](#)。
2. 在启动配置列表页面上方，选择 Web 应用所在的项目和地域。
3. 单击新建进入“创建启动配置”页面，根据指引完成启动配置的创建。

请注意在镜像中，需指定您刚创建的自定义镜像。

### 步骤3：为机器创建伸缩组

1. 登录弹性伸缩控制台，选择左侧导航栏中的 [伸缩组](#)。

2. 单击**新建**，根据以下信息创建伸缩组，并单击**下一步**。如下图所示：

### 新建伸缩组

1 基本配置 > 2 负载均衡配置 > 3 竞价实例分配 > 4 其他配置

名称 \*   
名称不超过55个字符，仅支持中文、英文、数字、下划线、分隔符-和小数点

所属项目

最小伸缩数 \*    ⓘ

起始实例数 \*    ⓘ

最大伸缩数 \*    ⓘ

启动配置 \*  [新建启动配置](#) ⓘ

支持网络 \*    
如果您尚无任何支持网络，您可以[新建私有网络](#)

支持子网 \*

<input type="checkbox"/> 子网ID	子网名称	可用区	支持IPv6
<input type="checkbox"/> subnet-		广州七区	否
<input type="checkbox"/> subnet-		广州七区	否

您可选择多个子网，自动扩容的机器随机地从您勾选的子网创建，达到跨子网容灾的效果。[设置建议](#)

- **名称**：按需起一个名字，本文以**应用服务器集群**为例。
- **所属项目**：可按需为伸缩组指定所属项目，未配置时默认归属“默认项目”。
- **最小伸缩数**：集群服务器数量的下限，示例网站填0即可。
- **起始实例数**：伸缩组刚创建时，**自动创建**的机器数量。一般不会刚创建伸缩组就自动创建机器，建议填0。
- **最大伸缩数**：集群服务器数量的上限，请按需填写。本文最大伸缩数以5为例，即伸缩组最多有5台机器。
- **启动配置**：选择刚才您创建的启动配置。
- **支持网络**：会话服务器的网络环境，一般选**基础网络**即可。
- **支持子网**：会话服务器所在子网，您可选择多个子网，自动扩容的机器随机地从您勾选的子网创建，达到跨子网容灾的效果。

3. (可选) 在“负载均衡配置”中, 选择集群所使用的负载均衡, 并配置挂载监听器、域名及实例端口权重, 单击**下一步: 竞价实例分配**。



4. (可选) 在“竞价实例分配”中, 配置竞价实例分配策略, 并单击**下一步: 其他配置**。

如需使用竞价实例, 则请开启“使用竞价实例”开关。相关参数说明如下:

- **按量基础实例数:** 伸缩组内按量计费实例必须满足的最小数量。当伸缩组扩容时, 首先扩容此部分的实例。
- **按量实例百分比:** 除按量计费基础实例数外, 按量实例所占的比例。可以指定0到100间的任意比例。
- **竞价实例创建策略:** 启动配置配置多机型时, 竞价实例创建的策略。
  - **容量优化策略:** 优先选择最可用的竞价实例机型, 以这种方式扩容可帮助您更好的利用竞价实例资源。
  - **成本优化策略:** 优先选择单核价格最低的竞价实例机型, 将从您指定的可用区中分配您的实例, 以这种方式扩容可帮助您最大限度节约成本。
- **竞价实例回收监测:** 开启后, 弹性伸缩会尝试使用新的实例主动替换伸缩组中即将被回收的竞价实例, 从而帮助您保持伸缩组内的实例数量及按量实例的比例。
- **按量实例补充竞价容量:** 开启后, 当您配置的机型竞价实例库存不足时为您尝试创建按量计费实例。

5. (可选) 在“其他步骤”中, 进行伸缩组其他相关配置, 您也可单击**完成**跳过此步骤。

- **移出策略:** 当伸缩组要减少实例且有多个选择时, 将根据移出策略来选择移出哪个实例。可选择:
  - **移出最旧的实例:** 删除最旧自动增加的机器。自动增加的机器删除完后, 删除最早手动增加的机器。通常选择该项。
  - **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。
- **实例创建策略:** 当伸缩组需增加实例且指定多个子网在不同可用区时, 将根据该策略增加实例。可选择:
  - **首选可用区(子网)优先:** 根据您配置的可用区(子网)顺序, 优先选择靠前的配置项, 失败后自动按顺序重试(适合以某个可用区为主, 其他可用区为辅的架构)。
  - **多可用区(子网)打散:** 系统将根据扩容时伸缩组内实例在不同可用区(子网)的分布情况, 选择相对较少的可用区(子网)创建新的实例(适合需要均匀分布实例的架构)。
- **标签配置:** 可通过标签来进行资源的分类管理, 详情请参见 [标签](#)。

6. 单击**完成**, 即可创建伸缩组。

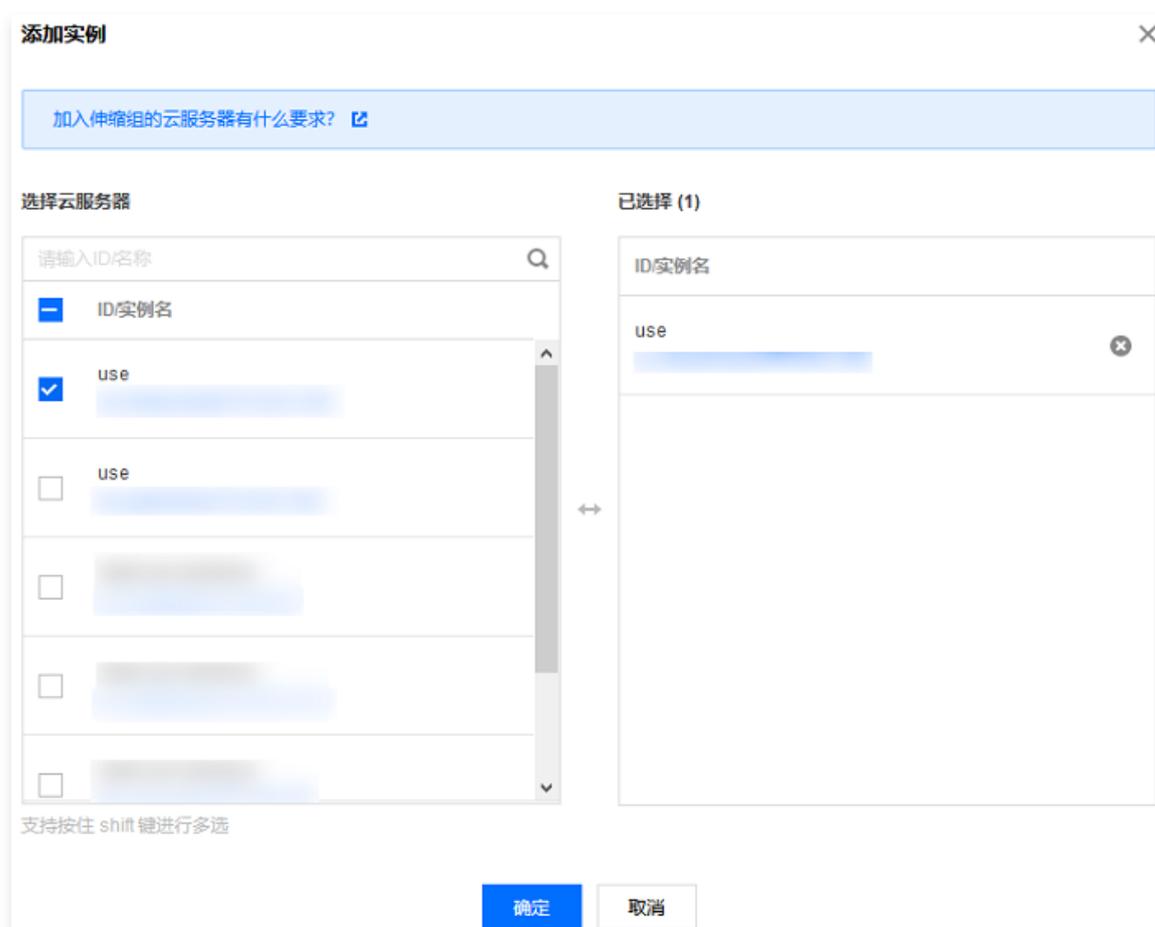
## 步骤4: 添加现有机器进伸缩组

1. 在 **伸缩组** 列表页面中, 选择伸缩组 ID, 进入伸缩组管理页。
2. 选择**关联实例**页签, 并单击**添加实例**。如下图所示:



3. 在弹出的添加实例窗口中，选择集群已有的云服务器加入伸缩组。如下图所示：

如果现在是非高峰期，集群中未充分利用的云服务器可以退还，节约成本。本文中两台名为 `use` 的云服务器，一台加入伸缩组，另一台可退还。



4. 成功添加后，对该云服务器设置免于缩容。

选择云服务器所在行右侧的设置移出保护，并在弹出框中单击确定。如下图所示：



已设置移出保护的云服务器在缩容活动中，伸缩组不会选择该云服务器缩容。此机器将永久保留在集群中进行服务，AS 不会更改它。

## 步骤5: 设置扩缩容策略 (重点)

AS 支持定时扩容及基于告警动态扩容、接收扩缩容通知、查看历史扩缩容详情等功能。您可结合实际情况进行使用，如下图所示：



1. 在“伸缩组管理”页面，选择**定时任务**页签，并单击**新建**。
2. 在“新建定时任务”窗口中，根据以下信息**设置一个20:00的定时扩容任务**，20:00云服务器数量增加至2台。如下图所示：

### 新建定时任务

名称 \*

伸缩组活动 \*

更改最小实例数为:

更改期望实例数为:

更改最大实例数为:

重复周期 按天  天, 执行一次

按天重复周期从每月1号开始, 例如周期为14天, 每月1号、15号、29号 (如有) 执行一次, 执行开始时间并不作为重复周期的计算起点 [详细说明](#)

执行开始时间 \*

重复结束时间 \*  [取消](#)

- **名称**: 输入任务名称, 本文以“20:00扩容”为例。
- **伸缩组活动**:
  - **更改最小实例数**: 伸缩组最小实例数, 保持为0不更改。
  - **更改期望实例数**: 根据**执行开始时间**每天执行一次, 本文中即为每日20:00都调整到2台云服务器。
  - **更改最大实例数**: 伸缩组最大实例数, 保持为5不变, 您可按需设置。
- **重复周期**: 伸缩组活动的周期, 本文设置每天执行一次。
- **执行开始时间**: 伸缩组活动重复开始时间, 即  开始活动, 并按照重复周期执行。
- **重复结束时间**: 伸缩组活动重复结束时间, 即  后不再重复。

### ⚠ 注意:

腾讯云的 CVM 需要1分钟左右创建, 如果自定义镜像较大, 可能需要更多时间。您可以将执行开始时间提早5分钟。

3. 参考以上步骤，再设置一个24:00的缩容任务，24:00云服务器数量减少至1台。如下图所示：

### 新建定时任务

名称

伸缩组活动

重复周期   天, 执行一次

按天重复周期从每月1号开始，例如周期为14天，每月1号、15号、29号（如有）执行一次，执行开始时间并不作为重复周期的计算起点[详细说明](#)

执行开始时间

重复结束时间

至此已完成配置。

网站的后台集群变为“1台固定应用服务器+1台高峰时定时创建的应用服务器”。

没加入伸缩组的其他集群机器，大部分时间未充分利用，可以退还掉节约成本。