

解决方案实践

# 基于 Slurm 构建泛 HPC 场景云上云下 资源协同解决方案

文档版本 1.0.0  
发布日期 2023-11-30



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.....	27
<b>4 附录</b> .....	<b>30</b>
<b>5 修订记录</b> .....	<b>31</b>

# 1 方案概述

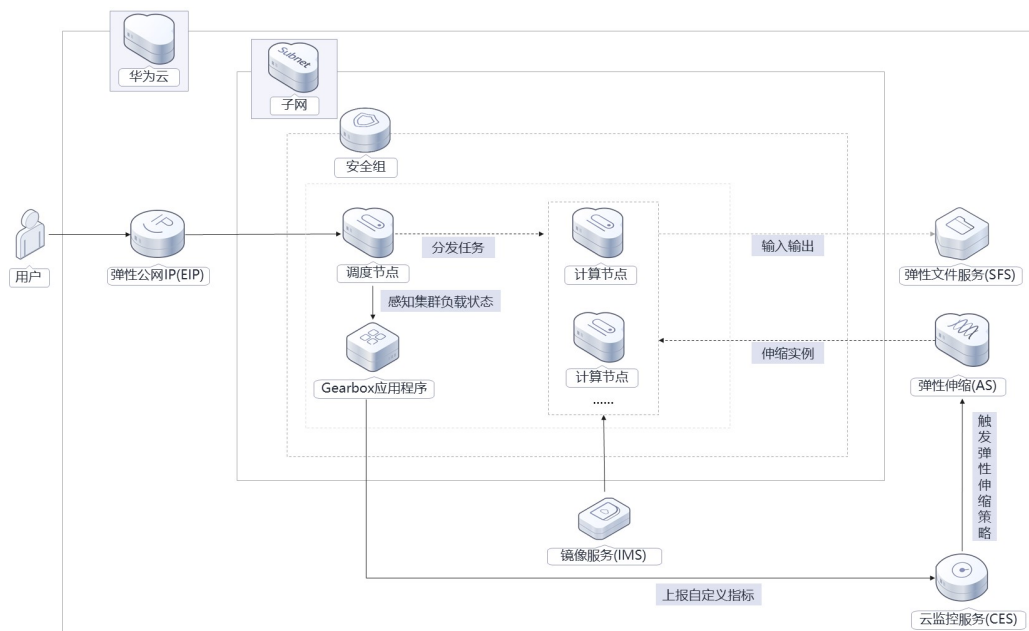
## 应用场景

该解决方案基于开源软件Slurm及华为云自研开源软件Gearbox，快速帮助用户在华为云上完成可自动弹性伸缩的HPC环境搭建。Slurm计算节点采用无配置模式，Gearbox程序与弹性伸缩 AS及云监控服务 CES对接，通过监控Slurm集群作业Job状态，实时自动弹性扩缩容Slurm集群环境；并自动完成弹性扩容出来的云服务器的注册，然后将其加入集群或从集群中注销并完成实例销毁。

## 方案架构

该解决方案部署架构如下图所示：

图 1-1 方案架构图



该解决方案将会部署如下资源：

- 创建2台Linux弹性云服务器，安装开源软件Slurm，在调度节点安装Gearbox程序、配置Java环境。

- 创建1条弹性公网IP，用于提供访问公网和被公网访问能力。
- 创建安全组，可以保护弹性云服务器的网络安全，通过配置安全组规则，限定云服务器的访问端口。
- 使用镜像服务IMS，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。
- 使用弹性伸缩AS，创建1个弹性伸缩组，通过设置弹性伸缩配置及伸缩策略来进行集群实例资源的弹性扩缩容。
- 使用云监控服务，Gearbox程序监测集群作业状态，计算自定义指标workload值，上报指标到云监控服务。
- 使用弹性文件服务SFS，挂载到所有弹性云服务器上，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HPC集群环境部署。

## 约束与限制

- 该解决方案部署前，需注册华为账号并开通华为云，完成实名认证。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。
- 请确保在部署解决方案前，使用的华为云账号有IAM的足够权限，具体请参考[3.1-创建rf\\_admin\\_trust委托](#)。
- 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，不同产品的花费仅供参考，具体请参考华为云官网[价格详情](#)，实际收费以账单为准。

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 ( )	<ul style="list-style-type: none"><li>● 区域：华东-上海一</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：CentOS 7.9 64bit</li><li>● 系统盘：高IO   100GB</li><li>● 购买量：1</li></ul>	389.00 元
弹性云服务器 ECS	<ul style="list-style-type: none"><li>● 区域：华东-上海一</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：CentOS 7.9 64bit</li><li>● 系统盘：高IO   100GB</li><li>● 购买量：1</li></ul>	2867.00 元
SFS Turbo标准型	<ul style="list-style-type: none"><li>● 区域：华东-上海一</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 标准型   500GB</li></ul>	550.00 元

华为云服务	配置示例	每月预估花费
弹性公网IP	<ul style="list-style-type: none"> <li>区域：华东-上海一</li> <li>计费模式：包年包月</li> <li>线路：动态BGP</li> <li>公网带宽：按带宽计费</li> <li>带宽大小：5Mbit/s</li> <li>购买量：1</li> </ul>	115.00 元
<b>合计</b>		<b>3921.00 元</b>

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	<ul style="list-style-type: none"> <li>按需计费：0.79元/小时</li> <li>区域：华东-上海一</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：CentOS 7.9 64bit</li> <li>系统盘：高IO   100GB</li> <li>购买量：1</li> </ul>	568.80 元
弹性云服务器 ECS	<ul style="list-style-type: none"> <li>按需计费：5.95元/小时</li> <li>区域：华东-上海一</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：CentOS 7.9 64bit</li> <li>系统盘：高IO   100GB</li> <li>购买量：1</li> </ul>	4284.00 元
SFS Turbo标准型	<ul style="list-style-type: none"> <li>按需计费：0.32元/小时</li> <li>区域：华东-上海一</li> <li>计费模式：按需计费</li> <li>SFS Turbo 标准型   500GB</li> </ul>	230.40 元



华为云服务	配置示例	每月预估花费
弹性公网IP	<ul style="list-style-type: none"><li>● 按需计费：0.34元/小时</li><li>● 区域：华东-上海一</li><li>● 计费模式：按需计费</li><li>● 线路：动态BGP</li><li>● 公网带宽：按带宽计费</li><li>● 带宽大小：5Mbit/s</li><li>● 购买量：2</li></ul>	244.80 元
<b>合计</b>		<b>5328.00 元</b>

# 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”，委托类型选择“云服务”，选择“RFS”，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，单击“下一步”。

图 3-5 选择策略



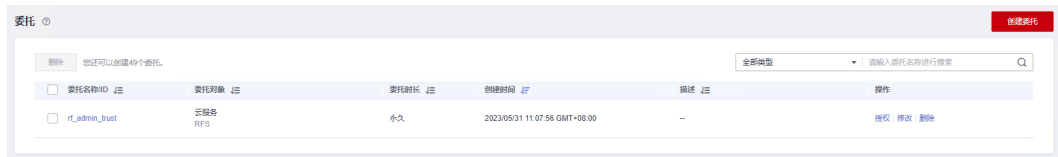
步骤5 选择“所有资源”，并单击下一步完成配置。

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



----结束

## 创建访问密钥

部署该方案之前，需要您在华为云控制台配置访问密钥对（AK、SK），并创建用于登录弹性云服务器账号的账号密钥对。

- 步骤1** 根据[官方文档](#)指引，在[访问密钥](#)新增访问密钥，并下载秘密访问密钥，密钥将用于启动Gearbox程序的配置项。（该解决方案部署成功后，建议删除该AKSK，并重新配置）。

图 3-8 创建 AK，SK



- 步骤2** 根据[官方文档](#)指引，在[密钥对管理](#)创建账号密钥对，在后续[快速部署](#)中使用，作为弹性扩容出来的ECS实例的登录密钥。

图 3-9 创建密钥对



图 3-10 密钥对成功创建



----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 参数说明

参数名称	类型	是否可选	参数解释	默认值
vpc_name	string	必填	虚拟私有云名称，该模板新建VPC，不允许重名。取值范围：1-54个字符，支持数字、字母、_(下划线)、-(中划线)、.(点)	scalable-hpc-cluster-with-slurm-demo
security_group_name	string	必填	安全组名称，该模板新建安全组。取值范围：1-64个字符，支持数字、字母、_(下划线)、-(中划线)、.(点)	scalable-hpc-cluster-with-slurm-demo

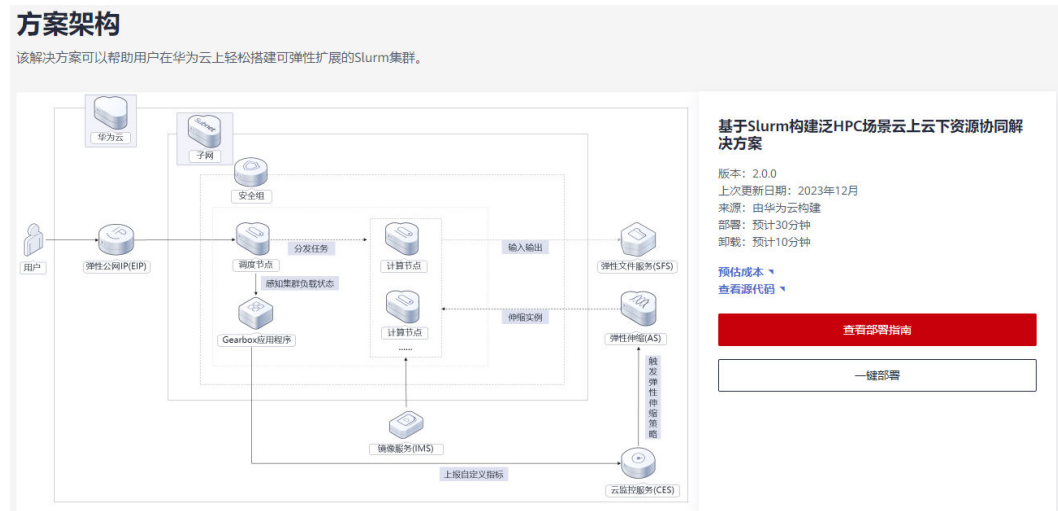
参数名称	类型	是否可选	参数解释	默认值
bandwidth_size	number	必填	带宽大小，取值范围：1-2,000Mbit/s。	5
ecs_name	string	必填	弹性云服务器名称，命名方式为{ecs_name}-master、{ecs_name}-node，不支持重名。取值范围：1-57个字符组成，包括小写字母、数字、连字符（-）。	scalable-hpc-cluster-with-slurm-demo
ecs_password	string	必填	弹性云服务器初始化密码，创建完成后，请参考部署指南登录ECS控制台修改密码。取值范围：长度为8-26个字符，密码至少包含大写字母、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_node_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
keypair_name	string	必填	已有SSH登录密钥对名称，当前只支持密钥对管理中的账号密钥对，请参考 <a href="#">密钥对管理</a> 获取。	空
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。	scalable-hpc-cluster-with-slurm-demo
sfs_turbo_size	number	必填	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：500-32,768。	500

参数名称	类型	是否可选	参数解释	默认值
as_name	string	必填	弹性伸缩资源名称，该模板使用新建弹性伸缩组，为集群提供弹性伸缩功能，不支持重名。取值范围：1-50 个字符组成，只能由英文字母、数字、中划线（-）组成。	scalable-hpc-cluster-with-slurm-demo
charge_mode	string	必填	计费模式，默认自动扣费，取值为prePaid（包年包月）或postPaid（按需计费）。	postPaid
charge_period_unit	string	必填	计费周期单位，当计费方式设置为prePaid时，此参数为必填项；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
charge_period	number	必填	计费周期，当计费方式设置为prePaid时，此参数为必填项；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
AK	string	必填	用户创建的访问密钥，用于配置gearbox软件，请参考 <a href="#">创建访问密钥步骤1</a> 创建并获取密钥。	空
SK	string	必填	用户创建的访问密钥，用于配置gearbox软件，请参考 <a href="#">创建访问密钥步骤1</a> 创建并获取密钥。	空
project_id	string	必填	用户方案部署所在区域的项目ID，请参考 <a href="#">统一身份认证--项目</a> 获取。	空

**步骤1** 登录华为云解决方案实践，选择“[基于Slurm构建泛HPC场景云上云下资源协同方案](#)”。

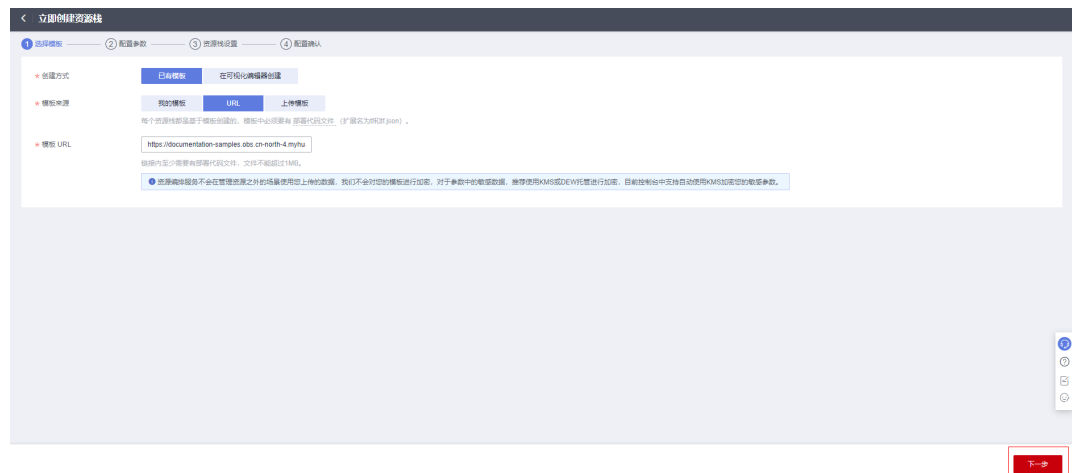


图 3-11 解决方案实施库



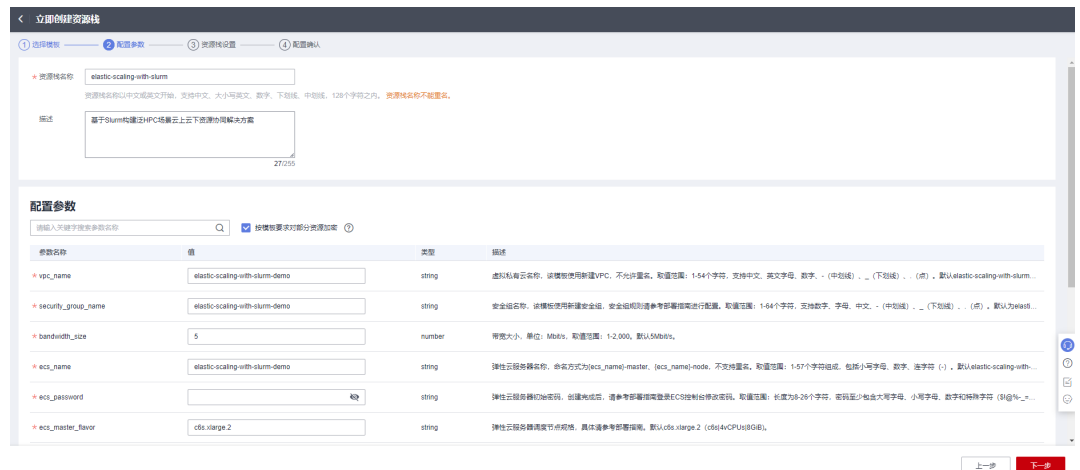
步骤2 在一键部署页面单击“一键部署”，跳转至该解决方案创建资源栈部署界面。

图 3-12 一键部署



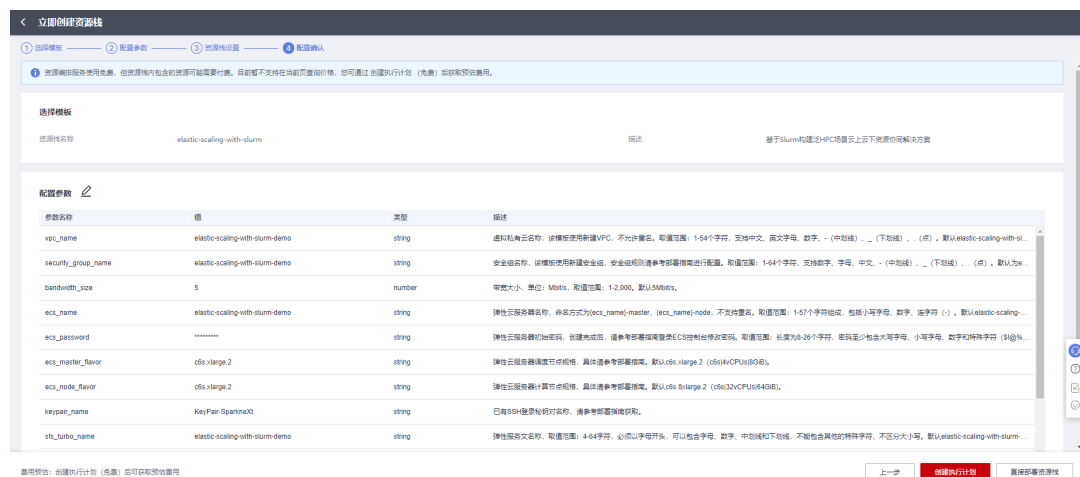
步骤3 单击“下一步”，参考表 3-1 完成自定义参数填写。

图 3-13 参数配置



步骤4 单击“下一步”，创建执行计划。

图 3-14 创建执行计划



步骤5 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



步骤6 待执行计划状态创建成功，单击“部署”，弹出页面提示等信息，单击“执行”。

图 3-16 部署

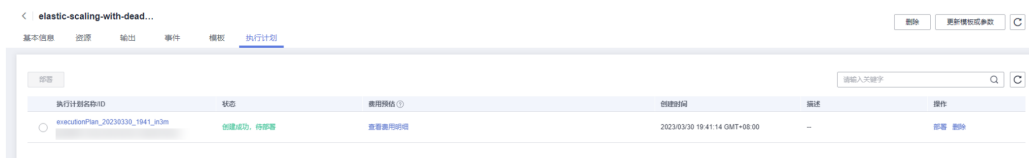


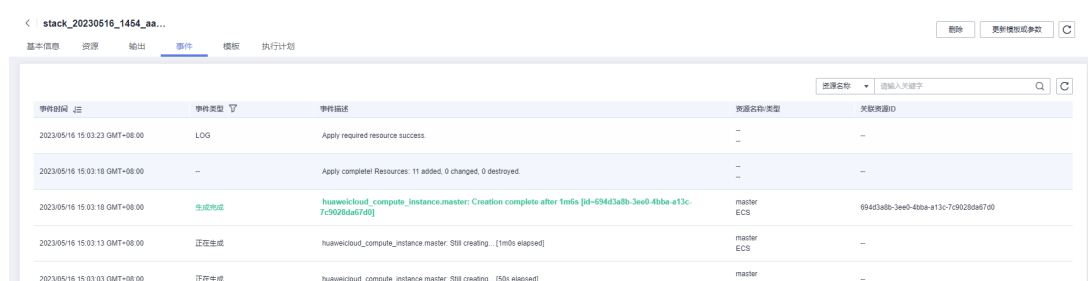
图 3-17 执行计划确认



**步骤7**（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8** 部署成功后，回显结果如下图。等待30分钟，集群内环境部署完成即可使用。

图 3-18 执行计划部署完成



注：集群内环境部署完成前请不要登录ECS进行其他操作，否则可能导致环境部署失败。

----结束

## 3.3 开始使用

本节将展开Slurm的基础使用说明，有关Slurm详细使用指导请参考[Slurm产品文档](#)。

该解决方案涉及的所有的业务端口均默认配置了同一子网可访问，如果另有需要，需根据实际情况修改安全组规则。涉及到的Slurm业务端口如下：

- 6817：Slurm调度节点的业务面端口
- 6818：Slurm计算节点的业务面端口

## 📖 说明

- 该解决方案默认使用的Slurm版本：22.05.3，部署的Java版本：jdk1.8.0\_202。
- 已在Slurm内默认创建了名为**cluster**的集群名。

## 安全组修改（可选）

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。

## 查看节点状态

- 步骤1** 在调度节点上查看Slurm软件运行状态，执行systemctl status munge命令和systemctl status slurmctld命令查看Munge服务和Slurmctld服务是否运行正常。

图 3-19 Munge 服务正常运行

```
[root@master ~]# systemctl status munge
● munge.service - MUNGE authentication service
   Loaded: loaded (/usr/lib/systemd/system/munge.service; enabled; vendor preset: disabled)
   Active: active (running) since Wed 2023-04-26 11:40:35 CST; 2 weeks 0 days ago
     Docs: man:munged(8)
   Process: 12632 ExecStart=/usr/sbin/munged (code=exited, status=0/SUCCESS)
   Main PID: 12634 (munged)
    Tasks: 4
   Memory: 1.3M
   CGroup: /system.slice/munge.service
           └─12634 /usr/sbin/munged

Apr 26 11:40:35 master systemd[1]: Starting MUNGE authentication service...
Apr 26 11:40:35 master systemd[1]: Started MUNGE authentication service.
[root@master ~]#
```

图 3-20 Slurmctld 服务正常运行

```
[root@master ~]# systemctl status slurmctld
● slurmctld.service - Slurm controller daemon
   Loaded: loaded (/usr/lib/systemd/system/slurmctld.service; disabled; vendor preset: disabled)
   Active: active (running) since Wed 2023-04-26 15:01:22 CST; 1 weeks 6 days ago
   Main PID: 12387 (slurmctld)
    Tasks: 17
   Memory: 16.7M
   CGroup: /system.slice/slurmctld.service
           └─12387 /usr/sbin/slurmctld -D -s
             └─12388 slurmctld: slurmscriptd

May 10 11:18:09 master slurmctld[12387]: slurmctld: _job_complete: JobId=3063_49(3112) done
May 10 11:18:14 master slurmctld[12387]: slurmctld: _job_complete: JobId=3063_50(3063) WEXITSTATUS 0
May 10 11:18:14 master slurmctld[12387]: slurmctld: _job_complete: JobId=3063_50(3063) done
May 10 11:18:22 master slurmctld[12387]: slurmctld: update_node: node as-slurm-node-copy-4kwk9x4w reason set to: scale in
May 10 11:18:22 master slurmctld[12387]: slurmctld: update_node: node as-slurm-node-copy-4kwk9x4w state set to DRAINED
May 10 11:25:19 master slurmctld[12387]: slurmctld: error: find_node_record: lookup failure for node "as-slurm-node-copy-4kwk9x4w"
May 10 11:25:19 master slurmctld[12387]: slurmctld: error: Unable to find node as-slurm-node-copy-4kwk9x4w to delete
May 10 11:25:19 master slurmctld[12387]: slurmctld: error: failed to delete node 'as-slurm-node-copy-4kwk9x4w'
May 10 11:25:19 master slurmctld[12387]: slurmctld: slurm_rpc delete_node for as-slurm-node-copy-4kwk9x4w: Invalid node n
May 10 11:30:59 master slurmctld[12387]: slurmctld: error: Nodes as-slurm-node-copy-4kwk9x4w not responding, setting DOWN
[root@master ~]#
```

如果Munge没有启动，执行：

```
systemctl start munge
```

如果Slurmctld没有启动，执行：

```
systemctl start slurmctld  
#查看集群内节点状态  
sinfo
```

**步骤2** 每次调度节点服务器重启后，需登录调度节点云服务器，输入用户名root以及密码进入云服务器，输入以上命令启动服务。输入命令“sinfo”查看集群节点信息。

图 3-21 Slurm 配置成功

```
[root@master ~]# sinfo  
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST  
p1*        up    infinite    1    idle node1  
[root@master ~]# _
```

如果“NODES”下无节点，可登录到计算节点上开启Slurmd服务。命令如下：

```
#开启node1节点slurmd服务  
ssh 192.168.0.2  
systemctl start slurmd  
exit
```

**步骤3** 在调度节点上查看Slurm服务及Gearbox程序运行状态。

进入Gearbox程序包所在目录下（默认为/usr/local/），启动Gearbox程序：

```
cd /usr/local/  
nohup java -jar gearbox-0.0.1-SNAPSHOT.jar --spring.config.name=gearboxConfig > /dev/null 2>&1 &
```

**步骤4** 在调度节点上查看Gearbox运行状态，执行ps aux|grep gearbox查看Gearbox服务是否正常运行。

```
ps aux|grep gearbox
```

图 3-22 Gearbox 服务正常运行

```
[root@master ~]# ps aux|grep gearbox  
root      4492  0.0  0.0 112812  980 pts/1    R+   14:40   0:00 grep --color=auto gearbox  
root      16314 0.2  4.0 5695736 324216 pts/1    SL   May08   6:10 java -jar gearbox-0.0.1-SNAPSHOT.jar --spring.config.name=Config  
[root@master ~]#
```

**步骤5** 设置Slurm root用户配额，此配额与弹性伸缩的ECS数量相关。

```
#修改配额  
sacctmgr modify user root set GrpTRES="node={用户所需的配额数量}"  
#查看配额信息  
sacctmgr list ass
```

注：Gearbox启动后会将没有计算任务的节点设为“drain”状态，使用时要重新激活，命令如下：

```
scontrol update NodeName={NODELIST下节点名称} State=idle
```

图 3-23 节点状态为 drain

```
[root@master local]# sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
p1*        up       infinite    1     drain node1
[root@master local]# _
```

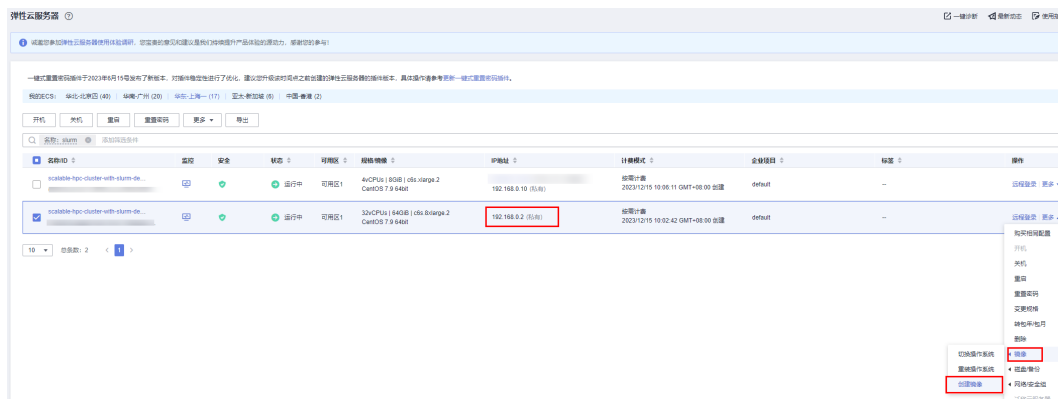
---结束

## 创建计算节点镜像

步骤1 确认调度集群正常后，登录[弹性云服务器](#)。

步骤2 找到计算节点（私有IP为192.168.0.2），依次单击右侧“更多>镜像>创建镜像”。

图 3-24 创建镜像



步骤3 镜像类型选择系统盘镜像，配置信息中名称可自行命名，本示例中名称为hpc-image。

图 3-25 选择系统盘镜像

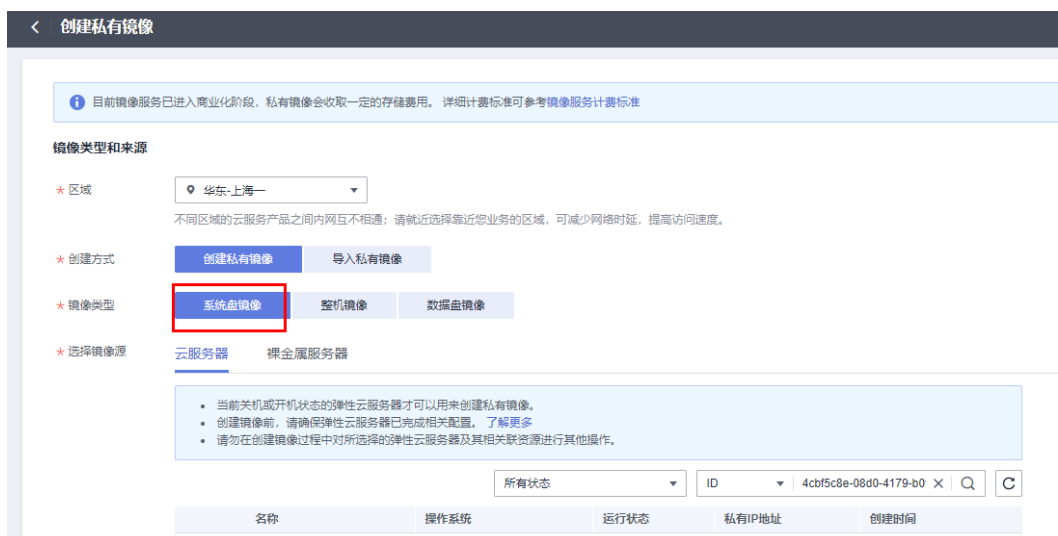


图 3-26 配置信息填写

配置信息

加密 未加密 ⓘ

\* 名称

\* 企业项目  ⓘ

标签 如果您需要使用同一标签标识多种云资源，即所有服务均可在标签输入框下拉选择同一标签。建议在TMS中创建预定义标签。 [查看预定义标签](#) ⓘ

您还可以添加10个标签。

描述

0/1,024

协议  我已经阅读并同意《镜像制作承诺书》和《镜像免责声明》

步骤4 单击“立即创建”后跳转到**镜像服务**控制台，可查看镜像创建进度。

图 3-27 镜像创建中

云服务器控制台

镜像服务 ⓘ

目前镜像服务已进入商业化阶段，私有镜像会收取一定的存储费用，删除创建的镜像后将不再计费。详细计费标准可参考**镜像服务**

公共镜像 私有镜像 共享镜像

私有镜像包括系统盘镜像、数据盘镜像和整机镜像，由现有运行的云服务器创建而来，或由外部导入而来。 [了解更多](#)

镜像支持云服务器快速发放，建议您优化不支持该功能的镜像。请在详情页面查看镜像是否支持快速发放。 [了解更多](#)

您还可以创建39个私有镜像。

Q 默认按照关键字搜索、过滤

名称/ID	状态	操作系统类型
<input type="checkbox"/> hpc-image	<input checked="" type="checkbox"/> 创建中 <div style="width: 20%;"></div>	Linux



图 3-28 镜像创建完成



----结束

## 配置弹性伸缩

- 步骤1** 打开华为云**镜像服务**控制台，确认镜像创建成功。
- 步骤2** 登录华为云弹性伸缩AS控制台，选择该方案创建的**弹性伸缩配置**，单击右侧“复制”。在弹窗中，选择“私有镜像”，选择**创建计算节点镜像**中创建的镜像，注意修改名称（名称不能带有下划线“\_”，否则无法自动删除实例），单击“确定”创建新的弹性伸缩配置。

图 3-29 创建新的弹性伸缩配置



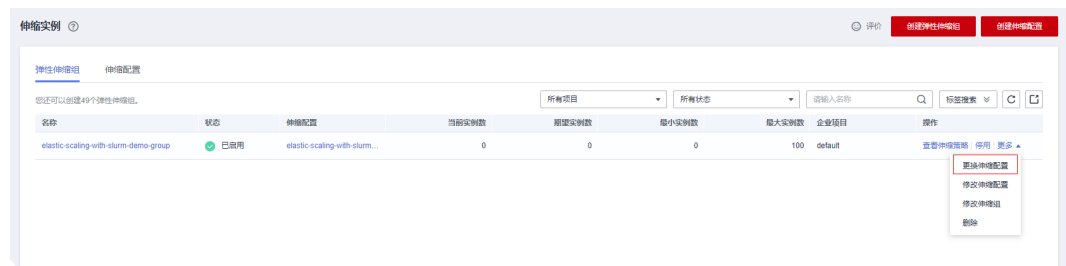
- 步骤3** 选择“密钥对”登录方式，本示例中密钥对名称为HPC-TEST（密钥对名称以3.1准备工作中创建的密钥对名称为准），并勾选相应选项。

图 3-30 设置登录方式



**步骤4** 打开**弹性伸缩组**列表, 选择该方案创建的弹性伸缩组, 单击右侧“更多>更换弹性伸缩配置”, 在弹窗中选择上一步骤创建的弹性伸缩配置, 单击“确定”。

图 3-31 更换弹性伸缩配置



**步骤5** 登录华为云弹性伸缩AS控制台, 查看该方案创建的弹性伸缩组, 单击“查看伸缩策略”。

图 3-32 弹性伸缩实例



**步骤6** 单击“添加伸缩策略”, 选择“告警策略”。根据业务需求配置相应的区间范围(1条伸缩策略最大支持添加10个策略区间), 配置完成后单击“确定”。

图 3-33 添加伸缩策略



**步骤7** 当策略区间增加的服务器数量不满足业务需要时，可以添加多个伸缩策略，单击“添加伸缩策略”。

图 3-34 添加策略



**步骤8** 安装业务软件，安装业务需要的软件到共享文件目录/data下面。

Slurm作业管理系统：

- Slurmd、Slurmdbd、Slurmctld日志文件目录：/var/log
  - Slurmdbd服务的配置文件：/etc/slurm/slurmdbd.conf
  - Slurmctld、Slurmd服务的配置文件：/etc/slurm/slurm.conf
- 详细使用参考[Slurm产品文档](#)，常用命令如下：

表 3-2 常用命令

命令	功能介绍	常用命令举例
sinfo	显示系统资源使用情况	sinfo
squeue	显示作业状态	squeue
srun	用于交互式作业提交	srun -N 2 -n 48 -p debug A.exe

命令	功能介绍	常用命令举例
sbatch	用于批处理作业提交	sbacth -n -2 -n 48 job.sh
salloc	用于分配模式作业提交	salloc -p debug
scancel	用于取消已提交的作业	scancel JOBID
scontrol	用于查询节点信息或正在运行的作业信息	scontrol show job JOBID
sacct	用于查看历史作业信息	sacct -u pp100 -S 03/01/17 -E 03/31/17 --field=joid,partitionjobname,user,nnodes,start,end,elapsed,state

----结束

## Gearbox 程序工作流程介绍

本文档以批任务为示例，演示Gearbox程序如何自动进行Slurm集群弹性扩缩容。

**步骤1** Gearbox程序启动后，设置Slurm root用户配额。

```
#修改配额
sacctmgr modify user root set GrpTRES="node={用户所需的配额数量}"
#查看配额信息
sacctmgr list ass
```

图 3-35 查看用户配额信息

```
[root@master ~]# sacctmgr list ass
Cluster Account User Partition Share Priority GrpJobs GrpTRES GrpSubmit GrpWall GrpTRESMins MaxJobs MaxTRES MaxTRESPerNode Ma
xSubmit MaxWall MaxTRESMins
-----
cluster root 1
cluster root root normal 1
cluster root root normal 1 node=2
[root@master ~]#
```

**步骤2** 提交批处理任务（仅供参考）。

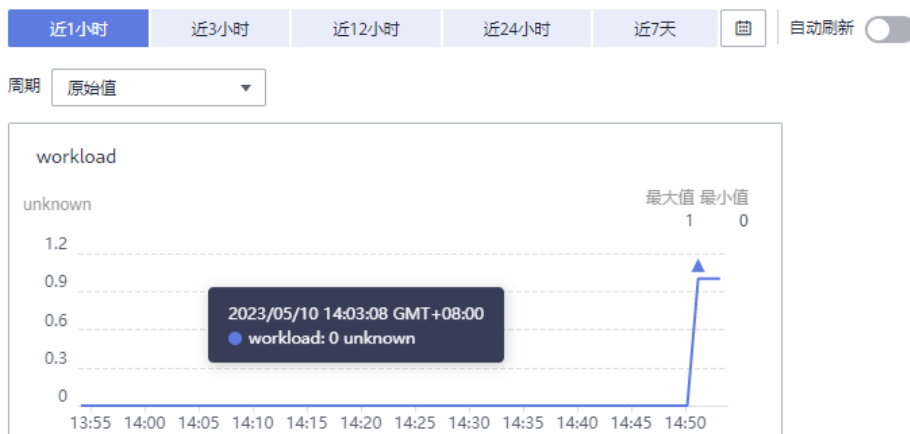
图 3-36 sbatch 提交 slurm 任务

```
[root@master ~]# cat myscript
#!/bin/sh
#SBATCH --job-name=ty
#SBATCH --array=1-50
#SBATCH -p p1
#SBATCH --output=/root/log/slurm-%A_%a.out

echo "job start..."
sleep 10
echo "job finished.."
[root@master ~]# sbatch myscript
Submitted batch job 3113
[root@master ~]#
```

**步骤3** 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 3-37 workload 值变化状态



**步骤4** 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 3-38 触发弹性扩容



**步骤5** 弹性扩容出来的ECS实例，使用的镜像为计算节点服务器镜像。登录方式为[3.3步骤1](#)创建的密钥对。登录该ECS实例，并运行Slurmd程序即可被纳管进入Slurm集群，参与作业。

图 3-39 弹性扩容的 ECS 实例



被纳入Slurm集群的可用ECS计算节点实例，Slurm调度节点将会下发作业任务给这些实例。

图 3-40 扩容出的实例执行作业任务

```
Every 1.0s: queue
      JOBID PARTITION   NAME   USER  ST      TIME  NODES NODELIST(REASON)
      3113_[31-50]    p1     ty    root  PD      0:00     1 (Resources)
      3113_29        p1     ty    root  R       0:02     1 as-slurm-node-copy-esiu6yqp
      3113_30        p1     ty    root  R       0:02     1 master
```

**步骤6** Gearbox将周期检查Slurm集群中计算节点状态。检查到有空闲的计算节点，将状态设置为“drain”；检查集群中状态为“drain”的计算节点，判断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，计算节点状态将变为“down”）；检查集群中状态为“down”或“drain”的计算节点，将该计算节点从Slurm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 3-41 将空闲实例状态修改为 drain

```
Every 1.0s: sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
p1         up       infinite    1     drain as-slurm-node-copy-esiu6yqp
p1         up       infinite    1     idle  master
p2*        up       infinite    0     n/a
```

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 3-42 空闲实例被移出



----结束

## gearbox 配置文件示例

下面内容为Gearbox配置文件示例，请存放在Gearbox安装目录下。配置 gearboxconfig.yaml文件时，请删除示例中的所有注释（即“#”所在行）。

```
user:
# console账号的AK
ak:
# console账号的SK
sk:
# 用户所在region的项目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-east-3.myhuaweicloud.com
# 预置伸缩组资源的伸缩组ID
group:
# 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-east-3.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:
# 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-east-3.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, 单位 秒
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
system:
# deadline/slurm
type: slurm
# 稳定节点的Node名称, 多节点英文逗号分割
stable-nodes: master
# 稳定节点所在的分区
stable-partition: p1
# 不稳定节点所在的分区
variable-partition: p1
# 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
# 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
```

## 3.4 快速卸载

**步骤1** 登录[资源编排服务RFS](#)资源栈, 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 3-43 解决方案卸载



**步骤2** 创建的计算节点镜像资源不会被资源栈删除，请进入镜像服务手动删除。



图 3-44 删除计算节点镜像



步骤3 创建的伸缩配置资源不会被资源栈删除，请进入伸缩配置手动删除。

图 3-45 删除伸缩配置



----结束

# 4 附录

## 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。
- 弹性公网IP：提供独立的公网IP资源，包括公网IP地址与公网出口带宽服务。可以与弹性云服务器、裸金属服务器、虚拟IP、弹性负载均衡、NAT网关等资源灵活地绑定及解绑。
- 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。
- Slurm：是一个开源，高度可扩展的集群管理工具和作业调度系统，用于各种规模的Linux集群。主要提供如下集中关键的特性：
  - 资源分配  
分配独占或者非独占的资源给用户，可以控制分配的时长，供用户运行作业。
  - 作业管理框架  
提供一个框架，可以帮助用户控制并行作业在所分配资源上的启动、运行和监控。
  - 队列  
提交的作业资源需求超出了可用资源，将作业放入队列。
  - 不同的作业调度策略  
提供资源预留，公平分享，回填等高级作业调度策略供使用。
  - 其他工具  
提供作业信息统计，作业状态诊断等工具。
- 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(Slurm)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
2023-05-30	第一次正式发布。
2023-11-30	修订方案内容及实施步骤。