

弹性伸缩 实践教学 产品文档



腾讯云

【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

实践教程

为伸缩组设置固定出口 IP

部署弹性 Web 应用服务

设置服务开机启动

实践教程

为伸缩组设置固定出口 IP

最近更新时间：2024-01-08 17:53:29

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：

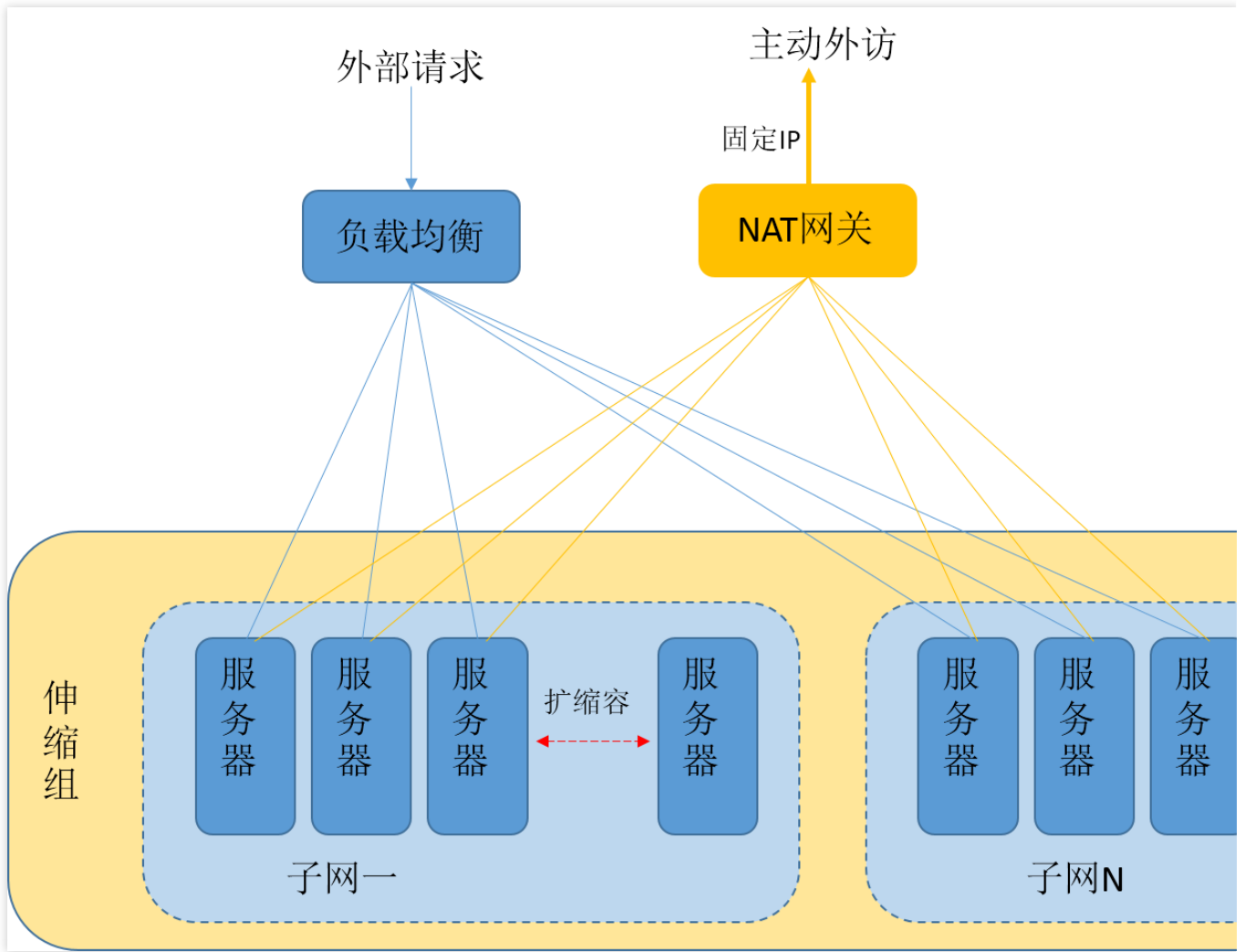
从负载均衡 CLB 接受请求。

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：

新增路由

i 路由策略用于控制子网内的流量走向，操作帮助请参考[配置路由策略](#)。

目的端	下一跳类型	下一跳	备注
<input type="text" value="0.0.0.0/0"/>	<input type="text" value="公网NAT网关"/>	<input type="text" value="创建NAT网关"/>	<input type="text" value="子网机器路径NAT网关外访"/>

[+新增一行](#)

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 支持网络

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](#) ⓘ

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 * ⓘ

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 checked="" type="checkbox"/> 子网ID	子网名称	可用区	支持IP
<input checked="" type="checkbox"/>	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

4. 单击下一步即可完成设置。

部署弹性 Web 应用服务

最近更新时间：2024-07-30 10:44:24

操作场景

本文档将指导您利用腾讯云 [弹性伸缩](#) 完成部署弹性 Web 应用服务。弹性伸缩帮助您以最合适的实例数量应对业务情况。当业务需求增加时，为您无缝地自动增加适量 CVM 实例。当业务需求下降时，为您自动减少不需要的 CVM 实例，提高设备利用率，为您节省部署和实例成本。

文档中提供的方案适合集群式部署的网站或 App。且示例的网站结构比较简单，只有应用服务器一个集群。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。

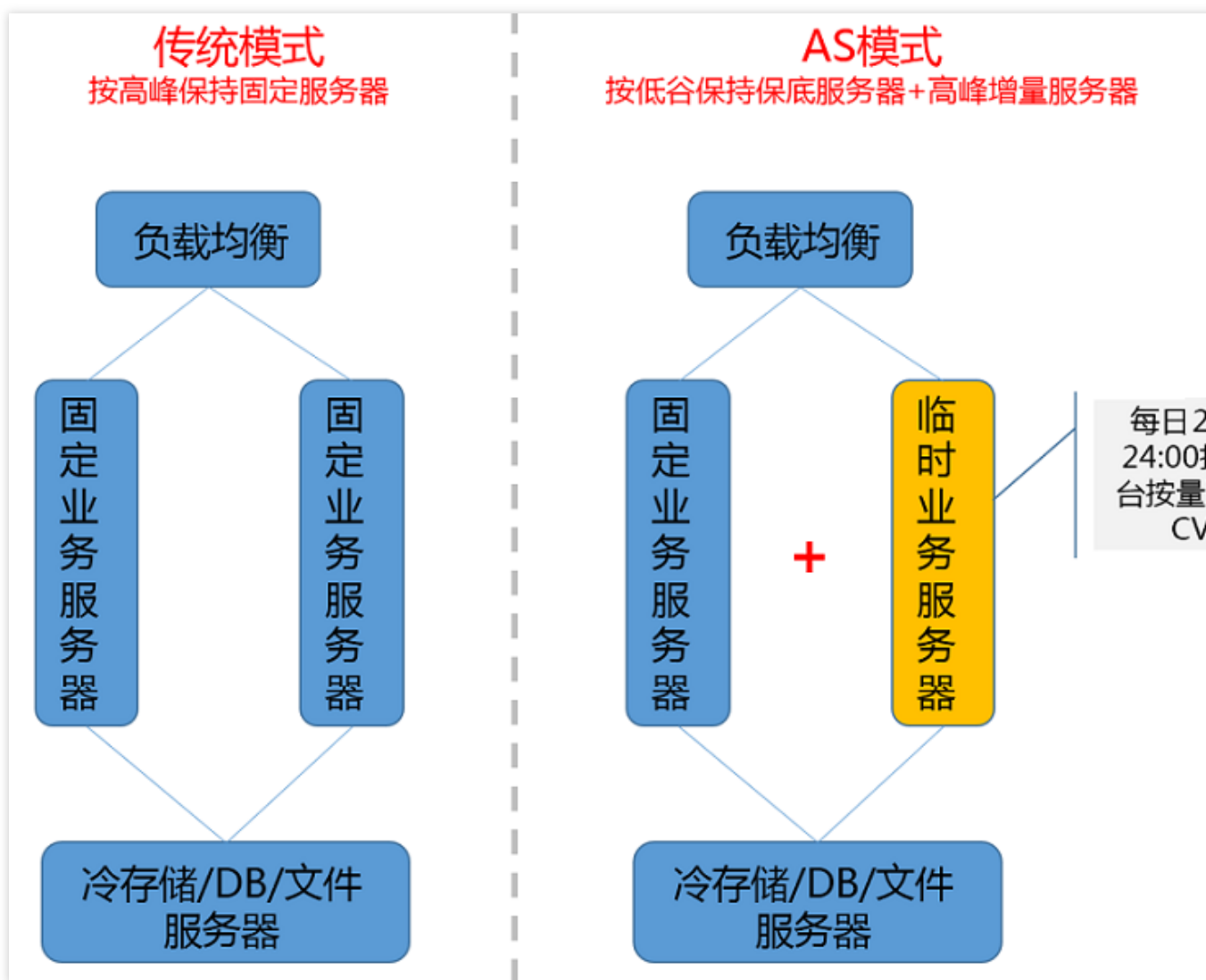
适用条件

网站使用集群的方式，且集群拥有超过1台以上的服务器。

网站有较长时间的空闲。大部分网站的高峰时间不超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作缩容处理，这种方式能帮助您节约大量成本。

本文档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。

高峰时段采用按量计费的 CVM，通过定时任务在20:00扩容1台 CVM 实例，并在24:00进行缩容。



步骤1：创建集群自定义镜像

基于一台集群中现有的机器制作镜像。关于创建自定义镜像的更多信息，请参见 [制作自定义镜像](#)。

1. 登录 [云服务器控制台](#)。
2. 在制作镜像前，需要关机实例。勾选需要关机的实例，单击上方**关机**。
3. 在需要制作镜像的实例右侧单击**更多**，单击**制作镜像**。
4. 在弹出框中，输入**镜像名称**和**镜像描述**，单击**确定**提交创建。创建自定义镜像需要一定的时间，请耐心等待。
5. 创建成功后，单击左侧导航栏中的 [镜像](#)，即可跳转至镜像列表查看详细信息。

步骤2：创建启动配置

扩容时弹性伸缩以 [启动配置](#) 为模板创建机器，因此需要事先通过 [启动配置](#) 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。本文以 [默认项目](#) 和 [广州](#) 为例。如下图所示：



3. 单击 [新建](#)，在弹出的 [创建启动配置](#) 页面选择机型及相关配置。
4. 选择机型。本文选择系列1标准型1颗 CPU1GB 内存。
5. 选择镜像。在 [镜像](#) 中选择 [自定义镜像](#)，指定您刚创建的镜像。本文以 [test](#) 为例。
6. 按需选择存储和网络，并单击 [下一步：设置主机](#)。
7. 设置信息。设置密码及安全组，并单击 [下一步：确认配置信息](#)。
8. 单击 [创建启动配置](#) 即可成功创建。

步骤3：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. 在 [伸缩组](#) 列表页面，单击 [新建](#) 进入 [新建伸缩组](#) 页面。
3. 在 [新建伸缩组](#) 页面，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](#) ⓘ

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 [去设置](#)

支持网络 * [新建私有网络](#)

如果您尚无任何支持网络，您可以[新建私有网络](#)

支持子网 *

子网ID	子网名称	可用区	支持IPv6
<input checked="" type="checkbox"/>			
<input type="checkbox"/>			
<input type="checkbox"/>			
<input type="checkbox"/>			

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](#)

名称：伸缩组的名称，本文以**应用服务器集群**为例。

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。本文做大伸缩数以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，本文以“基础网络”为例。

支持可用区：即选择扩容机器创建在哪个可用区中，此处按会话服务器所在的可用区勾选即可。

移出策略：选择默认。

4. 选择负载均衡。扩容出来的机器会自动挂载到您关联的负载均衡下，您可以选择已有负载均衡或者新建负载均衡。确认无误后，单击**完成**，完成创建。入下图所示：

负载均衡：选择集群的负载均衡，本文不选择负载均衡。

新建伸缩组

- ✓ 基本配置 >
- 2 负载均衡配置 >
- 3 竞价实例分配 >

负载均衡 ⓘ 请选择负载均衡 ↻

扩容出来的机器会自动挂载到您关联的负载均衡下，您可以选择已有负载均衡或**新建**。如需配置多个负载均衡，请在创建后编辑伸缩组。

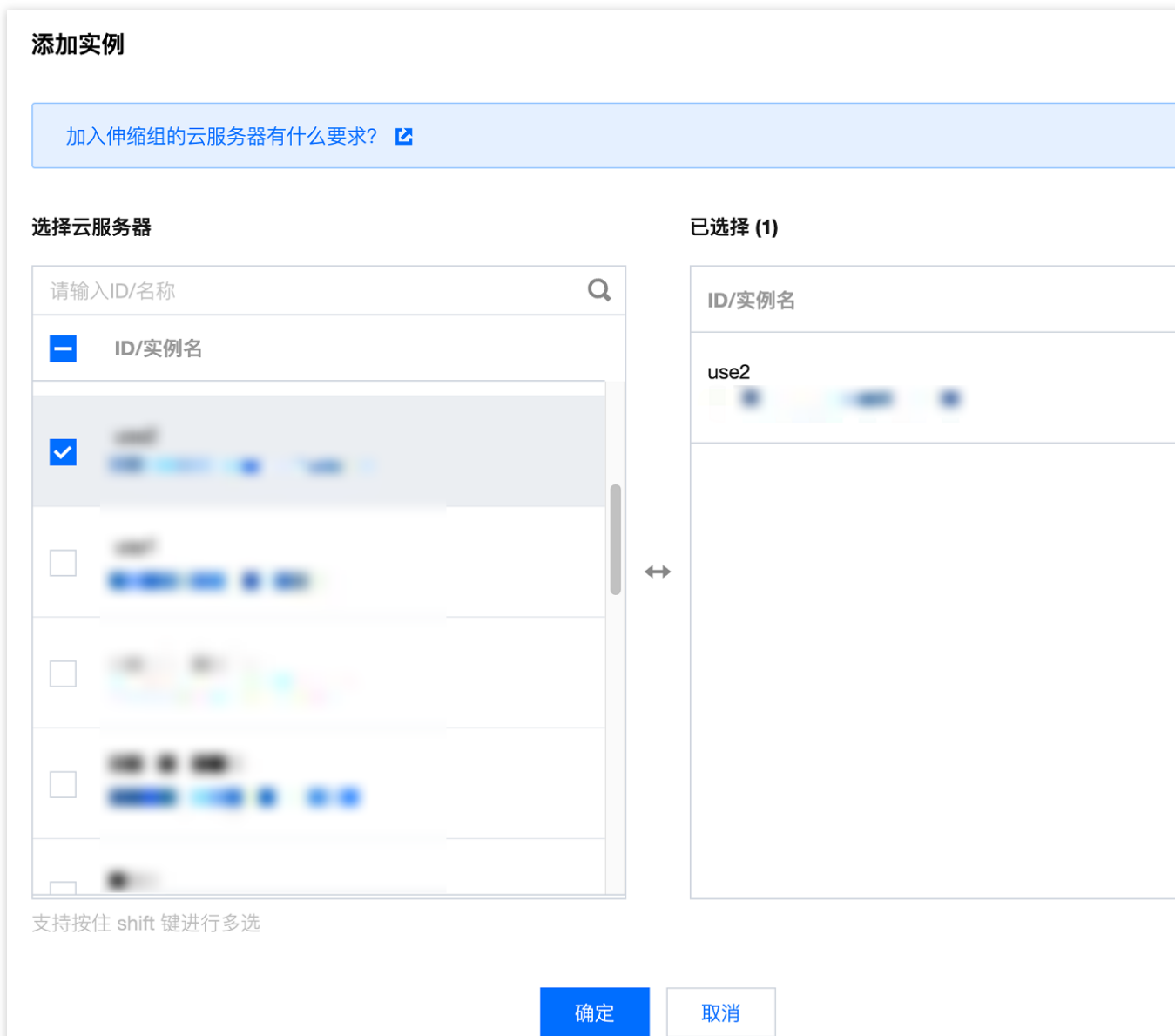
上一步
下一步：竞价实例分配
完成

步骤4：添加现有 CVM 实例进伸缩组

1. 在 [伸缩组](#) 列表页面，选择伸缩组 ID 进入该伸缩组管理页面。
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：
如果现在是非高峰时期，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**。

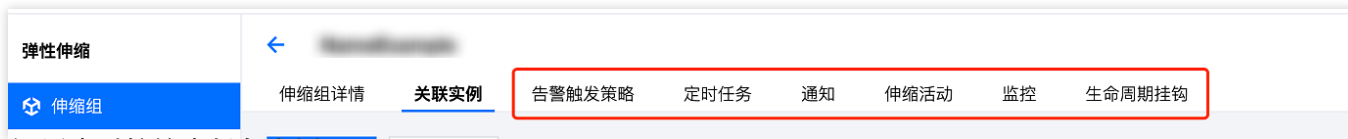
选择云服务器所在行右侧的**设置移出保护**，并在弹出框中单击**确定**。如下图所示：



5. 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

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：



本文以设置定时扩缩容任务为例。

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： - 0 +

更改期望实例数为： - 2 +

更改最大实例数为： - 5 +

重复周期 按天 ▾

每 - 1 + 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 * 2024-07-29 20:00 📅

重复结束时间 * 2025-07-29 20:00 📅 取消

确定 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 2022-07-29 20:00 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 2023-07-29 20:00 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 *

更改最小实例数为： - 0 +

更改期望实例数为： - 1 +

更改最大实例数为： - 5 +

重复周期 按天 ▼

每 - 1 + 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周期的计算起点[详细说明](#)

执行开始时间 * 2024-07-30 00:00

重复结束时间 * 2025-07-29 00:00 取消

确定 取消

4. 至此，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**。

设置服务开机启动

最近更新时间：2024-01-08 17:53:29

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

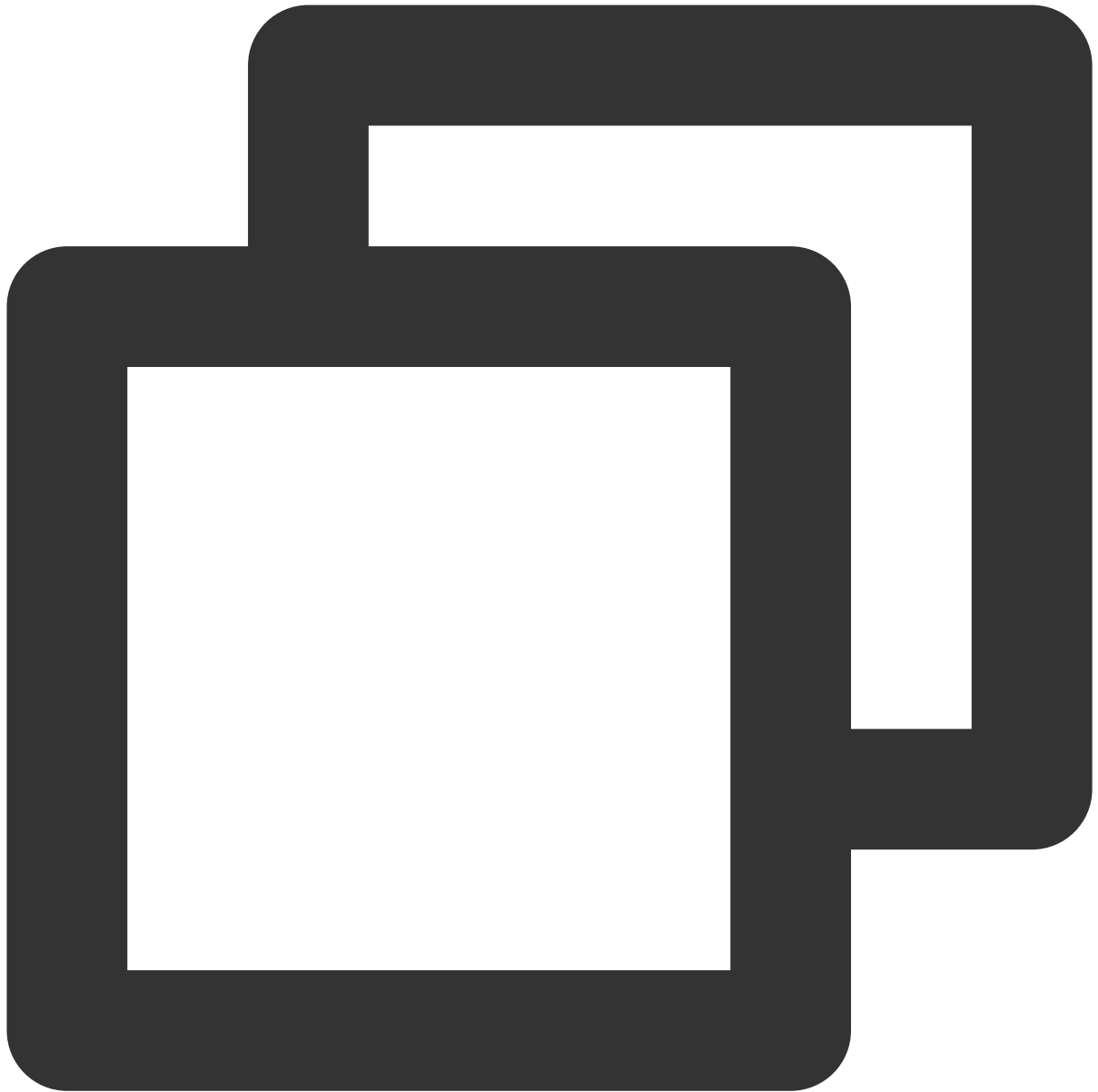
操作步骤

设置服务开机自启动

注意：

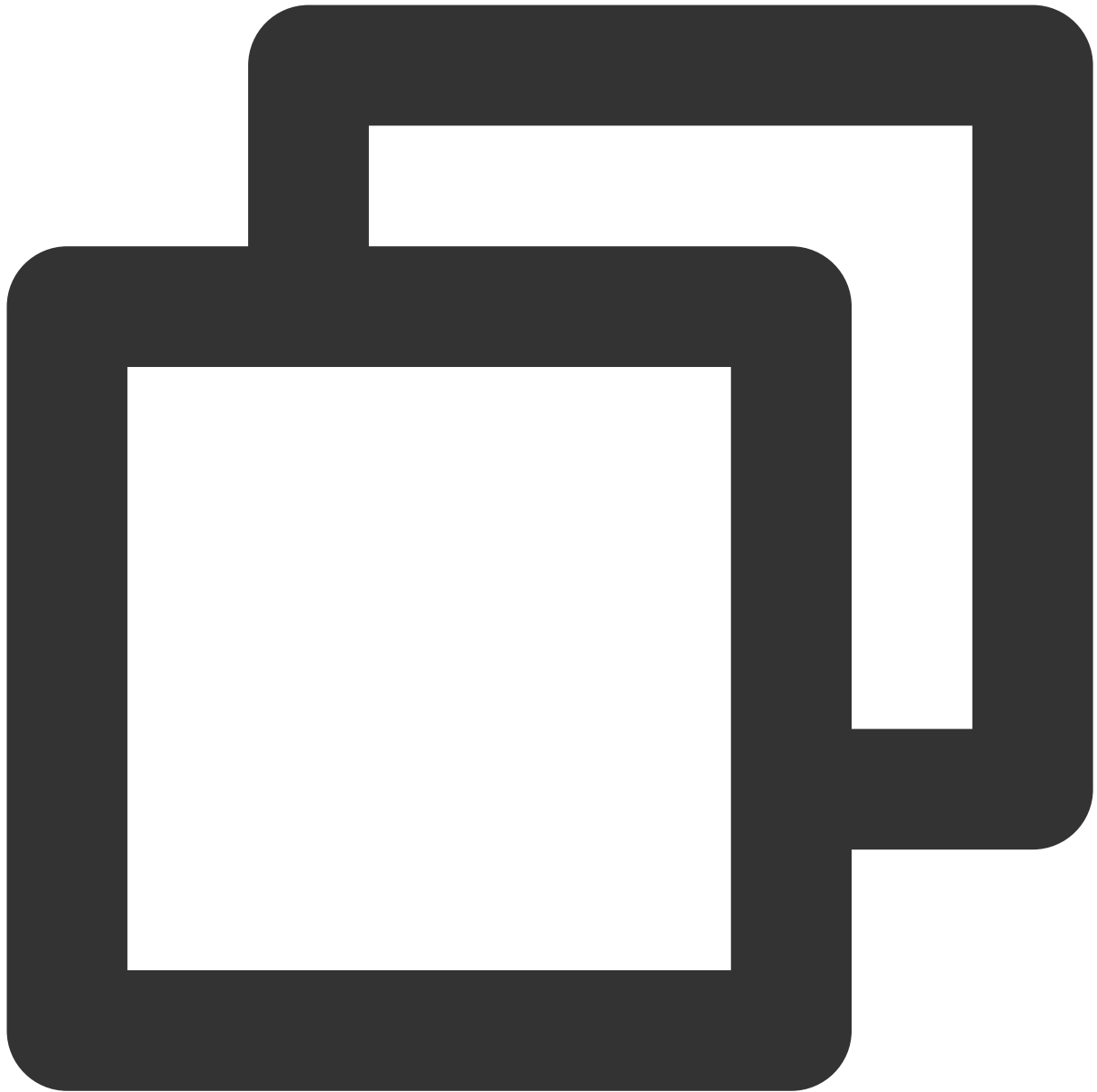
本文以操作系统为 CentOS 的自动扩容机器为例。

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
```

- 按 **i** 进入编辑模式，并按 **↓** 到达文件末尾位置。
- 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_affinit
/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. 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](#)