

高性能计算平台 操作指南



腾讯云

【 版权声明 】

©2013–2024 腾讯云版权所有

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

【 商标声明 】



及其它腾讯云服务相关的商标均为腾讯云计算（北京）有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标，依法由权利人所有。未经腾讯云及有关权利人书面许可，任何主体不得以任何方式对前述商标进行使用、复制、修改、传播、抄录等行为，否则将构成对腾讯云及有关权利人商标权的侵犯，腾讯云将依法采取措施追究法律责任。

【 服务声明 】

本文档意在向您介绍腾讯云全部或部分产品、服务的当时的相关概况，部分产品、服务的内容可能不时有所调整。您所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定，除非双方另有约定，否则，腾讯云对本文档内容不做任何明示或默示的承诺或保证。

【 联系我们 】

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

文档目录

操作指南

访问管理

概述

CAM 服务角色授权

存储资源管理

概述

存储选项填写说明

集群初始化脚本配置

集群扩缩容策略说明

使用自定义镜像加速扩容

操作指南

访问管理

概述

最近更新时间：2023-07-14 17:10:51

访问管理（Cloud Access Management，CAM）是腾讯云提供的 Web 服务，主要用于帮助用户对腾讯云账户下资源的访问权限的安全管理。您可以通过访问管理创建、管理和销毁用户或用户组，并使用身份管理和策略管理控制其他用户使用腾讯云资源的权限。策略能够授权或者拒绝用户使用指定资源完成指定任务，当您在使用访问管理时，可以将策略与一个用户或一组用户关联起来进行权限控制。

高性能计算平台已接入访问管理，您可以对高性能计算平台的相关资源进行权限控制。

相关概念

访问管理用户

访问管理 **用户** 是您在腾讯云中创建的一个实体，每一个访问管理用户仅同一个腾讯云账户关联。您注册的腾讯云账号身份为**主账号**，您可以通过 **用户管理** 来创建拥有不同权限的**子账号**进行协作。子账号的类型分为 **子用户**、**协作者** 以及 **消息接收人**。

策略

策略 是用于定义和描述一条或多条权限的语法规则，腾讯云的策略类型分为预设策略和自定义策略。

- **预设策略**：由腾讯云创建和管理的策略，是被用户高频使用的一些常见权限集合，如资源全读写权限等。预设策略操作对象范围广，操作粒度粗，且为系统预设，不可被用户编辑。
- **自定义策略**：由用户创建的策略，允许进行细粒度的权限划分。例如，为子账号关联一条使用策略，使其有权管理高性能计算平台的计算环境，而无权管理云数据库实例。

资源

资源（resource） 是策略的元素，描述一个或多个操作对象，例如高性能计算平台的集群。

高性能计算平台预设策略介绍

预设策略名	授权范围描述
QcloudTHPCFullAccess	高性能计算平台（THPC）全读写访问权限

可授权的资源类型

资源级权限指的是能够指定用户对哪些资源具有执行操作的能力。

在访问管理中对高性能计算平台可授权的资源类型如下：

资源类型	授权策略中的资源描述方法
cluster	qcs::thpc::uin/\${uin}:cluster/\${ClusterId}

高性能计算平台 API 接口支持资源级授权情况如下：

支持资源级授权的 API 列表

下表列出高性能计算平台支持资源级权限操作的 API，以及操作支持的资源路径。

- 设置资源路径时，您需要将 `${uin}`、`${ClusterId}` 等变量参数修改为您实际的参数信息，同时您也可以使用 `*` 通配符。
- 访问管理策略中的 `region`、`action`、`account`、`resource` 等相关概念请参见 [元素参考](#)。

API 接口: action	API 描述	资源路径: resource
DeleteCluster	删除集群	qcs::thpc::uin/\${uin}:cluster/\${ClusterId}

不支持资源级授权的 API 列表

⚠ 注意

下表中 API 操作不支持资源级权限。针对不支持资源级权限的高性能计算平台 API 操作，您仍可以向用户授予使用该操作的权限，但策略语句的资源元素必须指定为 `*`。

API 接口: action	API 描述
CreateCluster	创建集群

高性能计算平台访问管理策略示例

以开放高性能计算平台（THPC）全读写访问权限为例，策略如下：

```
{
  "version": "2.0",
  "statement": [
    {
      "action": "thpc:*",
```

```
"resource": "*",  
"effect": "allow"  
}  
]  
}
```

CAM 服务角色授权

最近更新时间：2023-07-14 17:10:51

操作场景

在使用高性能计算平台（TencentCloud High Performance Computing, THPC）的过程中，涉及使用相关云资源。为了能正常使用高性能计算平台的功能，您需要对高性能计算平台的服务角色进行授权，授权后才能使用相关云资源。

当您首次使用高性能计算平台时，需前往[访问管理](#)页面为当前账号授予腾讯云服务网格操作云服务器 CVM、弹性伸缩 AS、自动化助手 TAT 等云资源的权限。该权限授予通过关联预设策略

`QcloudAccessForTHPCLinkedRoleInAcrossService` 至服务角色 `THPC_QCSLinkedRoleInAcrossService` 完成。本文介绍如何通过访问管理控制台，完成授权操作。

操作步骤

1. 前往 [THPC_QCSLinkedRoleInAcrossService](#) 角色页面，并单击**同意授权**。如下图所示：



2. 完成身份验证后，即可成功授权。

存储资源管理

概述

最近更新时间：2023-09-18 10:29:01

创建 THPC 集群之后，可以给集群挂载额外的 CFS 或者 GooseFS 等文件系统，以提供不同场景下的存储能力。本文主要介绍集群存储选项的使用方法。

创建存储资源

在使用 THPC 存储挂载功能前，需要提前创建好存储资源。目前 THPC 支持的 CFS 和 GooseFS 文件系统（创建的存储资源必须与待挂载的 THPC 集群在同一 VPC 下）。

- 挂载 CFS：需开通 CFS 服务，创建 CFS 文件系统。具体操作请参见 [创建文件系统](#)。
- 挂载 CFS-TURBO：需开通 CFS-TURBO 服务，创建 TURBO 文件系统。具体操作请参见 [创建文件系统](#)。
- 挂载 GooseFS：需使用自建集群部署配置 GooseFS 服务。具体操作请参见 [自建集群部署 GooseFS](#)。

挂载存储选项

挂载存储选项有两种方式。

- 创建集群（[CreateCluster](#)）时，填写存储选项（[StorageOption](#)）参数，会为集群中所有节点挂载上存储资源。
- 集群创建完毕并且为运行中（RUNNING）状态的集群，通过添加集群存储选项（[AddClusterStorageOption](#)）接口进行添加集群存储选项，添加完之后，THPC 会尝试对当前集群存在的所有节点挂载上存储资源。后续通过 THPC 扩容出来的实例也会挂载上存储选项的资源。

[AddClusterStorageOption](#) 接口输入参数示例：

① 说明

- 如果本地 /opt/ 挂载点没有挂载文件系统，THPC 在集群创建过程中会自动创建一个标准型的 CFS 挂载到 /opt/ 目录下，用于存储 THPC 集群的调度器配置、THPC 集群相关依赖等。
- 如果 LocalPath 本地目录不存在，THPC 会自动创建此目录。
- 如果 RemotePath 远程目录不存在，THPC 会在 CFS 中为您创建子目录，CFS 支持 [子目录挂载](#)；注意 CFS-TURBO 不支持子目录挂载。
- 更多参数说明请参见 [CFSOptions 数据结构](#)。

```
{  
  "ClusterId": "hpc-rv7hahw2",
```



```

"StorageOption": {
  "CFSOptions": [{
    "LocalPath": "/data",
    "RemotePath": "172.30.3.90:/",
    "Protocol": "NFS 4.0",
    "StorageType": "SD"
  }]
}
    
```

检查存储选项

通过查询集群存储选项（[DescribeClusterStorageOption](#)）可以查询目前集群设置的存储选项，后续THPC会对扩容节点进行文件系统的挂载。

登录实例检查文件系统是否成功挂载。

1. 登录 [CVM 控制台](#)，找到需要检查的实例，单击登录。



2. 输入密码登录实例。
3. 登录实例后输入 `df -h` 命令。可以看到实例已经成功挂载了存储选项。

```
[root@manager ~]# df -h
Filesystem      Size  Used Avail Use% Mounted on
devtmpfs        3.8G   0  3.8G   0% /dev
tmpfs           3.8G  24K  3.8G   1% /dev/shm
tmpfs           3.8G  520K  3.8G   1% /run
tmpfs           3.8G   0  3.8G   0% /sys/fs/cgroup
/dev/vda1       50G   3.5G   44G   8% /
172.30.2.143:/  10G   32M   10G   1% /opt
tmpfs           773M   0  773M   0% /run/user/0
172.30.3.90:/   10G   32M   10G   1% /data
```

删除存储选项

通过删除集群存储选项（[DeleteClusterStorageOption](#)）接口进行删除集群存储选项，删除完之后，THPC 会尝试对当前集群存在的所有节点进行解挂载存储资源。后续扩容的节点都不会挂载上此存储选项。

[DeleteClusterStorageOption](#) 接口输入参数示例：

```
{
  "ClusterId": "hpc-5lyv94lq",
  "LocalPath": "/data"
}
```

调用成功后，可以看到实例 /data 已经不是本地挂载点：

```
[root@manager ~]# df -h
Filesystem      Size  Used Avail Use% Mounted on
devtmpfs        3.8G   0  3.8G   0% /dev
tmpfs           3.8G  24K  3.8G   1% /dev/shm
tmpfs           3.8G  520K  3.8G   1% /run
tmpfs           3.8G   0  3.8G   0% /sys/fs/cgroup
/dev/vda1       50G   3.5G   44G   8% /
172.30.2.143:/qxt67552/hpc-rv7hahw2/opt 10G   32M   10G   1% /opt
tmpfs           773M   0  773M   0% /run/user/0
[root@manager ~]#
```

存储选项填写说明

最近更新时间：2023-03-29 12:58:32

背景说明

THPC在使用创建集群（[CreateCluster](#)）或者添加集群存储选项（[AddClusterStorageOption](#)）接口时，需要使用到存储选项（[StorageOption](#)）参数。本文主要介绍不同的存储产品如何填写接口的存储选项（[StorageOption](#)）参数以使用集群存储挂载功能。

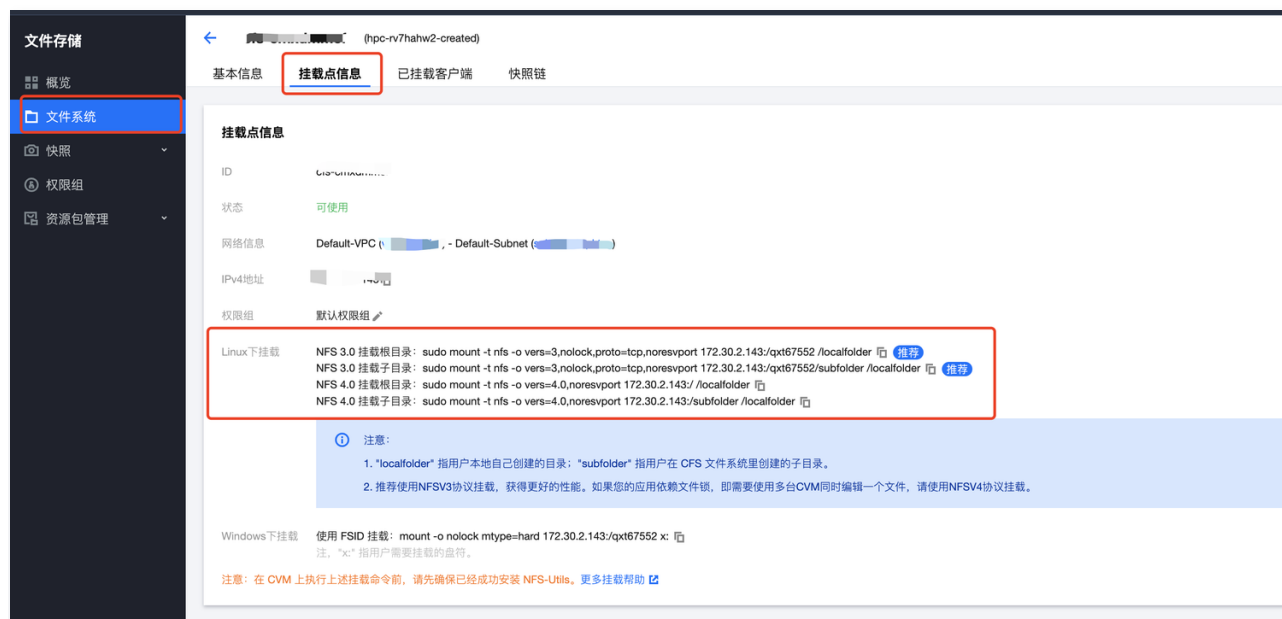
CFS 文件系统参数

CFS 的通用标准型对应 [StorageOption](#) 参数中的 [CFSOption](#) 复杂类型。

1. 进入 [CFS控制台](#)。选择左侧导航栏的[文件系统](#)，单击需要挂载的文件系统ID文件系统详情界面。



2. 进入详情界面后，单击上面标签栏的[挂载点信息](#)，可以找到 Linux 下的相关挂载命令。



3. 根据需求选择 NFS 3.0 或者 NFS 4.0 协议进行挂载操作。

NFS 3.0 挂载子目录命令如下:

```
sudo mount -t nfs -o vers=3,nolock,proto=tcp,noresvport  
172.30.2.143:/qxt67552/subfolder /localfolder
```

其中参数名与参数值对应关系如下表：

参数名	参数值	说明
RemotePath	172.30.2.143:/qxt67552/subfolder	远程存储目录
LocalPath	/localfolder	本地存储目录
Protocol	NFS 3.0	存储协议
StorageType	SD	存储类型。SD表示通用标准型

接口参数 json 格式示例如下：

```
{
  "LocalPath": "/localfolder",
  "RemotePath": "172.30.2.143:/qxt67552/subfolder",
  "Protocol": "NFS 3.0",
  "StorageType": "SD"
}
```

NFS 4.0 挂载子目录命令如下：

```
sudo mount -t nfs -o vers=4.0,noresvport 172.30.2.143:/subfolder /localfolder
```

其中参数名与参数值对应关系如下表：

参数名	参数值	说明
RemotePath	172.30.2.143:/subfolder	远程存储目录
LocalPath	/localfolder	本地存储目录
Protocol	NFS 4.0	存储协议
StorageType	SD	存储类型。SD表示通用标准型

接口参数 json 格式示例如下：

```
{
  "LocalPath": "/localfolder",
  "RemotePath": "172.30.2.143:/subfolder",
  "Protocol": "NFS 4.0",
  "StorageType": "SD"
}
```

CFS Turbo 文件系统参数

CFS 的 TURBO 对应 StorageOption 参数中的 CFSOption 复杂类型。

1.进入 **CFS控制台**。选择左侧导航栏的文件系统，单击需要挂载的文件系统ID文件系统详情界面。



2.进入详情界面后，单击上面标签栏的挂载点信息,可以找到 turbo 挂载命令。



3.填写 turbo 挂载命令。

```
sudo mount.lustre 172.30.2.180@tcp0:/4fe1839b/cfs /path/to/mount
```

其中参数名与参数值对应关系如下表：

参数名	参数值	说明
RemotePath	172.30.2.180@tcp0:/4fe1839b/cfs	远程存储目录。 turbo目前不支持子目录挂载，填写了cfs下的子目录，会挂载失败。例如：172.30.2.180@tcp0:/4fe1839b/cfs/data
LocalPath	/path/to/mount	本地存储目录。
Protocol	TURBO	存储协议。

StorageType	TB	存储类型。TB为turbo标准型， TP 为turbo性能型。
-------------	----	---------------------------------

接口参数 json 格式示例如下：

```
{
  "LocalPath": "/path/to/mount",
  "RemotePath": "172.30.2.180@tcp0:/4fe1839b/cfs",
  "Protocol": "TURBO",
  "StorageType": "TB"
}
```

GooseFS 参数

自建集群部署数据加速器 GooseFS 操作请参见 [使用自建集群部署 GooseFS](#)。GooseFS 对应 StorageOption 参数中的 GooseFSOption 复杂类型。

其中参数名与参数值对应关系如下表：

名称	参数值	说明
LocalPath	/path/to/mount	文件系统本地挂载路径。
RemotePath	/	RemotePath 为 GooseFS 中的目录，您需要确保该目录已经创建；如无特殊需求，可设置为根目录 '/'。
Masters	172.16.0.97:9200,172.16.0.79:9200,172.16.0.110:9200	Masters 为 GooseFS 加速集群的主节点地址。如果您选择高可用部署模式，主节点地址之间使用逗号分隔，例如：172.16.0.97:9200,172.16.0.79:9200,172.16.0.110:9200

接口参数 json 格式示例如下：

```
{
  "LocalPath": "/path/to/mount",
  "RemotePath": "/",
  "Masters": "172.16.0.97:9200,172.16.0.79:9200,172.16.0.110:9200"
}
```

GooseFSx 参数

创建全托管数据加速器 GooseFSx 请参见 [GooseFSx快速入门](#)。GooseFSx 对应 StorageOption 参数中的 GooseFSxOption 复杂类型。

1.进入GooseFSx控制台。进入GooseFSx 实例列表，点击需要挂载的 GooseFSx 实例 ID 进入集群详情页面，如图，GooseFSx 实例 ID 为 x_c60_54oi5id3。

数据加速器 GooseFS 实例列表 上海 1

数据加速器 GooseFSx 帮助

多个关键字用竖线“|”分隔，多个过滤标签用回车键分隔

ID/名称	监控	状态	总容量	可用区	所属VPC/子网	实例类型	创建/更新时间	标签	操作
x_c60_54oi5id3 默认名称_1677853957		运行中	4.5 TiB	上海二区		GooseFSx C60	2023-03-01 14:59:51 2023-03-01 15:49:25		管理 扩容 删除

共 1 条

10 条 / 页

2. 点击 POSIX 客户端，可以看到 GooseFSx 已经创建好的客户端管理节点列表，获取客户端管理节点内网 ip。

← x_c60_54oi5id3 默认名称_1677853957

数据加速器 GooseFSx 帮助

POSIX 客户端

① POSIX 客户端：已部署 GooseFSx POSIX 客户端软件的，并将 GooseFSx 挂载成本地目录的主机。
 POSIX 客户端管理节点：管理 POSIX 客户端的部署、挂载、删除等整个生命周期。在创建 GooseFSx 实例时，将会在您的腾讯云账号下自动创建3台腾讯云服务器（Cloud Virtual Machine, CVM）实例，实例规格不小于 4C8G，部署成 POSIX 客户端管理节点。POSIX 客户端管理节点随 GooseFSx 实例一同创建和销毁，您无需管理，但需注意，请勿销毁或变更 POSIX 客户端管理节点，否则会导致 POSIX 客户端工作异常。

POSIX 客户端列表

添加 批量删除

POSIX 客户端兼容性列表

主机ID/名称	内网 IP	状态	操作
暂无数据			

共 0 条

10 条 / 页

POSIX 客户端管理节点

主机ID/名称	内网 IP	操作
ins-9541zd41 x_c60_54oi5id3-customer-manager-1	172.16.0.9	查看初始密码
ins-r85he7l3 x_c60_54oi5id3-customer-manager-2	172.16.0.3	查看初始密码
ins-n8vintpn x_c60_54oi5id3-customer-manager-3	172.16.0.5	查看初始密码

共 3 条

10 条 / 页

3. GooseFSx 填写参数

参数名	参数值	说明
Masters	["172.16.0.6:55533", "172.16.0.13:55533", "172.16.0.14:55533"]	GooseFSx 全部管理节点的列表。GooseFSx 服务必须配置内网ip为 55533 端口。

LocalPath	/goosefsx/x_c60_54oi5id3_proxy/	/goosefsx/{Goosefsx实例ID}_proxy/ \
-----------	---------------------------------	-----------------------------------

⚠ 注意:

1. 一个 THPC 集群只支持挂载一个 GooseFSx 文件系统。
2. 目前暂时不支持指定本地目录挂载操作，GooseFSx在实例内挂载路径为/goosefsx/{Goosefsx实例ID}_proxy/ ，目前不支持其他子路径填写形式。
3. 删除 [GooseFSx 集群存储选项](#) 时，LocalPath 请填写 /goosefsx/{Goosefsx实例ID}_proxy/。

集群初始化脚本配置

最近更新时间：2023-12-27 11:36:11

背景说明

THPC 集群中的节点在创建过程中，提供用户自定义集群初始化脚本能力，满足客户对于节点开始执行任务之前，预部署需求。包括存储的挂载、数据预处理、配置变更和软件部署等操作。

初始化脚本配置方式

设置集群初始化脚本有以下两种方式：

1. [创建集群](#)时，通过 `InitNodeScripts` 参数进行设置。
2. 集群创建成功后，通过 [修改节点初始化脚本](#) 接口进行配置。

⚠ 注意：

创建集群时的初始化脚本信息对于创建的节点都生效。[修改节点初始化脚本](#) 接口配置初始化脚本之后，对后续新增的节点生效，对于已经存在的节点，不会执行集群初始化脚本。

操作步骤

1. 创建 COS 存储桶，上传初始化脚本到 COS 存储，脚本必须以 `#!/` 字符以及指向要读取脚本的解释器的路径（通常为 `#!/bin/bash`）开头。执行脚本支持以下环境变量，客户可以根据环境变量进行脚本执行的逻辑控制：

环境变量名	说明	参考值
THPC_NODE_ROLE	节点的角色	COMPUTE LOGIN MANAGER
THPC_CLUSTER_ID	集群ID信息	hpc-8hje8h12
THPC_SCHEDULER_TYPE	调度器类型	SLURM
THPC_ACCOUNT_TYPE	域名字服务类型	NIS

⚠ 注意：

初始化脚本需要提前上传到 cos 桶上进行存储。

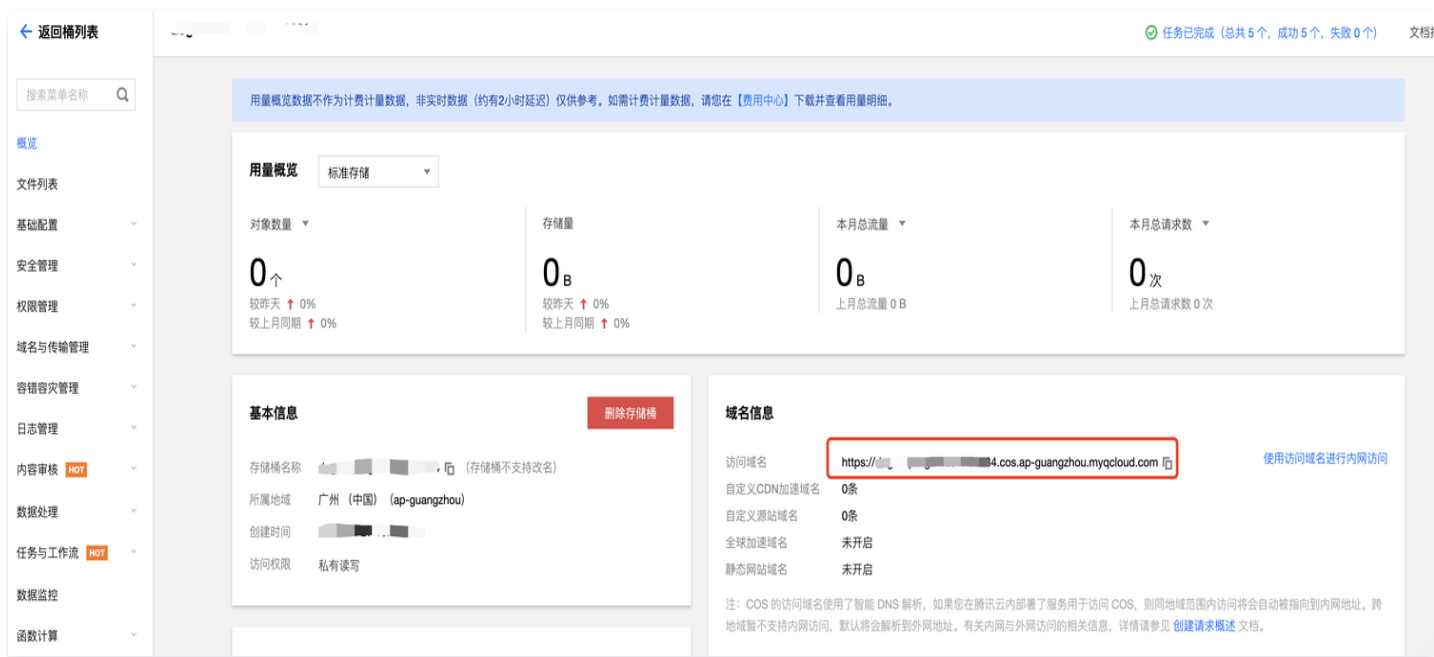
由于初始化脚本在加入集群之前进行执行，需要客户评估脚本执行超时时间，否则有可能脚本执行异常，导致实例长时间无法加入集群被使用。

示例脚本如下：

```
#!/bin/bash
echo ${THPC_NODE_ROLE} > /root/${THPC_NODE_ROLE}
echo ${THPC_CLUSTER_ID} > /root/${THPC_CLUSTER_ID}
echo ${THPC_SCHEDULER_TYPE} > /root/${THPC_SCHEDULER_TYPE}
echo ${THPC_ACCOUNT_TYPE} > /root/${THPC_ACCOUNT_TYPE}
```

2. 在创建集群时配置 InitNodeScripts 参数或者集群创建成功后调用 [修改节点初始化脚本](#) 接口配置；

其中 ScriptPath 参数需要填写以 cos:// 的形式开头，将图中访问域名的 https:// 替换成 cos://，后续添加添加上存储桶访问路径即可。



例如 `cos://test-appid.cos.ap-guangzhou.myqcloud.com/thpc_test/install.sh` 其中

`test-appid.cos.ap-guangzhou.myqcloud.com` 为存储桶访问域名，`/thpc_test/install.sh` 为脚本访问路径。

说明：

目前只支持拉取与集群相同账号的 COS 桶，跨账号的 COS 桶暂时不支持拉取。

3. 在创建集群或者节点创建时，集群中的节点会从 cos 地址中获取到脚本，并在 `/tmp/hpc_script/` 目录下进行执行。

脚本执行结果判断：

(1) 集群脚本在超时时间内，返回 `exit_code` 为0，则被视为执行成功。

(2) 集群脚本执行超时，或者返回的 `exit_code` 不为0，则被视为执行失败，执行失败的节点会被自动退还并销毁。

说明：

为了保证用户的操作能够在节点加入集群之后可以正常运行任务，初始化脚本在节点加入集群之前进行执行。执行自定义脚本会增加节点扩容所需时间。

集群扩缩容策略说明

最近更新时间：2023-12-27 11:36:12

背景说明

THPC 的弹性扩缩功能支持根据您的业务负载动态伸缩集群节点，帮助您节省集群成本。您可以根据需要选择不同的扩缩容模式，在 [创建集群](#) 接口中的 AutoScalingType 进行配置，推荐您使用 THPC_AS 模式进行扩缩容，支持自动匹配作业需要的实例资源，且弹性伸缩的实例与创建集群时配置的TAG一致。

说明：

创建弹性扩缩容集群请参见 [自动伸缩最佳实践](#)，目前仅支持 Slurm 集群开启 THPC_AS 模式。

THPC_AS 扩缩容策略

需把 AutoScalingType 设置为 THPC_AS 后，使用 [SetAutoScalingConfiguration](#) 接口配置扩缩容策略，以下是多个业务场景扩缩容效果的说明。

扩容匹配原则

THPC 主要通过实例的 CPU、内存规格与提交作业的资源申请规模进行匹配，优先匹配到运行作业后剩余资源最少的实例。THPC 会优先保证 CPU 资源剩余最少，如果实例运行作业后剩余 CPU 资源相同，会再匹配内存资源。

场景示例

场景 a: 作业所需4核4GB内存资源，扩容配置如下：

实例类型	vCPU	内存(GB)
A	4	16
B	8	8

THPC 会匹配到实例类型 A，因为 CPU 的优先级高，优先匹配 CPU 资源。

场景b: 作业所需4核4GB资源，扩容配置如下：

实例类型	vCPU	内存(GB)
A	4	16
B	4	8

THPC 会匹配到实例类型 B，因为在 CPU 相同的情况下，会匹配内存资源，B 实例运行上作业之后，剩余资源较少。且实例扩容的数量取决于此作业运行起来需要的节点数量。

⚠ 注意：

- 目前匹配流程中，不会考虑实例运行作业之后的剩余资源。
- 投递的作业所需资源在机型配置中没有满足的实例规格，则不会进行扩容。
- 投递作业时需要注意，登录实例操作系统看到的可使用的内存会小于实例规格的内存。例如：SA2.MEDIUM4 机型规格为4核4GB内存，扩容后的实例 OS 可用内存为3.6GB。

多机型、多可用区扩容

实例类型在每个可用区支持情况不同，您可能会面对所选机型在当前可用区售罄或相关配置问题，多可用区、多机型配置可帮助您有效的解决此类问题。

多机型配置功能能够有效降低售罄、实例调用失败带来的自动扩容失败风险。

场景示例

在 THPC_AS 策略中，可以对 CPU 核数和内存大小相同的机型进行多配置，以下为 SetAutoScalingConfiguration 接口中，队列名为 compute 所设置的ExpansionNodeConfig 机型列表、机型的资源信息、可用区以及对应的数组下标。

数组下标	机型	可用区	vCPU	内存(GB)
1	SA2.2XLARGE16	广州四区	8	16
2	SA2.2XLARGE16	广州五区	8	16
3	SA2.2XLARGE16	广州六区	8	16
4	SA3.2XLARGE16	广州五区	8	16
5	SA3.2XLARGE16	广州六区	8	16

假设客户投递一个作业，需要6核 CPU 和12GB内存，由于队列中最优匹配是需要8核、16GB内存的资源，那么弹性伸缩配置中，所有8核、16GB内存的资源都符合匹配，THPC 会根据配置的列表下标由低往高，依次尝试创建实例进行扩容。

比例扩容

由于扩容流程中，不会考虑扩容后实例的剩余资源，出于成本考虑，在扩容场景下，可以通过比例扩容的方式来逐步扩容，尽可能减少资源的浪费。

队列配置 中，涉及 [SetAutoScalingConfiguration](#) 接口的 [队列配置](#) 中的两个参数：

参数名称	参数说明

ScaleOutRatio	扩容比例设置。取值范围1~100。
ScaleOutNodeThreshold	比例扩容阈值。

当计算出的任务需要扩容的数量大于 ScaleOutNodeThreshold 时，实际扩容数量为（需要扩容数量 * ScaleOutRatio）/ 100 向上取整。

当计算出的任务需要扩容的数量小于 ScaleOutNodeThreshold 时，实际扩容数量为需要扩容数量（即 ScaleOutRatio=100）。

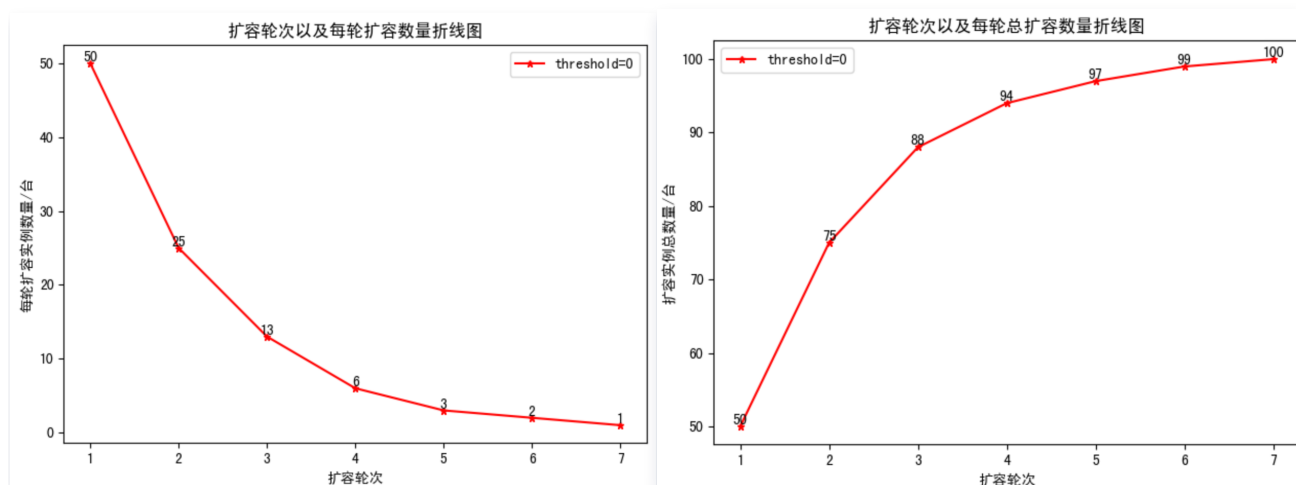
ScaleOutNodeThreshold 阈值有利于比例扩容在小数量扩容场景下，快速收敛。

场景示例

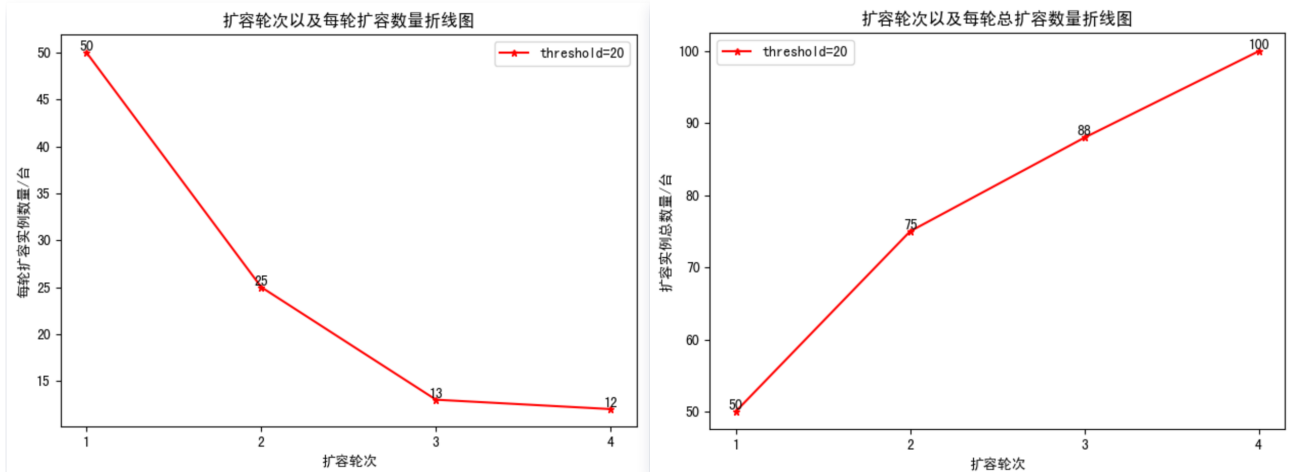
假设此时总共需要扩容100个节点加入集群，以下 A、B 配置的扩容比例设置为50，但是 ScaleOutNodeThreshold 阈值配置不同。

配置	ScaleOutRatio	ScaleOutNodeThreshold
A	50	0
B	50	20

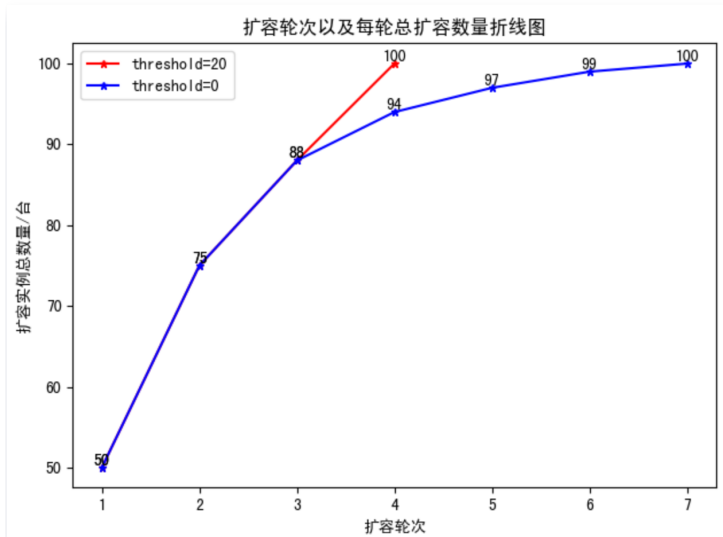
对于配置 A，100个节点总共需要自动扩容7轮，每轮节点扩容个数以及对应轮次如下图左，每轮节点扩容总数以及对应轮次如下图右：



对于配置 B，100个节点总共需要自动扩容4轮，当扩容至第4轮时，需要扩容的节点数量为12小于 ScaleOutNodeThreshold 参数值20，则按照需要扩容的节点数进行扩容。每轮节点扩容个数以及对应轮次如下图左，每轮节点扩容总数以及对应轮次如下图右：



二者每轮节点扩容总数以及对应轮次对比图如下，其中红色曲线为配置A，蓝色曲线为配置 B。可以看出，通过 ScaleOutNodeThreshold 参数能够加快在扩容实例数量较小的场景下收敛速度。



如果不需要比例扩容，可以将 ScaleOutRatio 参数默认设置为 100；如果对于扩容速度不敏感，可以将 ScaleOutNodeThreshold 设置为0。

THPC_AS 缩容策略说明

当节点连续空闲时间达到指定时间 ShrinkIdleTime，节点缩容。

使用自定义镜像加速扩容

最近更新时间：2023-12-27 11:36:12

背景说明

默认情况下，THPC 基于公共镜像进行扩容计算节点所需时间大概在3分钟到5分钟。如果您运行的任务对于扩容时间比较敏感，可以在扩容时选择预装 THPC 所需依赖包的自定义镜像，来加快 THPC 计算节点的扩容。

⚠ 注意：

- 后续 THPC 对于集群依赖软件进行升级，可能导致通过自定义镜像加速扩容的节点无法兼容，新特性无法使用等问题。
- 由于会跳过软件安装，如果自定义镜像中不存在相关依赖包，而集群操作会跳过安装，导致通过此自定义镜像创建的节点操作失败。

操作步骤

步骤一：创建集群

通过 [创建集群](#) 接口创建您所需的集群，通过 [查询集群列表](#) 接口查询到集群为 RUNNING 状态后，节点上已经安装上了 THPC 所需的安装包。

您需要创建至少一台计算节点，为自定义镜像制作做环境准备。管控节点制作的自定义镜像不能用于计算节点的扩容。

步骤二：制作自定义镜像前准备

登录集群中的计算节点：

1. 创建 thpc_preinstalled.json 文件（必选）

在创建自定义镜像前，需要在 `/root/` 目录下生成 `thpc_preinstalled.json` 文件。THPC 判断此文件存在，则会跳过扩容过程中所有的安装流程。

```
touch /root/thpc_preinstalled.json
```

2. 清理存储选项开机自挂载（建议）

THPC 在挂载存储选项流程中，会给计算节点注册上开机自挂载，如果该自定义镜像需要用于其他 THPC 集群使用，开机自挂载会导致此节点能够访问到预期之外的数据。

ⓘ 说明：

对于 CFS，开机自挂载信息在 `/etc/fstab` 文件中进行定义。文件中每条信息遵循格式：
`<file system> <mount point> <type> <options> <dump> <pass>`。可以根据 `mount`

point, 本地挂载点来判断挂载的 CFS。建议将 `/etc/fstab` 中所有挂载的 CFS 信息进行删除。如果该自定义镜像需要用于其他 THPC 集群使用, 强烈建议进行此步骤。

3. 清理集群免密信息 (建议)

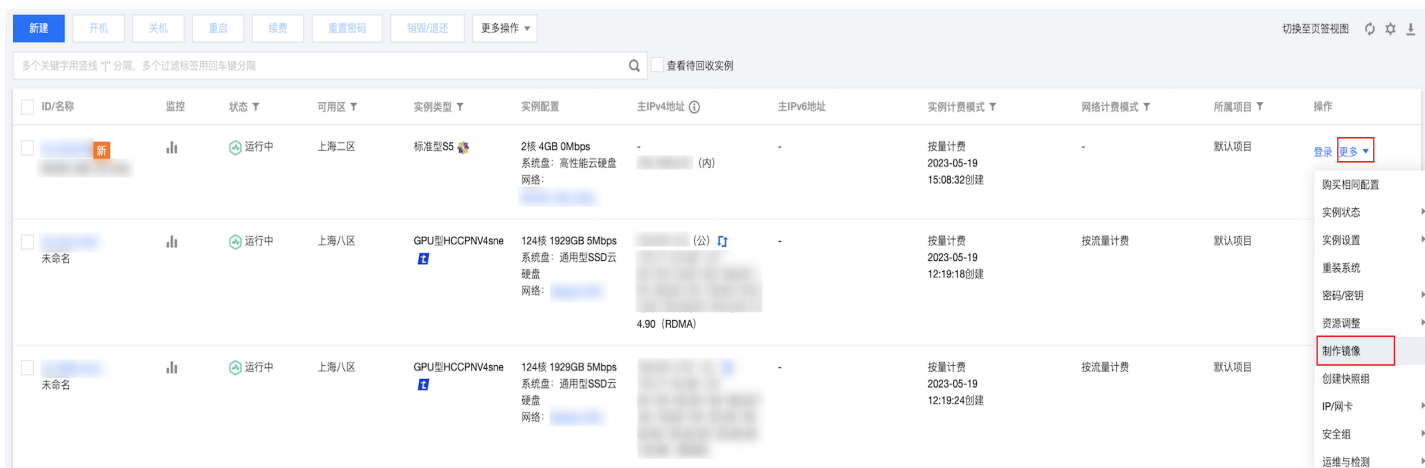
THPC 在初始化节点流程中, 会为集群中所有节点配置 SSH 免密, 如果该自定义镜像需要用于其他 THPC 集群, 免密信息会导致其他集群的节点能够免密登录此集群的节点。您需要将 `/root/.ssh/` 目录下所有的文件删除。

步骤三: 安装软件到自定义镜像中 (建议)

客户所需要的软件, 可以提前安装到此计算节点上, 后续据此计算节点制作的自定义镜像扩容的计算节点都会存在软件, 不需要再进行安装。建议安装的软件包括 GPU 驱动、作业所需的运行环境等。

步骤四: 制作自定义镜像 (必选)

在 CVM 控制台, 找到执行完上述步骤的计算节点实例, 选择**更多 > 制作镜像**。



填写镜像名称, 单击**制作镜像**。

制作自定义镜像



- 1. 创建自定义镜像的同时系统默认会创建相关快照，删除此快照之前需要先删除关联的镜像，**当前快照已经商业化**，保留镜像会产生一定的快照费用。账户欠费，会导致制作镜像失败。详细可见 [快照商业化FAQ](#)
- 2. 该实例支持不关机状态下制作镜像，镜像制作过程约需10分钟或者更长时间，与磁盘大小有关，请以实际情况为准。
- 3. Linux实例制作自定义镜像请确认 `/etc/fstab` 不包含**数据盘配置**，否则会导致使用该镜像创建的实例无法正常启动。如果有挂载数据盘，需要注释或删除`/etc/fstab`中自行配置的数据盘的相关配置。

您已选择1台实例 [收起](#)

ID/名称	实例类型	实例配置
██████████-cq	标准型S3	2核 4GB 5Mbps 系统盘：通用型SSD云硬盘 网络：██████████

镜像名称 *
只支持中文、字母、数字或连接符号"_"，且最多60个字符

标签 (选填) ×

+ 添加

镜像描述
0 / 256

你还可以输入256个字符

快照预热 对镜像关联的快照执行预热，预热后可显著提升创建实例速度

等待大概5分钟左右，在 [自定义镜像界面](#) 可以找到对应自定义镜像信息。



当镜像状态为正常状态，即可将对应镜像的 image-id 用到 [手动扩容](#) 或者 [自动扩容](#) 流程中。