

ALIBABA CLOUD

阿里云

云服务器ECS
部署与弹性

文档版本：20220603

 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.实例启动模板	06
1.1. 实例启动模板概述	06
1.2. 创建实例启动模板	06
1.3. 创建实例启动模板的新版本	08
1.4. 删除实例启动模板和版本	10
2.基于ECS实例创建伸缩组	12
3.资源编排	20
3.1. 资源编排概述	20
4.应用管理	22
4.1. 使用ECS应用管理平台	22
5.部署集	25
5.1. 部署集概述	25
5.2. 创建部署集	26
5.3. 在部署集内创建ECS实例	27
5.4. 调整实例所属部署集	29
5.5. 修改部署集信息	30
5.6. 删除部署集	30
6.弹性供应	32
6.1. 弹性供应概述	32
6.2. 管理弹性供应服务关联角色	35
6.3. 创建弹性供应组	36
6.4. 弹性供应组设置示例	41
6.5. 查看弹性供应组	48
6.6. 修改弹性供应组	48
6.7. 删除弹性供应组	49
7.Terraform	51

7.1. Terraform概述	51
7.2. 在Cloud Shell中使用Terraform	52
7.3. 在本地安装和配置Terraform	54
7.4. 创建一台ECS实例	55
7.5. 创建多台ECS实例	57
7.6. 部署Web集群	61

1. 实例启动模板

1.1. 实例启动模板概述

实例启动模板是一项持久化ECS实例配置的功能，可用于快速创建实例。实例启动模板中包含了用于创建实例的配置信息，可以存储除了密码以外的任意配置信息，包括密钥对、RAM角色、实例类型和网络设置等。

实例启动模板特性

实例启动模板（以下简称为模板）是用于快速创建实例的模板。模板中包含了您自主选择的用于创建实例的配置信息，所有配置为可选，能够满足不同场景的个性化需求。

模板可以存储除了密码以外的任意配置信息，包括密钥对、RAM角色、实例类型和网络设置等。如果模板中有充分的配置信息则可以一键创建实例。

实例启动模板不支持修改，但可以创建多个版本，每个版本可以配置不同的参数，通过版本管理体现实例配置的演进过程。您可以使用模板任意一个版本创建实例。

控制台操作

- [创建模板](#)
- [在一个模板中创建多个版本](#)
- [更改默认版本](#)
- [使用模板创建实例](#)
- [删除模板和版本](#)

API操作

- [CreateLaunchTemplate](#)
- [CreateLaunchTemplateVersion](#)
- [DescribeLaunchTemplates](#)
- [DescribeLaunchTemplateVersions](#)
- [ModifyLaunchTemplateDefaultVersion](#)
- [DeleteLaunchTemplate](#)
- [DeleteLaunchTemplateVersion](#)

1.2. 创建实例启动模板

本文介绍创建实例启动模板的注意事项、如何创建实例启动模板以及使用已创建的模板可以进行的相关操作。

背景信息

创建实例启动模板前，您需要了解以下注意事项：

- 一个账号在每个地域最多可创建30个实例启动模板。
- 创建实例启动模板时，所有参数可选。但是，如果实例启动模板中未包含必要参数（例如实例规格、镜像等），则在使用该模板创建实例时还需要补充必要参数。
- 模板一旦创建，不可修改，但支持创建新的版本。

您可以通过ECS控制台或者售卖页创建启动模板，也可以通过阿里云CLI、OpenAPI开发者门户和阿里云SDK等开发者工具调用[CreateLaunchTemplate](#)去创建实例启动模板。

通过ECS控制台创建启动模板

您可以提前创建启动模板，便于后续使用模板快速创建ECS实例、弹性伸缩组或者弹性供应组。

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择部署与弹性 > 实例启动模板。
3. 在顶部菜单栏左上角处，选择地域。
4. 单击创建模板。
5. 在启动模板页面，完成基础配置和高级配置。

详细的参数配置及说明信息，请参见[使用向导创建实例](#)。

 **说明** 如果您是首次创建模板，模板克隆区域不可用。如果您已经创建过模板，可以选择已有的模板和版本，并在此基础上修改配置。

6. 在配置确认页面，输入模板名称和模板版本描述，并单击创建启动模板。

- 所选配置：您可以单击基础配置和高级配置区域的图标修改基础配置和高级配置参数。

 **说明** 在基础配置和高级配置区域，会建议您配置用于创建实例的必要参数，以便后续一键创建实例，但所有参数是可选的，请根据实际需要进行配置。

- 模板保存方式：根据需要选择新建模板的保存方式。创建新模板是指新建一个全新模板将当前配置保存为新模板的默认版本。创建新版本需选择一个已有模板将当前配置保存为该模板的最新版本。
- 模板名称和模板版本描述：输入模板的名称和模板版本的描述信息，方便后期管理。
- 模板资源组：根据需要选择已有的资源组，方便对模板分组管理。

如需创建新的资源组，您可以单击[去创建](#)前往资源组页面，创建新的资源组。更多信息，请参见[资源组](#)。

基础配置 	付费模式：按量付费 镜像：CentOS 8.0 64位（安全加固） VPC：默认vpc 安全组：1) bw-12345678-90123456789012345678901234567890	可用区：华东1可用区I 系统盘：高效云盘 40GiB，随实例释放 交换机：bw-12345678-90123456789012345678901234567890	实例规格：共享标准型 s6 / ecs.s6-c1m4.small (1vCPU 4GiB) 网络：专有网络 公网带宽：按使用流量 5Mbps
高级配置 	登录凭证：密钥对 ec2-key-pair	实例名称：launch-advisor-20220303	
模板保存方式 创建新模板 创建新版本 			
模板名称 template0303		2~128个字符，以大小写字母或中文开头，可包含数字、点号（.）、下划线（_）或连字符（-）	
模板版本描述 启动模板		长度为2~256个字符，不能以http://或https://开头	
模板资源组 默认资源组		 	
如需创建新的资源组，您可以单击 去创建			

7. 在创建成功对话框中，单击查看模板前往控制台查看已创建的模板。

通过ECS售卖页创建启动模板

如果您需要创建实例，并且希望把该实例的配置保存为模板用于日后再次一键创建，可以在创建实例的同时创建启动模板。

1. 前往[ECS产品详情页](#)，并单击[立即购买](#)。
2. 在自定义购买页面，配置参数。
详细的参数配置及说明信息，请参见[使用向导创建实例](#)。
3. 在确认订单页面，单击保存为启动模板。
4. 在弹出的对话框中，选择创建新模板，输入模板名称和模板版本描述后，单击保存。
5. 在创建成功对话框中，单击查看模板前往控制台查看已创建的模板。

后续操作

创建完启动模板后，您可以使用启动模板进行如下相关操作：

相关操作	如何使用启动模板	具体操作
创建ECS实例	直接使用现有的启动模板一键创建ECS实例，省去重复选择各配置项的时间。	具体操作，请参见 使用实例启动模板创建实例 。
批量创建ECS实例	启动模板可以结合RunInstances使用，批量创建ECS实例，启动模板可以免除每次创建实例时都需要填入大量配置参数。 在RunInstances请求中指定LaunchTemplateId和LaunchTemplateVersion参数使用启动模板。	更多信息，请参见 RunInstances 。
创建弹性伸缩组	基于ECS实例创建弹性伸缩组时，您可以通过现有的启动模板快速创建。系统会提取模板中的配置信息去创建新的伸缩组，如果某些配置不符合业务要求，您可以在伸缩组创建过程中修改配置，例如修改伸缩配置所属的专有网络、交换机等。	具体操作，请参见 基于ECS实例创建伸缩组 。
创建弹性供应组	弹性供应组使用实例启动模板的指定版本作为配置来源，沿用配置来源中的镜像、安全组、登录凭证等属性。弹性供应组创建成功后，会在您指定的时间启动并尝试交付ECS实例集群，提升批量交付大量实例的效率。	具体操作，请参见 创建弹性供应组 。

相关文档

- [CreateLaunchTemplate](#)

1.3. 创建实例启动模板的新版本

一个实例启动模板可以拥有多个版本，本文介绍如何创建新版本和更改默认版本。

前提条件

您已经创建了实例启动模板。具体操作，请参见[创建实例启动模板](#)。

背景信息

创建实例启动模板的新版本前，您需要了解以下注意事项：

- 新建的实例启动模板版本号为1，您可以继续在该模板上创建新版本。

说明 版本号按创建先后顺序递增，无法自定义版本号。

- 每个实例启动模板最多可创建30个版本。
- 创建版本时，所有参数可选。但是，如果版本中未包含必要参数（例如，实例规格、镜像等），则在使用该版本的实例启动模板创建实例时还需要补充必要参数。
- 版本一旦创建，不可修改。

通过控制台创建启动模板的版本

如果您此时没有创建实例的需求，可以提前基于实例启动模板创建新版本，用于日后一键创建实例。

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择部署与弹性 > 实例启动模板。
3. 在顶部菜单栏左上角处，选择地域。
4. 在目标模板的操作列中，单击新建版本。

您也可以选中一个模板ID，查看该模板的配置信息，并在版本信息中单击新建版本。



5. 在启动模板页面，自定义配置参数。
您可以在模板克隆区域，选择已有的模板和版本，在此基础上配置参数。详细的参数配置及说明信息，请参见[使用向导创建实例](#)。
6. 在配置确认页面，选择创建新版本，并选择要保存到哪一个模板。
7. 单击创建启动模板。
8. 在弹出的对话框中，单击查看新版本前往控制台查看已创建的版本。

您也可以通过阿里云CLI、OpenAPI开发者门户和阿里云SDK等开发者工具调用[CreateLaunchTemplateVersion](#)创建实例启动模板的新版本。

通过售卖页创建启动模板的版本

如果您需要创建实例，并且希望把该实例的配置保存为模板用于日后再次创建，可以在创建实例的同时创建实例启动模板的新版本。

1. 前往[ECS产品详情页](#)，并单击立即购买。
2. 在自定义购买页面，配置参数。
详细的参数配置及说明信息，请参见[使用向导创建实例](#)。
3. 在确认订单页面，单击保存为启动模板。
4. 在弹出的对话框中，单击创建新版本，并选择要保存到哪一个模板。

2. 在左侧导航栏，选择部署与弹性 > 实例启动模板。
3. 在顶部菜单栏左上角处，选择地域。
4. 找到需要删除的模板，在操作列中，单击删除。
5. 在弹出的对话框中，单击确定。

相关文档

- [DeleteLaunchTemplateVersion](#)
- [DeleteLaunchTemplate](#)

2. 基于ECS实例创建伸缩组

弹性伸缩伸缩组

如果您基于ECS实例创建伸缩组，伸缩组创建完成后，系统会创建一个默认伸缩配置，简化启用伸缩组前的准备操作，让您更快地使用伸缩组开始扩缩容。

前提条件

- 如果需要为伸缩组关联传统型负载均衡CLB（原SLB）实例，请确保满足以下条件：
- 如果需要为伸缩组关联应用型负载均衡ALB服务器组，请确保满足以下条件：
 - 伸缩组的网络类型必须为VPC，且与ALB服务器组处于同一VPC。
 - ALB服务器组必须处于可用状态。
- 如果需要为伸缩组关联RDS实例，请确保满足以下条件：
 - 您持有一个或多个处于运行中状态的RDS实例。更多信息，请参见[云数据库RDS简介](#)。
 - RDS实例和伸缩组必须位于同一地域。

背景信息

在基于ECS实例创建伸缩组时，不限制作组内实例配置信息来源的ECS实例的计费方式，包年包月实例、按量付费实例、抢占式实例均可以用于创建伸缩组。伸缩组创建完成后，系统会创建一个默认伸缩配置，该伸缩配置中采用的计费方式为按量付费。

 **说明** 默认伸缩配置中采用的镜像为已有实例使用的基础镜像，而不是为实例创建一个自定义镜像。如果您希望伸缩配置中采用的镜像包括已有实例的当前所有系统配置和数据，请自行实例创建自定义镜像，然后使用该自定义镜像更新伸缩配置，具体操作，请参见[手动更新伸缩配置中的镜像](#)。

伸缩组创建完成后，支持自动创建ECS实例和手动添加已有ECS实例。自动创建ECS实例时，计费方式仅支持按量付费、抢占式实例；手动添加已有ECS实例时，不限制计费方式，包年包月、按量付费、抢占式实例均可。更多使用伸缩组时的限制说明，请参见[使用限制](#)。

操作步骤

1. 进入创建伸缩组页面。
 - i. 登录[ECS管理控制台](#)。
 - ii. 在左侧导航栏，选择**实例与镜像 > 实例**。
 - iii. 在顶部菜单栏左上角处，选择地域。
 - iv. 找到ECS实例，在**操作区域**，单击**更多 > 部署与弹性 > 创建弹性伸缩**。

2. 在创建伸缩组页面中，完成伸缩组设置，然后单击**确认**。

相比从零开始，在基于ECS实例创建伸缩组时，弹性伸缩自动填入以下信息：

- **组内实例配置信息来源**：已选择**选择已有实例**。系统提取已有ECS实例的配置信息创建一个默认伸缩配置，作为自动创建ECS实例的模板，提取的配置信息包括实例的实例规格、基础镜像、网络类型、安全组、登录密码、标签等。如果某些配置不符合业务要求，您可以在伸缩组创建完成后根据需要修改配置，例如修改伸缩配置的镜像等。
- **网络类型**：已选择和已有ECS实例一致的网络类型。
- **专有网络**：已选择和已有ECS实例一致的专有网络（如网络类型为专有网络）。
- **选择交换机**：已选择和已有ECS实例一致的交换机（如网络类型为专有网络）。

伸缩组所有配置项的说明如下表所示，已经自动填入的信息请保持默认，无需修改。

配置项	说明
伸缩组名称	2~64个字符，以大小写字母、数字或中文开头，可包含半角句号（.）、下划线（_）和短划线（-）。
伸缩组类型	指定伸缩组内提供计算能力的实例的类型，弹性伸缩会根据该配置为伸缩组扩缩容对应的实例。 ◦ ECS：伸缩组内的实例都为ECS实例。 ◦ ECI：伸缩组内的实例都为ECI实例。
组内实例配置信息来源	弹性伸缩使用该配置创建实例。伸缩组支持以下来源类型： ◦ 启动模板：包含除了密码以外的任意配置信息，如密钥对、RAM角色、实例类型和网络设置等。仅伸缩组类型为ECS时，支持配置该项。 您需要选择已创建的实例启动模板和实例启动模板版本。为增加模板的灵活性，您还可以通过扩展启动模板配置选择多个实例规格。为实例规格指定权重，请参见使用性能指标作为弹性伸缩的度量单位。 ◦ 选择已有实例：您需要选择已创建的实例，弹性伸缩会自动提取该实例的基础配置信息，创建一个默认伸缩配置。 当伸缩组类型为ECS时，从ECS实例提取的伸缩配置包括该ECS实例的实例规格、网络类型、安全组、基础镜像等基础配置，不包括登录密码、标签。其中，基础镜像是已有实例使用的基础镜像，不包括该实例的数据信息，如应用程序。如果您想使伸缩配置包括该实例的所有系统配置和数据信息，请创建该实例的自定义镜像，并使用自定义该镜像更新伸缩配置的镜像。具体操作，请参见手动更新伸缩配置中的镜像。 ◦ 从零开始创建：先不指定自动创建实例的模板。等伸缩组创建完成后，再继续创建伸缩配置或指定启动模板。  说明 从ECS控制台入口基于实例创建伸缩组时，弹性伸缩自动填入组内实例配置信息来源、网络类型相关的信息，请保持默认，无需修改。
标签	添加标签便于搜索和聚合伸缩组，更多信息，请参见标签概述。  说明 此处的标签适用于伸缩组，如果您需要为伸缩组内实例添加标签，请在伸缩配置或启动模板中设置。

配置项	说明
实例移出策略	当需要从伸缩组移出实例并且有多种选择时，按该策略选择需要移出的实例。如果按策略筛选后仍有多台实例满足要求，则随机移出一台。仅伸缩组类型为ECS时，支持配置该项。 该项支持两段设置先筛选和再从结果中移出，但不支持为这两段设置相同的选项。各选项的含义如下： 说明 当伸缩组类型为ECI时，默认先筛选最早伸缩配置对应的实例，再从筛选结果中移出最早创建的实例。 最早伸缩配置对应的实例：筛选添加时间最早的伸缩配置和启动模板对应的实例。手动添加的实例没有关联伸缩配置或启动模板，因此不会首先选出手动添加的实例。如果已移出全部关联的实例，仍需要继续移出实例，则随机移出手动添加的实例。 说明 最早伸缩配置对应的实例中提到的伸缩配置泛指组内实例配置信息来源，包括伸缩配置和启动模板。 启动模板的版本号低不代表添加时间早，例如在创建伸缩组时选择实例启动模板lt-foress的版本2，然后修改伸缩组，选择实例启动模板lt-foress的版本1，则对伸缩组来说，启动模板lt-foress的版本2是最早的。 最早创建的实例：筛选创建时间最早的实例。 最新创建的实例：筛选创建时间最新的实例。 --无策略--：仅在设置再从结果中移出时可选，表示不进行第二段筛选。 例如，如果第一段设置为先筛选最早伸缩配置对应的实例，则第二段设置仅支持以下选项： --无策略--：不进行第二段筛选。 最早创建的实例：在第一段筛选出的实例中，再筛选创建时间最早的实例。 最新创建的实例：在第一段筛选出的实例中，再筛选创建时间最新的实例。 说明 伸缩组移出实例还受伸缩组的扩缩容策略影响。移出ECS实例的更多信息，请参见设置移出实例的组合策略。
暂停的流程	方便您在执行某些操作前主动暂停指定的流程，例如暂停健康检查流程后再去停止实例，避免实例被视为不健康而自动移出伸缩组。支持暂停的流程如下： 扩容流程：伸缩组拒绝所有扩容动作。 缩容流程：伸缩组拒绝所有缩容动作。 健康检查：暂停将实例标记为不健康状态，并暂停移出健康的实例。 定时任务：到定时任务的执行时间后，不会触发关联的伸缩规则。 报警任务：报警任务进入报警状态后，不会触发关联的伸缩规则。 更多信息，请参见暂停和恢复伸缩组流程。
开启伸缩组保护	开启伸缩组保护后，您不能在弹性伸缩控制台或者通过API删除该伸缩组，有效避免误删除伸缩组。

配置项	说明
实例的健康检查	开启健康检查后，伸缩组会定期检查实例的运行状态，如果发现一台实例未处于运行中状态，则判定为不健康并移出该实例。更多信息，请参见 伸缩组内ECS实例的生命周期 。
组内最小实例数	当前实例数量低于下限时，伸缩组会自动添加实例，使得伸缩组内的实例数量等于下限。
组内最大实例数	当前实例数量超过上限时，伸缩组会自动移出实例，使得伸缩组内的实例数量等于上限。
组内期望实例数	填写组内期望实例数即可开启期望实例数功能，伸缩组会自动将实例数量维持在期望实例数。更多信息，请参见期望实例数。  说明 只能在新建伸缩组时开启期望实例数功能，不支持通过修改已有伸缩组开启该功能。
默认冷却时间（秒）	单位为秒，伸缩组发生伸缩活动后的默认冷却时间。在冷却时间内，伸缩组会拒绝由报警任务触发的伸缩活动请求，其他类型任务触发的伸缩活动可以绕过冷却时间立即执行，例如手动执行任务、定时任务。
网络类型	仅专有网络伸缩组支持扩缩容策略、实例回收模式、关联应用型负载均衡ALB服务器组配置。  说明 从ECS控制台入口基于实例创建伸缩组时，弹性伸缩自动填入组内实例配置信息来源、网络类型相关的信息，请保持默认，无需修改。 伸缩组的网络类型决定了伸缩组内实例的网络类型，即专有网络的伸缩组只能存在专有网络的实例，经典网络的伸缩组只能存在经典网络的实例。  说明 伸缩组创建完成后，不支持修改网络类型。

配置项	说明
扩缩容策略	仅伸缩组类型为ECS，且网络类型为专有网络时，支持配置该项。具体策略如下： 优先级策略：先选择的交换机优先级高。弹性伸缩优先在优先级最高的交换机所在可用区尝试扩缩容，如果无法扩缩容，则自动在下一优先级的交换机所在可用区进行扩缩容。  说明 当伸缩组类型为ECI时，默认扩缩容策略为优先级策略。 均衡分布策略：在伸缩组关联多个交换机且交换机分布在两个以上可用区时生效，支持在交换机所在的可用区之间均衡分布ECS实例。如果由于库存不足等原因导致可用区之间ECS实例的数量不均衡，您可以执行再均衡分布操作来平衡ECS实例的分布情况。具体操作，请参见ECS实例再均衡分布。 成本优化策略：在伸缩配置中指定了多个可选实例规格时生效。扩容时弹性伸缩按vCPU单价从低到高尝试创建ECS实例，缩容时按vCPU单价从高到低尝试移出ECS实例。如果伸缩配置中计费方式选择抢占式实例，优先创建抢占式实例。由于库存等原因无法创建各实例规格的抢占式实例时，再自动尝试创建按量付费实例。 如果您选择成本优化策略，还可以设置以下属性： 组内最小按量实例数（台）：伸缩组所需按量付费ECS实例的最小台数，默认为0台。如果伸缩组内的按量付费ECS实例的台数小于该值，将优先创建按量付费实例。 按量实例所占比例（%）：自动创建ECS实例时按量付费实例所占的比例，默认为70%。计算该值时，不包括组内最小按量实例数对应的台数。 最低价的多个实例规格（个）：价格最低的实例规格的个数，默认为1个。在伸缩配置中指定了多个可选实例规格时生效。创建抢占式实例时，伸缩组会在价格最低的几个实例规格之间均衡创建ECS实例。 开启抢占式实例补偿：开启抢占式实例补偿后，在抢占式实例被回收前5分钟，伸缩组会主动创建新的抢占式实例，并替换掉将被回收的抢占式实例。 使用按量实例补充抢占式容量：默认为开启状态。如果因价格、库存等原因无法满足所需要的抢占式实例容量时，开启该功能，伸缩组会尝试创建按量付费实例来满足当前抢占式实例的容量。

配置项	说明
实例回收模式	仅伸缩组类型为ECS，且网络类型为专有网络时，支持配置该项。具体回收模式如下： - 释放模式：移出的实例被释放，无资源被保留。在弹性扩张时，弹性伸缩创建新的实例加入伸缩组。  说明 当伸缩组类型为ECI时，默认回收模式为释放模式。 - 停机回收模式：移出的ECS实例被停用，并触发节省停机效果，部分资源被保留并计费。在弹性扩张时，弹性伸缩优先将停用的ECS实例加入伸缩组，再根据扩张数量决定是否创建新的ECS实例加入伸缩组。该模式可以提高扩缩容的效率，更多信息，请参见选择节省停机模式提高扩缩容效率。  注意 - 请不要在实例中保留应用数据、日志等，避免实例回收后数据丢失的风险。 - 停用的实例可能会被释放： - 如果您手动减少伸缩组的最大实例数，使现最大实例数比伸缩组内各状态（包括停用中）的实例总数小，弹性伸缩优先释放停用的ECS实例。 - 因为库存不足或者账号欠费等原因，停用的实例在加入伸缩组时可能会失败，失败后该实例会被释放。 - 停机回收模式的作用受按量付费实例节省停机模式的影响，更多信息，请参见按量付费实例节省停机模式的“启用条件”、“适用资源”和“触发效果”部分。
专有网络	选择已创建的专有网络。  说明 从ECS控制台入口基于实例创建伸缩组时，弹性伸缩自动填入组内实例配置信息来源、网络类型相关的信息，请保持默认，无需修改。
选择交换机	选择专有网络后必须选择交换机。一个交换机只能属于一个可用区，您可以指定多个属于不同可用区的交换机，从而达到多可用区的效果。多可用区可以规避单可用区库存不足的风险，提高扩容成功率。  说明 从ECS控制台入口基于实例创建伸缩组时，弹性伸缩自动填入组内实例配置信息来源、网络类型相关的信息，请保持默认，无需修改。

配置项	说明
添加已有实例	如果伸缩组类型为ECS，且组内实例配置信息来源为启动模板或选择已有实例时，支持配置该项。 如果同时设置期望实例数并添加已有实例，期望实例数会自动增加。例如，创建伸缩组时设置期望实例数为1，并添加2台已有实例，伸缩组创建完成后，2台已有实例添加至伸缩组，然后期望实例数变为3。 您可以将已有实例的生命周期托管给伸缩组，选择将实例的生命周期托管给伸缩组复选框即可。 - 如果托管给伸缩组，添加的已有实例因处于不健康状态等原因被自动移出伸缩组，或者被手动移出伸缩组时，会被自动释放。 - 如果未托管给伸缩组，添加的已有实例移出伸缩组时，不会被自动释放。  说明 支持将包年包月实例添加至伸缩组，但不支持托管包年包月实例。
关联传统型负载均衡 CLB（原SLB）	伸缩组关联CLB实例后，加入伸缩组的实例会自动添加为CLB实例的后端服务器，处理来自CLB实例转发的访问请求。 您可以指定实例需要加入的服务器组，支持以下两种服务器组： - 默认服务器组：用来接收前端请求的实例，如果监听没有设置虚拟服务器组或主备服务器组，默认将请求转发至默认服务器组中的实例。 - 虚拟服务器组：当您需要将不同的请求转发到不同的后端服务器上时，或需要通过域名和URL进行请求转发时，可以选择使用虚拟服务器组。 如果您同时指定了默认服务器组和多个虚拟服务器组，实例会同时添加至这些服务器组中。  说明 一个伸缩组支持关联的CLB实例和虚拟服务器组数量有限。如需查看或手动申请提升配额值，请前往配额中心。
关联应用型负载均衡 ALB服务器组	仅网络类型为专有网络时，支持配置该项。伸缩组关联ALB服务器组后，加入伸缩组的实例会自动添加为ALB服务器组的后端服务器，处理ALB实例分发的访问请求。您还需指定后端服务器的端口和权重，权重默认为50。权重越高，实例将被分配到越多的访问请求。如果权重为0，则实例不会收到访问请求。 如果您为伸缩组关联了多个ALB服务器组，加入伸缩组的实例会同时添加至这些服务器组中。  说明 一个伸缩组支持关联的ALB服务器组数量有限。如需查看或手动申请提升配额值，请前往配额中心。
关联RDS数据库实例	仅伸缩组类型为ECS，支持配置该项。伸缩组关联RDS数据库实例后，加入伸缩组的ECS实例的内网IP会自动加入RDS数据库实例的访问白名单，允许ECS实例和RDS数据库实例内网通信。  说明 一个伸缩组支持关联的RDS数据库实例数量有限。如需查看或手动申请提升配额值，请前往配额中心。

配置项	说明
设置消息接收通知	在伸缩活动成功、失败或者被拒绝时，弹性伸缩支持通过短信、站内信和邮件发送消息通知。更多信息，请参见 设置消息接收通知 。

3. 在创建伸缩组对话框，单击关闭。

4. 在伸缩组管理页面中，找到新创建的伸缩组，在操作列选择 > 启用。

启用伸缩组后，弹性伸缩根据组内实例配置信息来源、组内最小实例数等设置维护一组实例。例如，在组内最小实例数大于0时自动创建对应数量的ECS实例，确保组内ECS实例的数量不小于组内最小实例数。

后续步骤

您可以手动添加、移出或删除ECS实例，并通过定时任务、报警任务等方式实现自动扩缩容。具体操作，请参见：

- [手动添加ECS实例](#)
- [手动移出或删除ECS实例](#)
- [实现自动扩张](#)
- [实现自动收缩](#)

相关文档

- [创建伸缩组](#)
- [CreateScalingGroup](#)

3. 资源编排

3.1. 资源编排概述

资源编排ROS（Resource Orchestration Service）是一项简化云计算资源管理的服务，ROS编排引擎能根据资源栈模板创建和配置资源，自动化交付项目所需的ECS实例和RDS实例等云计算资源。

更多详情，请参见[什么是资源编排服务](#)。

功能特点

- 重复部署

无论是开发环境，还是测试和生产环境，您都可以使用同一套模板，只需在模板中指定不同的参数，即可实现差异化部署。例如，测试环境的ECS实例数量设置为2，生产环境的ECS实例数量设置为20。如果您需要在多个地域部署应用，一套模板即可实现重复部署，提高效率。

- 标准化部署

在实践中，不同应用环境的细微差异可能导致管理成本攀升，延误了问题诊断的时间。通过使用资源编排ROS重复部署，您可以将应用环境的配置统一化、标准化，减少差异。

- 全自动托管

您无需购买和维护执行资源栈模板所使用的资源，只需关注业务所需资源，即模板中定义的资源。如果您需要创建多个项目（对应多个资源栈）时，全托管的服务可以加速项目进度。

- 身份认证和操作审计

与访问控制RAM集成提供身份认证功能，您无需单独建立用户认证和权限划分体系。您还可以使用操作审计服务（ActionTrail）审查相关云产品全部的运维操作，包括资源编排ROS本身。

产品优势

- 基础设施即代码

资源编排ROS是阿里云官方提供的基础设施即代码IaC（Infrastructure as Code）云产品，帮助您在DevOps中快速地实践IaC理念。

- 提升效率

资源编排ROS提供了解决方案模板，将复杂解决方案（例如SAP部署方案）的部署时间降低近90%。也支持使用模版进行标准化地重复部署，提升部署效率。

- 优化架构

通过资源编排ROS一键部署经典云上案例的解决方案，降低上云门槛，优化云上架构。

- 内控合规

利用资源编排模版预设的云上环境，方便财务审计和IT合规审计。

- 节省成本

通过预先设定的ROS模板，定期部署或释放应用以及云上环境，实现按需使用和按需付费。

使用ROS

您可以使用ROS控制台或调用API创建资源栈模板，快速创建和管理资源。详情请参见 [资源编排文档](#)：

- [模板结构说明](#)

- [创建资源栈](#)
- [API概览](#)

此外，您还可以：

- 采用Git或SVN等源代码版本管理软件管理模板版本，再通过ROS API维护资源栈。
- 使用阿里云CLI创建资源栈，详情请参见[资源栈相关](#)。

产品支持

如果您有相关需求或反馈，可以通过钉钉（[钉钉通讯客户端下载地址](#)）添加产品支持群联系阿里云，钉钉群号为11783495。

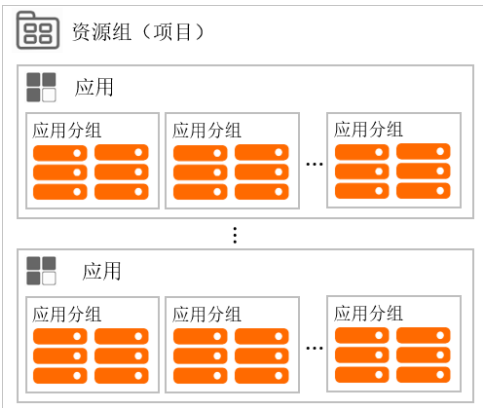
4.应用管理

4.1. 使用ECS应用管理平台

为了统一管理阿里云上的资源，ECS应用管理平台集成资源、安全、监控、事件、运维和发布等功能为一体，方便您高效地集中运维和管理资源。本文为您介绍如何使用ECS应用管理平台。

应用管理平台简介

本文中的应用管理指ECS应用管理，包括使用ECS+Docker的应用，且支持ACK形态。涉及的基本概念及关联关系如下所示：



概念	解释
资源组（项目）	指全地域下权限和财务管控的最小控制单元，是在阿里云账号下进行资源分组管理的一种机制。更多信息，请参见 资源组 或 基本概念 。
应用	指同一个软件包部署的最小分组单元，属于资源管理的最小控制单元。
应用分组	指应用下会部署多套环境，每一个应用分组对应不同的配置、地域、负载均衡等。根据业务环境或者业务场景不同，应用分组中的ECS实例可能关联的RDS或者SLB也不同。例如： • 预发环境和线上环境中，ECS实例会公用一个RDS，但却关联不同的SLB。 • 灰度环境和线上环境中，ECS实例会公用一个RDS，也会关联相同的SLB。 • 多地域分组场景下的RDS、SLB都不能通用。

一个资源组（项目）可以包含多个应用，每个应用可以由多个应用分组组成。

使用场景和流程

使用场景

假设您拥有多种云产品（例如云服务器ECS、云数据库RDS、负载均衡SLB等），每个云产品包含多个资源（例如云服务器ECS中的实例、块存储、云盘、快照等资源），针对多个资源有多个业务使用者，并且不同的使用者对资源的管理权限各不相同，例如有分组、分权、分账管理等需求。不同业务使用者（角色）包含以下三类：

角色名称	说明
系统管理员	具有管理应用基础信息、应用对应的标签（Tag）或资源组、应用的信息维护等权限。
应用管理者	具有资源创建权限、分配应用使用者权限、修改应用配置和应用配置下资源的权限、以及资源运维的权限、应用下的审批者。
应用使用者	仅具有资源查看和运维管理权限。

使用流程

使用ECS应用管理平台的简单流程如下所示：



流程	说明
快速创建应用	您可以提前配置应用管理员、应用联系人（信息通知）和应用使用者，根据业务使用者（角色）的不同权限，创建不同资源的应用并配置应用分组。例如，选择地域后，导入已有资源或者通过场景创建ECS资源。具体操作，请参见 步骤3 。
分组监控和管理资源	应用使用者和应用管理员到应用视图中可以直接看到对应的自己有权限的应用列表，进入具有权限的应用列表详情页，查看资源、安全、监控告警、事件等信息，并可以使用云效或者运维编排做常用运维操作。具体操作，请参见 步骤5 。
(可选) 修改和删除应用	您可以修改ECS应用的描述、备注等信息，应用对应的Tag和资源组不支持修改。如果当前应用不满足业务需要，您需要重新创建应用。具体操作，请参见 步骤6 。

操作步骤

- 1. 登录[ECS管理控制台](#)。
- 2. 在左侧导航栏，单击[应用管理](#)。
- 3. 快速创建ECS应用并分组管理已有资源。
如何创建应用和分组管理资源的具体操作，请参见[通过应用分组管理资源](#)。
- 4. 在[应用管理](#)页面，找到目标应用，点击应用名称进入应用详情页。
- 5. 在应用详情页，分组监控和管理ECS应用下的资源。

各功能模块的具体说明如下所示：

功能名称	说明
概览	您可以查看应用分组的费用详情，具体操作，请参见 查看应用分组费用分析 。
资源管理	您可以导入资源信息（例如导入云服务器ECS的实例资源），并对导入的资源进行相应的操作管理。例如，您可以对该实例资源进行移除、远程连接、执行命令、资源变配等操作。更多信息，请参见 通过应用分组管理资源 。
安全风险	您可以查看云平台的配置检查及资源的安全等信息。例如待处理的告警数、待修复的漏洞以及基线问题等。
监控告警	您可以配置资源的告警规则，具体操作，请参见 配置资源监控告警规则

功能名称	说明
事件	您可以查看系统事件的触发信息。如果您订阅了系统事件通知功能，就可以在该功能页签下，查看到计划内运维事件、非预期运维事件、实例费用事件、实例安全事件或实例因底层升级需迁移事件的触发信息，更多信息，请参见 ECS系统事件概述 和 设置事件通知 。
运维管理	您可以运维管理应用分组的资源，应用分组中可以使用OOS进行自动化运维，支持发送远程命令、下载文件和自定义任务，具体操作，请参见 运维应用分组的资源 。
发布管理	您可以实现代码从编译构建到部署应用的操作，通过云效流水线集成，将您的代码一键完成从代码编译构建到部署应用的操作。具体操作，请参见 创建云效流水线部署应用 。

6.（可选）在应用详情页，您可以根据业务需要修改或删除应用。

- 修改应用：如果您需要对应用的描述、备注等信息进行完善，可以对其执行修改操作。单击**编辑应用**，在弹出的对话框中，修改应用描述等信息后，单击**确定**。
- 删除应用：如果您不再需要某个应用时，可以对其执行删除操作。单击**删除应用**，在弹出的确认对话框中，单击**删除**。

删除应用时会解绑该应用下导入的资源，同时也会释放该场景下创造出来的资源（例如应用分组内的报警规则和事件通知等）。您无法通过该应用分组查看资源的监控数据，也无法收到该资源的报警通知，请谨慎操作。

 **说明** 您也可以直接在应用列表页，找到待删除的应用，在对应操作列，单击删除。

5.部署集

5.1. 部署集概述

部署集是控制ECS实例分布的策略，使您能在创建ECS实例的时候就设计容灾能力和可用性。

部署策略

您可以使用部署集将业务涉及到的ECS实例分散部署在不同的物理服务器上，以此保证业务的高可用性和底层容灾能力。在部署集内创建ECS实例时，会根据您事先设置的部署策略，ECS实例在指定地域下被分散启动。

部署集支持以下策略：

- 高可用策略

采用高可用策略后，部署集内所有ECS实例会在指定地域内严格分散在不同的物理服务器上。适用于需要将几台ECS实例相互隔离的应用架构，大幅降低服务不可用的几率。

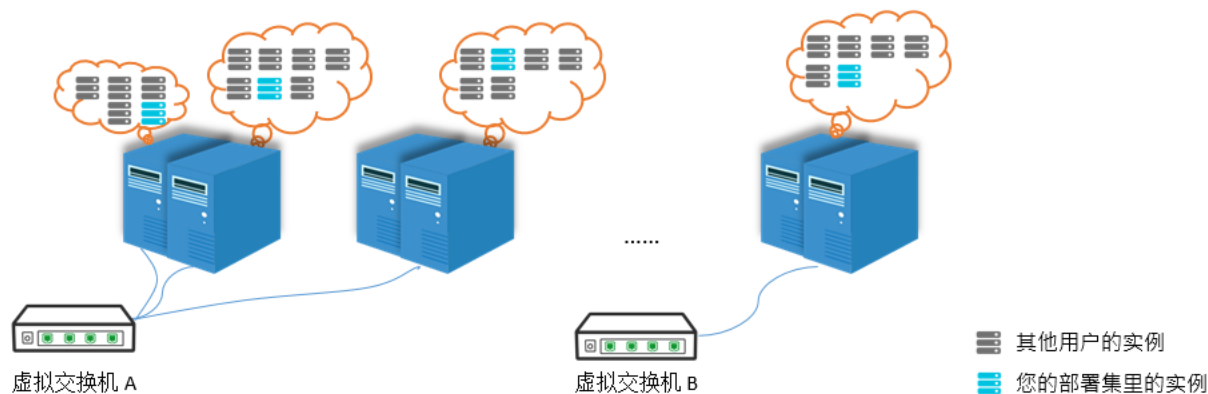
- 部署集组高可用策略

该策略支持将部署集划分为最多7个分组，多台ECS实例可以根据实际需要分散部署在不同的分组当中。不同分组的ECS实例会在指定地域内严格分散在不同的物理服务器上；相同分组的ECS实例不保障严格分散部署。

 **说明** 遇到地域内供货紧缺时，可能无法创建ECS实例，或者重启按量付费ECS实例（节省停机模式）失败。一般情况下，您可以等待一段时间后重试创建或重启操作。

部署示例

下图是利用部署集能力提升业务可靠性的典型示例，您的四台ECS实例分布在四台不同的物理服务器上。



如果您需要部署集内的ECS实例能实现低延时通信，建议您保持实例的网络类型一致性。例如，创建实例时为这些实例选择同一个专有网络VPC。

计费详情

使用部署集不会收取服务费用，但您创建和使用的ECS实例、磁盘、快照、镜像和公网带宽等服务将收取费用。更多信息，请参见[计费概述](#)。

使用限制

在使用部署集之前，您需要注意：

- 部署集之间不支持相互合并。
- 部署集内不能创建抢占式实例。
- 部署集不支持创建专有宿主机。
- 在部署集内创建ECS实例时，一个可用区内最多能创建20台ECS实例（该数值根据您使用云服务器ECS的情况而变化），一个阿里云地域下能创建的ECS实例数量为 $20 * \text{可用区数量}$ 。
- 部署集支持的实例规格族信息：
 - c7、g7、r7、c6、g6、r6、c5、g5、r5、c6e、g6e、r6e、c7se、g7se、r7se、r6se、c7t、g7t、r7t、c7a、g7a、r7a、c6a、g6a、r6a、g5ne、re6、re4、ic5
 - hfc7、hfg7、hfr7、hfc6、hfg6、hfr6、hfc5、hfg5
 - d2s、d2c、d1、d1ne、d1-c14d3、d1-c8d3
 - i3、i3g、i2、i2g、i2ne、i2gne、i1
 - se1ne、sn1ne、sn2ne、se1
 - ebmg5、sccgn6、scch5、sccg5、scch5s、sccg5s
 - s6、t6、xn4、mn4、n4
 - gn6i

您可以调用[DescribeDeploymentSetSupportedInstanceTypeFamily](#)获取部署集最新支持的实例规格族信息。

- 在部署集内创建ECS实例时，或者重启按量付费ECS实例（节省停机模式）时，供货紧缺依然会导致请求失败。更多信息，请参见[按量付费实例节省停机模式](#)。

更多有关部署集的使用限制及配额，请参见[使用限制部署集](#)章节。

相关操作

- [创建部署集](#)
- [在部署集内创建ECS实例](#)
- [调整实例所属部署集](#)
- [修改部署集信息](#)
- [删除部署集](#)

相关API

- 创建部署集：[CreateDeploymentSet](#)
- 将一台ECS实例加入一个部署集，或者将ECS实例从一个部署集调整到另外一个部署集：[ModifyInstanceDeployment](#)
- 查询部署集列表：[DescribeDeploymentSets](#)
- 修改部署集属性：[ModifyDeploymentSetAttribute](#)
- 删除部署集：[DeleteDeploymentSet](#)

5.2. 创建部署集

使用部署集将ECS实例分散部署在不同的物理服务器上，可提升业务的高可用性和底层容灾能力。本文介绍如何在ECS管理控制台上创建一个部署集。

操作步骤

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择部署与弹性 > 部署集。
3. 在顶部菜单栏左上角处，选择地域。
4. 在部署集列表页面上，单击创建部署集。
5. 在创建部署集页面上，输入部署集名称和描述。

部署集策略目前仅支持高可用，更多关于部署集策略的详情，请参见[部署策略](#)。



6. 单击确定。

后续步骤

创建了部署集后，您可以：

- [在部署集内创建实例](#)
- [将一台ECS实例加入当前部署集](#)

相关文档

- [CreateDeploymentSet](#)

5.3. 在部署集内创建ECS实例

本文介绍了如何通过ECS管理控制台在部署集内创建ECS实例。

前提条件

您在指定地域中已经创建了部署集。具体操作，请参见[创建部署集](#)。

背景信息

一个部署集在一个可用区内只能创建20台ECS实例，一个地域内可以创建（20*可用区）数量的ECS实例。该数值根据您使用云服务器ECS的情况而变化。

创建ECS实例时，您可以搭配使用实例启动模板或者批量创建功能，避免数量限制带来的不便。更多信息，请参见[实例启动模板概述](#)。

操作步骤

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择[部署与弹性](#) > [部署集](#)。
3. 在顶部菜单栏左上角处，选择地域。
4. 在部署集列表页面，找到目标部署集创建实例。

您可以通过以下方式创建实例：

- 在部署集的操作列中，单击[创建实例](#)。

创建部署集					刷新
部署集ID	部署集名称	部署集描述	实例数量	创建时间	操作
ds-bp1	e		0	2020年3月28日 01:09	创建实例 修改信息 删除
ds-bp1	y	t	0	2020年2月19日 03:33	创建实例 修改信息 删除
ds-bp1	k	test	0	2019年12月13日 24:31	创建实例 修改信息 删除

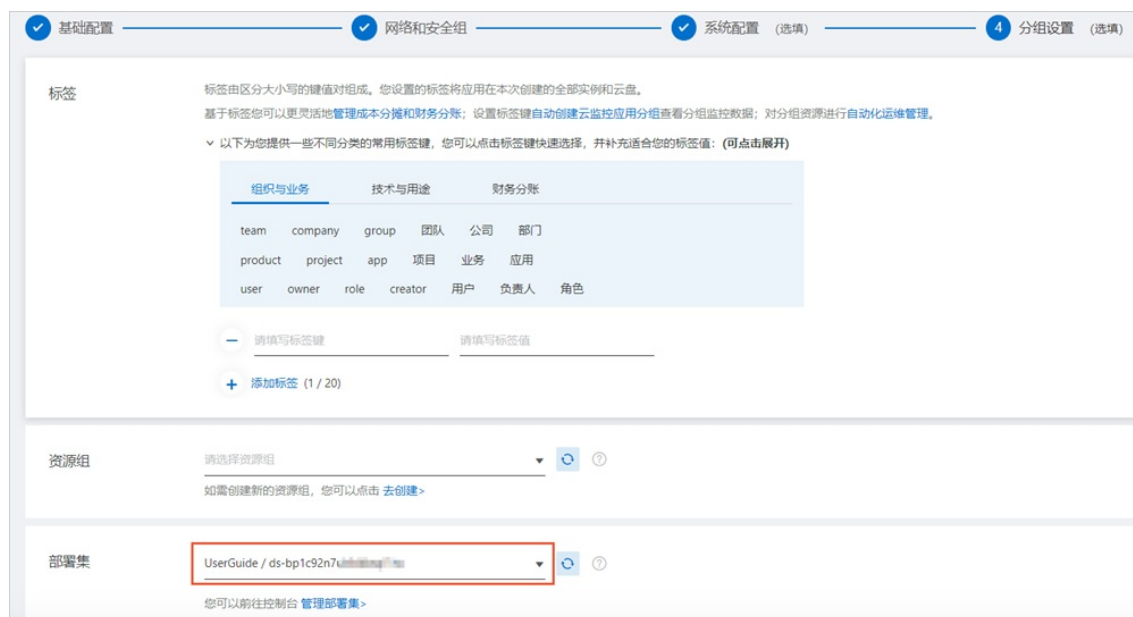
- 单击部署集ID，在实例列表处单击[创建实例](#)。
5. 在跳转后的自定义购买页面完成实例配置选项。
- 具体操作，请参见[使用向导创建实例](#)。创建过程中您需要注意的配置项：

基础配置：

- **地域：**ECS实例必须和目标部署集在同一个地域。
- **可用区：**部署集的每一个可用区内均不能超过20台ECS实例。
- **实例：**部署集支持的实例规格族信息：
 - c7、g7、r7、c6、g6、r6、c5、g5、r5、c6e、g6e、r6e、c7se、g7se、r7se、r6se、c7t、g7t、r7t、c7a、g7a、r7a、c6a、g6a、r6a、g5ne、re6、re4、ic5
 - hfc7、hfg7、hfr7、hfc6、hfg6、hfr6、hfc5、hfg5
 - d2s、d2c、d1、d1ne、d1-c14d3、d1-c8d3
 - i3、i3g、i2、i2g、i2ne、i2gne、i1
 - se1ne、sn1ne、sn2ne、se1
 - ebmg5、sccgn6、scch5、sccg5、scch5s、sccg5s
 - s6、t6、xn4、mn4、n4
 - gn6i

您也可以调用[DescribeDeploymentSetSupportedInstanceTypeFamily](#)获取部署集最新支持的实例规格族信息。

- （可选）您可以在[购买实例数量](#)处指定此次操作的ECS实例数量，并需要考虑部署集的当前可用区内已有的ECS实例数量。
- （可选）**系统配置 > 有序后缀：**批量创建ECS实例后，可以为实例名称和主机名添加有序后缀。有序后缀从001开始递增，最大不会超过999。
- **分组配置 > 部署集：**选择目标部署集即可。



- (可选) 确认订单 > 保存为启动模板：保存配置为启动模板，便于下次快速创建ECS实例。更多信息，请参见[实例启动模板概述](#)。

6. 确认无误后，单击创建实例。

7. 在左侧导航栏，单击部署与弹性 > 部署集，您可以查看在部署集内成功创建的实例信息。

后续步骤

创建ECS实例后，您可以：

- 在部署集内查看并管理ECS实例。更多信息，请参见实例相关操作文档。
- 更换ECS实例所在的部署集。具体操作，请参见[调整实例所属部署集](#)。

相关文档

- [RunInstances](#)

5.4. 调整实例所属部署集

通过调整实例所属部署集功能，您可以将一台已加入一个部署集的实例更换到另一个部署集，或者将一台未加入部署集的实例添加到部署集。

前提条件

目标实例必须处于已停止或者运行中状态。

操作步骤

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择实例与镜像 > 实例。
3. 在顶部菜单栏左上角处，选择地域。
4. 找到目标实例，在操作列中，单击更多 > 实例设置 > 调整实例所属部署集。
5. 在调整实例所属部署集对话框中，选择目标部署集，并设置是否强制调整。

强制调整的设置影响如下：

- 是：允许实例更换宿主机关，可能导致实例重启

否：不允许实例更换宿主机，只在当前宿主机上加入部署集，可能导致调整部署集失败。

6. 单击**确定**。

相关文档

- [ModifyInstanceDeployment](#)

5.5. 修改部署集信息

创建了部署集后，您可以重新修改部署集名称和描述。

前提条件

请确保您已经创建了部署集。具体操作，请参见[创建部署集](#)。

操作步骤

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择**部署与弹性 > 部署集**。
3. 在顶部菜单栏左上角处，选择地域。
4. 找到需要编辑的部署集。
5. 在下列两种方式中，任选一种方式修改信息：
 - 鼠标悬浮至**部署集名称**处，单击出现的图标，在弹出的对话框中重新填写部署集名称和描述。
 - 在目标部署集的**操作**列中，单击**修改信息**，在弹出的对话框中重新填写部署集名称和描述。
6. 单击**确定**。

相关文档

- [ModifyDeploymentSetAttribute](#)

5.6. 删除部署集

您可以删除不再需要的部署集以确保不超过使用限制。

前提条件

请确保部署集内没有实例。如果存在实例，必须释放实例后才能删除部署集。

操作步骤

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择**部署与弹性 > 部署集**。
3. 在顶部菜单栏左上角处，选择地域。
4. 在**部署集**页面，选择需要删除的部署集，在**操作**列单击**删除**。

部署集

设置资源管理策略 设置

创建部署集 操作

刷新

部署集ID	部署集名称	部署集描述	实例数量	创建时间	操作
☐ ds-xxxxxxs1v	hpc-null-deploymentset	autogen deploymentset, do not delete	0	2019年7月24日 07:24	创建实例 修改信息 删除
☐ ds-xxxxxxy1qx	hpc-null-deploymentset	autogen deploymentset, do not delete	0	2019年7月24日 07:18	创建实例 修改信息 删除
☐ ds-xxxxxxqq5	hpc-null-deploymentset	autogen deploymentset, do not delete	0	2019年7月24日 07:14	创建实例 修改信息 删除
☐ ds-xxxxxxn3x	InternalUse	DataProcessingCluster	0	2018年9月19日 00:32	创建实例 修改信息 删除
☐ ds-xxxxxxqkc	UserGuide	MyFirstDS	0	2018年9月18日 11:57	创建实例 修改信息 删除

5. 在弹出的对话框中，单击**确定**。

相关文档

- [DeleteDeploymentSet](#)

6. 弹性供应

6.1. 弹性供应概述

弹性供应是一个快速交付ECS实例集群的方案，简单配置后即可自动在多个可用区内交付不同计费方式（按量付费和抢占式实例）、多种实例规格的实例集合，提升批量交付大量实例的效率。弹性供应提供多种供应策略，有效缓解抢占式实例被回收对业务的影响，让您更稳定地使用低价算力。

弹性供应简介

弹性供应组根据您设置的资源池、目标容量、供应策略等属性自动交付实例集群，您无需逐一关注实例的创建过程或者计算单台实例的成本。以下为弹性供应组的主要属性：

- 资源池：单可用区和单实例规格的组合形成一个资源池，指定多个可用区、多个实例规格，弹性供应组在创建实例时即可有多个资源池可以选择。
- 目标容量：期望通过弹性供应组交付的算力，分为抢占式实例容量和按量付费实例容量，容量单位支持实例个数、vCPU个数或内存大小。

弹性供应组默认使用抢占式实例满足目标容量的需求，您可以手动指定按量付费实例容量满足最低算力的需求。如果指定了按量付费实例容量，优先创建按量付费实例，然后创建抢占式实例补齐剩余容量。但如果因库存不足导致不能创建抢占式实例，会创建按量付费实例满足目标容量的需求。

- 供应策略：创建实例时选择资源池的策略，例如尽量选择单价低的资源池以降低成本，尽量在不同可用区均衡创建实例以提高集群可用性。

应用场景

弹性供应组和抢占式实例的应用场景类似，即无状态应用的场景，例如可横向伸缩的Web站点服务、图像渲染、大数据分析、并行计算等。更多信息，请参见[抢占式实例概述](#)。

使用说明

在创建弹性供应组前，请确定以下内容：

- 确定弹性供应组的实例基本配置，即启动模板。更多信息，请参见[实例启动模板概述](#)。

弹性供应组沿用启动模板中指定的镜像、安全组、登录凭证等属性，但不会直接使用交换机（决定实例所在可用区）和实例规格属性，而是使用单独指定的交换机和实例规格。

- 确定弹性供应组的资源池，实现多可用区、多实例规格。

单可用区和单实例规格的组合形成一个资源池，当某些资源池库存不足时，弹性供应组会自动切换到其他资源池继续尝试创建实例。资源池越多，创建实例时的可选范围越大，创建成功率越高。

 **说明** 单可用区下仅支持指定一个交换机。如果指定了同一个可用区下的多个交换机，仅第一个交换机生效。

- 确定弹性供应组的目标容量以及不同类型容量的占比。

弹性供应组目标容量的单位支持实例个数、vCPU个数和内存大小，通过权重表示单台实例可满足的容量大小。权重规则如下：

- 如果目标容量单位为实例个数，各实例规格的权重都是相同的。
- 如果目标容量单位为vCPU个数，各实例规格的权重和vCPU个数有关，vCPU个数越多权重越大，满足目标容量所需实例的数量越少。

- 如果目标容量是实例规格多个维度因子（例如vCPU、内存）的组合，请您自行评估指定实例规格能为目标容量贡献的算力，然后手动为各实例规格设置权重，权重越大代表该实例规格能贡献的算力越大。

 **说明** 仅在调用API `CreateAutoProvisioningGroup` 创建弹性供应组时支持手动设置权重。

按量付费实例容量用于满足最低算力的需求，在此基础上使用抢占式实例可以大大降低成本。

● 确定供应策略。

弹性供应组支持的策略类型如下表所示。

策略类型	适用于	支持设置策略的方式	说明
容量优化策略 (capacity-optimized)	抢占式实例	- 控制台（抢占式实例）：供应策略设置为容量优化策略。 - API（抢占式实例）：SpotAllocationStrategy取值为 <i>capacity-optimized</i>。	综合考虑抢占式实例的回收率和价格，使用创建成功率高且成本相对较低的资源池创建实例，有效降低抢占式实例被回收的次数，保证容量稳定。
多可用区均衡策略 (diversified)	抢占式实例	- 控制台（抢占式实例）：供应策略设置为多可用区均衡策略。 - API（抢占式实例）：SpotAllocationStrategy取值为 <i>diversified</i>。	优先在多个可用区中均衡地创建实例，规避因单可用区库存不足导致实例创建失败，有效提高应用容灾能力。
成本最优策略 (lowest-price)	抢占式实例、 按量付费实例	- 控制台（抢占式实例）：供应策略设置为成本最优策略。 - API（抢占式实例）：SpotAllocationStrategy取值为 <i>lowest-price</i>。 - API（按量付费实例）：PayAsYouGoAllocationStrategy取值为 <i>lowest-price</i>。	使用成本最低的资源池创建实例，有效降低成本。此处成本最低指vCPU单价最低。 如果您调用 <code>CreateAutoProvisioningGroup</code> 创建弹性供应组，可以针对抢占式实例指定使用成本最优策略和使用的资源池数量（SpotInstancePoolsToUseCount），同时使用成本最低的多个资源池创建实例。 例如，目标容量为100台抢占式实例，SpotInstancePoolsToUseCount取值为5，则使用每个资源池创建20台实例。当单个资源池对应的抢占式实例被回收时，其他资源池对应的抢占式实例仍然可用，可以有效提高服务的可用性。

策略类型	适用于	支持设置策略的方式	说明
优先级策略 (prioritized)	按量付费实例	API（按量付费实例）：PayAsYouGoAllocationStrategy取值为 <i>prioritized</i> 。	按照资源池的优先级从高到低依次创建实例，在高优先级资源池的库存不足时，才会使用低优先级资源池创建实例。 需要调用 CreateAutoProvisioningGroup 创建弹性供应组，指定使用优先级策略和资源池的优先级（ <code>LaunchTemplateConfig.N.Priority</code> ）。

 **说明** 仅在调用API [CreateAutoProvisioningGroup](#)创建弹性供应组时支持为按量付费实例设置供应策略，否则默认应用成本最优策略。

- 确定价格上限。

弹性供应组支持针对单资源池和全部资源池设置抢占式实例的价格上限。如果超过了价格上限，弹性供应组即使没有达到目标容量也会停止创建实例，保证不会超过您的预算。

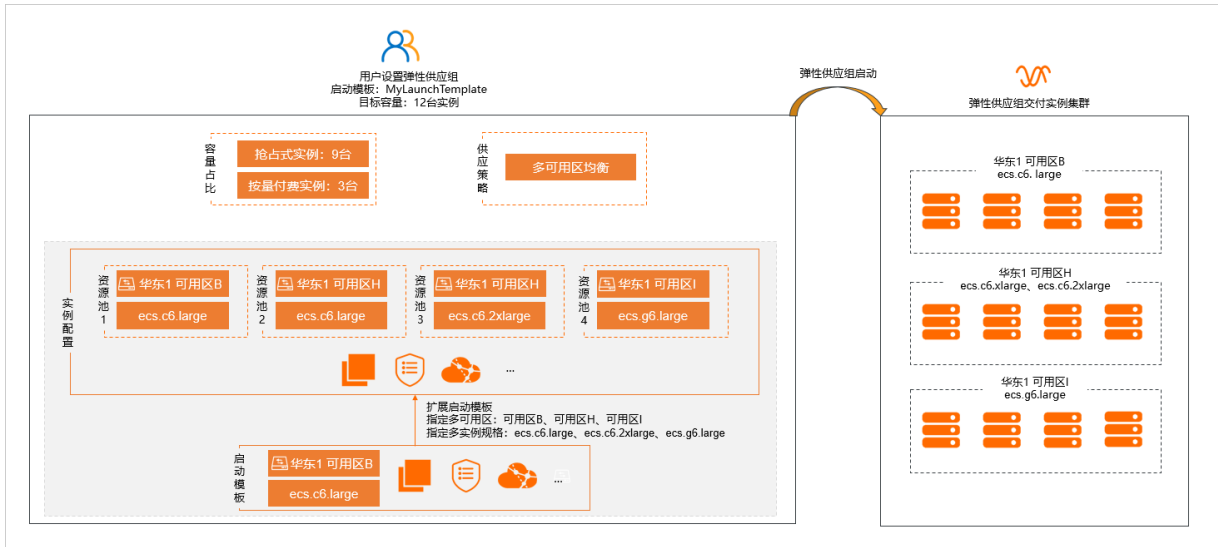
针对按量付费实例，您可以搭配预留实例券、节省计划等折扣计划，进一步降低成本。更多信息，请参见[预留实例券概述](#)和[节省计划概述](#)。

- 确定是否持续保持目标容量。

弹性供应组类型分为一次性交付和持续保持。持续保持类型的弹性供应组支持自动检查实例的健康状态，以及比较弹性供应组的实时容量和目标容量，当实例因不健康或库存不足被移除时，弹性供应组会自动创建新的实例补齐容量，保证实时容量满足算力需求。

弹性供应组启动后会自动根据供应策略选择合适的资源池，尝试交付满足目标容量需求的实例集群。例如，您希望基于启动模板MyLaunchTemplate交付一个包含12台实例的集群，并尽量提高集群可用性，则设置弹性供应组的流程如下：

1. 基于MyLaunchTemplate扩展出多个资源池。
2. 确定抢占式实例容量和按量付费实例容量的占比。
3. 选择提高集群可用性的供应策略。



计费

弹性供应为免费功能，但是您需要为创建出的ECS实例付费。弹性供应支持创建抢占式实例和按量付费实例，更多信息，请参见[抢占式实例概述](#)和[按量付费](#)。

警告 请确保账号可用额度充足。如果账号欠费，所有后付费的ECS实例（包括按量付费实例和抢占式实例）都会停机，停机时间点请参见[按量付费](#)。此时弹性供应组不能交付新的ECS实例，同时由于弹性供应组具有健康检查能力，已停机的ECS实例会被判定为不健康，然后被移出弹性供应组并释放。

使用限制

- 弹性供应组不支持跨地域供应实例。
- 每个弹性供应组最多指定一个启动模板的指定版本作为实例基本配置，但是您可以扩展模板中的实例规格形成多个资源池。
- 弹性供应组下最多支持设置20个资源池（即可用区+实例规格的组合）。
- 每个弹性供应组下最多创建1000台实例。

6.2. 管理弹性供应服务关联角色

本文介绍如何通过弹性供应服务关联角色授予弹性供应服务访问关联云资源的权限。

前提条件

如果您使用RAM用户，必须为RAM用户添加使用弹性供应服务的权限，才能够管理弹性供应服务关联角色。具体操作，请参见[为RAM用户授权](#)。

需要添加的权限内容如下：

说明 请将<account ID>替换为您的阿里云主账号的ID。

```
{
  "Statement": [
    {
      "Action": [
        "ram:CreateServiceLinkedRole"
      ],
      "Resource": "acs:ram:*:<account ID>:role/*",
      "Effect": "Allow",
      "Condition": {
        "StringEquals": {
          "ram:ServiceName": [
            "autoprovisioning.ecs.aliyuncs.com"
          ]
        }
      }
    }
  ],
  "Version": "1"
}
```

背景信息

弹性供应服务关联角色（AliyunServiceRoleForAutoProvisioning）是访问控制提供的一种服务关联角色，用于为弹性供应服务获取关联云服务的访问权限。通过AliyunServiceRoleForAutoProvisioning，弹性供应服务可以获得云服务器ECS、专有网络VPC、云数据库RDS、负载均衡SLB、运维编排OOS、消息服务MNS、云监控CMS服务的访问权限。更多服务关联角色的说明，请参见[服务关联角色](#)。

创建AliyunServiceRoleForAutoProvisioning

在您创建供应组时，系统会检查当前账号是否已有AliyunServiceRoleForAutoProvisioning，如果不存在则提示需要创建该角色，确认后就会自动创建。您也可以手动创建AliyunServiceRoleForAutoProvisioning，具体操作请参见[创建服务关联角色](#)。

服务关联角色拥有的权限由对应的云服务定义和使用，您不能为服务关联角色添加、修改或删除权限。您可以在角色详情中查看角色拥有的权限，具体操作请参见[查看RAM角色基本信息](#)。

删除AliyunServiceRoleForAutoProvisioning

如果您暂时不需要使用AliyunServiceRoleForAutoProvisioning，例如不需要供应组创建和管理资源、确定不使用该角色的影响等，可以删除AliyunServiceRoleForAutoProvisioning，具体操作请参见[删除RAM角色](#)。

 **说明** 删除AliyunServiceRoleForAutoProvisioning前，必须先删除当前账号下所有地域中的供应组，否则会提示删除失败。

删除AliyunServiceRoleForAutoProvisioning后，供应组不能再创建和管理资源。

6.3. 创建弹性供应组

本文介绍如何在控制台创建弹性供应组，弹性供应组会按照您的设置自动交付实例集群。

前提条件

- 已授予账号操作弹性供应组相关云资源的权限。

?

说明

首次进入弹性供应功能页面时，按提示授权系统RAM角色 AliyunECSAutoProvisioningGroupRole 即可。

- 已创建实例启动模板。具体操作，请参见[创建实例启动模板](#)。
弹性供应组使用实例启动模板的指定版本作为配置来源，沿用配置来源中的镜像、安全组、登录凭证等属性。

操作步骤

1. 进入弹性供应功能页面。
 - i. 登录[ECS管理控制台](#)。
 - ii. 在顶部菜单栏左上角处，选择地域。
 - iii. 在左侧导航栏，选择部署与弹性 > 弹性供应。
2. 单击创建供应组。
3. 在供应组名称区域，输入弹性供应组的名称。
名称长度为2~128个字符，必须以大小写字母或中文开头，不能以http://和https://开头。可以包含数字、下划线（_）、连字符（-）和句点（.）。
4. 在目标容量区域，设置容量相关选项。
容量相关选项决定弹性供应组交付的算力总和，以及抢占式实例和按量实例的算力占比，选项如下表所示。

容量相关选项	说明
目标容量	计划使用弹性供应组交付的算力，单位支持实例个数和vCPU个数。选中使用按量实例提供部分算力后，可以继续指定使用按量实例提供的算力。默认全部交付抢占式实例，设置按量实例容量大于0时才会交付按量实例。 ? 说明 如果目标容量是实例规格多个维度因子（例如vCPU、内存）的组合，请调用API CreateAutoProvisioningGroup创建弹性供应组，并为您指定的实例规格设置权重。
按量实例容量	目标容量中需要使用按量付费方式购买的实例个数或vCPU个数。由于抢占式实例存在回收机制，您可以通过指定使用按量实例提供的算力保证满足最低算力需求。

- 弹性供应组通过实例规格权重表示单台实例在弹性供应组中提供的容量大小。
- 如果目标容量单位为实例个数，各实例规格的权重都是相同的。
 - 如果目标容量单位为vCPU个数，各实例规格的权重和vCPU个数有关，vCPU个数越多权重越大，满足目标容量所需实例的数量越少。使用不同实例规格交付时的权重示例如下表所示。

实例规格	vCPU	权重
ecs.c6.large	2	2
ecs.c6.xlarge	4	4
ecs.c6.2xlarge	8	8

 **说明** 在控制台创建弹性供应组时，您无需为每个实例规格手动设置权重，弹性供应会根据vCPU个数自动赋予实例规格权重。

假设目标容量为24 vCPU，有多种交付方式可以满足要求，弹性供应组会根据每个实例规格对应的权重，自动组合交付。例如：

- 12台ecs.c6.large
- 8台ecs.c6.large、1台ecs.c6.2xlarge
- 4台ecs.c6.large、2台ecs.c6.xlarge、1台ecs.c6.2xlarge
- 如果目标容量是实例规格多个维度因子（例如vCPU、内存）的组合，请评估您指定的实例规格能为目标容量贡献的算力，然后为各实例规格设置权重，权重越大代表该规格能贡献的算力越大。假设某应用要求实例集群的总算力为20 vCPU、48 GiB，且单节点算力是2 vCPU 4 GiB的倍数，您可以设置弹性供应组的目标容量为48，并为备选实例规格设置权重，示例如下表所示。

实例规格	vCPU	内存	权重
ecs.c6.large	2	4 GiB	4
ecs.c6.xlarge	4	8 GiB	8
ecs.c6.2xlarge	8	16 GiB	16

有多种交付方式可以满足要求，弹性供应组会根据每个实例规格对应的权重，自动组合交付。例如：

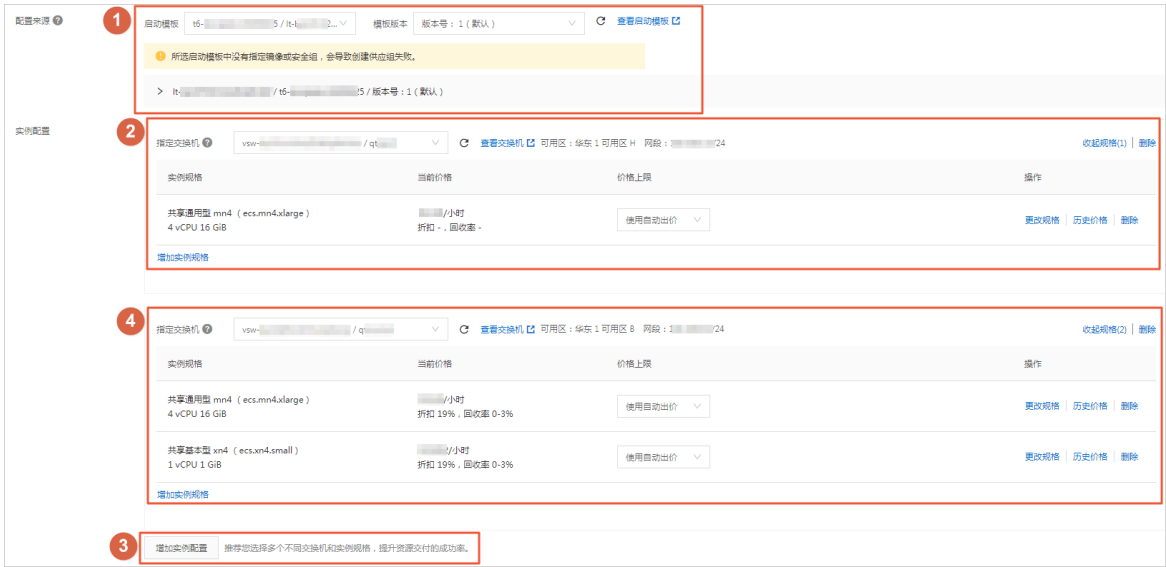
- 12台ecs.c6.large
- 8台ecs.c6.large、1台ecs.c6.2xlarge
- 4台ecs.c6.large、2台ecs.c6.xlarge、1台ecs.c6.2xlarge

 **说明** 如果目标容量是实例规格多个维度因子（例如vCPU、内存）的组合，请调用API [CreateAutoProvisioningGroup](#)创建弹性供应组，并为您指定的实例规格设置权重。

5. 在配置来源和实例配置区域，设置实例属性相关选项。

弹性供应组支持多可用区、多实例规格。如果单可用区下或单实例规格的库存不足导致实例创建失败，弹性供应组会自动尝试在其他可用区或使用其他实例规格创建实例，有效提高创建成功率。

您可以通过指定多个属于不同可用区的交换机实现多可用区，通过增加扩展实例规格实现多实例规格。示例如下图所示，更多示例，请参见[弹性供应组设置示例](#)。



该示例中包括了2个可用区、3个实例规格，上图中各序号对应的操作说明如下表所示。

序号	操作说明
①	选择一个实例启动模板的指定版本作为配置来源。 说明 创建实例时使用②、④中指定的交换机和实例规格，但沿用该配置来源中的镜像、安全组、登录凭证等属性。
②	完成实例配置，默认使用配置来源中指定的交换机和实例规格。您也可以自行选择其他交换机，选择和增加其他实例规格。 说明 弹性供应组至少需要一条实例配置。
③	增加1条实例配置，以实现多个可用区。
④	完成增加的实例配置，其中至少设置以下信息： - 指定交换机：选择实例所属的交换机，请确保该交换机和②中交换机所属的可用区不同。 注意 如果选择了同一个可用区下的多个交换机，仅第一个交换机生效。 - 增加实例规格：您可以选择多个实例规格提升实例创建成功率，本示例中选择了两个实例规格。选择实例规格时，页面中推荐了相同实例大小和相同实例配比的实例规格，帮助您筛选合适的实例规格，您也可以自行选择其他实例规格。

除设置交换机和实例规格外，您还可以为每个实例规格设置抢占式实例的价格上限，支持通过以下形式设置价格上限：

- 自动出价：自动按抢占式实例的实时市场价出价，保证不会因出价低于抢占式实例市场价导致实例创建失败，且抢占式实例市场价较低时成本也随之降低。
- 最高价（固定价格）：设置固定的价格，抢占式实例市场价高于固定价格时实例创建失败，保证不超

过预算。

- 最高价（同期按量价格）：设置相对按量实例价格的比例，保证相对按量实例可以节省一定比例的成本。例如设置为同期按量价格的50%，则抢占式实例市场价高于50%的按量实例价格时实例创建失败。

 **说明** 推荐在设置前参考抢占式实例市场价的趋势，避免因上限过低创建不了足量的实例，或因上限过高超过预算。单击操作列下的历史价格即可查看历史价格。

6. 在供应策略区域，设置创建实例时的策略。

可选项如下表所示。

可选项	说明
容量优化策略	综合考虑抢占式实例的历史价格波动和回收率，自动选择综合成本最优的实例规格创建实例。  说明 抢占式实例可能因价格、库存等因素被回收，优先考虑回收率低的实例规格。
成本最优策略	按vCPU单价从低到高尝试创建实例。
多可用区均衡策略	在实例配置中指定了不同可用区下的交换机时才有效果，在多个可用区之间均匀创建实例。 回收抢占式实例时会按实例规格来逐一回收。同一实例规格族内的实例资源共享，例如无法创建ecs.c6.large时，可能原因是c6规格族都即将库存不足，ecs.c6.xlarge等规格也可能无法创建。  说明 如果选用多可用区均衡策略，建议您配置不同的实例规格族，避免所有实例同时被回收，保证实例集群的高可用性。

7. 设置高级选项。

选项如下表所示。

高级选项	说明
供应组类型	一次性交付：弹性供应组启动后尝试交付指定容量的实例集群。只尝试一次，即使交付失败了也不会重试。 持续保持：弹性供应组启动后尝试交付指定容量的实例集群，并且会持续比较实时容量和目标容量，如果两者存在差距则自动创建实例或扩容。
启动时间	指定弹性供应组开始交付指定目标容量的时间，与结束时间结合使用，确定弹性供应组的有效时段。 立即启动：创建成功后立即开始交付实例集群。 指定启动时间：自定义弹性供应组的启动时间。

高级选项	说明
结束时间	与启动时间结合使用，确定弹性供应组的有效时段。 ◦ 无限期：除非您主动删除弹性供应组，弹性供应组始终存在。 ◦ 指定结束时间：自定义弹性供应组的过期时间。
全局抢占式实例的价格上限	弹性供应组内抢占式实例的全局价格上限，适用于所有实例规格。如果该全局价格上限和实例配置中单实例规格的价格上限不一致，以较低者为准。 ◦ 自动出价：自动按抢占式实例的实时市场价出价，保证不会因出价低于抢占式实例市场价导致实例创建失败，且抢占式实例市场价较低时成本也随之降低。 ◦ 设置价格：设置固定的价格，抢占式实例市场价高于固定价格时实例创建失败，保证不超过预算。

8. 单击创建供应组。

执行结果

弹性供应组创建成功后，会在您指定的时间启动并尝试交付实例集群。如果弹性供应组的类型为持续保持，弹性供应组还会持续维护实例集群，在抢占式实例被回收时自动尝试创建新的实例，缩小实时容量和目标容量的差距，并及时替换不健康的实例。

相关文档

- [CreateAutoProvisioningGroup](#)

6.4. 弹性供应组设置示例

本文介绍如何根据示例场景中的需求确定弹性供应组的相关配置。

机器学习场景示例

假设您计划在未来一周完成一项机器学习任务，用于分析住房抵押贷款的风险因素，对实例集群的需求如下：

- 目标地域：华东1（杭州）。
- 实例配备NVIDIA V100 GPU计算卡，单实例GPU显存不大于32 GB。
- 目标容量：20台实例。
- 尽量节省成本，只用抢占式实例，可以接受集群规模不达标。
- 任务结束后释放实例。

考虑以上需求，弹性伸缩组的设置示例如下表所示。

配置区域	配置参数	示例说明
容量配置	目标容量	由于目标容量为20台实例，且仅使用抢占式实例，设置如下： • 单位选择实例个数。 • 容量设置为20。

配置区域	配置参数	示例说明
实例配置	实例配置	由于要求实例配备NVIDIA V100 GPU计算卡，单机GPU显存不大于32 GB： 1. 符合要求的实例规格为ecs.gn6v-c8g1.2xlarge、ecs.gn6e-c12g1.3xlarge。 ? 说明 详细的实例规格信息，请参见实例规格族。 2. 查询在华东1（杭州）地域中，仅可用区H、可用区I中有ecs.gn6v-c8g1.2xlarge、ecs.gn6e-c12g1.3xlarge的库存。 ? 说明 您可以前往ECS实例可购买地域，查看实例在各地域的可购情况。 根据可用区和实例规格情况，即可增加实例配置： 1. 增加第1条实例配置： ◦ 指定华东1（杭州）可用区H中的交换机。 ◦ 增加实例规格ecs.gn6v-c8g1.2xlarge、ecs.gn6e-c12g1.3xlarge。 2. 增加第2条实例配置： ◦ 指定华东1（杭州）可用区I中的交换机。 ◦ 增加实例规格ecs.gn6v-c8g1.2xlarge、ecs.gn6e-c12g1.3xlarge。 增加实例配置后，形成以下资源池： • 可用区H+ecs.gn6v-c8g1.2xlarge • 可用区H+ecs.gn6e-c12g1.3xlarge • 可用区I+ecs.gn6v-c8g1.2xlarge • 可用区I+ecs.gn6e-c12g1.3xlarge
	供应策略	选择 成本最优策略 。弹性供应组启动后，会使用成本最低的资源池交付实例集群。
高级选项	供应组类型	由于需要尽量节省成本，可以接受集群不达标，因此选择 一次性交付 。
	启动时间、结束时间	按照计划的未来一周设置启动时间和结束时间。
	组内实例关停设置	• 由于需要任务结束后释放实例，选中供应组过期时关停供应组内实例。 • 由于需要尽量节省成本，选中超过容量时关停供应组内实例。

购票网站场景示例

假设您需要搭建一个购票网站，提供覆盖日常和高峰时段的购票服务，对实例集群的需求如下：

- 目标地域：华东1（杭州）。
- 单实例vCPU个数不大于8。
- 目标容量：80 vCPU。
- 最低容量：60 vCPU。
- 在集群最低容量基础上优化网站访问体验，尽量节省成本。

- 具有容灾能力。

考虑以上需求，弹性伸缩组的设置示例如下表所示。

配置区域	配置参数	示例说明
容量配置	目标容量	由于目标容量为80 vCPU，最低容量为60，设置如下： ● 单位选择vCPU个数。 ● 容量设置为80。 ● 选中使用按量实例提供部分算力。
	按量实例容量	容量设置为60，使用按量实例容量满足最低容量需求。

配置区域	配置参数	示例说明
实例配置	实例配置	实例规格族考虑适合搭建Web前端服务器的c6，由于要求单实例vCPU个数不大于8： 1. 符合要求的实例规格为ecs.c6.large、ecs.c6.xlarge、ecs.c6.2xlarge。  说明 详细的实例规格信息，请参见实例规格族。 2. 查询在华东1（杭州）地域中，可用区H、可用区I、可用区J中有ecs.c6.large、ecs.c6.xlarge、ecs.c6.2xlarge的库存。  说明 您可以前往ECS实例可购买地域，查看实例在各地域的可购情况。 根据可用区和实例规格情况，即可增加实例配置： 1. 增加第1条实例配置： ◦ 指定华东1（杭州）可用区H中的交换机。 ◦ 增加实例规格ecs.c6.large、ecs.c6.xlarge、ecs.c6.2xlarge。 2. 增加第2条实例配置： ◦ 指定华东1（杭州）可用区I中的交换机。 ◦ 增加实例规格ecs.c6.large、ecs.c6.xlarge、ecs.c6.2xlarge。 3. 增加第3条实例配置： ◦ 指定华东1（杭州）可用区J中的交换机。 ◦ 增加实例规格ecs.c6.large、ecs.c6.xlarge、ecs.c6.2xlarge。 增加实例配置后，形成以下资源池： • 可用区H+ecs.c6.large • 可用区H+ecs.c6.xlarge • 可用区H+ecs.c6.2xlarge • 可用区I+ecs.c6.large • 可用区I+ecs.c6.xlarge • 可用区I+ecs.c6.2xlarge • 可用区J+ecs.c6.large • 可用区J+ecs.c6.xlarge • 可用区J+ecs.c6.2xlarge
	供应策略	选择多可用区均衡策略。弹性供应组启动后，尝试在多个可用区中均衡地创建实例，规避单可用区库存不足导致实例创建失败，可以有效提高应用容灾能力。
	供应组类型	由于需要持续提供购票服务，因此选择持续保持。

配置区域	配置参数	示例说明
高级选项	启动时间、结束时间	由于需要持续提供购票服务，指定弹性供应组立即启动并无限期保留。
	组内实例关停设置	由于需要尽量节省成本，选中超过容量时关停供应组内实例。

由于目标容量单位为vCPU个数，每台实例的权重和实例规格的vCPU个数有关，因此各实例规格还有对应权重单价，示例如下表所示。

 **说明** 表中价格仅为示例，请以售卖页中的实时价格为准。

实例规格	vCPU	按量价格（元）	权重	权重单价（元）
ecs.c6.large	2	0.39/时	2	0.195/时
ecs.c6.xlarge	4	0.78/时	4	0.195/时
ecs.c6.2xlarge	8	1.56/时	8	0.195/时

弹性供应组在交付实例集群时，首先尝试满足多可用区均衡策略，在多个可用区中均衡地创建实例。同时尽量使用权重单价较低的方案，如果各实例规格的权重单价相同，则随机选择一种方案。

使用成本最低的多个资源池

如果您希望尽量降低成本，并降低单个资源池对应的抢占式实例被回收的影响，可以设置弹性供应组同时使用成本最低的多个资源池创建实例。

仅支持通过调用CreateAutoProvisioningGroup完成该设置，请保证SpotAllocationStrategy取值为 *lowest-price*，并设置SpotInstancePoolsToUseCount。例如，下表参数的效果为：

- 根据实例启动模板lt-bp1ivgo4p5now3px****扩展出5个资源池。
- 目标容量为30台实例，全部为抢占式实例。
- 创建抢占式实例时，使用成本最低的3个资源池，每个资源池各创建10台实例。

参数	示例值	说明
TotalTargetCapacity	30	目标容量30台实例。
SpotTargetCapacity	30	抢占式实例容量30台实例。
PayAsYouGoTargetCapacity	0	按量付费实例容量0台实例。
SpotAllocationStrategy	lowest-price	抢占式实例采用成本优化策略。
PayAsYouGoAllocationStrategy	lowest-price	按量付费实例采用成本优化策略。
SpotInstancePoolsToUseCount	3	使用成本最低的3个资源池。
LaunchTemplateId	lt-bp1ivgo4p5now3px****	作为配置来源的实例启动模板的ID。

参数	示例值	说明
LaunchTemplateConfig.1.VSwitchId	vsw-bp1ygryo03m39xhsy****	扩展启动模板1（即资源池1）的交换机。
LaunchTemplateConfig.1.InstanceType	ecs.c6e.large	扩展启动模板1（即资源池1）的实例规格。
LaunchTemplateConfig.1.WeightedCapacity	1	扩展启动模板1（即资源池1）的实例规格权重。
LaunchTemplateConfig.2.VSwitchId	vsw-bp16hgf8f3kvtcbyu****	扩展启动模板2（即资源池2）的交换机。
LaunchTemplateConfig.2.InstanceType	ecs.c6e.xlarge	扩展启动模板2（即资源池2）的实例规格。
LaunchTemplateConfig.2.WeightedCapacity	1	扩展启动模板2（即资源池2）的实例规格权重。
LaunchTemplateConfig.3.VSwitchId	vsw-bp1oeawdo9tj2gvjp****	扩展启动模板3（即资源池3）的交换机。
LaunchTemplateConfig.3.InstanceType	ecs.c6e.2xlarge	扩展启动模板3（即资源池3）的实例规格。
LaunchTemplateConfig.3.WeightedCapacity	1	扩展启动模板3（即资源池3）的实例规格权重。
LaunchTemplateConfig.4.VSwitchId	vsw-bp1oeawdo9tj2gvjp****	扩展启动模板4（即资源池4）的交换机。
LaunchTemplateConfig.4.InstanceType	ecs.g6e.xlarge	扩展启动模板4（即资源池4）的实例规格。
LaunchTemplateConfig.4.WeightedCapacity	1	扩展启动模板4（即资源池4）的实例规格权重。
LaunchTemplateConfig.5.VSwitchId	vsw-bp1oeawdo9tj2gvjp****	扩展启动模板5（即资源池5）的交换机。
LaunchTemplateConfig.5.InstanceType	ecs.g6e.2xlarge	扩展启动模板5（即资源池5）的实例规格。
LaunchTemplateConfig.5.WeightedCapacity	1	扩展启动模板5（即资源池5）的实例规格权重。

优先使用指定的资源池

在您希望优先使用指定的资源池创建实例时，例如购买了ecs.c6e.large的可用区级预留实例券，可以设置弹性供应组在创建按量付费实例时优先选择该规格，以享受折扣后抵扣。

仅支持通过调用CreateAutoProvisioningGroup完成该设置，请保证PayAsYouGoAllocationStrategy取值为*prioritized*，并将对应资源池的优先级设置为最高（即0）。例如，下表参数的效果为：

- 根据实例启动模板lt-bp1ivgo4p5now3px****扩展出5个资源池。
- 目标容量为20台实例，包括10台抢占式实例和10台按量付费实例。
- 创建按量付费实例时，优先使用ecs.c6e.large规格对应的资源池（即LaunchTemplateConfig.1）。

参数	示例值	说明
TotalTargetCapacity	20	目标容量20台实例。
SpotTargetCapacity	10	抢占式实例容量10台实例。
PayAsYouGoTargetCapacity	10	按量付费实例容量10台实例。
SpotAllocationStrategy	lowest-price	抢占式实例采用成本优化策略。
PayAsYouGoAllocationStrategy	prioritized	按量付费实例采用成本优化策略。
LaunchTemplateId	lt-bp1ivgo4p5now3px****	作为配置来源的实例启动模板的ID。
LaunchTemplateConfig.1.VSwitchId	vsw-bp1ygry03m39xhsy****	扩展启动模板1（即资源池1）的交换机。
LaunchTemplateConfig.1.InstanceType	ecs.c6e.large	扩展启动模板1（即资源池1）的实例规格。
LaunchTemplateConfig.1.WeightedCapacity	1	扩展启动模板1（即资源池1）的实例规格权重。
LaunchTemplateConfig.1.Priority	0	扩展启动模板1（即资源池1）的优先级，0为最高。
LaunchTemplateConfig.2.VSwitchId	vsw-bp16hgf8f3kvtcbyu****	扩展启动模板2（即资源池2）的交换机。
LaunchTemplateConfig.2.InstanceType	ecs.c6e.xlarge	扩展启动模板2（即资源池2）的实例规格。
LaunchTemplateConfig.2.WeightedCapacity	1	扩展启动模板2（即资源池2）的实例规格权重。
LaunchTemplateConfig.3.VSwitchId	vsw-bp1oeawdo9tj2gvjp****	扩展启动模板3（即资源池3）的交换机。
LaunchTemplateConfig.3.InstanceType	ecs.c6e.2xlarge	扩展启动模板3（即资源池3）的实例规格。
LaunchTemplateConfig.3.WeightedCapacity	1	扩展启动模板3（即资源池3）的实例规格权重。
LaunchTemplateConfig.4.VSwitchId	vsw-bp1oeawdo9tj2gvjp****	扩展启动模板4（即资源池4）的交换机。
LaunchTemplateConfig.4.InstanceType	ecs.g6e.xlarge	扩展启动模板4（即资源池4）的实例规格。

参数	示例值	说明
LaunchTemplateConfig.4.WeightedCapacity	1	扩展启动模板4（即资源池4）的实例规格权重。
LaunchTemplateConfig.5.VSwitchId	vsw-bp1oeawdo9tj2gvjp****	扩展启动模板5（即资源池5）的交换机。
LaunchTemplateConfig.5.InstanceType	ecs.g6e.2xlarge	扩展启动模板5（即资源池5）的实例规格。
LaunchTemplateConfig.5.WeightedCapacity	1	扩展启动模板5（即资源池5）的实例规格权重。

6.5. 查看弹性供应组

本文介绍如何查看弹性供应组的设置信息、已创建实例的信息和调度任务执行情况。

操作步骤

1. 登录[ECS管理控制台](#)。
2. 在左侧导航栏，选择部署与弹性 > 弹性供应。
3. 在顶部菜单栏左上角处，选择地域。
4. 找到待查看的弹性供应组，在操作区域单击详情。

供应组详情页面包含的信息如下表所示。

信息分类	描述
供应组设置	◦ 基本信息：弹性供应组标识相关的信息，包括弹性供应组的类型、状态、启动时间、结束时间等。 ◦ 供应组容量：弹性供应组的实时容量信息，包括总当前容量、总目标容量、抢占式实例当前容量、抢占式实例目标容量、按量实例当前容量、按量实例目标容量。 ◦ 容量相关设置：弹性供应组容量相关的设置，包括抢占式实例扩容策略（即供应策略）、供应组过期或超过目标容量时的关停设置等。
模板设置	◦ 启动模板原始设置：配置来源，包括实例启动模板的ID、实例启动模板版本等信息。 ◦ 启动模板的扩展设置：弹性供应组当前使用中的实例规格、创建实例的可用区等信息。
供应组内实例	弹性供应组已创建实例的列表。
供应组历史	弹性供应组已执行调度任务的记录。如果任务的成功率低，请检查弹性供应组的设置是否合理，例如可选资源池是否过少、价格上限是否过低。

6.6. 修改弹性供应组

本文介绍如何修改弹性供应组的设置，支持修改弹性供应组的名称、目标容量和部分容量相关的设置。

操作步骤

- 1. 登录ECS管理控制台。
- 2. 在左侧导航栏，选择部署与弹性 > 弹性供应。
- 3. 在顶部菜单栏左上角处，选择地域。
- 4. 找到待修改的弹性供应组，在操作区域单击修改。

支持修改的设置如下表所示。

信息分类	描述
基本信息	支持修改弹性供应组的名称。
供应组容量	支持修改以下容量： - 目标容量：计划使用弹性供应组交付的算力，单位支持实例个数和vCPU个数。选中使用按量实例提供部分算力后，可以继续指定使用按量实例提供的算力。 - 按量实例的目标容量：目标容量中需要使用按量付费方式购买的实例个数或vCPU个数。由于抢占式实例存在回收机制，您可以通过指定使用按量实例提供的算力保证满足最低算力需求。 说明 如果您调用APICreateAutoProvisioningGroup创建弹性供应组，目标容量单位支持实例个数、vCPU个数和内存大小。
容量相关设置	支持修改以下容量相关的设置： - 差额容量实例的默认付费类型：支持抢占式实例和按量付费实例。如果（抢占式实例目标容量+按量实例目标容量）<总目标容量，默认创建该类型的实例以满足总目标容量的要求。 - 设置全局的抢占式实例上限价格：支持修改固定价格，抢占式实例市场价高于固定价格时实例创建失败，保证不超过预算。

说明 无论弹性供应组的类型为一次性交付还是持续保持，修改容量或容量相关的设置后均会触发执行一次调度任务。

- 5. 单击确定。

相关文档

- ModifyAutoProvisioningGroup

6.7. 删除弹性供应组

本文介绍如何删除弹性供应组。弹性供应组可以缓解抢占式实例的回收机制带来的不稳定因素，如果不使用弹性供应组，使用抢占式实例时需要及时关注实例的状态。

操作步骤

- 1. 登录ECS管理控制台。

2. 在左侧导航栏，选择**部署与弹性 > 弹性供应**。
3. 在顶部菜单栏左上角处，选择地域。
4. 选择一种方式删除弹性供应组。
 - 单个删除：找到待删除的弹性供应组，在**操作区域**中单击**删除**。
 - 批量删除：选中待删除的一个或多个弹性供应组，单击页面底部的**删除供应组**。
5. 在**删除供应组**对话框中，设置是否删除弹性供应组内的实例。
 - 打开**供应组内实例是否删除**开关，则在删除弹性供应组后释放所有弹性供应组内的实例。
 - 关闭**供应组内实例是否删除**开关，则仅删除弹性供应组，保留所有弹性供应组内的实例。
6. 单击**确定**。

7.Terraform

7.1. Terraform概述

Terraform是一种开源工具，用于安全高效地预览、配置和管理云基础架构和资源。

什么是Terraform

HashiCorp Terraform是一个IT基础架构自动化编排工具，可以用代码来管理维护IT资源。它编写了描述云资源拓扑的配置文件中的基础结构，例如虚拟机、存储账户和网络接口。Terraform的命令行接口（Command Line Interface, CLI）提供一种简单机制，用于将配置文件部署到阿里云或其他任意支持的云上，并对其进行版本控制。更多信息，请参见[HashiCorp Terraform](#)。

同时，Terraform是一个高度可扩展的工具，通过Provider来支持新的基础架构。Terraform能够让您在阿里云上轻松使用简单的模板语言来定义、预览和部署云基础结构。您可以使用Terraform来创建、修改、删除ECS、VPC、RDS、SLB等多种资源。更多信息，请参见[配置语法](#)。

阿里云作为中国第一家与Terraform集成的云厂商，Alibaba Cloud Provider目前已经提供了超过163个Resource和113个Data Source，覆盖计算、存储、网络、负载均衡、CDN、容器服务、中间件、访问控制、数据库等超过35款产品，已经满足了大量大客户的自动化上云需求。更多信息，请参见[Alibaba Cloud Provider](#)。

从Terraform 0.12.2版本开始，阿里云支持将对象存储服务OSS作为标准的Remote State Backend，开始提供远端存储State的能力，在提高state安全性的同时，提升多人协作效率。更多信息，请参见[Remote State Backend](#)。

为了给开发者提供“开箱即用”的使用体验，阿里云提供了丰富多样的Modules和Examples，覆盖计算、存储、网络、中间件、数据库等多个产品和使用场景，欢迎大家使用和贡献自己的Module。更多信息，请参见[Modules](#)和[Examples](#)。

功能优势

- 将基础结构部署到多个云

Terraform适用于多云方案，将相类似的基础结构部署到阿里云、其他云提供商或者本地数据中心。开发人员能够使用相同的工具和相似的配置文件同时管理不同云提供商的资源。

- 自动化管理基础结构

Terraform能够创建配置文件的模板，以可重复、可预测的方式定义和预配ECS资源，减少人为因素导致的部署和管理错误。能够多次部署同一模板，创建相同的开发、测试和生产环境。

- 基础架构即代码（Infrastructure as Code）

可以用代码来管理维护资源。允许保存基础设施状态，从而使您能够跟踪对系统（基础设施即代码）中不同组件所做的更改，并与其他人共享这些配置。

- 降低开发成本

您通过按需创建开发和部署环境来降低成本。并且，您可以在系统更改之前进行评估。

应用场景

Terraform的应用场景，请参见[Terraform应用场景](#)。

使用Terraform

Terraform能够让您在阿里云上轻松使用简单模板语言来定义、预览和部署云基础结构。更多信息，请参见[简单模板语言](#)。以下为Terraform在ECS中预配资源的必要步骤：

1. 安装并配置Terraform。具体操作，请参见[在Cloud Shell中使用Terraform](#)或[在本地安装和配置Terraform](#)。
2. 使用Terraform创建一台或多台ECS实例。具体操作，请参见[创建一台ECS实例](#)和[创建多台ECS实例](#)。
3. （可选）使用Terraform部署Web集群。具体操作，请参见[部署Web集群](#)。

更多关于Terraform的使用教程，请参见[Terraform文档](#)。

更多信息

- [Terraform Alibaba provider](#)
- [Terraform Alibaba github](#)
- [Terraform Registry Alibaba Modules](#)

7.2. 在Cloud Shell中使用Terraform

阿里云Cloud Shell是一款帮助您运维的免费产品，预装了Terraform的组件，并配置好身份凭证（credentials）。因此您可直接在Cloud Shell中运行Terraform的命令。

完成以下操作，在Cloud Shell中使用Terraform：

1. 打开浏览器，访问Cloud Shell的地址<https://shell.aliyun.com>。

更多Cloud Shell入口及使用请参见[使用云命令行](#)。

2. 登录Cloud Shell。
3. 编写Terraform模板。

您可以使用vim命令直接编写模板，如果开通了OSS存储，您可以直接将配置模板上传到为Cloud Shell创建的bucket中。

示例：

```
provider "alicloud" {}

resource "alicloud_vpc" "vpc" {
  name      = "tf_test_foo"
  cidr_block = "172.16.0.0/12"
}

resource "alicloud_vswitch" "vsw" {
  vpc_id          = alicloud_vpc.vpc.id
  cidr_block      = "172.16.0.0/21"
  availability_zone = "cn-hangzhou-b"
}

resource "alicloud_security_group" "default" {
  name = "default"
  vpc_id = alicloud_vpc.vpc.id
}

resource "alicloud_instance" "instance" {
  # 可用区
  availability_zone = "cn-hangzhou-b"
  # 绑定安全组
  security_groups = alicloud_security_group.default.*.id
  # 实例规格
  instance_type      = "ecs.n2.small"
  # 系统盘类型
  system_disk_category = "cloud_efficiency"
  # 系统镜像
  image_id            = "ubuntu_140405_64_40G_cloudinit_20161115.vhd"
  # 实例名称
  instance_name       = "test_foo"
  # 所在交换机
  vswitch_id          = alicloud_vswitch.vsw.id
  # 公网带宽
  internet_max_bandwidth_out = 10
}

resource "alicloud_security_group_rule" "allow_all_tcp" {
  type          = "ingress"
  ip_protocol   = "tcp"
  nic_type      = "intranet"
  policy        = "accept"
  port_range    = "1/65535"
  priority      = 1
  security_group_id = alicloud_security_group.default.id
  cidr_ip       = "0.0.0.0/0"
}
```

4. 执行`terraform init`命令初始化配置。

5. 执行`terraform plan`命令预览配置。

```
shell@alicloud:~$ terraform plan
Refreshing Terraform state in-memory prior to plan...
The refreshed state will be used to calculate this plan, but will not be
persisted to local or remote state storage.

-----

An execution plan has been generated and is shown below.
Resource actions are indicated with the following symbols:
  + create

Terraform will perform the following actions:

+ alicloud_instance.master[0]
  id: <computed>
  availability_zone: "cn-beijing-f"
  host_name: "sample"
  image_id: "centos_7_04_64_20G_alibase_201701015.vhd"
  instance_charge_type: "PostPaid"
  instance_name: "instance_sample"
  instance_type: "ecs.t5-lc1m1.small"
  internet_charge_type: "PayByTraffic"
  internet_max_bandwidth_in: <computed>
  internet_max_bandwidth_out: "1"
  key_name: <computed>
  password: <sensitive>
  private_ip: <computed>
  public_ip: <computed>
  role_name: <computed>
  security_groups.#: <computed>
  spot_strategy: "NoSpot"
  status: <computed>
  subnet_id: <computed>
  system_disk_category: "cloud_efficiency"
  system_disk_size: "40"
  vswitch_id: "${alicloud_vswitch.vsw.id}"

+ alicloud_instance.master[1]
  id: <computed>
  availability_zone: "cn-beijing-f"
  host_name: "sample"
  image_id: "centos_7_04_64_20G_alibase_201701015.vhd"
  instance_charge_type: "PostPaid"
  instance_name: "instance_sample"
  instance_type: "ecs.t5-lc1m1.small"
  internet_charge_type: "PayByTraffic"
  internet_max_bandwidth_in: <computed>
```

6. 执行terraform apply创建ECS实例。

7.3. 在本地安装和配置Terraform

在使用Terraform的简单模板语言定义、预览和部署云基础结构前，您需要安装预配置Terraform。

操作步骤

1. 前往[Terraform官网](#)下载适用于您的操作系统的程序包。
2. 将程序包解压到 `/usr/local/bin`。

如果将可执行文件解压到其他目录，按照以下方法为其设置环境变量：

- Linux：请参见[在Linux系统设置环境变量](#)。

- Windows: 请参见[在Windows系统设置环境变量](#)。
 - macOS: 请参见[在macOS系统设置环境变量](#)。
3. 运行 `terraform` 验证路径配置。
- 将显示可用的Terraform选项的列表，类似如下所示，表示安装完成。

```
username:~$ terraform
Usage: terraform [-version] [-help] <command> [args]
```

4. 为提高权限管理的灵活性和安全性，建议您创建RAM用户，并为其授权。
- i. 登录[RAM控制台](#)。
 - ii. 创建名为 *Terraform* 的RAM用户，并为用户创建AccessKey。具体步骤请参见[创建RAM用户](#)和[获取AccessKey](#)。
 - iii. 为RAM用户授权。在本示例中，给用户 *Terraform* 授予 `AliyunECSFullAccess` 和 `AliyunVPCFullAccess` 权限，具体步骤请参见[为RAM用户授权](#)。
5. 创建环境变量，用于存放身份认证信息。

```
export ALICLOUD_ACCESS_KEY="LTAIUrZCw3*****"
export ALICLOUD_SECRET_KEY="zfwwWAMWIAiooj14GQ2*****"
export ALICLOUD_REGION="cn-beijing"
```

7.4. 创建一台ECS实例

本文介绍如何使用Terraform创建一台ECS实例。

前提条件

在开始之前，请您确保完成以下操作：

- 使用Terraform，您需要一个阿里云账号和访问密钥（AccessKey）。请在阿里云控制台中的[AccessKey管理页面](#)上创建和查看您的AccessKey。
- 已经安装并配置了Terraform，具体操作请参见[在本地安装和配置Terraform](#)和[在Cloud Shell中使用Terraform](#)。

操作步骤

1. 创建专有网络和交换机。
- i. 创建`terraform.tf`文件，输入以下内容，并保存在当前的执行目录中。

```
resource "alicloud_vpc" "vpc" {
  name      = "tf_test_foo"
  cidr_block = "172.16.0.0/12"
}

resource "alicloud_vswitch" "vsw" {
  vpc_id      = alicloud_vpc.vpc.id
  cidr_block  = "172.16.0.0/21"
  availability_zone = "cn-beijing-b"
}
```

- ii. 运行 `terraform apply` 开始创建。

- iii. 运行 `terraform show` 查看已创建的专有网络和交换机。

您也可以登录VPC控制台查看专有网络和交换机的属性。

2. 在上一步创建的专有网络中创建一个安全组，并添加一个允许任何地址访问的安全组规则。

- i. 在`terraform.tf`文件中增加以下内容。

```
resource "alicloud_security_group" "default" {
  name = "default"
  vpc_id = alicloud_vpc.vpc.id
}
resource "alicloud_security_group_rule" "allow_all_tcp" {
  type           = "ingress"
  ip_protocol    = "tcp"
  nic_type       = "intranet"
  policy         = "accept"
  port_range     = "1/65535"
  priority       = 1
  security_group_id = alicloud_security_group.default.id
  cidr_ip        = "0.0.0.0/0"
}
```

- ii. 运行 `terraform apply` 开始创建。

- iii. 运行 `terraform show` 查看已创建的安全组和安全组规则。

您也可以登录ECS控制台查看安全组和安全组规则。

3. 创建ECS实例。

- i. 在`terraform.tf`文件中增加以下内容。

```
resource "alicloud_instance" "instance" {
  # cn-beijing
  availability_zone = "cn-beijing-b"
  security_groups = alicloud_security_group.default.*.id
  # series III
  instance_type      = "ecs.n2.small"
  system_disk_category = "cloud_efficiency"
  image_id           = "ubuntu_18_04_64_20G_alibase_20190624.vhd"
  instance_name      = "test_foo"
  vswitch_id         = alicloud_vswitch.vsw.id
  internet_max_bandwidth_out = 10
  password           = "<replace_with_your_password>"
}
```

② 说明

- 在上述示例中，指定了 `internet_max_bandwidth_out= 10`，因此会自动为实例分配一个公网IP。
- 详细的参数解释请参见 [阿里云参数说明](#)。

- ii. 运行 `terraform apply` 开始创建。

- iii. 运行 `terraform show` 查看已创建的ECS实例。

iv. 运行`ssh root@<publicip>`，并输入密码来访问ECS实例。

```
provider "alicloud" {}
resource "alicloud_vpc" "vpc" {
  name      = "tf_test_foo"
  cidr_block = "172.16.0.0/12"
}
resource "alicloud_vswitch" "vsw" {
  vpc_id          = alicloud_vpc.vpc.id
  cidr_block      = "172.16.0.0/21"
  availability_zone = "cn-beijing-b"
}
resource "alicloud_security_group" "default" {
  name = "default"
  vpc_id = alicloud_vpc.vpc.id
}
resource "alicloud_instance" "instance" {
  # cn-beijing
  availability_zone = "cn-beijing-b"
  security_groups = alicloud_security_group.default.*.id
  # series III
  instance_type      = "ecs.n2.small"
  system_disk_category = "cloud_efficiency"
  image_id           = "ubuntu_18_04_64_20G_alibase_20190624.vhd"
  instance_name      = "test_foo"
  vswitch_id         = alicloud_vswitch.vsw.id
  internet_max_bandwidth_out = 10
}
resource "alicloud_security_group_rule" "allow_all_tcp" {
  type          = "ingress"
  ip_protocol   = "tcp"
  nic_type      = "intranet"
  policy        = "accept"
  port_range    = "1/65535"
  priority      = 1
  security_group_id = alicloud_security_group.default.id
  cidr_ip       = "0.0.0.0/0"
}
```

7.5. 创建多台ECS实例

本文介绍如何使用Terraform模块批量创建多台ECS实例。

前提条件

在开始之前，请您确保完成以下操作：

- 使用Terraform，您需要一个阿里云账号和访问密钥（AccessKey）。请在阿里云控制台中的[AccessKey管理页面](#)上创建和查看您的AccessKey。
- 已经安装并配置了Terraform，具体操作请参见[在本地安装和配置Terraform](#)和[在Cloud Shell中使用Terraform](#)。

操作步骤

1. 创建专有网络和交换机。

- i. 创建`terraform.tf`文件，输入以下内容，保存在当前的执行目录中。

```
resource "alicloud_vpc" "vpc" {
  name      = "tf_test_foo"
  cidr_block = "172.16.0.0/12"
}

resource "alicloud_vswitch" "vsw" {
  vpc_id      = alicloud_vpc.vpc.id
  cidr_block   = "172.16.0.0/21"
  availability_zone = "cn-beijing-b"
}
```

- ii. 运行 `terraform apply` 开始创建。
- iii. 运行 `terraform show` 查看已创建的专有网络和交换机。

您也可以登录VPC控制台查看专有网络和交换机的属性。

2. 在上一步创建的专有网络中创建一个安全组，并添加一个允许任何地址访问的安全组规则。

- i. 在`terraform.tf`文件中增加以下内容。

```
resource "alicloud_security_group" "default" {
  name = "default"
  vpc_id = alicloud_vpc.vpc.id
}

resource "alicloud_security_group_rule" "allow_all_tcp" {
  type            = "ingress"
  ip_protocol     = "tcp"
  nic_type        = "intranet"
  policy          = "accept"
  port_range      = "1/65535"
  priority        = 1
  security_group_id = alicloud_security_group.default.id
  cidr_ip         = "0.0.0.0/0"
}
```

- ii. 运行 `terraform apply` 开始创建。
- iii. 运行 `terraform show` 查看已创建的安全组和安全组规则。

您也可以登录ECS控制台查看安全组和安全组规则。

3. 使用Module创建多台ECS实例。在本示例中，创建3台ECS实例。

i. 在 `terraform.tf` 文件中增加以下内容。

```
module "tf-instances" {
  source              = "alibaba/ecs-instance/alicloud"
  region              = "cn-beijing"
  number_of_instances = "3"
  vswitch_id          = alicloud_vswitch.vsw.id
  group_ids            = [alicloud_security_group.default.id]
  private_ips          = ["172.16.0.10", "172.16.0.11", "172.16.0.12"]
  image_ids            = ["ubuntu_18_04_64_20G_alibase_20190624.vhd"]
  instance_type        = "ecs.n2.small"
  internet_max_bandwidth_out = 10
  associate_public_ip_address = true
  instance_name        = "my_module_instances_"
  host_name             = "sample"
  internet_charge_type  = "PayByTraffic"
  password              = "User@123"
  system_disk_category = "cloud_ssd"
  data_disks = [
    {
      disk_category = "cloud_ssd"
      disk_name      = "my_module_disk"
      disk_size      = "50"
    }
  ]
}
```

 **说明** 在上述示例中，同时指定 `associate_public_ip_address = true` 和 `internet_max_bandwidth_out = 10`，因此会自动为实例分配一个公网IP。详细的参数解释请参见 [参数说明](#)。

- ii. 运行 `terraform apply` 开始创建。
- iii. 运行 `terraform show` 查看已创建的ECS实例。
- iv. 运行 `ssh root@<publicip>`，并输入密码来访问ECS实例。

```
provider "alicloud" {}

resource "alicloud_vpc" "vpc" {
  name      = "tf_test_foo"
  cidr_block = "172.16.0.0/12"
}

resource "alicloud_vswitch" "vsw" {
  vpc_id      = alicloud_vpc.vpc.id
  cidr_block   = "172.16.0.0/21"
  availability_zone = "cn-beijing-b"
}

resource "alicloud_security_group" "default" {
  name      = "default"
  vpc_id    = alicloud_vpc.vpc.id
}

resource "alicloud_security_group_rule" "allow_all_tcp" {
  type            = "ingress"
  ip_protocol     = "tcp"
  nic_type        = "intranet"
  policy          = "accept"
  port_range      = "1/65535"
  priority        = 1
  security_group_id = alicloud_security_group.default.id
  cidr_ip         = "0.0.0.0/0"
}

module "tf-instances" {
  source              = "alibaba/ecs-instance/alicloud"
  region              = "cn-beijing"
  number_of_instances = "3"
  vswitch_id          = alicloud_vswitch.vsw.id
  group_ids           = [alicloud_security_group.default.id]
  private_ips         = ["172.16.0.10", "172.16.0.11", "172.16.0.12"]
  image_ids           = ["ubuntu_18_04_64_20G_alibase_20190624.vhd"]
  instance_type       = "ecs.n2.small"
  internet_max_bandwidth_out = 10
  associate_public_ip_address = true
  instance_name       = "my_module_instances_"
  host_name           = "sample"
  internet_charge_type = "PayByTraffic"
  password            = "User@123"
  system_disk_category = "cloud_ssd"
  data_disks = [
    {
      disk_category = "cloud_ssd"
      disk_name     = "my_module_disk"
      disk_size     = "50"
    }
  ]
}
```

相关文档

- [terraform-alicloud-ecs-instance](#)

7.6. 部署Web集群

本文介绍如何使用Terraform部署Web集群。

前提条件

在开始之前，请您确保完成以下操作：

- 使用Terraform，您需要一个阿里云账号和访问密钥（AccessKey）。请在阿里云控制台中的[AccessKey管理页面](#)上创建和查看您的AccessKey。
- 已经安装并配置了Terraform，具体操作请参见[在本地安装和配置Terraform](#)和[在Cloud Shell中使用Terraform](#)。

背景信息

在部署一个网站或者应用时，需要部署一系列的节点，SLB对各个节点分配请求，并根据访问数量或者资源使用的情况来自动伸缩。本文介绍将如何使用Terraform部署一个这样的Web集群。在本示例中，整个应用部署在一个可用区，并且只提供8080端口访问hello world网页。

操作步骤

1. 创建VPC网络和交换机。

i. 创建`terraform.tf`文件，输入以下内容，并保存在当前的执行目录中。

```
resource "alicloud_vpc" "vpc" {
  name      = "tf_test_foo"
  cidr_block = "172.16.0.0/12"
}
resource "alicloud_vswitch" "vsw" {
  vpc_id      = alicloud_vpc.vpc.id
  cidr_block  = "172.16.0.0/21"
  availability_zone = "cn-beijing-b"
}
```

ii. 运行 `terraform apply` 开始创建。

iii. 运行 `terraform show` 查看已创建的VPC和vSwitch。

您也可以登录VPC控制台查看VPC和vSwitch的属性。

2. 创建安全组，并将安全组作用于上一步创建的VPC中。

- i. 在`terraform.tf`文件中增加以下内容。

```
resource "alicloud_security_group" "default" {
  name = "default"
  vpc_id = alicloud_vpc.vpc.id
}

resource "alicloud_security_group_rule" "allow_all_tcp" {
  type           = "ingress"
  ip_protocol    = "tcp"
  nic_type       = "intranet"
  policy         = "accept"
  port_range     = "1/65535"
  priority       = 1
  security_group_id = alicloud_security_group.default.id
  cidr_ip        = "0.0.0.0/0"
}
```

- ii. 运行 `terraform apply` 开始创建。
 - iii. 运行 `terraform show` 查看已创建的安全组和安全组规则。
3. 创建负载均衡实例，为其分配公网IP。在本示例中，为负载均衡实例配置了从前端80端口到后端8080端口的映射，并输出公网IP用于后续测试。

- i. 创建`slb.tf`文件，并增加以下内容。

```
resource "alicloud_slb" "slb" {
  name          = "test-slb-tf"
  vswitch_id    = alicloud_vswitch.vsw.id
  internet      = true
}

resource "alicloud_slb_listener" "http" {
  load_balancer_id = alicloud_slb.slb.id
  backend_port     = 8080
  frontend_port    = 80
  bandwidth       = 10
  protocol         = "http"
  sticky_session   = "on"
  sticky_session_type = "insert"
  cookie           = "testslblistenercookie"
  cookie_timeout   = 86400
  health_check     = "on"
  health_check_type = "http"
  health_check_connect_port = 8080
}

output "slb_public_ip" {
  value = alicloud_slb.slb.address
}
```

- ii. 运行 `terraform apply` 开始创建。
 - iii. 运行 `terraform show` 查看已创建的负载均衡实例。
4. 创建弹性伸缩。

在本示例中，将创建以下资源：

- 伸缩组：在模版中指定伸缩最小为2，最大为10，并将伸缩组与新建的负载均衡实例绑定。由于伸缩

组的配置要求SLB必须有相应配置的监听器，因此模版中用depends_on属性指定了部署顺序。

- o 伸缩组配置：在模版中指定ECS实例的具体配置。在初始化配置（user-data）中生成一个Hello World的网页，并在8080端口提供服务。为简化操作，本示例中会为ECS实例分配公网IP，并且设置 `force_delete=true` 用于后续删除环境。
- o 伸缩规则：定义具体的伸缩规则。
- i. 创建 `ess.tf` 文件，并增加以下内容。

```
resource "alicloud_ess_scaling_group" "scaling" {
  min_size = 2
  max_size = 10
  scaling_group_name = "tf-scaling"
  vswitch_ids = alicloud_vswitch.vsw.*.id
  loadbalancer_ids = alicloud_slb.slb.*.id
  removal_policies = ["OldestInstance", "NewestInstance"]
  depends_on = ["alicloud_slb_listener.http"]
}

resource "alicloud_ess_scaling_configuration" "config" {
  scaling_group_id =
vswitch
6 of 7
alicloud_ess_scaling_group.scaling.id
  image_id = "ubuntu_18_04_64_20G_alibase_20190624.vhd"
  instance_type = "ecs.c5.large"
  security_group_id = alicloud_security_group.default.id
  active = true
  enable = true
  user_data = "#!/bin/bash\nnecho \"Hello, World\" > index.html\nnohup busybox httpd\n-f -p 8080&"
  internet_max_bandwidth_in = 10
  internet_max_bandwidth_out = 10
  internet_charge_type = "PayByTraffic"
  force_delete = true
}

resource "alicloud_ess_scaling_rule" "rule" {
  scaling_group_id = alicloud_ess_scaling_group.scaling.id
  adjustment_type = "TotalCapacity"
  adjustment_value = 2
  cooldown = 60
}
```

- ii. 运行 `terraform apply` 开始创建。

创建成功后，会输出SLB的公网IP。

等待大约两分钟，弹性伸缩将自动创建ECS实例。

- iii. 输入命令 `curl http://<slb public ip>` 进行验证。

如果看到 `Hello, World`，表示成功通过负载均衡实例访问ECS实例提供的网页。

5. 运行 `terraform destroy` 删除测试环境。经确认后，整个部署的环境将被删除。

使用Terraform可以便捷地删除和重新部署一个环境。如果您想重新部署，运行 `terraform apply` 即可。