

基因容器

# 用户指南

文档版本 01  
发布日期 2023-05-30



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 欢迎使用基因容器服务.....</b>	<b>1</b>
<b>2 环境管理.....</b>	<b>2</b>
2.1 环境选择.....	2
2.2 云容器实例环境.....	3
2.3 云容器引擎环境.....	6
<b>3 数据管理.....</b>	<b>16</b>
<b>4 工具管理.....</b>	<b>18</b>
<b>5 流程管理.....</b>	<b>23</b>
5.1 分析流程.....	23
5.2 流程设计器.....	28
<b>6 执行结果.....</b>	<b>34</b>
6.1 查看和管理执行结果.....	34
6.2 备份执行结果.....	37
<b>7 参考基因组.....</b>	<b>40</b>
<b>8 权限管理.....</b>	<b>42</b>
8.1 委托权限.....	42
8.2 创建用户并授权使用 GCS.....	44
8.3 GCS 自定义策略.....	45

# 1 欢迎使用基因容器服务

感谢您更深入的了解、学习并使用基因容器服务（GeneContainer Service，GCS）。

基因容器服务GCS提供云端基因分析解决方案，支持DNA、RNA、液态活检等主流生物基因分析场景。基因容器基于轻量级容器技术，结合大数据、深度学习算法，优化官方标准算法，为您提供灵活可定制的分析流程、秒级可伸缩的高可靠资源。

基因容器服务GCS的使用方式包括可视化界面、Rest API、SDK、命令行，您可以选择任意一种或是搭配使用。

- 本文档提供可视化界面使用方法，界面入口为<https://console.huaweicloud.com/gcs/>。使用流程如下：

图 1-1 GCS 使用流程



- SDK使用方法请参见[SDK参考](#)，当前仅支持Python语言的SDK。
- REST API使用方法请参见[API参考](#)。
- 命令行使用方法请参见[命令参考](#)。

## 使用建议

如果您需要对华为云上的资源进行权限的精细管理，建议您在GCS前，使用IAM服务创建IAM用户及用户组，并授权，以使得IAM用户获得GCS的操作权限。IAM操作请参见[创建用户并授权使用GCS](#)。

# 2 环境管理

## 2.1 环境选择

环境是基因容器服务所需要使用的计算资源的集合。基因容器的环境由云容器实例和云容器引擎提供。

- 云容器实例

云容器实例（Cloud Container Instance, CCI）服务提供 Serverless Container（无服务器容器）引擎，让您无需创建和管理服务器集群即可直接运行容器，详细介绍请参考[云容器实例用户指南](#)。

使用云容器实例，您不再需要关注集群和服务器，通过简单的配置即可快速创建容器负载。如果基因分析流程是成熟稳定的，建议您使用云容器实例环境，可以省去对资源的关注。云容器实例环境的相关操作请参见[云容器实例环境](#)。

- 云容器引擎

云容器引擎（Cloud Container Engine, CCE）服务提供高可靠高性能的企业级容器应用管理服务，支持Kubernetes社区原生应用和工具，简化云上自动化容器运行环境搭建，详细介绍请参考[云容器引擎](#)。

使用云容器引擎，您可以管理集群和节点资源。云容器引擎环境的相关操作请参见[云容器引擎环境](#)。

- Cromwell引擎

Cromwell 是 Broad Institute 开发的工作流管理系统。通过 Cromwell 可以将 WDL（Workflow Description Language）描述的 workflow 运行在CCI容器中。Cromwell 相关操作请参见[Cromwell引擎使用指导](#)。

- 病毒基因组分析平台

病毒基因组分析平台基于 [rampart](#) 实时读取分析病毒数据，并且提供web界面查看病毒分析的结果。

病毒基因组分析平台的相关操作请参见[病毒基因组分析平台入门](#)。

### 基本概念

- 弹性云服务器

弹性云服务器（Elastic Cloud Server, ECS）是一种可随时自助获取、可弹性伸缩的云服务器，帮助您打造可靠、安全、灵活、高效的应用环境。

- 集群  
集群是容器运行所需云资源的集合，包含了若干台服务器节点、虚拟私有云等云资源。
- 命名空间  
命名空间（namespace）是一种在多个用户之间划分资源的方法。不同的命名空间中的数据彼此隔离。命名空间对应一个虚拟私有云（Virtual Private Cloud, VPC），在创建命名空间时会关联已有VPC或创建一个新的VPC，并在VPC下创建一个子网。后续在该命名空间下创建的容器及其他资源都会在对应的VPC及子网之内。
- 对象存储服务  
对象存储服务（Object Storage Service, OBS）是华为云中基于对象的存储服务，可以为您提供海量、安全、高可靠、低成本的数据存储能力。
- 桶  
桶（Bucket）是OBS中存储对象的单位。对象存储提供了基于桶和对象的扁平化存储方式，桶中的所有对象都处于同一逻辑层级，去除了文件系统中的多层级树形目录结构。对象存储服务设置有三类存储类别，分别为：标准存储、低频访问存储、归档存储，从而满足客户业务对存储性能、成本的不同诉求。创建桶时可以指定桶的存储类别。桶的存储类别可以修改。基因容器中桶可用于存储分析前后产生的数据，包括原始基因数据、流程执行中间数据及执行结果数据。
- 弹性文件服务  
弹性文件服务（Scalable File Service, SFS）提供按需扩展的高性能文件存储，可供云上多个弹性云服务器共享访问。

## 2.2 云容器实例环境

云容器实例提供无服务器容器引擎，让您无需创建和管理服务器集群即可直接运行容器。

### 创建环境

开始基因分析前，请先创建环境。

- 步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“环境管理”，在右侧页面单击“创建环境”。
  - 步骤2** 设置“默认环境”：是/否。若当前没有环境，则将要创建的环境即为默认环境。默认环境有且只有一个。在有多个环境时，执行分析任务时如果不指定投递环境，则将任务投递至该默认环境。
  - 步骤3** 选择“环境类型”为“云容器实例 CCI”。
  - 步骤4** 选择“关联OBS存储”：OBS存储用于存储分析前后产生的数据，包括原始基因数据、流程执行中间数据及执行结果数据。
    - 如果您已有可用桶，在创建环境中，选择对应的桶即可。关联OBS存储，对应的是并行文件系统，并行文件系统详细说明请参见[并行文件系统](#)。
    - 如果没有可用桶或是需要新建桶，请单击“创建OBS存储”创建。此处创建的桶无法选择企业项目，默认创建default企业项目。
  - 步骤5** 命名空间选择：“已有命名空间”或“新建命名空间”。
- 如果您在CCI中已有可用命名空间，则在创建环境时，选择对应的命名空间即可。


如果您在CCI中没有可用的命名空间，或不想使用已有命名空间，请选择“新建命名空间”，并根据需求选择参数。

- 命名空间名称：新建命名空间的名称。
- 命名空间类型：“通用计算型”和“GPU加速型”。
  - 通用计算型：支持创建含CPU资源的容器实例及工作负载，适用于通用计算场景。
  - GPU加速型：支持创建含GPU资源的容器实例及工作负载，适用于深度学习、科学计算、视频处理等场景。
- 企业项目：

该参数针对企业用户使用。如需使用该功能，请联系客服申请开通。

企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理，默认项目为default。

请从下拉列表中选择所在的企业项目。更多关于企业项目的信息，请参见《[企业管理用户指南](#)》。
- 容器所属VPC：虚拟私有云是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。此处VPC列表会根据企业项目自动过滤，若没有可选VPC，单击“创建虚拟私有云”进行创建，参数配置请参考[创建虚拟私有云基本信息及默认子网](#)。
- 子网选择：子网是VPC内IP地址块，提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。若没有可选子网，单击“创建子网”进行创建。

**步骤6** 上传访问密钥：单击  [单击上传](#)，在弹出的对话框中上传已下载访问密钥（AK/SK），单击“确认”。若没有访问密钥，请前往“我的凭证”的[管理访问密钥](#)页面新增并下载访问密钥。

**步骤7** 单击“下一步”，确认环境信息配置后单击“提交”。单击“环境管理列表”将跳转“环境管理”页面，环境状态为“运行中”，环境已创建成功。

### 须知

使用CCI环境时，请确保流程中定义的cpu（任务所需CPU核数）为0.25C的倍数，memory（任务所需内存资源大小）为1G的倍数，resources（执行任务需要的资源）中cpu与memory的配比在1:2-1:8之间。resources、memory、cpu参数的详细介绍请参见[workflow](#)。

在使用CCI创建环境时，选择新建命名空间，如果提示security-group没有找到，请用用户去CCI服务创建相关命名空间

----结束

## 查看环境

创建环境后，可以通过查看环境确认环境的状态和信息。



**步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“环境管理”，单击“云容器实例”页签，在环境列表页可查看环境状态、命名空间名称、命名空间类型、VCP名称等。



图 2-1 CCI 环境列表页



在环境列表页您可执行查看命名空间详情、变更关联桶、更新密钥、清理环境等操作。

- 查看命名空间详情：单击命名空间名称，您将进入到**CCI控制台**，可查看命名空间的详细信息、容器所属虚拟私有云VPC、子网等。
- 变更关联桶：单击关联桶后的 ，在弹出的对话框中选择其他桶，然后单击“确认”。
- 更新密钥：单击访问密钥后的 ，在弹出的对话框中上传访问密钥（AK/SK）文件，然后单击“确认”。
- 清理环境：单击“清理环境”，可清理CCI环境，具体操作请参见**清理环境**。

**步骤2** 单击环境名称，进入环境管理详情页面。在环境管理详情页面可查看虚拟私有云VPC和子网的信息，还可执行导入对象存储卷、解关联对象存储卷等操作。

图 2-2 导入/解关联对象存储卷



- 导入对象存储卷：单击“导入”，在弹出的对话框中选择要导入的存储，然后单击“确认”。导入对象存储时，如果要导入的对象存储原始名称非法（例如包含小数点“.”），将不允许导入。请确保要导入的对象存储内的数据已备份或不再使用，且没有被其他服务（如云审计、云监控等）所使用。
- 解关联对象存储卷：勾选需要删除的存储，单击“解关联”，在弹出的对话框中单击“确认”。解关联操作会解除PVC与存储的绑定关系，存储不会被删除。若想再次使用该存储，请使用导入功能导入该存储。

----结束

## 清理环境

为了防止您为不需要的服务付费，建议您使用完成后清理环境。环境清理后基因容器服务将不再使用该命名空间，命名空间还将为您保留。如默认环境被清理，则最新创建的环境会自动变成默认环境。

- 步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“环境管理”。
- 步骤2** 单击“云容器实例”页签，在需要清理的环境中单击“清理环境”。
- 步骤3** 在弹出的对话框中输入DELETE，然后单击“确认”。

----结束

## 2.3 云容器引擎环境

云容器引擎是由一系列弹性云服务器组成的一个Kubernetes集群，基因容器所有执行的流程都会运行在这个集群中。如果您对集群并不了解，请查看[集群概述](#)，了解集群概念、集群和虚拟私有云及子网间关系、集群和节点的使用约束。

### 创建环境

开始基因分析前，请先创建环境。

- 步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“环境管理”，在右侧页面单击“创建环境”。
- 步骤2** 设置“默认环境”：是/否。若当前没有环境，则将要创建的环境即为默认环境。默认环境有且只有一个。在有多个集群时，执行分析任务时如果不指定投递集群，则将任务投递至该默认集群。
- 步骤3** 选择“环境类型”：“云容器引擎 CCE”。
- 步骤4** 选择“关联桶”：OBS存储用于存储分析前后产生的数据，包括原始基因数据、流程执行中间数据及执行结果数据。
  - 如果您已有可用桶，在创建环境中，选择对应的桶即可。关联OBS存储，对应对象存储或并行文件系统，并行文件系统详细说明请参见[并行文件系统](#)。
  - 如果没有可用桶或是需要新建桶，请单击“创建桶”，基因容器将为您创建一个桶。此处创建的桶无法选择企业项目，默认创建default企业项目。
- 步骤5** 设置“计算资源”：选择“已有资源”或“新建资源”。如果您在CCE中已有可用集群，则在创建环境时，选择对应的集群做为基因容器的计算资源即可。

#### 说明

为保证有足够的资源，请给基因环境规划独立的集群。


如果您在CCE中没有可用的集群，或不想使用已有集群，请选择“新建资源”，并根据需求选择集群参数。环境创建成功后，可对计费模式和集群规格进行变更，具体操作请参见[查看环境](#)。

- 计费模式：“按需计费”或“包年/包月”。
  - 按需计费：根据实际使用的资源按小时计费。
  - 包年/包月：包周期计费。包年/包月集群创建后不能删除，如需停止使用，请到用户中心执行退订操作。
- 集群名称：新建集群的名称。

- 集群版本：选择集群对应的版本，对应Kubernetes社区基线版本，建议选择最新版本。
- 集群管理规模：当前集群支持管理的最大节点规模。若选择50节点，表示当前集群最多可管理50个节点。
- 高可用：高可用集群包含多台管理节点，单管理节点故障时，集群依然可用。推荐使用高可用集群。
- 企业项目：

该参数针对企业用户使用。如需使用该功能，请联系客服申请开通。

企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理，默认项目为default。

请从下拉列表中选择所在的企业项目。更多关于企业项目的信息，请参见《[企业管理用户指南](#)》。
- 虚拟私有云：虚拟私有云是通过逻辑方式进行网络隔离，提供安全、隔离的网络环境。若没有可选VPC，单击“创建虚拟私有云”进行创建，参数配置请参考[创建虚拟私有云基本信息及默认子网](#)。
- 所在子网：通过子网提供与其他网络隔离的、可以独享的网络资源，以提高网络安全。若没有可选子网，单击“创建子网”进行创建。如果集群规模超过250台，那么子网大于255.255.255.0。
- 可用区：指在同一服务区内，电力和网络互相独立的地理区域。一个区域内有多个可用区，一个可用区发生故障后不会影响同一区域内的其它可用区，可用区之间通过内网访问。
- 节点规格：选择所需的规格（通用计算增强型/通用型计算型/GPU加速型），选择节点的CPU配额和内存配额。
  - 通用计算增强型：计算增强型实例具有性能稳定且资源独享的特点，满足计算性能高且稳定的企业级工作负载诉求。
  - 通用型计算型：通用计算型实例提供均衡的计算、存储以及网络配置，适用于大多数的使用场景。通用计算型实例可用于Web服务器、开发测试环境以及小型数据库应用等场景。
  - GPU加速型：GPU实例提供图形处理器（GPU）及较高的计算性能配置的实例，可以适用于图形渲染要求较高的应用，例如高清视频、图形渲染、远程桌面等场景。
- 节点个数：单击 ，选择添加的节点个数。
- 操作系统：选择节点对应的操作系统。推荐使用EulerOS 2.2。
- 系统盘：磁盘类型为“普通IO”、“高IO”或“超高IO”，磁盘大小可设置为40~1024GB。
  - 普通IO：由SATA存储提供资源的磁盘类型，该类型云硬盘的最大IOPS为1000；
  - 高IO：由SAS存储提供资源的磁盘类型，该类型云硬盘的最大IOPS为3000；
  - 超高IO：由SSD存储提供资源的磁盘类型，该类型云硬盘的最大IOPS为20000。
- 数据盘：磁盘类型为“普通IO”、“高IO”或“超高IO”，磁盘大小可设置为100~32678GB。
- 登录方式：“密钥对”或“密码”。
  - 密钥对：选择登录节点的密钥对。密钥对用于远程登录节点时的身份认证，若没有密钥对，可单击“创建SSH密钥对”来新建。

- 密码：输入登录节点的密码，并确认密码。

**步骤6** 单击“下一步”，确认环境信息配置后单击“提交”。系统将跳转到“环境管理”页面，环境创建预计需要花费6-10分钟时间。新建资源的环境您可单击“查看详情”，前往CCE页面查看集群创建详情。

---结束

## 查看环境



创建环境后，可以通过查看环境确认环境的状态和信息。

**步骤1** 登录**GCS控制台**，选择左侧导航栏的“环境管理”，单击“云容器引擎”页签。在环境列表页可查看环境状态、集群名称、计费模式、关联桶等，如图2-3。

图 2-3 CCE 环境列表页



在环境列表页您可执行变更关联桶、查看集群详情、查看集群监控、休眠/唤醒环境、清理环境等操作。

- 变更关联桶：单击关联桶后的 ，在弹出的对话框中选择其他桶，然后单击“确认”。
- 查看集群详情：单击集群名称，前往CCE界面查看集群详情。
- 查看集群监控：单击 ，进入集群监控页面查看集群节点、控制节点、工作负载和负载实例的CPU、内存和磁盘等的使用情况。
- 环境休眠/唤醒：单击“环境休眠/唤醒”，可将CCE环境休眠/唤醒，具体操作请参见**休眠/唤醒环境**。
- 清理环境：单击“清理环境”，可清理CCE环境，具体操作请参见**清理环境**。

**步骤2** 单击“云容器引擎”的环境名称，进入环境管理详情页面。

图 2-4 CCE 环境管理详情



在环境管理详情页面还可查看节点、控制节点、弹性伸缩、事件、存储等信息。

- 节点：节点是指接入到平台的计算资源，包括虚拟机、物理机等。您需确保所在项目节点资源充足，若节点资源不足，会导致创建工作负载等操作失败。

图 2-5 环境管理详情-节点



单击可监控节点的CPU、内存和磁盘的使用情况。

单击“节点管理”，可根据界面提示使用用户模式对节点进行管理，“按需付费节点”标签为必选，可勾选“禁止自动扩容”、“不再接收新任务”、“任务指定内存”标签。单击“确定”后，在节点页面可查看节点属性，如图2-5。

图 2-6 节点管理-用户模式



专业模式的管理方法请参考[CCE节点标签管理](#)。


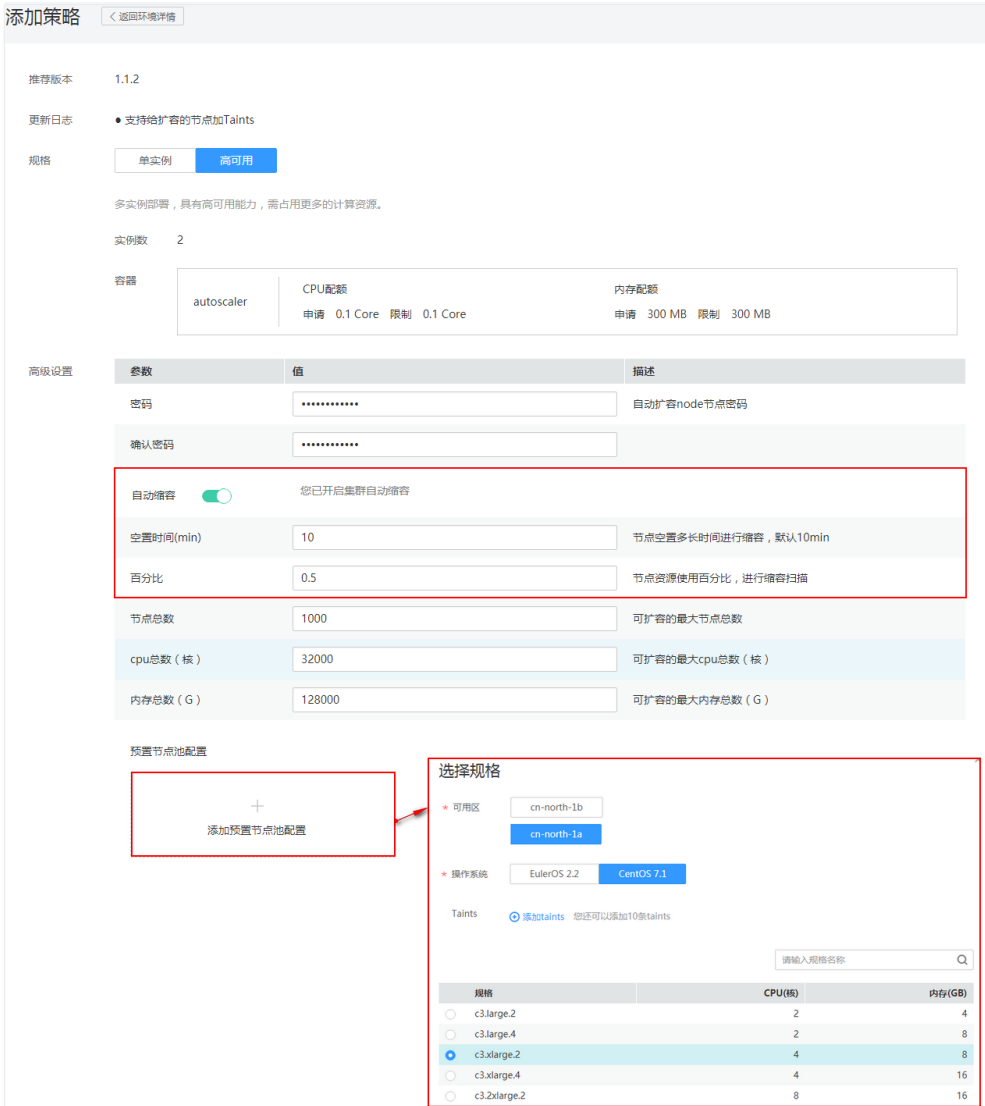
- 控制节点：控制节点由华为云托管，您不能登录。单击可监控控制节点的CPU、内存和磁盘的使用情况。如果控制节点出现异常，请联系华为云支持定位修复。
- 弹性伸缩：伸缩策略定义后，可自动触发节点的增减，从而降低人为反复调整资源以应对业务变化和高峰压力的工作量，帮助您节约资源和人力成本。单击“添加策略”，根据界面提示配置集群自动扩缩容参数。

图 2-7 添加策略



添加策略 < 返回环境详情

推荐版本 1.1.2

更新日志 ● 支持给扩容的节点加Taints

规格

多实例部署，具有高可用能力，需占用更多的计算资源。

实例数 2

容器	CPU配额	内存配额
autoscaler	申请 0.1 Core 限制 0.1 Core	申请 300 MB 限制 300 MB

高级设置

参数	值	描述
密码	*****	自动扩容node节点密码
确认密码	*****	
自动缩容	<input checked="" type="checkbox"/> 您已开启集群自动缩容	
空置时间(min)	10	节点空置多长时间进行缩容，默认10min
百分比	0.5	节点资源使用百分比，进行缩容扫描
节点总数	1000	可扩容的最大节点总数
cpu总数(核)	32000	可扩容的最大cpu总数(核)
内存总数(G)	128000	可扩容的最大内存总数(G)

预置节点池配置

+

添加预置节点池配置

**选择规格**

可用区

操作系统

Taints  添加taints 您还可以添加10条taints

请输入规格名称

规格	CPU(核)	内存(GB)
<input type="radio"/> c3.large.2	2	4
<input type="radio"/> c3.large.4	2	8
<input checked="" type="radio"/> c3.xlarge.2	4	8
<input type="radio"/> c3.xlarge.4	4	16
<input type="radio"/> c3.2xlarge.2	8	16

- 密码：自动扩容node节点密码。通过该密码登录node节点执行相关操作，用户名为root。
- 确认密码：再次输入自动扩容node节点密码。
- 自动缩容：是否启用集群缩容。
  - 空置时间（min）：启用集群缩容后，集群节点处于unneed状态多长时间后，删除节点（默认10min）。

- 百分比：启用集群缩容后，集群节点资源低于多少，会进行集群缩容（默认0.5）。
- 节点总数：可扩容的最大节点总数。
- cpu总数（核）：可扩容的最大cpu总数（核）。
- 内存总数（G）：可扩容的最大内存总数（G）。
- 预置节点池配置：默认资源组，最多可设置10个。集群扩容时，若没有其他可用分组，使用默认分组的资源规格进行扩容节点。
  - 可用区：指同一服务区内，电力和网络互相独立的地理区域。
  - 操作系统：节点对应的操作系统。
  - 节点规格：节点的CPU配置和内存配置。
- Taints：支持给该节点池扩容出来的节点加Taints来设置反亲和性，每个节点池最多配置10条Taints，每条Taints包含以下3个参数：

### 须知

Taints是可选项，默认为空。配置时需要配合Pod的toleration使用，否则可能导致扩容失败或者Pod无法调度到扩容节点。

- Key：必须以字母或数字开头，可以包含字母、数字、连字符、下划线和点，最长63个字符；另外可以使用DNS子域作为前缀，例如example.com/my-app，DNS子域最长253个字符。
- Value：必须以字符或数字开头，可以包含字母、数字、连字符、下划线和点，最长63个字符。
- Effect：只可选NoSchedule，PreferNoSchedule或NoExecute。

### 须知

Effect参数建议选择PreferNoSchedule，这样可使自动扩容的节点能够及时的缩容。

伸缩策略配置成功后，单击“卸载”可删除策略。

图 2-8 删除伸缩策略

节点	控制节点	弹性伸缩	事件	存储
版本	1.1.1		最新事件	安装中
状态	安装中		更新时间	2019/01/09 16:23:05
插件规格	CPU 0.1 Core   内存 300 MB   实例数 1		安装时间	2019/01/09 16:23:05
自动缩容	空置时间 10 min   百分比 50 %		预置节点池配置	可用区 cn-north-1b   节点规格 c3.large.2   Euler...
节点总数	1000 可扩容的最大节点总数		cpu总数（核）	32000 可扩容的最大cpu总数（核）
内存总数（G）	128000 可扩容的最大内存总数（G）			

- 事件：可以在“事件”页签下查看集群的事件。



- 存储：包括“对象存储卷”、“文件存储卷”和“云硬盘存储卷”，使用方法请参考[CCE存储管理](#)。存储卷创建成功后，单击“删除”可删除存储卷。

**步骤3** 单击集群名称，您将进入到[CCE控制台](#)，可查看集群详情信息。

----结束

## 休眠/唤醒环境

创建环境过程中用到的CCE、ECS、弹性IP等资源创建后即开始收费。当您在某段时间内不使用GCS时，需要休眠CCE环境。而由于CCI环境是按使用量计费，因此环境不使用时不产生费用，不需要休眠。

休眠后环境将保留，按需付费的CCE控制节点将暂停收费，弹性云服务器ECS、绑定的弹性IP等资源仍需收费。处于休眠状态的环境，可以执行唤醒环境操作，环境唤醒后，将继续收取控制环境资源费用。

**步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“环境管理”。

**步骤2** 选择需要休眠的环境，单击“环境休眠”，进入环境休眠详情页面，单击“确认”。

图 2-9 休眠环境



**步骤3** 环境休眠后，环境将保留。若需要继续使用环境，可单击“唤醒环境”来执行唤醒环境操作。



图 2-10 唤醒环境



----结束

## 清理环境

为了防止您为不需要的服务付费，建议您使用完成后清理环境。环境清理后基因容器服务将不再使用该集群，集群还将为您保留。已休眠的环境，请先唤醒再删除。如默认环境被清理，则最新创建的环境会自动变成默认环境。

- 步骤1** 登录GCS控制台，选择左侧导航栏的“环境管理”。
- 步骤2** 单击“云容器引擎”页签，在需要清理的环境中单击“清理环境”。

图 2-11 清理环境



- 步骤3** 在弹出的对话框中输入DELETE，环境清理后基因容器服务将不再使用该集群，集群还将为您保留。如需彻底删除该集群，请勾选“删除环境所创建的集群”，然后单击“确认”。

## 须知

- 只有通过GCS“新建资源”创建的按需计费环境才可以“删除环境所创建的集群”。
- 包年/包月集群如需停止使用，请到“费用中心”执行退订操作。

---结束

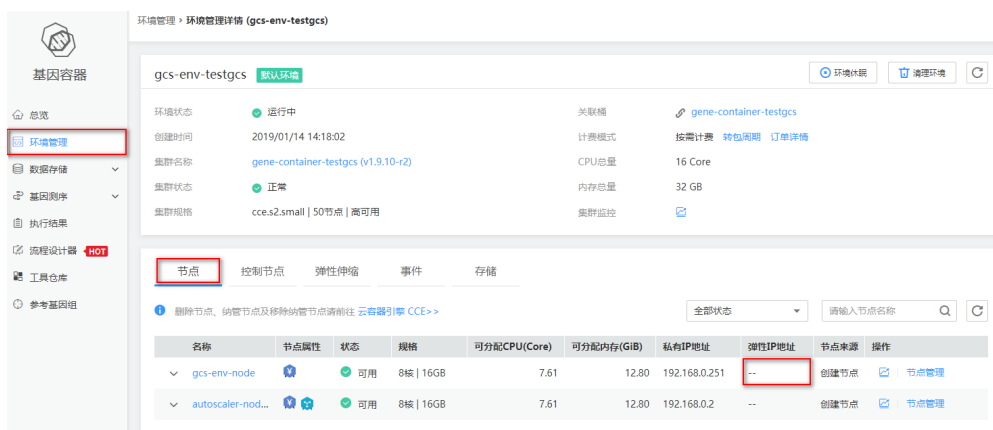
## 查看或删除中间数据

由于中间数据存储于弹性文件服务（以下简称SFS）上，并且系统不会自动检测并清理SFS数据，因此为了确保SFS上有足够的存储空间，请维护中间数据，定期删除数据。

**步骤1** 请确保CCE集群节点中配有可用的弹性IP，并已获取对应弹性IP地址。

登录**GCS控制台**，选择“环境管理 > 云容器引擎CCE”。单击环境名称，在环境管理详情页面查看集群中是否有某一个节点中已配置“弹性IP地址”（[图2-12](#)表示没有配置）。如没有配置，请进入**网络控制台**购买弹性公网IP，并绑定至集群任一节点上，详细操作请参见[为弹性云服务器申请和绑定弹性公网IP](#)。

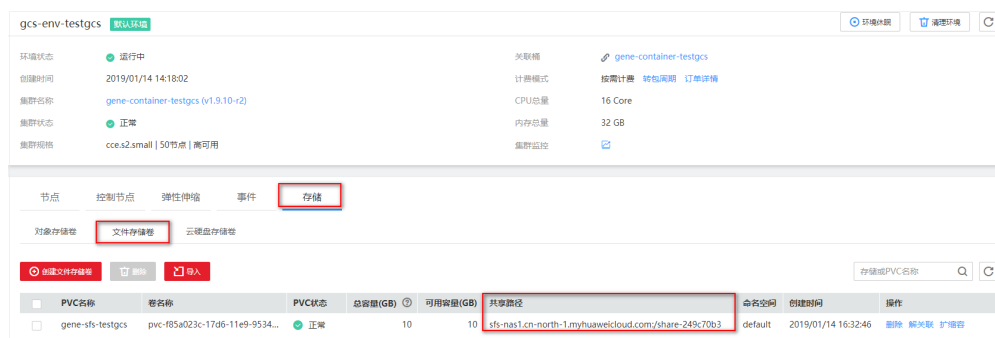
图 2-12 未配置弹性 IP 地址



**步骤2** 查找SFS文件，查看或删除中间文件。

1. 查找SFS文件共享路径位置。

登录**GCS控制台**，选择“环境管理 > 云容器引擎CCE”。单击环境名称，在环境管理详情页面单击“存储>文件存储卷”，“共享路径”列即为SFS文件共享路径位置。



2. 使用密钥对登录到有弹性IP的节点上（方法请参见[SSH密钥方式登录](#)），本示例使用Linux操作系统登录弹性云服务器。
3. 将卷挂载到节点中，以下以挂载目录至cde目录为例（挂载方法请参见[挂载文件系统到云服务器（Linux）](#)）。

```
##### Any questions please contact 4000-955-988 #####
root@test-57571 ~]# cd /
root@test-57571 /]# ll
total 57
-rwxrwxrwx 1 root root 7 Jun 6 2017 bin -> usr/bin
-rwxr-xr-x 4 root root 1024 Jun 1 16:52 boot
-rwxrwxrwx 2 root 99999 4096 Aug 16 20:58 cde
-rwx----- 7 root root 4096 Dec 8 2017 CloudResetPwdUpdateAgent
-rwx----- 7 root root 4096 Dec 8 2017 CloudrResetPwdAgent
-rwxr-xr-x 19 root root 3320 Aug 16 20:58 dev
-rwxr-xr-x 85 root root 4096 Aug 16 17:52 etc
-rwxr-xr-x 3 root root 4096 Aug 6 21:56 home
-rwxrwxrwx 1 root root 7 Jun 6 2017 lib -> usr/lib
-rwxrwxrwx 1 root root 9 Jun 6 2017 lib64 -> usr/lib64
-rwxr-xr-x 2 root root 4096 Dec 12 2016 media
-rwxr-xr-x 3 root root 4096 Aug 16 17:48 mnt
-rwxr-xr-x 6 root root 4096 Aug 16 17:52 opt
-r-xr-xr-x 213 root root 0 Aug 16 17:47 proc
-r-xr-x--- 4 root root 4096 Aug 6 22:02 root
-rwxr-xr-x 27 root root 940 Aug 16 17:52 run
-rwxrwxrwx 1 root root 8 Jun 6 2017 sbin -> usr/sbin
-rwxr-xr-x 2 root root 4096 Dec 12 2016 srv
-r-xr-xr-x 13 root root 0 Aug 16 17:47 sys
-rwxrwxrwt 11 root root 4096 Aug 16 21:24 tmp
-rwxr-xr-x 13 root root 4096 Jun 6 2017 usr
-rwxr-x--- 2 root root 4096 Aug 16 17:47 usrCustomdir
-rwxr-xr-x 22 root root 4096 Aug 16 17:52 var
```

4. 挂载完成后，可在目录中查看或删除中间文件。请定期删除过期的中间文件，确保剩余足够空间。

```
drwxrwxrwt 11 root root 4096 Aug 16 21:24 tmp
drwxr-xr-x 13 root root 4096 Jun 6 2017 usr
drwxr-x--- 2 root root 4096 Aug 16 17:47 usr
drwxr-xr-x 22 root root 4096 Aug 16 17:52 var
[root@test-57571 /]# cd cde
[root@test-57571 cde]# ll
total 4
-rw-r----- 1 root root 0 Aug 16 20:58 a.txt
-rw-r----- 1 root root 4 Aug 16 20:58 b.txt
-rw-r----- 1 root root 0 Aug 16 20:18 t.txt
```

----结束

# 3 数据管理

基因容器使用对象存储服务OBS和弹性文件服务SFS存储分析前后产生的数据，其中包括原始基因数据、流程执行中间数据及执行结果数据。CCE环境支持对象存储服务OBS和弹性文件服务SFS存储数据，而由于目前CCI环境不支持OBS挂载，CCI环境只能使用SFS存储数据。

基因容器的数据存储包括“公共数据”和“私有数据”。

- 公共数据中存放的是GCS提供的参考组数据，供您查看和使用。
- 私有数据中存放的是分析前后产生的数据。在“私有数据”页面您可以上传原始基因数据、查看流程执行的中间数据及执行结果数据。

## 数据上传

开始基因测试前，请先将用于分析的原始数据上传到已创建好的OBS桶里。

**步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“数据管理>私有数据”。

若CCE环境中存在默认环境，私有数据将展示默认环境关联桶的数据；若CCE环境中无默认环境，私有数据将展示按默认排序（gene-container开头的OBS桶在前，其他桶在后）后的第一个桶的数据。您可以通过切换桶，将数据上传到您需使用的OBS桶内。

**步骤2** 建议在OBS桶内，将不同的数据存放在不同的文件夹里。此处，建议为原始数据新建文件夹，并对应命名。

单击“新建文件夹”，将所有的原始数据传至该文件夹内。

- 页面上传：单击“页面上传”，直接上传文件。该方法适用于小于5M单文件上传。
- 工具上传：数据量较大时推荐使用OBS客户端上传数据。单击“工具上传”，根据提示登录OBS客户端，在OBS客户端中上传数据，具体方法请参考[OBS Browser入门](#)。该方法适用于批量上传多个文件或上传文件大于5M时。
- 数据快递：单击“数据快递”，进入[数据快递控制台](#)上传数据。该方法适用于TB或PB级数据上云。

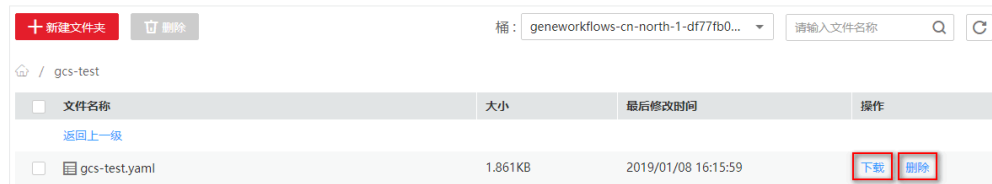
图 3-1 数据上传



**步骤3** 数据上传成功后，您还可以做如下操作：

- 下载：选择左侧导航栏的“数据管理>私有数据”，进入数据上传的文件夹中，单击“下载”，将数据下载到本地。
- 删除：选择左侧导航栏的“数据管理>私有数据”，进入数据上传的文件夹中，单击“删除”，删除数据。数据删除后不能恢复，请谨慎操作。

**图 3-2** 数据下载及删除



----结束

# 4 工具管理

工具是生物信息软件的镜像封装，工具既可以编排入流程串联使用，也可以独立使用。您可以基于基因容器为您提供业界主流的基因数据处理工具，当前已支持的公共工具请参见[公共工具的范围](#)。您可以直接使用公共工具来创建流程，如[执行工具](#)。此外，基因容器也支持添加自有工具并支持对接第三方数据处理服务，您可以添加私有工具来创建流程，这些工具都存放在“工具仓库”中。

## 公共工具的范围

基因容器当前已支持的公共工具有：

- busybox: BusyBox包含了一些简单的工具，例如ls、cat和echo等等，还包含了一些更大、更复杂的工具，例grep、find、mount以及telnet。

## 添加私有工具

您使用的工具如果不在GCS提供的公共工具中（公共工具的范围请参见[公共工具的范围](#)），您可以通过添加私有工具，使用自定义工具。

如要使用自有工具镜像，添加私有工具前，您需要前往[容器镜像服务SWR](#)上传镜像。

**步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“工具仓库”，在右侧页面单击“添加工具”。

**步骤2** 设置工具参数，其中带“\*”的参数为必填参数。

- 企业项目：  
该参数针对企业用户使用。如需使用该功能，请联系客服申请开通。  
企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理，默认项目为default。  
请从下拉列表中选择所在的企业项目。更多关于企业项目的信息，请参见[《企业管理用户指南》](#)。
- \* 工具镜像：单击“选择镜像”，可选择“我的镜像”、“共享镜像”、“镜像中心”或“第三方镜像”。
  - 我的镜像：您在容器镜像服务中上传的自有镜像。
  - 共享镜像：他人共享给您的镜像。
  - 镜像中心：公有镜像，例如nginx、alpine、busybox等。
  - 第三方镜像：即SWR容器镜像仓库与镜像中心之外的镜像仓库，使用方法请参考[如何使用第三方镜像](#)。

- \* 工具名称：新建工具的名称。
- \* 工具版本：新建工具的版本。同一工具支持多版本，添加成功后可查看版本情况，如图4-1。

图 4-1 工具版本



- 发布者：工具发布者名称。
- 工具LOGO：工具LOGO为60\*60px大小，上传图片支持PNG、JPG格式。
- \* 标签：工具的分类，也可添加自定义标签。添加成功后，可在“工具仓库>私有工具”的对应标签下找到该工具。
- 使用说明：工具的说明。
- 常用命令：多条命令以换行区分。

**步骤3** 单击“完成”，工具即添加成功。可在“私有工具”中查看添加的工具。

----结束

## 执行工具

您可以直接使用公共工具运行流程，也可以使用已添加的私有工具运行流程。

**步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“工具仓库”，选择公共工具或私有工具，单击工具右侧的“执行”。

**步骤2** 在弹出的对话框中，设置流程名称、内存、CPU，上传执行脚本（格式为.sh），单击“确定”，该任务将添加至自定义流程中。

系统将自动跳转到“基因分析 > 自定义流程”的运行流程页面，运行流程操作请参考[运行流程](#)。

图 4-2 执行工具



---结束

## 查看工具

您可以查看公共工具和私有工具的使用说明和示例等。

- 步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“工具仓库”，选择公共工具或私有工具。
- 步骤2** 单击工具名称，进入工具详情页面，该页面展示了工具的基本信息（名称、标签、创建时间、发布者、描述、企业项目）、使用说明、使用示例及工具版本，如[图4-3](#)。



图 4-3 工具详情

**bwa**

工具名称: bwa 标签: 基因组分析基础

创建时间: 2018/09/03 09:57:55 发布者: 基因容器团队

工具描述: BWA(Burrows Wheeler Aligner)与 Samtools, 是经典的基因组比对应用软件, 可以用来建立参考基因组的比对索引, 继而基于索引进行基因组的快速比对, 并进行 sam/bam 格式转换、序列排序等操作。

---

**使用说明**

--

---

**使用示例**

命令行方式:

1. 手动列出所有命令及参数排列组合, 可通过 \${var} 引用 input 定义的变量。
2. 将命令行存入 obs 桶中 scripts 目录, 将变量暴露到外部, 使用命令行模板, 迭代生成, 手动列出所有排列组合。
3. 将命令行存入 obs 桶中 scripts 目录, 将变量暴露到外部, 使用命令行模板, 迭代生成, 自动遍历组合。

```

1 job-xxx:
2 tool: bwa:0.7.12
3 type: GCS.Job
4 resource:
5   memory: 8G
6   cpu: 2C
7 commands:
8   - command: ${volumepathobs}/scri
9     - command: ${volumepathobs}/scri
10    - command: ${volumepathobs}/scri
11    - command: ${volumepathobs}/scri
                
```

```

1 job-xxx:
2 tool: bwa:0.7.12
3 type: GCS.Job
4 resource:
5   memory: 8G
6   cpu: 2C
7 commands_iter:
8   command: ${volumepathobs}/script
9   vars:
10    - ["sample1", 1, 25]
11    - ["sample0", 1, 25]
12    - ["sample1", 1, 25]
13    - ["sample0", 0, 25]
                
```

```

1 job-xxx:
2 tool: bwa:0.7.12
3 type: GCS.Job
4 resource:
5   memory: 8G
6   cpu: 2C
7 commands_iter:
8   command: ${volumepathobs}/script
9   vars_iter:
10    - ["sample1", "sample2"]
11    - [0, 1]
12    - [25]
                
```

---

**工具版本**

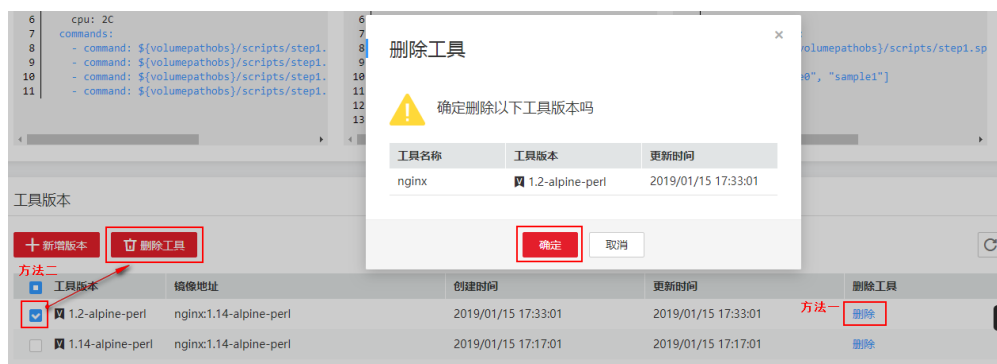
工具版本	镜像地址	创建时间	更新时间
<input type="checkbox"/> 0.7.12	100.125.0.198:20202/genecontainer/bwa:0.7.12-...	2018/09/03 09:57:55	2018/09/03 09:57:55

**步骤3** 在“私有工具”的工具详情页面, 您还可以做如下操作:

- **更新工具:** 单击“更新工具”, 可修改“发布者”、“工具描述”、“工具 LOGO”、“使用说明”和“常用命令”。更新时只更新最新版本工具信息。
- **删除工具版本:**
  - 方法一: 单击“删除”, 在弹出的对话框中单击“确定”。
  - 方法二: 勾选要删除的工具, 单击“删除工具”, 在弹出的对话框中单击“确定”。该方法可批量删除工具版本。

若删除后工具不存在版本, 则该工具将被删除。数据删除后不能恢复, 请谨慎操作。

图 4-4 删除工具版本



- 新增工具版本：单击“新增版本”，可新增工具版本。在新增版本页面设置工具参数，其中工具名称即为当前工具的名称，不可修改。工具版本名称不可与已有版本名称重复。

----结束

# 5 流程管理

## 5.1 分析流程

基因分析流程包含分析过程所需工具的执行先后信息以及数据输入输出等定义。流程由至少一个工具组成。流程中的各个工具由其前后顺序关系形成数据流，前序工具为后序工具提供输入。

基因容器流程分为示例流程和自定义流程，当前已有的示例流程有gcs-grammar，其余可以通过“自定义流程”页面来创建。基于示例流程，您可以快速完成分析任务。您也可以使用自定义流程进行分析。

### 创建自定义流程

当需要使用除示例流程之外的流程执行分析时，首先要创建自定义流程。

**步骤1** 登录[GCS控制台](#)，左侧导航栏中选择“基因分析>自定义流程”，在右侧页面中选择“GCS流程”或“WDL流程”，单击“创建流程”。

**步骤2** 创建流程有两种方式：

- 上传本地流程模板

流程模板是一种遵循GCS流程或WDL流程描述语法规则的文本文件，用于控制流程的详细步骤。基因容器的流程模板的编写请参见[流程语法参考](#)。

a. 设置流程信息。

- 企业项目：

该参数针对企业用户使用。如需使用该功能，请联系客服申请开通。



企业项目是一种云资源管理方式，企业项目管理服务提供统一的云资源按项目管理，以及项目内的资源管理、成员管理，默认项目为default。

请从下拉列表中选择所在的企业项目。更多关于企业项目的信息，请参见《[企业管理用户指南](#)》。

- 流程名称：输入流程名称，需确保该值唯一。

- 流程LOGO（可选）：上传流程LOGO。

- 选择文件/选择WDL流程文件：上传流程模板文件，上传后可预览流程模板内容。

- 发布者（可选）：输入发布者名称。
- 流程描述（可选）：输入流程描述信息。
- b. 单击“创建”，创建成功后系统将自动跳转到“自定义流程”页面，可查看到已创建成功的流程。
- 流程设计器创建（仅支持创建GCS流程）  
基因容器的流程设计器是一种用于创建、查看、修改流程的图形化工具，设计器的使用请参见[流程设计器](#)。
  - a. 在设计器中拖拽工具到画布中，连接各工具，指定工具间的先后顺序。
  - b. 参照设计器属性帮助栏的提示修改流程模板，单击设计器页面顶部的，校验模板是否正确。在界面左侧“操作日志”窗口中若出现“流程校验成功”，表示流程正确。
  - c. 单击设计器页面顶部的，选择“自定义流程”，选择企业项目，输入流程名称、流程Logo、发布者和流程描述，单击“保存”。  
保存成功后请到“自定义流程”页面，可查看到已创建成功的流程。

---结束

## 运行流程

您可以运行示例流程进行分析，也可以使用自定义流程进行分析。

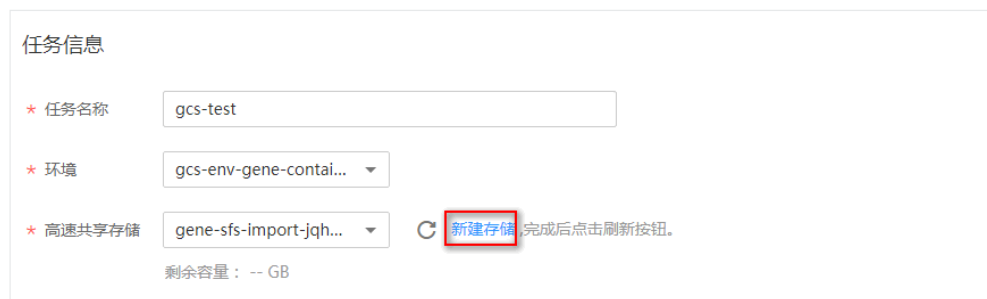
**步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“基因分析”，选择“示例流程”或“自定义流程”。

**步骤2** 在右侧页面中单击“开始分析”，按界面提示填写流程参数。

流程参数中包括“任务信息”、“配置”和“高级设置”。

- 任务信息：包括任务名称、环境和存储选择。“高速共享存储”对应的是文件存储服务SFS，用于存储流程中间数据。如果您已有可用的SFS，请在下拉框中选择。如果没有SFS，请“新建存储”，并选择容量大小。


图 5-1 任务信息



任务信息

\* 任务名称

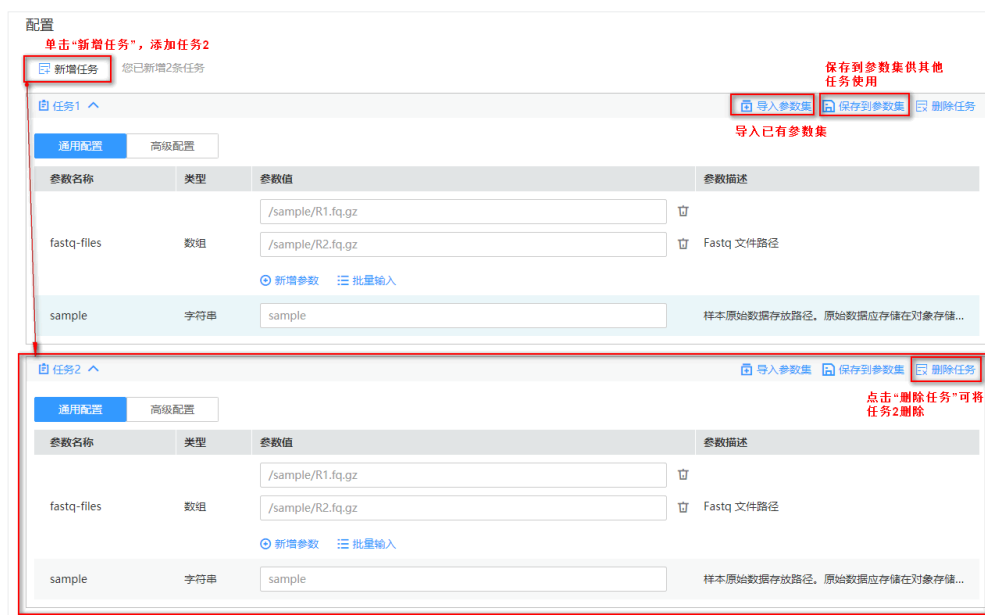
\* 环境

\* 高速共享存储   [新建存储](#) 完成后点击刷新按钮。

剩余容量：-- GB

- 配置：包括“通用配置”和“高级配置”，请根据界面提示信息完成参数配置。其中“通用配置”中的参数为必配项，需要根据您项目的实际情况修改。“高级配置”可以保留默认值，或根据您的需求修改。您还可以执行“导入参数集”、将参数“保存到参数集”、新建/删除任务等操作。

图 5-2 配置



- 高级设置：可选项，包括“超时时间”、“批次名称”、“优先级”、“订阅消息通知”、“镜像拉取策略”和“重试次数”。
  - 超时时间：设置一个可接受的流程执行成功的时间，如果到这个时间流程没有执行成功，就会停止并且报执行失败。
  - 批次名称：设置当前任务的批次，在执行结果页面可对批次进行过滤和筛选，便于批量操作同一个批次的任务。
  - 优先级：设置当前任务优先级，优先级范围为[0-999]，数值大的优先执行。
  - 订阅消息通知：订阅者可通过此功能在接收终端收到所订阅的任务事件通知消息（目前仅支持单个任务消息订阅）。订阅消息通知开启后，可勾选订阅事件（任务执行成功、任务执行失败）、添加订阅终端（短信或邮箱，最多可添加20个订阅终端）。

图 5-3 添加订阅终端



添加订阅终端后，可选择设置自定义消息模板。模板内容默认为空，最多可包含360个字符（中文字默认为3个字符，空格为1个字符）。模板中可使用占位符，{name}、{guid}、{status}占位符分别对应执行任务的名称、ID、状态，每个占位符最多使用一次，不支持其他字符写成占位符格式。  
模板内容示例：任务{name}，executionID为{guid}，任务执行状态为{status}。

消息内容示例：任务gcs-job-02-19-141146， executionID为6745c000-340d-11e9-a8c0-0255ac109xxx， 任务执行状态为Succeeded。

图 5-4 自定义消息模板



自定义消息模板

协议 ⓘ 邮箱

模板内容

任务{name}， executionID为{guid}， 任务执行状态为{status}。

70 / 360

可使用{guid}、{name}和{status}作为占位符，每个占位符至多使用1次。  
在使用该模板发送消息时，{guid}、{name}和{status}分别被替换成执行任务的ID、名称和状态。  
例如：任务{name}， executionID为{guid}， 任务执行状态为{status}。

确定 取消

- 镜像拉取策略：使用上次已拉取镜像（只有镜像不存在时，才会进行镜像拉取）、每次重新拉取（不管镜像是否存在都会进行一次拉取）。
- 重试次数：任务失败重试次数，重试次数范围为[0-9]，CCI环境默认重试0次，CCE环境默认重试3次。

**步骤3** 单击“下一步”。在“执行预览”页面，查看流程预览、执行任务预览和配置。

**步骤4** 确认无误后单击“开始”，流程执行任务提交成功。

分析流程执行时间需要数小时，具体时长与环境资源类型、环境资源大小、处理数据大小等相关，您可以单击“执行结果详情”进入执行结果页面查看执行过程。

----结束

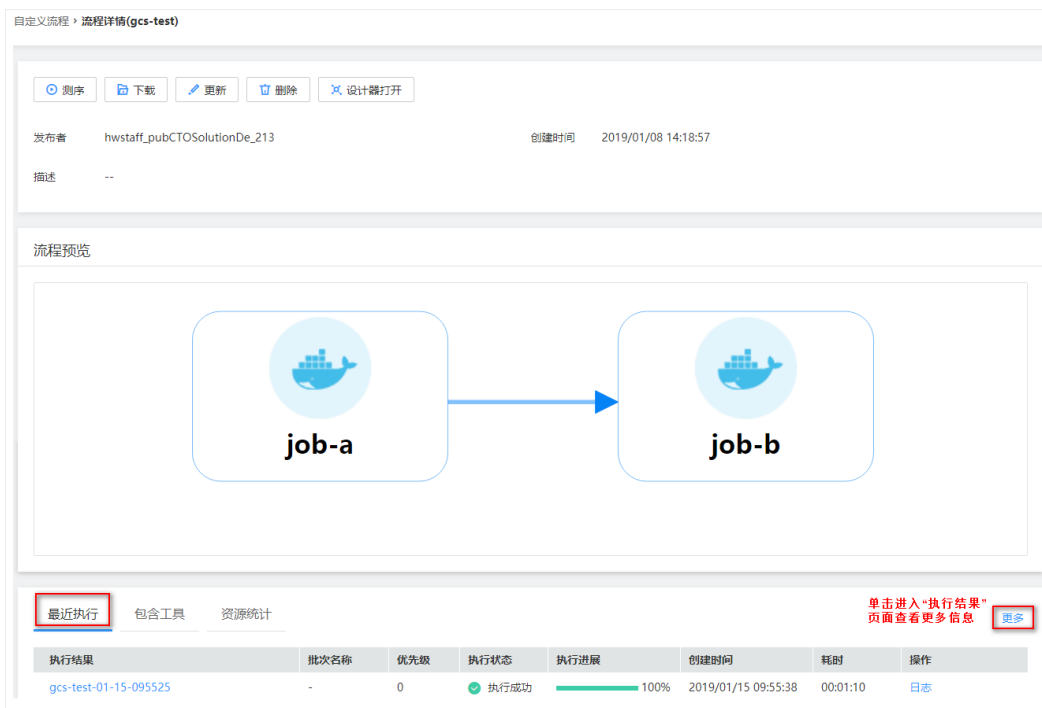
## 查看流程

您可以查看示例流程和自定义流程的信息和预览图等。

**步骤1** 登录GCS控制台，选择左侧导航栏的“基因分析”，选择“示例流程”或“自定义流程”。

**步骤2** 在右侧页面中单击流程的名称，进入流程详情页面，该页面展示了流程的基本信息（发布者、创建时间及描述）、流程预览图（预览图仅支持通过设计器生成的流程）、最近执行结果，如图5-5。

图 5-5 流程信息、预览图、最近执行结果



单击“包含工具”页签，可查看流程包含的工具。

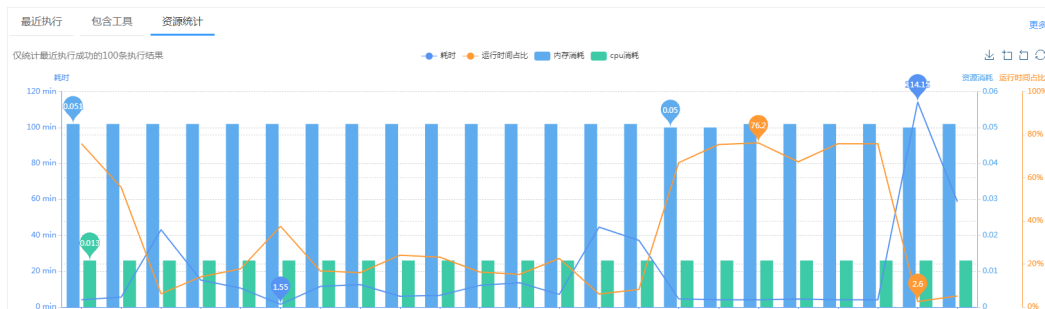
图 5-6 包含工具



单击“资源统计”页签，可查看资源统计图，包括耗时、运行时间占比、内存消耗、CPU消耗、GPU消耗等。

- 耗时：所有子任务实际运行时间与调度时间总和。
- 运行时间占比：所有子任务实际运行时间总和与耗时之比。

图 5-7 资源统计



**步骤3** 在流程详情页面，您还可以做如下操作：

- 分析：单击“分析”，将跳转到运行流程页面，运行流程操作请参考[运行流程](#)。

- 下载：单击“下载”，将流程的模板文件下载到本地。
- 更新：仅自定义流程支持更新操作。单击“更新”，可以通过上传本地流程模板更新流程LOGO、流程文件和流程描述，也可以通过流程设计器更新流程。
- 删除：仅支持自定义流程删除操作。单击“删除”，删除当前流程。删除操作无法恢复，请谨慎操作。
- 设计器打开：单击“设计器打开”，在设计器中查看、修改流程。

---结束

## 5.2 流程设计器

基因容器的流程中至少有一个工具，流程中每个工具会根据它们的先后顺序形成数据流，前一个工具会为后一个工具提供输入。

基因容器的流程设计器是一种用于创建、查看、修改流程的图形化工具。借助流程设计器，您可以拖拽工具到画布中，可视化连接各工具，指定工具间的先后顺序。

不管您是新手，还是熟悉基因容器流程编写的专家，设计器都能帮助您快速了解流程中各个工具的关系，并轻松修改流程。关于流程模板的编写方法请参阅[GCS流程语言参考](#)。

### 为什么使用设计器？

使用设计器具备以下优势：

- 可视化流程  
分析YAML格式的流程模板中的工具及其之间的关系是项艰难的任务。在设计器中，您可以查看流程中包含的工具的图形化表示，了解它们之间的关系。  
设计器在模板元数据中定义有关资源的信息，如资源大小、相对位置等。在打开流程模板时，设计器会自动添加此元数据，以便在保存模板时保留当前布局。当您在设计器中重新打开模板时，会显示和上次保存模板时完全相同的图表。
- 简化流程模板编写工作  
在文本编辑器中编写流程模板时，必须手动编辑YAML，过程繁琐又容易出错。在设计器中，通过拖放新工具就能将其添加到模板中，在工具间拖出连接线即可建立关系。模板编写完后，您还可以直接在设计器中验证模板。
- 借助集成的YAML编辑器简化编辑工作  
借助集成编辑器，您可以在设计器中执行所有模板修改操作。无需使用单独的文本编辑器来修改和保存模板。集成编辑器还提供了自动补全功能，它可列出资源的所有属性名称，无需您查找或记住。

### 设计器界面概述

基因容器的设计器界面由六个部分组成，分别是控制栏、资源栏、日志栏、设计台、流程模板栏和属性帮助栏。

设计器各部分详细功能说明如下：



图 5-8 设计器界面

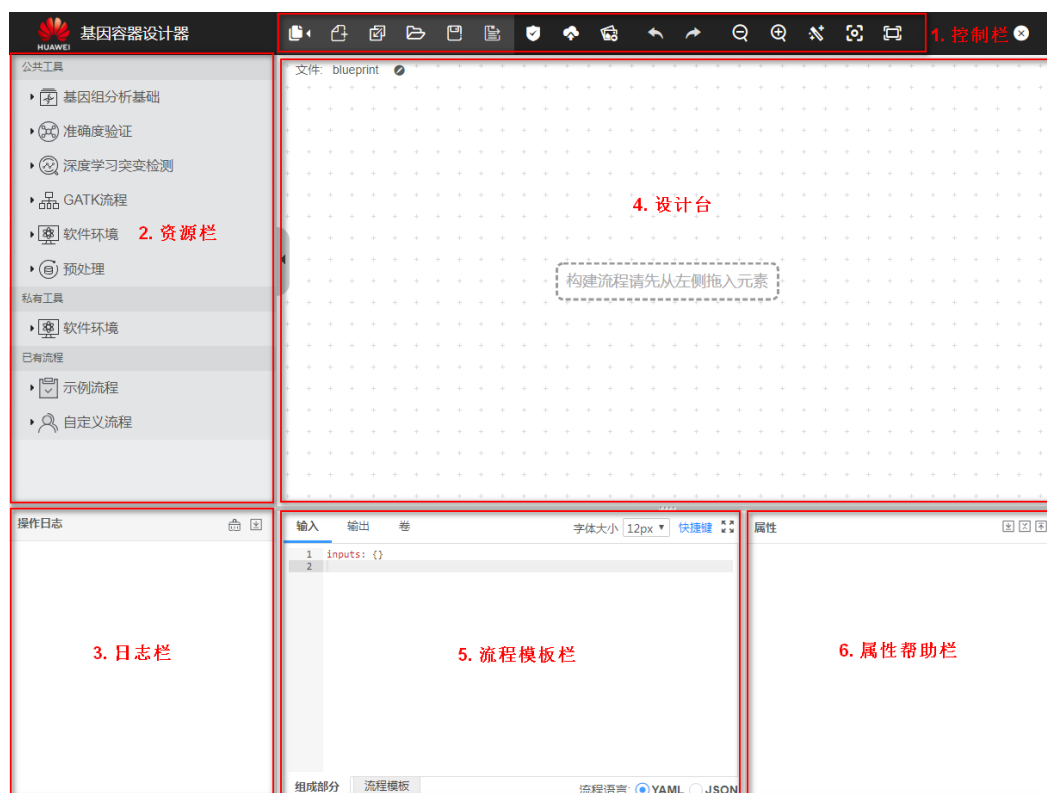


表 5-1 流程设计器页面说明

编号 (对应上图)	栏目	解释
1	控制栏	集中显示设计台的快捷控制操作。
2	资源栏	展示所有可被拖动设计的资源，包括“公共工具”、“私有工具”和“已有流程”。“已有流程”中包括“示例流程”和“自定义流程”。
3	日志栏	展示操作或设计模板过程中所触发的错误信息、提示信息等。 例如，将流程模板由JSON格式切换成YAML格式时，日志栏会提示您已将流程模板成功转换为YAML。
4	设计台	设计模板的操作画布。您可以把想要编排的资源拖动到设计台画布，并运用其生成的连接点进行逻辑关系的定义与关联。
5	流程模板栏	设计过程中动态生成相应的流程模板，支持流程模板修改和属性定义。关于流程模板的编写方法请参阅 <a href="#">GCS流程语言参考</a> 。
6	属性帮助栏	释放一个拖动资源或选中一个拖动资源时，属性栏会展示您所选资源所包含的属性名、属性类型，便于您在流程模板栏进行属性编写或定义时做参照。

## 流程资源/工具

基因分析流程使用的工具，是流程设计器设计编排的基本单位。



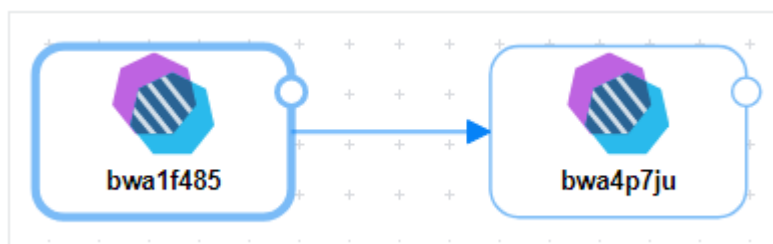
左侧工具拖拽到画布中后，出现一个图标，右键单击该图标。可查看到两个图标，单击  可复制一个相同的工具，单击  可删除该工具。

图 5-9 右键单击工具



工具上有一个“空心点”，可以通过空心点来连接其他资源。鼠标移动到空心点上，单击该空心点，会出现一个箭头，拖动鼠标，将箭头挪动到目标工具上，目标工具会变成绿色状态，放开鼠标。从工具空心点连接到其他工具上的连接线代表上游任务依赖下游任务。

图 5-10 连接线



## 设计器快捷键

操作	Windows	Mac
显示下一个错误信息	Alt-E	F4
显示上一个错误信息	Alt-Shift-E	Shift-F4
复制	Ctrl-C	Command-C
粘贴	Ctrl-V	Command-V
剪切	Ctrl-X	Command-X
全选	Ctrl-A	Command-A
定位到指定行	Ctrl-L	Command-L
展开/折叠当前元素	Alt-L Ctrl-F1	Command-Alt-L  Command-F1

操作	Windows	Mac
展开/折叠当前元素	Alt-Shift-L Ctrl-Shift-F1	Command-Alt-Shift-L  Command-Shift-F1
展开/折叠父元素	F2	F2
展开/折叠父元素	Alt-F2	Alt-F2
折叠父级元素之外的其他元素	Alt-0	Command-Option-0
展开所有折叠的元素	Alt-Shift-0	Command-Option-Shift-0
找到选中文本的下一个	Ctrl-K	Command-G
找到选中文本的上一个	Ctrl-Shift-K	Command-Shift-G
选中当前文本或找到下一个	Alt-K	Ctrl-G
选中当前文本或找到上一个	Alt-Shift-K	Ctrl-Shift-G
查找	Ctrl-F	Command-F
选中光标位置到文本开头的所有内容	Ctrl-Shift-Home	Command-Shift-Home  Command-Shift-Up
前往文本开头	Ctrl-Home	Command-Home  Command-Up
选中光标位置到上一行结尾的所有内容	Shift-Up	Shift-Up Ctrl-Shift-P
前往上一行	Up	Up Ctrl-P
选中光标位置到文本结尾的所有内容	Ctrl-Shift-End	Command-Shift-End  Command-Shift-Down
前往文本结尾	Ctrl-End	Command-End  Command-Down
选中光标位置到下一行结尾的所有内容	Shift-Down	Shift-Down Ctrl-Shift-N
前往下一行	Down	Down Ctrl-N
选中单词在光标位置左边的内容	Ctrl-Shift-Left	Option-Shift-Left
前往左边单词的开头	Ctrl-Left	Option-Left
选中光标位置到当前行开头的内容	Alt-Shift-Left Shift-Home	Command-Shift-Left  Ctrl-Shift-A
前往当前行开头位置	Alt-Left Home	Command-Left Home  Ctrl-A

操作	Windows	Mac
选中左边文本	Shift-Left	Shift-Left Ctrl-Shift-B
左移	Left	Left Ctrl-B
选中单词在光标位置右边的内容	Ctrl-Shift-Right	Option-Shift-Right
前往右边单词的开头	Ctrl-Right	Option-Right
选中光标位置到当前行结尾的内容	Alt-Shift-Right Shift-End	Command-Shift-Right Shift-End Ctrl-Shift-E
前往当前行结尾位置	Alt-Right End	Command-Right End Ctrl-E
选中右边文本	Shift-Right	Shift-Right
右移	Right	Right Ctrl-F
选中光标位置到当前页结尾的所有内容	Shift-PageDown	-
滚动到当前页结尾	-	Option-PageDown
前往当前页结尾	PageDown	PageDown Ctrl-V
选中光标位置到当前页开头的全部内容	Shift-PageUp	-
滚动到当前页开头	-	Option-PageUp
前往当前页开头	PageUp	PageUp
向上滚动	Ctrl-Up	-
向下滚动	Ctrl-Down	-
开始记录文本输入	Ctrl-Alt-E	Command-Option-E
回放文件记录	Ctrl-Shift-E	Command-Shift-E
删除当前行	Ctrl-D	Command-D
复制当前行	Ctrl-Shift-D	Command-Shift-D
对选中行进行排序	Ctrl-Alt-S	Command-Alt-S
yaml注释/解除	Ctrl-/	Command-/
对选中的数字加一	Ctrl-Shift-Up	Alt-Shift-Up
对选择数字减一	Ctrl-Shift-Down	Alt-Shift-Down
文本替换	Ctrl-H	Command-Option-F
撤销	Ctrl-Z	Command-Z

操作	Windows	Mac
重做	Ctrl-Shift-Z Ctrl-Y	Command-Shift-Z  Command-Y
复制当前行到上一行	Alt-Shift-Up	Command-Option-Up
复制当前行到下一行	Alt-Shift-Down	Command-Option-Down
当前行上移	Alt-Up	Option-Up
当前行下移	Alt-Down	Option-Down
向右删除	Delete	Delete Ctrl-D Shift-Delete
向左删除	Shift-Backspace  Backspace	Ctrl-Backspace Shift-Backspace Backspace  Ctrl-H
向左删除或者剪切	Shift-Delete	-
删除光标位置到当前行开头的 内容，有缩进	Alt-Backspace	Command-Backspace
删除光标位置到当前行结尾的 内容，有缩进	Alt-Delete	Ctrl-K Command-Delete
删除光标位置到当前行开头的 内容，取消缩进	Ctrl-Shift-Backspace	-
删除光标位置到当前行结尾的 内容	Ctrl-Shift-Delete	-
删除光标左边的单词	Ctrl-Backspace	Alt-Backspace Ctrl-Alt-Backspace
删除光标右边的单词	Ctrl-Delete	Alt-Delete
取消缩进	Shift-Tab Ctrl-[	Shift-Tab Ctrl-[
缩进	Tab	Tab
空格缩进	Ctrl-]	Ctrl-]
和后一文本交换位置	Alt-Shift-X	Ctrl-T
转换大写	Ctrl-U	Ctrl-U
转换小写	Ctrl-Shift-U	Ctrl-Shift-U
选中多行	Ctrl-Shift-L	Command-Shift-L

# 6 执行结果

## 6.1 查看和管理执行结果

分析任务的执行时间较长，一般需要数小时，详细的时长与环境资源类型、环境资源大小、处理数据大小等相关。您可以通过[GCS控制台](#)左侧导航栏的“执行结果”查看执行结果或是操作任务。

- 查看执行结果，您可以实时查看分析任务的执行过程、数据、状态、结果等信息，请参见[查看执行结果](#)。
- 管理执行结果，您可以根据需求对任务执行删除、启动、停止、编辑、更新优先级、查看日志、备份操作。备份操作请参见[备份执行结果](#)，其它操作请参见[管理执行结果](#)。

### 执行状态说明

从执行状态可以确定当前任务所属阶段，已有执行状态如下。

- 初始化：任务准备执行中。
- 执行中：任务正在执行中，可更新任务优先级、查看任务日志、删除任务、停止任务。
- 成功：任务执行成功，可查看任务日志、删除任务。
- 失败：任务执行失败，可查看任务日志、删除任务、重试执行任务。
- 已停止：任务已停止，可修改任务优先级、查看任务日志、删除任务、启动任务、更新流程的配置参数。
- 停止中：任务正在停止中。
- 删除中：任务正在删除中。

### 须知


任务数据包括“执行结果”中展示的基本数据及GCS在执行任务过程中，由GCS产生的过程数据。执行状态为成功、失败的任务，GCS会定期清理任务过程数据。

- 执行状态为成功的任务，1天后将由GCS自动清理过程数据，任务失效。失效后的任务，您仍然可以在“执行结果”列表中查看任务基本数据。
- 执行状态为失败的任务，3天后，将由GCS自动清理过程数据，任务失效。失效后的失败任务，无法再“重试”，任务基本数据仍然可以在“执行结果”中查看。

## 查看执行结果

执行结果页面可以查看所有任务，任务按创建时间排序，最新的排在最前面。

- 筛选查看结果。您可以通过页面右上角的筛选框对任务状态、任务名称、流程名称、批次名称进行过滤筛选查看任务。任务名称搜索支持模糊匹配，流程名称和批次名称搜索需要完全匹配。

- 筛选执行结果展示列。单击可以编辑执行结果展示列，勾选您想要展示的列即可完成筛选。

- 查看任务详情。您可以单击“任务名称”进入“执行结果详情”页面，查看任务对应流程名称、创建时间、耗时、任务进度图、监控等任务详情。

- 基本信息：包括流程名称、创建时间、执行状态、耗时、执行进展、批次名称、优先级、资源消耗。可以单击流程名称，查看详细的流程信息。执行状态的说明请参见[执行状态说明](#)。
- 任务进度图：显示当前任务所属流程所包括的子任务的执行情况。可以从进度视图、时间视图、内存视图、CPU视图、GPU视图、重试视图六个不同角度的视图查看。其中内存、CPU、GPU视图只有在流程中指定了内存、CPU、GPU大小才可以查看。重试视图只有执行状态为失败的任务才有。
- 流程事件：流程事件按事件时间展示了流程中重要步骤的监控信息（最新的事情展示在最前面）。
- 任务事件：任务事件按事件的时间展示任务中各子任务的操作进度（最新的事件展示在最前面）。任务事件中的“实例事件”页签显示每个子任务执行过程中的重要步骤。“Job名称”页签显示每个Job执行过程中的事件。您可以单击“查看YAML”查看流程文件。

GCS支持if/else和switch/case[条件分支](#)，当某子任务的condition值为false时，该子任务不会执行，在“任务事件”中该子任务状态显示为“忽略”。

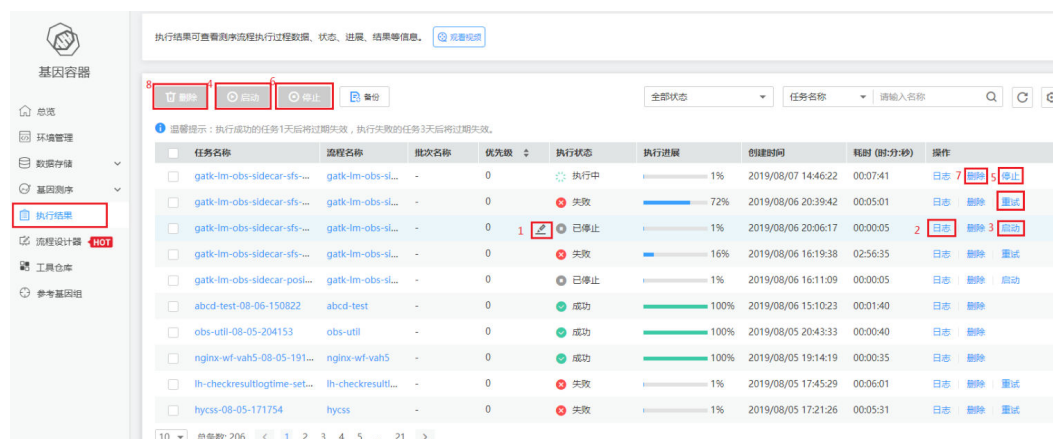


- 日志：GCS的日志免费存储在AOM服务中，日志存储时长不超过7天且占用空间不超过500MB。单击“查看”可以进入AOM服务的日志文件页面查看执行日志。

- 输入：显示在流程模板中声明的输入参数及其取值。
- 输出：显示在流程模板中声明的输出参数及其取值。
- 监控：显示各子任务的CPU、内存、网络和磁盘的监控数据。

## 管理执行结果

任务提交后，除查看执行结果外，您可以在执行结果页面对任务做操作。



- **重试任务**：对于执行失败的流程，您可以单击“重试”将执行过程中失败的任务再次执行。

### 说明

重试只会执行流程中失败的任务，不会重复执行成功的任务。例如一个流程中有10个任务，其中5个成功5个失败，那重试只会重新执行这5个失败的任务。

- **更新优先级**：更新任务的执行优先级，除执行成功和已过期失效的任务外，都支持更新优先级。

在执行结果列表，鼠标移动至需要更新优先级的任务的“优先级”列，单击“更新优先级”（上图中的1），设置当前任务优先级，范围为[0, 999]，数值越大则优先级越高，相同数值的任务并行执行。

- **查看任务日志**：查看执行结果的日志，所有任务均可查看任务日志。

在执行结果列表，单击“操作”列的“日志”（上图中的2），可以查看该执行结果的日志信息。

- **启动任务**：重新启动执行状态“已停止”的任务。

- 启动一个任务。在执行结果列表中，单击“操作”列的“启动”（上图中的3），重新启动任务。
- 批量启动任务。在执行结果列表中的“任务名称”勾选需要批量启动的任务，单击“启动”（上图中的4），重新启动任务。

- **停止任务**：停止执行状态为“执行中”的任务。

- 停止一个任务。在执行结果列表中，单击“操作”列的“停止”（上图中的5），停止任务。
- 批量停止任务。在执行结果列表中的“任务名称”勾选需要批量停止的任务，单击“停止”（上图中的6），停止任务。



**须知**

停止任务成功后即停止下发新Job至环境，已下发的Job将持续运行直至自然结束。若想停止Job，请在环境对应的服务内删除Job，或直接删除该任务。

- 删除任务：删除任务执行结果，所有任务均可删除，删除后不可恢复，请谨慎操作。
  - 删除一个任务。在执行结果列表，单击“操作”列的“删除”（上图中的7），删除任务。
  - 批量删除任务。在执行结果列表中的“任务名称”勾选需要批量删除的任务，单击“删除”（上图中的8），删除任务。

**须知**

删除任务仅能删除任务在GCS中的记录，并不会删除任务的存储在OBS或是SFS中的原始数据或是中间数据，如需删除，请手动删除。

- 编辑任务：执行状态为“已停止”的任务，可以更新流程的配置参数。单击“任务名称”，进入执行结果详情，在该页面单击“变更”，即可跳转至“变更流程”页面。在该页面根据需要修改参数，完成后单击“下一步”，确认无误后单击“完成”，流程更新任务提交成功。

## 6.2 备份执行结果

GCS的执行结果最多可以存放500条记录，当超过配额后，执行新的任务将会失败，配额查询方法请参见[我的配额](#)。为了保证能正常执行新的任务，您可以将历史执行结果备份至自有OBS桶，备份的同时，被备份的执行结果将从“执行结果”列表中删除，移动至您指定的OBS桶中。

此外，通过历史数据备份，您可以使用备份的历史数据做季度或是年度的业务分析。

### 注意事项

- 只有任务状态为“成功”、“失败”、“已停止”的执行结果可以备份。
- 单次备份最多备份50条执行结果。

### 备份执行结果

备份后的执行结果数据，将以CSV文件格式保存在指定OBS存储的指定路径下以CSV文件的方式，并从“执行结果”列表中删除。为保证新的分析任务可以正常执行，建议定期备份历史数据。

**步骤1** 登录[GCS控制台](#)，选择左侧导航栏的“执行结果”，在右侧页面单击“备份”。



**步骤2** 在弹出的容器中填写参数，并勾选需要备份的任务，完成后单击“确认”。

备份执行结果

1 选择OBS存储  创建OBS存储，完成创建后点击刷新按钮。

2 存储路径

文件名

3 根据执行结果状态筛选

4 选择备份的执行结果

任务名称	执行状态	创建时间
<input type="checkbox"/> gatk-lm-obs-sidecar-posix-08-06-161046	已停止	2019/08/06 16:11:09
<input type="checkbox"/> abcd-test-08-06-150822	成功	2019/08/06 15:10:23
<input type="checkbox"/> obs-util-08-05-204153	成功	2019/08/05 20:43:33
<input type="checkbox"/> nginx-wf-vah5-08-05-191341	成功	2019/08/05 19:14:19
<input type="checkbox"/> lh-checkresultlogtime-set-08-05-174401	失败	2019/08/05 17:45:29

5 总条数: 207 < 1 2 3 4 5 ... 42 >

5 确认 取消

- 选择OBS存储。选择待备份执行结果存储的OBS存储。可在下拉框中选择已有存储，或是创建新的OBS存储。
- 存储路径。待备份执行结果在OBS存储中的存储路径，建议存储在“backup/”路径下。斜杠（/）表示分隔并创建多层级文件夹。
- 文件名。系统指定，无法修改。备份完成后，可以通过该文件名在存储路径中找到备份的CSV文件。
- 选择通过执行结果状态过滤执行结果（上图中的3），从而快速备份特定状态的执行结果。

**步骤3** 备份成功后，界面将提示备份成功，您可以通过界面提示链接“数据存储列表”，或是在“数据管理”>“私有数据”查看备份结果。

通过**步骤2**中的文件名找到备份文件，单击“操作”中的“下载”，即可查看备份的执行结果数据。

结果数据中各列的数据说明如下，以下从A列起顺序说明。

- execution\_name: 任务名称。
- guid: 任务ID。
- project\_id: 项目ID，代表任务所属华为云区域（项目）。项目ID与区域（项目）的对应关系，请参见[获取项目ID](#)。
- status: 任务执行状态。
- workflow\_id: 任务所使用的流程的ID。
- workflow\_name: 任务所使用的流程名称。
- time\_duration: 任务耗时，单位为秒。
- environment: 任务所使用的环境，如果为CCI环境，则表示命名空间名称；如果为CCE环境，则表示集群名称。

- inputs: 任务输入参数。
- outputs: 任务输出参数。
- batch\_name: 批次名称。
- resources: 资源消耗。
- tae\_resources: 重试资源消耗。仅失败任务重试此项有数值。
- job\_status: 子任务状态。
- create\_at: 任务创建时间。
- update\_at: 任务状态最后更新时间。

----结束

# 7 参考基因组

基因参考组由正常和病理条件下非正常的分子表达所组成。基因容器提供主流基因参考组，作为基因分析参考。

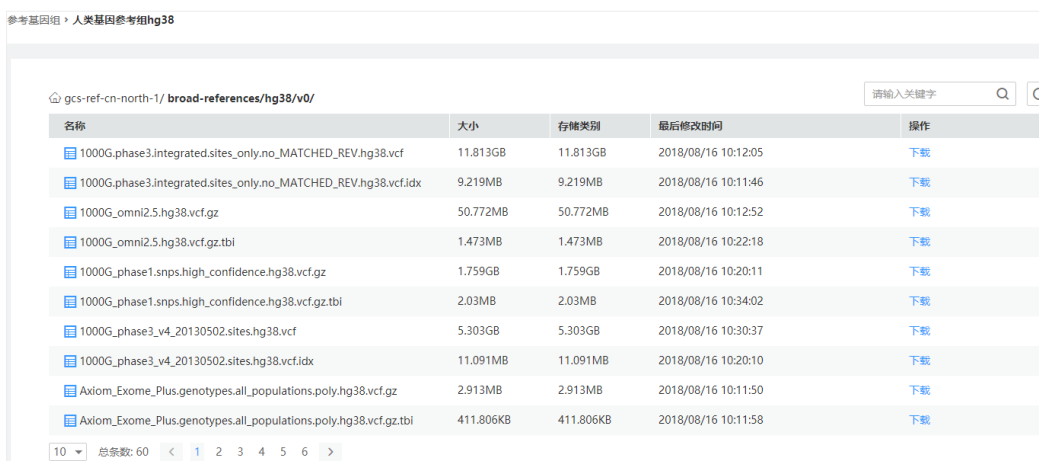
**步骤1** 登录[GCS控制台](#)，左侧导航栏中选择“参考基因组”。

图 7-1 参考基因组页面



**步骤2** 在右侧页面单击“查看”，可查看基因参考组的详细情况。

图 7-2 基因参考组详情



**步骤3** 单击“下载”，可下载参考组文件。当需要批量下载多个文件或者下载文件大于5MB时，请使用客户端工具OBS Browser。

----结束

# 8 权限管理

## 8.1 委托权限

### 须知

为正常提供服务功能，GCS老的委托方式是要求客户创建名为 `servicestage_admin_trust` 的委托，提供 `op_svc_servicestage` 账号获取创建资源的能力。

但由于 `op_svc_servicestage` 账号被多个服务共享，将导致您无法精细化控制对特定服务的授权行为，即授权 `servicestage` 的同时也会授权 GCS，反之亦然。

为解决该问题，GCS服务提供了新的委托方式，后续若您不再使用GCS服务，可直接删除委托“`gcs_admin_trust`”，该操作将只影响GCS，不影响您使用其他服务。

由于GCS在运行中对存储、容器等各类云服务资源都存在依赖关系（如：按照您的流程在指定的环境中下发任务）。因此在您使用GCS服务前，需要在IAM中创建名为“`gcs_admin_trust`”委托并授予“`op_svc_gcs_container`”账号Tenant Administrator权限，即允许GCS服务使用账户“`op_svc_gcs_container`”对您的其他云服务资源进行操作，从而更好地为您提供服务。该委托将仅用于GCS服务对所依赖的其他云服务资源进行调用。

具体操作步骤如下：

- 步骤1** 使用“管理员账号”登录[统一身份认证服务控制台](#)。
- 步骤2** 在统一身份认证服务的左侧导航窗格中，选择“委托”页签，单击右上方的“创建委托”。

图 8-1 创建委托



步骤3 在创建委托页面，设置委托名称为“gcs\_admin\_trust”。

步骤4 “委托类型”选择“普通帐号”，在“委托的帐号”中输入需要建立委托关系的帐号名“op\_svc\_gcs\_container”。

图 8-2 云服务委托名称



步骤5 选择“持续时间”，填写“描述”信息。

步骤6 单击“下一步”，进入给委托授权页面。

步骤7 勾选“Tenant Administrator”权限，单击“下一步”。

图 8-3 选择策略



步骤8 选择权限的作用范围为“所有资源”。

图 8-4 授权范围



步骤9 单击“确定”，委托创建完成，服务即可使用。

----结束

如果您不再使用GCS服务，可直接删除委托“gcs\_admin\_trust”，该操作将只禁止GCS服务操作您的服务，不影响您使用其他服务。

## 8.2 创建用户并授权使用 GCS

如果您需要对您所拥有的GCS进行精细的权限管理，您可以使用[统一身份认证服务](#)（Identity and Access Management，简称IAM），通过IAM，您可以：

- 根据企业的业务组织，在您的华为云账号中，给企业中不同职能部门的员工创建IAM用户，让员工拥有唯一安全凭证，并使用GCS资源。



- 根据企业用户的职能，设置不同的访问权限，以达到用户之间的权限隔离。
- 将GCS资源委托给更专业、高效的其他华为云账号或者云服务，这些账号或者云服务可以根据权限进行代运维。

如果华为云账号已经能满足您的要求，不需要创建独立的IAM用户，您可以跳过本章节，不影响您使用GCS服务的其它功能。

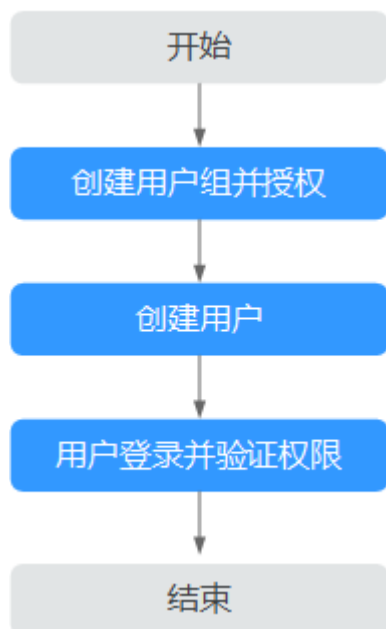
本章节为您介绍对用户授权的方法，操作流程如[示例流程](#)所示。

## 前提条件

给用户组授权之前，请您了解用户组可以添加的GCS系统策略，并结合实际需求进行选择，GCS支持的系统策略及策略间的对比，请参见：[GCS系统策略](#)。若您需要对除GCS之外的其它服务授权，IAM支持服务的所有策略请参见[权限策略](#)。

## 示例流程

图 8-5 给用户授权 GCS 权限流程



1. [创建用户组并授权](#)。  
在IAM控制台创建用户组，并授予基因容器服务只读权限“GCS ReadOnlyAccess”。
2. [创建用户并加入用户组](#)。  
在IAM控制台创建用户，并将其加入[1.创建用户组并授权](#)中创建的用户组。
3. [用户登录](#)并验证权限。  
新创建的用户登录控制台，验证基因容器服务的只读权限。

## 8.3 GCS 自定义策略

如果系统预置的GCS权限，不满足您的授权要求，可以创建自定义策略。自定义策略中可以添加的授权项（Action）请参考[权限策略和授权项](#)。

目前华为云支持以下两种方式创建自定义策略：

- 可视化视图创建自定义策略：无需了解策略语法，按可视化视图导航栏选择云服务、操作、资源、条件等策略内容，可自动生成策略。
- JSON视图创建自定义策略：可以在选择策略模板后，根据具体需求编辑策略内容；也可以直接在编辑框内编写JSON格式的策略内容。

具体创建步骤请参见：[创建自定义策略](#)。本章为您介绍常用的GCS自定义策略样例。

## GCS 自定义策略样例

- 示例1：授权创建和删除环境的权限

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "gcs:environments:create",
        "gcs:environments:delete"
      ]
    }
  ]
}
```

- 示例2：授权创建，查看流程，执行流程的权限

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "gcs:workflows:create",
        "gcs:workflows:list",
        "gcs:executions:create"
      ]
    }
  ]
}
```