

解决方案实践

# 渲染农场弹性伸缩解决方案

文档版本 01  
发布日期 2023-12-30



版权所有 © 华为技术有限公司 2024。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

## 商标声明



HUAWEI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。

# 安全声明

## 漏洞处理流程

华为公司对产品漏洞管理的规定以“漏洞处理流程”为准，该流程的详细内容请参见如下网址：

<https://www.huawei.com/cn/psirt/vul-response-process>

如企业客户须获取漏洞信息，请参见如下网址：

<https://securitybulletin.huawei.com/enterprise/cn/security-advisory>

---

# 目录

---

<b>1 方案概述</b> .....	<b>1</b>
<b>2 资源和成本规划</b> .....	<b>3</b>
<b>3 实施步骤</b> .....	<b>6</b>
3.1 准备工作.....	6
3.2 快速部署.....	10
3.3 开始使用.....	16
3.4 快速卸载.....	62
<b>4 附录</b> .....	<b>64</b>
<b>5 修订记录</b> .....	<b>69</b>

# 1 方案概述

## 应用场景

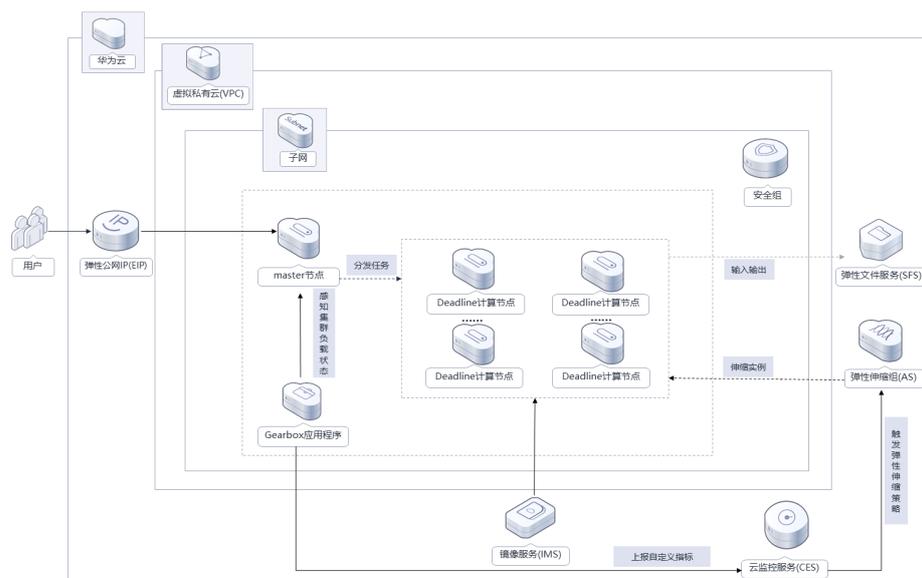
该解决方案基于渲染农场软件Deadline及华为云自研开源软件Gearbox构建弹性伸缩能力，快速帮助用户在华为云上完成可自动弹性伸缩的渲染农场环境搭建。Deadline调度控制节点云服务器采用无配置模式，Gearbox程序与弹性伸缩 AS 及云监控服务 CES 对接，通过监控Deadline集群作业Job状态，实时自动弹性扩缩容Deadline集群环境，并自动完成弹性扩容出来的云服务器注册并加入集群或从集群注销并完成实例销毁。适用于以下场景：

渲染农场集群、数据管理和集群调度等场景。

## 方案架构

该解决方案部署架构如下图所示：

图 1-1 方案架构



该解决方案会部署如下资源：

1. 创建两台**弹性云服务器 ECS**，安装渲染农场软件Deadline，在Master节点安装Gearbox程序、配置Java环境。
2. 创建两条**弹性公网IP EIP**，用于提供访问公网和被公网访问能力。
3. 创建安全组，可以保护弹性云服务器 ECS的网络安全，通过配置安全组规则，限定云服务器的访问端口。
4. 使用**镜像服务 IMS**创建计算节点服务器镜像，用于弹性扩容时使用该镜像配置计算节点服务器初始化环境。
5. 创建一个**弹性伸缩 AS**组，通过设置伸缩策略来弹性扩缩容。
6. 创建**云监控服务 CES**告警规则，该规则根据Gearbox程序监测到的作业状态，触发弹性伸缩 AS去弹性扩缩容。
7. 创建**弹性文件服务 SFS**，挂载到所有弹性云服务器 ECS上，为集群环境提供共享文件存储服务。

## 方案优势

- 动态扩缩容  
该解决方案配置弹性伸缩组，调度节点服务器内置Gearbox程序，该程序可周期性监测集群指标，并汇总指标数据上报云监控服务CES，由CES告警规则触发AS自动扩缩容，节约成本。
- 个性定制化  
该解决方案及内置Gearbox程序均为开源，用户可以免费用于商业用途，并可以在源码基础上进行定制化开发。
- 一键部署  
一键轻松部署，即可完成弹性扩缩容的Deadline集群环境部署。

## 约束与限制

- 部署该解决方案之前，您需注册华为账号并开通华为云，完成实名认证，且账号不能处于欠费或冻结状态。如果计费模式选择“包年包月”，请确保账户余额充足以便一键部署资源的时候可以自动支付；或者在一键部署的过程进入**费用中心**，找到“待支付订单”并手动完成支付。
- 确认租户配额充足：在“资源 > 我的配额”中查看以下配额是否充足，如配额不够，请提前提**工单**申请增加配额：
  - 计算：弹性云服务器实例数、CPU核心数、RAM容量；
  - 存储：云硬盘、弹性文件服务；
  - 网络：虚拟私有云、子网、弹性公网IP、安全组等。

# 2 资源和成本规划

该解决方案主要部署如下资源，支持包年包月和按需计费两种模式。不同产品的花费仅供参考，实际已收费账单为准，具体请参考华为云[官网价格](#)：

表 2-1 资源和成本规划（包年包月）

华为云服务	配置示例	每月预估花费
弹性云服务器 ECS (Master节点)	<ul style="list-style-type: none"><li>● 区域：华北-北京四</li><li>● 计费模式：包年包月</li><li>● 规格：X86计算   ECS   c6s.xlarge.2   4vCPUs   8GiB</li><li>● 镜像：Windows Server 2019 Datacenter 64bit（市场镜像，需付费，可选）</li><li>● 镜像2：CentOS 7.6 64bit（免费，默认）</li><li>● 系统盘：超高IO   100GB</li><li>● 购买量：1</li></ul>	454.00 元+（windows镜像费用48.00元，可选）
弹性云服务器 ECS (Slave节点)	<ul style="list-style-type: none"><li>● 区域：华北-北京四</li><li>● 计费模式：包年包月</li><li>● 规格：X86计算   ECS   c6s.8xlarge.2   32vCPUs   64GiB</li><li>● 镜像1：Windows Server 2019 Datacenter 64bit（市场镜像，需付费，可选）</li><li>● 镜像2：CentOS 7.6 64bit（免费，默认）</li><li>● 系统盘：超高IO   100GB</li><li>● 购买量：1</li></ul>	2932.00 元+（windows镜像费用384.00元，可选）

华为云服务	配置示例	每月预估花费
SFS Turbo性能型	<ul style="list-style-type: none"> <li>区域：华北-北京四</li> <li>计费模式：包年包月（默认创建按需计费，用户需要手动<a href="#">转包周期</a>）</li> <li>SFS Turbo 性能型   500GB</li> </ul>	1150.00 元
弹性公网IP	<ul style="list-style-type: none"> <li>区域：华北-北京四</li> <li>计费模式：按需计费</li> <li>线路：动态BGP</li> <li>弹性公网IP保有费：¥0.02/小时（未绑定云服务器时收取）</li> <li>公网带宽：按流量计费</li> <li>公网流量费用：¥0.80/GB</li> <li>带宽大小：300Mbit/s</li> <li>购买量：2</li> </ul>	根据流量实际使用情况计费
<b>合计</b>		<b>4536.00元</b> （未包含windows镜像费用、未包含实际使用流量费用）

表 2-2 资源和成本规划（按需计费）

华为云服务	配置示例	每月预估花费
弹性云服务器 ECS (Master节点)	<ul style="list-style-type: none"> <li>按需计费：0.88元/小时</li> <li>区域：华北-北京四</li> <li>计费模式：按需计费</li> <li>规格：X86计算   ECS   c6s.xlarge.2   4vCPUs   8GiB</li> <li>镜像1：Windows Server 2019 Datacenter 64bit（市场镜像，需付费，可选）</li> <li>镜像2：CentOS 7.6 64bit（免费，默认）</li> <li>系统盘：超高IO   100GB</li> <li>购买量：1</li> </ul>	0.88 * 24 * 30 = 633.60元 +(windows镜像费用42.84元/月，可选)

华为云服务	配置示例	每月预估花费
弹性云服务器 ECS (Slave节点)	<ul style="list-style-type: none"> <li>● 按需计费: 6.04元/小时</li> <li>● 区域: 华北-北京四</li> <li>● 计费模式: 按需计费</li> <li>● 规格: X86计算   ECS   c6s.8xlarge.2   32vCPUs   64GiB</li> <li>● 镜像1: Windows Server 2019 Datacenter 64bit (市场镜像, 需付费, 可选)</li> <li>● 镜像2: CentOS 7.6 64bit (免费, 默认)</li> <li>● 系统盘: 超高IO   100GB</li> <li>● 购买量: 1</li> </ul>	6.04 * 24 * 30 = 4348.80元+(windows镜像费用391.68元/月, 可选)
SFS Turbo性能型	<ul style="list-style-type: none"> <li>● 按需计费: 1.29元/小时</li> <li>● 区域: 华北-北京四</li> <li>● 计费模式: 按需计费</li> <li>● SFS Turbo 性能型   500GB</li> </ul>	1.29 * 24 * 30 = 928.80元
弹性公网IP	<ul style="list-style-type: none"> <li>● 区域: 华北-北京四</li> <li>● 计费模式: 按需计费</li> <li>● 线路: 动态BGP</li> <li>● 弹性公网IP保有费: ¥0.02/小时 (未绑定云服务器时收取)</li> <li>● 公网带宽: 按流量计费</li> <li>● 公网流量费用: ¥0.80/GB</li> <li>● 带宽大小: 300Mbit/s</li> <li>● 购买量: 2</li> </ul>	根据流量实际使用情况计费
<b>合计</b>		<b>5911.2元</b> (未包含windows镜像费用、未包含实际使用流量费用)

# 3 实施步骤

- 3.1 准备工作
- 3.2 快速部署
- 3.3 开始使用
- 3.4 快速卸载

## 3.1 准备工作

### 创建 rf\_admin\_trust 委托（可选）

**步骤1** 进入华为云官网，打开[控制台管理](#)界面，鼠标移动至个人账号处，打开“统一身份认证”菜单。

图 3-1 控制台管理界面



图 3-2 统一身份认证菜单



步骤2 进入“委托”菜单，搜索“rf\_admin\_trust”委托。

图 3-3 委托列表



- 如果委托存在，则不用执行接下来的创建委托的步骤
- 如果委托不存在时执行接下来的步骤创建委托

步骤3 单击步骤2界面中的“创建委托”按钮，在委托名称中输入“rf\_admin\_trust”，选择“普通账号”，委托的账号，输入“op\_svc\_IAC”，单击“下一步”。

图 3-4 创建委托



步骤4 在搜索框中输入“Tenant Administrator”权限，并勾选搜索结果。

图 3-5 选择策略



步骤5 选择“所有资源”，并单击下一步完成配置。

图 3-6 设置授权范围



步骤6 “委托”列表中出现“rf\_admin\_trust”委托则创建成功。

图 3-7 委托列表

委托名称	委托环境	委托角色	创建时间	状态	操作
ft_admin_trust	ops_huaweicloud	永久	2022-05-06 17:02:56 GMT+08:00	--	授权 修改 删除

----结束

## 创建访问密钥&获取项目 ID

部署该方案之前，需要您在华为云控制台配置访问密钥对，并创建弹性云服务器SSH密钥对。

- 步骤1** 根据[官方文档](#)指引，在控制台--我的凭证--访问密钥中配置访问密钥AK并下载秘密访问密钥SK，密钥用于启动Gearbox程序的配置项。（该解决方案部署成功后，需要删除该临时AKSK，并重新配置）

图 3-8 创建 AK, SK



- 步骤2** 在控制台--我的凭证--API凭证中，选择本部署方案所在区域的项目ID，项目ID用于配置启动Gearbox程序的配置项

- 步骤3** 参考[通过管理控制台创建密钥对](#)，用于设置弹性伸缩实例创建的登录密钥。

图 3-9 创建密钥对



----结束

## 3.2 快速部署

本章节主要帮助用户快速部署该解决方案。

表 3-1 参数填写说明

参数名称	类型	是否必填	参数解释	默认值
ak	String	必填	Access Key ID，访问密钥，参考 <a href="#">3.1节步骤1获取</a> 。	空
sk	String	必填	Secret Access Key，与ak配合使用。参考 <a href="#">3.1节步骤1获取</a> 。	空
project_id	String	必填	终端节点的租户ID。参考 <a href="#">3.1节步骤2获取</a> 。	空
vpc_name	String	必填	虚拟私有云名称，该模板使用新建VPC，不允许重名。取值范围：1-54个字符，支持中文、英文字母、数字、-（中划线）、.（点）。	elastic-scaling-with-deadline-demo
security_group_name	String	必填	安全组名称，该模板使用新建安全组，安全组规则请参考部署指南进行配置。取值范围：1-64个字符，支持数字、字母、中文、-（中划线）、.（点）。	elastic-scaling-with-deadline-demo
bandwidth_size	Number	必填	弹性公网带宽大小，该模板计费方式为按流量计费（根据具体业务选择按带宽或按流量）。单位：Mbit/s，取值范围：1-300Mbit/s。	300
as_name	String	必填	弹性伸缩资源名称前缀，命名方式为{as_name}-group，不支持重名。取值范围：1-50个字符组成，只能由中文、英文字母、数字、中划线（-）组成。	elastic-scaling-with-deadline-demo
keypair_name	String	必填	已有SSH登录密钥对名称，请参考部署指南获取，参考 <a href="#">3.1节步骤3获取</a>	空

参数名称	类型	是否必填	参数解释	默认值
sfs_turbo_name	String	必填	弹性服务文件名称，取值范围:4-64字符，必须以字母开头，可以包含字母、数字、中划线和下划线，不能包含其他的特殊字符，不区分大小写。	elastic-scaling-with-deadline-demo
sfs_turbo_size	Number	选填	弹性服务文件大小，单位：GB，为集群环境提供共享文件存储服务。取值范围：500-3,2768。	500
ecs_name	String	必填	弹性云服务器名称，命名方式为{ecs_name}-master、{ecs_name}-slave，不支持重名。取值范围：1-58个字符组成，包括小写字母、数字、连字符（-）。	elastic-scaling-with-deadline-demo
ecs_password	String	必填	弹性云服务器初始化密码，创建完成后，请参考部署指南登录ECS控制台修改密码。取值范围：长度为8-26个字符，密码至少包含大写字母、小写字母、数字和特殊字符（\$!@%-_+=[]:./^,{}?）中的三种，密码不能包含用户名或用户名的逆序。管理员账户默认为root。	空
ecs_master_flavor	String	必填	弹性云服务器调度节点规格，具体请参考官网 <a href="#">弹性云服务器规格清单</a> 。	c6s.xlarge.2
ecs_slave_flavor	String	必填	弹性云服务器计算节点规格，具体请参考官网 <a href="#">弹性云服务器规格清单</a> 。	c6s.8xlarge.2
ecs_os_type	String	必填	弹性云服务器操作系统类型，有效值为：windows-server-2019-dataCenter、centos7.6	centos7.6
charging_mode	String	必填	弹性云服务器计费方式，默认自动扣费，有效值为：prePaid（包年包月）或postPaid（按需计费）。	postPaid
charging_unit	String	选填	计费周期单位，当计费方式设置为postPaid时，此参数失效。有效值为：month（包月）和year（包年）。	month

参数名称	类型	是否必填	参数解释	默认值
charging_perio d	Number	选填	计费周期，当计费方式设置为postPaid时，此参数失效。当计费周期单位设置为month，取值范围为1~9；当计费周期单位设置为year，取值范围为1~3。	1

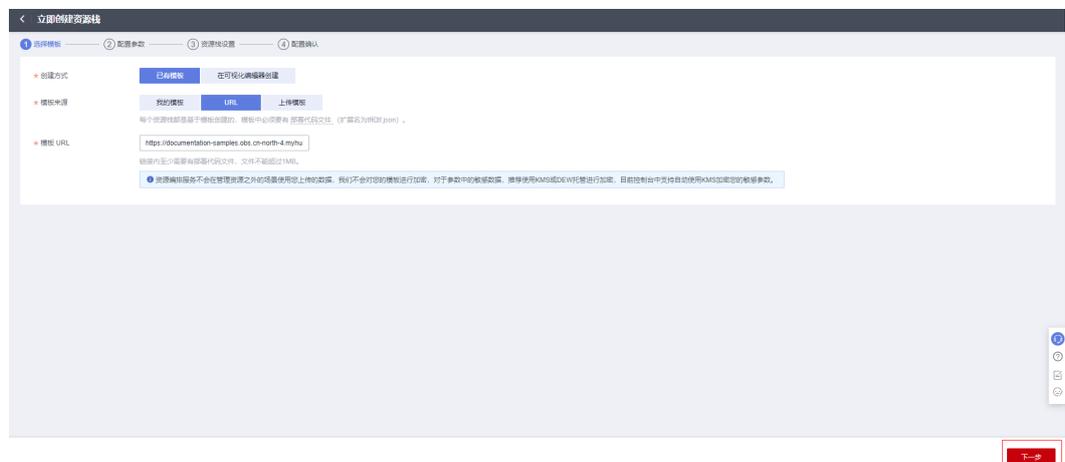
**步骤1** 登录[华为云解决方案实践](#)实施库，选择“渲染农场弹性伸缩解决方案”跳转至该解决方案详情页面。

图 3-10 解决方案实施库



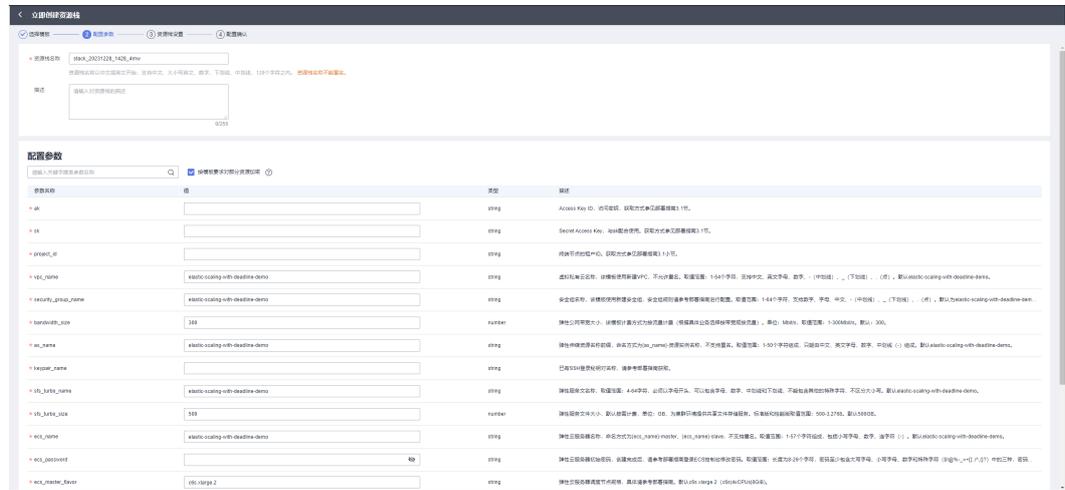
**步骤2** 在一键部署页面单击“一键部署”，跳转至该解决方案创建资源栈部署界面。

图 3-11 一键部署



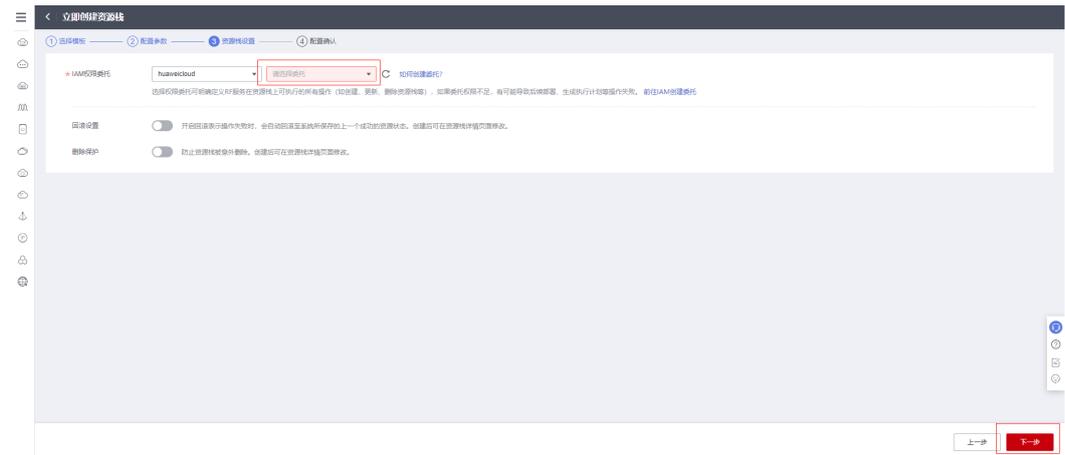
步骤3 单击“下一步”，参考表1参数填写说明完成自定义参数填写。

图 3-12 参数配置



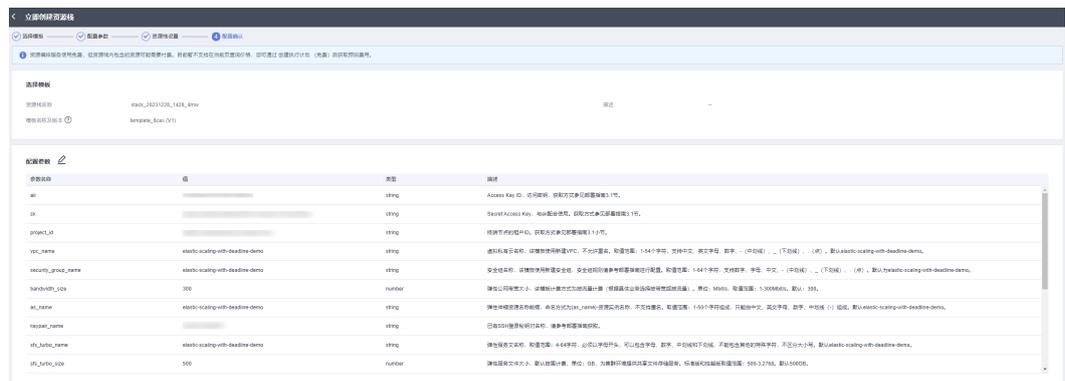
步骤4 在资源栈设置界面，单击“下一步”。

图 3-13 选择创建好的 rf\_admin\_trust 委托。可选



步骤5 单击“下一步”，创建执行计划。

图 3-14 创建执行计划



**步骤6** 单击“创建执行计划”，根据提示输入执行计划名称等，单击“确定”。

图 3-15 创建执行计划



图 3-16 资源栈创建结果



**步骤7** 待执行计划状态创建成功，单击资源栈名称进入，单击“部署”，弹出页面提示等信息，单击“执行”。

图 3-17 部署资源栈

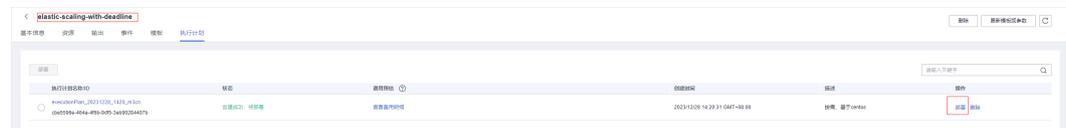


图 3-18 执行计划确认



**步骤8** (可选) 如果计费模式选择“包年包月”, 在余额不充足的情况下(所需总费用请参考表2-1)请及时登录[费用中心](#), 手动完成待支付订单的费用支付。

**步骤9** 部署成功后, 回显结果如下:

图 3-19 执行计划部署完成



## 📖 说明

- 云服务器名称后缀为master是Master节点，云服务器名称后缀为slave则是Slave节点。
- 方案部署成功预计需要20分钟，其中云服务器完成环境安装配置需要10-15min，请在此期间内不要登录服务器。
- 该解决方案使用3389端口远程登录弹性云服务器ECS，默认对该方案创建的VPC子网网段放开，请参考[修改安全组规则](#)，配置IP地址白名单，以便能正常访问服务。  
安全组实际是网络流量访问策略，包括网络流量入方向规则和出方向规则，通过这些规则为安全组内具有相同保护需求并且相互信任的云服务器、云容器、云数据库等实例提供安全保护。  
如果您的实例关联的安全组策略无法满足使用需求，比如需要添加、修改、删除某个TCP端口，请参考以下内容进行修改。
  - 添加安全组规则：根据业务使用需求需要开放某个TCP端口，请参考[添加安全组规则](#)添加加入方向规则，打开指定的TCP端口。
  - 修改安全组规则：安全组规则设置不当会造成严重的安全隐患。您可以参考[修改安全组规则](#)，来修改安全组中不合理的规则，保证云服务器等实例的网络安全。  
删除安全组规则：当安全组规则入方向、出方向源地址/目的地址有变化时，或者不需要开放某个端口时，您可以参考[删除安全组规则](#)进行安全组规则删除。

----结束

## 3.3 开始使用

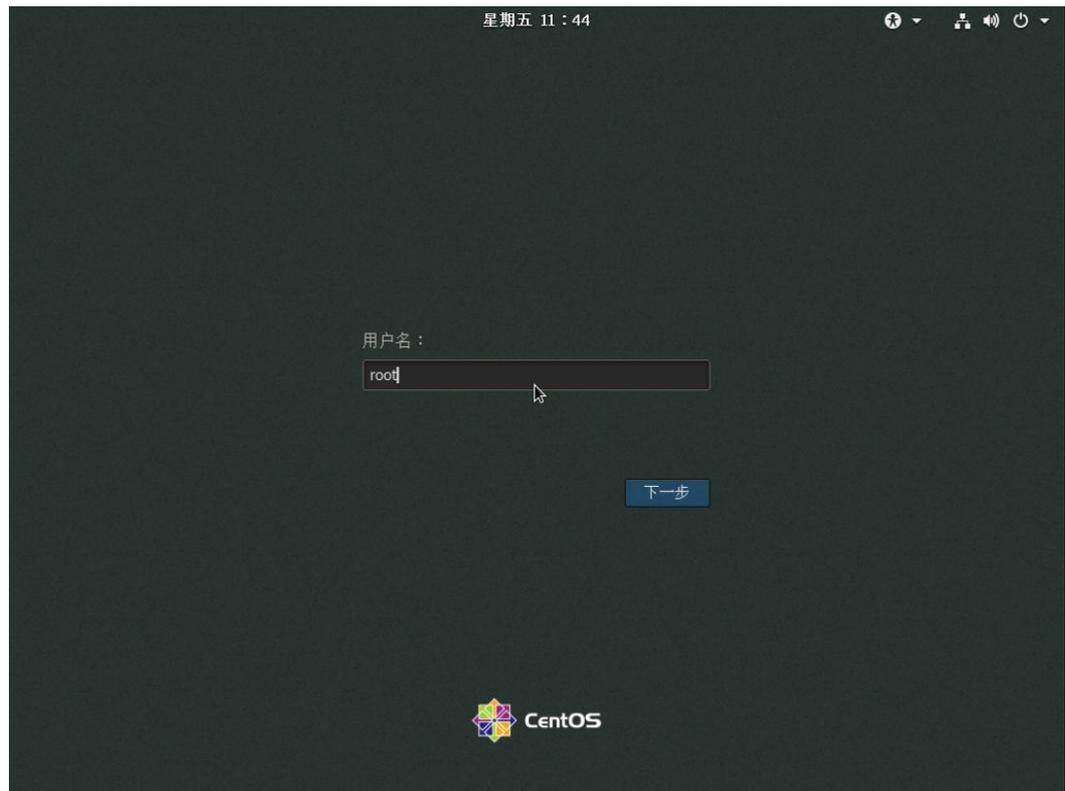
请根据云服务器操作系统类型选择部署指导方案。

### Centos 环境安装与部署

**步骤1** 检查Master节点基础环境安装配置情况。

登录Master节点。centos系统登录时请采用"root"身份登录。

图 3-20 root 身份登录



使用如下命令查看环境初始化日志文件，检查相关软件及环境是否安装配置完成。  
`cat /root/envInstallConfig.log`

图 3-21 环境安装配置日志

```
Thu Dec 14 14:31:21 CST 2023 start env install and config
Thu Dec 14 14:31:21 CST 2023 install java
Thu Dec 14 14:31:43 CST 2023 install and config nfs
Thu Dec 14 14:31:47 CST 2023 stop and disable firewalld
Thu Dec 14 14:31:48 CST 2023 install centos GUI
Thu Dec 14 14:35:28 CST 2023 dowbload and unzip Deadline
Thu Dec 14 14:36:01 CST 2023 dowbload mongodb
Thu Dec 14 14:36:01 CST 2023 install DeadlineRepository
Thu Dec 14 14:36:54 CST 2023 share DeadlineRepository /opt/Thinkbox/DeadlineRepository10
Thu Dec 14 14:36:54 CST 2023 create directory /DeadlineRepository
Thu Dec 14 14:36:54 CST 2023 mount master DeadlineRepository to node /DeadlineRepository10
Thu Dec 14 14:36:54 CST 2023 install DeadlineClient
Thu Dec 14 14:37:42 CST 2023 config DeadlineClient systemENVPPath
Thu Dec 14 14:37:42 CST 2023 mount SFS-Turbo exportLocation
Thu Dec 14 14:37:42 CST 2023 download gearbox
Thu Dec 14 14:37:42 CST 2023 config gearbox application.yaml
finish env install and config
```

### 说明

- Master节点会自动安装Deadline软件、共享DeadlineRepository目录、为DeadlineClient添加系统环境变量、安装Java环境、安装配置Gearbox软件、自动挂载SFS-Turbo、安装Centos桌面等。详见envInstallConfig.log日志。
- Deadline相关软件存放在/root/thinkbox目录下。
- DeadlineRepository目录/opt/Thinkbox/DeadlineRepository 目录将会被通过NFS协议共享出来，便于与Slave节点通信
- DeadlineClient软件配置的Repository目录为/DeadlineRepository10。
- Gearbox软件存放在/root/gearbox目录下。
- SFS-Turbo挂载点为/root/assets， /root/assets目录用于存储渲染资产。

### 步骤2 检查Deadline集群状态。

登录Master节点查看集群状态。

图 3-22 启动 Deadline Monitor 10

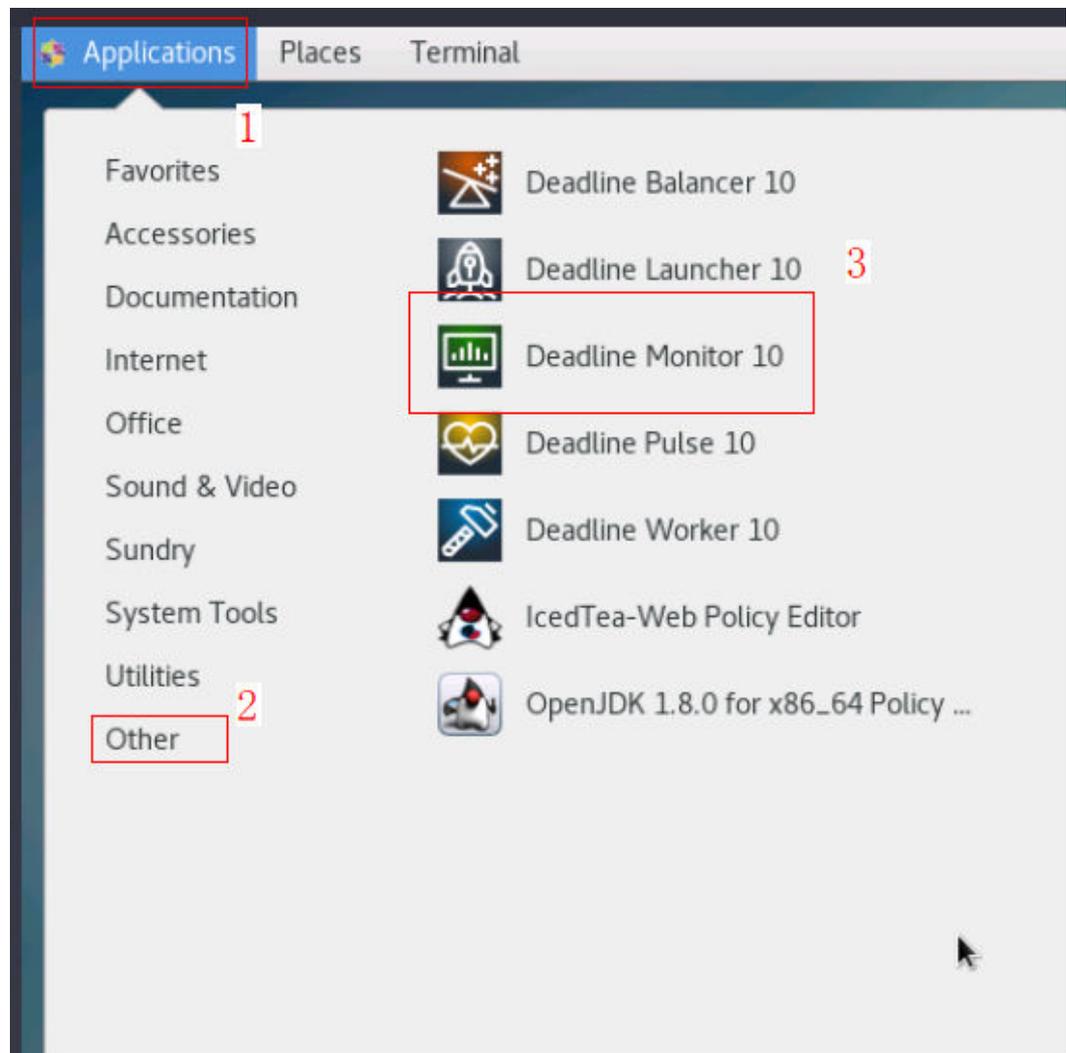


图 3-23 确定登录账户

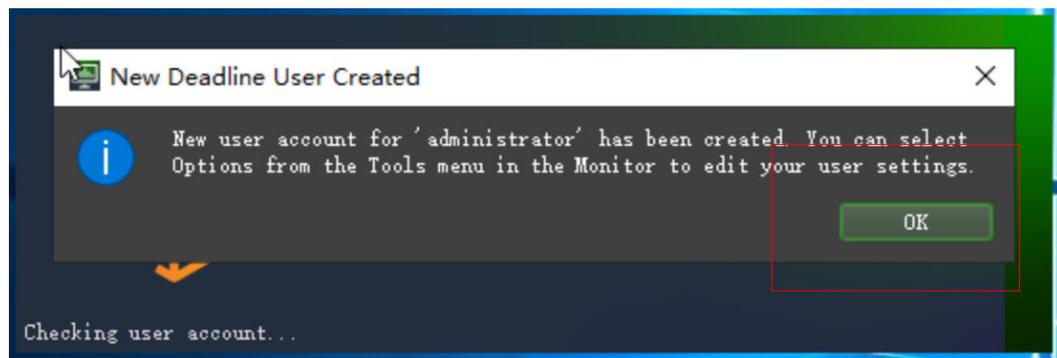
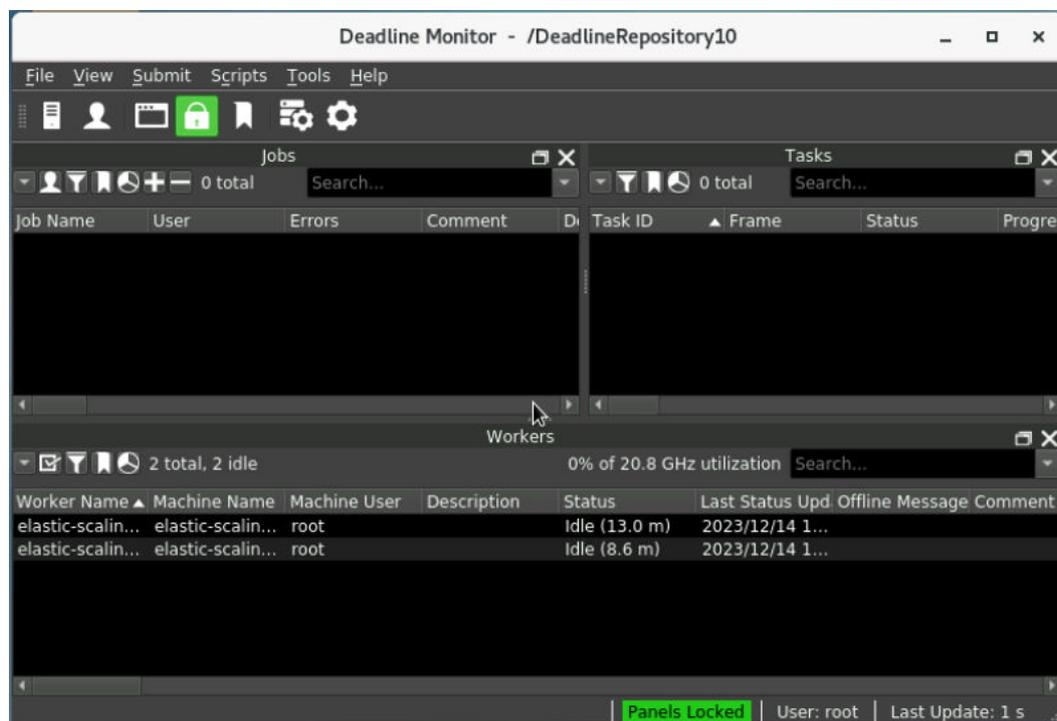


图 3-24 查看节点状态



**步骤3** 配置Gearbox软件。最新版可从[Gearbox的Gitee仓库地址](#)下载，新版[Gearbox配置文件地址](#)（可选）。

登录Master节点对Gearbox软件进行配置。请保持Deadline Monitor开启。

1、切换到 /root/gearbox 目录下。

```
cd /root/gearbox
```

2、使用gedit 或者 vim 命令打开Gearbox的配置文件 application.yml，对可选项根据需要自行配置。说明：ak、sk、project-id、group等必要启动项已完成自动配置。

配置项说明如下：

```
user:  
# AK硬编码或者明文存储都存在较大的安全风险,建议在配置文件或者环境变量中密文存储,使用时解密,确保安全  
ak:  
# SK硬编码或者明文存储都存在较大的安全风险,建议在配置文件或者环境变量中密文存储,使用时解密,确保安全  
sk:  
# 待使用局点的租户ID  
project-id:
```

```
# 代理地址, 端口, 用户名密码等, 无需代理可不配置
proxy-address:
proxy-port:
proxy-user-name:
proxy-password:
as:
# AS服务在待使用局点的终端节点域名
endpoint: as.cn-north-4.myhuaweicloud.com
# 预置伸缩组资源的伸缩组ID
group: f12d57a3-ed8a-4a16-8c3f-4817ab7171f2
# 查询伸缩实例时单页最大返回数量, 默认100可不修改
list-instance-limit: 100
# 删除伸缩实例时最大数量限制, AS服务最大支持50, 可不修改
delete-instance-limit: 50
ecs:
# ECS服务在待使用局点的终端节点域名
endpoint: ecs.cn-north-4.myhuaweicloud.com
metric:
# 自定义监控指标命名空间, 可不修改
namespace: test.HPC
# 自定义指标的指标名
name: workload
# 自定义指标维度名, 可不修改
dimension-name: autoscaling_group
# 自定义指标维度ID, 可配置成伸缩组的ID, 该值不影响功能
dimension-id: f12d57a3-ed8a-4a16-8c3f-4817ab7171f2
# 指标上报的TTL参数, 可不修改
report-ttl: 172800
# CES服务在待使用局点的终端节点域名
metric-report-endpoint: ces.cn-north-4.myhuaweicloud.com
task:
# 节点状态检查周期, 单位 秒
health-audit-period: 30
# 自定义指标上报周期, 单位 秒
metric-report-period: 60
# 检测是否需要缩容的周期, 单位 秒
scale-in-period: 5
# 自动删除待缩容节点周期, 单位 秒
delete-instance-period: 5
# 新扩容节点自发现周期
discover-instance-period: 20
# 对比伸缩组与slurm节点数量周期, 单位 秒
diff-instance-and-node-period: 60
# 内部缓存刷新周期, 单位 秒
refresh-cache-period: 100
# 配置检查任务运行周期, 单位 秒
configuration-check-period: 300
system:
# deadline/slurm
type: deadline
# 稳定节点的Node名称, 多节点英文逗号分割
stable-nodes: deadline-master
# 稳定节点所在的分区
stable-partition: dyn1
# 不稳定节点所在的分区
variable-partition: dyn1
# 空闲时间阈值, 空闲时间超过该值的节点将被缩容, 单位 秒
scale-in-time: 1
# 判断job是否属于排队状态的时间限制条件, 排队时间超过该值的job被认为是在排队并参与指标计算, 建议为0
job-wait-time: 1
# 新节点注册的超时时间, 超过该时间依然未注册成功将被AS删除, 单位: 分钟, 建议为10
register-timeout-minutes: 10
# 弹性节点使用的cpu核数
cpu: 4
# 弹性节点使用的内存大小, 该字段为预留字段, 可设置为任意大于0的数值
memory: 12600
gpu: 1
# 仅slurm集群配置有效, 默认是true, 在计算Workload时忽略含有GPU需求作业中的CPU需求量
ignore-cpu-request-of-gpu-job: true
```

```
# 是否使用powershell执行指令，只有windows环境下生效
use-powershell-executor: true
# 支持多分区资源协同
partition:
# 特性整体开发，默认值为false，对接系统为slurm且该配置开启后后续配置生效；
multi-partition-scaling: false
# 提交到多个分区的作业负载，在多个分区上的映射策略 MAPPING_TO_FIRST_PARTITION，仅累积到第一个分区中，其余分区不计算(默认)；MAPPING_BY_WEIGHT，按照各分区的权重计算负载的映射；
load-mapping-policy: MAPPING_TO_FIRST_PARTITION
# 分区配置列表
partition-configuration-list:
-
# 分区名
name: partition1
# 分区中节点资源类型，支持CPU、GPU二选一；
type: CPU
# 分区在负载映射中的计算权重，整数，取值范围[0,100]；
weight: 10
# 分区对应的租户id
project-id:
# 分区对应的伸缩组id，伸缩组与分区需要一一对应
scaling-group-id: XXX
#节点的资源规格，当资源类型为CPU时可填节点CPU和内存规格，当节点资源类型为GPU时可填节点规格为GPU数量；
node-spec:
# 节点CPU数量
cpu: 4
# 节点内存容量，单位GB；
mem: 8
# 加速卡数量
gpu: 2
# 空闲一定时间后触发缩容，单位秒，未配置时使用全局配置，配置后针对本分区覆盖全局空闲缩容时间；
scale-in-after-idle-time: 100
# 自定义指标维度名
dimension-name: autoscaling_group
# 自定义指标维度ID
dimension-id: f12d57a3-ed8a-4a16-8c3f-4817ab7171f2
# 云端API调用的地址端点列表，未配置时使用全局配置
ecs-endpoint: ecs.cn-north-4.myhuaweicloud.com
as-endpoint: as.cn-north-4.myhuaweicloud.com
ces-endpoint: ces.cn-north-4.myhuaweicloud.com
```

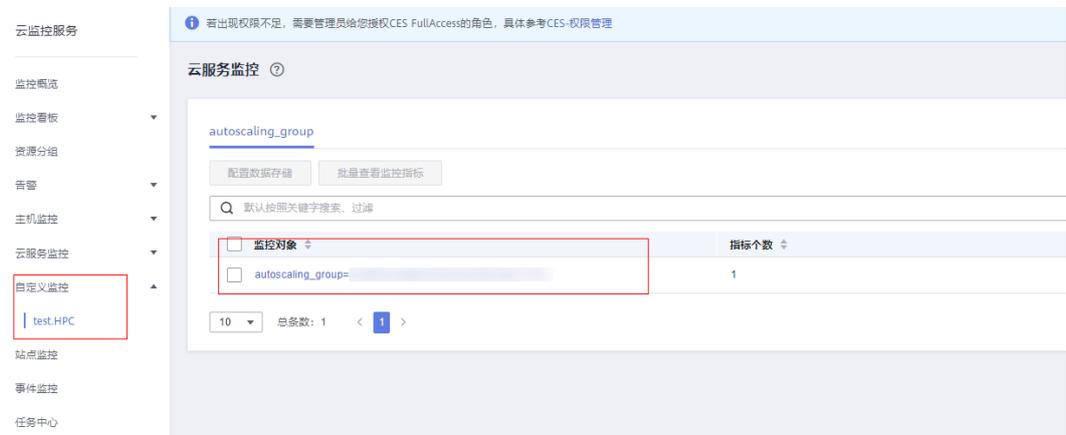
### 3、测试gearbox是否配置正确。使用如下命令：

```
cd /root/gearbox #进入Gearbox所在目录。
java -jar gearbox-0.0.1-SNAPSHOT.jar #请根据实际jar包名来启动Gearbox软件。
```

**说明：** Gearbox 启动后会在终端打印日志，请检查有无报错信息。

### 4、登录[华为云监控服务CES控制台](#)查看自动定义指标workload是否上报成功。

图 3-25 查看自定义指标



说明：无提交任务，故workload指标为0。

图 3-26 查看自定义指标



**步骤4** 自行安装业务软件。在Master节点和Slave节点安装所需业务软件，并配置业务软件开机自启。

**步骤5** 测试Slave节点。完成业务软件安装后，在Master节点提交测试任务，检查Slave节点能否正常渲染。**说明：请将渲染资产请存放在/root/assets 目录下，便于所有渲染节点共享。**

**步骤6** 将Slave节点打包为私有镜像。

登录Slave节点，完成制作私有镜像的准备工作。

1. 制作镜像前清除cloud-init缓存。  
#使用如下命令删除cloud-init缓存，保证基于此镜像创建的新虚拟机可以正确注入机器名、登录密码等初始化数据。**注意：不要在这个过程中重启机器，否则需要重新执行下面命令。**  
`rm -rf /var/lib/cloud/*`  
`rm -rf /var/log/cloud-init*`
2. 自行删除Slave节点中的敏感信息。
3. 镜像预处理操作完成后，登录[华为云服务器控制台](#)，将Slave节点关机，打包Slave节点镜像。

找到本方案创建的云服务器，单击右侧“更多>镜像>创建镜像”进行打包。

图 3-27 打包私有镜像



图 3-28 打包私有镜像

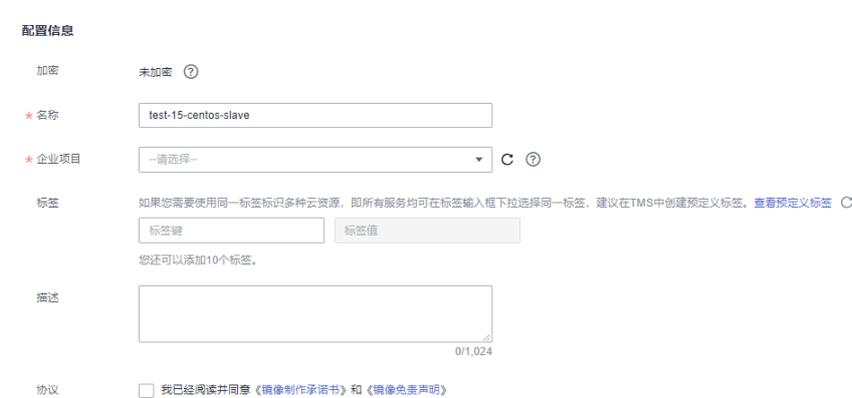


图 3-29 创建私有镜像



在创建私有镜像界面，填写名称，选择企业项目，勾选协议，单击立即创建，等待镜像创建完成。

图 3-30 创建私有镜像



打开[华为云镜像服务控制台](#)，等待大约5-10min后即可查看镜像创建结果。

图 3-31 计算节点镜像



### 步骤7 修改伸缩配置

登录[华为云弹性伸缩AS控制台](#)，选择本方案部署时创建的伸缩组。

图 3-32 修改伸缩配置



说明：修改伸缩配置时，AS会自动为“名称”添加“\_copy”的后缀，请将此后缀删除，因为伸缩配置名称带有下划线“\_”时，会导致节点扩容异常。

图 3-33 修改伸缩配置的名称

修改伸缩配置



建议选择C6、C6s、C7，规格可选32U64G、64U128G。

图 3-34 配置镜像规格 可选。



此处镜像选择上一步骤中创建的镜像，磁盘根据业务需要自行选择，安全组选择方案部署时创建的安全组。

图 3-35 配置镜像、磁盘、安全组等内容

修改伸缩配置



图 3-36 登录方式选择"密钥对", 密钥对选择部署方案时创建的密钥对。在 高级配置--->现在配置--->实例自定义数据注入--->以文本形式 注入如下内容。

```
#!/bin/sh  
reboot
```

说明：因Deadline以节点hostname来标识渲染节点，而新节点启动时会保留原始镜像节点的hostname信息，故此处重启以使Deadline能够正常标识渲染节点。

图 3-37 配置登录方式、密钥对



步骤8 按需配置伸缩组。可选。

登录[华为云弹性伸缩AS控制台](#)，右侧选择“更多”--->"修改伸缩组"

图 3-38 修改弹性伸缩组



说明：最大实例数、期望实例数、最小实例数根据业务需要自行配置，冷却时间建议小于30s，其他配置可保持默认。

图 3-39 修改伸缩组配置

**修改伸缩组**

伸缩组为非启用状态、实例数为0且没有正在进行的伸缩活动时，可以修改伸缩组的子网。

- \* 名称: elastic-scaling-with-deadline-demo-group
- \* 最大实例数(台): 50
- \* 期望实例数(台): 0
- \* 最小实例数(台): 0
- \* 冷却时间(秒): 10
- \* 可用区: 可用区1, 可用区2
- \* 多可用区扩展策略:  均衡分布  选择优先
- 负载均衡:  不使用  使用弹性负载均衡
- \* 健康检查方式: 云服务器健康检查

说明：下面配置根据业务需要自行配置。

图 3-40 修改伸缩组配置

**修改伸缩组**

- \* 多可用区扩展策略:  均衡分布  选择优先
- 负载均衡:  不使用  使用弹性负载均衡
- \* 健康检查方式: 云服务器健康检查
- \* 健康检查间隔: 10秒
- \* 健康状况检查宽限期(秒): 60
- \* 实例移除策略: 根据较早创建的配置较早创建的实例
- \* 企业项目: default
- 弹性公网IP:  释放  不释放
- 数据盘:  删除  不删除
- 委托: -请选择- [新建委托](#)

新的委托仅对修改后创建的云服务器生效，已有云服务器的委托请前往弹性云服务器控制台修改。

**确定** **取消**

### 步骤9 启动Gearbox软件。

Master节点上，在保持Deadline Monitor开启的情况下，启动Gearbox软件。使用如下命令进入Gearbox所在目录并启动程序。

```
cd /root/gearbox  
java -jar gearbox-0.0.1-SNAPSHOT.jar
```

图 3-41 gearbox 启动

```
[root@elastic-scaling-with-deadline-demo-master gearbox]# java -jar gearbox-0.0.1-SNAPSHOT.jar  
  
:: Spring Boot ::  
 (v2.7.3)  
  
2023/12/14 15:11:04.055 [background-preinit] INFO o.h.validator.internal.util.Version.<clinit>(Version.java:21)  
- HV000001: Hibernate Validator 6.2.4.Final  
2023/12/14 15:11:04.065 [main] INFO com.gearbox.AsGearboxApplication.logStarting(StartupInfoLogger.java:55) - S  
tarting AsGearboxApplication v0.0.1-SNAPSHOT using Java 1.8.0_392 on elastic-scaling-with-deadline-demo-master w  
ith PID 28041 (/root/gearbox/gearbox-0.0.1-SNAPSHOT.jar started by root in /root/gearbox)  
2023/12/14 15:11:04.066 [main] INFO com.gearbox.AsGearboxApplication.logStartupProfileInfo(SpringApplication.ja  
va:634) - No active profile set, falling back to 1 default profile: "default"  
2023/12/14 15:11:05.284 [main] INFO o.s.b.w.e.tomcat.TomcatWebServer.initialize(TomcatWebServer.java:108) - Tom  
cat initialized with port(s): 8080 (http)  
2023/12/14 15:11:05.297 [main] INFO o.a.coyote.http11.Http11NioProtocol.log(DirectJDKLog.java:173) - Initializi  
ng ProtocolHandler ["http-nio-8080"]  
2023/12/14 15:11:05.297 [main] INFO o.a.catalina.core.StandardService.log(DirectJDKLog.java:173) - Starting ser  
vice [Tomcat]  
2023/12/14 15:11:05.298 [main] INFO o.a.catalina.core.StandardEngine.log(DirectJDKLog.java:173) - Starting Serv  
let engine: [Apache Tomcat/9.0.65]  
2023/12/14 15:11:05.484 [main] INFO o.a.c.c.c.[Tomcat].[/].log(DirectJDKLog.java:173) - Initializin  
g Spring embedded WebApplicationContext  
2023/12/14 15:11:05.484 [main] INFO o.s.b.w.s.c.ServletWebServerApplicationContext.prepareWebApplicationContext
```

### 步骤10 配置伸缩组的伸缩策略。

登录[华为云弹性伸缩AS控制台](#)，单击弹性伸缩组名称，进入伸缩组详情界面。

图 3-42 添加策略



策略类型选择"告警策略"，告警规则选择"现在创建"，监控类型选择"自定义监控"，触发条件选择"workload"、"原始值"。

图 3-43 添加伸缩策略

添加伸缩策略

策略名称: as-policy-6328

策略类型: 告警策略 (选中) | 定时策略 | 周期策略

告警规则: 现在创建 (选中) | 使用已有

告警规则名称: as-alarm-6335 (自动匹配)

监控类型: 系统监控 | 自定义监控 (选中)

资源类型: test.HPC (自动匹配)

维度: autoscaling\_group

监控对象: f12d57a3-ed8a-4a16-8c3f...

触发条件: workload (原始值) > 1 (数字可自定义)

图 3-44 添加伸缩策略

连续出现次数: 1 (自定义)

告警策略类型: 简单策略 (选中) | 区间策略

执行动作: 增加 (自定义) | 1 | 个实例

冷却时间(秒): 5 (建议小于30s)

确定 | 取消

**步骤11** 提交测试任务。登录Master节点提交测试任务，保持Gearbox处于运行状态。**说明：**由于Master节点配置较低，建议渲染时关闭Deadline Worker。

**步骤12** CES控制台查看指标上报情况。

登录[华为云监控服务CES控制台](#)，查看部署方案创建的自定义监控。

图 3-45 自定义监控



监控指标workload值表示集群需要扩容的节点数量，如workload为24时，表示需要扩容24个计算节点。

图 3-46 查看监控指标 workload



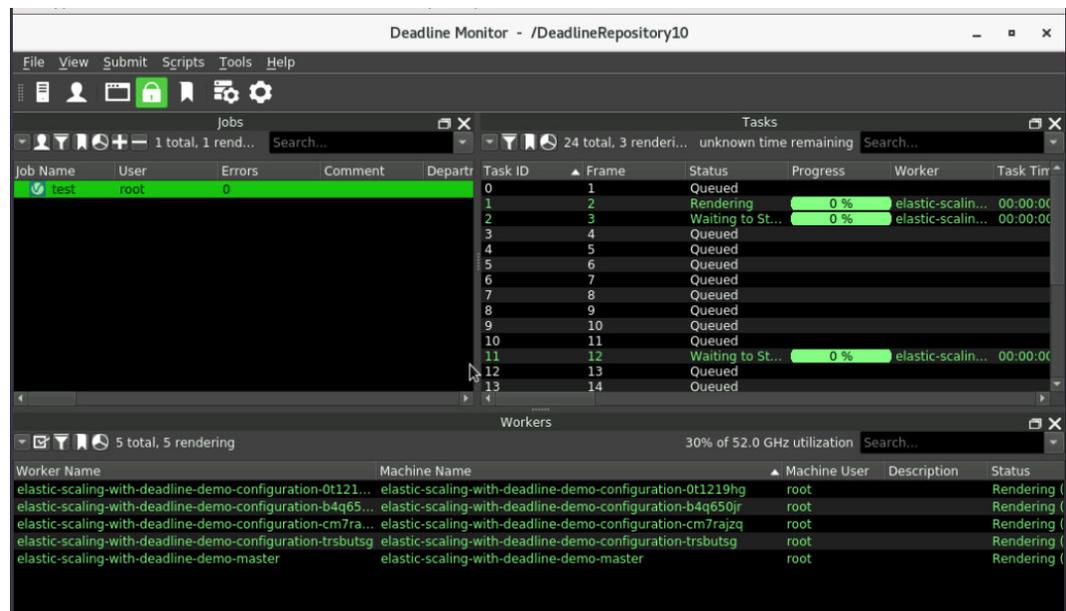
**步骤13** 查看节点伸缩情况。登录[华为云弹性伸缩AS控制台](#)，单击本方案部署时创建的伸缩组名，进入实例监控台。

图 3-47 某个时间段监控面板状态。



**步骤14** Deadline Monitor查看节点扩容情况。登录Maste节点，启动Deadline Monitor查看任务运行情况，渲染节点状态。

图 3-48 某个时间段 Monitor 监控面板状态



**步骤15** 结束任务

1. 登录[华为云监控服务CES控制台](#)查看指标上报情况。

图 3-49 workload 指标开始归零



2. 登录[华为云弹性伸缩AS控制台](#)，单击本方案部署时创建的伸缩组名，进入实例监控台。

从下图可以看出，本次测试一共扩容出来六台渲染节点，手动结束渲染作业后，节点正常完成缩容。

说明：测试时并未将机器扩容至任务所需数量。

图 3-50 整个过程活动伸缩记录

### 📖 说明

- Deadline调度节点(Master)弹性云服务器每次重启之后，都需要执行重启Gearbox程序，方可恢复环境。Gearbox启动命令为：  
java -jar gearbox-0.0.1-SNAPSHOT.jar --spring.config.name=Config（当配置文件名不是application.yml时可使用此命令，Config是Gearbox配置文件名称，不要带后缀）
- 本部署方案中SFS-Turbo 选用的是通用型性能版，如果此版本无法满足您的业务对存储的要求，则您可在[弹性文件服务SFS控制台](#)创建新的SFS-Turbo。并参考[弹性文件服务挂载方式，自动挂载](#)。  
建议：
  - 线上线下载混合云渲染采用缓存型SFS-Turbo。
  - 全部都在云上渲染时，业务对存储要求较高可采用SFS-Turbo HPC型。

----结束

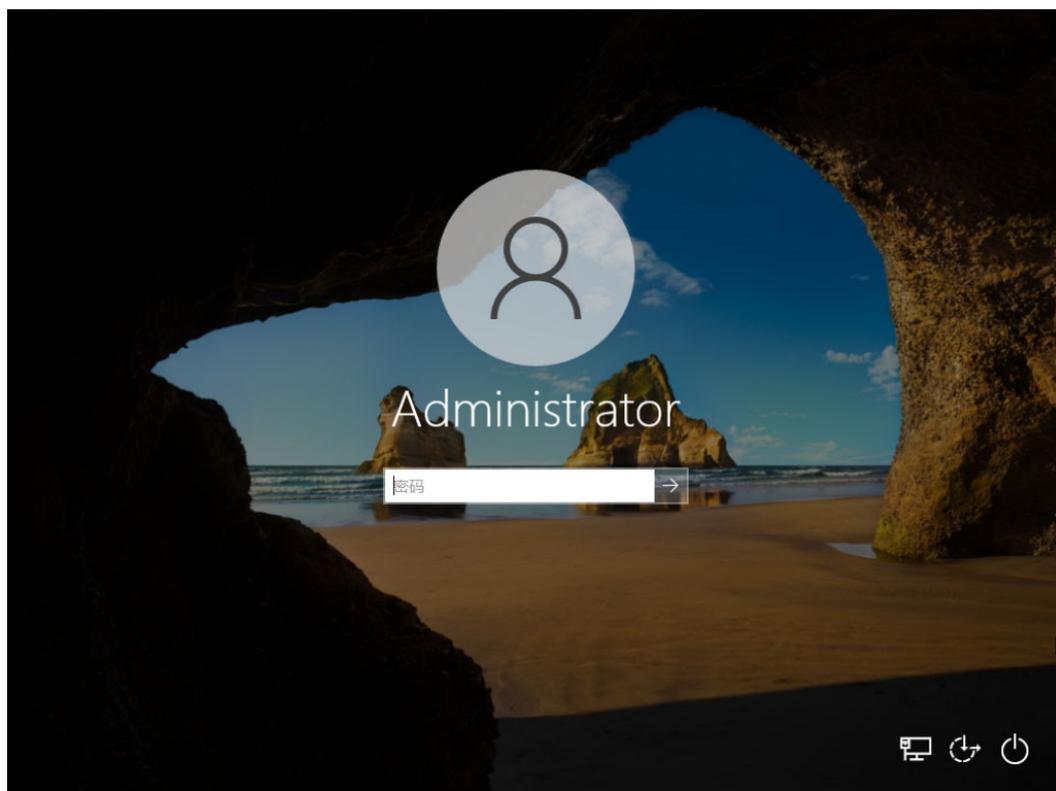
## Windows 环境安装与部署

### 步骤1 检查环境预装情况。

1. 登录[华为云服务器控制台](#)，选择"远程登录"--->选择"密码"登录Master节点。

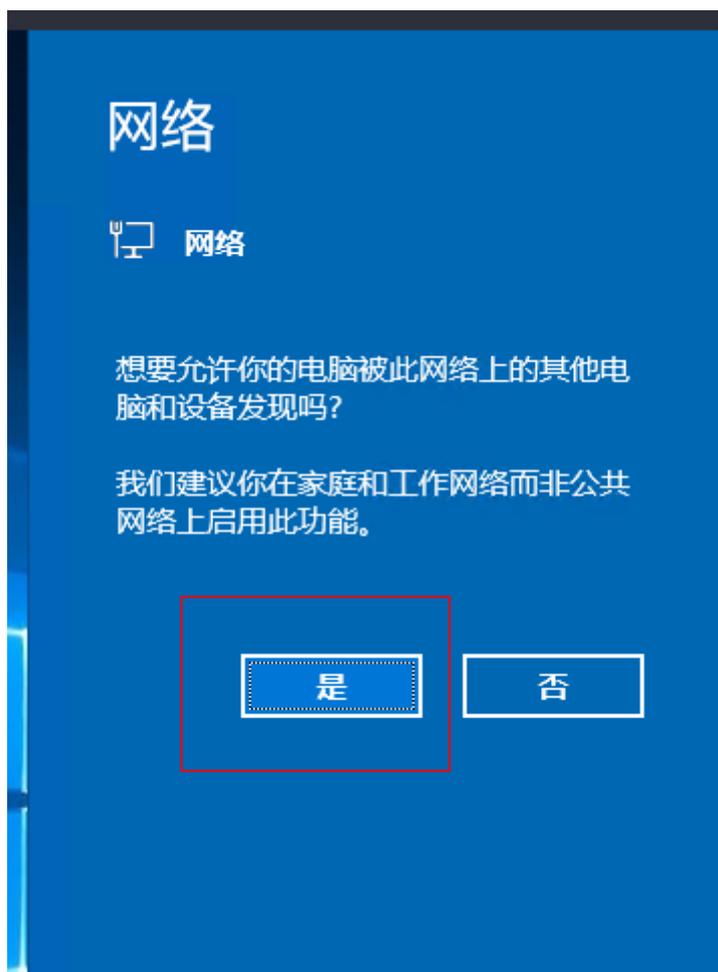
windows默认采用Administrator账户登录。

图 3-51 windows Administrator 登录



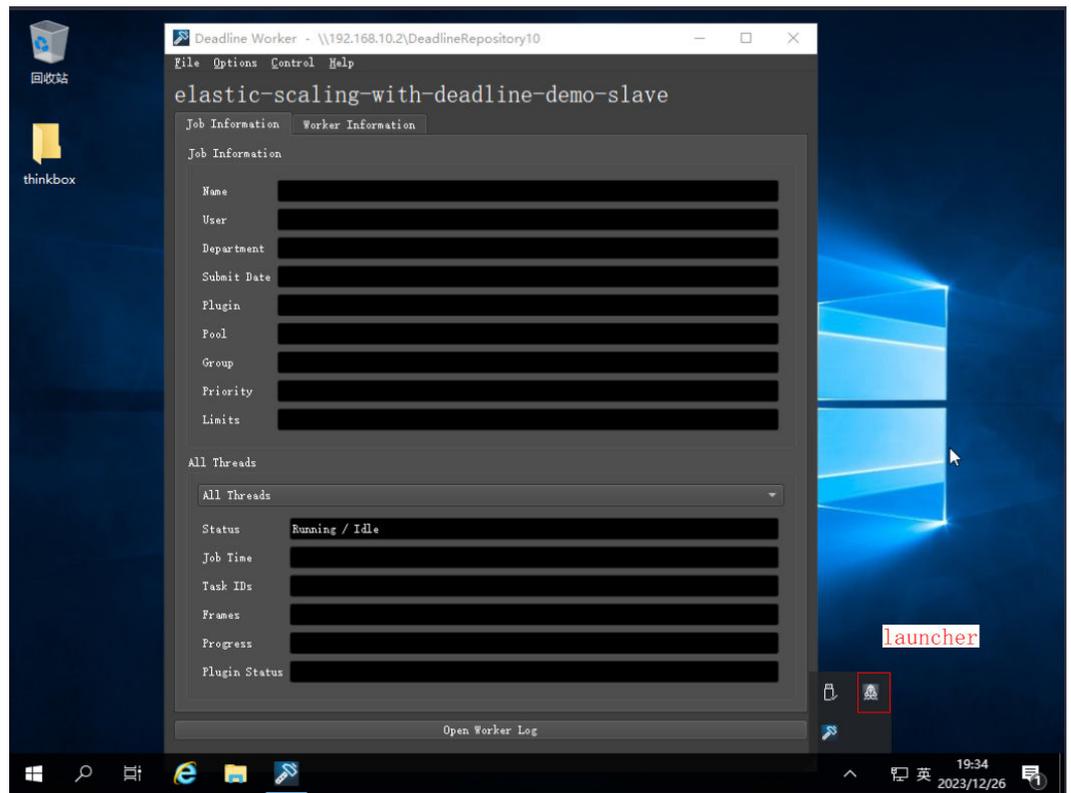
2. 进入系统后会出现如下弹窗，请选择“是”，便于后续完成Slave节点发现Master节点。Master节点的Deadline Worker在用户进入系统后会自动启动

图 3-52 是否共享网络



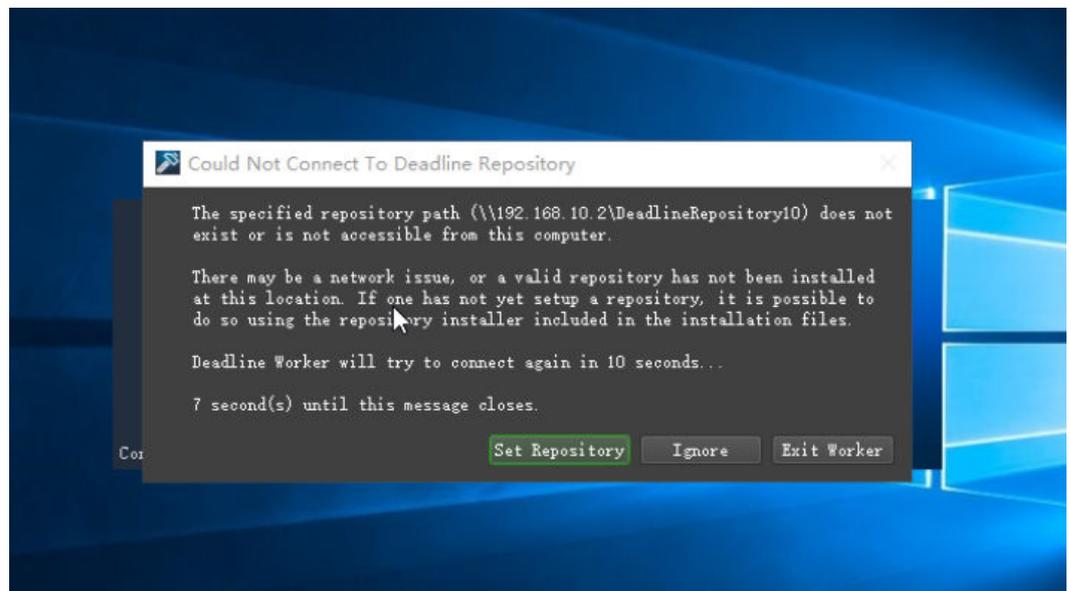
3. 检查Master节点能否发现Slave节点。  
进入[华为云服务器控制台](#)，登录Slave节点，使用远程登录，密码方式进入系统，针对共享网络弹窗请选择"是"。  
Deadline Worker 在用户登录进系统后会自动启动。  
如下图所示，表示Worker启动成功。

图 3-53 Deadline Worker 启动成功



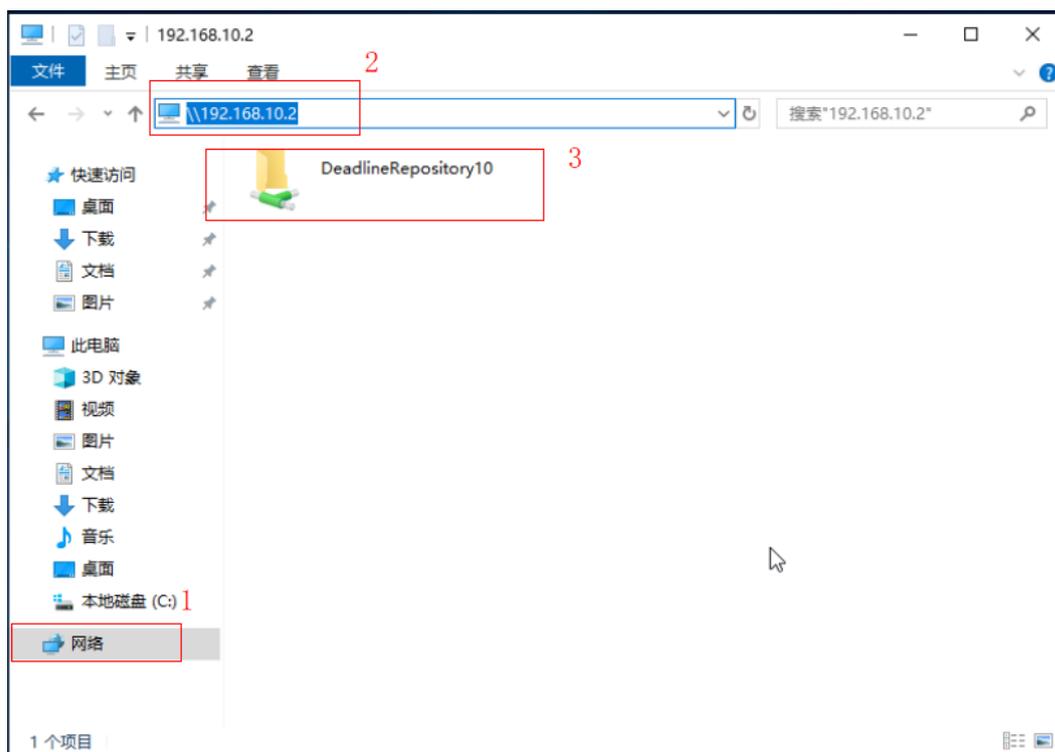
如遇下图问题要求配置Respositroy。

图 3-54 配置 Respository



请先在Slave节点上检查是否可以发现Master节点的共享目录DeadlineRepository10。下图表示Slave节点可以发现共享目录。

图 3-55 检查共享目录

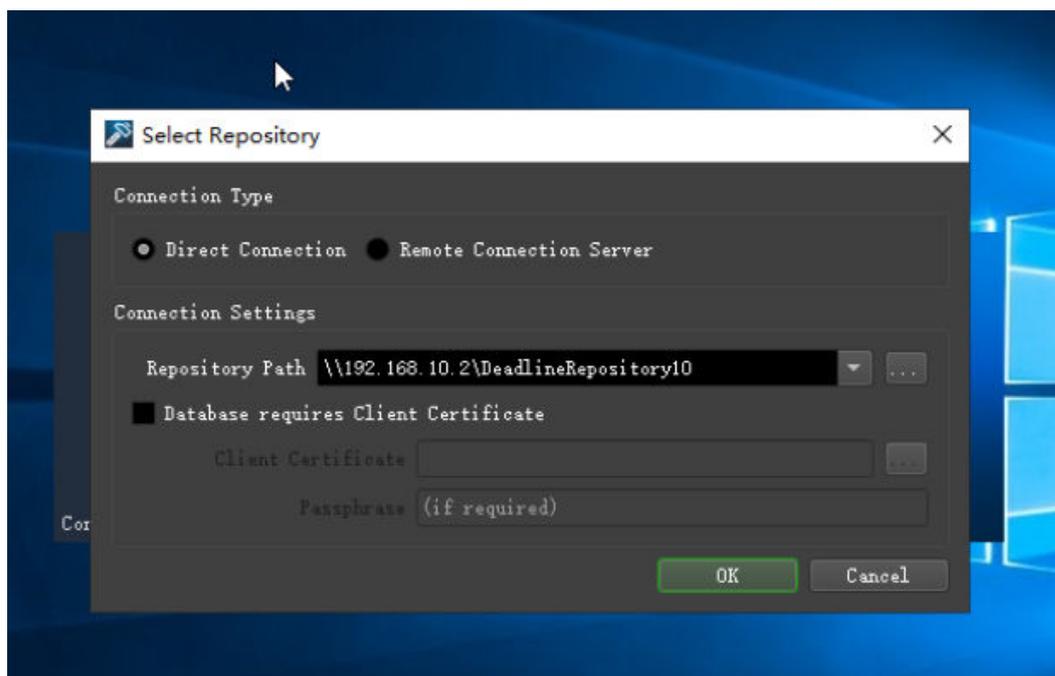


如果Slave节点无法发现共享目录，则请登录Master节点，检查C:\DeadlineRepository10 目录是否存在，如果不存在则请执行installRepo.bat 脚本重新安装Repository软件，脚本执行大约需要5min，执行完成后窗口会自行关闭；

或者检查C:\DeadlineRepository10目录是否共享成功，如图36表示共享成功；如果共享失败，则请手动共享 C:\DeadlineRepository10目录。

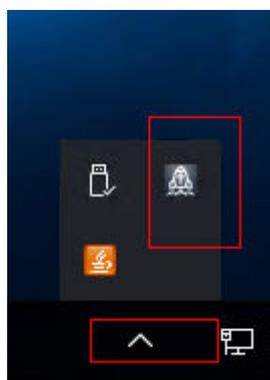
问题排除后，保持默认配置，单击"OK"，然后Deadline Worker便会启动。

图 3-56 配置 Repository



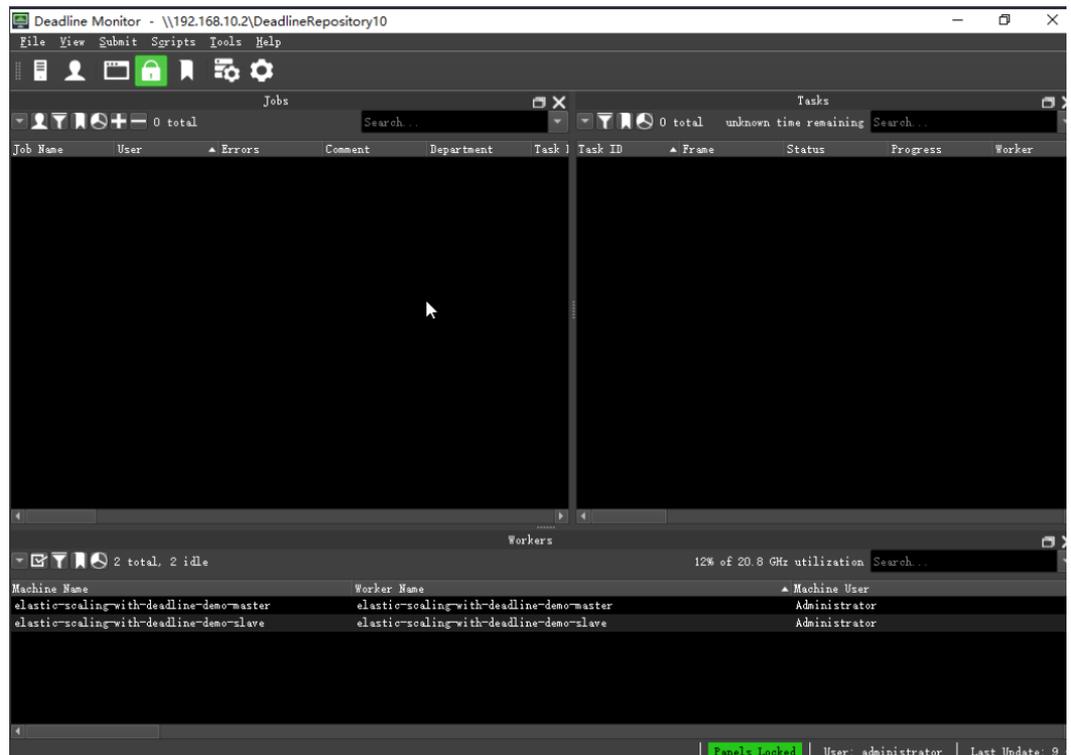
登录Master节点，启动Deadline Monitor查看集群状态。右键下面图标-->单击 Launcher Monitor 来启动Deadline Monitor。

图 3-57 启动 Deadline Monitor



如图39，Monitor面板可以发现Master和Slave两个节点。

图 3-58 集群状态



**步骤2** 配置Gearbox软件。最新版可从[Gearbox的Gitee仓库地址](#)下载，新版[Gearbox配置文件地址](#)（可选）。

登录调度节点Master对Gearbox进行配置。请保持Deadline Monitor开启。

1. 进入桌面gearbox目录下，对Gearbox的配置文件application.yml进行编辑，对可选项根据需要自行配置。说明：ak、sk、project-id、group等必要的启动项已完成自动配置。

配置项说明如下：

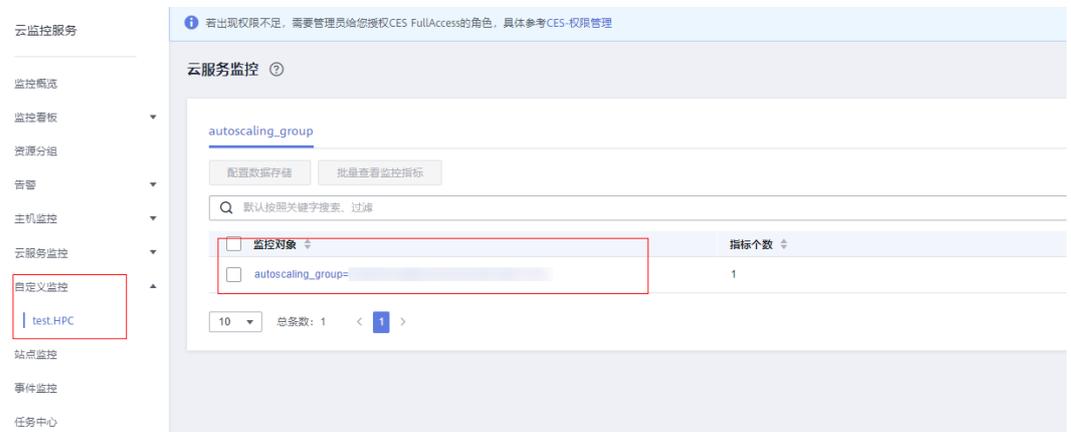
```
user:
  # AK硬编码或者明文存储都存在较大的安全风险,建议在配置文件或者环境变量中密文存储,使用时解密,确保安全
  ak:
  # SK硬编码或者明文存储都存在较大的安全风险,建议在配置文件或者环境变量中密文存储,使用时解密,确保安全
  sk:
  # 待使用局点的租户ID
  project-id:
  # 代理地址, 端口, 用户名密码等, 无需代理可不配置
  proxy-address:
  proxy-port:
  proxy-user-name:
  proxy-password:
as:
  # AS服务在待使用局点的终端节点域名
  endpoint: as.cn-north-4.myhuaweicloud.com
  # 预置伸缩组资源的伸缩组ID
  group: f12d57a3-ed8a-4a16-8c3f-4817ab7171f2
  # 查询伸缩实例时单页最大返回数量, 默认100可不修改
  list-instance-limit: 100
  # 删除伸缩实例时最大数量限制, AS服务最大支持50, 可不修改
  delete-instance-limit: 50
ecs:
  # ECS服务在待使用局点的终端节点域名
  endpoint: ecs.cn-north-4.myhuaweicloud.com
```

```
metric:
# 自定义监控指标命名空间，可不修改
namespace: test.HPC
# 自定义指标的指标名
name: workload
# 自定义指标维度名，可不修改
dimension-name: autoscaling_group
# 自定义指标维度ID，可配置成伸缩组的ID，该值不影响功能
dimension-id: f12d57a3-ed8a-4a16-8c3f-4817ab7171f2
# 指标上报的TTL参数，可不修改
report-ttl: 172800
# CES服务在待使用局点的终端节点域名
metric-report-endpoint: ces.cn-north-4.myhuaweicloud.com
task:
# 节点状态检查周期，单位 秒
health-audit-period: 30
# 自定义指标上报周期，单位 秒
metric-report-period: 60
# 检测是否需要扩容的周期，单位 秒
scale-in-period: 5
# 自动删除待扩容节点周期，单位 秒
delete-instance-period: 5
# 新扩容节点自发现周期
discover-instance-period: 20
# 对比伸缩组与slurm节点数量周期，单位 秒
diff-instance-and-node-period: 60
# 内部缓存刷新周期，单位 秒
refresh-cache-period: 100
# 配置检查任务运行周期，单位 秒
configuration-check-period: 300
system:
# deadline/slurm
type: deadline
# 稳定节点的Node名称，多节点英文逗号分割
stable-nodes: deadline-master
# 稳定节点所在的分区
stable-partition: dyn1
# 不稳定节点所在的分区
variable-partition: dyn1
# 空闲时间阈值，空闲时间超过该值的节点将被扩容，单位 秒
scale-in-time: 1
# 判断job是否属于排队状态的时间限制条件，排队时间超过该值的job被认为是在排队并参与指标计算,建议为0
job-wait-time: 1
# 新节点注册的超时时间，超过该时间依然未注册成功将被AS删除，单位：分钟，建议为10
register-timeout-minutes: 10
# 弹性节点使用的cpu核数
cpu: 4
# 弹性节点使用的内存大小，该字段为预留字段，可设置为任意大于0的数值
memory: 12600
gpu: 1
# 仅slurm集群配置有效，默认是true，在计算Workload时忽略含有GPU需求作业中的CPU需求量
ignore-cpu-request-of-gpu-job: true
# 是否使用powershell执行指令，只有windows环境下生效
use-powershell-executor: true
# 支持多分区资源协同
partition:
# 特性整体开发，默认值为false，对接系统为slurm且该配置开启后后续配置生效；
multi-partition-scaling: false
# 提交到多个分区的作业负载，在多个分区上的映射策略 MAPPING_TO_FIRST_PARTITION，仅累积到第一个分区中，其余分区不计算(默认)；MAPPING_BY_WEIGHT，按照各分区的权重计算负载的映射；
load-mapping-policy: MAPPING_TO_FIRST_PARTITION
# 分区配置列表
partition-configuration-list:
-
# 分区名
name: partition1
# 分区中节点资源类型，支持CPU、GPU二选一；
type: CPU
```

```
# 分区在负载映射中的计算权重，整数，取值范围[0,100];
weight: 10
# 分区对应的租户id
project-id:
# 分区对应的伸缩组id，伸缩组与分区需要一一对应
scaling-group-id: XXX
# 节点的资源规格，当资源类型为CPU时可填节点CPU和内存规格，当节点资源类型为GPU时可填节点规格为GPU数量;
node-spec:
# 节点CPU数量
cpu: 4
# 节点内存容量，单位GB;
mem: 8
# 加速卡数量
gpu: 2
# 空闲一定时间后触发缩容，单位秒，未配置时使用全局配置，配置后针对本分区覆盖全局空闲缩容时间;
scale-in-after-idle-time: 100
# 自定义指标维度名
dimension-name: autoscaling_group
# 自定义指标维度ID
dimension-id: f12d57a3-ed8a-4a16-8c3f-4817ab7171f2
# 云端API调用的地址端点列表，未配置时使用全局配置
ecs-endpoint: ecs.cn-north-4.myhuaweicloud.com
as-endpoint: as.cn-north-4.myhuaweicloud.com
ces-endpoint: ces.cn-north-4.myhuaweicloud.com
```

2. Gearbox配置完成后，在gearbox目录下双击startGearbox.bat脚本启动Gearbox程序进行测试。**说明：Gearbox程序启动后会在终端打印日志信息，请检查有无报错信息。**
3. 登录[华为云监控服务CES控制台](#)查看自动定义指标workload是否上报成功。

图 3-59 自定义指标



**说明：未提交任务，故workload指标为0。**

图 3-60 查看自定义指标



**步骤3** 挂载SFS-Turbo。

登录调度节点Master，参照如下步骤挂载SFS-Turbo

1. 安装NFS服务端和客户端

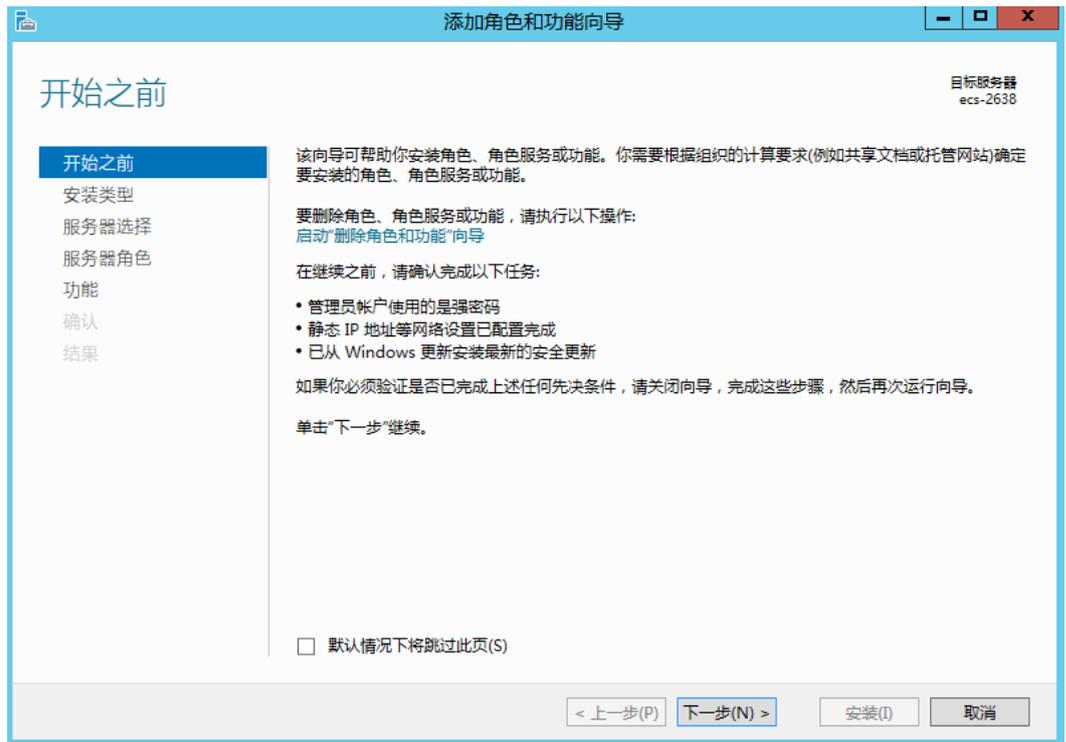
单击左下角“服务器管理器”，弹出“服务器管理器”界面，

图 3-61 服务器管理



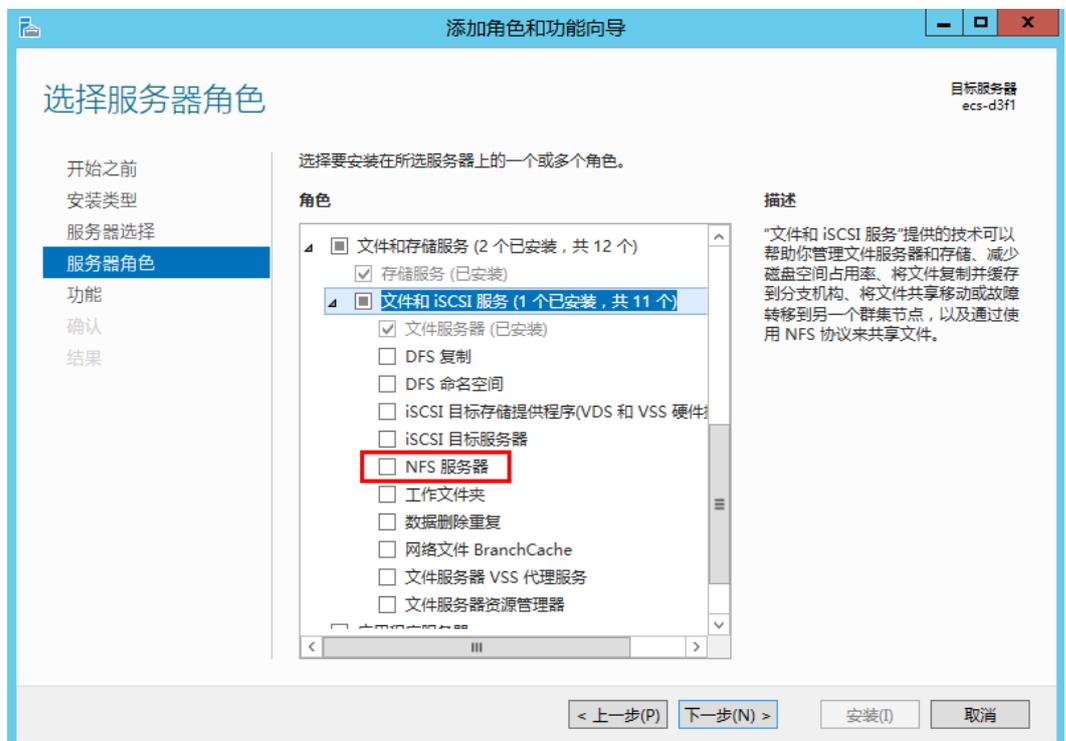
单击“添加角色和功能”

图 3-62 添加角色和向导



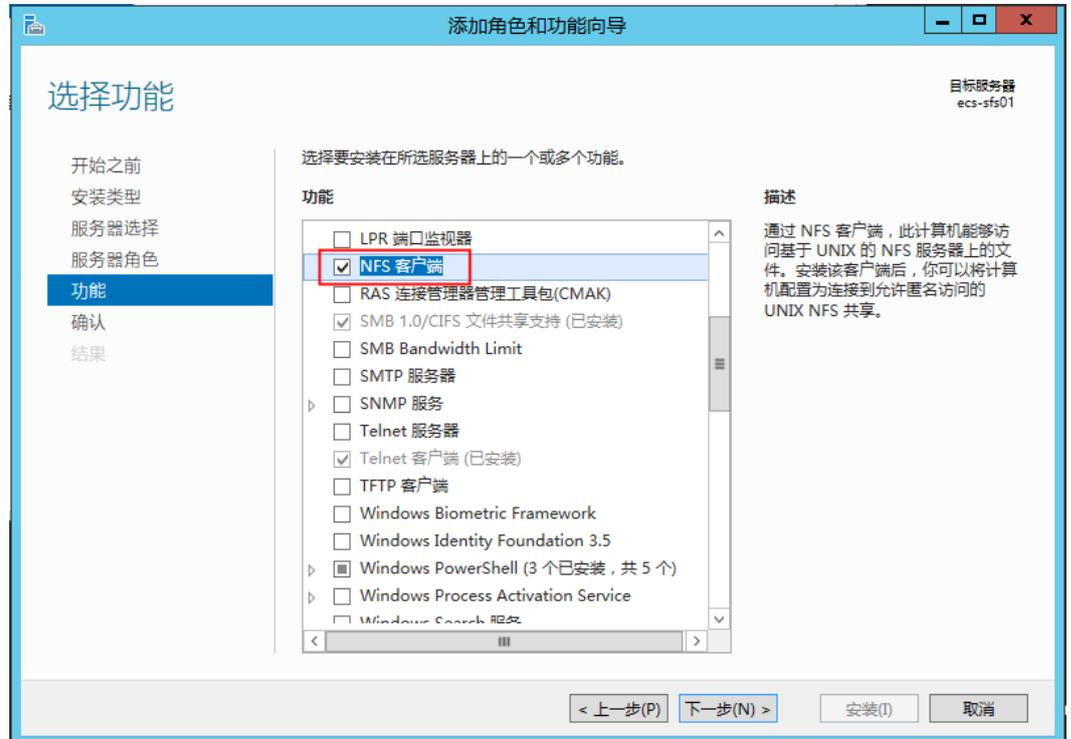
根据系统提示操作单击“下一步”，在“服务器角色”，选择“NFS服务器”，

图 3-63 选择 nfs 服务器



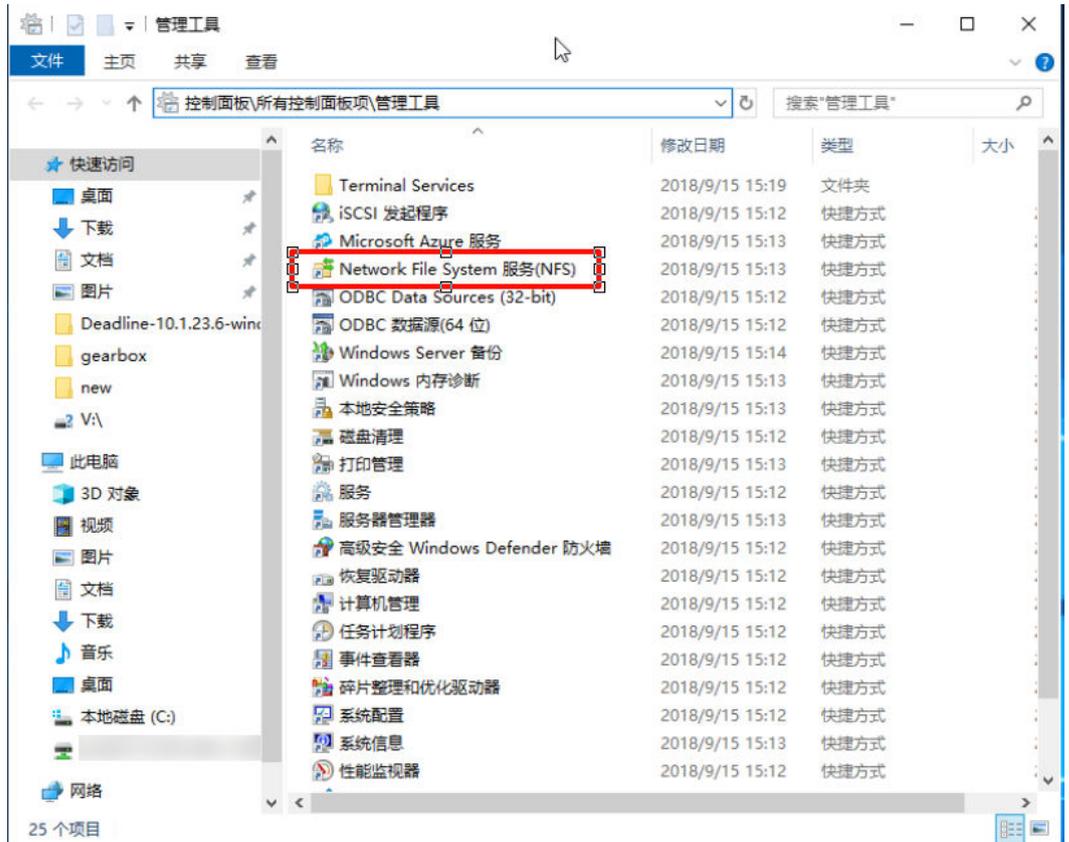
单击“下一步”，在“功能”里选择“NFS客户端”，单击“下一步”。确认无误后单击“安装”。安装结束后，如果首次安装NFS客户端，需要按照系统提示重新启动客户端并重新登录云服务器。

图 3-64 选择 nfs 客户端



2. 修改NFS传输协议。  
选择“控制面板 > 所有控制面板> 管理工具 > Network File System 服务 (NFS)”。

图 3-65 管理工具



右键单击“NFS客户端”选择“属性”，修改传输协议为“TCP协议”，同时选中“使用硬装载”

图 3-66 网络文件系统服务

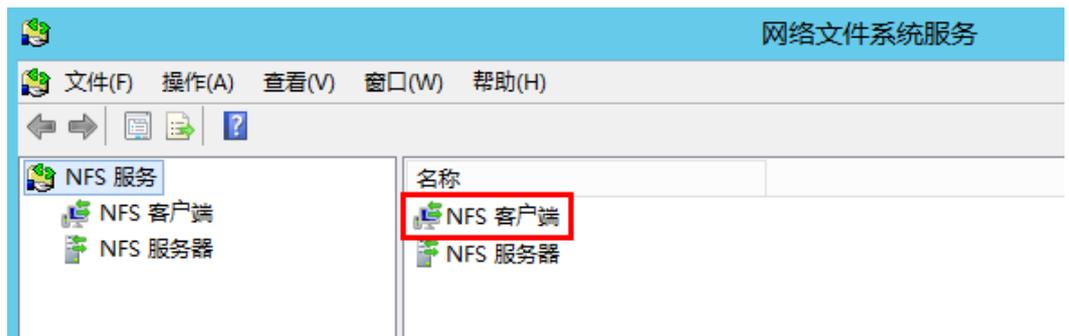
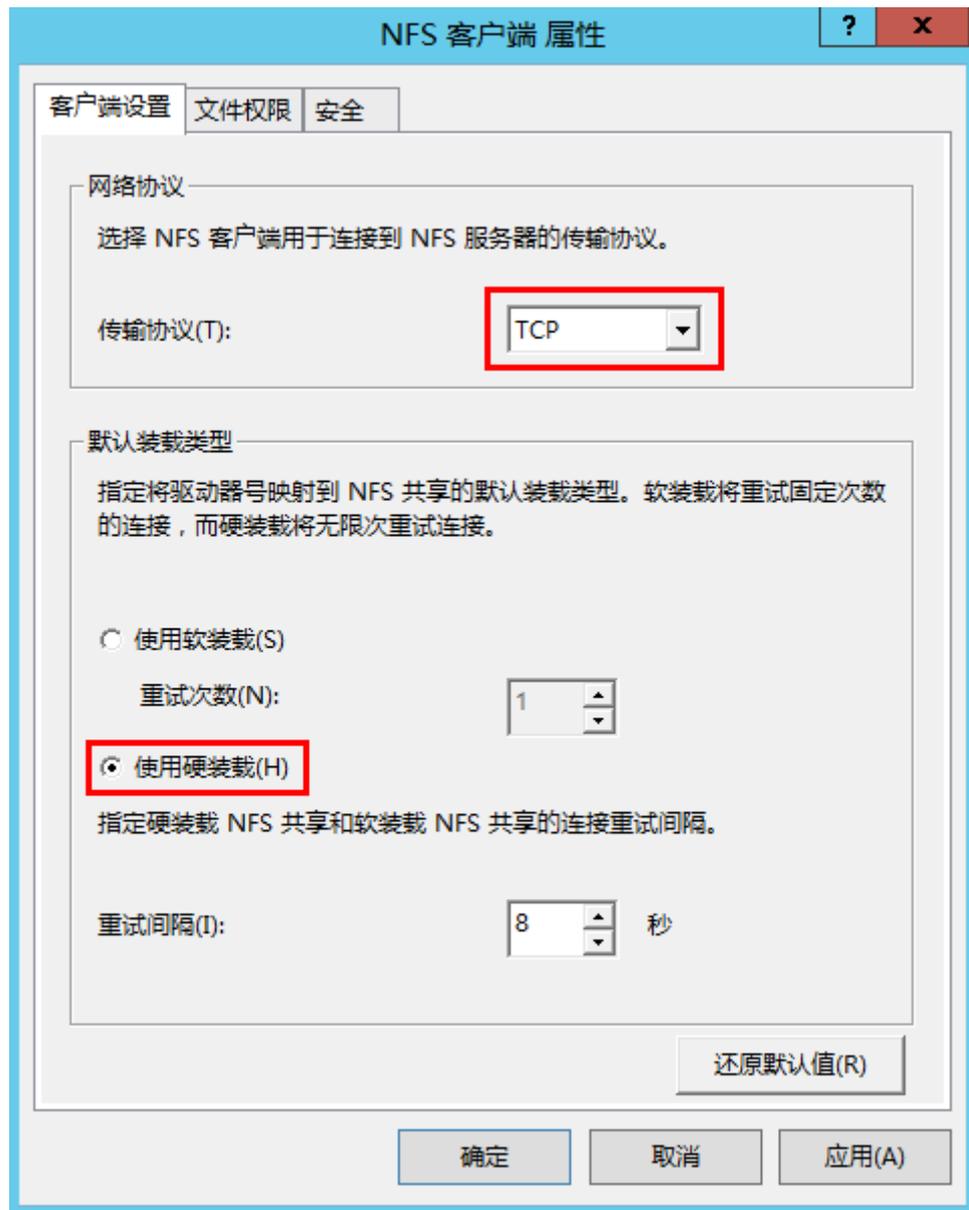


图 3-67 nfs 客户端属性



### 3. 挂载文件系统

- 方式一：双击桌面上的auto\_mount.bat 脚本进行自动挂载SFS-Turbo文件系统。
- 方式二：手动挂载

启动cmd终端，使用如下命令挂载，"!" 必须有。登录[SFS-Turbo控制台](#)，查看挂载ip。

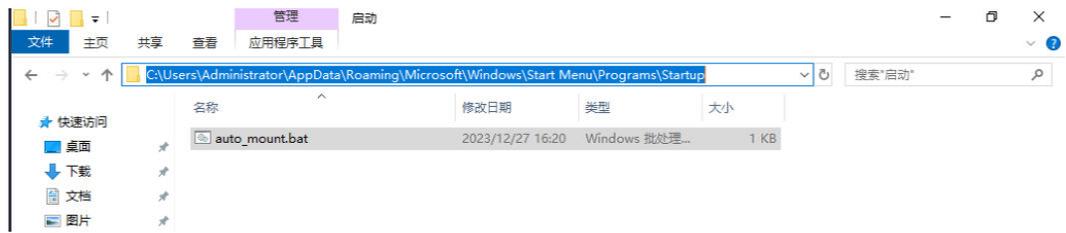
建议命令执行完成后，在"此电脑"中查看是否存在 X 盘符，并可以进入X盘创建文件/文件夹。

```
mount -o nolock -o casesensitive=yes ip:!/ X: #此处ip为sfs-turbo挂载ip
```

### 4. 开机自动挂载SFS-Turbo。也可参考官网[自动挂载SFS-Turbo](#)。

- 将auto\_mount.bat脚本移动到windows开机软件自启动目录，路径如下图所示：

图 3-68 软件自启动目录



**步骤4** 配置Slave节点环境。

登录Slave节点，进行如下内容的配置。

1. 配置开机免密登录。

打开cmd终端，输入如下命令

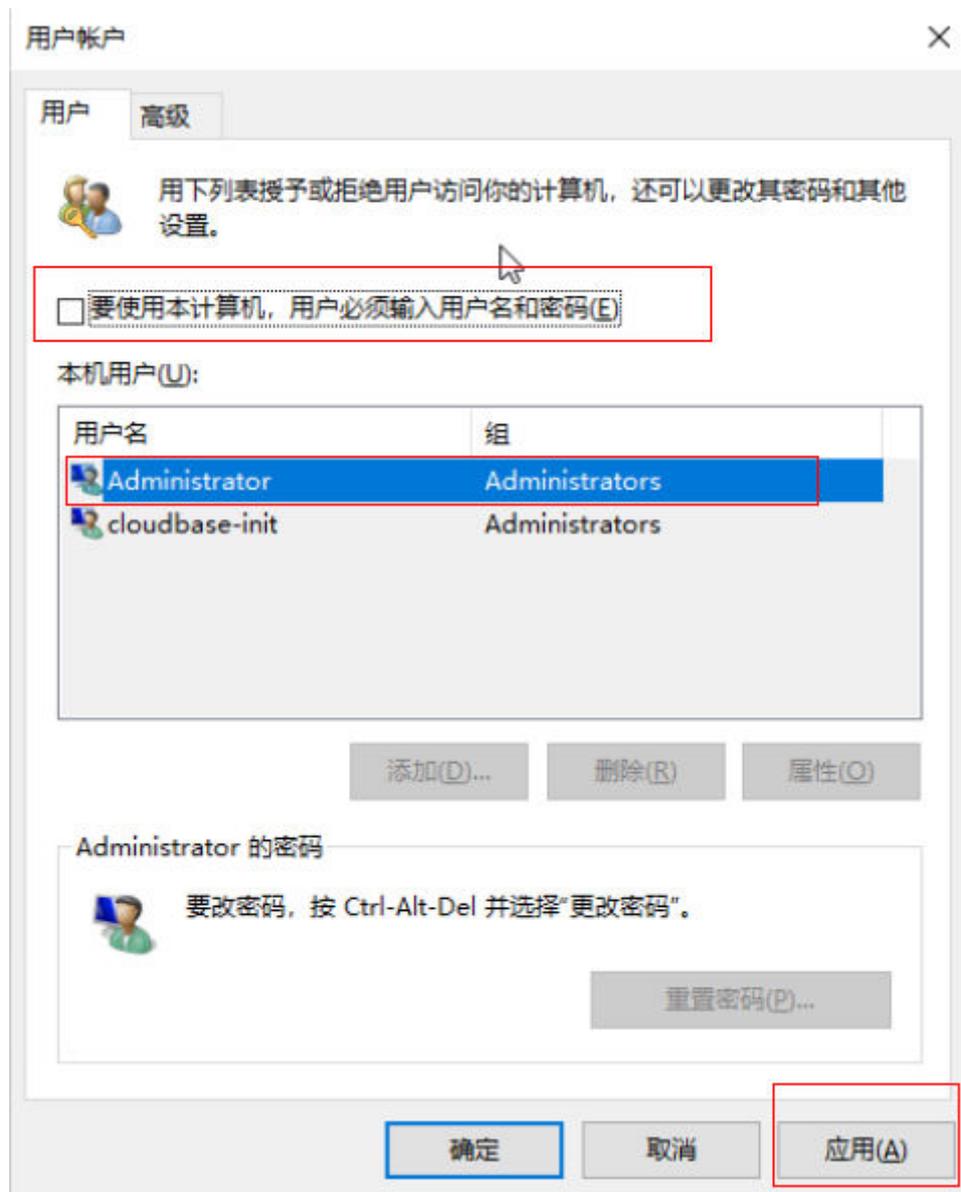
```
netplwiz
```

图 3-69 配置开机免密登录



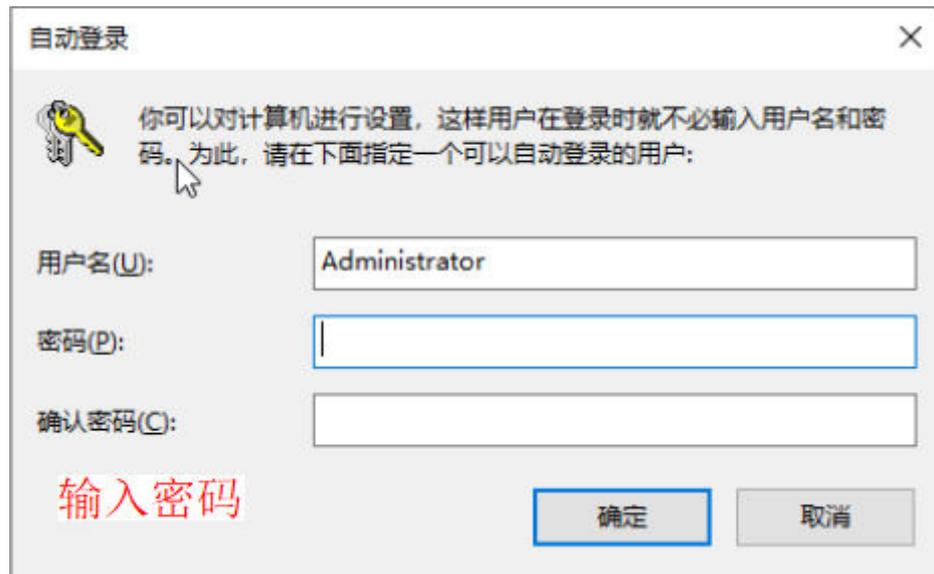
取消下图中Administrator对应勾选框。

图 3-70 配置免密登录



输入密码配置自动登录。

图 3-71 配置自动登录



2. Slave节点挂载SFS-Turbo。参考**步骤3：挂载SFS-Turbo。**

**步骤5** 为Master节点和Slave节点自行安装业务软件，并配置开机自启。

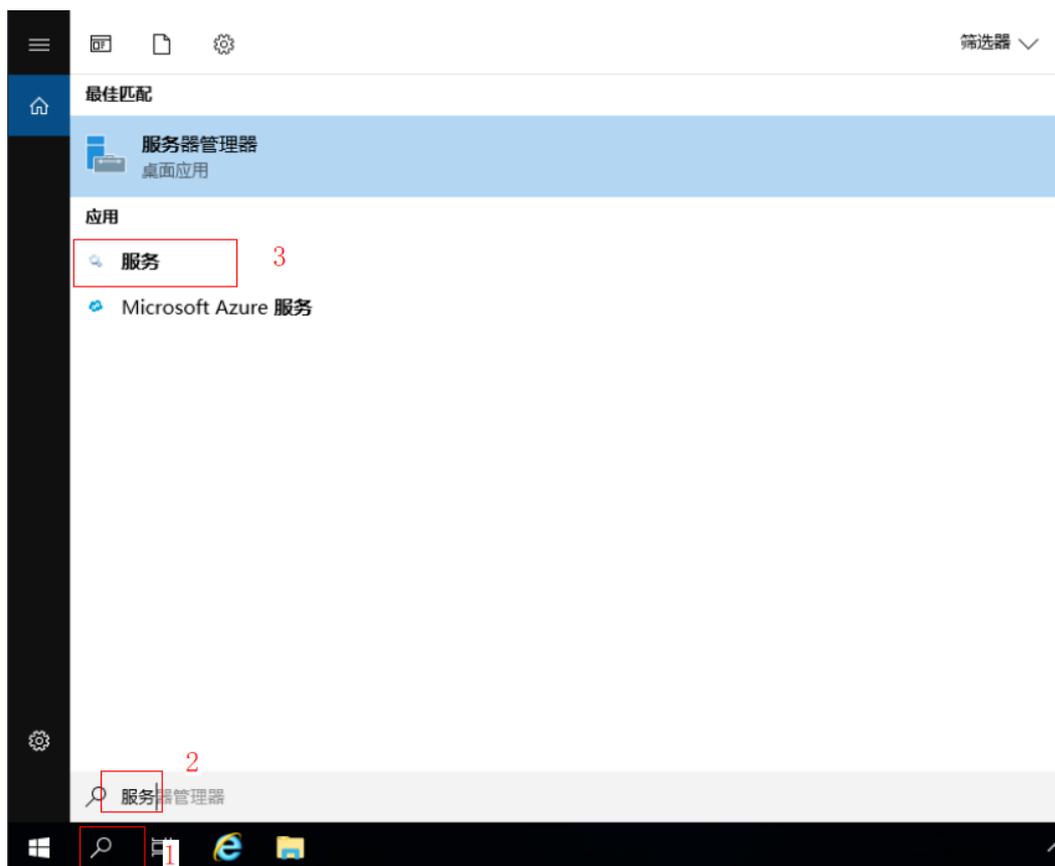
**步骤6** 测试Slave节点。安装完成业务软件后，在Master节点提交测试任务，检查Slave节点能否正常渲染。**说明：渲染资产请存放在SFS-Turbo挂载盘符 X: 下。**

**步骤7** 将Slave节点打包为私有镜像。

Slave节点测试正确后，登录Slave节点开始制作私有镜像。

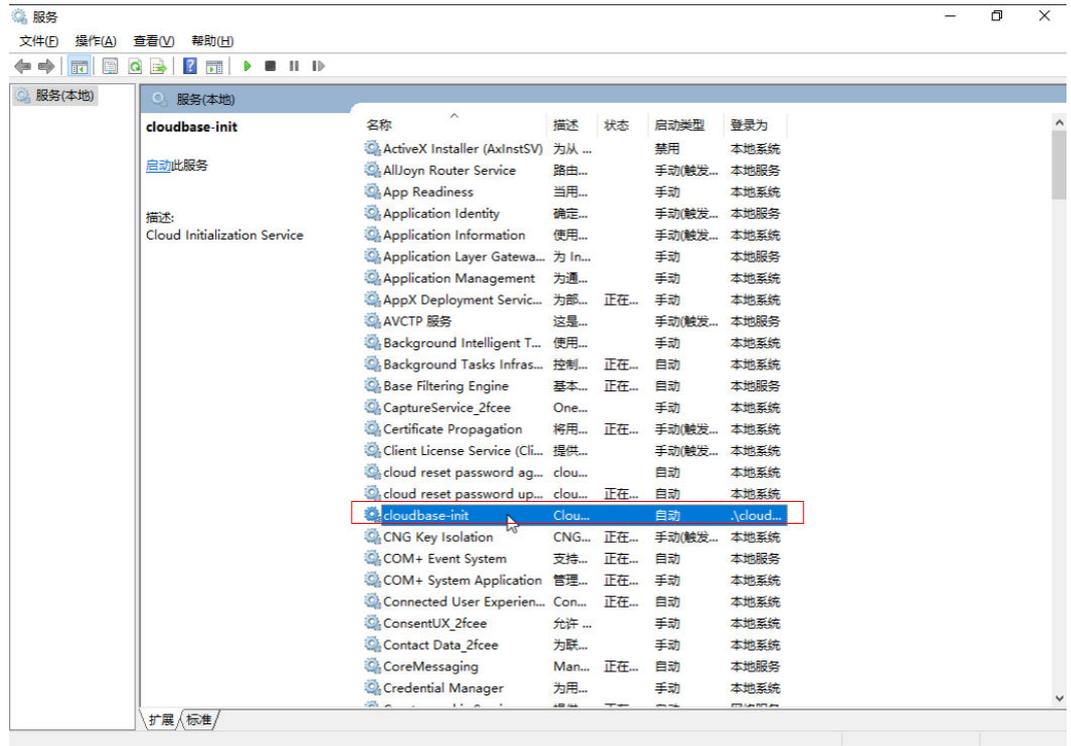
1. 配置cloudbase-init 本地账户登录。  
在windows搜索窗口搜索“服务”。

图 3-72 搜索服务



如下图所示。在服务列表中，找到cloudbase-init服务。单击 cloudbase-init 服务进入服务的配置窗口。

图 3-73 选择 cloudbase-init 服务



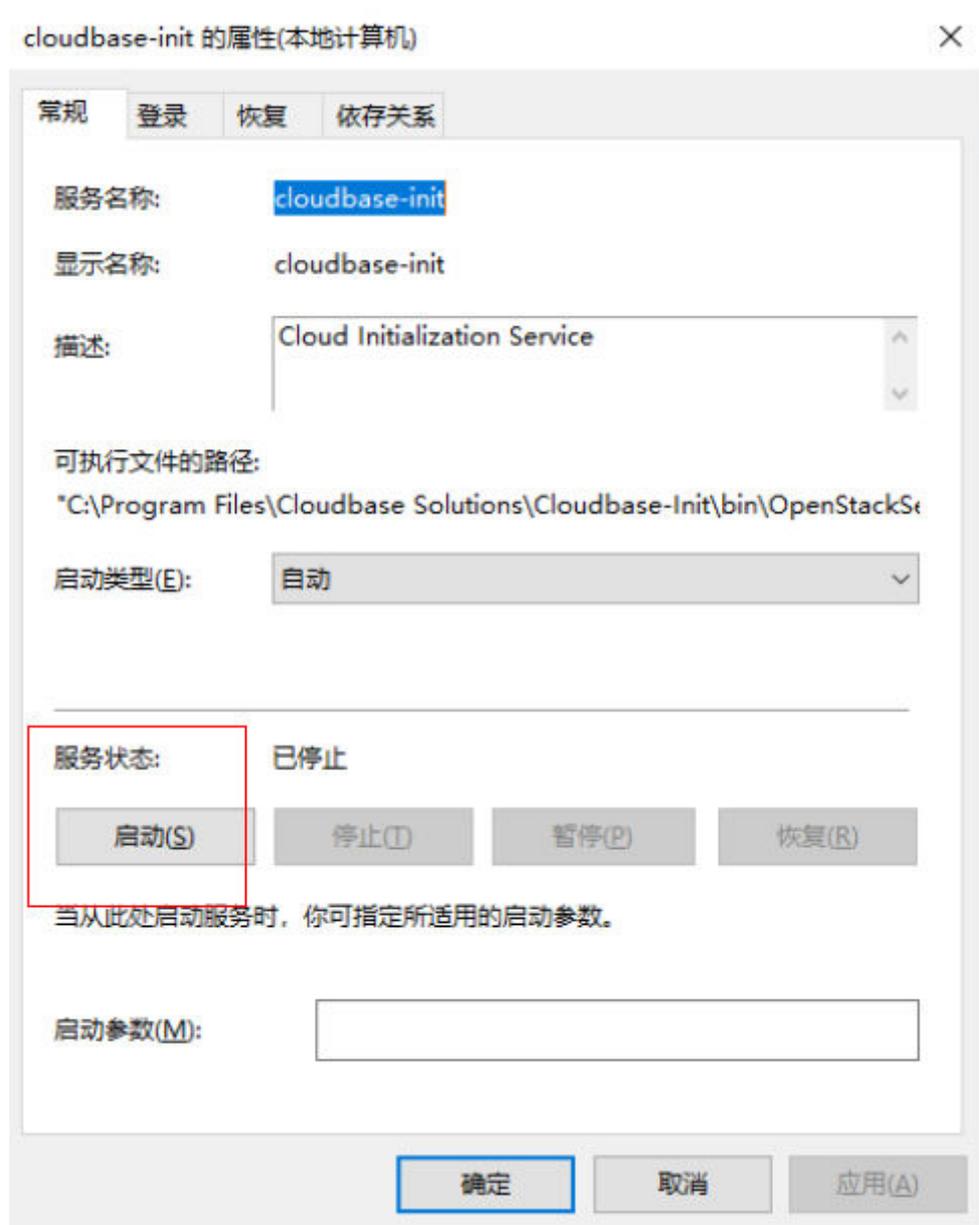
选择"登录"---->登录身份选择"本地系统账户"

图 3-74 cloudbase-init 配置本地账户



修改登录账户后，启动cloudbase-init，检查有无报错。选择 "常规"----> "启动"。说明：请保持启动类型为"自动"

图 3-75 启动 cloudbase-init



2. 自行删除Slave节点中的敏感信息。
3. 镜像预处理操作完成后, 登录[华为云服务器控制台](#), 将Slave节点关机, 开始打包Slave节点镜像。  
找到Slave节点对应的云服务器, 单击右侧“更多>镜像>创建镜像”进行打包。

图 3-76 打包私有镜像



图 3-77 创建私有镜像



在创建私有镜像界面，填写名称，选择企业项目，勾选协议，单击立即创建，等待镜像创建完成。

图 3-78 配置信息



打开[华为云镜像服务控制台](#)，等待大约5-10min，查看镜像是否创建成功。

图 3-79 镜像创建结果查询



步骤8 修改伸缩配置。

登录[华为云弹性伸缩AS控制台](#)，选择本方案部署时创建的伸缩组，右侧选择“更多”--->“修改伸缩配置”。

说明：修改伸缩配置时，AS会自动为“名称”添加“\_copy”的后缀，请将此后缀删除，因为伸缩配置名称带有下划线“\_”时，会导致节点扩容异常。

图 3-80 修改伸缩配置的名称



建议选择C6、C6s、C7，规格选择32U64G、64U128G。

图 3-81 配置镜像规格



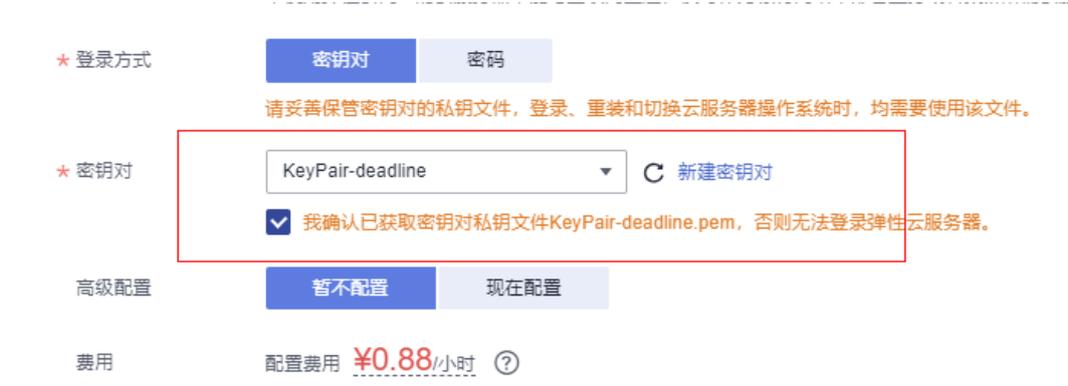
镜像选择上一步骤中创建的镜像，磁盘类型和容量根据业务需要自行配置，安全组选择方案部署时创建的安全组。

图 3-82 配置镜像、磁盘、安全组



登录方式选择"密钥对"。密钥对使用部署方案时创建的密钥对。

图 3-83 配置登录方式和密钥对



**步骤9** 按需配置伸缩组。可选

登录[华为云弹性伸缩AS控制台](#)，选择右侧“更多”--->"修改伸缩组"。

图 3-84 修改弹性伸缩组



说明：最大实例数、期望实例数、最小实例数根据业务需要自行配置，冷却时间建议小于30s，其他配置可保持默认。

图 3-85 修改伸缩组配置

**修改伸缩组**

1 伸缩组为非启用状态、实例数为0且没有正在进行的伸缩活动时，可以修改伸缩组的子网。

\* 名称: elastic-scaling-with-deadline-demo-group

\* 最大实例数(台): 50

\* 期望实例数(台): 0

\* 最小实例数(台): 0

\* 冷却时间(秒): 10

\* 可用区: 可用区1, 可用区2

\* 多可用区扩展策略:  均衡分布  选择优先 ?

负载均衡:

\* 健康检查方式: 云服务器健康检查

说明：下面配置根据业务需要自行配置。

图 3-86 修改伸缩组配置

**修改伸缩组**

\* 多可用区扩展策略:  均衡分布  选择优先 ?

负载均衡:

\* 健康检查方式: 云服务器健康检查

\* 健康检查间隔: 10秒

\* 健康状况检查宽限期(秒): 60

\* 实例移除策略: 根据较早创建的配置较早创建的实例

\* 企业项目: default

弹性公网IP:

数据盘:

委托: 请选择-  ?

新的委托仅对修改后创建的云服务器生效，已有云服务器的委托请前往[弹性云服务器控制台](#)修改。

步骤10 启动Gearbox软件。

登录Master节点，开启Deadline Monitor。

进入桌面上gearbox文件夹，双击bat脚本启动gearbox程序。

### 步骤11 配置伸缩组的伸缩策略。

登录[华为云弹性伸缩AS控制台](#)，单击弹性伸缩组名称，进入伸缩组详情界面。

图 3-87 添加策略



策略类型选择"告警策略"，告警规则选择"现在创建"，监控类型选择"自定义监控"，触发条件选择"workload"、"原始值"。

图 3-88 添加伸缩策略



图 3-89 添加伸缩策略

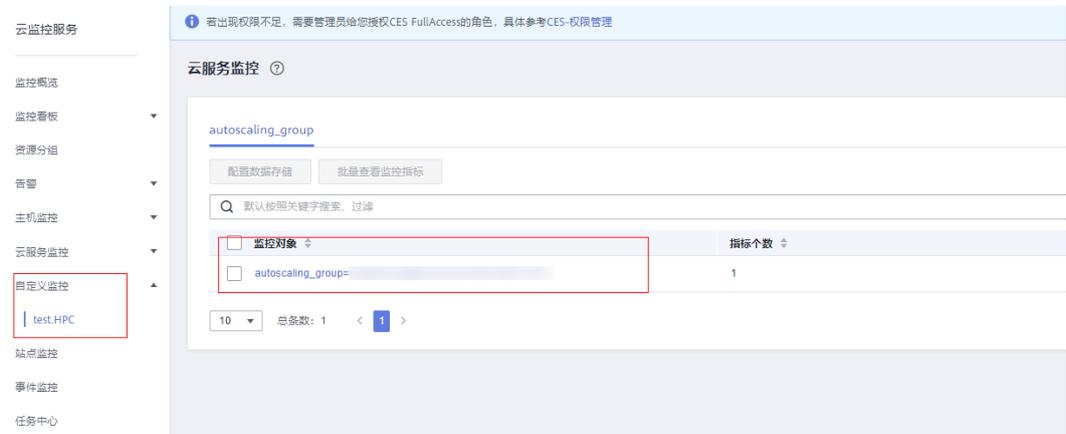


### 步骤12 提交测试任务。登录Master节点，在Deadline Monitor提交测试任务。保持Gearbox处于运行状态。说明：由于Master节点配置较低，建议渲染时关闭Dealine Worker。

**步骤13** CES控制台查看指标上报情况。

登录[华为云监控服务CES控制台](#)，查看该解决方案创建的自定义监控。

**图 3-90** 自定义监控



监控指标workload值表示集群需要扩容的节点数量，如下图workload为23时，表示需要扩容23个计算节点。

**图 3-91** 某时刻指标监控情况



**步骤14** AS查看节点伸缩情况。

登录[华为云弹性伸缩AS控制台](#)，单击本方案部署时创建的伸缩组名，进入实例监控台。

如下图，可见AS已经开始扩容节点。

图 3-92 某时刻节点扩容情况



步骤15 Deadline Monitor查看节点扩容情况。

登录Master节点，在Deadline Monitor查看任务运行情况，渲染节点状态。

如下图，Master节点成功发现扩容出来的渲染节点。

图 3-93 某时刻渲染节点状态



步骤16 结束任务。

如步骤13，登录[华为云监控服务CES控制台](#)，查看指标上报情况。

图 3-94 workload 指标开始下降归零



如步骤14，登录[华为云弹性伸缩AS控制台](#)，查看节点缩容情况。

图 3-95 节点完成缩容



如下图，整个渲染过程AS弹性伸缩节点记录。

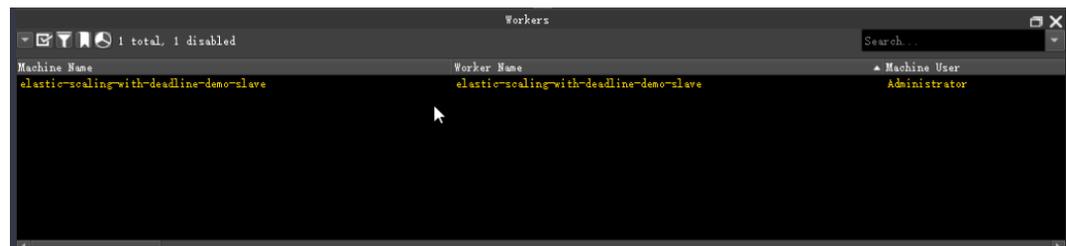
图 3-96 伸缩记录

The screenshot shows the '活动历史' (Activity History) page for the AS group. The table lists several scaling events, all with a status of '成功' (Success). The '描述' (Description) column is highlighted with a red box, showing details for manual scaling and auto-scaling events.

活动ID	状态	活动类型	描述
92eb92ae-5705-4619-abae-80d8f711679	成功	实例伸缩	手动触发，减少4个实例。
9f0d2f5f-78c8-4a17-a61e-c6508c39eadd	成功	实例伸缩	手动触发，减少1个实例。
a83c8eb6-3b26-460a-b2ea-2659a990a966	成功	实例伸缩	告警策略触发，增加2个实例。 策略名称: as-policy-9328
40eb0f0c-e551-4330-a049-60d22b352e1d	成功	实例伸缩	手动触发，减少1个实例。
23a181b8-c207-484b-a98a-ec53c0f883ef	成功	实例伸缩	告警策略触发，增加2个实例。 策略名称: as-policy-9328
f05bf5c-6292-490c-a949-641dae9e92f2	成功	实例伸缩	告警策略触发，增加2个实例。 策略名称: as-policy-9328

登录Master节点，查看Deadline Monitor 控制面板。

图 3-97 面板中只保留 Slave 节点



### 📖 说明

- Deadline调度节点(master)弹性云服务器每次重启之后，都需要执行重启Gearbox程序，方可恢复环境。Gearbox启动命令为：  
java -jar gearbox-0.0.1-SNAPSHOT.jar --spring.config.name=Config（当配置文件名不是application.yml时可使用此命令，Config是Gearbox的配置文件，命令中不可以带后缀）
- 本部署方案中SFS-Turbo 选用的是通用型性能版，如果此版本无法满足您的业务对存储的要求，则您可在[弹性文件服务SFS](#)控制台创建新的SFS-Turbo。并参考[弹性文件服务挂载方式，自动挂载](#)。

建议：

- 线上线下载混合云渲染采用缓存型SFS-Turbo。
- 全部都在云上渲染时，业务对存储要求较高可采用SFS-Turbo HPC型

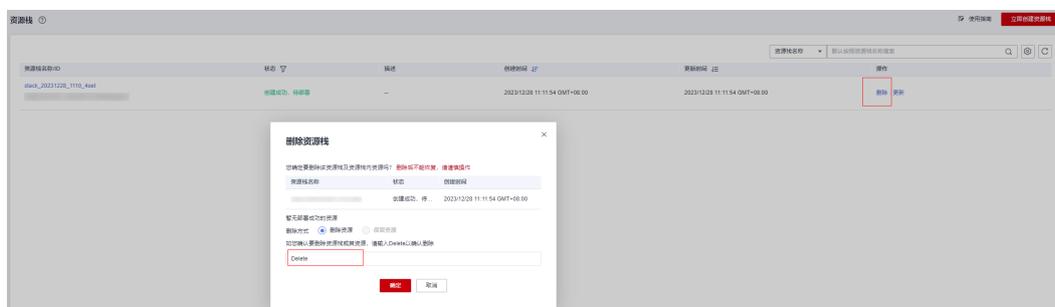
----结束

## 3.4 快速卸载

### 一键卸载

**步骤1** 登录[资源编排服务 RFS](#)控制台，找到该解决方案创建的资源栈，单击资源栈名称最右侧 "删除" 按钮，在弹出的 "删除资源栈" 提示框输入 "Delete"，单击 "确定" 进行解决方案卸载。

图 3-98 解决方案卸载



**步骤2** 手动删除旧的弹性伸缩配置

因每次修改弹性伸缩配置本质上都是创建新的弹性伸缩配置，所以需要手动删除旧的弹性伸缩配置。

登录[华为云弹性伸缩AS控制台](#)，进入 "伸缩配置" 页面将旧的伸缩配置进行删除。

图 3-99 删除旧配置



**步骤3** 根据需要删除制作的Slave节点私有镜像

登录[华为云镜像服务控制台](#)，删除创建的私有镜像。

----结束

# 4 附录

---

## Gearbox 程序工作流程介绍

本方案以Maya渲染任务为示例，演示Gearbox程序如何自动进行Deadline集群弹性扩容。

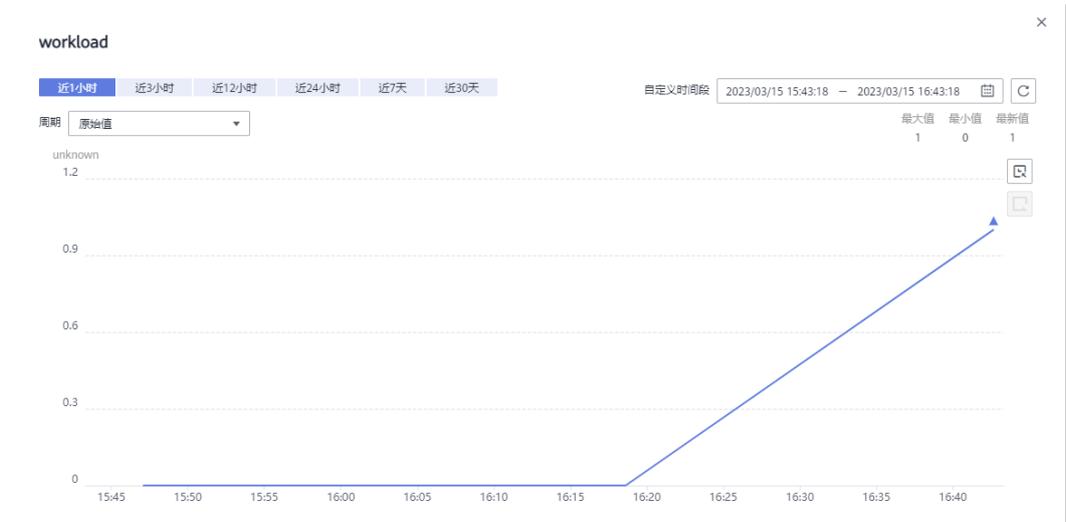
**步骤1** 提交Maya任务（仅供参考）。

图 4-1 任务示例



**步骤2** 查看[云监控告警状态](#)。当作业状态为“queued”时，代表作业job在排队中，当前集群计算节点资源不足。此时查看自定义监控中workload值>1，将通知弹性伸缩实例进行自动扩容。

图 4-2 workload 值变化状态



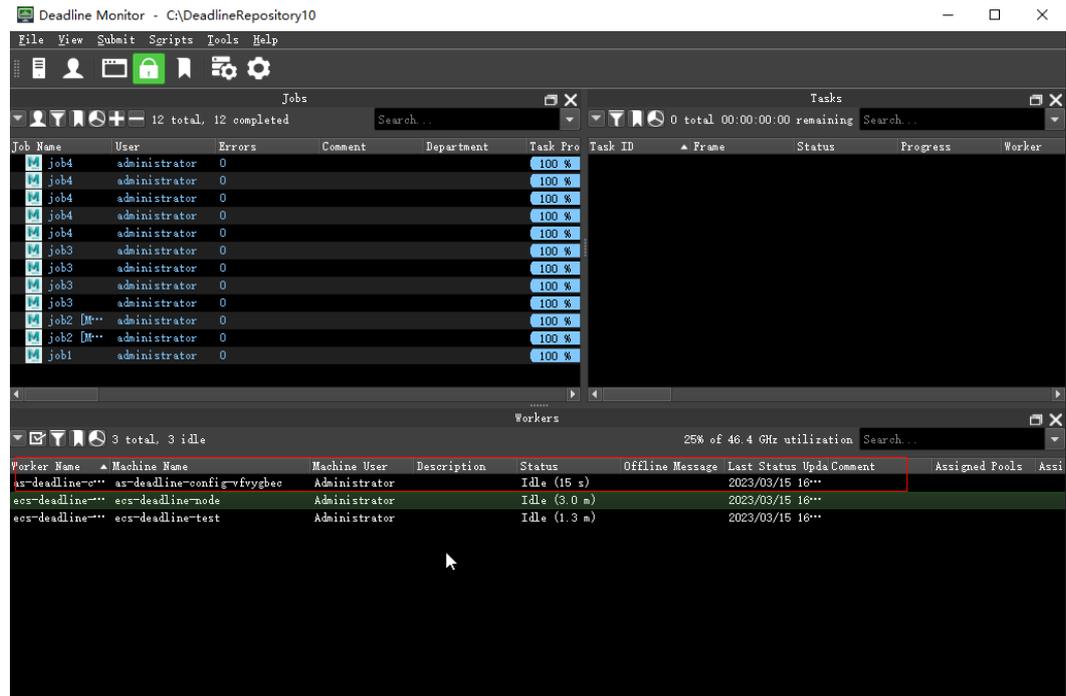
**步骤3** 登录[弹性伸缩AS控制台](#)，单击打开该方案创建的弹性伸缩组，查看弹性伸缩正在创建实例。

图 4-3 触发弹性扩容



**步骤4** 弹性扩容的ECS实例，使用固定资源池中计算节点镜像创建，登录密码与原服务器保持一致，登录创建好的实例，并运行Deadline程序即可被纳管进入Deadline集群，参与作业。

图 4-4 弹性扩容的 ECS 实例

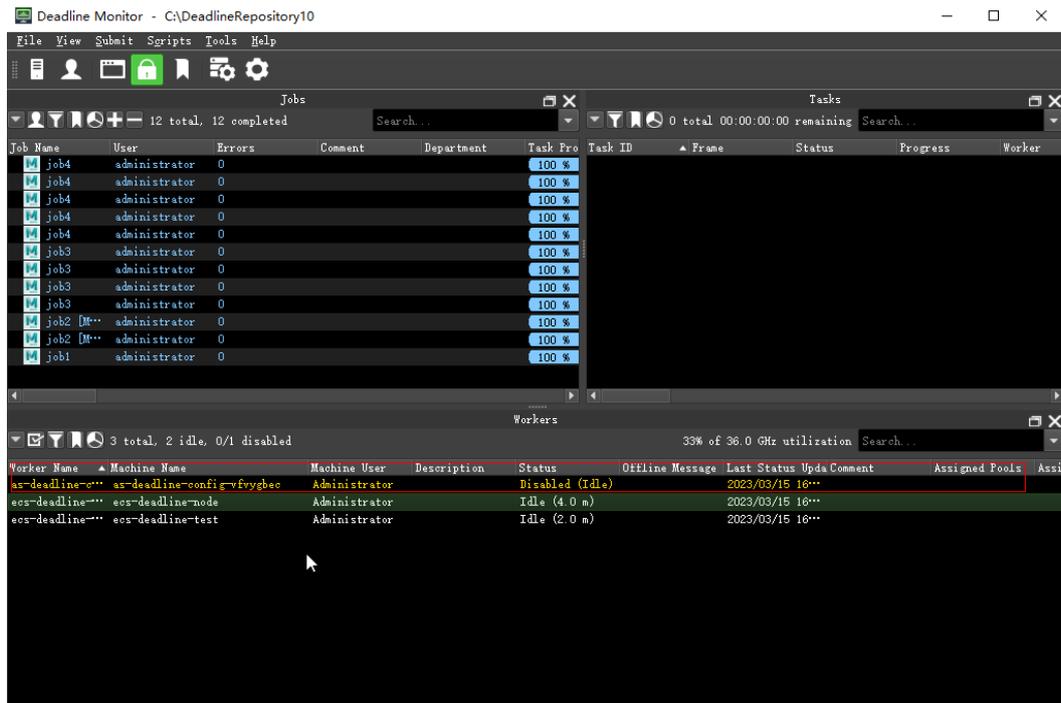


**步骤5** 被纳管进入Deadline集群的可用ECS计算节点实例，Deadline调度节点(master)将会下发作业任务给这些实例。

**步骤6** Gearbox将周期检查Deadline集群中计算节点node状态。检查空闲的node节点并node状态设置为DRAIN；检查集群中状态为DRAIN的node，是否还有任务正在该计算节点运行，如果没有，将该实例删除（成功后，node状态将变为DOWN）；检查集群中状态为DOWN或DRAIN\*的node，将node记录从Deadline集群移除。

（参数说明：drain(节点故障)，alloc(节点在用)，idle(节点可用)，down(节点下线)，mix(节点部分占用，但仍有剩余资源)

图 4-5 将空闲实例状态修改为 drain



步骤7 登录[弹性伸缩AS控制台](#)，单击打开该方案创建的弹性伸缩组，查看弹性伸缩已删除实例。

图 4-6 空闲实例被移出



----结束

## 名词解释

基本概念、云服务简介、专有名词解释

- 弹性云服务器ECS：是一种可随时自助获取、可弹性伸缩的云服务器，可帮助您打造可靠、安全、灵活、高效的应用环境，确保服务持久稳定运行，提升运维效率。
- 弹性文件服务SFS Turbo：为用户提供一个完全托管的共享文件存储，能够弹性伸缩至320TB规模，具备高可用性和持久性，为海量的小文件、低延迟高IOPS型应用提供有力支持。
- **Deadline**：Deadline是AWS下属的Thinkbox软件公司的一款渲染农场软件。
- **Gearbox**：Gearbox是一个华为云自研开源的资源协同系统。Gearbox系统与调度平台(Deadline)协同,协调云下、云上资源使用。支持自定义业务负载指标,提供基于指标的自动扩容能力,感知业务任务状态,闲置自动释放,提供无损的弹性伸缩能力。

# 5 修订记录

表 5-1 修订记录

发布日期	修订记录
2023-03-30	第一次正式发布。
2024-01-30	第二次发布