

弹性云服务器

用户指南

文档版本 01
发布日期 2025-01-26



版权所有 © 华为云计算技术有限公司 2025。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云计算技术有限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，华为云计算技术有限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。

目录

1 通过 IAM 授予使用 ECS 的权限	1
1.1 创建用户并授权使用 ECS	1
1.2 ECS 自定义策略	2
2 实例	5
2.1 ECS 实例概述	5
2.2 选择 ECS 计费模式	7
2.2.1 包年/包月计费	7
2.2.2 按需计费	7
2.2.3 竞价计费	8
2.2.3.1 竞价计费型实例	8
2.2.4 预留实例	11
2.2.4.1 预留实例概述	11
2.2.4.2 开通和购买预留实例	16
2.2.4.3 修改预留实例属性	19
2.2.5 按需转包年/包月	20
2.2.6 包年/包月转按需	22
2.3 购买 ECS	24
2.3.1 ECS 购买方式导航	24
2.3.2 自定义购买 ECS	25
2.3.3 购买竞价计费型 ECS	50
2.3.4 购买竞享计费型 ECS	53
2.3.5 通过私有镜像购买 ECS	55
2.3.6 通过智能购买助手购买 ECS	57
2.3.7 通过共享 VPC 购买 ECS	62
2.3.8 相同配置 ECS	63
2.4 登录 Windows ECS	64
2.4.1 Windows ECS 登录方式概述	64
2.4.2 通过控制台 VNC 登录 Windows ECS	65
2.4.3 通过本地 Windows 主机登录 Windows ECS	66
2.4.4 通过本地 Linux 主机登录 Windows ECS	73
2.4.5 通过本地 macOS 主机登录 Windows ECS	75
2.4.6 通过移动设备登录 Windows ECS	78
2.5 登录 Linux ECS	83

2.5.1 Linux ECS 登录方式概述.....	83
2.5.2 通过 CloudShell 登录 Linux ECS.....	84
2.5.3 通过 VNC 登录 Linux ECS.....	88
2.5.4 通过 SSH 密钥方式登录 Linux ECS.....	90
2.5.5 通过 SSH 密码方式登录 Linux ECS.....	96
2.5.6 通过 macOS 主机登录 Linux ECS.....	98
2.5.7 通过移动设备登录 Linux ECS.....	99
2.6 管理 GPU 加速型 ECS 的 GPU 驱动.....	111
2.6.1 GPU 驱动概述.....	111
2.6.2 Tesla 驱动及 CUDA 工具包获取方式.....	112
2.6.3 手动安装 GPU 加速型 ECS 的 GRID 驱动.....	113
2.6.4 手动安装 GPU 加速型 ECS 的 Tesla 驱动.....	123
2.6.5 卸载 GPU 加速型 ECS 的 GPU 驱动.....	137
2.7 管理 ECS 配置信息.....	141
2.7.1 更改时区.....	141
2.7.2 开启/关闭 ECS 超线程.....	144
2.7.3 使用实例自定义数据和元数据.....	146
2.7.3.1 通过元数据获取 ECS 基本信息.....	147
2.7.3.2 通过实例自定义数据配置 ECS 实例.....	155
2.7.4 修改 ECS 名称.....	163
2.7.5 将 ECS 迁移至专属主机.....	163
2.7.6 管理云服务器组.....	164
2.7.7 配置同一 VPC 内 ECS 通过主机名进行通信.....	167
2.7.8 启、停弹性云服务器.....	168
2.8 变更 ECS 规格（vCPU 和内存）.....	170
2.8.1 变更单台 ECS 规格.....	170
2.8.2 XEN 实例变更为 KVM 实例（Windows）.....	174
2.8.3 XEN 实例变更为 KVM 实例（Linux-自动配置）.....	179
2.8.4 XEN 实例变更为 KVM 实例（Linux-手动配置）.....	183
2.9 变更 ECS 操作系统.....	187
2.9.1 重装操作系统.....	188
2.9.2 切换操作系统.....	189
2.10 查看弹性云服务器信息.....	193
2.10.1 查看创建状态.....	193
2.10.2 查看失败任务.....	193
2.10.3 查看弹性云服务器详细信息.....	194
2.10.4 导出弹性云服务器列表信息.....	195
2.10.5 搜索云服务器.....	195
3 镜像管理.....	203
3.1 镜像概述.....	203
3.2 创建镜像.....	205
4 磁盘管理.....	207

4.1 ECS 磁盘概述.....	207
4.2 为 ECS 新增磁盘.....	207
4.3 为 ECS 挂载磁盘.....	208
4.4 初始化数据盘.....	209
4.4.1 初始化概述.....	210
4.4.2 初始化 Linux 数据盘（容量小于等于 2TiB）.....	211
4.4.3 初始化 Linux 数据盘（容量大于 2TiB）.....	217
4.4.4 初始化 windows 数据盘.....	222
4.5 添加“包年/包月”磁盘.....	230
4.6 在线卸载磁盘.....	230
4.7 扩容云硬盘.....	233
4.8 扩容本地盘.....	233
4.9 开启高级磁盘功能.....	234
5 弹性网卡管理.....	236
5.1 弹性网卡概述.....	236
5.2 绑定弹性网卡.....	237
5.3 解绑弹性网卡.....	239
5.4 切换虚拟私有云.....	240
5.5 修改私有 IP 地址.....	241
5.6 管理虚拟 IP 地址.....	242
5.7 开启网卡多队列功能.....	248
5.8 开启网卡 IPv6 功能.....	252
5.9 动态获取 IPv6 地址.....	254
6 弹性公网 IP 管理.....	273
6.1 弹性公网 IP 概述.....	273
6.2 绑定弹性公网 IP.....	274
6.3 解绑弹性公网 IP.....	275
6.4 更换弹性公网 IP.....	276
6.5 修改弹性公网 IP 的带宽.....	277
6.6 无公网 IP 的弹性云服务器访问 Internet.....	278
7 安全管理.....	282
7.1 提升云服务器安全性方法概述.....	282
7.2 安全组.....	288
7.2.1 安全组概述.....	288
7.2.2 默认安全组和规则.....	291
7.2.3 安全组配置示例.....	294
7.2.4 配置安全组规则.....	299
7.2.5 更改安全组.....	304
7.3 主机安全.....	305
7.4 云堡垒机.....	307
7.5 管理项目和企业项目.....	311

7.6 敏感操作保护.....	312
8 备份管理.....	316
8.1 云备份概述.....	316
8.2 备份弹性云服务器.....	321
9 密码和密钥对管理.....	325
9.1 重置 ECS 密码.....	325
9.1.1 密码使用场景介绍.....	325
9.1.2 在控制台重置 ECS 密码.....	326
9.2 管理 ECS 一键式重置密码插件.....	328
9.2.1 获取 ECS 一键式重置密码插件.....	328
9.2.2 安装单台 ECS 的一键式重置密码插件（单台操作）.....	334
9.2.3 更新单台 ECS 的一键式重置密码插件（单台操作）.....	340
9.2.4 通过脚本批量更新 ECS 一键式重置密码插件（Linux）.....	344
9.2.5 通过脚本批量更新 ECS 一键式重置密码插件（Windows）.....	347
9.3 管理 ECS 密钥对.....	351
9.3.1 密钥对使用场景介绍.....	351
9.3.2 （推荐）通过管理控制台创建密钥对.....	352
9.3.3 通过 PuTTY Key Generator 工具创建密钥对.....	353
9.3.4 导入密钥对.....	357
9.3.5 获取和清除云服务器密码（Windows）.....	359
9.3.5.1 获取 Windows 弹性云服务器的密码.....	359
9.3.5.2 清除 Windows 弹性云服务器的密码.....	361
10 启动模板管理.....	362
10.1 启动模板概述.....	362
10.2 创建启动模板.....	362
10.3 管理启动模板.....	363
11 智能购买组管理.....	365
11.1 智能购买组概述.....	365
11.2 创建智能购买组.....	366
11.3 管理智能购买组.....	368
12 事件管理.....	370
12.1 事件概述.....	370
12.2 查询事件.....	373
12.3 响应事件.....	374
12.3.1 响应实例重部署事件.....	374
12.3.2 响应本地盘换盘事件.....	379
12.3.3 响应系统维护事件.....	386
12.4 换盘预处理.....	389
12.4.1 本地盘换盘预处理（磁盘增强型实例）.....	389
12.4.2 本地盘换盘预处理（超高 I/O 型实例）.....	392

12.4.3 本地盘换盘预处理（裸金属类型实例）	395
12.5 实例重部署预处理	398
13 操作系统 dump 管理	399
13.1 配置操作系统 dump	399
13.2 查看操作系统 dump	406
14 自助运维管理	408
14.1 配置 ECS 自助运维自定义策略	408
14.2 为 ECS 配置定时运维任务	410
15 资源与标签	415
15.1 标签管理	415
15.1.1 标签概述	415
15.1.2 添加标签	416
15.1.3 使用标签检索资源	419
15.1.4 删除标签	420
15.2 配额调整	422
16 使用 CES 监控 ECS	424
16.1 监控弹性云服务器	424
16.2 弹性云服务器支持的基础监控指标	425
16.3 弹性云服务器支持的操作系统监控指标（安装 Agent）	431
16.4 弹性云服务器支持的进程监控指标（安装 Agent）	472
16.5 设置告警规则	475
16.6 查看监控指标	476
17 使用 CTS 审计 ECS 操作事件	478
17.1 ECS 支持云审计的关键操作	478
17.2 查询 ECS 审计事件	479
18 QingTian Enclave 管理	483
18.1 QingTian Enclave 简介	483
18.1.1 什么是 QingTian Enclave	483
18.1.2 QingTian Enclave 概念	485
18.2 快速入门	487
18.3 典型使用案例	489
18.3.1 工作流程	489
18.3.2 构建 QingTian Enclave 镜像	490
18.3.3 启动 QingTian Enclave	491
18.4 密码学证明	491
18.4.1 PCR 简介	492
18.4.2 证明文档	493
18.4.3 文档签名验证	494
18.4.4 结合华为云 KMS 服务	495
18.5 QingTian Enclave 应用开发	495

18.5.1 Linux 系统上 QingTian Enclave 应用的开发.....	496
18.5.2 QingTian Enclave 网络代理工具.....	499
18.5.3 QingTian Enclave 日志转发工具.....	505
18.6 QingTian CLI (qt CLI)	510
18.6.1 安装 qt CLI.....	511
18.6.2 qt-enclave-env 工具介绍.....	511
18.6.3 qt enclave 子命令介绍.....	511
18.7 错误码.....	516
18.8 常见问题.....	519
18.8.1 通用类.....	519
18.8.2 开发部署类.....	521

1 通过 IAM 授予使用 ECS 的权限

1.1 创建用户并授权使用 ECS

如果您需要对您所拥有的ECS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ECS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ECS资源委托给更专业、高效的其他华为账号或者云服务，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ECS服务的其他功能。

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。

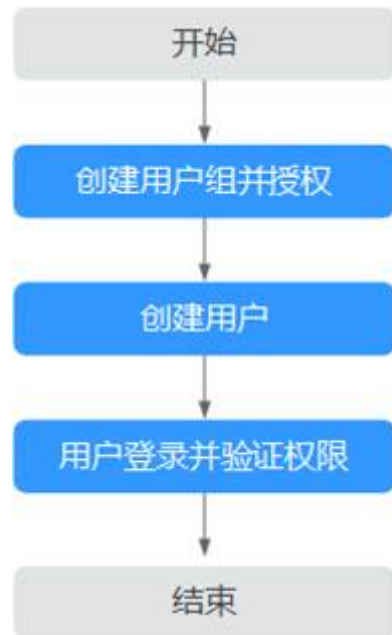
前提条件

给用户组授权之前，请您了解用户组可以添加的ECS系统策略，并结合实际需求进行选择。

ECS支持的系统策略及策略间的对比，请参见：[ECS系统策略](#)。若您需要对除ECS之外的其他服务授权，IAM支持服务的所有策略请参见[系统权限](#)。

示例流程

图 1-1 给用户授权 ECS 权限流程



1. 创建用户组并授权

在IAM控制台创建用户组，并授予弹性云服务只读权限“ECS 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中创建的用户组。

3. 用户登录并验证权限

新建的用户登录控制台，切换至授权区域，验证弹性云服务的只读权限。

- 在“服务列表”中选择弹性云服务器，进入ECS主界面，单击右上角“购买弹性云服务器”，尝试购买弹性云服务器，如果无法购买弹性云服务器（假设当前权限仅包含ECSReadOnlyAccess），表示“ECSReadOnlyAccess”已生效。
- 在“服务列表”中选择除弹性云服务器外（假设当前策略仅包含ECSReadOnlyAccess）的任一服务，若提示权限不足，表示“ECSReadOnlyAccess”已生效。

1.2 ECS 自定义策略

如果系统预置的ECS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《ECS API参考》中“策略及授权项说明”章节。

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ECS自定义策略样例。

ECS 自定义策略样例

- 示例1：授权仅允许用户批量开机、关机、重启云服务器

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ecs:cloudServerFlavors:get",
        "ecs:cloudServers:reboot",
        "ecs:cloudServers:start",
        "ecs:cloudServers:get",
        "ecs:cloudServers:list",
        "ecs:cloudServers:stop"
      ]
    }
  ]
}
```

- 示例2：授权仅允许用户批量关闭云服务器、删除云服务器

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ecs:cloudServers:get",
        "ecs:cloudServers:delete",
        "ecs:cloudServers:list",
        "ecs:cloudServers:stop"
      ]
    }
  ]
}
```

- 示例3：授权仅允许用户通过VNC进行登录

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ecs:cloudServerFlavors:get",
        "ecs:cloudServers:vnc",
        "ecs:cloudServers:get",
        "ecs:cloudServers:list"
      ]
    }
  ]
}
```

- 示例4：授权拒绝用户删除云服务器

拒绝策略需要同时配合其他策略使用，否则没有实际作用。用户被授予的策略中，一个授权项的作用如果同时存在Allow和Deny，则遵循Deny优先。

如果您给用户授予ECSFullAccess的系统策略，但不希望用户拥有ECSFullAccess中定义的删除云服务器权限，您可以创建一条拒绝删除云服务的自定义策略，然后将ECSFullAccess和拒绝策略授予用户，根据Deny优先原则，则用户可以对ECS执行除了删除云服务器外的所有操作。拒绝策略示例如下：

```
{
  "Version": "1.1",
```

```
"Statement": [  
  {  
    "Effect": "Deny",  
    "Action": [  
      "ecs:cloudServers:delete"  
    ]  
  }  
]
```


2 实例

2.1 ECS 实例概述

实例概述

实例即弹性云服务器，是由CPU、内存、操作系统、云硬盘组成的基础计算组件。

弹性云服务器创建成功后，您就可以像使用自己的本地PC或物理服务器一样，在云上使用弹性云服务器，打造一个高效、可靠、安全的计算环境。弹性云服务器的开通是自助完成的，您只需要指定CPU、内存、操作系统、规格、登录鉴权方式即可，同时也可以根据您的需求随时调整您的弹性云服务器规格。

云平台提供了多种实例类型供您选择，不同类型的实例可以提供不同的计算能力和存储能力。同一实例类型下可以根据CPU和内存的配置选择不同的实例规格。

- 关于实例类型的信息，请参考[实例类型](#)。
- 更多实例规格清单请参考[规格清单（x86）](#)和[规格清单（鲲鹏）](#)。

实例配置

ECS的配置信息包含CPU、内存、操作系统、存储容量、网络等基础资源。

表 2-1 实例配置

配置信息	说明	相关参考
规格	弹性云服务器提供x86和鲲鹏架构的多种类型的实例规格。 通过规格定义实例的vCPU个数、内存大小、内网基准/最大带宽、可达到的最大收发包能力、是否支持IPv6等配置。	<ul style="list-style-type: none">• 规格清单（x86）• 规格清单（鲲鹏）

配置信息	说明	相关参考
镜像	镜像包含了ECS运行所必需的操作系统、应用软件、初始化应用程序等。 镜像分为公共镜像、私有镜像、共享镜像、市场镜像。	镜像概述
存储	ECS通过挂载的云盘和本地盘提供存储能力： <ul style="list-style-type: none">云盘：通过专属分布式存储池或云硬盘创建，支持作为系统盘和数据盘。本地盘：部分规格自带本地盘，仅支持作为数据盘。	ECS磁盘概述
网络	虚拟私有云为ECS构建隔离的、用户自主配置和管理的虚拟网络环境，可以在VPC中定义网卡、私网IP地址段、安全组等网络特性。 不同VPC中的ECS网络默认不互通。	弹性网卡概述

实例选型

表 2-2 实例选型

选型方式	说明	相关参考
规格类型选型	提供按CPU架构类型、vCPUs个数和内存大小、规格名称模糊搜索以及实例类型的选型方式。 适用于实例类型，想要根据具体配置要求进行规格选型的用户。	自定义购买ECS
业务场景选型	提供按Web应用、网站应用/电商、游戏、数据库等业务场景以及不同场景业务量的推荐规格。 适用于有明确业务场景需求，想要根据业务场景、业务量等参考指标进行规格选型的用户。	

实例变更

ECS创建成功后，您可以根据业务需要变更ECS的vCPUs和内存、操作系统、存储和带宽。

- 变更vCPUs和内存：[变更单台ECS规格](#)。
- 变更操作系统：[切换操作系统](#)。

- 变更存储：[为ECS新增磁盘、扩容云硬盘](#)。
- 变更带宽：[修改弹性公网IP的带宽](#)。

2.2 选择 ECS 计费模式

2.2.1 包年/包月计费

什么是包年/包月

包年/包月的计费模式是一种预付费方式，按订单的购买周期计费，适用于可预估资源使用周期的场景，价格比按需计费模式更优惠。

更多计费相关信息，请参见[包年/包月](#)。

包年/包月计费模式弹性云服务器使用说明

1. “包年/包月”的弹性云服务器创建后不能删除，如需停止使用，请在弹性云服务器列表页，单击“操作”列下的“更多 > 退订”，执行退订操作。
2. 在创建弹性云服务器时购买的系统盘，卸载后，如果继续作为系统盘使用，则只能挂载给原弹性云服务器，不能挂载其他弹性云服务器的系统盘作为自己的系统盘；如果作为数据盘使用，则可以挂载给任意弹性云服务器。
3. 在创建弹性云服务器时购买的数据盘，卸载后，如果需要重新挂载，则只能挂载给原弹性云服务器作为数据盘使用。

适用包年/包月的资源

包年/包月的资源包括：

- ECS实例，包括vCPU和内存。
 - 镜像，包括预付费的市场镜像。
 - 购买云服务器时同时购买的云硬盘价格，包括系统盘和数据盘。
 - 购买云服务器时同时购买的带宽价格。
- IP和固定带宽统一计费，不单独计费，计费标准参考按固定带宽计费。

如果您要创建一个包年/包月的弹性云服务器，配置费用即包括了以上几种资源的使用。

您可以使用ECS[价格计算器](#)快速了解弹性云服务器配置费用。

2.2.2 按需计费

什么是按需计费

按需付费是后付费方式，可以随时开通/删除弹性云服务器，支持秒级计费，系统会根据云服务器的实际使用情况每小时出账单，并从账户余额里扣款。

更多计费相关信息，请参见[按需计费](#)。

📖 说明

按需付费的弹性云服务器关机再次开机时，可能会出现由于资源不足引起的启动失败，请过一段时间再次启动，或更改弹性云服务器规格。

计费样例

ECS的“按需计费”是秒级计费，ECS产品价格详情中标出了每小时价格，您可以将每小时价格除以3600，即得到每秒价格。

示例，某一按需实例价格为0.68美元/小时，购买一台按需实例根据实际使用时长、按秒计费。

- 使用30分钟，根据实际使用时长按秒计费： $(0.68/3600) \times 30 \times 60=0.34$ 美元
- 使用1小时30分钟，根据实际使用时长按秒计费： $(0.68/3600) \times 90 \times 60=1.02$ 美元

适用按需计费的资源

按需计费的ECS资源包括：

- ECS实例，包括vCPU核数和内存容量。
- 镜像，包括市场镜像，以及来源于市场镜像的共享镜像或自定义镜像。
- 购买云服务器时同时购买的云硬盘价格，包括系统盘和数据盘。
- 购买云服务器时同时购买的带宽价格。

您可以使用ECS[价格计算器](#)快速了解弹性云服务器配置费用。

2.2.3 竞价计费

2.2.3.1 竞价计费型实例

什么是竞价计费型实例？

华为云将可用的计算资源，按照一定折扣进行售卖，其价格随市场供需关系实时变化，这种打折销售、价格实时变化的计费模式称为“竞价计费”。

采用竞价模式购买的弹性云服务器，称之为“竞价计费型实例”，简称“竞价实例”。

在该计费模式下，您可以以折扣价购买并使用弹性云服务器，性能与常规云服务器无异。但是当库存资源不足，或市场价格上浮、并超过您的预期价格时，系统会自动释放您的云服务器资源，对这些折扣售卖的弹性云服务器进行中断回收。与“按需计费”和“包年/包月”方式购买的弹性云服务器相比，在提供同等性能的前提下，竞价计费型实例可提供超低折扣，降低用云成本。

工作原理

不同时期，市场对某一规格的云服务器供需情况不同，从而引起市场价格的波动。您可以在市场价格较低时，购买并使用竞价计费型实例，降低计算成本。

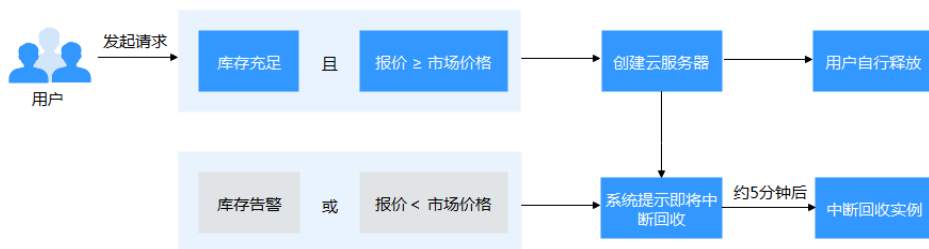
购买竞价计费型实例时，您需要为指定规格的弹性云服务器设定价格上限，即报价。该报价是您愿意为该弹性云服务器支付的最高价格，报价越高，购买成功的机会越大。

- 当“报价 \geq 市场价格”且库存资源充足时，购买竞价计费型实例成功。系统将按当前市场价格进行计费。
- 当“报价 $<$ 市场价格”时，无法购买竞价计费型实例。

竞价计费型实例购买成功后，您可以像使用其他计费模式的云服务器一样，登录并使用弹性云服务器，没有任何差别。但是，系统会定期比价，并检查资源池库存情况：

- 当“报价 \geq 市场价格”且库存资源充足时，您可以继续持有并使用弹性云服务器。
- 当“报价 $<$ 市场价格”或库存资源不足时，系统发出即将释放云服务器资源的通知，并在约5分钟后自动删除弹性云服务器。

图 2-1 竞价计费型实例生命周期



应用场景

适用场景

如下业务适合使用竞价计费型实例策略：图像渲染、无状态Web服务、基因测序、离线分析、函数计算、批量计算、样本分析、CI/CD与测试等。

说明

当市场价格高于您的报价或者库存资源不足时，竞价计费型实例会被中断回收，请务必做好数据备份工作。

不适用场景

为避免中断回收机制对业务带来影响，对于需要长时间作业、或稳定性要求极高的服务，请勿采用竞价计费型实例策略。

使用须知

- 当前仅支持虚拟化类型为KVM的弹性云服务器使用竞价计费型实例策略。具体的规格支持清单，请以控制台显示的资源为准。
- 在不同可用区，同一规格的云服务器市场价格变化可能不同。
- 使用竞价计费型实例策略的弹性云服务器，不支持切换操作系统功能。
- 使用竞价计费型实例策略的弹性云服务器，不支持自动恢复功能。
- 使用竞价计费型实例策略的弹性云服务器，不支持变更规格。
- 不支持使用市场镜像创建竞价计费型实例。
- 竞价计费型实例不支持转“包年/包月”方式。
- 竞价计费型实例不支持卸载系统盘。
- 竞价计费型实例在中断回收时：
 - 不允许使用该竞价计费型实例制作系统盘镜像和整机镜像，但允许使用该云服务器的数据盘制作数据盘镜像。

- 不允许删除。
- 竞价计费型实例在中断回收时，数据盘和弹性公网IP默认不释放。请设置通知提醒，来决定是否释放关联的数据盘和弹性公网IP。具体方法，请参见“[购买竞价计费型实例](#)”。

计费规则

请认真阅读竞价计费型实例的计费规则，详情请参见“[竞价计费（竞价模式）](#)”。

计费样例

- **样例1：当市场价高于用户出价，竞价计费型实例被释放，根据用户实际使用时长、按秒计费，费用单价为成交时的市场价格。**
8: 30 市场价0.02美元/小时，用户出价0.04美元/小时，按照市场价格计费0.02美元/小时
9: 00 市场价0.03美元/小时
10: 00 市场价0.04美元/小时
10: 30 市场价0.05美元/小时，高于用户的出价，标记释放，提醒用户即将中断
用户共产生3个账单周期，实际扣费明细如下：
8: 30~9: 00 使用30分钟，按秒计费= $0.02/3600*30*60=0.01$ 美元
9: 00~10: 00 使用1小时，按9:00起始成交价0.03美元/小时，计费= 0.03 美元/小时*1小时= 0.03 美元
10: 00~10: 30 使用30分钟，按秒计费= $0.04/3600*30*60=0.02$ 美元
合计0.06美元，实际使用时长2小时。
- **样例2：竞价计费型实例资源不足的情况下，竞价计费型实例被释放，根据用户实际使用时长、按秒计费，费用单价为成交时的市场价格。**
8: 30 市场价0.02美元/小时，用户出价0.06美元/小时，按照市场价格计费0.02美元/小时
9: 00 市场价0.03美元/小时
10: 00 市场价0.04美元/小时
10: 30 市场价0.05美元/小时，虽然市场价低于用户出价，但由于竞价计费型实例资源不足，标记释放，提醒用户即将中断
用户共产生3个账单周期，实际扣费明细如下：
8: 30~9: 00 使用30分钟，按秒计费= $0.02/3600*30*60 = 0.01$ 美元
9: 00~10: 00 使用1小时，按9:00起始成交价0.03美元/小时，计费= 0.03 美元/小时*1小时= 0.03 美元
10: 00~10: 30 使用30分钟，按秒计费= $0.04/3600*30*60=0.02$ 美元
合计0.06美元，实际使用时长2小时。

购买竞价计费型实例

您可以通过控制台方式或调用API方式，购买竞价计费型实例。

- 控制台方式，请参见“[购买竞价计费型实例](#)”。
- API方式，请参见《弹性云服务器API参考》的“[创建云服务器](#)”章节。

中断回收

华为云始终可能会中断您的竞价计费型实例，终止实例的使用。竞价计费型实例在中断回收时，不允许使用该竞价计费型实例制作镜像。

中断回收竞价计费型实例的可能原因包括：

- 市场价格高于您的报价
- 库存资源不足

📖 说明

- 由于库存资源不足或者用户出价低于市场价导致的首小时内中断，竞价计费型实例不收费。
- 竞价计费型实例的第一个结算周期内（结算周期为自然小时）均按开机计费。
- 如果竞价计费型实例中断回收的5分钟内跨过了整点，那么整点后该竞价计费型实例按照市场价格计费。
- 开机过程中，竞价计费型实例每小时更新一次价格。重启或关机再开机后，竞价计费型实例以开机时的市场价格计费。

因此，请务必做好数据备份工作。系统会提前发出即将释放云服务器资源的通知，如需设置通知提醒，请参见“[购买竞价计费型实例](#)”。

常见问题

请参见[FAQ](#)。

2.2.4 预留实例

2.2.4.1 预留实例概述

什么是预留实例

预留实例（Reserved Instance，简称RI）并非真正的实例，而是一种可以应用于按需实例的折扣券。当您购买的按需实例和预留实例的[属性匹配](#)时，按需实例即可享受账单折扣。预留实例与按需实例组合使用即可以充分利用按需资源的灵活性，同时较按需购买资源更加优惠。

📖 说明

- 无论您的预留实例是否和按需实例组合使用，预留实例都将正常计费。
- 预留实例暂不支持与SQL Server组合使用。
- 预留实例当前处于公测阶段，申请公测请单击[预留实例公测](#)。

表 2-3 预留实例与按需实例、包年/包月实例对比

功能对比	基本概念	使用方法
预留实例	应用于按需实例的折扣券。	需匹配按需实例使用，当购买的预留实例与按需实例 属性匹配 时，按需实例即可享受账单折扣。

功能对比	基本概念	使用方法
按需实例	按需计费方式的弹性云服务器资源。按实际使用时长计费，可以随时开通/删除弹性云服务器。	按需实例是由CPU、内存、操作系统、云硬盘组成的最基础的计算组件，购买成功后就可以在云上使用弹性云服务器。
包年/包月实例	包年/包月计费方式的弹性云服务器资源，按订单的购买周期计费，适用于可预估资源使用周期的场景。	包年/包月实例是由CPU、内存、操作系统、云硬盘组成的最基础的计算组件，购买成功后就可以在云上使用弹性云服务器。
竞价计费型实例	采用竞价计费模式购买的弹性云服务器，称之为“竞价计费型实例”。	竞价计费型实例是由CPU、内存、操作系统、云硬盘组成的最基础的计算组件，购买成功后就可以在云上使用弹性云服务器。

- 购买预留实例详细操作步骤请参考：[开通和购买预留实例](#)
- 修改预留实例详细操作步骤请参考：[修改预留实例属性](#)

什么是预留实例与按需实例属性匹配

我们定义在区域内（未指定可用区）购买的预留实例为区域级预留实例；可用区内购买的预留实例为可用区级预留实例。

- 区域级预留实例的属性匹配：是指区域级预留实例与按需ECS实例的规格系列、CPU/内存比和操作系统三个参数完全一致。
当您变更ECS实例的规格或操作系统，并想要继续享受相同优惠时，需重新购买与ECS实例属性匹配的预留实例。
- 可用区级预留实例属性匹配：是指可用区级预留实例与按需ECS实例的规格名称和操作系统两个参数完全一致。
当您变更ECS实例的规格或操作系统，并想要继续享受相同优惠时，需重新购买与ECS实例属性匹配的预留实例。

应用场景

如果是短期使用完就释放云服务器，我们推荐您使用按需购买的方式，如果计划使用云服务器时间接近一年和三年，推荐您使用预留实例。预留实例可以为与其属性匹配按需实例提供折扣。

例如，如果购买可用区1中的2个Linux操作系统，租期为1年的s3.2xlarge预留实例，则可用区1中最多两个正在运行的Linux操作系统的s3.2xlarge按需实例可享受预留实例折扣。

工作原理

您的账号中有一个正在运行的按需实例，如果购买了一个与该按需实例属性相匹配的预留实例，按需实例在预留实例生效后即可享受购买优惠（预留实例购买成功后，从下个整点开始生效）。

[表2-4](#)列举了预留实例的属性，您在购买时可以根据这些属性筛选预留实例：

表 2-4 预留实例的属性

参数名称	说明
区域/可用区	<ul style="list-style-type: none">区域：区域内（未指定可用区）购买预留实例不支持预留容量。可用区：可用区内购买预留实例支持预留容量。
规格	<ul style="list-style-type: none">购买区域级预留实例的规格应与按需实例系列、CPU/内存比一致。购买可用区级预留实例应与按需实例的规格一致。 <p>说明 开通和购买预留实例后，不支持直接修改规格，但可以对预留实例进行拆分或合并，详细内容请参见修改预留实例属性。</p>
操作系统	开通云服务器时必须选择与您的预留实例的操作系统相匹配的镜像。例如，如果选择使用Linux操作系统的预留实例，请选择Linux操作系统的公共镜像或私有镜像开通云服务器。
期限	预留实例的购买期限。一年定义为 31536000 秒 (365 天)。
类别	标准：预留实例在购买期限内支持修改部分属性，例如实例大小；但不支持修改实例类型。
预付方式	无预付。

可用区级预留实例

如果是可用区内购买的预留实例，即AZ级预留实例，则可为该可用区中与预留实例同规格名称和同操作系统的按需实例提供折扣。

例如，如果购买可用区一中的2个Linux操作系统的、租期为1年的c3.xlarge.2预留实例，则可用区一中最多2个正在运行的Linux操作系统的c3.xlarge.2按需实例可享受折扣。

区域级预留实例

如果是区域内购买预留实例，具备以下特点：

- 可用区灵活性：区域内购买的预留实例可以为该区域内任何可用区的按需实例提供折扣。
- 实例大小灵活性：实例大小灵活性是指当开通与预留实例同规格系列、同CPU/内存比和同操作系统的云服务器时，系统自动匹配预留实例使用折扣。实例大小灵活性取决于实例大小的标准因子。实例大小灵活性不适用于可用区级预留实例。

实例大小灵活性基于标准因子应用于实例系列中各种规模的实例（从最小到最大）。表 2-5描述了实例类型中的各种规格及相应的每小时标准因子。

说明

仅在开通与区域级预留实例同规格系列、同CPU/内存比和同操作系统的云服务器时，系统自动匹配预留实例使用折扣。

例如：由于CPU/内存比不同，区域级的预留实例c3.large.4不能匹配c3.large.2的云服务器使用。

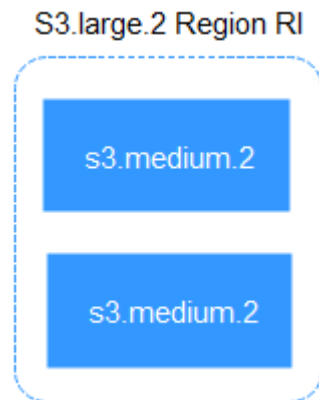
表 2-5 标准因子

实例规格	标准因子
small	1
medium	1
large	2
xlarge	4
2xlarge	8
4xlarge	16
6xlarge	24
7xlarge	28
8xlarge	32
9xlarge	36
12xlarge	48
14xlarge	56
15xlarge	60
16xlarge	64
26xlarge	104
52xlarge	208
nxlarge	n x 4

例如，s3.large.2实例具有标准因子2。如果购买了1个华为云香港Region中Linux操作系统、租期为1年的s3.large.2区域级预留实例。

- 如果在该区域中有两个正在运行的Linux操作系统s3.medium.2按需实例，则可将预留实例的优惠全部应用于这两个实例。

图 2-2 预留实例的应用举例 1



- 如果在该区域中有一个正在运行的Linux操作系统s3.xlarge.2按需实例，s3.xlarge.2实例具有标准因子4，则按需实例中的50%可应用预留实例的优惠。

图 2-3 预留实例的应用举例 2

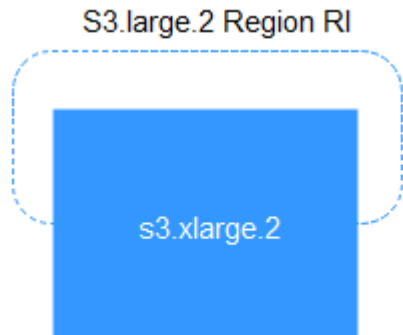


表 2-6 区域级预留实例与可用区级预留实例对比

预留实例类型	可用区灵活性	实例大小灵活性	资源预留
区域级预留实例	支持 区域内购买的预留实例可以为该区域内任何可用区的按需实例提供折扣。	支持 当开通与预留实例同规格系列、同CPU/内存比和同操作系统的云服务器时，系统自动匹配预留实例使用折扣。	不支持 实例资源不足时无法创建相应的按需实例。
可用区级预留实例（不推荐）	不支持 必须匹配相同可用区下的按需实例。	不支持 必须匹配同规格名称和同操作系统的按需实例。	支持 可预留实例资源。

预留实例的使用样例

如果用户当前具有以下按需实例：

- 在可用区1中，5个s3.large.2的Windows操作系统云服务器。
- 在可用区2中，3个m3.xlarge.2的Windows操作系统云服务器。
- 在可用区3中，1个c3.xlarge.2的Windows操作系统云服务器。

您在区域A中购买以下预留实例：

- 在可用区1中，5个s3.large.2预留实例，操作系统为Windows，租期为1年。
- 在区域A中，6个m3.large.2的区域级预留实例，操作系统为Windows，租期为1年。
- 在区域A中，1个c3.large.2的区域级预留实例，操作系统为Windows，租期为1年。

预留实例的应用方式如下：

- 5个s3.large.2可用区级预留实例应用于5个s3.large.2实例使用，因为它们之间的属性（可用区，操作系统类型，实例类型）相匹配。
- m3.large.2区域级预留实例具备可用区灵活性和实例大小灵活性。
m3.large.2等效于2个标准因子。6个m3.large.2区域级预留实例等于12个标准因子（6 x 2）。当前账号有3个正在运行的m3.xlarge.2实例，等效于12个标准因子（3 x 4）。这种情况下，6个m3.large.2区域级预留实例可以匹配3个m3.xlarge.2实例使用。
- 1个c3.large.2的区域级预留实例，具备可用区和实例大小灵活性，可应用于c3.xlarge.2按需实例。
c3.large.2预留实例等效于2个标准因子（1 x 2），而c3.xlarge.2按需实例需要4个标准因子（1 x 4）的预留实例来匹配。因此，c3.xlarge.2按需实例的50%的费用可应用预留实例的折扣，剩余的c3.xlarge.2用量按需计费，不可应用预留实例的折扣。

2.2.4.2 开通和购买预留实例

预留实例并非真正的实例，而是一种可以应用于按需实例的账单折扣。当您购买的按需实例和预留实例的属性相匹配时，按需实例即可享受账单折扣。

预留实例适用于可预估资源使用周期的场景，当按需实例和预留实例的属性相匹配时，按需实例即可享受账单折扣。

- 预留实例产品介绍请参考：[预留实例概述](#)
- 修改预留实例详细操作请参考：[修改预留实例属性](#)



约束与限制

- 您可在预留实例列表页左上方查看当前区域的预留实例配额。当前每个用户在每个区域预留实例配额为20。
- 系统会在每个月重置预留实例配额。
- 预留实例配额仅在购买时减少，变更、拆分、合并、退订不改变当前剩余配额（剩余配额=总配额-已使用配额）。

开通预留实例

购买预留实例前，请先联系客服为您开通预留实例购买权限。

购买预留实例

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 单击左侧导航栏“预留实例”。
5. 单击“购买预留实例”。
系统进入购买页。
6. 确认当前“区域”。
如果所选区域没有符合您需求的预留实例，请重新选择区域。

7. （可选）如需查看支持预留容量的预留实例，勾选“显示预留容量的产品”，则系统将提示您选择预留实例的可用区。
 - 可用区内购买预留实例支持预留容量。
 - 区域内（未指定可用区）购买预留实例不支持预留容量。
8. （可选）如需购买支持预留容量的预留实例，请选择“可用区”。

您可以选择购买指定可用区的预留实例，可用区内购买预留实例支持预留容量。
9. 选择“规格类型”。

云服务平台提供了多种类型的预留实例供您选择，针对不同的应用场景，可以选择不同规格的预留实例。
10. 规格筛选。

可通过“规格筛选”功能，输入规格名称、操作系统、预留实例购买期限、预留实例类别、预付方式来搜索目标规格。
详细参数说明请参考[表2-7](#)

表 2-7 预留实例的属性

参数名称	说明
区域/可用区	<ul style="list-style-type: none">• 区域：区域内（未指定可用区）购买预留实例不支持预留容量。• 可用区：可用区内购买预留实例支持预留容量。
规格	<ul style="list-style-type: none">• 购买区域级预留实例的规格应与按需实例系列、CPU/内存比一致。• 购买可用区级预留实例应与按需实例的规格一致。 <p>说明 开通和购买预留实例后，不支持直接修改规格，但可以对预留实例进行拆分或合并，详细内容请参见修改预留实例属性。</p>
操作系统	开通云服务器时必须选择与您的预留实例的操作系统相匹配的镜像。例如，如果选择使用Linux操作系统的预留实例，请选择Linux操作系统的公共镜像或私有镜像开通云服务器。
期限	预留实例的购买期限。一年定义为 31536000 秒 (365 天)。
类别	标准：预留实例在购买期限内支持修改部分属性，例如实例大小；但不支持修改实例类型。
预付方式	无预付。

11. 选择规格。

云服务平台提供了多种类型的预留实例供您选择，针对不同的应用场景，可以选择不同规格的预留实例。您可以在列表中查看已上线的预留实例规格。

有效费率：预留实例的总费用分摊到整个期限内按需实例的每小时单价。

每台预付价格：购买预留实例需预先支付的费用。

每小时费率：预留实例的总费用减去预付费用后分摊到整个期限内按需实例的每小时费用。

12. 设置“预留实例名称”。

名称可自定义，但需符合命名规则：只能由中文字符、英文字母、数字及“_”、“-”组成。输入长度的范围为1-128字符。

13. 设置您购买预留实例的数量。

- 购买量：系统会显示您当前还可以购买的预留实例数量。
- 标准因子总数：标准因子由所购买的预留实例规格决定，可用于衡量预留实例大小的灵活性。
- 预付费用+按需费用：用户购买当前配置预留实例的价格包含预付费费用和按需费用两部分。预付费为预留实例每台预付价格乘以台数。按需费用为每台按需使用支付的费用乘以台数。

具体价格详情，可单击“了解计费详情”进行查看。

14. 单击“立即购买”。

如果您对价格有疑问，可以单击“了解计费详情”来了解产品价格。

15. 在规格确认页面，您可以查看资源详情并提交订单。

如果您确认规格和价格无误，单击“去支付”，并根据界面提示完成支付。

16. 根据界面提示，返回预留实例列表查看已购买的预留实例。

后续操作

- **开通预留实例配套的按需资源：**

在预留实例的操作列下单击“开通云服务器”。系统会自动跳转至购买弹性云服务器页面，并默认选择的规格与预留实例规格一致。

📖 说明

- 如操作系统不一致、或者云服务器不是按需计费都会导致预留实例无法使用。当开通与预留实例同系列、同CPU/内存比的云服务器时，系统自动匹配预留实例使用折扣。
- **查看预留实例的使用情况：**
 - a. 在预留实例列表页，勾选待查看预留实例。
 - b. 在预留实例列表下方的延展视图中，显示已选择该预留实例。

图 2-4 已选择预留实例

名称ID	状态	规格	操作系统	区域	可用区	类型	期限	数量	剩余时间	操作
6-7a0 91154-495281772-0	生效中	2vCPUs 4GB c3.large.2	Linux	区域级	-	标准	1年	3	286天	开通云服务器 再次购买 退订
	生效中	2vCPUs 4GB c3.large.2	Linux	区域级	-	标准	1年	1	286天	开通云服务器 再次购买 退订
	生效中	2vCPUs 8GB s3.large.4	Linux	区域级	-	标准	1年	1	254天	开通云服务器 再次购买 退订
	生效中	1vCPUs 2GB s3.medium.2	Linux	区域级	-	标准	1年	2	249天	开通云服务器 再次购买 退订
	生效中	2vCPUs 4GB s3.large.2	Linux	区域级	-	标准	1年	1	254天	开通云服务器 再次购买 退订
	生效中	1vCPUs 2GB s3.medium.2	Linux	区域级	-	标准	1年	1	254天	开通云服务器 再次购买 退订
	生效中	1vCPUs 2GB s3.medium.2	Linux	区域级	-	标准	1年	1	254天	开通云服务器 再次购买 退订
	生效中	1vCPUs 2GB s3.medium.2	Linux	区域级	-	标准	1年	1	234天	开通云服务器 再次购买 退订
	生效中	1vCPUs 1GB s3.small.1	Linux	区域级	-	标准	1年	1	234天	开通云服务器 再次购买 退订
	生效中	1vCPUs 1GB s3.small.1	Linux	区域级	-	标准	1年	2	234天	开通云服务器 再次购买 退订

已选择: n-7a0

- c. 单击该预留实例的延展视图区域，展开预留实例详情。

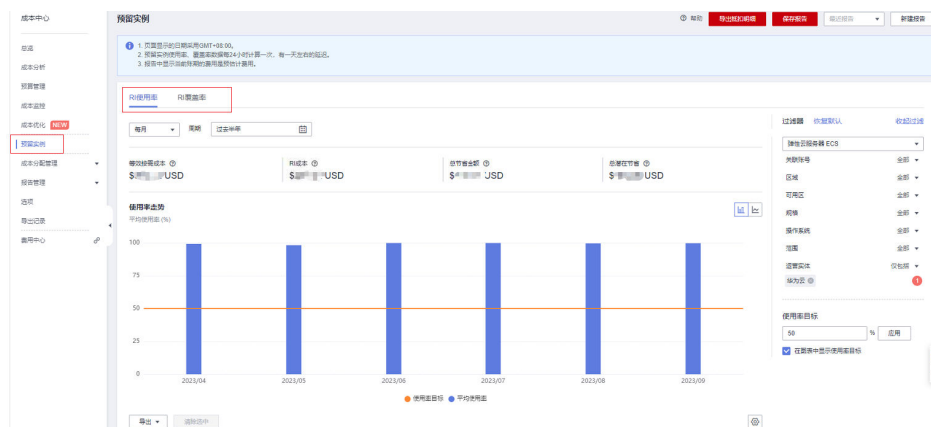
图 2-5 预留实例详情



名称	n7a0	ID	01154-495201172-0
状态	主态中	所属区域	华东-1
规格	2x CPU 4GB c1 large 2	可用区	可用区A
操作系统	Linux	类别	标准
标准因子	1	期限	1年
支付方式	后付费	剩余时间	266天
每小时价格	¥0.0000	创建时间	2023/07/17 20:00:00 GMT+08:00
每小时费用	¥0.0000	到期时间	2024/07/17 23:59:59 GMT+08:00
实例使用情况	查看实例使用		

- d. 在预留实例详情中，单击“查看RI使用率”，跳转至“成本中心”查看该预留实例的详细使用情况。

图 2-6 预留实例详细使用情况



2.2.4.3 修改预留实例属性

操作场景

当预留实例的类型不满足您当前的计算需求时，您可以修改预留实例的属性并将其继续应用于当前的按需实例。


预留实例类别为“标准”时，支持修改预留实例的可用区、范围、实例大小。

- 预留实例产品介绍请参考：[预留实例概述](#)
- 购买预留实例详细操作请参考：[开通和购买预留实例](#)

约束与限制

- 当预留实例的属性（操作系统、预付方式、类别、期限、过期时间、区域、同系列和CPU/内存比、折扣）全部相同时才可进行实例规格的合并。
- 修改前后标准因子总数需保持一致。
- 每次仅支持批量修改5组预留实例。
- 您可将1个预留实例拆分为多个，但多个预留实例只能合并为1个。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。

- 单击“☰”，选择“计算 > 弹性云服务器 > 预留实例”。
- 在预留实例的列表页，勾选待修改的预留实例，单击左上方的“修改预留实例”。
- 在预留实例的修改列表中分别显示修改前与修改后的预留实例的属性，根据需要修改预留实例的各个属性：

表 2-8 修改预留实例的常用操作

可修改属性	说明
拆分或合并	例如当前账号下有6个s3.xlarge.2的预留实例，一个s3.xlarge.2的标准因子为4，那么当前6个预留实例等效于24个标准因子。那么可以将其合并为3个s3.2xlarge.2的预留实例。或拆分为24个s3.medium.2的预留实例。
区域级预留实例与可用区级预留实例的相互转换	如果购买的预留实例为区域级预留实例，那么可以通过修改预留实例属性，修改为可用区级预留实例。

须知

标准因子总数是指预留实例的个数乘以预留实例的标准因子。预留实例修改前后标准因子总数需保持不变。

例如修改前：6个s3.large.4的预留实例，修改前标准因子总数为 $6 \times 2 = 12$ 。可以拆分为2个s3.xlarge.4的预留实例和4个s3.medium.4的预留实例，修改后标准因子总数为 $2 \times 4 + 4 \times 1 = 12$ 。

- 确认修改后的预留实例属性后单击“提交申请”。

2.2.5 按需转包年/包月

操作场景

- 按需计费：按需计费是后付费模式，按弹性云服务器的实际使用时长计费，可以随时开通/删除弹性云服务器。
- 包年/包月：包年/包月是预付费模式，按订单的购买周期计费，适用于可预估资源使用周期的场景，价格比按需计费模式更优惠。

如果您需要长期使用当前弹性云服务器，可以将按需购买的云服务器转为包年/包月计费模式，节省开支。按需转包年/包月的具体操作，请参见本节内容。

说明

- 如果弹性云服务器的绑定资源（云硬盘等磁盘、弹性公网IP）支持转包年/包月，则随弹性云服务器同步转换计费模式。
- 如果弹性云服务器的绑定资源（云硬盘等磁盘、弹性公网IP）不支持转包年/包月，则保留原计费模式，详细内容，请参见[绑定资源计费模式转换规则](#)。
- ISO镜像创建的弹性云服务器不支持按需转包年/包月。

绑定资源计费模式转换规则

弹性云服务器的绑定资源主要包括各类磁盘和弹性公网IP，当按需计费的弹性云服务器转包年/包月时，绑定资源的计费模式转换规格如表2-9和表2-10所示。

表 2-9 磁盘计费模式转换规则

磁盘类型	计费模式	是否共享	是否支持随ECS按需转包年/包月	处理措施
本地盘	不涉及	否	不涉及	无。
DSS/DESS下磁盘	包年/包月（存储池计费模式）	否	不涉及	无。
云硬盘	按需计费	否	是（极速型SSD V2类型磁盘不支持）	无。
云硬盘	按需计费	是	否	在控制台的云硬盘页面进行按需转包年/包月操作。 详细内容，请参见 云硬盘计费说明 。
云硬盘	包年/包月	否	否	已经是包年/包月计费模式，无需处理。
云硬盘	包年/包月	是	否	已经是包年/包月计费模式，无需处理。

表 2-10 弹性公网 IP 计费模式转换规则



资源	计费模式	计费方式	带宽类型	是否支持随ECS按需转包年/包月	处理措施
弹性公网IP	按需计费	按带宽计费	独享带宽	是	无。
弹性公网IP	按需计费	按流量计费	独享带宽	否	在控制台的弹性公网IP页面，先将按需按流量计费转换为按需按带宽计费，再将按需计费转换为包年/包月。 详细内容，请参见 变更弹性公网IP计费方式 。

资源	计费模式	计费方式	带宽类型	是否支持随ECS按需转包年/包月	处理措施
弹性公网IP	按需计费	按带宽计费	共享带宽	否	在控制台的弹性公网IP页面进行按需转包年/包月操作。 详细内容，请参见 变更弹性公网IP计费方式 。
弹性公网IP	包年/包月	按带宽计费	独享带宽/共享带宽	否	已经是包年/包月计费模式，无需处理。

前提条件

- 弹性云服务器的计费模式是“按需计费”。
- 弹性云服务器状态为“运行中”或“关机”。

操作步骤

- 登录管理控制台。
- 单击管理控制台左上角的，选择区域和项目。
- 单击“”，选择“计算 > 弹性云服务器”。
- 在弹性云服务器列表页，选中目标云服务器。
- 单击“操作”列下的“更多 > 转包年/包月”。

说明

系统支持批量转包年/包月操作，即一次将多台按需计费的弹性云服务器变更为包年/包月计费模式的。操作如下：

- 勾选待转包年/包月的多台弹性云服务器。
- 单击云服务器列表上方常用操作栏的“更多 > 转包年/包月”。
- 根据界面提示，确认转包年/包月的弹性云服务器信息后，选择弹性云服务器的购买时长，并支付订单。

2.2.6 包年/包月转按需

操作场景

包年/包月是预付费模式，按订单的购买周期计费，适用于可预估资源使用周期的场景。

如果您需要更灵活的计费方式，按照弹性云服务器的实际使用时长计费，您可以将实例的计费方式转为按需付费。

📖 说明


包年/包月转按需，支持两种方式：

- 即时转按需：按需的资费模式立即生效。
- 到期转按需：包年/包月资费模式到期后，按需的资费模式才会生效。


约束与限制

- 只有通过实名认证的客户，才可以执行包年/包月转按需操作。
- 在续费管理页面，订单状态是“使用中”的云服务器资源才能执行包年/包月转按需。
- 解决方案组合产品不支持包年/包月转按需。

即时转按需

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中，勾选需要进行操作的弹性云服务器。
5. 单击弹性云服务器列表左上角的“更多 > 即时转按需”。
6. 单击“是”，跳转至费用中心的“包年/包月转按需”页面。
7. 根据界面提示，勾选需要转按需的资源。
8. 确认退款信息后，单击“转按需”。
9. 再次根据界面信息确认要转按需的资源后，单击“确定”。

到期转按需（云服务器控制台）

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中，勾选需要进行操作的弹性云服务器。
5. 单击弹性云服务器列表左上角的“更多 > 到期转按需”。
6. 单击“是”，跳转至费用中心的“到期转按需”页面。
7. 根据界面提示，勾选需要转按需的资源。
8. 单击“到期转按需”。

到期转按需（费用中心）

1. 登录管理控制台。
2. 单击“控制台”页面上方“费用与成本”，在下拉菜单中选择“续费管理”。系统进入“续费管理”页面。
3. 自定义查询条件。
 - 可在“到期转按需项”页签查询已经设置到期转按需的资源。
 - 可对“手动续费项”、“自动续费项”、“到期不续费项”页签的资源设置到期转按需的操作。

图 2-7 续费管理



4. 设置包年/包月资源到期后转按需。

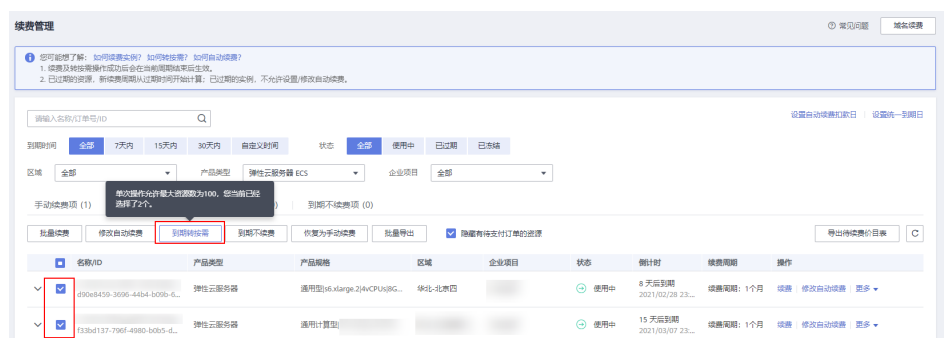
- 单个资源到期转按需：选择需要更改计费方式的弹性云服务器，单击操作列“更多 > 到期转按需”。

图 2-8 单个资源转按需



- 批量资源到期转按需：选择需要更改计费方式的弹性云服务器，单击列表左上角的“到期转按需”。

图 2-9 多个资源转按需



5. 查看资费变更的相关信息后，单击“转按需”。

2.3 购买 ECS

2.3.1 ECS 购买方式导航

云平台提供了多种购买ECS的方式，满足您在不同业务场景的需求。

表 2-11 实例购买方式

购买方式	适用场景
自定义购买ECS	购买ECS的通用操作指导，您可以根据业务场景灵活设置ECS的计费模式、规格类型、镜像类型、存储、网络、安全组、公网访问等配置信息。
购买竞价计费型ECS	购买竞价计费竞价模式ECS的操作指导，相当于以一定折扣购买按需计费ECS。 适用于对成本控制要求比较高，期望以较低折扣使用同等性能应用或服务的场景。
购买竞享计费型ECS	购买竞价计费竞享模式ECS的操作指导，相当于以一定折扣购买按需计费ECS。 适用于对成本控制要求比较高，期望以较低折扣使用同等性能应用或服务的场景。
通过私有镜像购买ECS	私有镜像包含操作系统、预装的公共应用以及用户的私有应用，通过私有镜像购买ECS可以节省重复配置的时间。 适用于有习惯使用的操作系统和应用程序的用户。
通过智能购买助手购买ECS	智能购买助手包含启动模板和智能购买组。 <ul style="list-style-type: none">启动模板：包含了用于创建ECS的除密码之外的任意配置信息，可以批量购买相同配置的ECS。智能购买组：通过自定义配置可以批量购买不同可用区、不同实例类型以及不同计费模式的ECS。
通过共享VPC购买ECS	通过其他账号共享的VPC和子网购买ECS，实现网络资源的共享和统一管理，提升资源管控效率、降低运维成本。
相同配置ECS	通过“购买相同配置”功能可以快速购买与当前配置相同的ECS，节省配置时间。

2.3.2 自定义购买 ECS

操作场景

弹性云服务器（Elastic Cloud Server，ECS）是由CPU、内存、操作系统、云硬盘组成的一种可随时获取、弹性可扩展的云服务器。弹性云服务器购买成功后，您就可以像使用自己的本地PC或物理服务器一样，在云上使用弹性云服务器。

您只需要指定CPU、内存、操作系统、规格、登录凭证方式即可获得一台弹性云服务器。

下面介绍如何在管理控制台自定义购买弹性云服务器。

操作流程

操作步骤	说明
准备工作	<ul style="list-style-type: none">注册华为账号并开通华为云，为账户充值。创建VPC、子网、安全组、密钥对、专属计算集群、CloudPond等相关资源。
步骤一：进入ECS购买页	进入“购买弹性云服务器”页面，开始设置购买参数。
步骤二：设置购买参数	根据业务需要，设置购买ECS的参数。
步骤三：确认配置并购买ECS	根据配置概要确认配置详情，完成ECS的购买。

准备工作

- 注册账号并实名认证。
在购买ECS之前，请先[注册华为账号并开通华为云、实名认证](#)。
如果您已开通华为云并进行实名认证，请忽略此步骤。
- 为账户充值。
请保证账户有足够的资金，以免购买ECS失败。具体操作，请参见[账户充值](#)。
- 规划VPC和子网等网络资源。
当您开始购买ECS时，系统会创建默认的VPC（vpc-default）和子网（subnet-default）。
如果您不想使用默认的VPC和子网，建议您提前在对应区域创建新的VPC和子网。
更多信息，请参见[虚拟私有云和子网规划建议](#)。
- 创建安全组并配置安全组规则。
当您开始购买ECS时，系统会创建默认的安全组（default、Sys-WebServer和Sys-FullAccess）。默认安全组的更多信息，请参见[默认安全组和规则](#)。
如果默认安全组和规则不能满足您的业务需要，建议您提前修改安全组规则。具体操作，请参见[配置安全组规则](#)。
- 创建密钥对。
如果采用密钥对方式登录ECS，建议您提前创建密钥对。具体操作，请参见[（推荐）通过管理控制台创建密钥对](#)。
- 创建专属物理资源。
如果您想要将ECS运行在隔离的物理硬件中，请先申请专属计算集群，再购买ECS。
具体操作，请参见[申请专属计算集群](#)。
- 注册边缘小站。
如果您想要将ECS部署在CloudPond上，需提前注册边缘小站。
具体操作，请参见[线上注册边缘小站](#)。

步骤一：进入 ECS 购买页

登录控制台，进入[购买弹性云服务器](#)页面。

步骤二：设置购买参数

请根据您的业务需求完成ECS购买参数的设置。

基础配置：计费模式、区域、可用区

1. 选择“计费模式”。

计费模式与ECS的付费方式、计费周期、资源状态变化等有关，请根据实际业务场景进行选择。

表 2-12 计费模式参数配置说明

选项	说明	场景及约束	更多参考
包年/包月	预付费模式，先付费再使用。 根据订单的购买周期计费。	适用于可预估资源使用周期的场景，价格比按需计费模式更优惠。对于长期使用用户，推荐该方式。 购买后如果不再使用，不能直接删除，需进行资源退订。	包年/包月
按需计费	后付费模式，先使用再付费。 根据实际使用时长秒级计费，按小时结算。	适用于计算资源需求波动的场景，可以随时开通，随时删除。	按需计费
竞价计费	后付费模式，先使用再付费，相对于按需计费有一定的折扣。 包含竞价模式和竞享模式，根据实际使用时长秒级计费，按小时结算。	适用于对成本控制要求比较高，期望以较低折扣使用同等性能应用或服务的场景。 存在中断机制，实例生命周期不可控，不适用需长时间作业、或稳定性要求极高的服务。	<ul style="list-style-type: none">• 竞价计费（竞价模式）• 竞价计费（竞享模式）

2. （可选）设置“预留实例”。

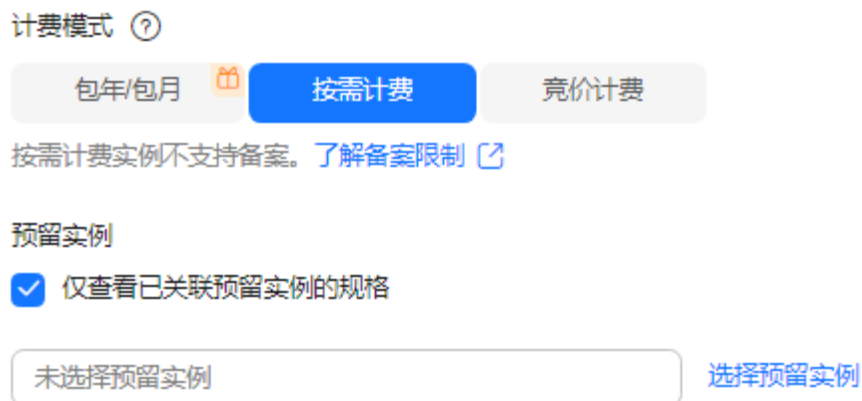
仅当“计费模式”选择“按需计费”，且开通了预留实例的公测时显示该参数。如果按需计费ECS需要和预留实例组合使用，可勾选“仅查看已关联预留实例的规格”并选择预留实例。

说明

预留实例当前处于公测阶段，申请公测请单击[预留实例公测](#)。

更多信息，请参见[预留实例概述](#)。

图 2-10 设置“预留实例”



- (可选) 设置“竞价类型”。
仅当“计费模式”选择“竞价计费”时显示该参数。支持设置为“竞价模式”或“竞享模式”。

图 2-11 设置“竞价类型”。



- 竞价模式：价格随市场价格波动，更多信息，请参见[购买竞价计费型ECS](#)。
 - 竞享模式：价格取决于保障周期，更多信息，请参见[购买竞享计费型ECS](#)。
- 选择“区域”。
区域指ECS的物理数据中心所在的位置，建议您就近选择靠近您业务的区域，可减少网络时延，提高访问速度。

📖 说明

- 不同区域的云服务产品之间内网互不相通。
 - ECS购买成功后，不支持直接切换区域，请谨慎选择。
 - 当您为CloudPond购买云服务器时，默认区域为当前小站所在的区域，且不支持更换为其他区域。
- 选择“可用区”。
可用区指在同一区域下，电力、网络隔离的物理区域，可用区之间内网互通，不同可用区之间物理隔离。

表 2-13 可用区参数配置说明

选项	说明	场景及约束	更多参考
随机分配	不同可用区上线的云服务器类型和规格不同。系统使用哈希算法基于用户的UUID (Universally Unique Identifier) 以及选择的规格, 分配ECS所在的可用区。	如果需要查看当前平台支持的全部云服务器类型和规格, 请选择“随机分配”。	区域和可用区
可用区n	当前区域支持的具体可用区, “n”表示可用区序号。	如果需要将ECS创建在指定可用区, 请选择具体的可用区。 <ul style="list-style-type: none"> 如果您需要提高应用的高可用性, 建议您将弹性云服务器创建在不同的可用区。 如果您需要较低的网络时延, 建议您将弹性云服务器创建在相同的可用区。 	
边缘可用区	边缘可用区部署在用户数据中心, 为CloudPond用户专属。	当您为CloudPond购买云服务器时, 默认可用区为“边缘可用区”, 且不支持更换为其他可用区。	边缘小站与区域和可用区是什么关系?

实例：规格选型、价格上限（竞价实例）、保障周期（竞享实例）

1. 选择实例选型方式。

表 2-14 实例参数配置说明

选项	说明	场景及约束	更多参考
规格类型选型	提供按不同规格属性的选型： <ul style="list-style-type: none"> • CPU架构：x86计算、鲲鹏计算。 • 实例筛选：支持按指定vCPU、内存或规格名称搜索规格。 • 规格类型：支持按实例类型和代系名称筛选规格。 	适用于熟悉CPU架构、vCPU、内存、实例类型和规格代系名称等属性, 想要根据具体配置要求进行规格选型的用户。	ECS实例概述

选项	说明	场景及约束	更多参考
业务场景选型	提供按业务场景、子场景及子场景业务量的推荐规格。	适用于有明确业务场景需求，想要根据业务场景、业务量等参考指标进行规格选型的用户。	

2. (可选) 选择“CPU架构”。
当实例选择“规格类型选型”方式时，该参数可见。

表 2-15 CPU 架构参数配置说明

选项	说明	场景及约束	更多参考
x86计算	采用复杂指令集CISC (Complex Instruction Set Computer)，支持几乎所有常用软件。 由于指令执行较为复杂所以每条指令花费的时间较长。	适用于Windows系列、仅x86兼容的商业软件等强平台相关场景。	规格清单 (x86)
鲲鹏计算	采用精简指令集RISC (Reduced Instruction Set Computer)，自研芯片，性价比更高。 因指令类型较少，所以能够以更高的速度执行操作，使计算机的结构更加简单、合理地提高运行速度。	适用于如下场景： <ul style="list-style-type: none"> ● 电商、大数据、科学计算等弱平台相关场景。 ● 手机仿真等ARM原生场景。 	规格清单 (鲲鹏)

3. (可选) 选择“业务场景”和“子场景”。
当实例选择“业务场景选型”方式时，该参数可见。

说明

不同区域、不同可用区上线的规格不同，推荐规格请以控制台显示为准。

表 2-16 业务场景选项

业务场景	子场景	说明	推荐规格 (示例)
Web应用	传统办公	高安全及高可靠性，预期并发访问量小于200，适用于OA、ERP、CRM等传统办公场景。	C7、C6
	企业展示型网站	高性价比，基本水平的vCPU性能、平衡的计算、内存和网络资源。	<ul style="list-style-type: none"> ● C7、C6 ● S7、S6

业务场景	子场景	说明	推荐规格（示例）
	个人应用搭建	高性价比，基本水平的vCPU性能、平衡的计算、内存和网络资源。	<ul style="list-style-type: none"> ● S7、S6 ● T6
	开发测试环境	基本水平的vCPU性能、平衡的计算、内存和网络资源，同时可根据工作负载的需要实现性能的突增，具有短期发挥更高性能的能力。	S7、S6
	前端服务器	基本水平的vCPU性能、平衡的计算、内存和网络资源，适合Apache、Nginx、IIS等前端服务场景。	S7、S6
	后端服务器	CPU/内存配比高，具备高性能、低时延、高性价比的特点。适用于Tomcat、JBoss等后端服务场景。	C7、C6、C6s
网站应用/电商	10万pv，1000活跃用户	高性价比，弹性可用的ECS可随取随用，避免资源浪费。	<ul style="list-style-type: none"> ● S7、S6 ● T6
	20万pv，2000活跃用户	在促销、秒杀、爆款、直播等场景下，满足电商网站对云服务器的高性能、快速弹性、高稳定的业务诉求。	<ul style="list-style-type: none"> ● C7、C6、C6s ● S7、S6
	50万pv，5000活跃用户	在促销、秒杀、爆款、直播等场景下，满足电商网站对云服务器的高性能、快速弹性、高稳定的业务诉求。	C7、C6、C6s
游戏	游戏	具有高性能、高稳定性、低时延、高性价比的特点，满足游戏业务对服务器性能、可靠性、网络、弹性伸缩的高要求。	C7、C6、C6s
数据库	计算	提供稳定的高性能计算能力。	C7、C6、C6s
	存储	自带高存储带宽和IOPS的本地盘，具有高存储IOPS以及读写带宽的优势，在海量数据存储场景下，具备更高的性价比。	<ul style="list-style-type: none"> ● M7、M6 ● D7、D6
	网络	数据访问量大，要求快速的数据交换和处理。依赖于较高的PPS和低网络时延，提供较高的TPS吞吐量。	<ul style="list-style-type: none"> ● E7、E6 ● M7、M6 ● C7、C6、C6s

业务场景	子场景	说明	推荐规格（示例）
数据分析	管理节点	高网络性能要求，调动大量的计算资源，提升数据处理速度。	C7、C6、C6s
	计算节点	高计算性能、高稳定性要求，各个节点的计算能力要保持均衡。	<ul style="list-style-type: none">• M7、M6• C7、C6、C6s• I7、I7、I3、I3• S7、S6
	存储节点	大量的数据读写，对存储带宽有很高诉求，在海量数据存储场景下，具备更高的性价比。	D7、D6、D3
高性能计算	高性能计算	计算资源使用量大。利用高性能云服务器的高性价比计算资源构建计算集群，可大幅降低成本。	H3
图形渲染	动画渲染	精细度要求高的渲染场景，基于CPU的渲染计算，计算稳定性高。	aC7、C6、C3
	网络影视渲染	基于GPU渲染器的渲染计算，运算速度快。	G6
AI/机器学习	AI训练	搭配NVIDIA智能网卡，适用于深度学习训练、科学计算、计算流体动力学、计算金融、地震分析、分子建模、基因组学等领域。	P2vs、P2v、P2s
	AI推理	搭配NVIDIA智能网卡，适用于图像分类识别、语音识别、自然语言处理、视频编解码、机器学习、轻量级训练。	Pi2

4. 设置“实例筛选”。

a. 选择规格。

- 搜索规格：输入vCPUs、内存或规格名称直接搜索规格。
当实例选择“规格类型选型”方式时，支持搜索规格。
- 通过列表选择规格：通过实例类型和规格代系名称在列表中选择已上线的规格。
如果您有特定的规格配置需求，选择规格前，请先查看规格的特点、适用场景、指标数据等信息，具体内容请参见[实例类型](#)。

说明

- 购买弹性云服务器时，用户不能选择已售罄的CPU和内存资源。用户可通过勾选“隐藏售罄的规格”将列表中已售罄的规格隐藏。
- b. 设置规格显示范围。
- 仅展示最新系列：勾选后，仅显示新上线的云服务器类型与规格，若不勾选，则显示云服务平台当前提供的所有云服务器类型与规格。
 - 隐藏售罄的规格：勾选后，不显示当前已售罄的规格。
5. （可选）设置“价格上限”。
- 仅当“计费模式”选择“竞价计费”，“竞价类型”选择“竞价模式”时，该参数可见。
- 自动设置：推荐使用，指将按需计费的价格设置为价格上限。即您愿意为弹性云服务器支付的最高价格是按需计费的价格。
 - 手动设置：需要您自己设定竞价计费实例的价格上限。该上限值需大于等于当前的市场价格，小于等于按需计费的价格。
- 更多信息，请参见[购买竞价计费型ECS](#)。
6. （可选）设置“保障周期”。
- 仅当“计费模式”选择“竞价计费”，“竞价类型”选择“竞享模式”时，该参数可见。
- 保障周期：购买竞享实例的最小单位时间，保障周期不同价格不同。
在保障周期内由云平台中断竞享实例，当前周期不计费。在保障周期内用户删除竞享实例，当前周期按使用时长计费。
 - 配置单价：无需设置。
 - （可选）保障周期数量：仅当“保障周期”设置为“6小时”时显示该参数。
- 更多信息，请参见[购买竞享计费型ECS](#)。

操作系统：镜像、安全防护

1. 选择“镜像”。
镜像是一个包含了操作系统及必要配置的弹性云服务器模板，使用镜像可以创建弹性云服务器。

表 2-17 镜像参数配置说明

选项	说明	更多参考
公共镜像	常见的标准操作系统镜像，具有高度稳定性，皆为正版授权，所有用户可见，包括操作系统以及预装的公共应用。 您可以根据实际需求自助配置应用环境或相关软件。	公共镜像概述
私有镜像	由用户创建或导入的个人镜像，仅用户自己可见。包含操作系统、预装的公共应用以及用户的私有应用，可以节省重复配置ECS的时间。	创建私有镜像 通过私有镜像购买ECS

选项	说明	更多参考
共享镜像	由其他账号共享的私有镜像，提供跨账号使用同一镜像创建ECS的能力。	共享镜像概述
市场镜像	仅当计费模式为“包年/包月”或“按需付费”时可选。 由云商店提供的预装操作系统、应用环境和各类软件的优质第三方镜像。无需配置，可一键部署，满足建站、应用开发、可视化管理等个性化需求。	云商店

2. （可选）设置“开启安全防护”。

选择部分操作系统的公共镜像时，系统默认开启主机安全服务（Host Security Service, HSS）的安全防护功能，开启后，默认提供一个月免费的基础防护并自动安装主机安全Agent，支持口令检测、漏洞检测等防护功能。

主机安全服务可以为云服务器提供风险预防、入侵检测、高级防御、安全运营、网页防篡改等安全防护功能，帮助构建云服务器安全防护体系，降低云服务器面临的主要安全风险。

说明

一个月的基础防护免费使用期限结束后，该防护配额将自动释放，停止相应的实时防护能力。

如您需要保留或升级原有安全防护能力，建议您重新购买主机安全服务。详细情况，请参见主机安全的[免费试用HSS基础版到期后怎么办？](#)。

购买后您可以在主机安全服务控制台切换不同版本。各版本之间的差异请参考[服务版本差异](#)。

3. （可选）选择“许可证类型”。

仅当选择的镜像为计费镜像时，会显示该参数。用于设置操作系统或软件的许可证类型。

当前支持选择“使用自带许可证（BYOL）”，使用用户已有操作系统的许可证，无需重新申请。

更多关于许可证类型的信息，请参见[许可证类型](#)。

存储与备份：系统盘、数据盘、备份

1. （可选）设置“磁盘类型”。

仅当已在[专属分布式存储](#)页面申请了存储池时，会显示该参数。

根据磁盘使用的存储资源是否独享，磁盘划分为“云硬盘”和“专属分布式存储”两类，其中，专属分布式存储是为您提供的独享存储资源。

- 专属分布式存储：用于在已申请的存储池中创建磁盘。
- 云硬盘：用于使用公共存储资源创建磁盘。

说明

- 使用专属分布式存储资源创建磁盘时，待创建磁盘的磁盘类型需和申请的存储池资源类型保持一致，如都是“高IO”类型。
- 更多关于专属分布式存储的内容，请参见[《专属分布式存储服务用户指南》](#)。

2. 设置“系统盘”。

系统盘用于存储云服务器的操作系统，创建云服务器时自带系统盘，且系统盘自动初始化。

📖 说明

购买“包年/包月”云服务器时购买的系统盘，卸载后，如果继续作为系统盘使用，则只能挂载给原云服务器；如果作为数据盘使用，则可以挂载给任意云服务器。

表 2-18 系统盘参数配置说明

选项	说明	场景及约束	更多参考
磁盘类型	根据磁盘的IO性能划分磁盘类型。	根据性能，磁盘可分为极速型SSD V2、极速型SSD、通用型SSD V2、超高IO、通用型SSD、高IO、普通IO。 不同类型云硬盘的性能和价格有所不同，您可根据应用程序要求选择您所需的云硬盘。	磁盘类型及性能介绍
系统盘大小	系统盘容量大小，单位为GiB。	云硬盘根据磁盘容量计费，请根据业务需要选择合适的容量大小。 说明 对于P1型、P2型弹性云服务器，系统盘大小需大于等于15GB，建议大于40GB。	
IOPS	云硬盘每秒进行读写的操作次数。 仅当“磁盘类型”选择“通用型SSD V2”时，支持设置该参数。	建议您参考取值范围并根据业务需要，设置IOPS。	
吞吐量	云硬盘每秒成功传送的数据量，即读取和写入的数据量。 仅当“磁盘类型”选择“通用型SSD V2”时，支持设置该参数。	建议您参考取值范围并根据业务需要，设置吞吐量。	

3. （可选）设置系统盘的“高级设置”。
“高级设置”用于设置系统盘的“SCSI”和“加密”属性。

表 2-19 系统盘高级设置参数配置说明

选项	说明	场景及约束	更多参考
SCSI	用于设置系统盘为SCSI磁盘模式。 该参数无需设置，默认勾选。	磁盘模式包含： <ul style="list-style-type: none">• VBD（虚拟块存储设备，Virtual Block Device）模式• SCSI（小型计算机系统接口，Small Computer System Interface）模式 系统默认的磁盘模式为VBD，如果选择了SCSI，能支持SCSI指令透传，可应用于需要支持SCSI指令的场景。 说明 磁盘模式在购买磁盘时配置，购买完成后无法修改。	磁盘模式介绍
加密	用于加密系统盘。 <ul style="list-style-type: none">• 如果选择了加密镜像，系统盘会自动加密，无需再次设置。• 如果未选择加密镜像，可通过勾选“加密”进行系统盘加密操作。	加密系统盘，可以为您的数据提供强大的安全防护。加密磁盘生成的快照及通过这些快照创建的磁盘将自动继承加密功能。 系统盘加密的具体参数，请参见4。	云硬盘加密

4. （可选）设置磁盘加密参数。

仅当磁盘的“高级设置”勾选了“加密”时，会显示该参数。

说明

为了使用加密特性，需单击“创建委托”授权EVS获取KMS密钥用来加解密云硬盘。

如果您有授权资格，则可直接授权，如果权限不足，需先联系拥有Security Administrator权限的用户授权，然后再重新操作。更多信息，请参见。

加密参数包括：

- 委托名称：用于授权EVS获取KMS密钥来加解密云硬盘。当“委托名称”为“EVSAccessKMS”时，表示已授权成功。

说明

为了使用加密特性，需单击“创建委托”授权EVS获取KMS密钥用来加解密云硬盘。

如果您有授权资格，则可直接授权，如果权限不足，需先联系拥有Security Administrator权限的用户授权，然后再重新操作。更多信息，请参见。

- KMS加密：用于设置获取KMS密钥的方式。
 - 从KMS密钥中选择：通过密钥名称选择KMS密钥。
 - 输入KMS密钥ID：通过密钥ID选择KMS密钥。

- （可选）密钥名称：加密云硬盘使用的密钥名称。仅当“KMS加密”选择“从KMS密钥中选择”时，会显示该参数。

您可以选择使用已有的密钥，或者单击“查看密钥列表”，在KMS控制台创建新的密钥。默认为evs/default。

- 密钥ID：该加密数据盘使用的密钥的ID。

5. 设置“数据盘”

数据盘用于存储用户数据。数据盘可以在购买ECS的时候购买，由系统自动挂载给云服务器。也可以在购买了云服务器之后，单独购买云硬盘并挂载给云服务器。

说明

- 数据盘挂载至云服务器后，需要[初始化云硬盘](#)才可以正常使用。
- 购买“包年/包月”云服务器时购买的非共享数据盘，卸载后，如果重新挂载，则只能挂载给原云服务器作数据盘使用。
- 购买“包年/包月”云服务器时购买的数据盘，不支持单独续订、退订、开通自动续费、转按需付费方式、以及释放功能。

表 2-20 数据盘参数配置说明

选项	说明	场景及约束	更多参考
磁盘类型	根据磁盘的IO性能划分磁盘类型。	根据性能，磁盘可分为极速型SSD V2、极速型SSD、通用型SSD V2、超高IO、通用型SSD、高IO、普通IO。 不同类型云硬盘的性能和价格有所不同，您可根据应用程序要求选择您所需的云硬盘。	磁盘类型及性能介绍
数据盘大小	数据盘容量大小，单位为GiB。	云硬盘根据磁盘容量计费，请根据业务需要选择合适的容量大小。	
IOPS	云硬盘每秒进行读写的操作次数。 仅当“磁盘类型”选择“通用型SSD V2”时，支持设置该参数。	建议您参考取值范围并根据业务需要，设置IOPS。	
吞吐量	云硬盘每秒成功传送的数据量，即读取和写入的数据量。 仅当“磁盘类型”选择“通用型SSD V2”时，支持设置该参数。	建议您参考取值范围并根据业务需要，设置吞吐量。	
数量	数据盘的数量。	请根据业务需要选择需要增加的数据盘数量。 购买ECS时，您最多可以为ECS添加23块数据盘。	

6. （可选）设置数据盘的“高级设置”。
- “高级设置”用于设置系统盘的“SCSI”、“共享盘”和“加密”属性。

表 2-21 系统盘高级设置参数配置说明

选项	说明	场景及约束	更多参考
SCSI	用于设置数据盘为SCSI磁盘模式。 该参数无需设置，默认勾选。	磁盘模式包含： <ul style="list-style-type: none"> • VBD（虚拟块存储设备，Virtual Block Device）模式 • SCSI（小型计算机系统接口，Small Computer System Interface）模式 系统默认的磁盘模式为VBD，如果选择了SCSI，能支持SCSI指令透传，可应用于需要支持SCSI指令的场景。 <p>说明 磁盘模式在购买磁盘时配置，购买完成后无法修改。</p>	磁盘模式介绍
共享盘	用于设置数据盘为共享盘。	设置数据盘为共享盘后，该共享盘可以挂载至多个ECS。	共享云硬盘
加密	用于加密数据盘。	加密数据盘，可以为您的数据提供强大的安全防护。 数据盘加密的具体参数，请参见4。	云硬盘加密
用数据盘镜像创建磁盘	用于设置用数据盘镜像创建数据盘。	当选择Windows或Linux的镜像创建ECS时，系统支持使用数据盘镜像创建数据盘。 勾选“用数据盘镜像创建磁盘”，在弹窗中选择您制作的数据盘镜像。 <p>说明 同一个数据盘镜像，只能给一个数据盘使用，不能重复选择。 使用整机镜像创建云服务器，或数据盘使用SCSI、共享盘、加密功能时，均不支持与数据盘镜像创建磁盘同时使用。</p>	创建私有镜像

7. （可选）设置“开启备份”。
- 仅当需要备份ECS或云硬盘时，设置该参数。
- 备份可以帮助您在服务器故障时恢复数据，为了您的数据安全，建议您开启备份。

📖 说明

- 使用云备份功能需要进行单独计费，计费详情请参见[云备份是如何计费的?](#)
 - CloudPond不支持设置“云备份”。
8. （可选）设置“云备份”。
- 仅当勾选“开启备份”时，会显示该参数。
- 您可以根据实际情况选择设置云备份的方式。
- 现在购买：设置云备份相关参数。
 - i. 设置“云备份存储库名称”：只能由中文字符、英文字母、数字、下划线、中划线组成，且长度小于等于64个字符。例如：vault-f61e。默认的命名规则为“vault_xxxx”。
 - ii. 设置“云备份存储库容量”：此容量为备份云服务器所需的容量。存储库的空间不能小于云服务器的空间。取值范围为[云服务器总容量，10485760]GB。
 - iii. 设置“备份策略”：在下拉列表中选择备份策略，或进入云备份控制台查看或编辑备份策略。
 - 使用已有：选择已有云备份存储库并设置备份策略。
 - i. 选择“云备份存储库”：在下拉列表中选择已有的云备份存储库。
 - ii. 设置“备份策略”：在下拉列表中选择备份策略，或进入云备份控制台查看或编辑备份策略。

网络：虚拟私有云、主网卡、扩展网卡、源/目的检查

1. 设置“虚拟私有云”和“主网卡”。
- 虚拟私有云（VPC）为弹性云服务器构建隔离的、用户自主配置和管理的虚拟网络环境，可以在VPC中定义安全组、VPN、IP地址段、带宽等网络特性。不同虚拟私有云里面的弹性云服务器网络默认不通。

图 2-12 “网络”参数设置



您可以在下拉列表中选择已有的虚拟私有云，或者根据业务需要创建新的虚拟私有云。系统会默认关联主网卡并设置私有IP地址的分配方式。

更多信息，请参见[虚拟私有云和子网规划建议](#)。

📖 说明

- 弹性云服务器使用的VPC网络DHCP不能禁用。
- 如果想要使用其他账号的共享VPC和子网，请先接受共享邀请，详细操作，请参见[接受/拒绝共享邀请](#)。
有关VPC子网共享的更多信息，请参见《虚拟私有云用户指南》的“[共享VPC](#)”相关内容。

2. 设置“主网卡”。

主网卡用于系统的默认路由，不允许删除，随ECS自动创建。当设置了“虚拟私有云”后，系统会默认关联主网卡并设置私有IP地址的分配方式。

- a. 选择“主网卡”：当虚拟私有云下有多个子网时，可以在下拉列表选择其他子网作为主网卡。
- b. 设置IPv4地址分配方式：默认选择“自动分配IP地址”。
 - 自动分配IP地址：系统自动为主网卡分配私有IPv4地址。
 - 手动分配IP地址：为主网卡设置指定私有IPv4地址。设置前，可单击“查看已使用IP地址”进行查看，避免冲突。

📖 说明

为主网卡设置指定私有IPv4地址时，如果是批量创建多台弹性云服务器：

- 此时，该IP地址为起始IP地址。
 - 请确保IP地址在子网范围内且连续可用。
 - 其他子网不能与指定IP的子网相同。
- c. （可选）设置IPv6地址分配方式：默认选择“暂不分配IPv6地址”。

仅当选择的规格的“IPv6”列为“是”、且子网开启了IPv6功能时，该参数可见。

- 暂不分配IPv6地址：暂不为网卡分配私有IPv6地址。
- 自动分配IPv6地址：系统自动为网卡分配私有IPv6地址。
当选择“自动分配IPv6地址”后，网卡的IP地址为IPv6类型。在同一VPC内，ECS通过IPv6地址在双栈ECS之间进行内网访问。
如需访问公网，需要创建EIP共享带宽，并将实例的IPv6地址添加至共享带宽即可。
详细内容，请参见[添加弹性公网IP到共享带宽](#)。

📖 说明

- 子网开启IPv6功能的详细内容，请参见[IPv6网络](#)。
- ECS创建成功后，需检查是否获取IPv6地址，如果未获取需配置ECS动态获取IPv6地址，具体操作请参见[动态获取IPv6地址](#)。

3. （可选）设置“新增扩展网卡”。

扩展网卡是可以单独创建的网卡。当您需要为ECS配置多个网卡时，可以新增多张扩展网卡，并指定扩展网卡的IP地址。

扩展网卡无法和外部正常通信，需要配置策略路由，详细信息，请参见[为多网卡ECS配置策略路由](#)。

📖 说明

云服务器可绑定的扩展弹性网卡数量由云服务器实例规格决定，具体请参见[规格清单 \(x86\)](#)和[规格清单 \(鲲鹏\)](#)。

- a. 选择“扩展网卡1”：当虚拟私有云下有多个子网时，可以在下拉列表选择其他子网作为扩展网卡。
- b. 设置IPv4地址分配方式：默认选择“自动分配IP地址”。
 - 自动分配IP地址：系统自动为扩展网卡分配私有IPv4地址。
 - 手动分配IP地址：为扩展网卡设置指定私有IPv4地址。设置前，可单击“查看已使用IP地址”进行查看，避免冲突。
 - 使用已有弹性网卡：设置指定弹性网卡作为扩展网卡，可以通过下拉列表选择指定的弹性网卡。
- c. （可选）设置IPv6地址分配方式：默认选择“暂不分配IPv6地址”。

仅当选择的规格的“IPv6”列为“是”、且子网开启了IPv6功能时，该参数可见。

- 暂不分配IPv6地址：暂不为网卡分配私有IPv6地址。
- 自动分配IPv6地址：系统自动为网卡分配私有IPv6地址。
当选择“自动分配IPv6地址”后，网卡的IP地址为IPv6类型。在同一VPC内，ECS通过IPv6地址在双栈ECS之间进行内网访问。
如需访问公网，需要创建EIP共享带宽，并将实例的IPv6地址添加至共享带宽即可。
详细内容，请参见[添加弹性公网IP到共享带宽](#)。

📖 说明

- 子网开启IPv6功能的详细内容，请参见[IPv6网络](#)。
 - ECS创建成功后，需检查是否获取IPv6地址，如果未获取需配置ECS动态获取IPv6地址，具体操作请参见[动态获取IPv6地址](#)。
4. 设置“源/目的检查”。

开启“源/目的检查”，系统会检查弹性云服务器发送的报文中源IP地址是否正确，如果不正确，则不允许发送该报文。通过该功能，有助于防止伪装报文攻击，提升安全性。默认情况下“源/目的检查”的状态为“ON”，无需设置。

📖 说明

源/目的检查设置对使用已有弹性网卡无效，仅对同云服务器一起创建的网卡生效。

安全组：安全组、安全组规则

1. 设置“安全组”。

在下拉列表中选择可用的安全组。支持为ECS选择多个安全组（建议不超过5个），此时，ECS的访问规则遵循几个安全组规则的并集。

首次创建ECS时，系统会自动创建默认安全组default、Sys-WebServer和Sys-FullAccess，详细信息，请参见[默认安全组和规则](#)。

您可以单击“展开安全组规则”查看具体的“入方向规则”和“出方向规则”。安全组规则的配置会影响ECS的正常访问与使用，配置方法请参见[为安全组添加安全组规则](#)。常用端口与协议的用途如下，请按需开启：

- 80端口：浏览网页的默认端口，主要用于HTTP服务。
 - 443端口：网页浏览端口，主要用于HTTPS服务。
 - ICMP协议：用于ping云服务器之间的通信情况。
 - 22端口：用于Linux云服务器的SSH方式登录。
 - 3389端口：用于Windows云服务器的远程桌面登录。
2. （可选）新建安全组。
- 当下拉列表中可选择的安全组不满足业务需求时，可单击“新建安全组”进行创建。

图 2-13 新建安全组

×

创建安全组

* 名称

* 模板

入方向规则

- 远程登录和ping:
 - SSH (22) RDP (3389) FTP (20-21)
 - Telnet (23) ICMP (全部)
- Web服务:
 - HTTP (80) HTTPS (443) HTTP_ALT (8080)
- 数据库:
 - MySQL (3306) MS SQL (1433) PostgreSQL (5432)
 - Oracle (1521) Redis (6379)

描述

您可以勾选常见协议端口，在入方向快速添加规则放行对应的协议及端口。若您未勾选任何协议端口，则不会放行任何端口。您可在安全组创建后，根据实际访问需求添加或修改安全组规则。

0/255

[隐藏模板规则](#)

入方向 出方向

优先级	策略	类型	协议端口	源地址
1	允许	IPv4	全部	sg-dc8b
1	允许	IPv6	全部	sg-dc8b

新建安全组的相关参数说明如下：

表 2-22 安全组参数说明

参数	参数说明	取值样例
名称	<p>必选参数。输入安全组的名称。要求如下：</p> <ul style="list-style-type: none"> 长度范围为1~64位。 名称由中文、英文字母、数字、下划线（_）、中划线（-）、点（.）组成。 <p>说明 安全组名称创建后可以修改，建议不要重名。</p>	sg-AB
模板	<p>必选参数。模板自带安全组规则，方便您快速创建安全组。默认选择“快速添加规则”。</p> <p>提供如下几种模板：</p> <ul style="list-style-type: none"> 通用Web服务器：默认会配置放通22、3389、80、443端口和ICMP协议。 开放全部端口：默认放通所有协议和所有端口。 <p>说明 开放全部端口有一定安全风险，请谨慎选择。</p> <ul style="list-style-type: none"> 快速添加规则：通过勾选常见协议端口，在入方向快速添加规则放通对应的协议及端口。 	快速添加规则
入方向规则	<p>可选参数。仅当“模板”选择“快速添加规则”时，会显示该参数。</p> <p>当前支持快速添加的协议和端口如下，请根据业务场景进行选择：</p> <ul style="list-style-type: none"> 远程登录和ping：SSH（22）、RDP（3389）、FTP（20-21）、Telnet（23）、ICMP（全部） Web服务：HTTP（80）、HTTPS（443）、HTTP_ALT（8080） 数据库：MySQL（3306）、MS SQL（1433）、PostgreSQL（5432）、Oracle（1521）、Redis（6379） 	-
描述	<p>可选参数。安全组的描述信息。</p> <p>描述信息内容不能超过255个字符，且不能包含“<”和“>”。</p>	-
查看/隐藏模板规则	安全组规则的展示区域，您可以查看或隐藏当前组的入方向规则和出方向规则。	-

- （可选）配置安全组规则。
如需修改当前安全组的规则，可单击“配置安全组规则”进行修改。
更多详细信息，请参见[配置安全组规则](#)。
- （可选）展开/隐藏安全组规则。
您可以展开或隐藏安全组规则区域，在此区域包含：

- 已选安全组：展示当前已选的安全组。如果选择多个安全组，支持通过单击“上移”或“下移”调整安全组的优先级。
- 安全组规则：支持查看“入方向规则”和“出方向规则”。

公网访问：弹性公网 IP、公网带宽、带宽大小、释放方式

1. 设置“弹性公网IP”。

弹性公网IP是指将公网IP地址和路由网络中关联的弹性云服务器绑定，以实现虚拟私有云内的弹性云服务器通过固定的公网IP地址对外提供访问服务。

您可以根据实际情况选择以下三种方式：

- 现在购买：自动为每台弹性云服务器分配独享带宽的弹性公网IP，带宽值可以由您设定。
- 使用已有：为弹性云服务器分配已有弹性公网IP。使用已有弹性公网IP时，不能批量创建弹性云服务器。
- 暂不购买：弹性云服务器不能与互联网互通，仅可作为私有网络中部署业务或者集群所需弹性云服务器进行使用。

说明

对于包年/包月方式购买的弹性云服务器，不支持“现在购买”弹性公网IP。如需使用弹性公网IP，请绑定已有的弹性公网IP，或先购买按需方式的弹性公网IP，然后再绑定。

2. 设置弹性公网IP的“线路”。

- 全动态BGP：可根据设定的寻路协议第一时间自动优化网络结构，以保持客户使用的网络持续稳定、高效。
- 静态BGP：网络结构发生变化，运营商无法在第一时间自动调整网络设置以保障用户的体验度。

3. （可选）设置“公网带宽”。

仅当“弹性公网IP”选择“现在购买”时，会显示该参数。一个带宽只能被一个弹性公网IP使用。

- 按带宽计费：独享带宽，按照购买的带宽大小计费。
- 按流量计费：独享带宽，按照实际使用的流量来计费。
- 加入共享带宽：一个带宽中可以加入多个弹性公网IP，多个弹性公网IP共用一个带宽。

说明

- 一个共享带宽支持添加的弹性公网IP个数有限，如果配额不足，可以选择切换使用其他共享带宽，或者申请扩大共享带宽的EIP配额。
- 包年/包月方式购买的EIP，不支持使用共享带宽。
- 包年/包月方式购买的共享带宽，到期后系统自动删除，并给该共享带宽中添加的EIP创建按流量计费的独占带宽。

4. （可选）设置“带宽大小”。

仅当“弹性公网IP”选择“现在购买”时，会显示该参数。根据业务需要，选择所需的带宽大小，单位Mbit/s。

5. （可选）选择弹性公网IP。

仅当“弹性公网IP”选择“使用已有”时，会显示该参数。可以在下拉列表中选择可用的弹性公网IP。

云服务器管理：“云服务器名称”、登录凭证、密码/密钥对、标签

1. 设置“云服务器名称”。

本步骤中设置的“云服务器名称”，会被设置为云服务器操作系统中初始的主机名（hostname）。

云服务器名称需符合命名规则：只能由中文字符、英文字母、数字及“_”、“-”、“.”组成。

📖 说明

Windows操作系统云服务器的名称建议不超过15个字符，且不要重名，否则部分Windows应用将无法使用。

弹性云服务器操作系统中的主机名（hostname）的命名规则遵循 [RFC 952](#)和[RFC 1123](#)规范。

在将“云服务器名称”设置为虚拟机操作系统中的主机名时，为避免出现未知问题，建议使用a-z或0-9以及中划线“-”组成的名称命名。

- “_”将在弹性云服务器内部被转换为“-”。
 - 连续使用的“-”和“_”将在弹性云服务器内部被转换为1个“-”。
 - 以“.”、“-”、“_”和中文字符开头将在弹性云服务器内部被忽略，从第一个非以上字符开始。
 - 不作为开头的“.”和中文字符及后面内容将在弹性云服务器内部被忽略。
- 批量购买多台云服务器时，支持自动增加数字后缀命名或者自定义规则命名。
- 自动增加数字后缀命名：多台云服务器的名称自动按序增加4位数字后缀。例如：输入ecs，从ecs-0001开始命名；若已有ecs-0010，则从ecs-0011开始命名。
 - 自定义规则命名：通过name_prefix[begin_number,bits]name_suffix格式可给各个云服务器设置不同的名称。begin_number取值区间为[0,9999]，bits取值区间为[1,4]。例如：ecs[66,3]abc，若创建2台云服务器最终生成的名称分别为ecs066abc、ecs067abc。
- 允许重名：允许创建的云服务器名称相同。如果是批量创建，勾选“允许重名”后，批量创建的多台云服务器名称全部相同，不再按序增加。

2. 设置“登录凭证”。

登录凭证用于设置登录ECS的方式。

表 2-23 登录凭证参数配置说明

选项	说明	场景及约束	更多参考
密码	使用设置初始密码方式作为ECS的登录凭证方式，您可以通过用户名密码方式登录ECS。 Linux ECS设置的是root用户的初始密码，Windows ECS设置的是Administrator用户的初始密码。	建议您使用复杂度高的密码，保证密码符合要求，防止恶意攻击。密码设置规则如表2-24所示。 说明 系统不会定期自动修改弹性云服务器密码。为安全起见，建议您定期修改密码。	密码使用场景介绍

选项	说明	场景及约束	更多参考
密钥对	使用密钥对作为ECS的登录凭证方式。 您可以选择使用已有的密钥，或者单击“新建密钥对”创建新的密钥。	密钥对登录凭证方式的安全性更高。 说明 如果选择使用已有的密钥，请确保您已在本地获取该文件，否则，将影响您正常登录弹性云服务器。	（推荐）通过管理控制台创建密钥对
创建后设置	暂不设置ECS的密码。	待ECS创建成功后，您需要单击“操作”列下的“重置密码”，根据界面提示，为ECS设置密码。	在控制台重置ECS密码

表 2-24 密码设置规则

参数	规则
密码	<ul style="list-style-type: none"> 密码长度范围为8到26位。 密码至少包含以下4种字符中的3种： <ul style="list-style-type: none"> 大写字母 小写字母 数字 Windows操作系统云服务器支持的特殊字符：!@\$%^_-=+[{()}];./?~#* Linux操作系统云服务器支持的特殊字符：!@\$%^_-=+[{()}];./?~#* 密码不能包含用户名或用户名的逆序。 Windows操作系统的云服务器，不能包含用户名中超过两个连续字符的部分。

3. （可选）设置密码。

仅当“登录凭证”选择“密码”时，会显示该参数。

根据表2-24完成“密码”和“确认密码”的设置，两次输入需保持一致。

4. （可选）设置“密钥对”。

仅当“登录凭证”选择“密钥对”时，会显示该参数。通过下拉列表选择可用的密钥对，或参考[（推荐）通过管理控制台创建密钥对](#)新建密钥对。

说明

如果选择使用已有的密钥，请确保您已在本地获取该文件，否则，将影响您正常登录弹性云服务器。

5. （可选）设置“标签”。

如需通过标签管理ECS资源时，可以为ECS添加标签。

标签用于标识弹性云服务器，可以方便识别和管理您拥有的ECS资源。您最多可以给弹性云服务器添加10个标签。

更多信息，请参见[标签概述](#)。

📖 说明

创建弹性云服务器时添加的标签，将同步添加至一同创建的EIP和云硬盘（包括系统盘、数据盘）上。如果云服务器使用的是已有EIP，则该标签不会在EIP上标识。

如您的组织已经设定弹性云服务器的相关标签策略，则需按照标签策略规则为弹性云服务器添加标签。如果添加的标签不符合标签策略规则，则可能会导致弹性云服务器创建失败，请联系组织管理员了解标签策略详情。

云服务器创建成功后，您可以在弹性云服务器、EIP和云硬盘详情页，查看到对应的标签。

高级配置：云监控、云服务器组、云服务器描述、实例自定义数据注入、委托、CPU选项

1. 设置“云监控”。

选择部分操作系统的公共镜像时，系统推荐您配套使用主机监控。主机监控采集丰富的操作系统层面监控指标，用于监控服务器资源使用情况，也可查询数据用于资源的故障排查。

勾选“开启详细云监控”后，通过在云服务器中安装Agent插件，支持对云服务器的CPU、内存、网络、磁盘、进程等指标的1分钟详细监控。

2. 设置“云服务器组（可选）”

云服务器组内的弹性云服务器将遵循反亲和策略，尽量分散地创建在不同主机上。创建云服务器组的详细操作，请参见[管理云服务器组](#)。

📖 说明

包含本地盘的弹性云服务器无法在创建后加入弹性云服务器组，如需使用弹性云服务器组功能，请在创建时选择弹性云服务器组。

3. 设置“云服务器描述”信息。

4. 设置实例自定义数据注入。

用于在创建ECS时，通过注入自定义数据配置ECS，例如，通过脚本初始化操作系统。配置后，ECS首次启动时会自行注入数据信息。

- 以文本形式：在下方文本框内输入用户数据内容。

- 以文件形式：主要用于创建云服务器时注入的脚本文件或其他文件。

实例自定义数据注入的详细操作，请参见[通过实例自定义数据配置ECS实例](#)。

5. 设置“委托”。

当需要与其他账号共享云服务器资源，或者委托更专业的人或团队来代为管理时，租户管理员可以在IAM创建委托并授予云服务器资源的管理权限。被委托方使用自己的用户登录系统后，切换到您的账号下管理资源，避免您将自己的安全凭证（密码）共享给其他账号，确保了您的账号安全。

如果您在IAM上创建了委托，可以通过单击下拉列表选择委托名称，获取相应权限。更多关于委托的信息，请参见[委托其他账号管理资源](#)。

6. 设置“CPU选项”。

- 当需要为云服务器实例开启或关闭超线程时，勾选“指定CPU选项”。

关于超线程的详细介绍，请参见[开启/关闭ECS超线程](#)。

- 每核心线程数

当勾选“指定CPU选项”时，显示该参数。在“每核心线程数”下拉框进行设置。

- 1: 关闭超线程
- 2: 开启超线程，默认开启

购买量：购买时长/使用时长、自动续费、购买数量、企业项目

1. （可选）设置“购买时长”。
仅当“计费模式”选择“包年/包月”时，会显示该参数。用于设置ECS的购买时长，最短为1个月，最长为1年。
2. （可选）设置“自动续费”。
仅当“计费模式”选择“包年/包月”时，会显示该参数。
勾选“自动续费”，在“包年/包月”资源到期后，可以自动进行续费。
 - 按月购买的资源：每次续费1个月，次数不限。
 - 按年购买的资源：每次续费1年，次数不限。关于自动续费的更多说明，请参考[自动续费规则说明](#)。
3. （可选）设置“使用时长”。
仅当“计费模式”选择“按需计费”或“竞价计费”时，会显示该参数。勾选“设置定时删除时间”后，可以在设定时间自动删除云服务器。

须知

设置定时删除时间后，系统将在设定时间自动删除云服务器，请提前做好数据备份，谨慎操作。

设置的删除时间需晚于当前时间1小时，且不能超过3年。设置后还可以在实例删除前修改定时删除时间。

系统每5分钟执行一次定时删除任务，按照系统删除服务器的时间停止计费。

4. 设置“购买数量”。
如需批量购买ECS，可以设置购买ECS的数量。设置后，会批量创建多台相同配置的ECS。
系统会显示您当前还可以购买的ECS数量。为了保证所有资源的合理分配，如果您需要的ECS数量超过当前可以购买的最大数值，可以[申请扩大配额](#)。

说明

您也可以在购买页底部进行设置：

- 当“计费模式”选择“包年/包月”时，可设置“购买量”和“购买时长”。
- 当“计费模式”选择“按需计费”或“竞价计费”时，可设置“购买量”。

设置完成后，可以通过“配置费用”查看具体收费项。如果您对价格有疑问，可通过[了解计费详情](#)来了解产品价格。

步骤三：确认配置并购买 ECS

1. 在页面右侧的“配置概要”中，查看ECS配置详情。
未设置的必选字段会通过红色字体进行提示，需在参数配置区域进行设置。
2. （可选）保存为启动模板。

仅当需要通过本配置创建启动模板，用于创建新的ECS时，可执行本步骤。更多信息，请参见[通过智能购买助手购买ECS](#)。

包含以下任一配置项时均不支持保存为启动模板：计费模式为包年/包月、竞价类型为竞享模式、配置了主机安全、登录凭证为密码。

3. 勾选协议并单击“立即购买”。

ECS创建成功后会默认开机。

2.3.3 购买竞价计费型 ECS

操作场景

竞价计费型实例采用竞价计费模式，以一定折扣购买并使用弹性云服务器，性能与常规云服务器无异。但是当库存资源不足，或市场价格上浮、并超过您的预期价格时，系统会自动释放您的云服务器资源，对这些折扣售卖的弹性云服务器进行中断回收。

与“按需计费”和“包年/包月”方式购买的弹性云服务器相比，在提供同等性能的前提下，竞价计费型实例可提供超低折扣，降低用云成本。

更多关于竞价计费型实例的介绍，请参见[竞价计费（竞价模式）](#)。

约束与限制

- 当前仅支持虚拟化类型为KVM的弹性云服务器使用竞价计费型实例策略。具体的规格支持清单，请以控制台显示的资源为准。
- 在不同可用区，同一规格的云服务器市场价格变化可能不同。
- 使用竞价计费型实例策略的弹性云服务器，不支持切换操作系统功能。
- 使用竞价计费型实例策略的弹性云服务器，不支持自动恢复功能。
- 使用竞价计费型实例策略的弹性云服务器，不支持变更规格。
- 不支持使用市场镜像创建竞价计费型实例。
- 竞价计费型实例不支持转“包年/包月”方式。
- 竞价计费型实例不支持卸载系统盘。
- 竞价计费型实例在中断回收时：
 - 不允许使用该竞价计费型实例制作系统盘镜像和整机镜像，但允许使用该服务器的数据盘制作数据盘镜像。
 - 不允许删除。
- 竞价计费型实例在中断回收时，数据盘和弹性公网IP默认不释放。请设置通知提醒，来决定是否释放关联的数据盘和弹性公网IP。

购买云服务器

请按照[购买弹性云服务器](#)、[Windows ECS登录方式概述](#)或[Linux ECS登录方式概述](#)的描述，购买并登录竞价计费型实例。但是，如下参数的设置请特别注意：

创建竞价计费型实例时，如下参数的设置请特别注意：

- 计费模式：选择“竞价计费”。
相对于按需计费模式，以更低的折扣按实际使用时长计费。但是，华为云会根据市场价格或供需资源变化随时自动释放竞价计费型实例。

- 出价模式：包括“自动设置”和“手动设置”。
 - 自动设置：推荐使用，指将按需计费的价格设置为价格上限。即您愿意为弹性云服务器支付的最高价格是按需计费的价格。
 - 手动设置：需要您自己设定竞价计费型实例的价格上限。该上限值需大于等于当前的市场价格，小于等于按需计费的价格。
- 单击“立即购买”，确认规格和价格无误后，勾选同意协议，并单击“提交”。

📖 说明

竞价计费型实例购买成功后，始终可能会被系统中断回收。因此，请务必做好数据备份工作。

配置中断回收提醒（可选）

竞价计费型实例购买成功后，您可以像使用云上其他云服务器一样，正常使用竞价计费型实例。但是，华为云会根据市场价格或供需资源变化随时自动释放竞价计费型实例，终止实例的使用。

当“报价 < 市场价格”或库存资源不足时，系统发出即将释放云服务器资源的通知，并在约5分钟后自动中断回收弹性云服务器。您可以通过如下方式配置中断回收提醒：

通过“云审计服务”、“消息通知服务”，查看并设置中断回收提醒。操作步骤类似如下，具体操作请以《[云审计服务用户指南](#)》为准：

步骤1 开通“云审计服务”，具体操作请参见“[开通云审计服务](#)”。

用户开通云审计服务后，系统会自动识别当前云平台上所开通的云服务，自动抓取各云服务的各项关键操作并主动向云审计服务上报各项关键操作的审计日志。

步骤2 配置中断回收通知。

云审计服务在记录某些特定关键操作发生时，支持对这些关键操作通过消息通知服务实时向相关订阅者发送通知，该功能由云审计服务触发，消息通知服务（SMN）完成通知发送。


1. 登录控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“管理与监管 > 云审计服务”。
4. 单击左侧导航树的“关键操作通知”。
5. 单击页面右上角的“创建关键操作通知”，并按[表2-25](#)填写参数配置。

表 2-25 创建关键操作通知

参数类型	参数	参数配置
基本信息	通知名称	用户自定义，如spottest。
配置操作	操作类型	选择“自定义操作”。
	操作列表	依次选择“ECS > server > interruptServer”，并单击“添加”。
配置用户	指定用户	选择“不指定”，此时，云审计服务默认指定对所有用户发起的关键操作进行通知。

参数类型	参数	参数配置
配置SMN主题	发送通知	选择“发送”。
	SMN主题	在下拉列表中选择SMN发送通知的主题，用于华为云中中断回收竞价计费型实例时，系统发送通知使用。如果没有合适的SMN主题，请新建主题： 1. 单击“消息通知服务”，跳转至消息通知服务页面。 2. 单击“创建主题”，根据界面提示新建一个主题，如“spot”。具体操作请参见“ 创建主题 ”。 3. 主题创建成功后，单击“主题”列表页“操作”列下的“添加订阅”，接收发布至主题的消息。具体操作请参见“ 订阅主题 ”。

配置成功后，当删除实例时，您将提前5分钟收到系统发出的提醒。

步骤3 （可选）查看中断回收实例。

1. 选择“管理与监管 > 云审计服务”。
2. 单击左侧导航树的“事件列表”。
3. 事件列表支持通过筛选来查询对应的操作事件。请展开“筛选”下拉框，按[表 2-26](#)选择查询条件，并单击“查询”。

表 2-26 查询中断回收实例

参数	参数配置
事件来源	ECS
资源类型	server
筛选类型	按事件名称 > interruptServer
操作用户	所有操作用户
事件级别	所有事件级别

4. 在需要查看的事件左侧，展开该记录的详细信息。
5. 在需要查看的记录右侧，单击“查看事件”，弹窗显示该操作事件结构的详细信息。

----结束

2.3.4 购买竞享计费型 ECS

操作场景

竞享实例采用竞价计费模式，以一定折扣购买并使用弹性云服务器，性能与常规云服务器无异。但是当库存资源不足，系统会自动释放您的云服务器资源，对这些折扣售卖的弹性云服务器进行中断回收。

与“按需计费”和“包年/包月”方式购买的弹性云服务器相比，在提供同等性能的前提下，竞享实例可提供超低折扣，降低用云成本。

更多关于竞享实例的介绍，请参见[竞价计费（竞享模式）](#)。

约束与限制

- 竞享计费型实例目前支持通用计算增强型实例规格。
- 华为云尽可能保障竞享实例正常运行，但在系统资源不足等极端情况下会主动释放实例，请务必做好数据备份工作。
- 竞享计费型实例目前只支持部分区域及规格，请以购买页面显示的资源为准
- 竞享计费型实例保障周期不同价格不同。
- 竞享计费型实例不支持转换为按需和包年/包月实例。
- 竞享计费型实例不支持变更实例规格。
- 竞享计费型实例不支持变更操作系统。
- 竞享计费型实例不支持自动恢复功能。
- 竞享计费型实例不支持卸载系统盘。
- 竞享计费型实例在中断回收时：
 - 不允许使用竞享实例制作系统盘镜像和整机镜像，但允许使用该云服务器的数据盘制作数据盘镜像。
 - 不允许删除。
- 竞享计费型实例在中断回收时，数据盘和弹性公网IP默认不释放。请设置通知提醒，来决定是否释放关联的数据盘和弹性公网IP。

购买云服务器

请按照[购买弹性云服务器](#)、[Windows ECS登录方式概述](#)或[Linux ECS登录方式概述](#)的描述，购买并登录竞享计费型实例。但是，如下参数的设置请特别注意：

创建竞享实例时，如下参数的设置请特别注意：

- 计费模式：选择“竞价计费”。
相对于按需计费模式，以更低的折扣按实际使用时长计费。但是当库存资源不足，系统会自动释放您的云服务器资源，请务必做好数据备份。
- 竞价类型：选择“竞享模式”。
- 选择“保障周期”和“保障周期数量”。
保障周期是购买竞享实例的最小单位时间。在保障周期内由云平台中断竞享实例，当前周期不计费。在保障周期内用户删除竞享实例，当前周期按使用时长计费。
- 单击“立即购买”，确认规格和价格无误后，勾选同意协议，并单击“提交”。

📖 说明

竞享实例购买成功后，始终可能会被系统中断回收。因此，请务必做好数据备份工作。

配置中断回收提醒

竞享实例购买成功后，您可以像使用其他云服务器一样，正常使用竞享实例。但是，当使用时间到达您的购买时长或在系统资源不足等极端情况下，系统会自动回收实例。您可以通过如下方式配置中断回收提醒：

通过“云审计服务”、“消息通知服务”，查看并设置中断回收提醒。操作步骤类似如下，具体操作请以《云审计服务用户指南》为准：

步骤1 开通“云审计服务”，具体操作请参见“[开通云审计服务](#)”。

用户开通云审计服务后，系统会自动识别当前云平台上所开通的云服务，自动抓取各云服务的各项关键操作并主动向云审计服务上报各项关键操作的审计日志。

步骤2 配置中断回收通知。

云审计服务在记录某些特定关键操作发生时，支持对这些关键操作通过消息通知服务实时向相关订阅者发送通知，该功能由云审计服务触发，消息通知服务（SMN）完成通知发送。


1. 登录控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“管理与监管 > 云审计服务”。
4. 单击左侧导航树的“关键操作通知”。
5. 单击页面右上角的“创建关键操作通知”，并按[表2-27](#)填写参数配置。

表 2-27 创建关键操作通知

参数类型	参数	参数配置
基本信息	通知名称	用户自定义，如spottest。
配置操作	操作类型	选择“自定义操作”。
	操作列表	依次选择“ECS > server > interruptServer”，并单击“添加”。
配置用户	指定用户	选择“不指定”，此时，云审计服务默认指定对所有用户发起的关键操作进行通知。
配置SMN主题	发送通知	选择“发送”。

参数类型	参数	参数配置
	SMN主题	在下拉列表中选择SMN发送通知的主题，用于华为云中中断回收竞价计费型实例时，系统发送通知使用。如果没有合适的SMN主题，请新建主题： 1. 单击“消息通知服务”，跳转至消息通知服务页面。 2. 单击“创建主题”，根据界面提示新建一个主题，如“spot”。具体操作请参见“ 创建主题 ”。 3. 主题创建成功后，单击“主题”列表页“操作”列下的“添加订阅”，接收发布至主题的消息。具体操作请参见“ 订阅主题 ”。

配置成功后，当删除实例时，您将提前5分钟收到系统发出的提醒。

步骤3 （可选）查看中断回收实例。

1. 选择“管理与监管 > 云审计服务”。
2. 单击左侧导航树的“事件列表”。
3. 事件列表支持通过筛选来查询对应的操作事件。请展开“筛选”下拉框，按[表 2-28](#)选择查询条件，并单击“查询”。

表 2-28 查询中断回收实例

参数	参数配置
事件来源	ECS
资源类型	server
筛选类型	按事件名称 > interruptServer
操作用户	所有操作用户
事件级别	所有事件级别

4. 在需要查看的事件左侧，展开该记录的详细信息。
5. 在需要查看的记录右侧，单击“查看事件”，弹窗显示该操作事件结构的详细信息。

---结束

2.3.5 通过私有镜像购买 ECS

操作场景

私有镜像是由用户创建或导入的个人镜像，仅用户自己可见。包含操作系统、预装的公共应用以及用户的私有应用，可以节省重复配置ECS的时间。

使用公共镜像和私有镜像创建云服务器云主机的区别是：

- 公共镜像：创建的云服务器云主机包含所需操作系统和预装的公共应用，需要您自行安装应用软件。
- 私有镜像或共享镜像：创建的云服务器云主机包含操作系统、预装的公共应用以及用户的私有应用。

说明

您可以选择使用加密镜像，更多关于镜像加密的信息，请参见[镜像加密](#)。

约束与限制

- 如果选择使用整机镜像创建弹性云服务器，整机镜像关联的云硬盘不支持使用“数据盘镜像创建磁盘”功能。
- 当整机镜像为可用区内可用时（状态栏为“正常”，且提示“AZx可用”），那么该整机镜像只能在该可用区中创建云服务器，且云服务器系统盘和数据盘的加密属性会和原始云服务器的系统盘和数据盘的加密属性保持一致，不支持在创建时修改系统盘和数据盘的SCSI模式、共享属性、加密属性。
- 当整机镜像为区域内可用时（状态栏为“正常”，且未提示“AZx可用”），那么该整机镜像可以在整个区域内创建云服务器，且云服务器系统盘和数据盘的加密属性会和原始云服务器的系统盘和数据盘的加密属性保持一致，支持在创建时修改数据盘的SCSI模式、共享属性、加密属性。
- 不支持选择使用ISO文件注册的ISO镜像创建弹性云服务器，必须完成操作系统及驱动安装将ISO镜像做成系统盘镜像后才能使用。
- ISO镜像创建的云服务器只建议用来装机，部分功能受限。如：不支持挂载磁盘操作。
- 为了确保使用私有镜像创建的弹性云服务器开启网卡多队列功能，建议您在创建私有镜像时完成网卡多队列的配置。弹性云服务器开启网卡多队列功能可以将弹性云服务器中的网卡中断分散给不同的CPU处理，以满足网卡的需求，从而提升网络PPS和带宽性能。

详细操作请参考[如何设置镜像的网卡多队列属性?](#)

操作步骤（通过 ECS 购买页）

1. 登录控制台，进入[购买弹性云服务器](#)页面。
2. 根据业务需要，设置计费模式、区域、可用区等基础配置、选择实例。
3. 在“操作系统”区域，选择“私有镜像”页签。

图 2-14 私有镜像参数设置



4. 通过下拉列表，选择私有镜像。
5. 继续设置其他购买ECS相关参数，完成ECS的购买。

其余参数的设置的详细内容，请参见[自定义购买ECS](#)。

操作步骤（通过 IMS 列表页）

1. 登录控制台，进入[私有镜像](#)页面。
2. 在镜像所在行的“操作”列下，单击“申请服务器”，进入“购买弹性云服务器”页面。

图 2-15 申请云服务器



名称ID	状态	操作系统类型	镜像系统	镜像类型	磁盘容量 (GiB)	加密	创建时间	操作
image-example 60cae765-5511-4652-9...	正常	Linux	CentOS 7.5 64bit	ECS系统盘镜像(96)	40	否	2023/09/31 19:33:14	申请服务器 修改 更多

3. 根据业务需要，在“购买弹性云服务器”页面完成参数的设置。
参数设置的详细内容，请参见[自定义购买ECS](#)。

说明

- 使用系统盘私有镜像创建ECS时，ECS的规格可以重新设置，系统盘的类型也可以重新选择，但系统盘的容量只能设置为大于镜像的系统盘。
- 使用整机私有镜像创建ECS时，如果整机镜像中包含了一块或多块数据盘，系统会自动设置好数据盘参数。您可以增加系统盘和数据盘容量，但不能缩小。
- 如果整机镜像中包含多块数据盘，需要加载一段时间才能正常显示系统盘信息，请耐心等待。

后续操作

通过私有镜像购买的ECS，可能因缺少插件而无法重置密码，请检查安装一键式重置密码插件。

详细内容，请参见[安装单台ECS的一键式重置密码插件（单台操作）](#)。

2.3.6 通过智能购买助手购买 ECS

操作场景

智能购买助手支持跨计费模式和可用区部署不同类型实例，包含启动模板和智能购买组。

- 启动模板是一种用于快速购买ECS的工具，包含了用于创建ECS的除密码之外的任意配置信息，可以批量购买相同配置的ECS。
更多信息，请参见[启动模板管理](#)。
- 智能购买组是一种新的算力获取方式，通过启动模板的自定义配置，可以批量购买不同可用区、不同实例类型以及不同计费模式的ECS。
更多信息，请参见[智能购买组管理](#)。

约束与限制

- 支持的区域包括：亚太-新加坡、中国-香港。
- 仅支持跨可用区部署ECS实例，不支持跨区域部署ECS实例。
- 通过智能购买组部署ECS实例，其目标容量限制：
 - 以实例个数为目标容量：最多创建500台实例。

- 以vCPUs为目标容量：最大创建40000vCPUs。
- 每个智能购买组最多指定一个启动模板。
- 智能购买组免费，但通过智能购买组创建的实例正常计费。

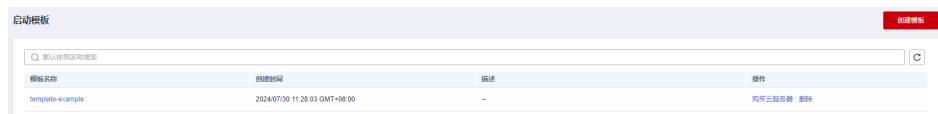
前提条件

已创建实例启动模板。具体操作，请参见[创建启动模板](#)。

通过启动模板快速购买 ECS

1. 在启动模板列表中，单击“操作”列的“购买云服务器”，进入“购买弹性云服务器”页面。

图 2-16 启动模板列表



2. 在“购买弹性云服务器”页面，已根据启动模板的配置信息，预设了购买ECS的参数，您可以根据业务需要进行修改。
例如，修改“登录凭证”方式为“密码”，并设置ECS的密码。
参数设置的详细内容，请参见[自定义购买ECS](#)。
3. 在购买页右侧的“配置概要”中，确认配置信息。
4. 确认无误后，勾选协议并单击“立即购买”。
ECS创建成功后会默认开机。

通过智能购买组跨计费模式和可用区购买 ECS

1. 登录控制台，进入[创建智能购买组](#)页面。
2. 设置智能购买组的名称、目标容量以及按需实例容量。

图 2-17 智能购买组

智能购买组名称	<input type="text" value="alg-example"/>
目标容量	<input type="text" value="实例个数"/> <input type="button" value="-"/> <input type="text" value="5"/> <input type="button" value="+"/> <input type="button" value="?"/>
	<input checked="" type="checkbox"/> 使用按需实例提供部分算力
按需实例容量	<input type="text" value="-"/> <input type="text" value="3"/> <input type="text" value="+"/> 实例个数

表 2-29 参数设置说明

参数	示例	说明
智能购买组名称	alg-example	智能购买组的名称。 长度为2~64个字符，只能包含中文、字母、数字、下划线（_）和中划线（-）。
目标容量	实例个数，5	用于设置使用智能购买组购买的实例目标总容量。 您可以通过目标容量指定智能购买组交付的算力总和，支持设置为实例个数或vCPU数量。每个智能购买组的目标容量限制： <ul style="list-style-type: none"> ● 如果以实例个数为目标容量最多创建500台实例。 ● 如果以vCPUs为目标容量最大创建40000vCPUs。
按需实例容量	3	仅当勾选“使用按需实例提供部分算力”时，会显示该参数。 用于设置按需实例的容量，取值小于等于“目标容量”。

3. 设置启动模板、实例配置及实例分配策略。

图 2-18 实例配置信息



表 2-30 参数设置说明

参数	示例	说明
启动模板	alg-example	用于设置智能购买组在创建实例时，选择启动模板作为配置信息的来源。通过“展开启动模板详情”可查看所选启动模板的具体配置信息。
实例配置	可用区1, s7.medium.2	用于在启动模板配置信息基础上自定义“指定可用区”和“规格”。支持添加多个实例自定义配置，推荐您选择多个不同可用区和实例规格，提升资源交付的成功率。
实例分配策略	最低价格策略	用于设置实例分配遵循的策略。 <ul style="list-style-type: none"> • 最低价格策略：智能购买组购买的所有实例的价格总和最低。 • 多可用区均衡策略：智能购买组购买的实例在多个可用区之间均匀分配算力。 • 容量优化策略：智能购买组购买的实例按照大规格优先的原则进行购买。即目标容量为实例个数时智能购买组的vCPU容量最大；目标容量为vCPU大小时智能购买组的实例个数最小。

4. 设置智能购买组类型、启动/结束时间等参数。

图 2-19 智能购买组信息智能购买组信息智能购买组信息

智能购买组类型 一次性交付 持续保持 ?

启动时间 立即启动 指定启动时间 ?

结束时间 无限期 指定结束时间 ?

全局竞价实例的价格上限 设定价格上限 ?

全局竞价实例的价格上限 \$ /小时

组内释放设置 智能购买组过期时释放实例 ? 超过容量时释放实例 ?

表 2-31 参数设置说明

参数	示例	说明
智能购买组类型	一次性交付	用于设置智能购买组的类型。 <ul style="list-style-type: none">• 一次性交付：智能购买组仅在启动时尝试一次性创建目标容量的实例，即使未达到目标容量也不再重试。• 持续保持：智能购买组实时监控当前集群的容量与目标容量，如未达到目标容量继续创建实例，直至满足目标容量。
启动时间	立即启动	用于设置智能购买组开始创建实例的时间。 <ul style="list-style-type: none">• 立即启动：智能购买组创建成功后立即启动。• 指定启动时间：自定义智能购买组的启动时间。
结束时间	无限期	与启动时间相结合，用于设置智能购买组的有效时间。 <ul style="list-style-type: none">• 无限期：智能购买组长期有效。• 指定结束时间：自定义智能购买组的过期时间。
全局竞价实例的价格上限	0.12	仅当勾选了“设定价格上限”时，会显示该参数。用于设置智能购买组内竞价实例的价格上限。 当智能购买组内某一竞价实例的市场价格高于该价格上限时，该竞价实例将被中断回收。如果同时设置了单个竞价实例的价格上限和全局竞价实例的价格上限，则以单个竞价实例的价格上限为准。
组内释放设置	<ul style="list-style-type: none">• 智能购买组过期时释放• 实例超过容量时释放实例	当设置了智能购买组的结束时间时，该参数有效。 <ul style="list-style-type: none">• 智能购买组过期时释放：当智能购买组超过有效时间，即删除组内实例。• 实例超过容量时释放实例：当智能购买组的实例个数/vCPU之和超过目标容量时，即删除超出容量的实例。

5. 单击“立即创建”，开始创建实例。

2.3.7 通过共享 VPC 购买 ECS

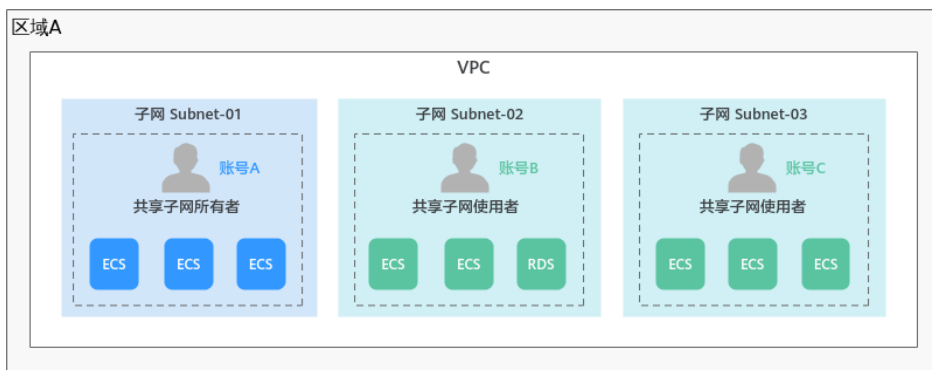
操作场景

购买ECS时，您可以通过共享VPC功能，使用其他账号共享的VPC和子网，以实现网络资源的共享和统一管理，提升资源管控效率、降低运维成本。

例如，为了规范管理网络资源，某企业使用账号A作为IT管理账号，用于管理基础公共资源，包括VPC、子网等。同时，账号A将多个子网共享给其他账号共同使用。

- 账号A：IT管理账号，作为资源所有者，创建VPC及子网，并将多个子网分别共享给其他账号使用。
- 账号B：业务账号，作为资源使用者，使用账号A共享的子网2创建ECS。
- 账号C：业务账号，作为资源使用者，使用账号A共享的子网3创建ECS。

图 2-20 业务规划示意图



本章节介绍通过共享VPC购买ECS的操作指导。有关VPC子网共享的更多信息，请参见《虚拟私有云用户指南》的“[共享VPC](#)”相关内容。

约束与限制

- 单个资源使用者最多可同时接收100个共享子网，当共享子网数量超过100个时，使用者将无法接收到超出数量的共享子网。
- 共享VPC功能不计费，资源使用者只需要为自己所创建的资源付费。

前提条件

作为资源所有者的账号A已创建共享VPC和子网，并指定资源使用者为账号B。创建共享的详细操作，请参见[创建共享](#)。

操作步骤

1. 以账号B登录控制台，进入[资源访问管理](#)页面，接受共享邀请。
详细内容，请参见[接受/拒绝共享邀请](#)。
2. 进入[购买弹性云服务器](#)页面。
3. 根据业务需要，完成购买ECS的基础配置、网络配置以及高级配置等各项参数的设置。
设置“网络”相关参数时，选择由账号A共享的VPC和子网。

图 2-21 设置网络参数



其余配置，请参见[购买弹性云服务器](#)。

2.3.8 相同配置 ECS

操作场景

对于已购买成功的弹性云服务器，如需再次购买相同配置的，建议您使用“购买相同配置”功能，快速购买同一配置的弹性云服务器，节省时间。

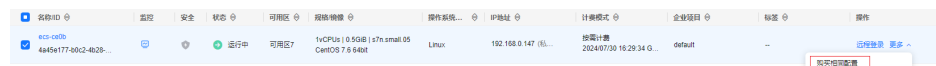
约束与限制

超大内存型弹性云服务器，暂不支持使用“购买相同配置”功能。

操作步骤

1. 登录控制台，进入[弹性云服务器](#)页面。
2. 选择目标弹性云服务器，并单击“操作”列下的“更多 > 购买相同配置”。

图 2-22 购买相同配置



系统跳转至“购买弹性云服务器”页面，并自动复制已选择云服务器的参数配置。

3. 根据需要调整待购买ECS的参数配置。
参数设置的详细内容，请参见[购买弹性云服务器](#)。
4. 确认无误后，勾选协议并单击“立即购买”。

说明

为安全起见，当复制的弹性云服务器存在如下场景时，系统不会自动复制相应的参数配置，需用户手动配置：

- 数据盘个数超过10个时，需手动添加磁盘数量。
- 网卡个数超过5个时，需手动添加其余网卡。
- 安全组个数超过5个时，需手动添加其余安全组。
- 使用数据盘镜像创建的磁盘，需重新选择数据盘镜像。
- 使用整机镜像创建的弹性云服务器，默认只显示整机镜像关联的磁盘，需手动添加新增磁盘。
- 使用磁盘加密功能的云硬盘，需重新勾选“加密”。
- “高级配置”中的相关功能，需重新设置。

2.4 登录 Windows ECS

2.4.1 Windows ECS 登录方式概述

约束与限制

- 只有运行中的云服务器才允许用户登录。
- Windows操作系统用户名“Administrator”。
- 忘记密码，请先通过“重置密码”功能设置登录密码。
重置密码：选中待重置密码的云服务器，并选择“操作”列下的“更多 > 重置密码”。详细操作，请参见[在控制台重置ECS密码](#)。
- 对于密钥方式鉴权的弹性云服务器，需先通过管理控制台提供的获取密码功能，将创建弹性云服务器时使用的私钥文件解析为密码。
- GPU实例中，部分G系列实例不支持云平台提供的远程登录功能，需要自行安装VNC Server进行登录。详细信息请参见[GPU加速型](#)。推荐使用MSTSC方式登录弹性云服务器。
- 使用MSTSC方式访问GPU加速型弹性云服务器时，使用WDDM驱动程序模型的GPU将被替换为一个非加速的远程桌面显示驱动程序，造成GPU加速能力无法实现。因此，如果需要使用GPU加速能力，您必须使用不同的远程访问工具，如VNC工具。如果使用管理控制台提供的“远程登录”功能无法满足您的访问需求，请自行在弹性云服务器上安装符合要求的远程访问工具（如Tight VNC）。
Tight VNC下载地址：<https://www.tightvnc.com/download.php>

登录方式概述

请根据需要选择登录方式，登录云服务器。

表 2-32 Windows 云服务器登录方式一览

云服务器操作系统	本地主机操作系统	连接方法	条件
Windows	Windows	使用mstsc方式登录云服务器。 在本地主机单击“开始”菜单，输入mstsc命令，打开远程桌面对话框。 通过本地Windows主机登录Windows ECS 。	云服务器器绑定弹性公网IP (通过内网登录云服务器时可以不绑定弹性公网IP，例如VPN、云专线等内网网络连通场景。)
	Linux	安装远程连接工具，例如rdesktop，执行连接命令。 通过本地Linux主机登录Windows ECS 。	

云服务器操作系统	本地主机操作系统	连接方法	条件
	macOS系统	安装远程连接工具，例如Microsoft Remote Desktop for Mac在macOS系统上登录。 通过本地macOS主机登录Windows ECS。	
	移动设备	安装远程连接工具，例如Microsoft Remote Desktop在移动设备上登录。 通过移动设备登录Windows ECS。	
	Windows	使用管理控制台远程登录方式： 通过控制台VNC登录Windows ECS。	不依赖弹性公网IP

相关链接

- [忘记密码怎么办？](#)
- [Windows云服务器如何配置多用户登录？](#)
- [无法登录到Windows云服务器怎么办？](#)

2.4.2 通过控制台 VNC 登录 Windows ECS

操作场景



本节为您介绍如何通过控制台提供的远程登录功能（即VNC方式）登录到弹性云服务器上。

当您无法使用MSTSC或其他远程登录工具连接弹性云服务器时，可以选择通过VNC方式进行登录，该登录方式主要用于紧急运维场景，可查看弹性云服务器状态或进行相关维护操作。

前提条件

如果您的弹性云服务器是采用密钥方式鉴权，请在登录前先使用密钥文件解析登录密码。密钥文件解析密码请参考[获取Windows弹性云服务器的密码](#)。

登录 Windows 弹性云服务器

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 获取弹性云服务器密码。
VNC方式登录弹性云服务器时，需已知其密码，然后再采用VNC方式登录。

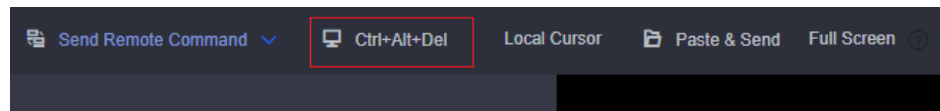
- 当您的弹性云服务器是采用密码方式鉴权时，请直接使用创建云服务器时设置的密码进行登录。
 - 当您的弹性云服务器是采用密钥方式鉴权时，密码获取方式请参见[获取Windows弹性云服务器的密码](#)。
5. 选择要登录的弹性云服务器，单击“操作”列下的“远程登录”。

图 2-23 远程登录



6. 在弹出的“登录Windows弹性云服务器”窗口中，选择“其他方式”下的VNC方式，单击“立即登录”。
7. （可选）如果界面提示“Press Ctrl+Alt+Delete to unlock”，请单击远程登录操作面板上方的“Ctrl+Alt+Del”按钮进行登录。

图 2-24 单击“Ctrl+Alt+Del”



8. 根据界面提示，输入弹性云服务器密码。

相关链接

- [忘记密码怎么办？](#)
- [Windows云服务器如何配置多用户登录？](#)
- [无法登录到Windows云服务器怎么办？](#)

2.4.3 通过本地 Windows 主机登录 Windows ECS

操作场景

本节为您介绍如何在本机使用远程登录工具MSTSC登录Windows弹性云服务器。

前提条件

- 弹性云服务器状态为“运行中”。
- 如果弹性云服务器采用密钥方式鉴权，已获取Windows弹性云服务器的密码，获取方式请参见[获取Windows弹性云服务器的密码](#)。
- 弹性云服务器已经绑定弹性公网IP，绑定方式请参见[绑定弹性公网IP](#)。

使用MSTSC方式通过内网登录云服务器时可以不绑定弹性公网IP，例如VPN、云专线等内网网络连通场景。

- 所在安全组入方向已开放3389端口，配置方式请参见[配置安全组规则](#)。
- 使用的登录工具与待登录的弹性云服务器之间网络连通。例如，默认的3389端口没有被防火墙屏蔽。
- 弹性云服务器开启远程桌面协议RDP（Remote Desktop Protocol）。使用公共镜像创建的弹性云服务器默认已打开RDP。打开RDP方法请参考[开启远程桌面协议RDP](#)。

使用 MSTSC 方式登录 Windows 弹性云服务器

本地主机为Windows操作系统，那么可以使用Windows自带的远程桌面连接工具MSTSC登录Windows云服务器。

以下演示了使用MSTSC方式登录Windows 2012操作系统云服务器的操作步骤。

图 2-25 MSTSC 方式登录云服务器

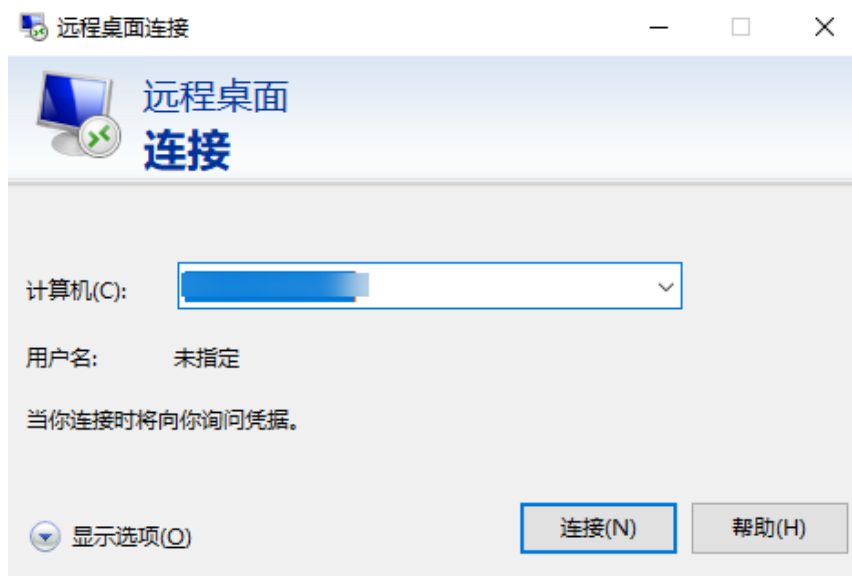


详细操作步骤如下：

1. 在本地主机单击“开始”菜单。
2. 在“搜索程序和文件”中，输入“mstsc”，单击mstsc打开远程桌面连接工具。

3. 在“远程桌面连接”的对话框中，单击“选项”。

图 2-26 显示选项

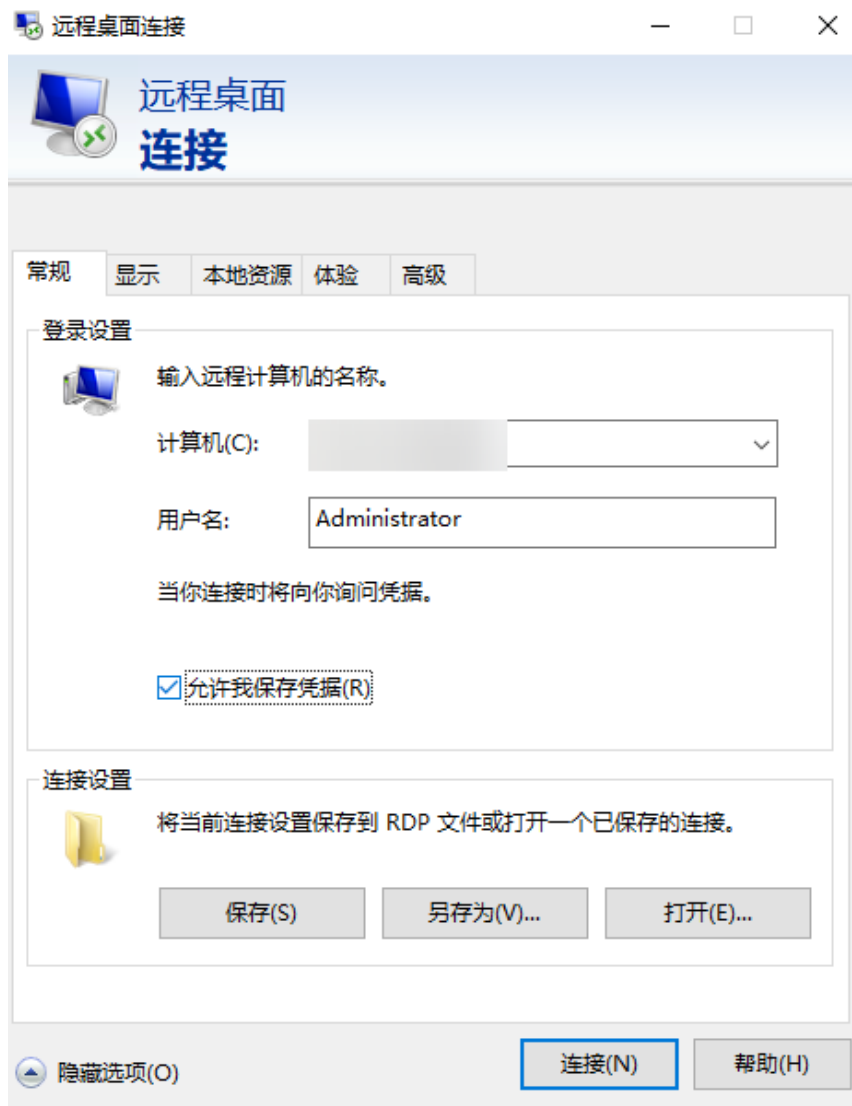


4. 输入待登录的云服务器的弹性公网IP和用户名，默认为Administrator。

说明

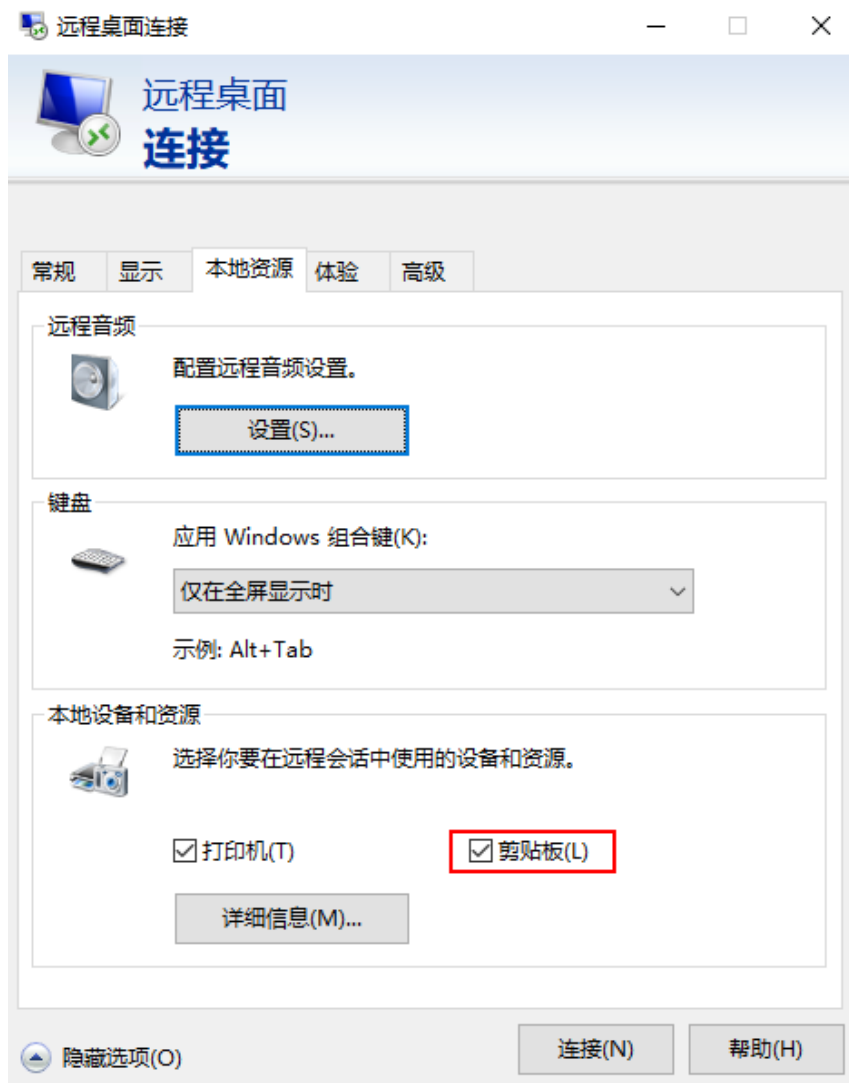
如需再次登录时不再重复输入用户名和密码，可勾选“允许我保存凭据”。

图 2-27 远程桌面链接



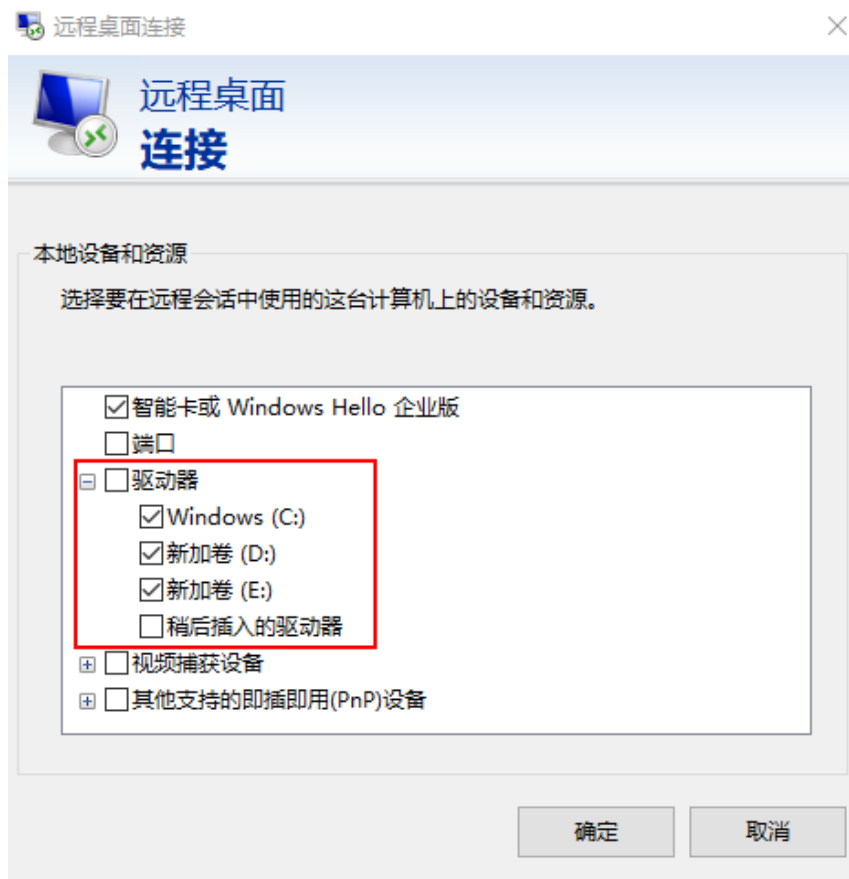
5. （可选）如需在远程会话中使用本地主机的资源，请单击“本地资源”选项卡完成如下配置。
如需从本地主机复制到云服务器中，请勾选“剪贴板”。

图 2-28 勾选剪贴板



如需从本地主机复制文件到云服务器中，单击“详细信息”，勾选相应的磁盘。

图 2-29 勾选驱动器



6. (可选) 如需调整远程桌面窗口的大小, 可以选择“显示”选项卡, 再调整窗口大小。

图 2-30 调整窗口大小



7. 单击“确定”，根据提示输入密码，登录云服务器。
为安全起见，首次登录云服务器，需更改密码。
8. （可选）通过远程桌面连接（Remote Desktop Protocol，RDP）方式登录云服务器后，如果需要使用RDP提供的“剪切板”功能，将本地的大文件（文件大小超过2GB）复制粘贴至远端的Windows云服务器中，由于Windows系统的限制，会导致操作失败。
具体的解决方法，请参考[使用远程桌面链接方式复制文件](#)。

开启远程桌面协议 RDP

首次登录弹性云服务器时，请先使用VNC方式登录弹性云服务器，打开RDP（Remote Desktop Protocol），然后再使用mstsc方式连接。

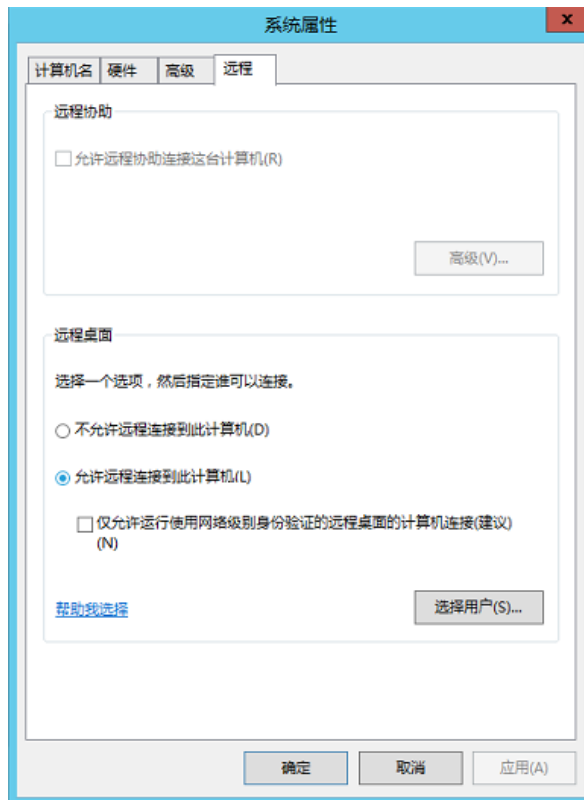
说明

使用公共镜像创建的云服务器，默认已打开RDP。

1. VNC方式登录弹性云服务器。
登录方法请参见[通过控制台VNC登录Windows ECS](#)。

- 单击“开始”菜单，选择“控制面板 > 系统和安全 > 系统 > 远程设置”。系统进入“系统属性”页面。

图 2-31 系统属性



- 选择“远程”页签，在“远程桌面”栏，选择“允许远程连接到此计算机”。
- 单击“确定”。

相关链接

- [忘记密码怎么办？](#)
- [Windows云服务器如何配置多用户登录？](#)
- [无法登录到Windows云服务器怎么办？](#)

2.4.4 通过本地 Linux 主机登录 Windows ECS

操作场景

本节为您介绍如何在Linux操作系统主机上登录Windows云服务器。

前提条件

- 云服务器状态为“运行中”。
- 云服务器已经绑定弹性公网IP。
使用MSTSC方式通过内网登录云服务器时可以不绑定弹性公网IP，例如VPN、云专线等内网网络连通场景。

- 所在安全组入方向已开放3389端口。
- 使用的登录工具与待登录的云服务器之间网络连通。例如，默认的3389端口没有被防火墙屏蔽。
- 云服务器开启远程桌面协议RDP (Remote Desktop Protocol)。使用公共镜像创建的云服务器默认已打开RDP。打开RDP方法请参考[开启远程桌面协议RDP](#)。

操作步骤

如果本地主机为Linux操作系统，您可以使用远程连接工具（例如rdesktop）连接Windows实例。

1. 执行以下命令，检查云服务器是否安装rdesktop。

rdesktop

如果提示“command not found”说明未安装rdesktop。请参考[rdesktop工具官方](#)获取rdesktop安装包安装rdesktop。

2. 输入以下命令登录云服务器。

rdesktop -u 用户名 -p 密码 -g 分辨率 弹性公网IP地址

例如：**rdesktop -u administrator -p password -g 1024*720 121.xx.xx.xxx**

表 2-33 远程登录命令参数

参数	说明
-u	用户名，Windows实例默认用户名是Administrator。
-p	登录Windows实例的密码。
-f	默认全屏，需要用 Ctrl+Alt+Enter 组合键进行全屏模式切换。
-g	分辨率，中间用星号（*）连接，可省略，省略后默认为全屏显示。例如： 1024*720
弹性公网IP地址	需要远程连接的服务器IP地址。需要替换为您的Windows实例的弹性公网IP地址。

开启远程桌面协议 RDP

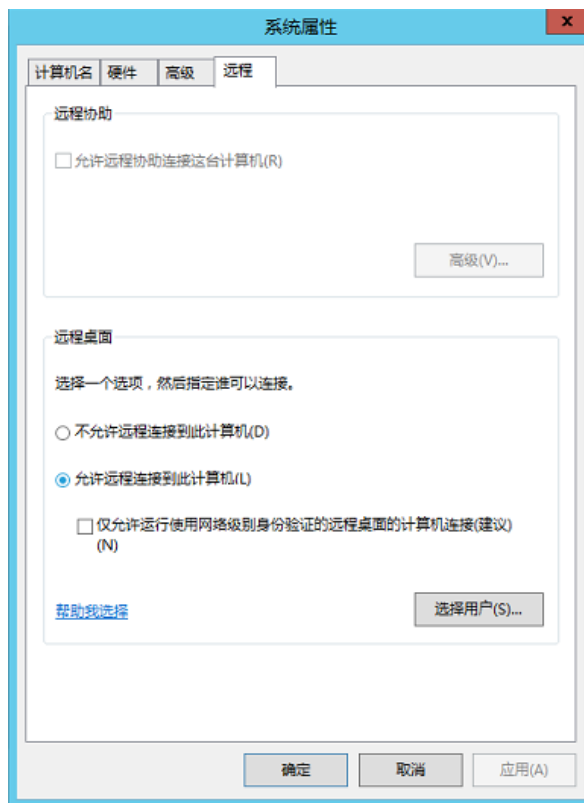
首次登录弹性云服务器时，请先使用VNC方式登录弹性云服务器，打开RDP（Remote Desktop Protocol），然后再使用mstsc方式连接。

说明

使用公共镜像创建的云服务器，默认已打开RDP。

1. VNC方式登录弹性云服务器。
登录方法请参见[通过控制台VNC登录Windows ECS](#)。
2. 单击“开始”菜单，选择“控制面板 > 系统和安全 > 系统 > 远程设置”。
系统进入“系统属性”页面。

图 2-32 系统属性



3. 选择“远程”页签，在“远程桌面”栏，选择“允许远程连接到此计算机”。
4. 单击“确定”。

2.4.5 通过本地 macOS 主机登录 Windows ECS

操作场景

本节操作以使用“Microsoft Remote Desktop for Mac”工具远程连接“Windows Server 2012 R2 数据中心版 64位”操作系统云服务器为例，介绍macOS系统登录Windows云服务器的操作步骤。

前提条件

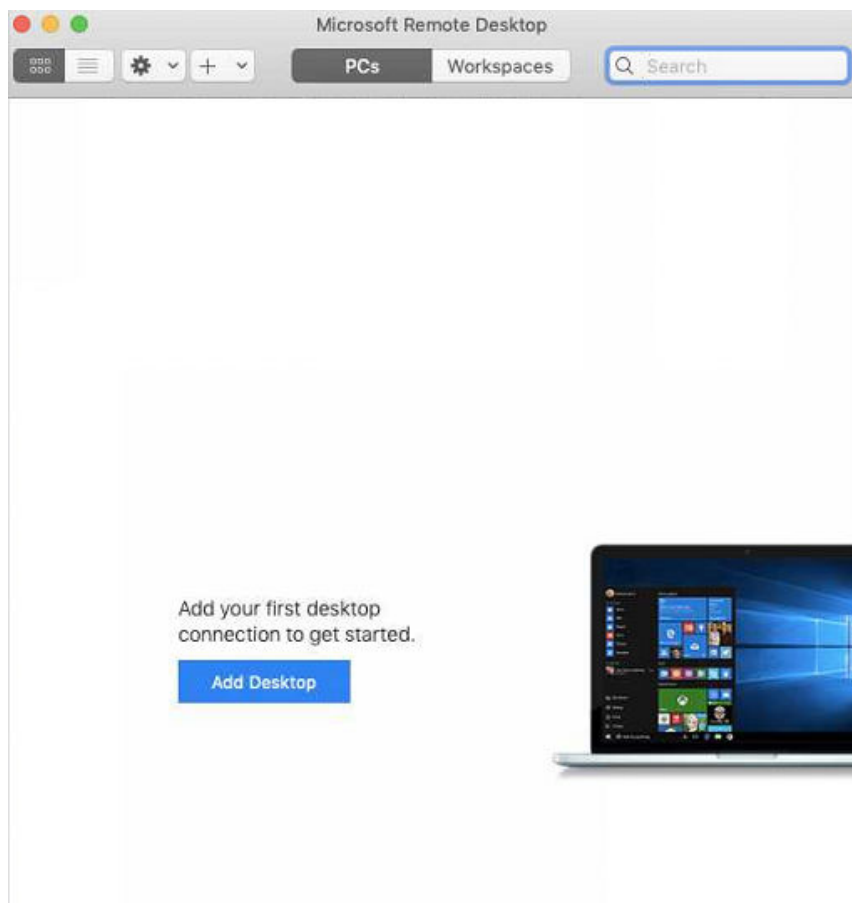
- 云服务器状态为“运行中”。
- 已获取Windows云服务器用户名和密码。忘记密码请参考[在控制台重置ECS密码重置密码](#)。
- 弹性云服务器已经绑定弹性公网IP，绑定方式请参见[绑定弹性公网IP](#)。
- 所在安全组入方向已开放3389端口，配置方式请参见[配置安全组规则](#)。
- 已安装Microsoft Remote Desktop for Mac或其他macOS系统适用的远程连接工具。

微软官方已停止提供Remote Desktop客户端的下载链接，您可单击[Microsoft Remote Desktop Beta](#)下载Beta版本使用。

操作步骤

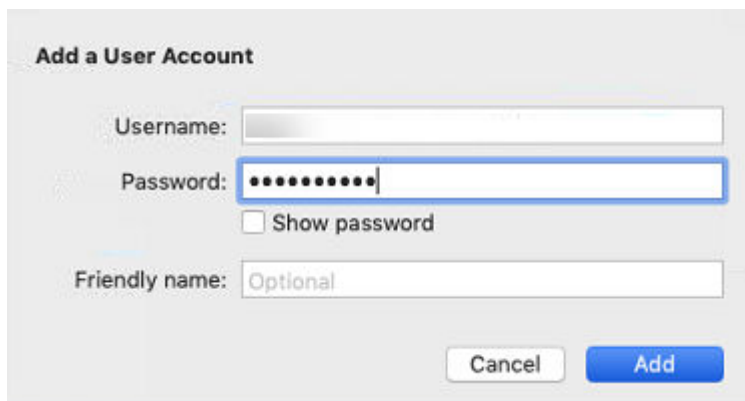
1. 启动Microsoft Remote Desktop。
2. 单击“Add Desktop”。

图 2-33 Add Desktop



3. 在“Add PC”页面，设置登录信息。
 - PC name: 输入需要登录的Windows实例的弹性公网IP地址。
 - User account: 在下拉列表中选择“Add user account”。弹出“Add user account”对话框。
 - i. 输入Windows实例账号“administrator”，并输入实例的登录密码，单击“Add”。

图 2-34 Add user account



Add a User Account

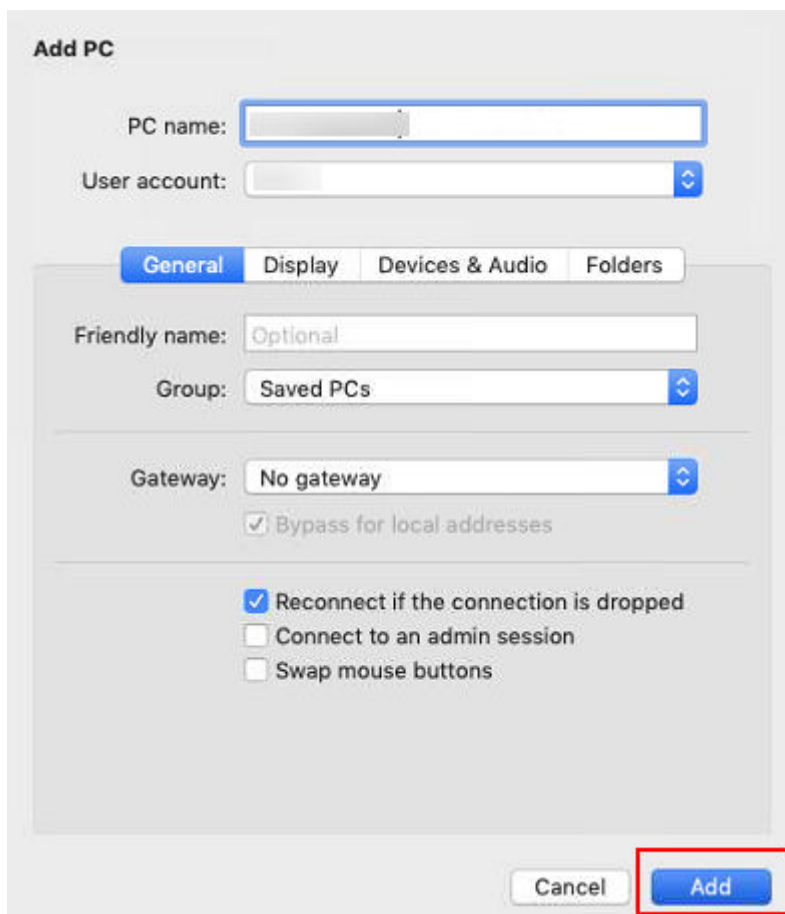
Username:

Password:

Show password

Friendly name:

图 2-35 Add PC



Add PC

PC name:

User account:

General | Display | Devices & Audio | Folders

Friendly name:

Group:

Gateway:

Bypass for local addresses

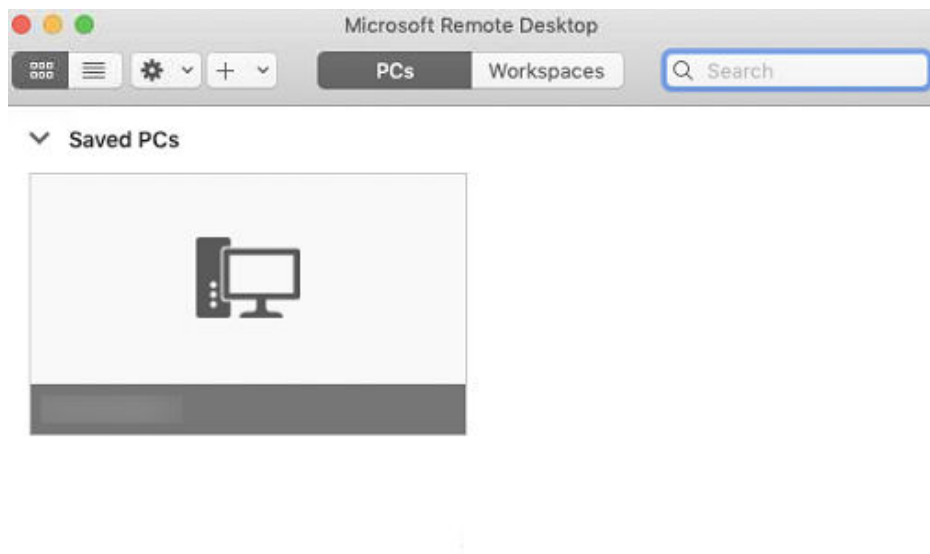
Reconnect if the connection is dropped

Connect to an admin session

Swap mouse buttons

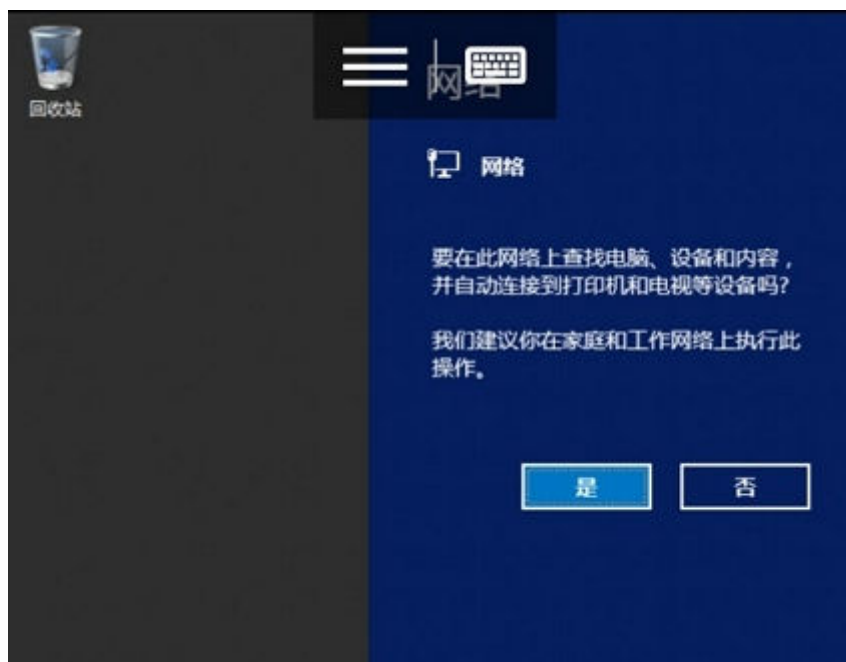
4. 在“Remote Desktop”页面，双击需要登录的Windows实例图标。

图 2-36 双击登录 Windows 实例



5. 确认信息后，单击“Continue”。至此，您已经登录Windows实例。

图 2-37 登录成功



2.4.6 通过移动设备登录 Windows ECS

操作场景

本节操作以“Windows Server 2012 R2 数据中心版 64位”操作系统为例，介绍如何通过Microsoft Remote Desktop客户端登录Windows实例。

前提条件

- 云服务器状态为“运行中”。
- 已获取Windows云服务器用户名和密码。忘记密码请参考[在控制台重置ECS密码重置密码](#)。
- 云服务器已经绑定弹性公网IP，绑定方式请参见[绑定弹性公网IP](#)。
- 所在安全组入方向已开放3389端口，配置方式请参见[配置安全组规则](#)。
- 移动设备已通过“应用市场”搜索并下载远程连接工具，本文以“Microsoft Remote Desktop”为例进行介绍。

操作步骤


1. 启动RD Client。
2. 在“Remote Desktop”页面右上角，单击图标，选择“Desktop”。

图 2-38 Remote Desktop

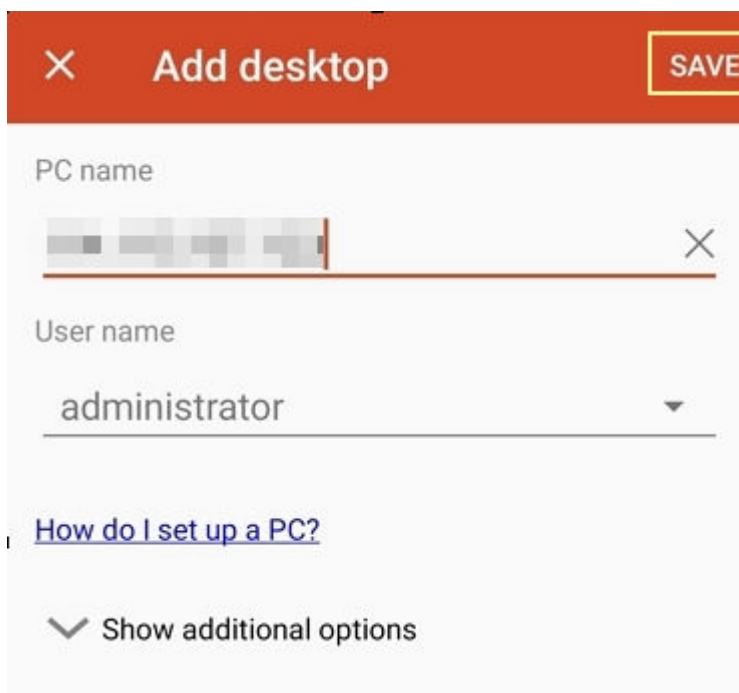


3. 在“Add desktop”页面，设置登录信息后，单击“SAVE”。
 - PC name：输入需要登录的Windows实例的弹性公网IP地址。
 - 按以下步骤设置“User name”：
 - i. 单击“User name”，在下拉列表中选择“Add user account”。弹出“Add user account”对话框。
 - ii. 输入Windows实例账号“administrator”，并输入实例的登录密码，单击“SAVE”。

图 2-39 输入登录信息

The image shows a mobile application interface for adding a desktop. The main dialog is titled "Add desktop" and has a "SAVE" button in the top right corner. Below the title, there are two input fields: "PC name" and "User name". The "User name" field is currently set to "administrator". A sub-dialog titled "Add user account" is overlaid on top of the main dialog. This sub-dialog has two input fields: "User name" (set to "administrator") and "Password" (masked with seven dots). At the bottom of the sub-dialog, there are two buttons: "CANCEL" and "SAVE". The "SAVE" button is highlighted with a red rectangular border.

图 2-40 保存填写的登录信息



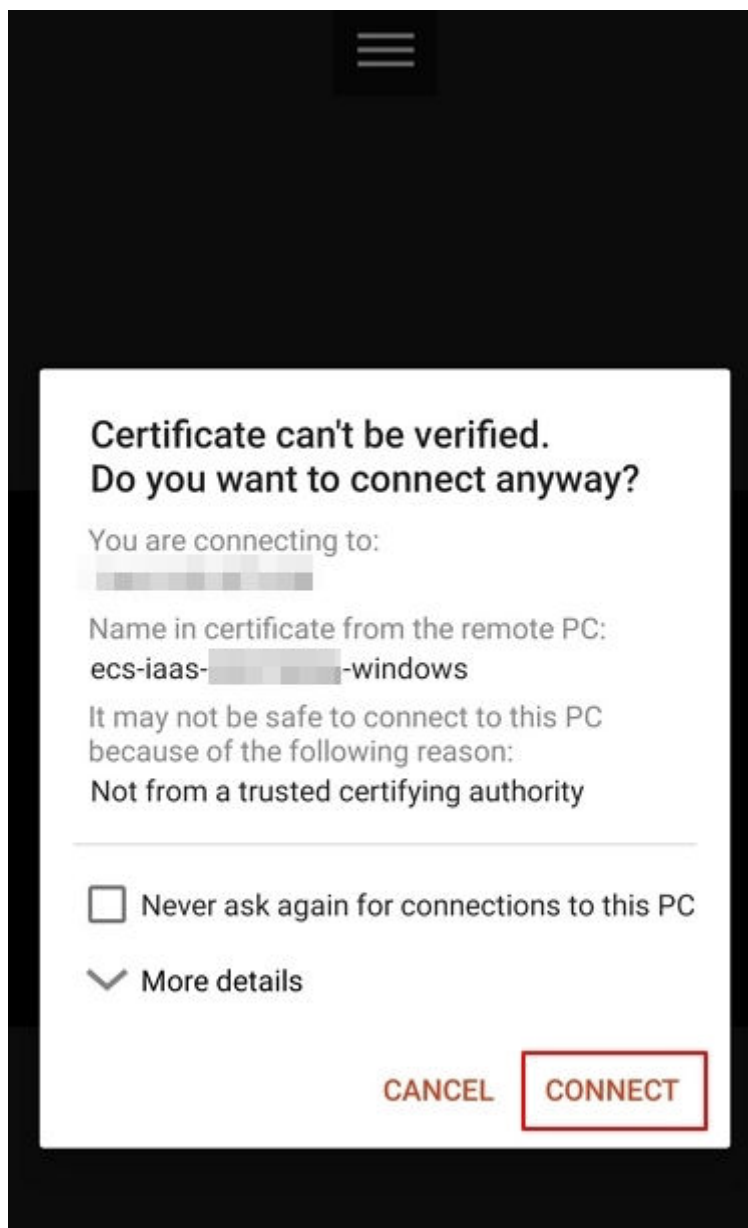
4. 在“Remote Desktop”页面，单击需要登录的Windows实例图标。

图 2-41 登录 Windows 实例



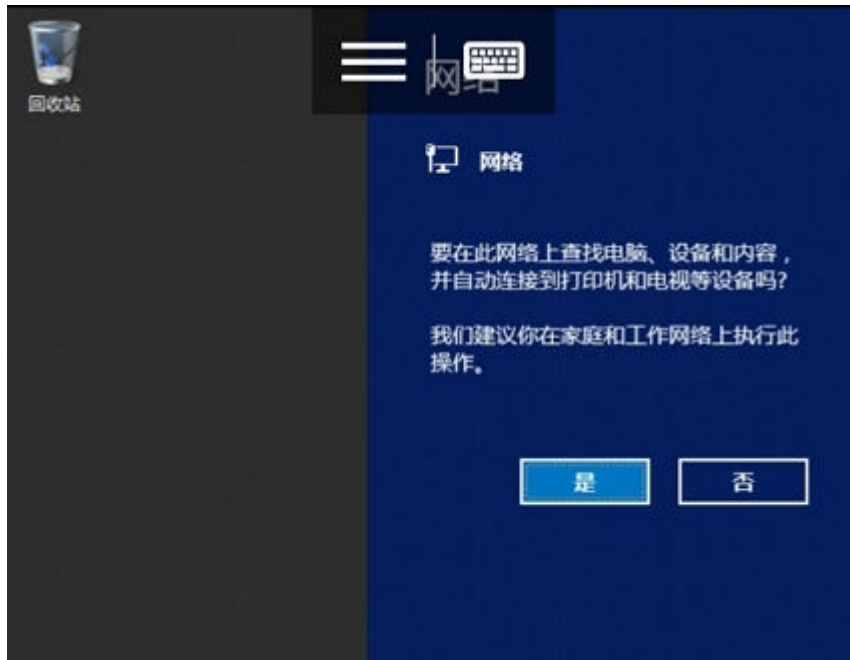
5. 确认信息后，单击“CONNECT”。

图 2-42 CONNECT



至此，您已经登录Windows实例。

图 2-43 登录成功



2.5 登录 Linux ECS

2.5.1 Linux ECS 登录方式概述

约束与限制

- 只有运行中的弹性云服务器才允许用户登录。
- Linux操作系统用户名“root”。
- 忘记密码，请先通过“重置密码”功能设置登录密码。
重置密码：选中待重置密码的云服务器，并选择“操作”列下的“更多 > 重置密码”。详细操作，请参见[在控制台重置ECS密码](#)。

登录方式概述

请根据需要选择登录方式，登录云服务器。

表 2-34 Linux 云服务器登录方式一览

云服务器操作系统	本地主机操作系统	连接方法	条件
Linux	Windows	(推荐使用)使用控制台提供的CloudShell登录云服务器。 通过CloudShell登录Linux ECS。	云服务器绑定弹性公网IP。 (通过内网登录云服务器时可以不绑定弹性公网IP,例如VPN、云专线等内网网络连通场景。)
	Windows	使用PuTTY、Xshell等远程登录工具: <ul style="list-style-type: none">• 密码方式鉴权: SSH密码方式登录(本地使用Windows操作系统)。• 密钥方式鉴权: SSH密钥方式登录(本地使用Windows操作系统)。	
	Linux	使用命令连接: <ul style="list-style-type: none">• 密码方式鉴权: SSH密码方式登录(本地使用Linux操作系统)。• 密钥方式鉴权: SSH密钥方式登录(本地使用Linux操作系统)。	
	移动设备	使用Termius、JuiceSSH等SSH客户端工具登录云服务器: 通过移动设备登录Linux ECS。	
	macOS系统	使用系统自带的终端(Terminal): 通过macOS主机登录Linux ECS。	
	Windows	使用管理控制台远程登录方式: 通过VNC登录Linux ECS。	不依赖弹性公网IP。

相关链接

- [忘记密码怎么办?](#)
- [无法登录到Linux云服务器怎么办?](#)

2.5.2 通过 CloudShell 登录 Linux ECS

操作场景

本节为您介绍通过控制台提供的CloudShell登录云服务器的操作步骤。

登录成功后,如需使用CloudShell界面提供的复制、粘贴功能,请参见[CloudShell常用操作](#)。

约束限制

支持通过CloudShell方式登录弹性云服务器的区域，请参见[功能总览](#)中“连接Linux实例”的发布区域。

前提条件

- 云服务器状态为“运行中”。
- 请确保安全组已开放登录端口，默认使用22端口，如需使用其他端口可登录云服务器后重新设置。
修改远程登录端口请参考[修改远程登录端口](#)。配置安全组规则请参考[配置安全组规则](#)。
- 如果在创建云服务器时未设置密码，请先重置密码后再登录云服务器。
- 使用CloudShell连接云服务器支持公网连接和私网连接两种方式。当使用CloudShell通过私网远程连接云服务器时，需要具有Security Administrator权限的用户进行服务授权。
 - 若用户有Security Administrator权限，在首次使用时，会弹出如下授权页面，单击“同意授权”即可。
服务授权区域级生效，仅需在每个区域首次使用时执行一次，授权后，用户即可在当前区域使用。

图 2-44 服务授权




- 若用户无Security Administrator权限，页面将提示“请添加Security Administrator管理员”，可以联系管理员（具有admin权限的用户）为该用户添加Security Administrator权限后重试。
授权方式为：
 - i. 创建用户组，并给用户组授权添加Security Administrator权限，详细操作请参考[创建用户组并授权](#)。
 - ii. 将用户添加到用户组中，详细操作请参考[用户组添加用户](#)。

📖 说明

使用CloudShell通过公网远程连接云服务器时，无需进行服务授权。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。

- 单击“☰”，选择“计算 > 弹性云服务器”。
- 选择要登录的云服务器，单击“操作”列下的“远程登录”。
- 在弹出的“登录Linux云服务器”窗口中，单击“使用CloudShell登录”。
- 在CloudShell界面配置云服务器信息。
首次登录，默认会打开CloudShell配置向导，输入云服务器的相关参数进行连接。

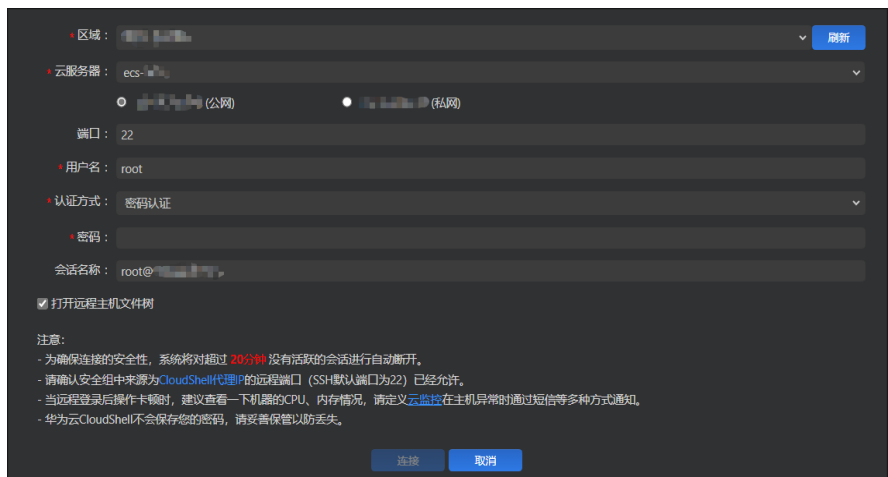
📖 说明

您可以选择使用云服务器的弹性公网IP或私有IP进行登录。

- 若勾选云服务器的公网IP（弹性公网IP）。
 - i. 在CloudShell配置向导中设置云服务器的端口（默认22）、用户名、认证方式、密码（或密钥）等参数。
 - ii. 单击“连接”登录云服务器。

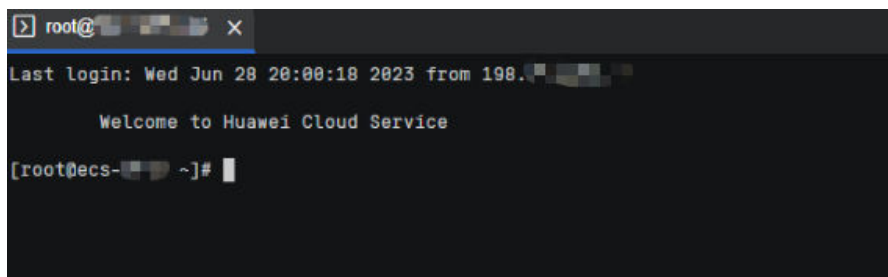
如果单击“连接”没有反应，可能是云服务器未设置登录密码或密码错误，请重置密码后重新登录。

图 2-45 CloudShell 配置向导（公网 IP）



连接成功后，CloudShell界面提示如下。

图 2-46 操作结果



- 若勾选云服务器的私网IP（私有IP）。
 - i. 单击“跳转”，打开新的CloudShell配置向导页。

图 2-47 CloudShell 配置向导（私网 IP）1

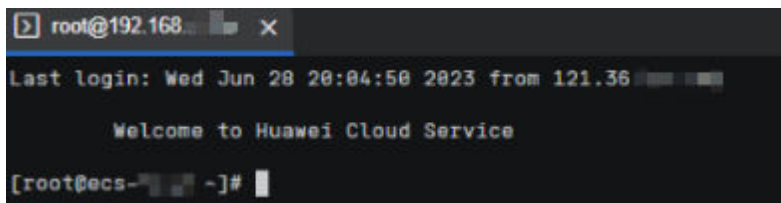


- ii. 在CloudShell配置向导中设置云服务器的端口（默认22）、用户名、认证方式、密码（或密钥）等参数。
- iii. 单击“连接”登录云服务器。

如果单击“连接”没有反应，可能是云服务器未设置登录密码或密码错误，请重置密码后重新登录。

连接成功后，CloudShell界面提示如下。

图 2-48 操作结果



CloudShell 常用操作

- **新建会话**
选择“终端 > 新建会话”，即可配置新的终端连接会话。
- **快捷键**
您可以使用快捷键编辑输入的命令。

表 2-35 CloudShell 快捷键

快捷键	功能
Ctrl+L	将当前行移到第一行
Ctrl+U	清除当前行
Ctrl+H	向前删除一个字符
Ctrl+A	光标移动到句首
Ctrl+E	光标移动到句末

- **复制、粘贴**
CloudShell支持直接在终端中进行复制粘贴。您既可以通过右键来复制粘贴，也可以直接用“Ctrl+C”、“Ctrl+V”等快捷键实现。
- **浏览输出历史**

对于跨屏内容，可以滚动终端查看历史输出。默认情况下，终端只会记录最近1000行输出。

- **多终端分区布局**

您可以在同一个页面中创建多个CloudShell终端，并可以直接拖动窗口，随意组合成您喜欢的布局。

2.5.3 通过 VNC 登录 Linux ECS

操作场景

本节为您介绍如何通过控制台提供的远程登录功能（即VNC方式）登录到弹性云服务器上。

当您无法使用其他远程登录工具连接弹性云服务器时，可以选择通过VNC方式进行登录，该登录方式主要用于紧急运维场景，可查看弹性云服务器状态或进行相关维护操作。

登录成功后，如需使用VNC界面提供的复制、粘贴功能，请参见[后续处理](#)。

说明

对于“密钥对”方式创建的Linux弹性云服务器，如需使用控制台提供的“远程登录”功能（VNC方式），需先使用“[SSH密钥方式](#)”登录，并设置密码，然后才能使用VNC方式登录。

约束与限制

- 使用VNC方式登录弹性云服务器时，系统不支持直接的复制粘贴操作，这将降低您使用云服务器的效率。如无特殊要求，建议使用SSH方式登录云服务器，具体操作请参见[通过SSH密钥方式登录Linux ECS](#)、[通过SSH密码方式登录Linux ECS](#)。

前提条件

对于“密钥对”鉴权方式的弹性云服务器，已使用SSH密钥方式登录Linux弹性云服务器，并设置密码。

登录 Linux 弹性云服务器



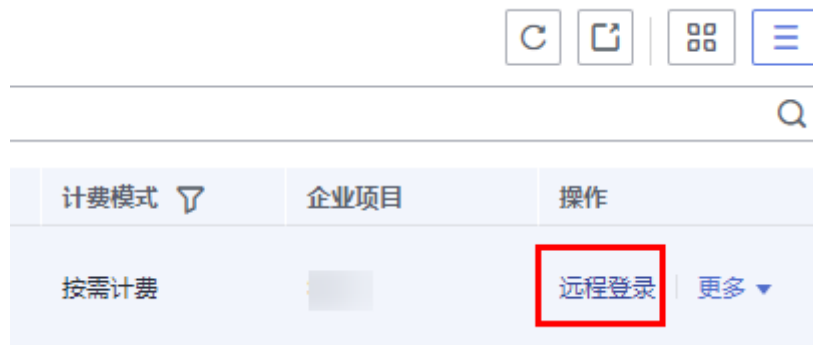
1. 登录管理控制台。
2. 单击管理控制台左上角的  ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 选择要登录的弹性云服务器，单击“操作”列下的“远程登录”。

图 2-49 远程登录



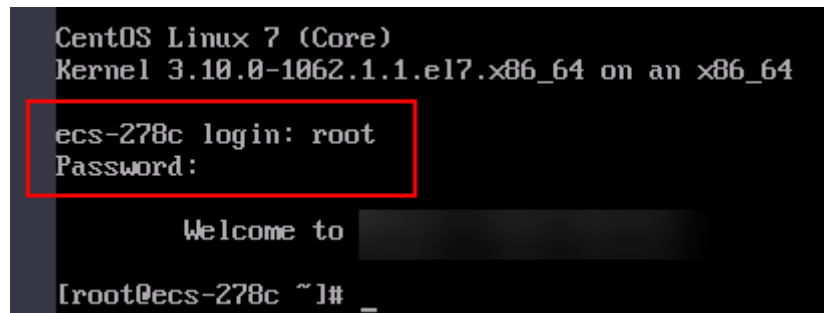
5. （可选）如果界面提示“Press CTRL+ALT+DELETE to log on”，请单击远程登录操作面板上方的“Ctrl+Alt+Del”按钮进行登录。

说明

请勿使用物理键盘按“CTRL+ALT+DELETE”，该操作不生效。

6. 根据界面提示，输入弹性云服务器密码。

图 2-50 输入用户名和密码（假设用户名为 root）



后续处理

系统支持从本地拷贝命令行到弹性云服务器，实现本地数据与弹性云服务器之间的单向复制、粘贴功能，方法如下：

1. 使用VNC方式成功登录弹性云服务器。
2. 单击页面上方的“复制粘贴”。

图 2-51 复制粘贴



3. 使用快捷键Ctrl+C，复制本地计算机的数据。
4. 使用快捷键Ctrl+V，将本地数据粘贴至“复制粘贴”窗口。
5. 单击“发送”。
将复制的数据发送至命令行窗口。

📖 说明

对于使用图形化界面的Linux弹性云服务器，在使用VNC窗口提供的“复制粘贴”功能时，会小概率出现数据丢失的情况。这是由于弹性云服务器CPU核数过低，图形化界面占用较多CPU资源导致。此时，建议您每次发送的字符数不超过5个，或者从图形化界面切换至命令行界面（也称“文本界面”），然后再使用“复制粘贴”功能。

相关链接

- [忘记密码怎么办？](#)
- [无法登录到Linux云服务器怎么办？](#)

2.5.4 通过 SSH 密钥方式登录 Linux ECS

操作场景

本节操作介绍在Windows和Linux环境中使用SSH密钥对方式远程登录Linux云服务器的操作步骤。

前提条件

- 已获取创建该弹性云服务器时使用的密钥对私钥文件，创建密钥对请参见（[推荐](#)）[通过管理控制台创建密钥对](#)。
- 弹性云服务器已经绑定弹性公网IP，绑定方式请参见[查看弹性云服务器详细信息](#)。

- 已配置安全组入方向的访问规则，配置方式请参见[配置安全组规则](#)。
- 使用的登录工具（如PuTTY）与待登录的弹性云服务器之间网络连通。例如，默认的22端口没有被防火墙屏蔽。

本地使用 Windows 操作系统

如果您本地使用Windows操作系统登录Linux弹性云服务器，可以按照下面方式登录弹性云服务器。

方式一：使用PuTTY登录

以PuTTY为例介绍如何登录弹性云服务器。使用PuTTY登录弹性云服务器前，需要先将私钥文件转化为.ppk格式。

1. 判断私钥文件是否为.ppk格式。
 - 是，执行7。
 - 否，执行2。
2. 在以下路径中下载PuTTY和PuTTYgen。

<https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>

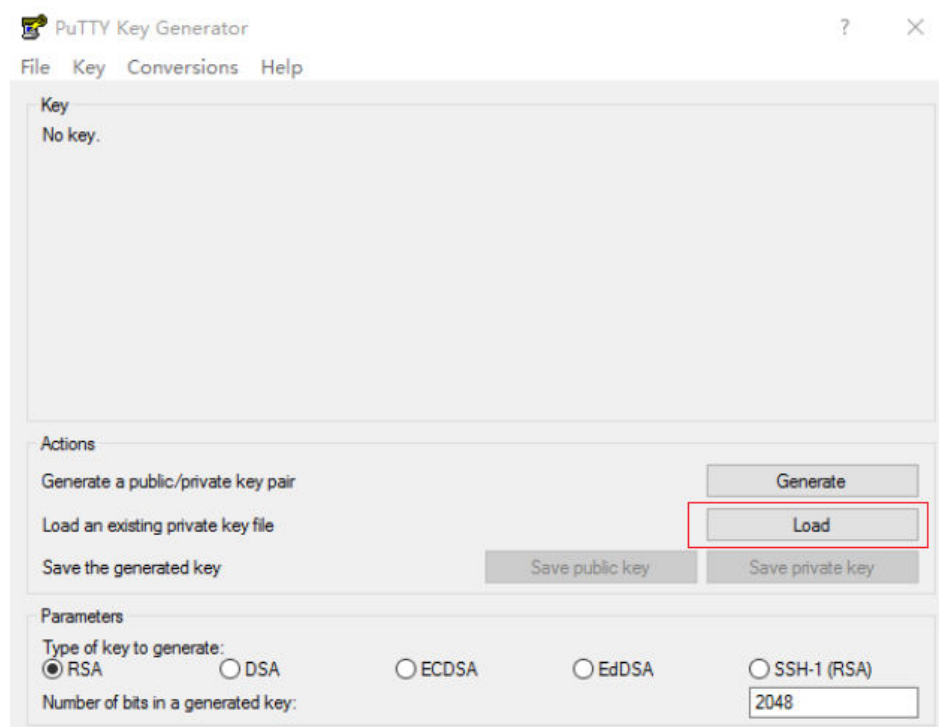
📖 说明

PuTTYgen是密钥生成器，用于创建密钥对，生成一对公钥和私钥供PuTTY使用。

3. 运行PuTTYgen。
4. 在“Actions”区域，单击“Load”，并导入创建弹性云服务器时保存的私钥文件。

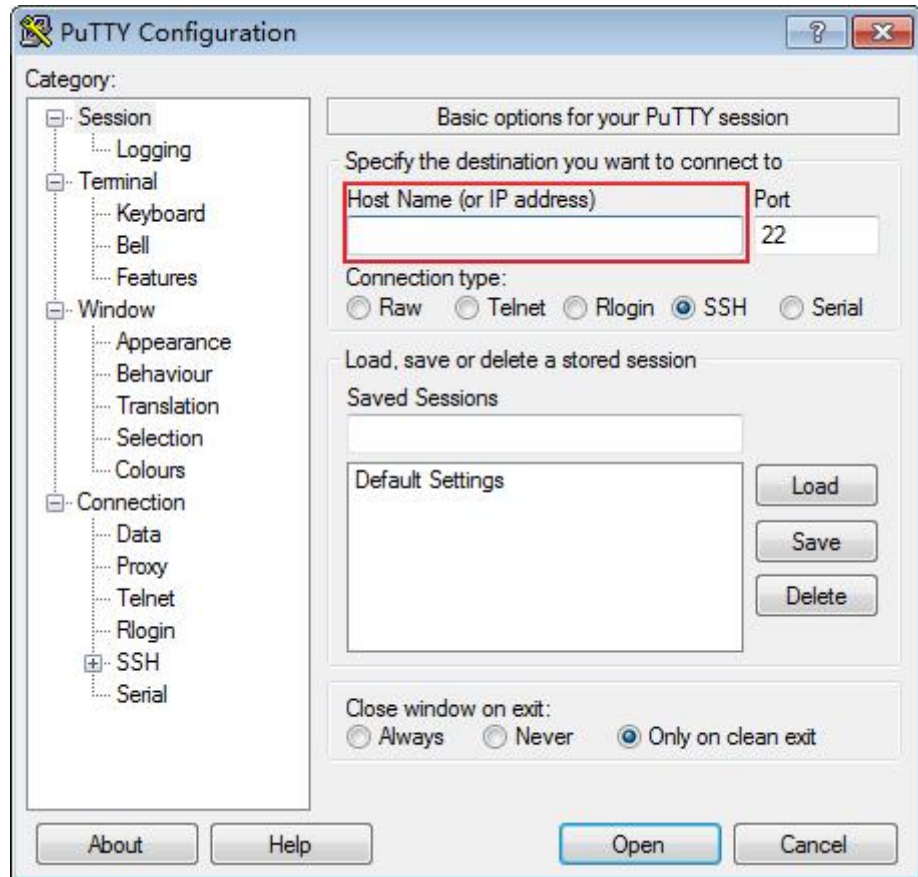
导入时注意确保导入的格式要求为“All files (*.*)”。

图 2-52 导入私钥文件



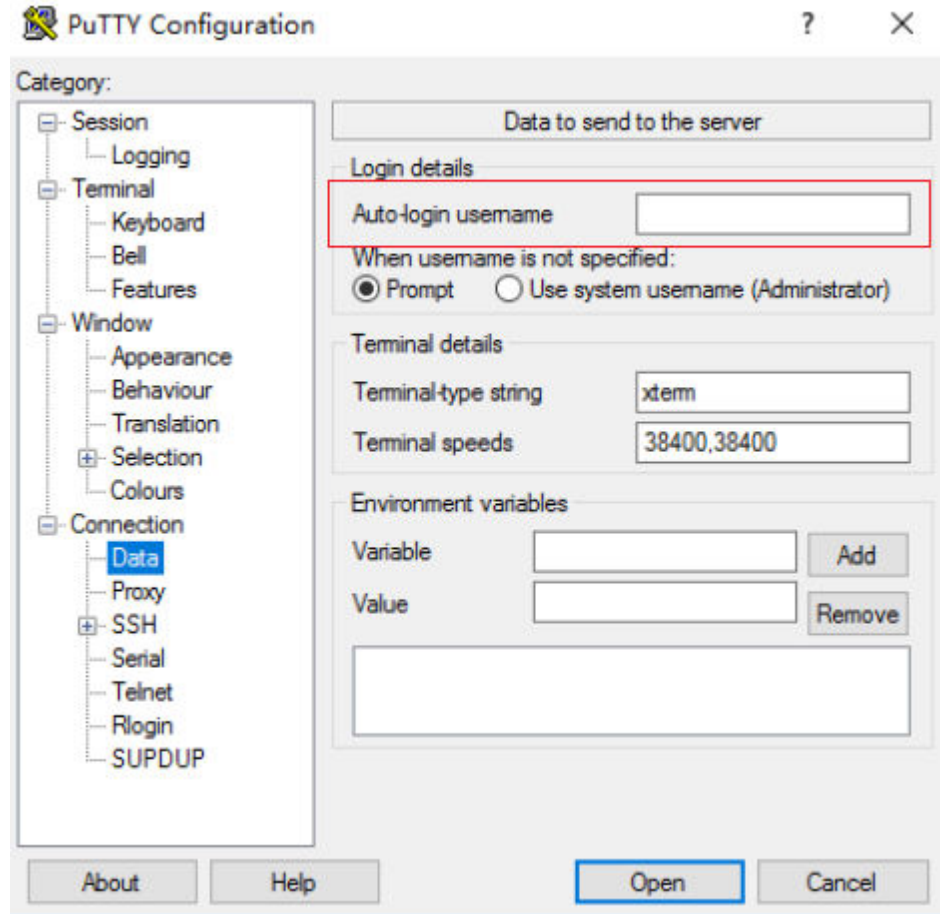
5. 在“Actions”区域，单击“Save private key”。
6. 保存转化后的私钥到本地。例如：kp-123.ppk
7. 双击“PUTTY.EXE”，打开“PuTTY Configuration”。
8. 单击“Session”，在“Host Name (or IP address)”下的输入框中输入弹性云服务器的弹性公网IP。

图 2-53 配置弹性公网 IP



9. 选择“Connection > data”，在Auto-login username处输入镜像的用户名。

图 2-54 输入用户名

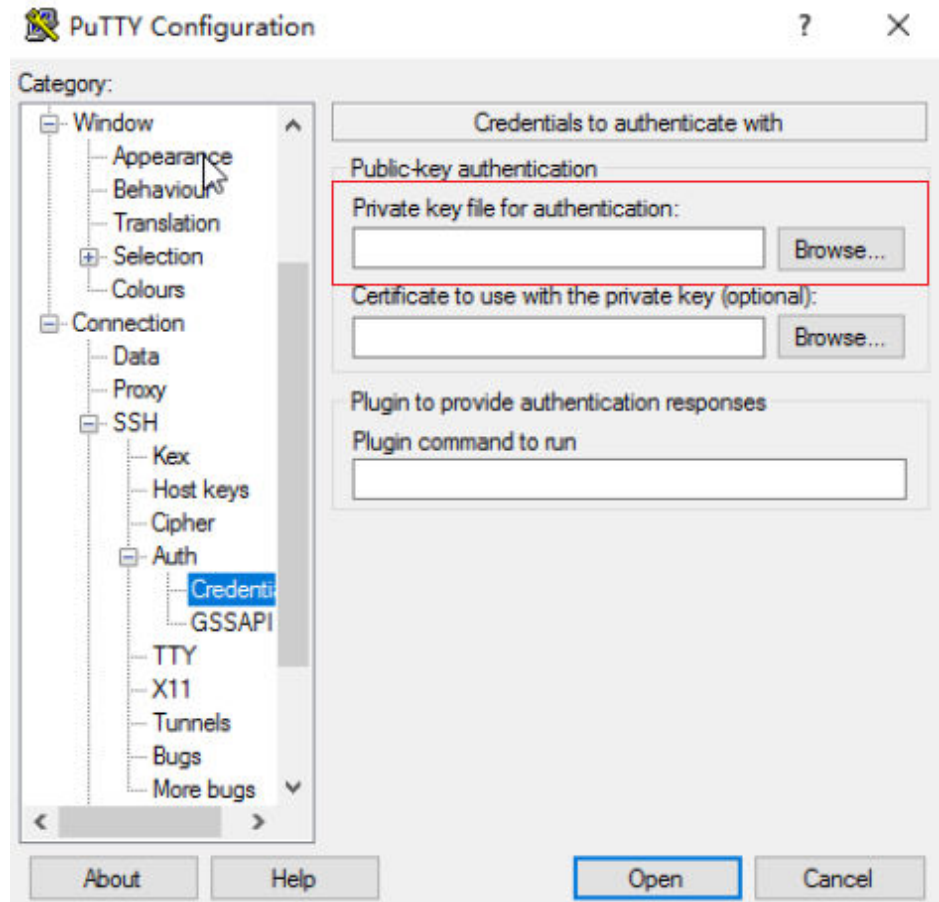


说明

使用“SSH密钥方式”登录弹性云服务器时：

- 如果是“CoreOS”的公共镜像，镜像的用户名为“core”。
 - 如果是“非CoreOS”的公共镜像，镜像的用户名为“root”。
10. 选择“Connection > SSH > Auth > Credentials”，在配置项“Private key file for authentication”中，单击“Browse”，选择6转化的密钥。

图 2-55 导入密钥



11. 单击“Open”。
登录弹性云服务器。

方式二：使用Xshell登录

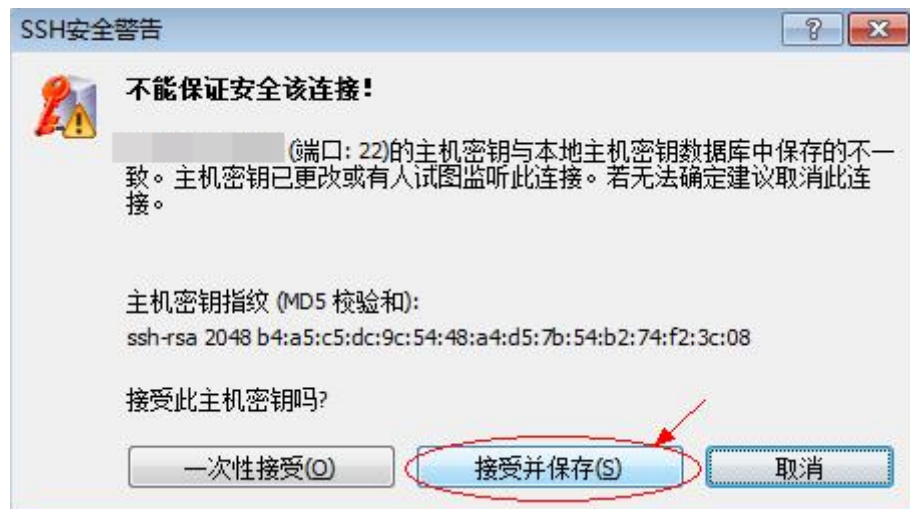
1. 打开Xshell工具。
2. 通过弹性公网IP，执行以下命令，SSH远程连接弹性云服务器。
ssh 用户名@弹性公网IP

📖 说明

使用“SSH密钥方式”登录弹性云服务器时：

- 如果是“CoreOS”的公共镜像，镜像的用户名为“core”。
 - 如果是“非CoreOS”的公共镜像，镜像的用户名为“root”。
3. （可选）如果系统弹窗提示“SSH安全警告”，此时需单击“接受并保存”。

图 2-56 SSH 安全警告



4. 选择“Public Key”，并单击“用户密钥(K)”栏的“浏览”。
5. 在“用户密钥”窗口中，单击“导入”。
6. 选择本地保存的密钥文件，并单击“打开”。
7. 单击“确定”，登录弹性云服务器。

本地使用 Linux 操作系统

如果您本地使用Linux操作系统登录Linux弹性云服务器，可以按照下面方式登录。下面步骤以私钥文件是kp-123.pem为例进行介绍。

1. 在您的linux计算机的命令中执行如下命令，变更权限。

```
chmod 400 /path/kp-123.pem
```

📖 说明

上述命令的path为密钥文件的存放路径。

2. 执行如下命令，登录弹性云服务器。

```
ssh -i /path/kp-123.pem 默认用户名@弹性公网IP
```

假设Linux弹性云服务器的默认用户名是root，弹性公网IP为123.123.123.123，则命令如下：

```
ssh -i /path/kp-123.pem root@123.123.123.123
```

📖 说明

- path为密钥文件的存放路径。
- 弹性公网IP地址为弹性云服务器绑定的弹性公网IP地址。

后续处理

- 以SSH密钥方式登录弹性云服务器后，可以通过设置密码（执行passwd命令），后续使用VNC方式登录Linux弹性云服务器。

相关链接

- [忘记密码怎么办？](#)

- [无法登录到Linux云服务器怎么办？](#)

2.5.5 通过 SSH 密码方式登录 Linux ECS

操作场景

本节操作介绍在Windows和Linux环境中使用SSH密码方式远程登录Linux云服务器的操作步骤。

前提条件

- 弹性云服务器状态为“运行中”。
- 弹性云服务器已经绑定弹性公网IP，绑定方式请参见[绑定弹性公网IP](#)。
- 所在安全组入方向已开放22端口，配置方式请参见[配置安全组规则](#)。
- 使用的登录工具（如PuTTY）与待登录的弹性云服务器之间网络连通。例如，默认的22端口没有被防火墙屏蔽。

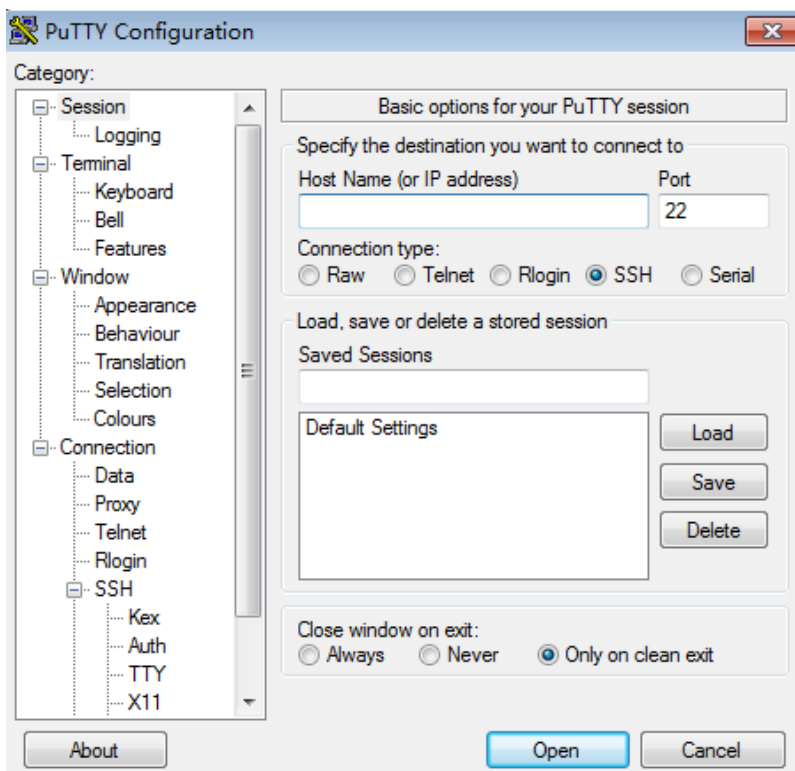
本地使用 Windows 操作系统

如果本地主机为Windows操作系统，可以按照下面方式登录云服务器。

下面步骤以PuTTY为例。

1. 在以下路径中下载PuTTY和PuTTYgen。
<https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>
2. 运行PuTTY。
3. 单击“Session”。
 - a. Host Name (or IP address): 输入云服务器的弹性公网IP。
 - b. Port: 输入 22。
 - c. Connection Type: 选择 SSH。
 - d. Saved Sessions: 任务名称，在下次使用putty时就可以单击保存的任务名称，即可打开远程连接。

图 2-57 单击“Session”



4. 单击“Window”，在“Translation”下的“Received data assumed to be in which character set:”选择“UTF-8”。
5. 单击“Open”。
如果首次登录服务器，PuTTY会显示安全警告对话框，询问是否接受服务器的安全证书。单击“是”将证书保存到本地注册表中。
6. 建立到云服务器的SSH连接后，根据提示输入用户名和密码登录云服务器。

说明

如果是公共镜像（包括CoreOS），首次登录时，登录用户名、密码如下：

- 用户名：root
- 密码：购买云服务器时，您设置的密码

若购买时云服务器未设置密码，请参考[在控制台重置ECS密码](#)进行设置。

本地使用 Linux 操作系统

如果本地主机为Linux操作系统，您可以在计算机的命令行中通过以下操作登录弹性云服务器。

1. 在您的linux计算机的命令行中执行如下命令，登录弹性云服务器。

```
ssh xx.xx.xx.xx
```

说明

xx.xx.xx.xx表示弹性云服务器绑定的弹性公网IP。

2. 根据界面提示信息，验证远程主机SSH指纹，并输入“yes”。

```
The authenticity of host 'xx.xx.xx.xx (xx.xx.xx.xx)' can't be established.  
ECDSA key fingerprint is SHA256:rnKuzrUSYS03MCoaXXXXXXXXXXXXXXXXXXXXXXXXXXXX.  
ECDSA key fingerprint is MD5:cf:64:5b:5e:74:30:xx:xx:xx:xx:xx:xx:xx:xx:xx:xx.
```

```
Are you sure you want to continue connecting (yes/no)? yes  
Warning: Permanently added 'xx.xx.xx.xx' (ECDSA) to the list of known hosts.
```

3. 根据界面提示，输入待登录的弹性云服务器的密码，完成登录操作。
root@xx.xx.xx.xx's password:

```
Welcome to Huawei Cloud Service
```

相关链接

- [忘记密码怎么办？](#)
- [无法登录到Linux云服务器怎么办？](#)

2.5.6 通过 macOS 主机登录 Linux ECS

操作场景

本节为您介绍如何在macOS系统主机上登录Linux云服务器。

前提条件

- 云服务器状态为“运行中”。
- 若您选择SSH密码方式登录，需在购买ECS时设置“登录凭证”为“密码”并已取得Linux云服务器用户名和密码。
忘记密码请参考[在控制台重置ECS密码](#)重置密码。
- 若您选择SSH密钥方式登录，需在购买ECS时设置“登录凭证”为“密钥对”并已取得Linux云服务器的私钥文件。
Linux云服务器的私钥文件在创建密钥对时生成，若私钥文件不慎遗失，可以通过[重置密钥对](#)的方式，重新为弹性云服务器绑定密钥对。若已授权托管私钥，可根据需要将托管的私钥导出使用，详细内容请参见[导出私钥](#)。
- 弹性云服务器已经绑定弹性公网IP，绑定方式请参见[绑定弹性公网IP](#)。
- 所在安全组入方向已开放22端口，配置方式请参见[配置安全组规则](#)。

操作步骤

您可以通过macOS系统自带的终端（Terminal）登录Linux云服务器。

- SSH密码方式
 - a. 打开系统自带的终端（Terminal），执行以下命令，登录云服务器。

```
ssh 用户名@弹性公网IP
```

说明

如果是公共镜像（包括CoreOS），用户名为“root”。

- SSH密钥方式
 - a. 打开系统自带的终端（Terminal），执行以下命令，变更权限。下面步骤以私钥文件是kp-123.pem为例进行介绍。

```
chmod 400 /path/kp-123.pem
```

📖 说明

Linux云服务器的私钥文件在创建密钥对时生成，若私钥文件不慎遗失，可以通过[重置密钥对](#)的方式，重新为弹性云服务器绑定密钥对。若已授权托管私钥，可根据需要将托管的私钥导出使用，详细内容请参见[导出私钥](#)。

上述命令的path为密钥文件的存放路径。

- b. 执行以下命令，登录云服务器。

```
ssh -i /path/kp-123.pem 用户名@弹性公网IP
```

📖 说明

- 如果是“CoreOS”的公共镜像，用户名为“core”。
- 如果是“非CoreOS”的公共镜像，用户名为“root”。

后续处理

- 以SSH密钥方式登录弹性云服务器后，可以通过设置密码（执行passwd命令），后续使用VNC方式登录Linux弹性云服务器。

2.5.7 通过移动设备登录 Linux ECS

操作场景

本节操作介绍如何在移动设备上连接Linux实例。

- 以Termius为例介绍如何在iOS设备上连接 Linux 实例，详细操作请参考[IOS设备上登录Linux云服务器](#)。
- 以JuiceSSH为例介绍如何在Android设备上连接 Linux 实例，详细操作请参考[Android设备上登录Linux云服务器](#)。

前提条件

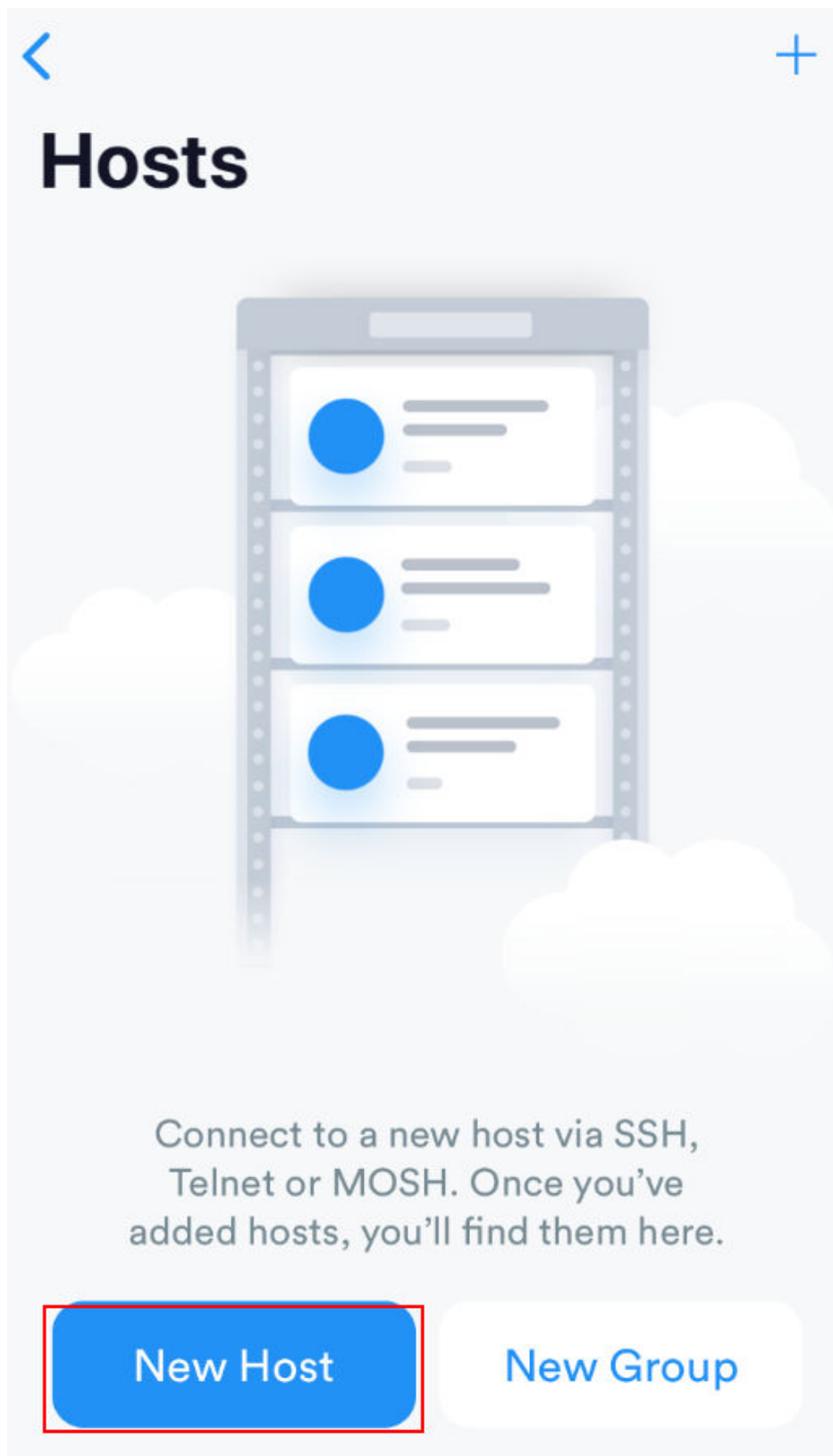
- 云服务器状态为“运行中”。
- 已获取Linux云服务器用户名和密码。忘记密码请参考[在控制台重置ECS密码重置密码](#)。
- 弹性云服务器已经绑定弹性公网IP，绑定方式请参见[绑定弹性公网IP](#)。
- 所在安全组入方向已开放22端口，配置方式请参见[配置安全组规则](#)。

IOS 设备上登录 Linux 云服务器

如果您使用iOS设备，请确保已经安装了SSH客户端工具，以Termius为例。本示例中使用CentOS 7.6操作系统，使用用户名和密码进行认证。

1. 启动Termius，单击New Host。

图 2-58 New Host



2. 在SSH页面上，输入连接信息后，单击 Save。需要输入的连接信息包括：
 - Alias: 指定Host名称，如本例中，设置为ecs01。

- Hostname: 输入需要连接的 Linux 实例的公网IP地址。
- Use SSH: 打开SSH登录配置。
- Host: 输入需要连接的Linux实例的公网IP地址。
- Port: 输入端口号22。
- 用户名: 输入用户名root。
- 密码: 输入实例登录密码。

图 2-59 输入连接信息

Cancel New Host Save

1 Alias

2 Hostname

Group >

Tags >

Backspace as CTRL+H

SSH / MOSH

3 Use SSH

Use Mosh (Beta)

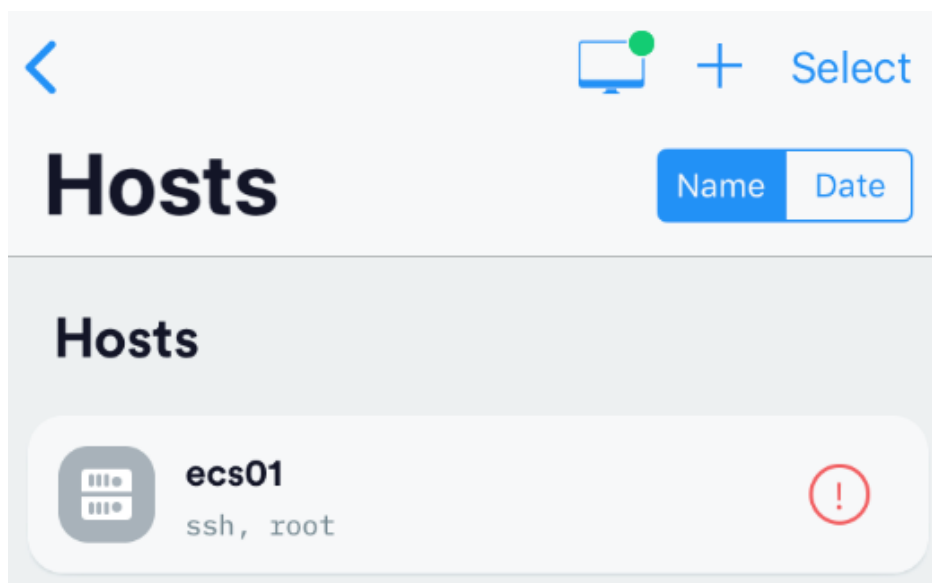
4 Port 22
Default

5 Username root

6 Password ●●●●●●●●

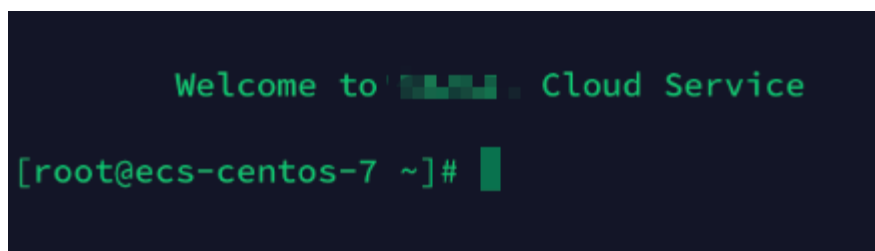
3. 单击右上角的“Save”，保存登录信息，在Hosts页面，单击连接的名称远程连接服务器。

图 2-60 保存登录信息



当出现如图2-61所示页面时，您已经成功地连接了Linux云服务器。

图 2-61 已连接



Android 设备上登录 Linux 云服务器

如果您使用Android设备，请确保已经安装了JuiceSSH。本示例中使用CentOS 7.6操作系统，通过用户名和密码进行认证。

1. 启动JuiceSSH，单击“连接”。

图 2-62 启动 JuiceSSH




2. 在“连接”页面单击  图标。


图 2-63 连接



无连接

你还没有配置任何连接。点击下面的按钮开始。



3. 在“新建连接”页面上，添加基本设置和高级设置的信息并保存。需要添加的信息如下：
 - 昵称：指定登录会话的名称，如本例中，设置为“linux_test”。
 - 类型：使用默认值“SSH”。
 - 地址：输入需要登录的Linux实例的弹性公网IP地址。
 - 按以下步骤设置认证：
 - i. 单击“认证”，在下拉列表里单击“新建”。
 - ii. 在“新建认证”页面上，添加如下信息后，单击图标。

- 昵称：可选项，您可以根据需要设置一个身份名称，方便后续管理。如本示例中，设置为“linux_test”。
- 用户名：输入用户名“root”。
- 密码：单击“设置（可选）”后，输入实例的登录密码，单击“确定”。

图 2-64 新建认证

← 新建认证 ✓

认证信息

昵称： linux_test

用户名： root

密码： 更新 / 清除

私钥： 设置（可选）

代码片段

JuiceSSH 高级版用户可以自动创建一个代码片段，该代码可用于添加公钥到服务器上的 ~/.ssh/authorized_keys 文件并设置正确的权限。

生成代码片段

- 端口：输入端口号“22”。

图 2-65 端口号

← 新建连接 ✓

基本设置

昵称： linux_test

类型： SSH

地址： [blurred]

认证： linux_test

高级设置

端口： 22

连接方式： (可选)

运行代码片段： (可选)

Backspace 模式： 默认发送 (DEL)

连接组

添加到组

4. 在“连接”页面单击创建好的连接。

图 2-66 单击连接



5. 确认提示信息后，单击“接受”。

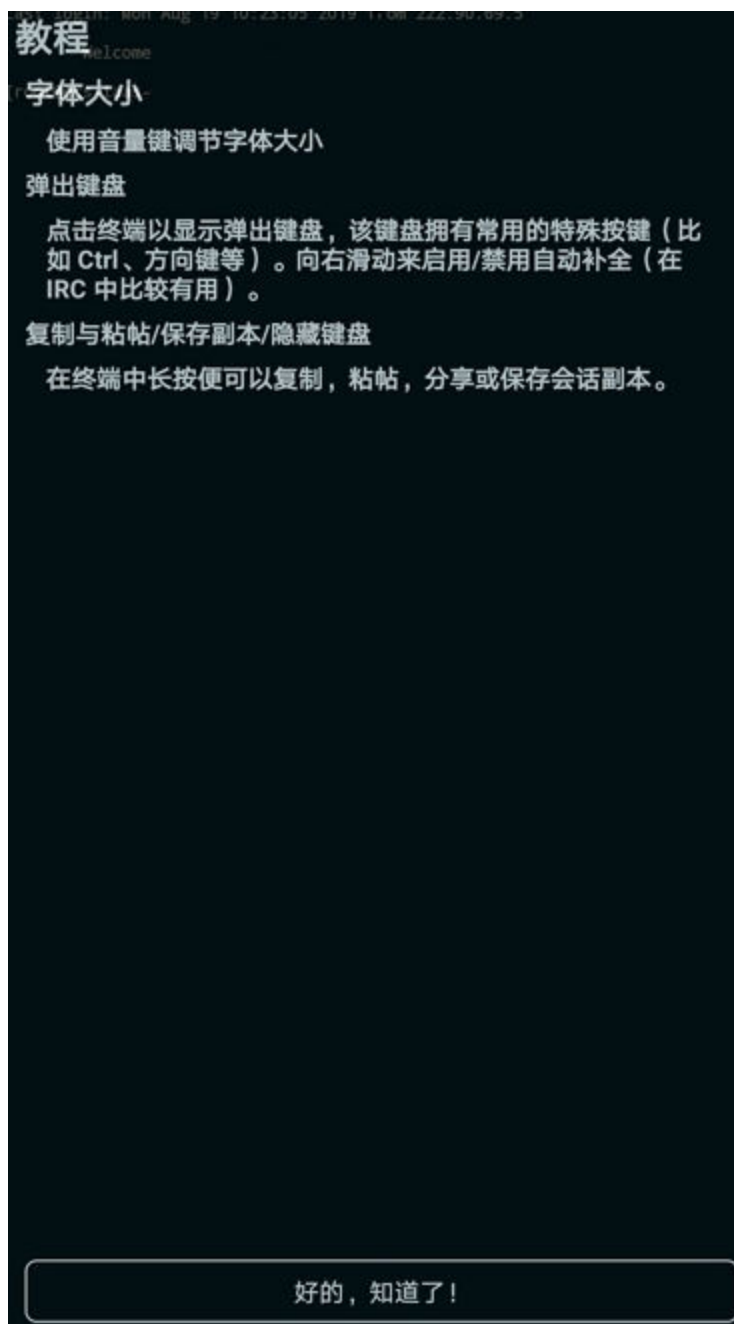


图 2-67 确认提示信息



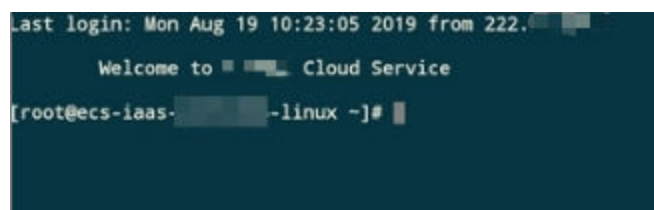
6. （可选）第一次连接时，JuiceSSH会提示您如何设置字体大小，如何弹出键盘等。确认信息后，单击“好的，我知道了！”。

图 2-68 教程



至此，您已经成功登录Linux实例。

图 2-69 登录 Linux 实例



2.6 管理 GPU 加速型 ECS 的 GPU 驱动

2.6.1 GPU 驱动概述

GPU 驱动概述

在使用GPU加速型实例前，请确保实例已安装GPU驱动以获得相应的GPU加速能力。

GPU加速型实例支持两种类型的驱动：GRID驱动和Tesla驱动。

- 如果需要使用OpenGL/DirectX/Vulkan等图形加速能力，则需要安装GRID驱动并自行购买和配置使用GRID License。此外，GRID驱动配合vDWS类型License，也支持CUDA，用来满足既需要计算加速也需要图形加速的场景。
 - 使用公共镜像创建的图形加速型（G系列）实例默认已安装特定版本的GRID驱动，但GRID License需自行购买和配置使用，请提前确认GPU加速型实例是否已经预装或者预装版本是否符合需求。
 - 使用私有镜像创建的GPU加速型实例，如需安装GRID驱动请参考[手动安装GPU加速型ECS的GRID驱动](#)。
- 如果需要通过实现计算加速能力，则需要安装Tesla驱动。
 - 使用公共镜像创建的计算加速型（P系列）实例默认已安装特定版本的Tesla驱动。
 - 使用私有镜像创建的GPU加速型实例，如需安装Tesla驱动请参考[手动安装GPU加速型ECS的Tesla驱动](#)。

表 2-36 GPU 驱动支持的加速能力

驱动类型	License	CUDA	OpenGL	DirectX	Vulkan	典型应用场景	说明
GRID驱动	需要	支持	支持	支持	支持	3D渲染、图形工作站、游戏加速	付费使用，需要购买License，满足图形图像类应用加速用途。
Tesla驱动	不需要	支持	不支持	不支持	不支持	科学计算、深度学习训练和推理	通常搭配使用NVIDIA CUDA SDK，可免费下载使用，满足通用计算类应用加速用途。

2.6.2 Tesla 驱动及 CUDA 工具包获取方式

操作场景

使用GPU加速型云服务器时，需确保已安装Tesla驱动和CUDA工具包，否则无法实现计算加速功能。本节内容提供Tesla驱动及CUDA工具包下载地址，请根据实例的类型，选择具体的驱动版本。

Tesla驱动及CUDA工具包安装操作指导请参考[手动安装GPU加速型ECS的Tesla驱动](#)。

Tesla 驱动下载地址

请单击[NVIDIA驱动下载](#)，根据实例的类型，选择NVIDIA产品类型、产品系列和产品。

表 2-37 Tesla 驱动产品类型对应关系

实例类型	产品类型 (Product Type)	产品系列 (Product Series)	产品 (Product)
P2s	Tesla	V-Series	V100
P2v	Tesla	V-Series	V100
Pi2	Tesla	T- Series	T4
Pi1	Tesla	P-Series	P4

CUDA 工具包下载地址

请从[CUDA软件包下载](#)获取CUDA软件包，您需要根据实例类型和驱动版本，选择对应的CUDA Toolkit软件包产品。

📖 说明

驱动版本与CUDA Toolkit版本存在对应关系，如二者版本不匹配，可能导致驱动无法使用。版本对应关系，请参见[NVIDIA驱动下载](#)。

下面以Tesla T4下载驱动软件包及CUDA Toolkit为例进行介绍。

1. Tesla T4安装驱动软件包时，选择Linux操作系统，并指定CUDA Toolkit软件版本为11.6。

图 2-70 指定 CUDA Toolkit 软件版本

Manual Driver Search

Search by product, product type or series

Data Center / Tesla

T-Series

Tesla T4

Linux 64-bit

11.6

English (US)

Find

2. 选择对应的版本，下载CUDA软件包。

2.6.3 手动安装 GPU 加速型 ECS 的 GRID 驱动

操作场景

GPU加速型实例如需使用OpenGL/DirectX/Vulkan等图形加速能力则需要安装GRID驱动并自行购买和配置使用GRID License。此外，GRID驱动配合vDWS类型License，也支持CUDA，用来满足既需要计算加速也需要图形加速的场景。

- 使用公共镜像创建的图形加速型（G系列）实例默认已安装特定版本的GRID驱动，但GRID License需自行购买和配置使用。
- 使用私有镜像创建的GPU加速型实例，则需要安装GRID驱动并自行购买和配置使用GRID License。

本节操作介绍如何安装GRID驱动，购买或者申请GRID License，以及如何配置License服务器。

安装GRID驱动操作步骤：

1. [购买GRID License](#)
2. [下载GRID驱动及License软件包](#)

- 3. [部署和配置License Server](#)
- 4. [安装GRID驱动并配置License](#)

说明

- NVIDIA支持用户申请90天试用版License。
- 不同规格的GPU实例介绍和应用场景请参见[GPU加速型](#)。

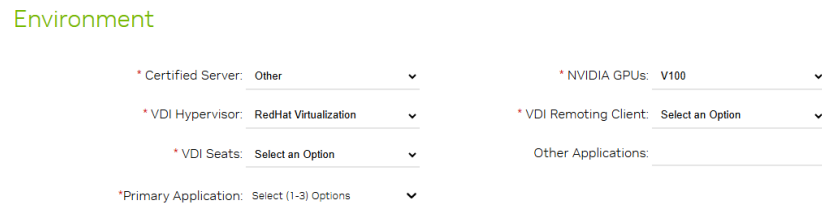
购买 GRID License

- 购买License
如果需要正式版本License，请联系NVIDIA或者所在国家/地区的NVIDIA代理商。
- 申请试用版License。
打开[NVIDIA官方网站](#)，填写相关信息。
注册账号和申请试用版License的注意事项请参见[NVIDIA官方帮助页](#)。

说明

试用版License的使用方法和正式版本的License一致，可以保留试用版账号激活正式版本的License，无需重新注册。试用版License有限期限为90天，账号过期将无法使用，请尽快购买正式版本。

图 2-71 申请试用版 License



下载 GRID 驱动及 License 软件包

- 1. 请根据[表2-38](#)对应操作系统下载驱动安装包。
了解更多GRID驱动信息请参考[NVIDIA vGPU驱动](#)。

说明

- GPU直通型实例：根据需求选择GRID驱动版本。
- GPU虚拟化型实例：请严格按照下表选择合适的驱动版本下载使用。

表 2-38 GPU 实例类型支持的 GRID 驱动版本

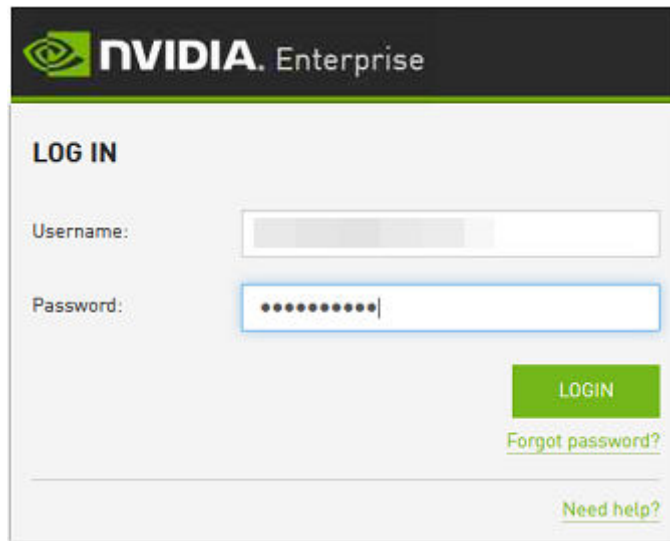
实例类型	GPU挂载方式	操作系统	驱动版本	CPU架构
G5.8xlarge.4	GPU直通型实例	<ul style="list-style-type: none">• CentOS 7.6 64bit• CentOS 7.5 64bit• Ubuntu 20.04 64bit• Ubuntu 18.04 64bit	按需选择版本	x86_64

实例类型	GPU挂载方式	操作系统	驱动版本	CPU架构
P2s	GPU直通型实例	<ul style="list-style-type: none">• Huawei Cloud EulerOS 2.0 64bit• CentOS 8.2 64bit• CentOS 7.9 64bit• CentOS 7.8 64bit• CentOS 7.7 64bit• CentOS 7.6 64bit• CentOS 7.5 64bit• Ubuntu 22.04 Server 64bit• Ubuntu 20.04 Server 64bit• Ubuntu 18.04 Server 64bit• Ubuntu 16.04 Server 64bit	按需选择版本	x86_64
P2v	GPU直通型实例	<ul style="list-style-type: none">• CentOS 7.4 64bit• EulerOS 2.2 64bit• Ubuntu 20.04 Server 64bit• Ubuntu 18.04 Server 64bit• Ubuntu 16.04 Server 64bit	按需选择版本	x86_64

实例类型	GPU挂载方式	操作系统	驱动版本	CPU架构
Pi2	GPU直通型实例	<ul style="list-style-type: none">• Huawei Cloud EulerOS 2.0 64bit• CentOS 8.2 64bit• CentOS 8.1 64bit• CentOS 8.0 64bit• CentOS 7.9 64bit• CentOS 7.8 64bit• CentOS 7.7 64bit• CentOS 7.6 64bit• CentOS 7.5 64bit• Ubuntu 22.04 Server 64bit• Ubuntu 20.04 Server 64bit• Ubuntu 18.04 Server 64bit• Ubuntu 16.04 Server 64bit	按需选择版本	x86_64
Pi1	GPU直通型实例	<ul style="list-style-type: none">• CentOS 7.3 64bit• Ubuntu 20.04 Server 64bit• Ubuntu 16.04 Server 64bit• Ubuntu 14.04 Server 64bit	按需选择版本	x86_64

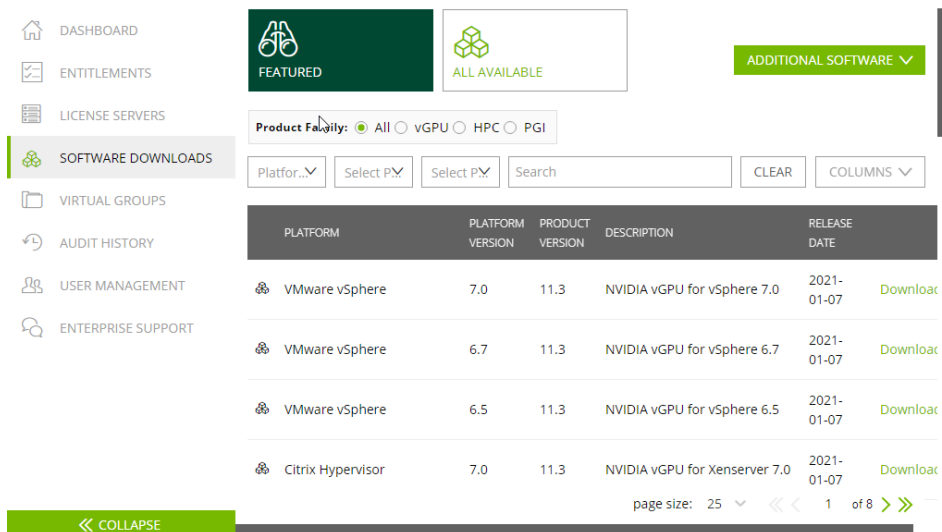
- 注册成功后，登录[NVIDIA官网](#)，填写账户信息。
- 判断是否为首次注册使用NVIDIA。
 - 是，执行4。
 - 否，执行6。
- 在注册NVIDIA用户成功的邮件中，查询PAK，如[图2-72](#)所示。

图 2-74 登录 NVIDIA 官网



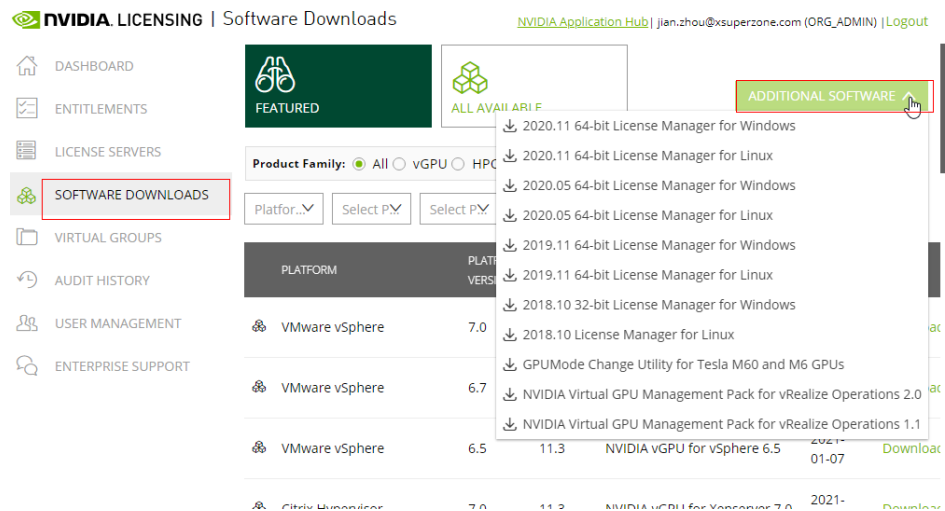
7. 根据界面提示，进入NVIDIA网站，并选择“SOFTWARE DOWNLOADS”。

图 2-75 打开 NVIDIA 网站



8. 对照表2-38选择相应版本的GRID驱动下载。
9. 解压缩GRID驱动包，并选择和弹性云服务器操作系统匹配的驱动进行安装。
10. 在“SOFTWARE DOWNLOADS”页面，单击“ADDITIONAL SOFTWARE”，下载License软件包。

图 2-76 选择 SOFTWARE DOWNLOADS



部署和配置 License Server

以CentOS 7.5操作系统的云服务器为例演示部署和配置License Server。

说明

- 云服务器规格不小于2vCPU，内存不小于4GiB。
- 请提前记录云服务器MAC地址。
- 如用作生产用途，建议采用高可用模式部署，主备高可用模式部署License Server 请参考 [NVIDIA官方License Server高可用部署文档](#)。

1. 配置网络：

- 如使用VPC网络访问License Server：请确保License Server和使用GRID驱动的GPU加速型实例处在同一个VPC子网内。
- 如使用公网IP访问License Server：请配置License Server所在的安全组，增加方向规则：TCP 7070和TCP 8080。

2. 安装License Server。

- 执行以下命令，解压缩安装包。其中“安装程序.zip”为10获取到的安装包名称。

unzip 安装程序.zip

- 执行以下命令，为安装程序添加执行权限。

chmod +x setup.bin

- 以root用户运行安装程序。

sudo ./setup.bin -i console

- 在Introduction部分，单击回车键继续。

```
=====
Introduction
-----

InstallAnywhere will guide you through the installation of License Server.

It is strongly recommended that you quit all programs before continuing with
this installation.

Respond to each prompt to proceed to the next step in the installation. If
you want to change something on a previous step, type 'back'.

You may cancel this installation at any time by typing 'quit'.

PRESS <ENTER> TO CONTINUE: █
```

- e. 在License Agreement部分，通过单击回车键进行翻页，翻页结束后接受许可协议。
当您达成许可协议时，系统会提示您接受许可协议条款，请输入“Y”，并单击回车键。

```
DO YOU ACCEPT THE TERMS OF THIS LICENSE AGREEMENT? (Y/N): Y █
```

- f. 在Choose Install Folder部分，单击回车键，保持默认的License Server软件安装路径。
- g. 在Choose Local Tomcat Server Path部分，输入Tomcat的本地路径，默认为/var/lib/tomcat版本号，例如：/var/lib/tomcat8。
- h. 在Choose Firewall Options部分，确认需要在防火墙中打开的端口，单击回车键，保持默认选项即可。

```
Choose Firewall Options
-----

The license server listens on port 7070. This port must be opened in the
firewall for other machines to obtain licenses from this server.

The license server's management interface listens on port 8080. Leave this
port closed to prevent unauthorized access to the management interface.

->1- License server (port 7070)
   2- Management interface (port 8080)

ENTER A COMMA-SEPARATED LIST OF NUMBERS REPRESENTING THE DESIRED CHOICES, OR
PRESS <ENTER> TO ACCEPT THE DEFAULT: █
```

- i. 在Pre-Installation Summary部分，确认信息并单击回车键启动安装。

```
Pre-Installation Summary
-----

Please Review the Following Before Continuing:

Product Name:
  License Server

Install Folder:
  /opt/flexnetls/nvidia

Link Folder:
  /root/NVIDIA Corporation/License Server

Disk Space Information (for Installation Target):
  Required:    105,216,774 Bytes
  Available:  35,501,248,512 Bytes

PRESS <ENTER> TO CONTINUE: █
```

- j. 在Install Complete部分，单击回车键，结束安装。

```
Install Complete
-----

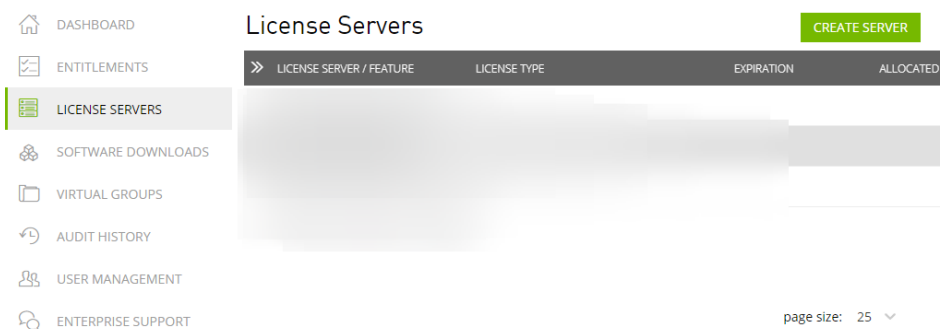
License Server has been successfully installed to:

  /opt/flexnetls/nvidia

PRESS <ENTER> TO EXIT THE INSTALLER:
```

- 3. 获取License文件。
 - a. 新建页签，登录[NVIDIA网站](#)，选择“LICENSE SERVERS”。

图 2-77 选择 LICENSE SERVERS



- b. 单击“CREATE SERVER”。
 - c. 在“Create License Server”界面，根据界面提示配置参数。

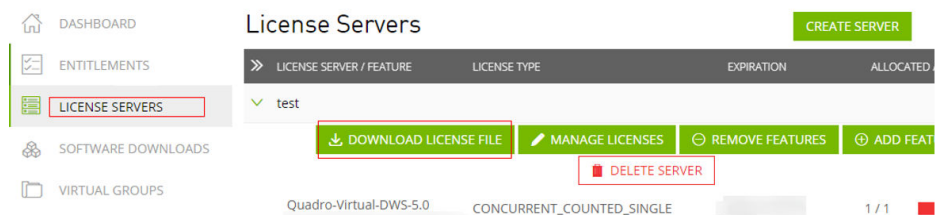
图 2-78 填写 LICENSE SERVERS 信息

表 2-39 LICENSE SERVERS 信息填写说明

参数	说明
Server Name	自定义需要的License Server名称。
Description	License Server的描述信息。
MAC Address	填写用于搭建License Server的ECS实例的MAC地址。 您可以登录实例，使用ipconfig -a命令进行查询。
Feature	在Licenses框中输入需要的license数目，单击“ADD”。 如果是主备部署的情况需要把备服务器的名称填入 Failover License Server，MAC地址填入 Failover MAC Address中。

- d. 单击“CREATE LICENSE SERVER”。
- e. 下载License文件。

图 2-79 下载 License 文件

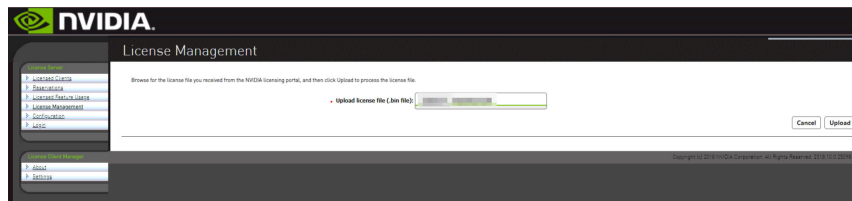


4. 在Web浏览器中，根据安装时配置的管理页面链接，访问License Server管理界面的主页。

默认访问链接为：<http://弹性公网IP地址:8080/licserver>。

5. 在左侧导航树中，单击“License Server > License Management”。
6. 使用License服务器配置菜单导入，并单击“Upload”上传许可证*.bin文件，完成License Server的配置。

图 2-80 上传许可证文件



安装 GRID 驱动并配置 License

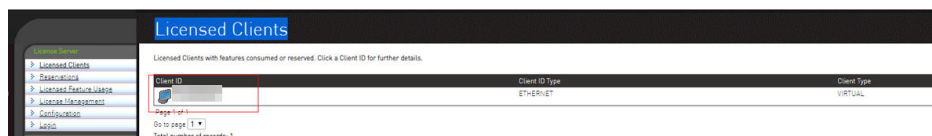
1. 以Windows操作系统GPU加速型实例为例，选择合适版本的GRID驱动进行安装。

说明

微软的远程登录协议不支持使用GPU的3D硬件加速能力，如需使用请安装VNC/PCoIP/NICE DCV等第三方桌面协议软件，并通过相应客户端连接GPU实例，使用GPU图形图像加速能力。

2. 使用第三方桌面协议连接后，在Windows控制面板中打开NVIDIA控制面板。
3. 在一级许可证服务器中填入部署的License server的IP和端口，并单击应用。当出现“您的系统已获GRID vGPU许可”则代表安装GRID驱动成功，并且可以在License Server管理控制台Licensed Clients中看到已安装GRID驱动并使用了License的GPU实例的MAC地址。

图 2-81 License Server 管理控制台



2.6.4 手动安装 GPU 加速型 ECS 的 Tesla 驱动

操作场景

GPU加速型云服务器，需要安装Tesla驱动和CUDA工具包以实现计算加速功能。

- 使用公共镜像创建的加速型（P系列）实例默认已安装特定版本的Tesla驱动。
- 使用私有镜像创建的GPU加速型云服务器，需在创建完成后安装Tesla驱动，否则无法实现计算加速功能。

本节操作介绍GPU云服务器安装Tesla驱动及CUDA工具包的操作步骤。

安装须知

- 云服务器已绑定弹性公网IP。
- 云服务器未安装Tesla驱动以及CUDA工具包。

📖 说明

- 从NVIDIA官网下载CUDA工具包进行安装，通常会同时自动安装一个和CUDA版本匹配的Tesla驱动，免去用户单独安装Tesla驱动的步骤。但是对NVIDIA驱动版本有强制要求或依赖时，请务必先单独从NVIDIA官网下载并安装匹配的Tesla驱动，然后再安装CUDA工具包。
- 如果云服务器已经安装了Tesla驱动，请检查当前驱动版本是否可用。如需安装新版本的驱动请卸载旧版本的Tesla驱动，避免因驱动程序冲突导致安装失败。

安装指引：

- [Tesla驱动及CUDA工具包获取方式](#)
- 安装NVIDIA驱动
 - [Linux操作系统云服务器安装Tesla驱动](#)
 - [Windows操作系统云服务器安装Tesla驱动](#)
- 安装CUDA驱动
 - [Linux操作系统安装CUDA工具包](#)
 - [Windows操作系统安装CUDA工具包](#)

Linux 操作系统云服务器安装 Tesla 驱动

以下操作以Ubuntu 20.04 64bit操作系统，GPU实例安装CUDA 10.1对应的Tesla驱动为例。

📖 说明

Linux内核版本和驱动的版本存在兼容性关系。如果驱动安装失败，请检查驱动安装日志(安装日志通常位于/var/log/nvidia-installer.log)。日志中如提示失败原因为驱动编译错误，例如get_user_pages参数不匹配，说明当前内核版本与驱动版本不兼容，请选择合适的内核版本和驱动版本重新安装。建议内核版本与驱动版本发布时间不要相差过大。

1. 登录云服务器。
2. 根据操作系统选择命令更新系统软件。
 - Ubuntu
更新软件安装源：`apt-get -y update`
安装必要程序：`apt-get install gcc g++ make`
 - CentOS
更新软件安装源：`yum -y update --exclude=kernel* --exclude=centos-release* --exclude=initscripts*`
安装必要程序：`yum install -y kernel-devel-`uname -r` gcc gcc-c++`
3. 下载NVIDIA驱动包。
单击[NVIDIA驱动下载](#)根据实例的类型，选择驱动。

图 2-82 选择 NVIDIA 驱动

Manual Driver Search

Search by product, product type or series 🔍

Data Center / Tesla ▼ ⓘ

T-Series ▼

Tesla T4 ▼

Linux 64-bit ▼


11.6 ▼

English (US) ▼

Find

4. 根据需求选择驱动版本，以下操作以选择Tesla 418.67为例。

图 2-83 选择驱动版本

Tesla Driver for Linux x64					View
Driver Version: 418.67	CUDA Toolkit: 10.1	Release Date: Tue May 07, 2019	File Size: 107.23 MB	Info: 	

5. 单击需要下载的驱动右侧的“View”，进入对应下载界面。
6. 右键单击“Download”，复制下载的连接地址。
7. 在云服务器内部执行如下命令进行下载。

wget 复制的连接地址

例如：**wget http://us.download.nvidia.com/tesla/418.67/NVIDIA-Linux-x86_64-418.67.run**

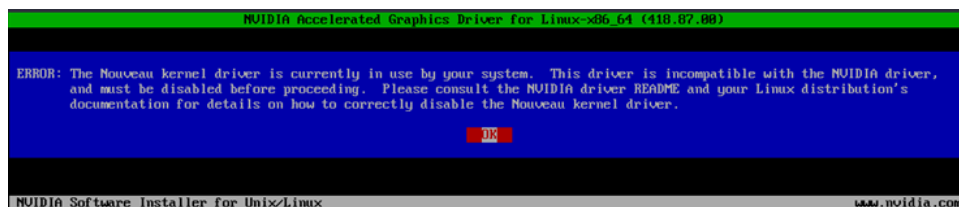
图 2-84 获取安装包

```
root@ecs-474b:~# wget http://us.download.nvidia.com/tesla/418.67/NVIDIA-Linux-x86_64-418.67.run
--2020-03-26 17:59:31-- http://us.download.nvidia.com/tesla/418.67/NVIDIA-Linux-x86_64-418.67.run
Resolving us.download.nvidia.com (us.download.nvidia.com)... 129.227.66.140, 129.227.66.139
Connecting to us.download.nvidia.com (us.download.nvidia.com)|129.227.66.140:80... connected.
HTTP request sent, awaiting response... 301 Moved Permanently
Location: https://us.download.nvidia.cn/tesla/418.67/NVIDIA-Linux-x86_64-418.67.run [following]
--2020-03-26 17:59:34-- https://us.download.nvidia.cn/tesla/418.67/NVIDIA-Linux-x86_64-418.67.run
Resolving us.download.nvidia.cn (us.download.nvidia.cn)... 60.222.11.61, 60.222.11.11, 123.134.184.166, ...
Connecting to us.download.nvidia.cn (us.download.nvidia.cn)|60.222.11.61:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 107232512 (102M) [application/octet-stream]
Saving to: 'NVIDIA-Linux-x86_64-418.67.run'

NVIDIA-Linux-x86_64-418.67.run 100%[=====] 102.26M  1.07MB/s  in 5m 2s
2020-03-26 18:04:40 (346 KB/s) - 'NVIDIA-Linux-x86_64-418.67.run' saved [107232512/107232512]
```

8. 执行以下命令安装驱动。
sh NVIDIA-Linux-x86_64-418.67.run
9. (可选) 如果执行驱动安装命令后出现如下提示信息, 需要禁用nouveau驱动。

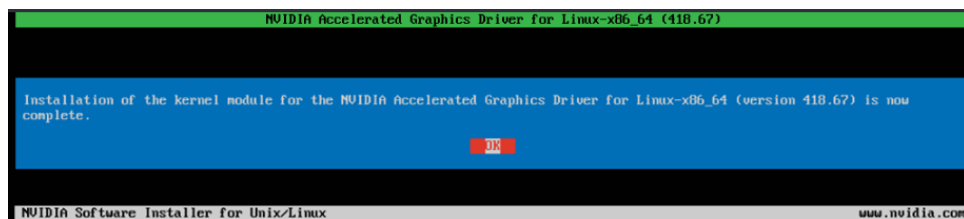
图 2-85 禁用 nouveau 驱动



- a. 执行以下命令, 查看是否安装Nouveau驱动。
lsmod | grep nouveau
 - 如果回显信息中包含Nouveau驱动信息, 说明Nouveau驱动已安装, 则需要禁用Nouveau驱动, 请执行步骤**9.b**。
 - 如果回显信息没有Nouveau驱动信息, 说明Nouveau驱动已被禁用, 请执行步骤**10**。
 - b. 执行如下命令编辑blacklist.conf文件。
如果没有“/etc/modprobe.d/blacklist.conf”文件, 请新建一个。
vi /etc/modprobe.d/blacklist.conf
添加如下语句添加至文件结尾。

```
blacklist nouveau
options nouveau modeset=0
```
 - c. 执行以下命令, 备份并新建一个initramfs。
 - Ubuntu系统:
sudo update-initramfs -u
 - CentOS系统:
mv /boot/initramfs-\$(uname -r).img /boot/initramfs-\$(uname -r).img.bak
dracut -v /boot/initramfs-\$(uname -r).img \$(uname -r)
 - d. 执行以下命令, 重启云服务器。
reboot
10. 根据安装提示, 连续三次选择“OK”。完成驱动的安装。

图 2-86 NVIDIA 驱动安装完成



11. 执行命令设置systemd。
systemctl set-default multi-user.target
12. 执行**reboot**，重启云服务器。
13. 登录云服务器，执行**nvidia-smi**，如果回显信息中包含了已安装的驱动版本，说明驱动安装成功。

图 2-87 查看 NVIDIA 驱动的版本

```
root@ecs-474b:~# nvidia-smi
Thu Mar 26 20:05:17 2020

+-----+
| NVIDIA-SMI 418.67      Driver Version: 418.67      CUDA Version: 10.1   |
+-----+
| GPU   Name           Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf    Pwr:Usage/Cap|  Memory-Usage | GPU-Util  Compute M. |
+-----+-----+
|    0   Tesla          Off          | 00000000:21:01.0 Off |                    0 |
| N/A   52C    P0      29W / 70W   |  0MiB / 15079MiB |    0%      Default   |
+-----+-----+

+-----+
| Processes:                      GPU Memory |
|  GPU       PID    Type   Process name                     Usage  |
+-----+-----+
| No running processes found      |
+-----+

root@ecs-474b:~#
```

Windows 操作系统云服务器安装 Tesla 驱动

以下操作以Windows Server 2016 Standard 64bit操作系统，GPU实例安装Tesla驱动为例。

1. 登录云服务器。
2. 下载NVIDIA驱动包。

单击[NVIDIA驱动下载](#)根据实例的类型，选择驱动版本。

图 2-88 选择驱动类型（Windows）

Manual Driver Search

Search by product, product type or series

Data Center / Tesla

T-Series

Tesla T4

Windows Server 2016

10.1

English (US)

Find

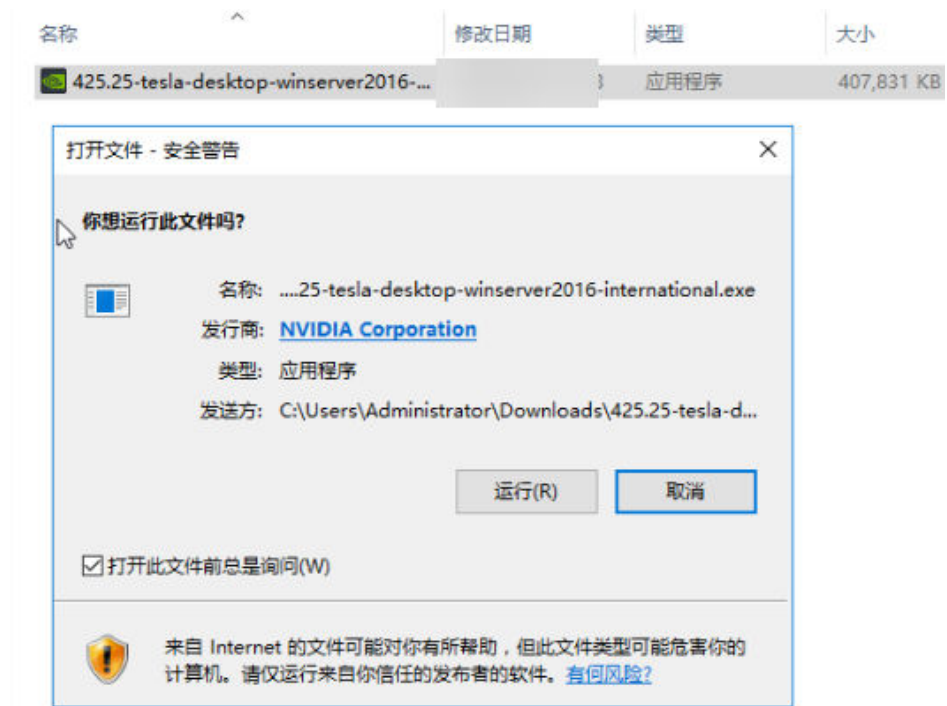
3. 根据需求选择驱动版本，本节操作以安装Tesla 425.25为例。

图 2-89 选择驱动版本（Windows）

Tesla Driver for Windows					View
Driver Version: 425.25	CUDA Toolkit: 10.1	Release Date: Tue May 07, 2019	File Size: 417.62 MB	Info: Production Branch WHQL	

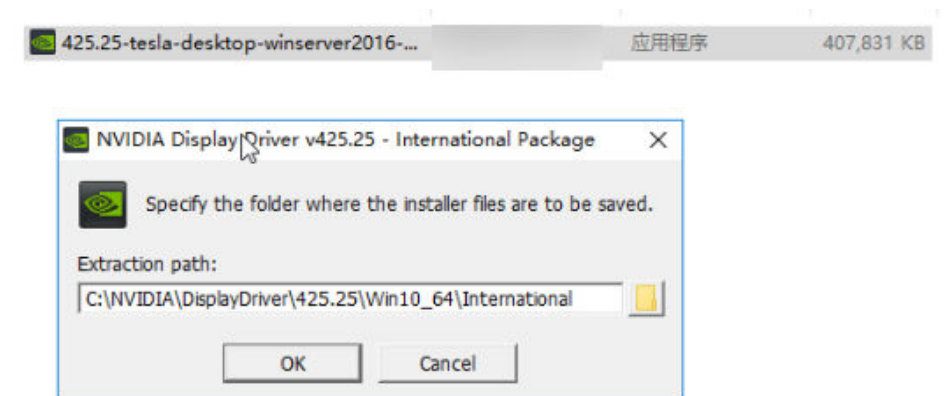
4. 单击需要下载的驱动，进入对应下载界面。
5. 单击“Download”，下载安装包。
6. 双击驱动安装驱动，单击“运行”。

图 2-90 运行 NVIDIA 驱动安装程序



7. 选择安装路径，单击“OK”。

图 2-91 选择 NVIDIA 驱动安装路径



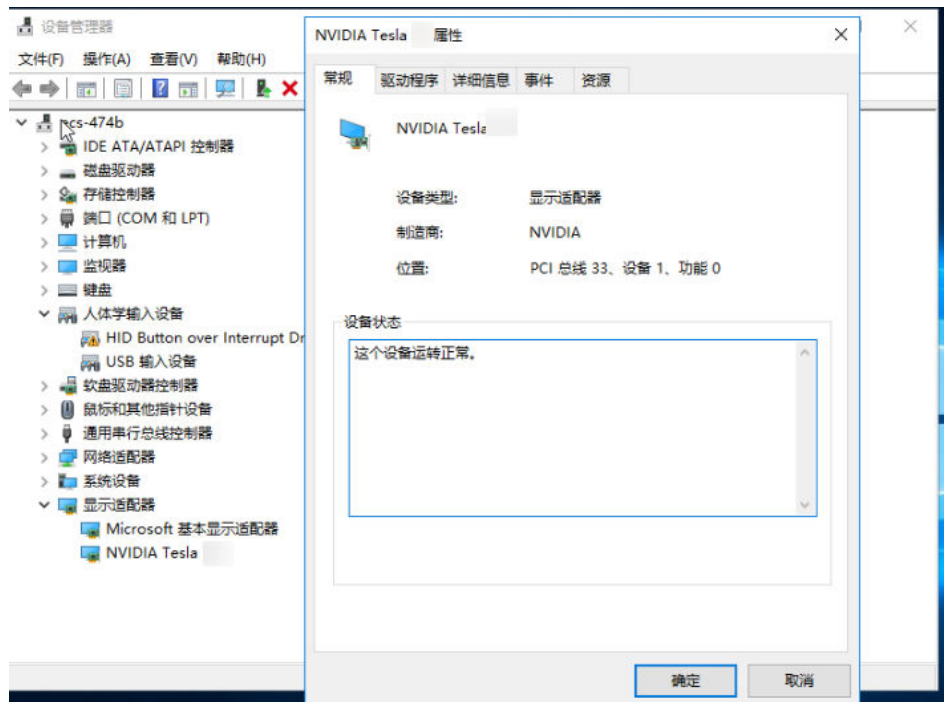
8. 根据安装提示完成NVIDIA程序的安装。

图 2-92 完成 NVIDIA 驱动的安装



9. 重启云服务器。
10. 检查驱动是否安装成功。
 - a. 进入“设备管理器”，查看“显示适配器”。

图 2-93 显示适配器



- b. 打开云服务器cmd窗口，执行以下命令：
cd C:\Program Files\NVIDIA Corporation\NVSMI

nvidia-smi

如果回显信息中包含了已安装的驱动版本，说明驱动安装成功。

图 2-94 查看 NVIDIA 驱动安装版本

```
C:\Program Files\NVIDIA Corporation\NVSMI>
C:\Program Files\NVIDIA Corporation\NVSMI>nvidia-smi
2020

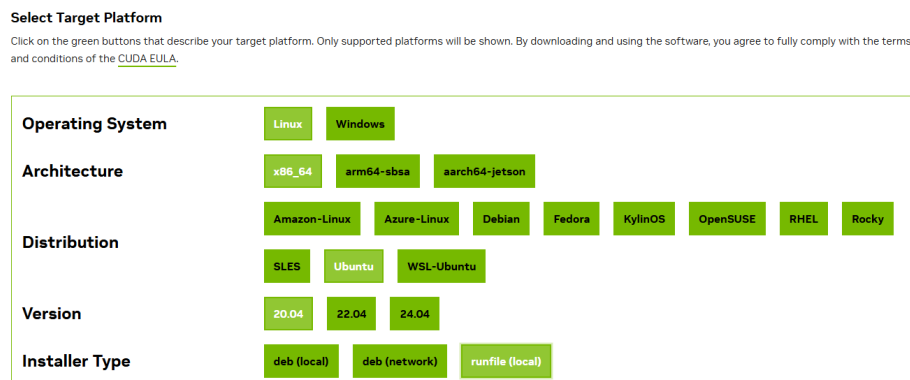
+-----+
| NVIDIA-SMI 425.25      | Driver Version: 425.25      | CUDA Version: 10.1      |
+-----+-----+-----+
| GPU  Name            | TCC/WDDM | Bus-Id      | Disp.A | Volatile Uncorr. ECC |
| Fan  Temp  Perf    | Pwr:Usage/Cap |          | Memory-Usage | GPU-Util  Compute M. |
+-----+-----+-----+-----+-----+
|   0   Tesla          | TCC      | 00000000:21:01.0 Off |         |      0               |
| N/A   27C    P8     |  9W / 70W |          0MiB / 15205MiB |         |      0%    Default  |
+-----+-----+-----+-----+-----+
| Processes:                                     | GPU Memory Usage          |
| GPU      PID  Type  Process name                               |                          |
+-----+-----+-----+-----+-----+
| No running processes found                    |                          |
+-----+-----+-----+-----+-----+
```

Linux 操作系统安装 CUDA 工具包

以下操作以Ubuntu 20.04 64bit操作系统，GPU实例使用run包安装CUDA 10.1为例。

1. 登录云服务器。
2. 根据操作系统选择命令更新系统软件。
 - Ubuntu
更新软件安装源：`apt-get -y update`
安装必要程序：`apt-get install gcc g++ make`
 - CentOS
更新软件安装源：`yum -y update --exclude=kernel* --exclude=centos-release* --exclude=initscripts*`
安装必要程序：`yum install -y kernel-devel-`uname -r` gcc gcc-c++`
3. 在CUDA下载页面中，按照[Tesla驱动及CUDA工具包获取方式](#)中的对应的索引项在页面中进行选择。

图 2-95 选择 CUDA 的版本



4. 选择完成后，页面会自动呈现出Ubuntu 20.04 64bit对应的CUDA 10.1的下载地址，复制下载地址。

5. 在云服务器内部执行如下命令进行下载。

wget 复制的链接地址

例如：**wget https://developer.nvidia.com/compute/cuda/10.1/Prod/local_installers/cuda_10.1.105_418.39_linux.run**

图 2-96 Linux 云服务器下载 CUDA

```
root@ecs-474b:~# wget https://developer.nvidia.com/compute/cuda/10.1/Prod/local_installers/cuda_10.1.105_418.39_linux.run
-- https://developer.nvidia.com/compute/cuda/10.1/Prod/local_installers/cuda_10.1.105_418.39_linux.run
Resolving developer.nvidia.com (developer.nvidia.com)... 129.227.66.139, 129.227.66.140
Connecting to developer.nvidia.com (developer.nvidia.com)|129.227.66.139|443... connected.
HTTP request sent, awaiting response... 302 Found
Location: https://developer.download.nvidia.com/compute/cuda/10.1/secure/Prod/local_installers/cuda_10.1.105_418.39_linux.run?FFJqELqU0UEH-9_YUp4zpYc3t1_j8q9J5clzHL2kX5Uo-f721Gy8tP3hm2Jd_j2QntuNupR-7nmp4Eb6Q_jd7YB4BhuW56D14U3prth0jz2Sx6Xugw950k9416Sb9he5n1SJJHhbPXUqULzUhhijXhuBUCkMBHsvvxn4X_uk_k2AY1oWbtU7D2sMaF8oU8 [following]
-- https://developer.download.nvidia.com/compute/cuda/10.1/secure/Prod/local_installers/cuda_10.1.105_418.39_linux.run?FFJqELqU0UEH-9_YUp4zpYc3t1_j8q9J5clzHL2kX5Uo-f721Gy8tP3hm2Jd_j2QntuNupR-7nmp4Eb6Q_jd7YB4BhuW56D14U3prth0jz2Sx6Xugw950k9416Sb9he5n1SJJHhbPXUqULzUhhijXhuBUCkMBHsvvxn4X_uk_k2AY1oWbtU7D2sMaF8oU8
Resolving developer.download.nvidia.com (developer.download.nvidia.com)... 129.227.66.139, 129.227.66.140
Connecting to developer.download.nvidia.com (developer.download.nvidia.com)|129.227.66.139|443... connected.
HTTP request sent, awaiting response... 301 Moved Permanently
Location: https://developer.download.nvidia.cn/compute/cuda/10.1/secure/Prod/local_installers/cuda_10.1.105_418.39_linux.run?FFJqELqU0UEH-9_YUp4zpYc3t1_j8q9J5clzHL2kX5Uo-f721Gy8tP3hm2Jd_j2QntuNupR-7nmp4Eb6Q_jd7YB4BhuW56D14U3prth0jz2Sx6Xugw950k9416Sb9he5n1SJJHhbPXUqULzUhhijXhuBUCkMBHsvvxn4X_uk_k2AY1oWbtU7D2sMaF8oU8 [following]
-- https://developer.download.nvidia.cn/compute/cuda/10.1/secure/Prod/local_installers/cuda_10.1.105_418.39_linux.run?FFJqELqU0UEH-9_YUp4zpYc3t1_j8q9J5clzHL2kX5Uo-f721Gy8tP3hm2Jd_j2QntuNupR-7nmp4Eb6Q_jd7YB4BhuW56D14U3prth0jz2Sx6Xugw950k9416Sb9he5n1SJJHhbPXUqULzUhhijXhuBUCkMBHsvvxn4X_uk_k2AY1oWbtU7D2sMaF8oU8
Resolving developer.download.nvidia.cn (developer.download.nvidia.cn)... 60.222.11.61, 60.222.11.10, 123.134.184.165, ...
Connecting to developer.download.nvidia.cn (developer.download.nvidia.cn)|60.222.11.61|443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 242314285 (2.3G) [application/octet-stream]
Saving to: 'cuda_10.1.105_418.39_linux.run'

cuda_10.1.105_418.39_linux.run 100%|=====>| 2.26G 34.2MB/s in 74s
```

6. CUDA工具包下载完成后，按照NVIDIA官网安装指引进行安装。
按照NVIDIA官网的Installation Instructions进行安装。
7. 执行以下命令安装CUDA。
sh cuda_10.1.243_418.87.00_linux.run
8. 在安装界面输入“accept”，回车进入下一步。

图 2-97 安装 CUDA_1

```
End User License Agreement
-----

Preface
-----

The Software License Agreement in Chapter 1 and the Supplement
in Chapter 2 contain license terms and conditions that govern
the use of NVIDIA software. By accepting this agreement, you
agree to comply with all the terms and conditions applicable
to the product(s) included herein.

NVIDIA Driver

Description

This package contains the operating system driver and

Do you accept the above EULA? (accept/decline/quit):
accept
```

9. 选择“Install”，回车进入安装过程。

图 2-98 安装 CUDA_2

```
CUDA Installer
- [X] Driver
  [X] 418.39
+ [X] CUDA Toolkit 10.1
  [X] CUDA Samples 10.1
  [X] CUDA Demo Suite 10.1
  [X] CUDA Documentation 10.1
  Install
  Options

Up/Down: Move | Left/Right: Expand | 'Enter': Select | 'A': Advanced options
```

图 2-99 安装 CUDA 完成

```
=====
= Summary =
=====
Driver:      Installed
Toolkit:    Installed in /usr/local/cuda-10.1/
Samples:    Installed in /root/, but missing recommended libraries

Please make sure that
- PATH includes /usr/local/cuda-10.1/bin
- LD_LIBRARY_PATH includes /usr/local/cuda-10.1/lib64, or, add /usr/local/cuda-10.1/lib64 to /etc/ld.so.conf and run ldconfig
as root

To uninstall the CUDA Toolkit, run cuda-uninstaller in /usr/local/cuda-10.1/bin
To uninstall the NVIDIA Driver, run nvidia-uninstall

Please see CUDA_Installation_Guide_Linux.pdf in /usr/local/cuda-10.1/doc/pdf for detailed information on setting up CUDA.
Logfile is /var/log/cuda-installer.log
root@ecs-474b:~#
```

10. 执行以下命令进入/usr/local/cuda-10.1/samples/1_Uilities/deviceQuery 目录下。
cd /usr/local/cuda-10.1/samples/1_Uilities/deviceQuery
11. 执行**make**命令，自动编译deviceQuery程序。
12. 执行以下命令查看CUDA是否安装成功。
./deviceQuery
如回显信息中有CUDA的版本信息证明CUDA安装成功。

图 2-100 deviceQuery 回显信息

```
root@ecs-474b:/usr/local/cuda-10.1/samples/1_Utilities/deviceQuery# ./deviceQuery
./deviceQuery Starting...

CUDA Device Query (Runtime API) version (CUDA static linking)

Detected 1 CUDA Capable device(s)

Device 0: "Tesla M40"
  CUDA Driver Version / Runtime Version      10.1 / 10.1
  CUDA Capability Major/Minor version number: 7.5
  Total amount of global memory:             15080 MBytes (15812263936 bytes)
  (40) Multiprocessors, ( 64) CUDA Cores/MP: 2560 CUDA Cores
  GPU Max Clock rate:                       1590 Mhz (1.59 GHz)
  Memory Clock rate:                        5001 Mhz
  Memory Bus Width:                         256-bit
  L2 Cache Size:                            4194304 bytes
  Maximum Texture Dimension Size (x,y,z)    1D=(131072), 2D=(131072, 65536), 3D=(16384, 16384, 16384)
  Maximum Layered 1D Texture Size, (num) layers 1D=(32768), 2048 layers
  Maximum Layered 2D Texture Size, (num) layers 2D=(32768, 32768), 2048 layers
  Total amount of constant memory:          65536 bytes
  Total amount of shared memory per block:  49152 bytes
  Total number of registers available per block: 65536
  Warp size:                                32
  Maximum number of threads per multiprocessor: 1024
  Maximum number of threads per block:     1024
  Max dimension size of a thread block (x,y,z): (1024, 1024, 64)
  Max dimension size of a grid size (x,y,z): (2147483647, 65535, 65535)
  Maximum memory pitch:                    2147483647 bytes
  Texture alignment:                       512 bytes
  Concurrent copy and kernel execution:     Yes with 3 copy engine(s)
  Run time limit on kernels:                No
  Integrated GPU sharing Host Memory:       No
  Support host page-locked memory mapping:  Yes
  Alignment requirement for Surfaces:       Yes
  Device has ECC support:                   Enabled
  Device supports Unified Addressing (UVA): Yes
  Device supports Compute Preemption:      Yes
  Supports Cooperative Kernel Launch:      Yes
  Supports MultiDevice Co-op Kernel Launch: Yes
  Device PCI Domain ID / Bus ID / location ID: 0 / 33 / 1
  Compute Mode:
    < Default (multiple host threads can use ::cudaSetDevice() with device simultaneously) >

deviceQuery, CUDA Driver = CUDART, CUDA Driver Version = 10.1, CUDA Runtime Version = 10.1, NumDevs = 1
Result = PASS
root@ecs-474b:/usr/local/cuda-10.1/samples/1_Utilities/deviceQuery#
```

13. 执行以下命令查看CUDA的版本。

```
/usr/local/cuda/bin/nvcc -V
```

图 2-101 查看 CUDA 的版本

```
root@ecs-474b:/usr/local/cuda-10.1/samples/1_Utilities/deviceQuery# /usr/local/cuda/bin/nvcc -V
nvcc: NVIDIA (R) Cuda compiler driver
Copyright (c) 2005-2019 NVIDIA Corporation
Built on Fri_Feb__8_19:08:17_PST_2019
Cuda compilation tools, release 10.1, V10.1.105
root@ecs-474b:/usr/local/cuda-10.1/samples/1_Utilities/deviceQuery#
```

14. 执行以下命令启动持久模式。

```
sudo nvidia-smi -pm 1
```

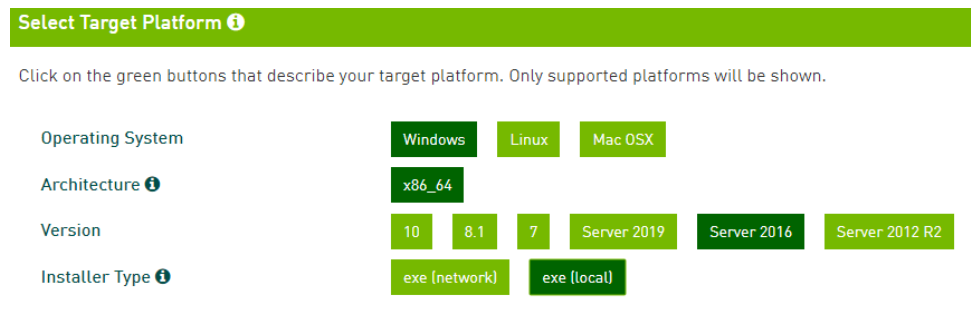
启用持久模式设置可以优化Linux实例上GPU设备的性能。

Windows 操作系统安装 CUDA 工具包

以下操作以Windows Server 2016 Standard 64bit操作系统GPU实例安装CUDA 10.1为例。

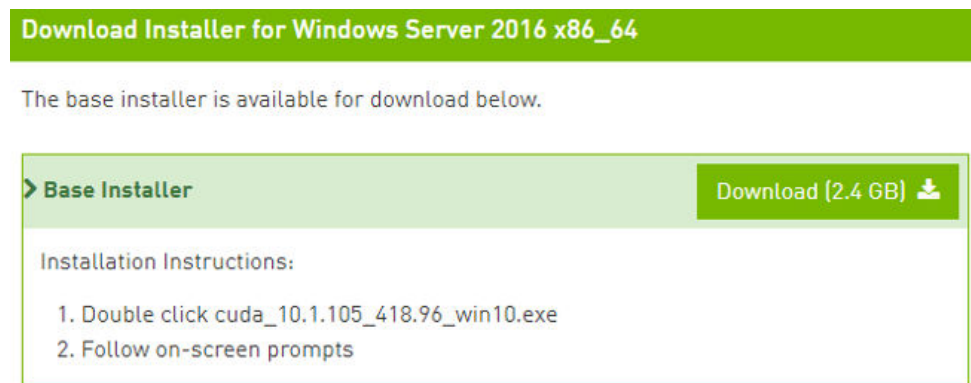
1. 登录云服务器。
2. 在CUDA下载页面中，按照[CUDA工具包下载地址](#)中的对应的索引项在页面中进行选择。

图 2-102 选择 CUDA 版本



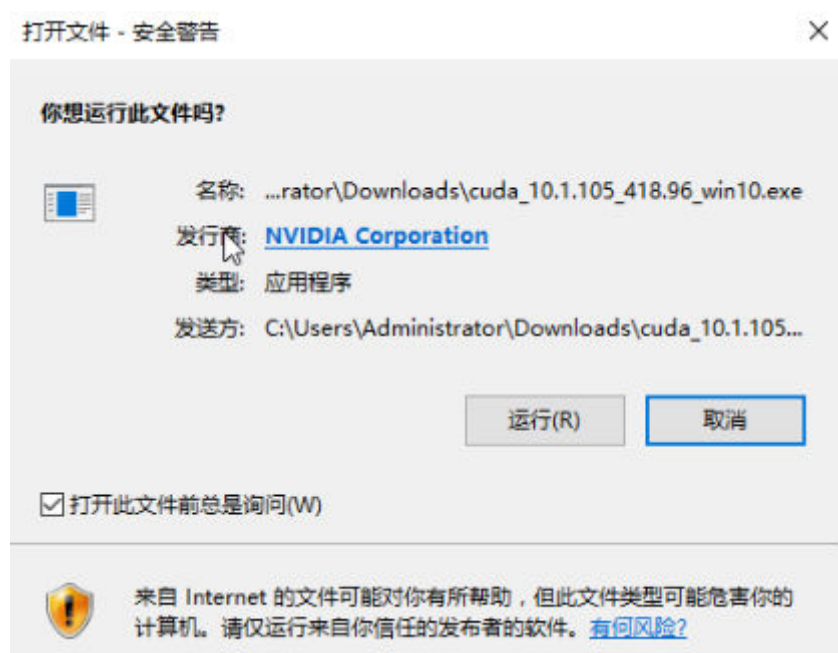
3. 选择完成后，页面会自动呈现出Windows Server 2016 Standard 64bit对应的CUDA 10.1的下载地址。

图 2-103 Windows 云服务器下载 CUDA



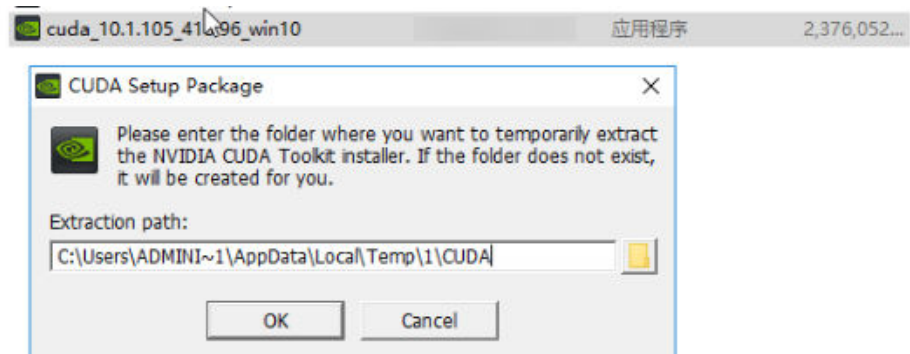
4. 单击“Download”下载CUDA工具包。
5. 双击打开安装文件，单击“运行”安装CUDA工具包。

图 2-104 Windows 云服务器安装 CUDA



6. 选择安装地址，在“CUDA Setup Package”界面，单击“OK”

图 2-105 选择 CUDA 安装路径



7. 根据安装提示完成CUDA的安装。

图 2-106 CUDA 安装完成



8. 检查CUDA是否安装成功。
打开cmd命令窗口，执行以下命令。
nvcc -V
如果回显信息中出现CUDA的版本信息，说明CUDA安装成功。

图 2-107 CUDA 安装成功

```
C:\Users\Administrator>nvcc -V
nvcc: NVIDIA (R) Cuda compiler driver
Copyright (c) 2005-2019 NVIDIA Corporation
Built on Fri_Feb__8_19:08:26_Pacific_Standard_Time_2019
Cuda compilation tools, release 10.1, V10.1.105

C:\Users\Administrator>_
```

2.6.5 卸载 GPU 加速型 ECS 的 GPU 驱动

操作场景

当GPU加速型云服务器需手动卸载GPU驱动时，可参考本文档进行操作。

GPU驱动卸载命令与GPU驱动的安装方式和操作系统类型相关，例如：

- [Windows操作系统卸载驱动](#)
- [Linux操作系统卸载驱动](#)

Windows 操作系统卸载驱动

以Windows Server 2016 数据中心版 64位操作系统为例，介绍GPU加速型云服务器卸载NVIDIA驱动（驱动版本462.31）的操作步骤。

1. 登录弹性云服务器。
2. 单击“开始”，打开“控制面板”。
3. 在控制面板中，单击“卸载程序”。

图 2-108 单击卸载程序



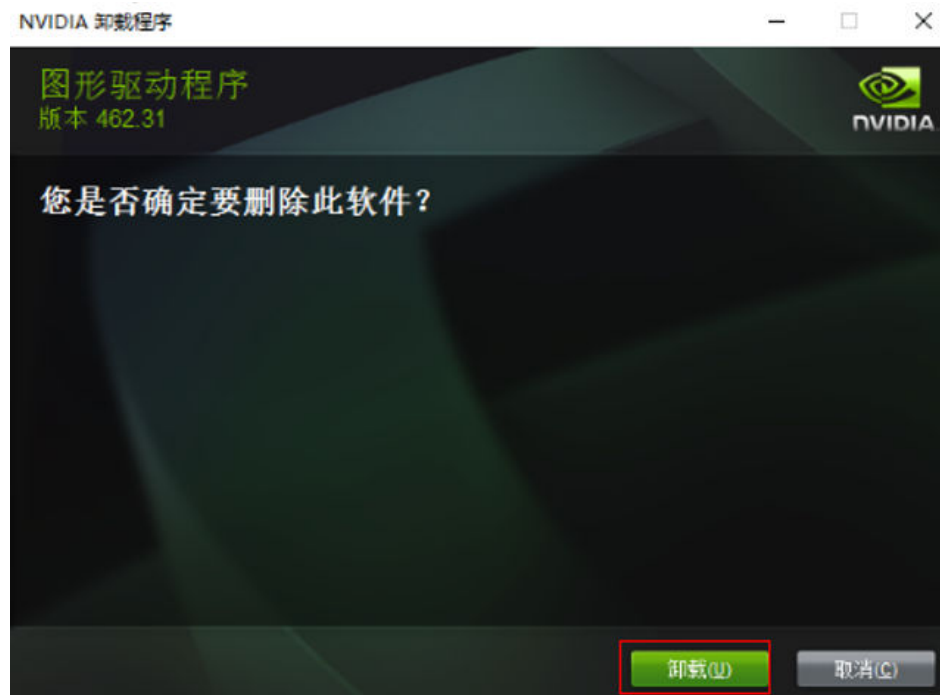
4. 右键单击要卸载的NVIDIA驱动，单击“卸载/更改”。

图 2-109 卸载驱动



5. 在弹出的“NVIDIA 卸载程序”界面，单击“卸载”。

图 2-110 NVIDIA 卸载程序



6. 卸载完成后，单击“稍后启动”。
7. 检查驱动是否卸载成功。
 - a. 在控制面板，单击“设备管理器”。
如果“显示适配器”中没有NVIDIA显卡，表明驱动卸载成功。

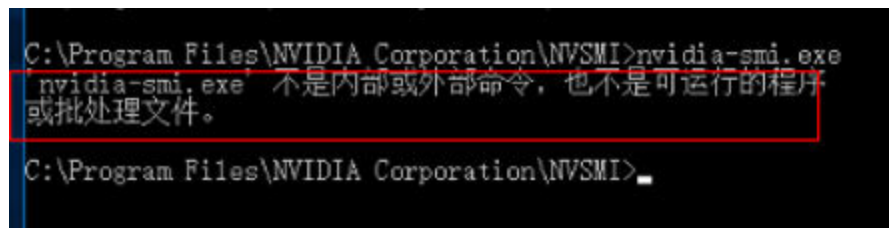
图 2-111 查看显示适配器



b. 打开云服务器cmd窗口，执行以下命令：

```
cd C:\Program Files\NVIDIA Corporation\NVSMI  
nvidia-smi.exe
```

图 2-112 命令执行结果



如果回显信息为该文件不存在，则说明驱动卸载成功。

确认NVIDIA驱动卸载完成后，可以先不重启云服务器，直接安装新的NVIDIA驱动。

Linux 操作系统卸载驱动

对于使用run安装包安装的NVIDIA Tesla驱动，建议您按照以下步骤卸载NVIDIA Tesla驱动。

说明

如果您使用run安装包安装NVIDIA Grid驱动，则卸载NVIDIA驱动即可，只需执行下面的步骤1。

以Ubuntu20.04 server 64bit为例，介绍Tesla 460.73.01、CUDA 11.2卸载的操作步骤。

1. 卸载NVIDIA驱动。
 - a. 执行以下命令，查询“nvidia-uninstall”所在路径。

whereis nvidia-uninstall

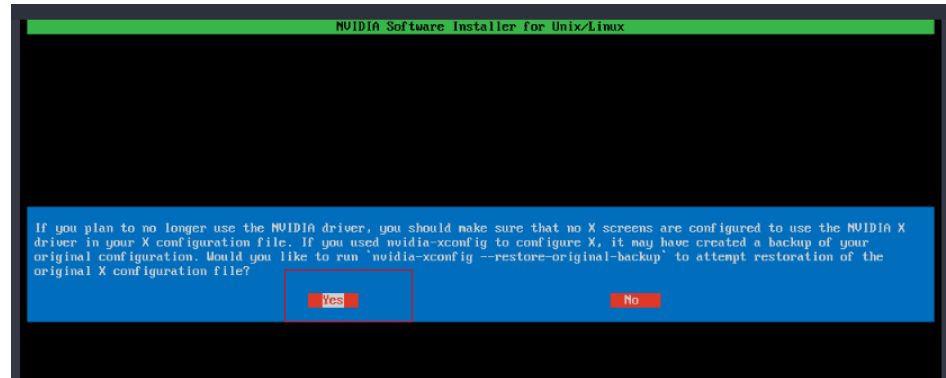
“nvidia-uninstall”一般在“/usr/bin/”目录下。

图 2-113 查询卸载路径



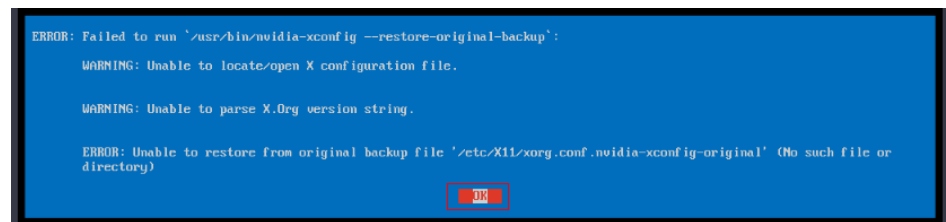
- b. 执行以下命令，基于查询的“nvidia-uninstall”路径，卸载驱动。
/usr/bin/nvidia-uninstall
 - c. 选择” Yes” ，单击回车键。

图 2-114 卸载驱动选择界面



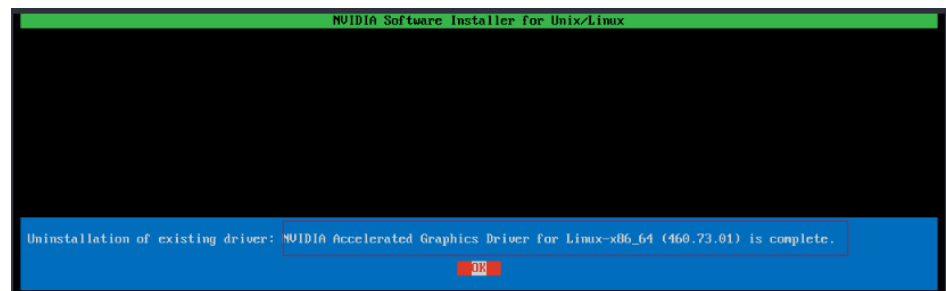
- d. 选择” OK” ，单击回车键。

图 2-115 卸载过程选择界面



- e. 驱动卸载成功，单击回车键。

图 2-116 卸载成功界面



2. 卸载CUDA库和cuDNN库

当需要升级CUDA驱动版本时，需要卸载对应的CUDA库后，再安装对应的CUDA版本。

- a. 执行以下命令，卸载CUDA库。

```
/usr/local/cuda/bin/cuda-uninstaller
```

“cuda-uninstaller”一般在“/usr/local/cuda/bin”目录下。

说明

不同CUDA版本的卸载命令可能存在差异，如果未找到“cuda-uninstaller”文件，可在“/usr/local/cuda/bin/”目录下查看是否存在“uninstall_cuda”开头的文件。

如果有，则将命令中的“cuda-uninstaller”替换为“uninstall_cuda”开头的文件名。

- b. 在卸载界面，勾选全部选项后，移动光标至“Done”，单击回车键。

图 2-117 卸载 CUDA 驱动



CUDA库卸载成功，会返回”Successfully uninstalled”。

- c. 移除CUDA库和cuDNN库：

```
rm -rf /usr/local/cuda-11.2
```

2.7 管理 ECS 配置信息

2.7.1 更改时区

操作场景

弹性云服务器默认设置的时区，是您制作镜像时选择的时区。如需修改，请参见本节内容，将其更改为所需的本地时间或网络中的其他时区。

弹性云服务器登录成功后，如果发现弹性云服务器系统时间与本地时间不一致，建议更改时区，将弹性云服务器系统时间与本地时间进行同步。

更改 Linux 弹性服务器的时区

针对不同的Linux操作系统，更改时区的操作略有不同。本节以CentOS 6.x 64bit为例，介绍更改Linux弹性云服务器时区的操作方法。

1. 登录弹性云服务器。
2. 执行以下命令，切换至root账号。
3. 执行以下命令，查询支持使用的时区。

```
su - root
```

```
ls /usr/share/zoneinfo/
```

其中，“/usr/share/zoneinfo”目录显示了时区数据文件的目录结构。您可以通过该目录结构，查找您所需时区的文件。

“/usr/share/zoneinfo”目录中显示的信息，部分为时区，部分为目录。其中，目录包含了针对特定城市的时区文件，您可以在该目录查找要用于弹性云服务器的所在城市时区。

示例如下：

- 假设您需要使用中国上海所在的时区，则需先执行命令 `ls /usr/share/zoneinfo/Asia` 打开目录。
其时区文件目录为 “/usr/share/zoneinfo/Asia/Shanghai”
- 假设您需要使用法国巴黎所在的时区，则需先执行命令 `ls /usr/share/zoneinfo/Europe` 打开目录，其时区文件目录为 “/usr/share/zoneinfo/Europe/Paris”。

4. 设置新时区。
 - a. 执行以命令，打开 “/etc/sysconfig/clock” 文件。
 - b. 查找ZONE条目，将其更改为所需的时区文件名称。

示例如下：

- 假设更改为中国上海所在时区，需将ZONE条目更改为：
ZONE="Asia/Shanghai"
- 假设更改为法国巴黎所在时区，需将ZONE条目更改为：
ZONE="Europe/Paris"

5. 按“ESC”，执行以下命令，保存并退出文件。
6. 执行以下命令，查询弹性云服务器中是否已经存在“/etc/localtime”文件。

```
:wq
```

```
ls /etc/localtime
```

- 是，执行7。
- 否，跳转8。

7. 执行以下命令，删除已有的“/etc/localtime”文件。

```
rm /etc/localtime
```

8. 执行以下命令，在“/etc/localtime”与时区文件之间创建一个符号链接，使得弹性云服务器在引用本地时间信息时找到此时区文件。

```
ln -sf /usr/share/zoneinfo/Asia/city1 /etc/localtime
```

9. 执行以下命令，重启系统，使得所有服务和应用程序接受新时区信息。

reboot

10. 重新登录弹性云服务器，使用root账号执行以下命令，查询弹性云服务器的时区，确认是否更改成功。

ls -lh /etc/localtime

回显信息如下所示：

```
# ls -lh /etc/localtime  
lrwxrwxrwx 1 root root 33 Nov 27 11:01 /etc/localtime -> /usr/share/zoneinfo/Asia/city1
```

更改 Windows 弹性云服务器的时区

1. 登录弹性云服务器。
2. 左键单击任务栏右下方的时间，选择“更改日期和时间设置”。系统进入“日期和时间”页面

图 2-118 日期和时间



3. 单击“更改时区”。系统进入“时区设置”页面。
4. 在“设置时区”栏的下拉框中选择待更换的时区。
5. 单击“确定”，完成Windows弹性云服务器的时区更换。

2.7.2 开启/关闭 ECS 超线程

操作场景

购买x86架构弹性云服务器时，您可以通过设置“CPU选项”，开启或关闭超线程。若不设置，则默认开启超线程。

对于支持超线程的x86架构弹性云服务器实例：

- 开启超线程：适用于需要CPU内核在同一时间并行处理更多的信息和后台任务的场景，开启多线程可以大幅提升计算体验。
- 关闭超线程：适用于计算密集型，且关闭超线程时性能优于开启超线程的场景，例如材料计算等HPC场景。

您可以在购买x86架构弹性云服务器或变更规格时，开启或关闭超线程。

背景知识

x86架构弹性云服务器的处理器运用超线程HT（Hyper-Threading）技术，允许在CPU的每个物理内核上公开两个执行上下文，即一个物理内核包含两个虚拟的“逻辑内核”，可以处理不同的软件线程。vCPU（virtual CPU）即为虚拟的“逻辑内核”。

规格名称展示vCPU数，即逻辑内核数。在x86架构弹性云服务器上可以查看实际的CPU逻辑内核数。查看方法，请参见[查看ECS实际的CPU逻辑内核数](#)。

当前绝大多数x86架构规格已经默认开启了超线程，如果在创建x86架构弹性云服务器或者变更规格时关闭了超线程，则在弹性云服务器上查看到的CPU核数是规格的Flavor名称中展示的vCPU数量的一半。

例如，对于c7.xlarge.2，其vCPU数，即逻辑内核数为4，4个逻辑内核对应了底层4个HT（超线程）。若关闭了超线程，则c7.xlarge.2的vCPU数，即逻辑内核数为2，2个逻辑内核对应了底层2个物理核。

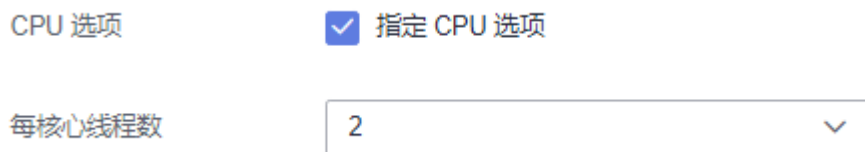
约束与限制

- x86弹性云服务器购买完成后无法修改超线程状态，需要通过变更规格的方式修改ECS实例的超线程状态。
- 该操作不涉及收费。
- 支持开启或关闭超线程的规格，请参见[规格清单（X86）](#)。

开启/关闭超线程（购买时）

1. 登录控制台，进入[购买弹性云服务器](#)页面。
根据业务需要，完成基础配置、网络配置以及高级配置。详细内容，请参见[购买弹性云服务器](#)。
2. 勾选“现在配置”，展开“高级选项”。
3. 勾选“指定CPU选项”。

图 2-119 设置 CPU 选项



4. 设置“每核心线程数”。
当勾选“指定CPU选项”时，显示该参数。在“每核心线程数”下拉框进行设置。
 - 1: 关闭超线程
 - 2: 开启超线程，默认开启
5. 单击“下一步：确认配置”，确认参数并完成弹性云服务器的购买。
购买后，您可以通过[查询云服务器详情](#)中的“cpu_options”参数查询超线程状态。

开启/关闭超线程（变更规格时）


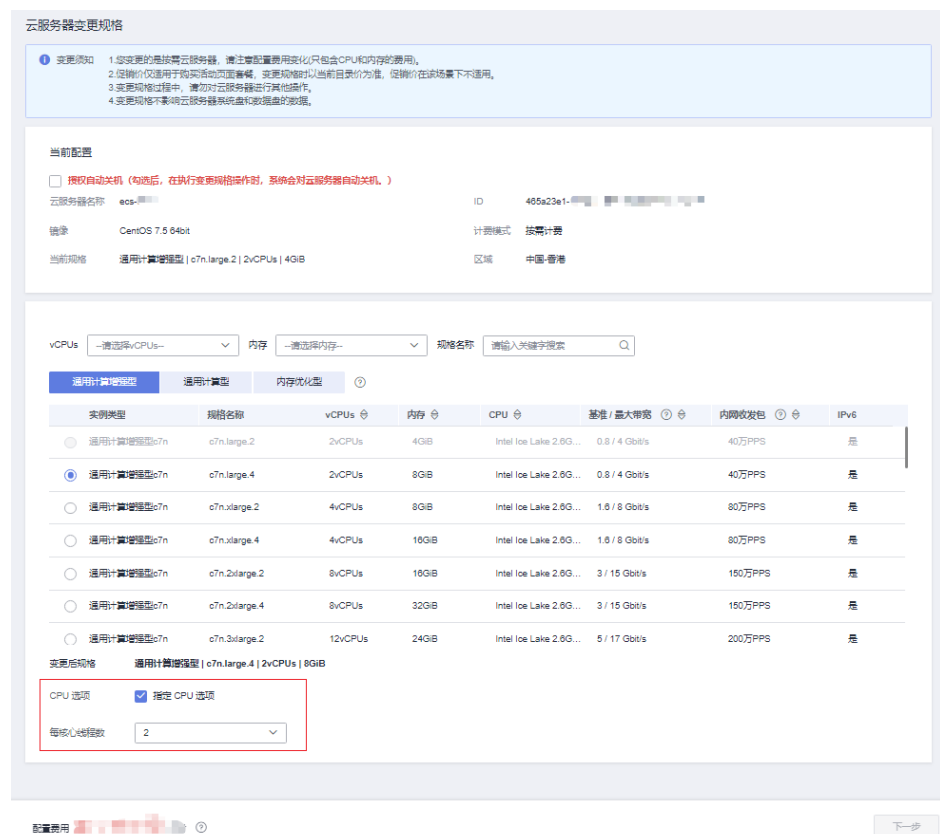
1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表，单击待变更的弹性云服务器“操作”列下的“更多 > 变更规格”。
系统进入“云服务器变更规格”页面。
5. 根据界面提示，选择变更后的弹性云服务器类型、vCPU和内存。
变更规格前请先将云服务器关机，或根据页面提示勾选“系统自动关机后变更规格”。

图 2-120 变更规格



6. 单击“下一步”。
7. 确认变更后的配置无误后，阅读并勾选同意服务协议，单击“提交申请”开始变更。
变更完成后，您可以通过[查询云服务器详情](#)中的“cpu_options”参数查询超线程状态。

查看 ECS 实际的 CPU 逻辑内核数

您可以登录ECS后，在ECS上查看实际的CPU逻辑内核数。

- 对于Linux系统ECS
 - a. [登录ECS](#)。
 - b. 执行以下命令，查询ECS的CPU逻辑内核数。

lscpu

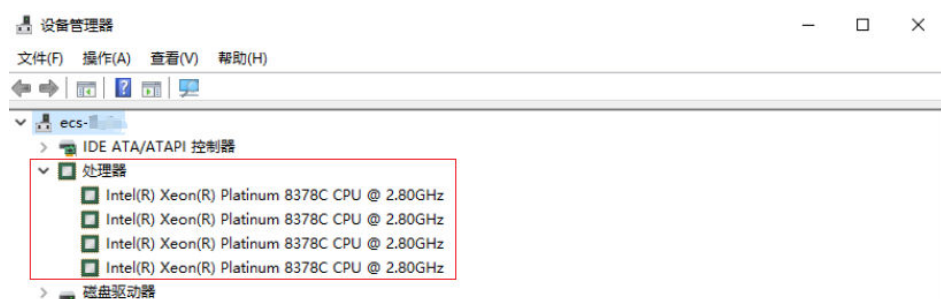
如[图2-121](#)所示，“CPU(s)”表示CPU逻辑内核数。

图 2-121 查询结果

```
[root@ecs-~]# lscpu
Architecture:          x86_64
CPU op-mode(s):        32-bit, 64-bit
Byte Order:             Little Endian
CPU(s):                 4
On-line CPU(s) list:   0-3
Thread(s) per core:    2
Core(s) per socket:    2
Socket(s):              1
NUMA node(s):          1
Vendor ID:              GenuineIntel
```

- 对于Windows系统ECS
 - a. [登录ECS](#)。
 - b. 打开“控制面板 > 设备管理器”，展开“处理器”，即可查看当前ECS实际的CPU逻辑内核数（线程）。

图 2-122 查询结果



2.7.3 使用实例自定义数据和元数据

2.7.3.1 通过元数据获取 ECS 基本信息

操作场景

弹性云服务器元数据包含了弹性云服务器在云平台的基本信息，例如云服务ID、主机名、网络信息等。弹性云服务器元数据支持两种风格，可以分别通过兼容Openstack和兼容EC2的API获取，如表2-40所示。对于支持的元数据类型，本节详细介绍了其URI和使用方法。

使用须知

如果元数据中包含了敏感数据，您应当采取适当的措施来保护敏感数据，比如限制访问范围、加密等。

防火墙配置示例如下所示：

- Windows

如果您不希望管理员之外的用户访问实例自定义数据，请以管理员身份开启防火墙后，在PowerShell 中执行下面的命令，阻止非管理员用户访问实例自定义数据：

```
PS C:\>$RejectPrincipal = New-Object -TypeName System.Security.Principal.NTAccount ("Everyone")
```

```
PS C:\>$RejectPrincipalSID = $RejectPrincipal.Translate([System.Security.Principal.SecurityIdentifier]).Value
```

```
PS C:\>$ExceptPrincipal = New-Object -TypeName System.Security.Principal.NTAccount ("Administrator")
```

```
PS C:\>$ExceptPrincipalSID = $ExceptPrincipal.Translate([System.Security.Principal.SecurityIdentifier]).Value
```

```
PS C:\>$PrincipalSDDL = "O:LSD:(D;;CC;;;$ExceptPrincipalSID)(A;;CC;;;$RejectPrincipalSID)"
```

```
PS C:\>New-NetFirewallRule -DisplayName "Reject metadata service for $($RejectPrincipal.Value), exception: $($ExceptPrincipal.Value)" -Action block -Direction out -Protocol TCP -RemoteAddress 169.254.169.254 -LocalUser $PrincipalSDDL
```

- Linux

如果您不希望root之外的用户访问自定义实例自定义数据，请以root用户权限执行下面的命令，阻止非root用户访问自定义实例自定义数据：

```
iptables --append OUTPUT --proto tcp --destination 169.254.169.254 --match owner ! --uid-owner root --jump REJECT
```

弹性云服务器元数据类别支持列表

表2-40中不包含兼容EC2类型的ami-id、ami-launch-index、ami-manifest-path、block-device-mapping/、instance-action、instance-id、reservation-id、ramdisk-id、kernel-id元数据项，以上元数据项均无实际含义，不建议使用。

表 2-40 弹性云服务器元数据类别支持列表

元数据类型	实例元数据项	约束说明
OpenStack类型	/meta_data.json	查询弹性云服务器的元数据信息。 元数据的关键字段请参见表2-41。
OpenStack类型	/password	查询弹性云服务器的密码。 Windows系统使用keypairs创建弹性云服务器初始化时cloudbase-init用于保存密文密码。
OpenStack类型	/user_data	查询弹性云服务器的自定义数据。 用户根据需要自行指定脚本和配置文件用于弹性云服务器初始化，详细操作请参考 通过实例自定义数据配置ECS实例 。 若linux虚拟化使用密码方式，则保存注入密码的脚本。
OpenStack类型	/network_data.json	查询弹性云服务器的网络信息。
OpenStack类型	/securitykey	获取临时的AK、SK。 对弹性云服务器获取临时的AK、SK，需要在IAM上对op_svc_ecs账户进行委托授权，并对相应的弹性云服务器资源进行授权委托管理。 说明 委托的授权权限由用户自行设置，建议根据实际业务需要进行最小化授权。 ECS服务本身不会使用委托对资源进行操作。
OpenStack类型	/spot/instance-action	查询竞价实例关机提示。
兼容EC2类型	/meta-data/ hostname	查询弹性云服务器的主机名称。 请参考以下链接为弹性云服务器去掉后缀.novalocal: 弹性云服务器的主机名带后缀.novalocal
兼容EC2类型	/meta-data/ local-hostname	该字段含义同hostname一致。
兼容EC2类型	/meta-data/ public-hostname	该字段含义同hostname一致。
兼容EC2类型	/meta-data/ instance-type	查询弹性云服务器的规格名称。
兼容EC2类型	/meta-data/ local-ipv4	查询弹性云服务器的固定IP地址。 多网卡情况下，只显示主网卡的地址。

元数据类型	实例元数据项	约束说明
兼容EC2类型	/meta-data/ placement/ availability-zone	查询弹性云服务器的AZ信息。
兼容EC2类型	/meta-data/ public-ipv4	查询弹性云服务器的弹性公网IP地址。 多网卡情况下，只显示主网卡的弹性公网IP地址。
兼容EC2类型	/meta-data/ public-keys/0/ openssh-key	查询弹性云服务器的公钥。
兼容EC2类型	/user-data	查询弹性云服务器的自定义数据。
兼容EC2类型	/meta-data/ security-groups	查询弹性云服务器所使用的安全组名称。

表 2-41 metadata 关键字段

参数	参数类型	描述
uuid	String	弹性云服务器的ID。
availability_zone	String	弹性云服务器所在可用区。
meta	Dict	元数据信息，包括镜像名称、镜像ID、VPC ID等信息。
hostname	String	弹性云服务器主机名。 请参考以下链接为弹性云服务器去掉后缀.novalocal: 弹性云服务器的主机名带后缀.novalocal
enterprise_project_id	String	查询弹性云服务器的企业项目ID信息。

前提条件

- 已登录弹性云服务器。
- 请确保安全组出方向规则满足如下要求，否则访问元数据请求失败：
 - 协议：TCP
 - 端口范围：80
 - 远端地址：169.254.0.0/16

说明

如果您使用的是默认安全组出方向规则，则已经包括了如上要求，可以正常访问元数据。默认安全组出方向规则请参见[默认安全组和规则](#)。

Metadata (OpenStack 元数据 API)

用于查询弹性云服务器的元数据。

- URI
/169.254.169.254/openstack/latest/meta_data.json
- 方法
支持GET请求。
- 示例:

Linux操作系统：以使用CURL工具为例，介绍查询弹性云服务器元数据的方法。

curl http://169.254.169.254/openstack/latest/meta_data.json

Windows操作系统：以使用Invoke-RestMethod工具为例，介绍查询弹性云服务器元数据的方法。

Invoke-RestMethod http://169.254.169.254/openstack/latest/meta_data.json | ConvertTo-Json

```
{
  "random_seed": "rEocCViRS+dNwlydGlxJHUp+00poeUsAdBFkbPbYQTmpNwpoEb43k9z+96TyrekNKS
+iLYDdRny4kKGoNPEVBCC05Hg1TcDbIAPfJwgJS1okqEtlcofUhKmL3K0fto
+5KXEDU3GNuGwyZXjdVb9HQWU+E1jztAJjqsahnU+g/tawABTVySLBKlAT8fMGax1mTGgArucn/
WzDcy19DGioKPE7F8lLtSQ4Ww3VCLK5VYB/h0x+4r7IVHrPmYX/
bi1Yhm3Dc4rRYNaTjdOV5gUOsbo3oAeQkmKwQ/
NO0N8qw5Ya4l8ZUW4tMav4mOsRySOOB35v0bvaJc6p
+50DTbWNex5A2MLiEhTP3vsPrmvk4LRF7CLz2J2TGIM14OoVBw7LARwmv9cz532zHki/c8tlhRzLmOTXh/
wL36zFW10DeuReUGmxth7IGNmRMQKV6+mil78jm/KMPpgAdK3vwYF/
GcelOFJD2HghMUUCeMbwYnvijLTejuBpwhJMNIHA/NvlEsxJDxqBCoss/Jfe+yCmUFyxovJ
+L8oNkTzkmCNzw3Ra0hiKchGhqK3BleToV/kVx5DdF081xrEA
+qyoM6CVyfJtEoz1zLRyoo9bJ65Eg6Jd8dj1UCVsDqRY1pljgzE/
Mzsw6AaaCVhaMJL7u7YMVdyKzA6z65Xtvujz0Vo=",
  "uuid": "ca9e8b7c-f2be-4b6d-a639-f10b4d994d04",
  "availability_zone": "lt-test-1c",
  "enterprise_project_id": "0",
  "hostname": "ecs-ddd4.novalocal",
  "launch_index": 0,
  "instance_type": "s3.medium.2",
  "meta": {
    "metering.image_id": "3a64bd37-955e-40cd-ab9e-129db56bc05d",
    "metering.imagetype": "gold",
    "metering.resourcespeccode": "s3.medium.2.linux",
    "metering.cloudServiceType": "hws.service.type.ec2",
    "image_name": "CentOS 7.6 64bit",
    "metering.resourcetype": "1",
    "vpc_id": "3b6c201f-aeb3-4bce-b841-64756e66cb49",
    "os_bit": "64",
    "cascaded.instance_extrainfo": "pcibrige:1",
    "os_type": "Linux",
    "charging_mode": "0"  },
  "region_id": "xxx",
  "project_id": "6e8b0c94265645f39c5abbe63c4113c6",
  "name": "ecs-ddd4"
}
```

User data (OpenStack 元数据 API)

用于查询弹性云服务器的自定义数据。该值仅在创建弹性云服务器时指定，不支持修改。

- URI
/169.254.169.254/openstack/latest/user_data

- 方法
支持GET请求。
- 示例

Linux操作系统：

```
curl http://169.254.169.254/openstack/latest/user_data
```

Windows操作系统：

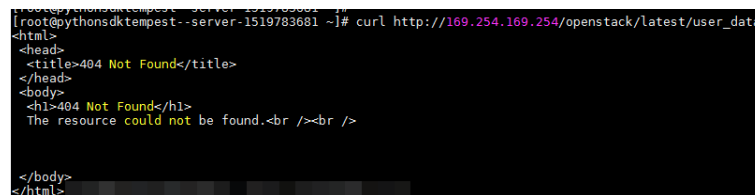
```
Invoke-RestMethod http://169.254.169.254/openstack/latest/user_data
```

```
ICAgICAgDQoiQSBjbG91ZCBkb2VzIG5vdCBrbm93IHdoeSBpdCBtb3ZlcyBpbjBqdXN0IHN1Y2ggYSBkaXJlY  
3Rpb24gYW5kIGF0IHN1Y2ggYSBzcGVlZC4uLkI0IGZlZWxzIGFuIGltcHVsc2lubi4uLnRoZXN1Y2ggYSBkaXJlY  
BsYWNlIHJlY2VzIG5vdy4gQnV0IHRob3R0ZSBza3kga25vd3MgdGhllHJlYXNvbnMgYW5kIHRob3R0ZSBwYXR0Z  
XJucyBiZWphbmQgYWxslGNsb3VkcycgYW5kIHlvdSB3aWxslGtub3csIHRob3R0ZSBwYXR0ZSBwYXR0ZSBwYXR0Z  
Vyc2VsZiBoaWdodGVub3VnaCB0byBzZWUgYmV5b25kIGhvcml6b25zLiINCg0KLVJpY2hhcmQgQmFjaA=  
=
```

📖 说明

如果创建弹性云服务器时未注入自定义数据，此时，该接口的查询结果是404，如图2-123所示。

图 2-123 404 Not Found



```
iroot@pythonsdktempest-1519783681 ~]# curl http://169.254.169.254/openstack/latest/user_data
<html>
<head>
<title>404 Not Found</title>
</head>
<body>
<h1>404 Not Found</h1>
The resource could not be found.<br /><br />
</body>
</html>
```

Network data (OpenStack 元数据 API)

查询弹性云服务器的网络信息，支持查询云服务器挂载的全部网卡的信息，包括网卡所在子网的DNS地址、网络带宽、网卡ID、网卡私有IP地址、网卡弹性公网IP地址、网卡的MAC地址。

- URI
/openstack/latest/network_data.json
- 方法
支持GET请求
- 示例

📖 说明

instance_max_bandwidth、instance_min_bandwidth的单位是Mbps，如果取值为-1，说明对带宽大小无限制。

Linux操作系统：

```
curl http://169.254.169.254/openstack/latest/network_data.json
```

Windows操作系统：

```
Invoke-RestMethod http://169.254.169.254/openstack/latest/  
network_data.json | ConvertTo-Json
```

```
{  
  "services": [{  
    "type": "dns",  
    "address": "xxx.xx.x.x"  }  
]
```

```
},
{
  "type": "dns",
  "address": "100.125.21.250"
}],
"qos":{
  "instance_min_bandwidth": 100,
  "instance_max_bandwidth": 500
},
"networks": [{
  "network_id": "67dc10ce-441f-4592-9a80-cc709f6436e7",
  "type": "ipv4_dhcp",
  "link": "tap68a9272d-71",
  "id": "network0"
}],
"links": [{
  "vif_id": "68a9272d-7152-4ae7-a138-3ef53af669e7",
  "public_ipv4": "100.100.xx.xx",
  "ethernet_mac_address": "fa:16:3e:f7:c1:47",
  "mtu": null,
  "local_ipv4": "192.169.10.10",
  "type": "cascading",
  "id": "tap68a9272d-71"
}]
}
```

Security Key (OpenStack 元数据 API)

获取临时的AK、SK。

📖 说明

- 您如果需要在弹性云服务器获取临时的AK、SK，需要在IAM对弹性云服务器进行云服务委托授权，并在相应的弹性云服务器的详情页面“管理信息 > 委托”栏目中添加委托。
详细操作请参考[委托其他云服务管理资源](#)。
- 临时的AK、SK有效期一个小时。临时AK、SK更新时间早于失效时间10分钟，10分钟内新旧临时AK、SK均可用。
- 使用临时AK、SK时，需要在消息的header中增加'X-Security-Token':{securitytoken}。其中，securitytoken就是调用接口返回的值。

- URI

/openstack/latest/securitykey

- 方法

支持GET请求

- 示例

Linux操作系统：

curl http://169.254.169.254/openstack/latest/securitykey

Windows操作系统：

Invoke-RestMethod http://169.254.169.254/openstack/latest/securitykey

Instance Action (OpenStack 元数据 API)

用于查询竞价实例关机提示。

📖 说明

如果您的竞价实例即将被中断，该接口会返回预计关机的时间。

Instance Type (EC2-兼容的 API)

用于查询弹性云服务器的规格名称。

- URI
/169.254.169.254/latest/meta-data/instance-type
- 方法
支持GET请求。
- 示例
Linux操作系统：
curl http://169.254.169.254/latest/meta-data/instance-type
Windows操作系统：
Invoke-RestMethod http://169.254.169.254/latest/meta-data/instance-type
s3.medium.2

Local IPv4 (EC2-兼容的 API)

用于查询弹性云服务器的固定IP地址。多网卡情况下，只显示主网卡的地址。

- URI
/169.254.169.254/latest/meta-data/local-ipv4
- 方法
支持GET请求。
- 示例
Linux操作系统：
curl http://169.254.169.254/latest/meta-data/local-ipv4
Windows操作系统：
Invoke-RestMethod http://169.254.169.254/latest/meta-data/local-ipv4
192.1.1.2

Availability Zone (EC2-兼容的 API)

用于查询弹性云服务器的AZ信息。

- URI
/169.254.169.254/latest/meta-data/placement/availability-zone
- 方法
支持GET请求。
- 示例
Linux操作系统：
curl http://169.254.169.254/latest/meta-data/placement/availability-zone
Windows操作系统：
Invoke-RestMethod http://169.254.169.254/latest/meta-data/placement/availability-zone
az1.dc1

Public IPv4 (EC2-兼容的 API)

用于查询弹性云服务器的弹性公网IP地址。多网卡情况下，只显示主网卡的弹性公网IP地址。

- URI
/169.254.169.254/latest/meta-data/public-ipv4
- 方法
支持GET请求。
- 示例
Linux操作系统：
curl http://169.254.169.254/latest/meta-data/public-ipv4
Windows操作系统：
Invoke-RestMethod http://169.254.169.254/latest/meta-data/public-ipv4
46.1.1.2

Public Keys (EC2-兼容的 API)

用于查询弹性云服务器的公钥。

- URI
/169.254.169.254/latest/meta-data/public-keys/0/openssh-key
- 方法
支持GET请求。
- 示例
Linux操作系统：
curl http://169.254.169.254/latest/meta-data/public-keys/0/openssh-key
Windows操作系统：
Invoke-RestMethod http://169.254.169.254/latest/meta-data/public-keys/0/openssh-key
ssh-rsa AAAAB3NzaC1yc2EAAAADAQABAAQADi5Fw5k8Fgzajn1zJwLoV3+wMP+6CyvsSilc/
hioggSnYu/AD0Yqm8vVO0kWlun1rFbdO+QUZKyVr/OPUjQSw4SRh4qsTKf/+eFoWTjplFvd1WCBZzS/
WRenxlwR00KkczHSJro763+wYcwKieb4eKRxaQoQvoFgVjLBULXAJH4eKoKTVNtMXAvPP9aMy2SLgsJNt
Mb9ArfziAibIQynq7UIflN3VclzPeiWrqtzjyOp6CPUXnL0lVPTvbLe8sUteBsjZwL6K4i
+Y0lf3ryqnmQgC21yW4Dzu+kwk8FVT2MgWkCwiZd8gQ/+uJzrJFyMfUOBikIOBfuUENIJUhaB
Generated-by-Nova

相关链接

[Linux操作系统云服务器无法获取元数据怎么办？](#)

2.7.3.2 通过实例自定义数据配置 ECS 实例

使用场景

当您有如下需求时，可以考虑使用实例自定义数据注入功能来配置弹性云服务器：

- 需要通过脚本简化弹性云服务器配置
- 通过脚本初始化系统

- 已有脚本，在创建弹性云服务器的时候一并上传到服务器
- 其他可以使用脚本完成的功能

使用限制

- Linux:
 - 用于创建弹性云服务器的镜像安装了Cloud-init组件。
 - 实例自定义数据大小限制：小于等于32KB。
 - 如果通过文本方式上传实例自定义数据，数据只能包含ASCII码字符；如果通过文件方式上传实例自定义数据，可以包含任意字符，同时，要求文件大小小于等于32KB。
 - 必须是公共镜像，或继承于公共镜像的私有镜像，或自行安装了Cloud-init组件的私有镜像。
 - 必须满足相应Linux弹性云服务器自定义脚本类型的格式要求。
 - 使用的VPC网络必须开启DHCP，安全组出方向规则保证80端口开放。
 - 选择“密码”登录方式时，不支持实例自定义数据注入功能。
- Windows
 - 用于创建弹性云服务器的镜像安装了Cloudbase-init组件。
 - 实例自定义数据大小限制：小于等于32KB。
 - 如果通过文本方式上传实例自定义数据，数据只能包含ASCII码字符；如果通过文件方式上传实例自定义数据，可以包含任意字符，同时，要求文件大小小于等于32KB。
 - 必须是公共镜像，或继承于公共镜像的私有镜像，或自行安装了Cloudbase-init组件的私有镜像。
 - 使用 VPC网络必须开启DHCP，安全组出方向规则保证80端口开放。

使用方法

1. 根据实际需要写User-Data脚本，脚本格式必须符合User-Data脚本，详细规范请参见[相关链接](#)。
2. 创建弹性云服务器时，需将实例自定义数据脚本内容粘贴到“高级选项 > 现在配置 > 实例自定义数据注入”的文本框，或选择实例自定义数据文件上传。

📖 说明

实例自定义数据注入支持使用文件和文本两种方式：

文本：将实例自定义数据脚本内容粘贴到文本框中。

文件：将实例自定义数据脚本内容保存至文本文件后再将上传。

图 2-124 实例自定义数据注入



3. 系统创建弹性云服务器，弹性云服务器启动后自动执行Cloud-init或Cloudbase-init组件，读取实例自定义数据脚本。

关于 Linux 云服务器的实例自定义数据脚本

Linux弹性云服务器的自定义实例自定义数据脚本（简称“脚本”）采用开源的Cloud-init架构实现，该架构以云服务器的元数据为数据来源，对弹性云服务器进行各项自动化配置。自定义脚本类型兼容开源Cloud-init，详情请参考Cloud-init开源说明：<http://cloudinit.readthedocs.io/en/latest/topics/format.html>。

- 脚本的执行时间：弹性云服务器的状态变为“running”之后，执行文件/etc/init之前。

📖 说明

脚本默认以root用户权限执行。

- 脚本类型：包括“User-Data Script”和“Cloud-Config Data”。

表 2-42 Linux 云服务器脚本类型介绍

-	User-Data Script	Cloud-Config Data
描述	以脚本作为自定义配置的手段，如shell和python脚本。	Cloud-init中预定义的对云服务器部分服务进行配置的手段，如yum源、SSH 密钥。
格式	首行必须是“#!”（如“#!/bin/bash和#!/usr/bin/env python”），且前面不能有空格。 首次启动时，将在rc.local-like级别上执行，rc.local-like意味着“在启动序列中非常靠后”。	首行必须是“#cloud-config”，且前面不能有空格。
限制	在采用Base64编码前，脚本内容（包括首行在内）不能超过32KB。	在采用Base64编码前，脚本内容（包括首行在内）不能超过32KB。
频率	仅在首次启动弹性云服务器时执行一次。	配置的服务不同，执行频率也会不同。

- 如何查看注入Linux弹性云服务器的自定义数据？

- 登录弹性云服务器。
- 执行以下命令，以root用户权限查看实例自定义数据。
curl http://169.254.169.254/openstack/latest/user_data

- 脚本使用示例：

该示例介绍如何以不同格式输入Linux弹性云服务器的脚本，并查看脚本的运行结果。

示例一：脚本类型为User-Data Script

创建弹性云服务器时，选择“文本”形式并输入实例自定义数据脚本，脚本如下：

```
#!/bin/bash  
echo "Hello, the time is now $(date -R)" | tee /root/output.txt
```

创建成功后，连接并启动弹性云服务器，执行命令**cat [file]**，查看脚本的运行结果。

```
[root@XXXXXXXX ~]# cat /root/output.txt
Hello, the time is now Mon, 16 Jul 2016 16:03:18+0800
```

示例二：脚本类型为Cloud-Config Data

创建弹性云服务器时，选择“文本”形式并输入实例自定义数据脚本，脚本如下：

```
#cloud-config
bootcmd:
- echo 192.168.1.130 us.archive.ubuntu.com >> /etc/hosts
```

创建成功后，连接并启动弹性云服务器，执行命令`cat /etc/hosts`查看脚本的运行结果。

图 2-125 查看运行结果

```
localhost localhost.localdomain localhost4 localhost4.localdomain4
::1 localhost.localdomain localhost6 localhost6.localdomain6
192.168.1.130 us.archive.ubuntu.com
```

关于 Windows 弹性云服务器的实例自定义数据脚本

Windows弹性云服务器的实例自定义数据脚本（简称“脚本”）采用开源的 cloudbase-init架构实现，该架构以弹性云服务器的元数据为数据来源，对弹性云服务器进行各项自动化配置，为Windows弹性云服务器提供执行初始化脚本的能力。自定义脚本类型兼容开源Cloudbase-init，详情请参考Cloudbase-init开源说明：<https://cloudbase-init.readthedocs.io/en/latest/userdata.html>。

- 脚本类型：包括“Batch批处理程序”和“PowerShell脚本”。

表 2-43 Windows 云服务器脚本类型介绍

-	Batch批处理程序	PowerShell脚本
格式	以“rem cmd”开头且作为首行，且前面不能有空格。	以“#ps1”开头且作为首行，且前面不能有空格。
限制	在采用Base64编码前，脚本内容（包括首行在内）不能超过32KB。	在采用Base64编码前，脚本内容（包括首行在内）不能超过32KB。

- 如何查看注入Windows弹性云服务器的实例自定义数据？
 - 登录弹性云服务器。
 - 在浏览器栏访问如下地址，查看注入的实例自定义数据。

http://169.254.169.254/openstack/latest/user_data

- 脚本使用示例

该示例介绍如何以不同格式输入Windows弹性云服务器的脚本，并查看脚本的运行结果。

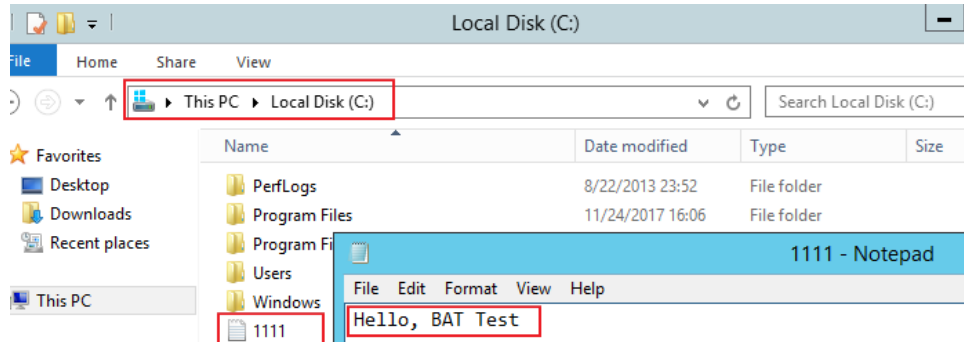
示例一：Batch批处理程序

创建弹性云服务器时，选择“文本”形式并输入实例自定义数据，脚本如下：

```
rem cmd
echo "Hello, BAT Test" > C:\1111.txt
```

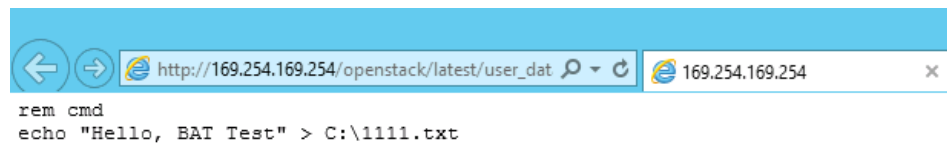
创建成功后，连接并启动弹性云服务器，查看脚本的运行结果：在C:\盘添加了一个名称为“1111”的文本文件，如图2-126所示。

图 2-126 新建文本文档_Batch



如需查看注入Windows弹性云服务器的实例自定义数据，可以在浏览器访问：http://169.254.169.254/openstack/latest/user_data，如图2-127所示。

图 2-127 查看实例自定义数据_Batch



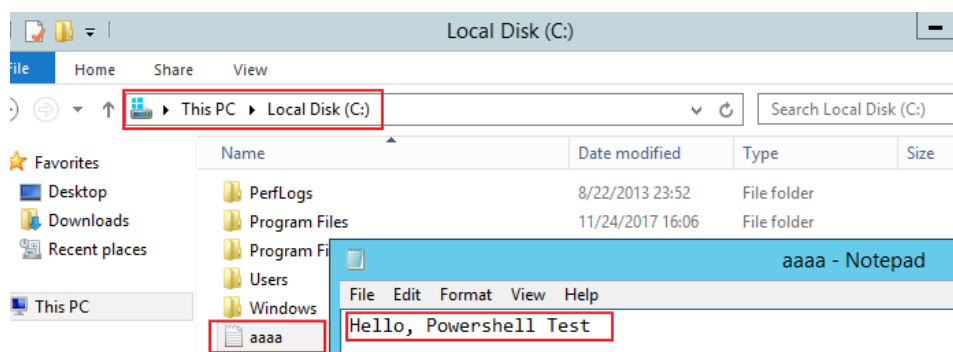
示例二：PowerShell脚本

创建弹性云服务器时，选择“文本”形式并输入实例自定义数据，脚本如下：

```
#ps1  
echo "Hello, Powershell Test" > C:\aaaa.txt
```

创建成功后，连接并启动弹性云服务器，查看脚本的运行结果：在C:\盘添加了一个名称为“aaaa”的文本文件。

图 2-128 新建文本文档_PowerShell



如需查看注入Windows弹性云服务器的实例自定义数据，可以在浏览器访问：http://169.254.169.254/openstack/latest/user_data，如图2-129所示。

图 2-129 查看实例自定义数据_PowerShell



案例 1

该样例介绍如何通过实例自定义数据注入，为Linux弹性云服务器简化云服务器配置。

当您需要设置vim的“语法高亮显示”、“制表符占用4个空格”和“显示行数”属性时，就可以写一个脚本.vimrc，并将其注入到“/root/.vimrc”。在您创建完云服务器时，您的vim配置便完成了。这样做可以大幅提高系统配置的效率，特别是在您一次性创建多台弹性云服务器时。

实例自定义数据注入示例：

```
#cloud-config
write_files:
- path: /root/.vimrc
  content: |
    syntax on
    set tabstop=4
    set number
```

案例 2

该样例介绍如何通过实例自定义数据注入，设置Linux弹性云服务器密码。

说明

新密码必须符合密码复杂度要求，密码规则如表2-44所示。

表 2-44 密码设置规则

参数	规则
密码	<ul style="list-style-type: none">密码长度范围为8到26位。密码至少包含以下4种字符中的3种：<ul style="list-style-type: none">大写字母小写字母数字Windows操作系统云服务器支持的特殊字符：!@\$%^_-=+[{()};,./?~#*Linux操作系统云服务器支持的特殊字符：!@\$%^_-=+[{()};,./?~#*密码不能包含用户名或用户名的逆序。Windows操作系统的云服务器，不能包含用户名中超过两个连续字符的部分。

实例自定义数据注入示例：

使用密文密码（推荐使用）：

```
#!/bin/bash
echo 'root:$6$V6azyeLwcD3CHlpY$BN3VVq18fmCkj66B4zdHLWevqcXlig' | chpasswd -e;
```

其中，\$6\$V6azyeLwcD3CHlpY\$BN3VVq18fmCkj66B4zdHLWevqcXlig为密文密码，生成方法如下：

1. 执行以下命令，生成加密盐值

```
python -c "import crypt, getpass, pwd;print crypt.mksalt()"
```

回显信息如下：

```
$6$V6azyeLwcD3CHlpY
```

2. 执行以下命令，根据盐值生成密文密码

```
python -c "import crypt, getpass, pwd;print crypt.crypt('Cloud.1234','\$6\
$V6azyeLwcD3CHlpY')"
```

回显信息如下：

```
$6$V6azyeLwcD3CHlpY$BN3VVq18fmCkj66B4zdHLWevqcXlig
```

弹性云服务器创建成功后，您可以使用该密码登录。

案例 3

该样例介绍如何通过实例自定义数据注入，重置Linux弹性云服务器密码。

假设您需要将root用户密码重置为”*****“。

📖 说明

新密码必须符合密码复杂度要求，密码规则如表2-45所示。

表 2-45 密码设置规则

参数	规则
密码	<ul style="list-style-type: none">● 密码长度范围为8到26位。● 密码至少包含以下4种字符中的3种：<ul style="list-style-type: none">- 大写字母- 小写字母- 数字- Windows操作系统云服务器支持的特殊字符：!@\$%^_+=+[{()];,./?~#*- Linux操作系统云服务器支持的特殊字符：!@\$%^_+=+[{()];,./?~#*● 密码不能包含用户名或用户名的逆序。● Windows操作系统的云服务器，不能包含用户名中超过两个连续字符的部分。

实例自定义数据注入示例（请严格遵守下面示例中的缩进标准）：

```
#cloud-config
chpasswd:
```

```
list: |
  root:*****
expire: False
```

弹性云服务器创建成功后，您可以使用重置的密码登录弹性云服务器。为了保证密码安全，建议您在第一次登录弹性云服务器后，修改root用户密码。

案例 4

该样例介绍如何通过实例自定义数据注入，为Windows弹性云服务器新增用户并设置密码。

假设您要创建的用户名为abc、密码为“*****”的普通用户，并将其添加至administrators用户组。

说明

新密码必须符合密码复杂度要求，密码规则如表2-45所示。

实例自定义数据注入示例：

```
rem cmd
net user abc ***** /add
net localgroup administrators abc /add
```

弹性云服务器创建成功后，您可以使用新创建的用户名和密码登录弹性云服务器。

案例 5

该样例介绍如何通过实例自定义数据注入，为Linux弹性云服务器更新系统软件包，并且开启httpd相关服务。注入成功后，您的弹性云服务器就可以使用httpd服务了。

实例自定义数据注入示例：

```
#!/bin/bash
yum update -y
service httpd start
chkconfig httpd on
```

案例 6

该样例介绍如何通过实例自定义数据注入，激活Linux弹性服务器的root用户远程登录权限。注入成功后，您可以使用SSH密钥方式，以root账户登录弹性云服务器。

实例自定义数据注入示例：

```
#cloud-config
disable_root: false
runcmd:
- sed -i 's/^PermitRootLogin.*$/PermitRootLogin without-password/' /etc/ssh/sshd_config
- sed -i '/^KexAlgorithms.*$/d' /etc/ssh/sshd_config
- service sshd restart
```

相关链接

更多关于实例自定义数据注入案例的介绍，请参见Cloud-init/Cloudbase-init官网：

- <https://cloudinit.readthedocs.io/en/latest/>
- <https://cloudbase-init.readthedocs.io/en/latest/>




2.7.4 修改 ECS 名称

操作场景



弹性云服务器创建成功后，您可以根据需求，修改云服务器的名称。

系统支持批量修改多台弹性云服务器的名称，修改完成后，这些弹性云服务器的名称相同。

修改单台弹性云服务器名称

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 将鼠标移动至目标云服务器的“名称/ID”列。
5. 单击弹性云服务器名称右方的“”，根据界面提示，修改云服务器名称。
允许重名：勾选后，允许修改后的名称与其他云服务器名称相同。如果未勾选，且设置的名称与其他云服务器名称相同，此时，系统将提示您该名称已被使用，您需要更换其他名称。
6. 单击“确定”，新名称生效。

批量修改弹性云服务器的名称

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 勾选待修改名称的弹性云服务器。
5. 单击云服务器列表页上方的“更多 > 修改名称”。
6. 设置新名称。
7. 单击“确定”。

批量修改云服务器时，修改后的云服务器名称相同，如全部为ecs-test。

2.7.5 将 ECS 迁移至专属主机

操作场景

云平台支持弹性云服务器在专属主机与公共资源池之间迁移。

在控制台的“弹性云服务器”页面，您可以将公共资源池的弹性云服务器迁移至专属主机，成为专属主机上的弹性云服务器。

本章节介绍公共资源池的弹性云服务器迁移至专属主机的操作指导。



📖 说明

- 迁移云服务器需保证有可用的专属主机资源，否则无法进行迁移操作。
- 在控制台的“专属主机”页面还可以进行如下两个场景的迁移操作，具体请参见《专属主机用户指南》的“[专属主机上云服务器迁移](#)”章节：
 - 创建在专属主机上的弹性云服务器迁移至其他专属主机。
 - 创建在专属主机上的弹性云服务器迁移至公共资源池。

约束与限制

- 关机状态的弹性云服务器才能执行迁移操作。
- 仅当有可用的专属主机资源时，迁移操作才能成功。
- 迁移后云服务器ID保持不变。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在待迁移的弹性云服务器的“操作”列下，单击“更多 > 运维与监控 > 调整所属主机”。
5. 在“调整所属主机”页面中，选择对应的专属主机。

📖 说明

如果无可选择的专属主机，需先完成专属主机资源的创建，详细内容，请参见[购买专属主机](#)。

6. 单击“确定”，将公共资源池的弹性云服务器迁移至专属主机。

2.7.6 管理云服务器组

操作场景

云服务器组是对云服务器的一种逻辑划分，云服务器组中的弹性云服务器遵从同一策略。

当前仅支持反亲和性，即同一云服务器组中的弹性云服务器分散地创建在不同的主机上，提高业务的可靠性。

您可以使用云服务器组将业务涉及到的云服务器分散部署在不同的物理服务器上，以此保证业务的高可用性和底层容灾能力。

云服务器组支持以下操作：

- [创建云服务器组](#)
- [添加云服务器到云服务器组](#)
 - 在创建弹性云服务器时，将弹性云服务器加入云服务器组。
详细操作，请参考[步骤三：高级配置](#)。
 - 在弹性云服务器创建成功后，添加到云服务器组。

- [从云服务器组移出云服务器](#)
- [删除云服务器组](#)

约束与限制



- 当前云服务器组支持反亲和性策略。
- 反亲和性策略云服务器组中，云服务器置放层级为物理机层级。
- 创建的云服务器组个数上限可调整，请联系客服扩大配额。
- 不同区域每个云服务器组最多支持添加的云服务器个数各不相同，请在云服务器组列表页查看可添加的云服务器个数。如[图2-130](#)所示。

图 2-130 查看可添加的云服务器数量



创建云服务器组

您可以申请创建一个云服务器组，同一云服务器组中的弹性云服务器遵从相同策略，云服务器组与云服务器组之间没有关联关系。



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“弹性云服务器 > 云服务器组”。
5. 在“云服务器组”页面，单击“创建云服务器组”。
6. 输入云服务器组的名称。
7. 选择云服务器组的策略。
8. 单击“确定”。

添加云服务器到云服务器组

为提升业务可靠性，您可以添加弹性云服务器到云服务器组，添加后，该弹性云服务器与云服务器组中的其他云服务器分散地创建在不同主机上。

说明

- 添加云服务器到云服务器组后，会重新分配该云服务器所在的主机，使其与云服务器组中的其他云服务器分散地创建在不同主机上。那么当弹性云服务器再次开机时，可能会出现由于资源不足引起的启动失败，请将云服务器移出云服务器组后重新启动。
- 弹性云服务器包含本地盘（如磁盘增强型、H2型、P1型、P2型）、NVMe SSD本地磁盘（如超高I/O型）、GPU卡（如G3型）、FPGA卡（如fp1型、fp1c型），则无法在创建后加入云服务器组，如需使用云服务器组功能，请在创建时选择云服务器组。



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“弹性云服务器 > 云服务器组”。
5. 单击“操作”列下的“添加云服务器”。
6. 在“添加云服务器”页面，选择待添加的弹性云服务器。
7. （可选）根据界面提示，为弹性云服务器执行关机操作。

若“运行中”状态的云服务器添加至云服务器组失败，可执行本操作关机。

- a. 单击“操作”列的“关机”。
 - b. 选择“关机方式”。
 - 关机：正常关机。
 - 强制关机：该操作会导致云服务器中未保存的数据丢失，请谨慎操作。
 - c. 单击“是”，完成关机操作。
8. 单击“确定”，将弹性云服务器加入云服务器组。

从云服务器组移出云服务器

弹性云服务器从云服务器组移出后，与云服务器组中的其他云服务器不再遵从云服务器组策略。


1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“弹性云服务器 > 云服务器组”。
5. 展开云服务器组，查看云服务器组中包含的弹性云服务器。
6. 单击目标云服务器“操作”列下的“移出云服务器组”。
7. 单击“是”。

将弹性云服务器从云服务器组中移出。

删除云服务器组

删除云服务器组会解除组内云服务器的策略约束。

1. 登录管理控制台。

2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“弹性云服务器 > 云服务器组”。
5. 单击待删除云服务器组“操作”列下的“删除”。
6. 单击“是”。

2.7.7 配置同一 VPC 内 ECS 通过主机名进行通信

对于同一VPC内的弹性云服务器，可以通过主机名称进行通信。此时，您需要配置主机名与IP地址之间的映射关系。较之通过IP地址进行通信，主机名方式进行通信更为方便。

约束限制

仅适用于Linux弹性云服务器。

操作步骤

假设VPC内共有2台弹性云服务器：ecs-01和ecs-02。通过如下操作，ecs-01和ecs-02可以通过主机名互相通信。

步骤1 分别登录ecs-01和ecs-02，获取2台弹性云服务器的私有IP地址。

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 弹性云服务器”。
3. 在弹性云服务器列表页，通过“私有IP地址”栏，查看弹性云服务器的私有IP地址。

【示例】假设查询的私有IP地址如下：

ecs-01: 192.168.0.1

ecs-02: 192.168.0.2

步骤2 分别获取2台弹性云服务器的主机名。

1. 登录弹性云服务器。
2. 执行以下命令，查询弹性云服务器的主机名。

```
sudo hostname
```

【示例】假设查询的主机名如下：

ecs-01: hostname01

ecs-02: hostname02

步骤3 建立主机名与IP地址之间的映射关系，并添加同一VPC内其他弹性云服务器的信息。

1. 登录弹性云服务器ecs-01。
2. 执行以下命令，切换至root权限。

```
sudo su -
```

3. 执行以下命令，编辑hosts文件。

```
vi /etc/hosts
```

- 按“i”，进入编辑模式。
- 按照如下格式添加语句，建立映射关系。
私有IP地址 主机名
【示例】需添加的语句为：
192.168.0.1 hostname01
192.168.0.2 hostname02
- 按“Esc”退出编辑模式。
- 执行以下命令，保存并退出。
:wq
- 登录ecs-02。
- 重复执行[步骤3.2](#)~[步骤3.7](#)。

步骤4 测试验证能否通过主机名正常通信。

分别登录同一VPC内的弹性云服务器，执行以下命令，使用ping命令ping添加的主机，看ping包是否能正常送达。

```
ping hostname
```

----结束

2.7.8 启、停弹性云服务器


在弹性服务器的使用过程中，您可以进行开机、关机、重启或删除/退订等操作。

- 大量弹性云服务器同时开机或关机时，会加重主机的负载。如果需要同时对大量弹性云服务器执行开机或关机操作，建议分批进行，避免对其他弹性云服务器的业务造成影响。
- 重启/关机弹性云服务器时，如果弹性云服务器长时间处于“正在重启”/“正在关机”状态，可以执行强制重启/强制关机操作。强制重启/强制关机操作会导致弹性云服务器中未保存的数据丢失，请谨慎操作。

说明

对于裸金属类型的规格实例（规格名称中包含physical），请勿通过操作系统的shutdown，poweroff，half等命令进行操作，可能会导致命令无效或关机后无法启动。


开机

- 登录管理控制台。
- 单击管理控制台左上角的，选择区域和项目。
- 选择“计算 > 弹性云服务器”。
- 在弹性云服务器列表中，勾选需要进行操作的弹性云服务器实例。
- 单击弹性云服务器列表左上角的“开机”。
- 参考界面提示信息，单击“是”，完成弹性云服务器的“开机”操作。

说明

如果弹性云服务器处于“开机中”中间状态超过30分钟，则说明出现异常，需要联系管理员处理。

关机

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中，勾选需要进行操作的弹性云服务器实例。
5. 单击弹性云服务器列表左上角的“关机”。
6. 在“关机”页面，您可以根据需要选择“关机方式”。

须知


“强制关机”方式会导致云服务器中未保存的数据丢失，请谨慎操作。

7. 单击“是”，完成弹性云服务器的“关机”操作。

说明

如果弹性云服务器处于“关机中”中间状态超过30分钟，则说明出现异常，需要联系管理员处理。

重启

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中，勾选需要进行操作的弹性云服务器实例。
5. 单击弹性云服务器列表左上角的“重启”。

须知


勾选“强制重启”会导致云服务器中未保存的数据丢失，请谨慎操作。

6. 单击“是”，完成弹性云服务器的“重启”操作。

说明

如果弹性云服务器处于“重启中”或“强制重启中”中间状态超过30分钟，则说明出现异常，需要联系管理员处理。

删除/退订

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中，勾选需要进行操作的弹性云服务器实例。
5. 单击弹性云服务器列表左上角的“更多 > 删除”或“更多 > 退订”。

- 删除按需计费的弹性云服务器。
 - i. 勾选“删除云服务器绑定的弹性公网IP地址”和“删除云服务器挂载的数据盘”，同步删除弹性云服务器绑定的资源。
 - ii. 单击“下一步”，确认删除的资源明细。
 - iii. 单击“确定”，完成弹性云服务器的“删除”操作。
- 退订包年/包月的弹性云服务器。
 - i. 单击“是”，跳转至费用中心的“退订资源”页面。
 - ii. 根据界面提示，勾选需要退订的资源，并选择退订原因。
 - iii. 确认退订信息无误后，勾选“资源退订后……”提示信息。
 - iv. 单击“退订”，再次根据界面信息确认要退订的资源。
 - v. 再次单击“退订”，完成包年/包月资源的退订操作。

📖 说明

如果弹性云服务器处于“正在删除”中间状态超过30分钟，则说明出现异常，需要联系管理员处理。

2.8 变更 ECS 规格（vCPU 和内存）

2.8.1 变更单台 ECS 规格

操作场景

当您购买的弹性云服务器规格无法满足业务需要时，可参考本章节变更规格，升级vCPU、内存。对于部分类型的弹性云服务器，您还可以在变更规格时，更换弹性云服务器的类型。

- “XEN实例”变更为“KVM实例”，需先手动配置弹性云服务器，安装对应的驱动，然后再变更规格。否则，规格变更后的弹性云服务器不可用（操作系统可能无法启动）。“XEN实例”变更为“KVM实例”操作指引如下。Linux操作系统优先推荐使用自动配置的方法变更规格。
 - [XEN实例变更为KVM实例（Windows）](#)
 - [XEN实例变更为KVM实例（Linux-自动配置）](#)
 - [XEN实例变更为KVM实例（Linux-手动配置）](#)

📖 说明

- 弹性云服务器的实例包括：
 - XEN实例：S1、C1、C2、M1型弹性云服务器。
 - KVM实例：参考[规格清单](#)，查询对应规格的虚拟化类型。

变更须知

- 变更规格过程中，弹性云服务器需要关机，建议您选择业务空闲时间进行操作。
- 变更规格过程中，不要对弹性云服务器做任何操作，例如关机、重启等，否则会导致变更规格失败。
- 变更弹性云服务器规格时，用户不能选择已售罄的CPU和内存资源。

- 弹性云服务器规格（CPU或内存）变小，会影响弹性云服务器的性能。
- 对于部分类型的弹性云服务器，暂不支持规格变更操作。[实例类型](#)章节列举了提供的云服务器类型，并对每种云服务器类型的功能、使用进行介绍，具体请以各类型云服务器的“使用须知”为准。
- 当云硬盘状态为“正在扩容”时，不支持变更所挂载的弹性云服务器规格。
- Windows操作系统的弹性云服务器，执行变更规格操作前建议您参考[磁盘脱机怎么办？](#)修改Windows操作系统SAN策略，避免变更规格后磁盘处于脱机状态。
- 使用了收费镜像的包年/包月模式的弹性云服务器，无法进行规格降配，即不支持变更到费用更低的规格。

费用说明

变更规格会引起费用的变化，具体费用说明请参见[变更资源费用说明](#)。

变更前准备

变更规格后，可能会出现网卡漂移现象，请在变更规格前做如下操作：

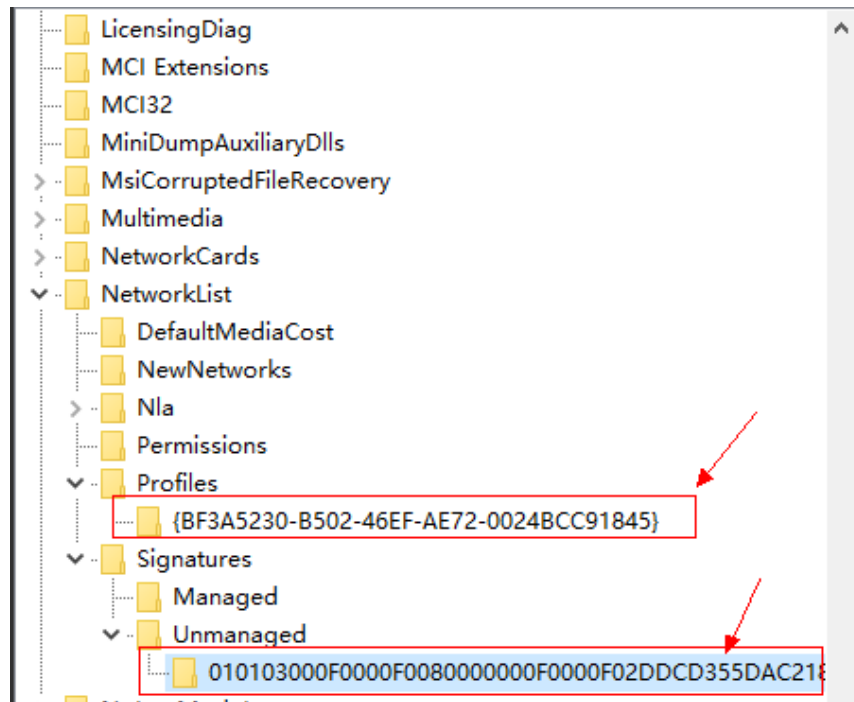
说明

出现网卡漂移的现象，是由于创建云服务器使用的镜像的操作系统中打开了网卡保留规则，才导致该云服务器变更规格后出现网卡漂移。

关于网卡漂移的更多信息，请参见[弹性云服务器变更规格后网卡漂移怎么办？](#)。

- Linux系统：
在云服务器中执行以下命令，删除网络规则目录下，文件名同时包含**persistent**和**net**的规则文件
rm -fr /etc/udev/rules.d/*net*persistent*.rules
rm -fr /etc/udev/rules.d/*persistent*net*.rules
- Windows系统：
在云服务器中删除如下注册表下的目录。
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion
\NetworkList\Profiles
HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows NT\CurrentVersion
\NetworkList\Signatures\Unmanaged

图 2-131 注册表



步骤 1: 变更规格


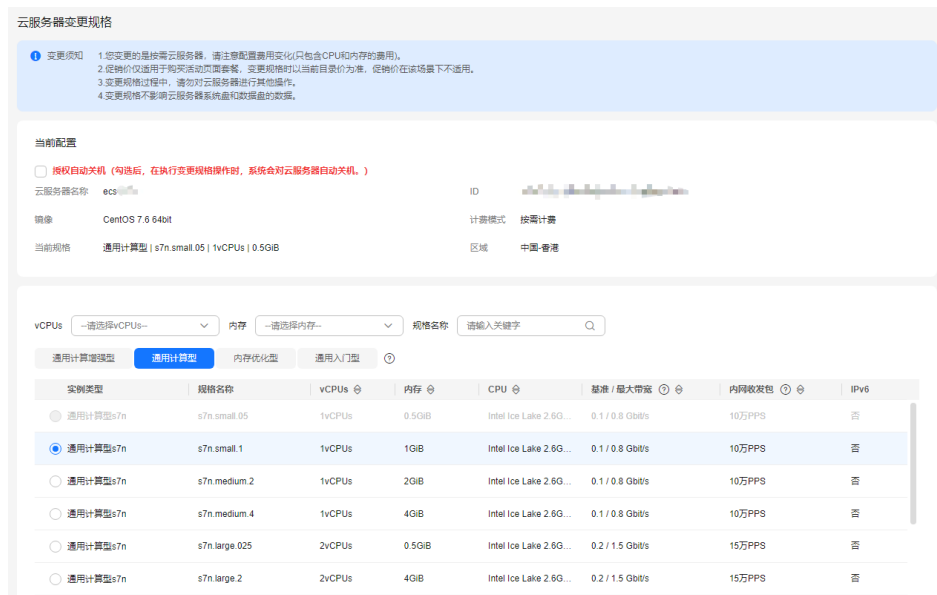
1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 单击“操作”列下的“更多 > 变更规格”。
系统进入“云服务器变更规格”页面。
5. 根据界面提示，选择变更后的弹性云服务器类型、vCPU和内存。
变更规格前请先将云服务器关机，或根据页面提示勾选“授权自动关机”。

图 2-132 变更规格



6. 单击“下一步”。
7. 确认变更后的配置无误后，阅读并勾选同意服务协议，单击“提交申请”。
8. 查询规格是否变更成功。

申请变更规格后，可以通过“异常任务”栏查看规格是否变更成功。

- a. 查看控制台是否显示“异常任务”栏，查看操作请参见[查看失败任务](#)。
 - 是，执行**8.b**。
 - 否，变更规格成功。
- b. 打开“异常任务”，根据弹性云服务器的“名称/ID”、“操作时间”和“任务”，确认列表中是否有刚刚执行的变更规格任务。
 - 是，变更规格失败，失败原因请参见[后续处理](#)。
 - 否，变更规格成功。

步骤 2：检查磁盘挂载状态

变更规格时，可能会发生磁盘挂载失败的情况，因此，变更规格后，需检查磁盘挂载状态是否正常。如果正常，则变更成功。

- Windows弹性云服务器
详细操作请参考[Windows云服务器变更规格后数据盘脱机怎么办?](#)
- Linux弹性云服务器
详细操作请参考[Linux云服务器变更规格后数据盘脱机怎么办?](#)

后续处理

如果变更规格失败，请到云审计页面查看失败原因。具体操作如下：

1. 登录管理控制台。

2. 选择“管理与监管 > 云审计服务”。
3. 在左侧导航栏，选择“云审计 > 事件列表”。
4. 在“事件名称”栏，根据“资源ID”查找到名称为“resizeServer”的规格变更失败事件。
其中，“资源ID”为规格变更失败的弹性云服务器ID。
5. 单击“操作”列下的“查看事件”，查看失败原因。
如果无法根据日志解决问题，可联系客服。

2.8.2 XEN 实例变更为 KVM 实例（Windows）

操作场景

XEN实例变更为KVM实例前，需要确保Windows弹性云服务器已安装了PV driver和UVP VMTools。

本节指导您安装PV driver和UVP VMTools，将XEN实例变更为KVM实例。

说明

- 弹性云服务器的实例包括：
 - XEN实例：S1、C1、C2、M1型弹性云服务器。
 - KVM实例：参考[规格清单](#)，查询对应规格的虚拟化类型。

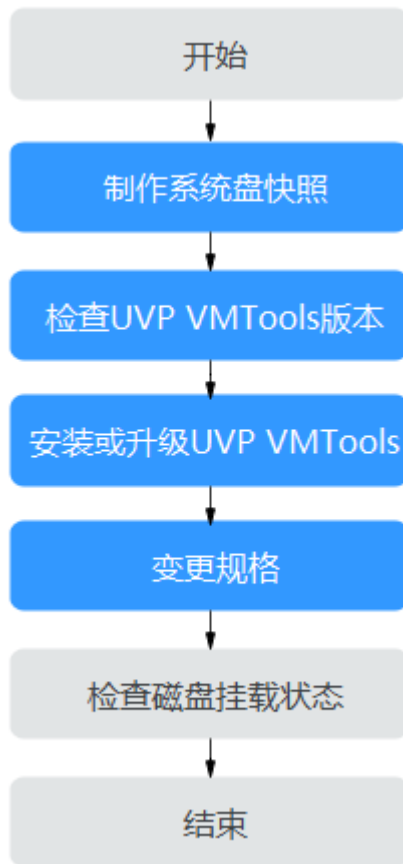
约束与限制

- 变更规格过程中，弹性云服务器需要关机，建议您选择业务空闲时间进行操作。
- Windows操作系统云服务器如果存在跨区卷，不支持变更规格，否则可能会导致数据丢失。
- 对于XEN实例，当挂载的VBD磁盘超过24块时，不支持将规格变更为KVM实例。
- 系统支持将“XEN实例”变更为“KVM实例”，不支持将“KVM实例”变更为“XEN实例”。

操作流程

XEN实例变更为KVM实例的操作流程如[图2-133](#)所示。

图 2-133 Windows 云服务器变更流程



具体的变更操作如表2-46所示。

表 2-46 XEN 实例变更为 KVM 实例

序号	任务
步骤1	步骤1：制作系统盘快照
步骤2	步骤2：检查UVP VMTools版本
步骤3	步骤3：安装或升级UVP VMTools
步骤4	步骤4：变更规格
步骤5	(可选) 步骤5：检查磁盘挂载状态

步骤 1：制作系统盘快照

如果云服务器未安装驱动就执行变更规格的操作，云服务器可能无法正常使用，造成您的系统盘数据丢失。因此，建议您先制作系统盘快照，防止数据丢失。

1. 制作系统盘快照前请对云服务器完成自检。

对云服务器执行关机、开机操作，确保云服务器重启后业务可以正常运行。再启动制作系统盘快照。

2. 制作系统盘快照的操作，请参见《云硬盘用户指南》的“用户指南 > [创建快照](#)”章节。

📖 说明

变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。

步骤 2：检查 UVP VMTools 版本

变更规格前请先检查 UVP VMTools 版本。

1. 登录弹性云服务器。
2. 下载驱动检查脚本

下载驱动检查脚本，用管理员权限执行脚本，等待检查结果。

下载地址：https://latin-server-resize.obs.na-mexico-1.myhuaweicloud.com/windows/server_resize/check_kvm_drivers.vbs

驱动检查脚本在确认驱动安装成功后会自动为云服务器设置标签，标记驱动安装成功，没有设置标签的云服务器无法执行变更规格的操作。

- 检查结果为“Check version success!”说明驱动版本符合要求，且为云服务器设置标签标记驱动安装成功。您可以执行[步骤4：变更规格](#)。
- 检查结果为“Check version success but set metadata failed! Please run this script again later.”说明驱动版本符合要求，但为云服务器设置标签失败，请稍后重试。
- 检查结果为“Check version failed! Please install drivers at first.”说明驱动版本不符合要求，请参见[步骤3：安装或升级 UVP VMTools](#)安装或升级 UVP VMTools。

步骤 3：安装或升级 UVP VMTools

安装或升级 UVP VMTools 时，如果云服务器中已安装 PV Driver，会对其版本进行校验，为了避免在云服务器上安装 UVP VMTools 失败，需确保 PV Driver 版本满足要求。本小节介绍检查安装 PV Driver、UVP VMTools 的操作步骤。

⚠️ 注意

安装 PV Driver 或 UVP VMTools 前请确保云服务器满足以下条件：

- 云服务器的系统盘的剩余空间必须大于 2GB。
- 为了避免在云服务器上安装驱动失败，安装前请先卸载第三方虚拟化平台的工具（例如：Citrix Xen Tools、VMware Tools），相关卸载方法请参考对应的工具的官方文档。
- 禁用防病毒软件或入侵检测软件，您可以在驱动安装完成后，再启用这些软件。

1. 安装前请先检查 PV Driver 版本是否满足 UVP VMTools 依赖。

进入目录“C:\Program Files (x86)\Xen PV Drivers\bin”，打开文件“version.ini”，查看 PV driver 版本号。

```
pvdriverVersion=5.0.104.010
```

- 如果存在目录，且驱动版本高于 5.0，表示 PV driver 已安装且满足版本要求，请执行[6](#)安装 UVP VMTools。

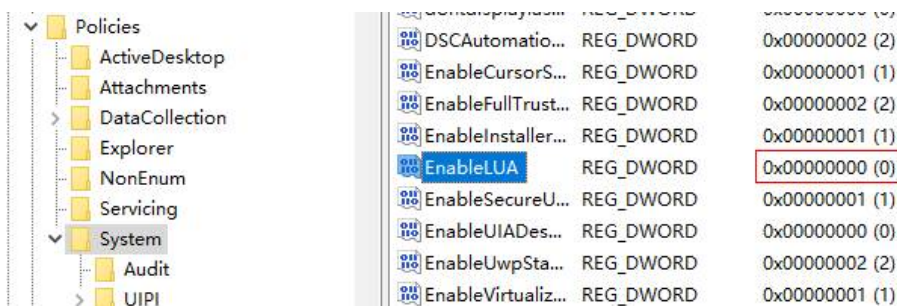
- 反之，表示PV driver未正常安装或版本不满足要求，请执行以下步骤卸载旧版本PV driver，安装新版本的PV driver。
2. 记录当前云服务器的UAC (User Account Control) 配置。

说明

PV Driver版本低于5.0时，为实现静默升级功能（防止升级过程中弹窗过多），安装PV Driver时在注册表中加入了Disable LUA操作，卸载PV Driver时在注册表中加入了Enable LUA操作（5.0及以上版本已优化解决此问题）。为避免对您的业务造成影响，如果PV Driver版本低于5.0，请在卸载前先记录UAC配置，并在安装新版本PV Driver后检查并恢复注册表中Enable LUA配置。了解更多UAC配置请参考[微软官方文档](#)。

- a. 打开“运行”窗口，输入“regedit”，打开“注册表编辑器”。
- b. 记录HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Policies\System\EnableLUA 键值。

图 2-134 EnableLUA



3. 卸载旧版本PV driver。
 - a. 在云服务器云主机操作系统界面，选择“开始 > 控制面板”。
 - b. 单击“卸载程序”。
 - c. 按照提示，卸载“GPL PV Drivers for Windows x.x.x.xx”。
 - d. 在控制台重启云服务器。
4. 安装新版本PV driver。
 - a. 下载PV Driver的安装包。



PV Driver下载地址：<https://ecs-instance-driver.obs.cn-north-1.myhuaweicloud.com/pvdriver-windows.zip>
 - b. 解压PV driver软件包。
 - c. 双击“pvdriver-windows.iso”。
 - d. 运行“Setup.exe”，根据界面提示安装PV Driver。

请耐心等待驱动安装完成，请勿重复单击“Setup.exe”。
 - e. 根据提示重启云服务器，使PV driver生效。
5. 检查并恢复UAC配置。
 - a. 打开“运行”窗口，输入“regedit”，打开“注册表编辑器”。
 - b. 查看HKEY_LOCAL_MACHINE\SOFTWARE\Microsoft\Windows\CurrentVersion\Policies\System\EnableLUA 键值，与卸载PV Driver前记录的键值比较，如果不同请将键值修改为2中记录的值。
6. 安装或升级UVP VMTools。
 - a. 下载UVP VMTools的安装包。

UVP VMTools下载地址：<https://ecs-instance-driver.obs.cn-north-1.myhuaweicloud.com/vmtools-windows.zip>

- b. 解压UVP VMTools的安装包。
- c. 双击“vmtools-windows.iso”。
- d. 运行“Setup.exe”，根据界面提示安装UVP VMTools。
安装程序会自动匹配当前操作系统版本，并识别新安装或升级场景。
请耐心等待驱动安装完成，请勿重复单击“Setup.exe”
- e. 根据提示重启云服务器，使UVP VMTools驱动生效。
- f. 驱动安装完成并重启云服务器后，需检查驱动是否安装成功。检查方法请参考**步骤2：检查UVP VMTools版本**。

步骤 4：变更规格

1. 登录控制台。
2. 单击控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表，查询待变更弹性云服务器状态。
如果不是关机状态，单击“操作”列下的“更多 > 关机”。
5. 单击“操作”列下的“更多 > 变更规格”。
系统进入“云服务器变更规格”页面。
6. 根据界面提示，选择变更后的云服务器类型、vCPU和内存。
7. （可选）选择“专属主机”。
对于在专属主机上创建的弹性云服务器，系统支持更换云服务器所在的专属主机。
此时，您可以单击下拉列表，选择更换专属主机。如果下拉列表中无可用的专属主机，说明专属主机所剩资源不足，不能用于创建变更规格后的弹性云服务器。
8. 勾选复选框“我确认已完成对弹性云服务器的配置”，确认已完成**步骤3：安装或升级UVP VMTools**。
9. 单击“确定”。

说明

- 单击“确定”后云平台自动为您制作系统盘快照。变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。
- 如果变更规格失败后，弹性云服务器无法使用，可能会需要重装操作系统来恢复云服务器，请注意重装操作系统会清除系统盘数据，但不影响数据盘的数据。

（可选）步骤 5：检查磁盘挂载状态

XEN实例变更为KVM实例时，可能会发生磁盘脱机，因此，变更规格后，需检查磁盘挂载状态是否正常。如果正常，则变更成功。

- Windows弹性云服务器
详细操作请参考**Windows云服务器变更规格后数据盘脱机怎么办？**

后续处理

如果控制台上云服务器列表页，显示弹性云服务器已变更规格成功，但是远程登录云服务器后，操作系统无法启动，此时，请联系客服进行恢复，或重装操作系统进行恢复。重装操作系统的操作指导，请参见[重装操作系统](#)。

📖 说明

重装操作系统会清除系统盘数据，但不影响数据盘的数据。

变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。

2.8.3 XEN 实例变更为 KVM 实例（Linux-自动配置）

操作场景

Linux操作系统XEN实例变更为KVM实例前，必须已完成必要的驱动安装和配置。

本节操作指导您使用自动化脚本的方式为Linux云服务器安装驱动、配置磁盘自动挂载等，并将XEN实例变更为KVM实例。

📖 说明

- XEN实例：S1、C1、C2、M1型弹性云服务器。
- KVM实例：参考[规格清单](#)，查询对应规格的虚拟化类型。
- 为了同时支持XEN虚拟化和KVM虚拟化，Linux弹性云服务器的正常运行需依赖于xen-pv驱动、virtio驱动等。XEN实例变更为KVM实例前，需要确保Linux弹性云服务器已完成相关配置，包括安装驱动、配置磁盘自动挂载等。

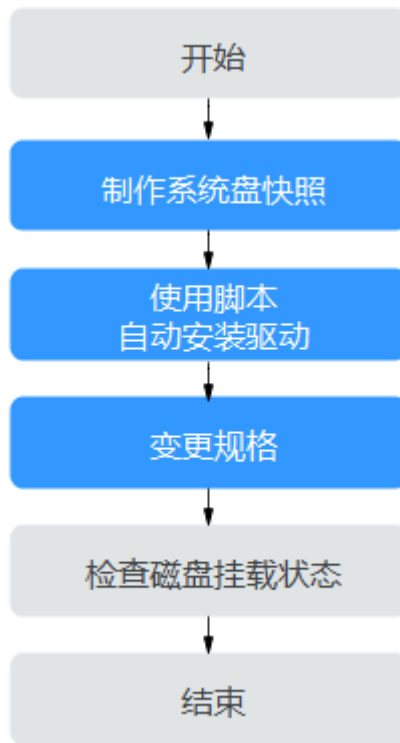
约束与限制

- 变更规格过程中，弹性云服务器需要关机，建议您选择业务空闲时间进行操作。
- Linux操作系统云服务器如果存在由多个物理卷组成的LVM逻辑卷或组建了RAID磁盘阵列，均不支持变更规格，否则可能会导致数据丢失。
- 对于XEN实例，当挂载的VBD磁盘超过24块时，不支持将规格变更为KVM实例。
- 系统支持将“XEN实例”变更为“KVM实例”，不支持将“KVM实例”变更为“XEN实例”。

操作流程

XEN实例变更为KVM实例的操作流程如[图2-135](#)所示。

图 2-135 Linux 云服务器变更流程（自动配置）



具体的变更操作如表2-47所示。

表 2-47 XEN 实例变更为 KVM 实例（自动配置）

序号	任务
步骤1	步骤1：制作系统盘快照
步骤2	步骤2：使用脚本自动安装驱动
步骤3	步骤3：变更规格
步骤4	（可选）步骤4：检查磁盘挂载状态

步骤 1：制作系统盘快照

如果云服务器未安装驱动就执行变更规格的操作，云服务器可能无法正常使用，造成您的系统盘数据丢失。因此，建议您先制作系统盘快照，防止数据丢失。

1. 制作系统盘快照前请对云服务器完成自检。
对云服务器执行关机、开机操作，确保云服务器重启后业务可以正常运行。再启动制作系统盘快照。
2. 制作系统盘快照的操作，请参见《云硬盘用户指南》的“用户指南 > [创建快照](#)”章节。

📖 说明

变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。

步骤 2：使用脚本自动安装驱动

对于支持使用脚本方式配置的弹性云服务器，请参见本部分内容。部分操作系统的弹性云服务器暂不支持使用自动化脚本方式进行配置，请直接参见[XEN实例变更为KVM实例（Linux-手动配置）](#)。

1. 登录弹性云服务器。
2. 执行以下命令，将驱动安装脚本下载到root目录下。

```
curl URL > ~/resize_ecs_modify_linux.sh
```

其中，URL为规格变更优化脚本的下载地址。

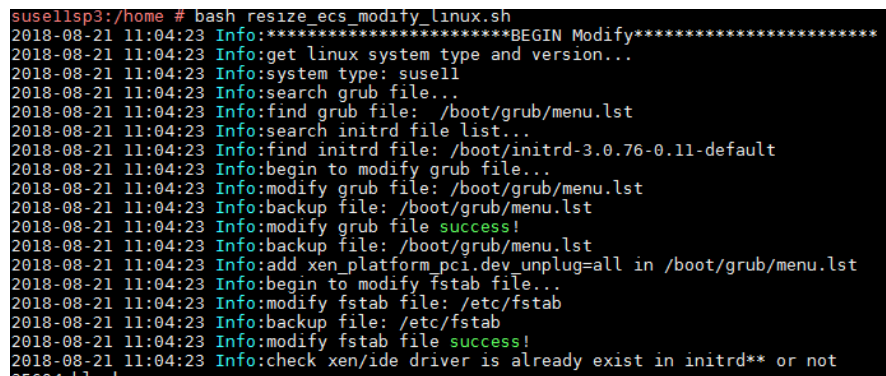
请根据云服务器所在区域选择优化脚本的下载地址：

下载地址：https://latin-server-resize.obs.na-mexico-1.myhuaweicloud.com/linux/server_resize/resize_ecs_modify_linux.sh

3. 执行以下命令，运行脚本，该脚本会自动检查并安装原生xen-pv驱动、virtio驱动。

```
bash resize_ecs_modify_linux.sh
```

图 2-136 运行脚本



```
suselisp3:/home # bash resize_ecs_modify_linux.sh
2018-08-21 11:04:23 Info:*****BEGIN Modify*****
2018-08-21 11:04:23 Info:get linux system type and version...
2018-08-21 11:04:23 Info:system type: suse11
2018-08-21 11:04:23 Info:search grub file...
2018-08-21 11:04:23 Info:find grub file: /boot/grub/menu.lst
2018-08-21 11:04:23 Info:search initrd file list...
2018-08-21 11:04:23 Info:find initrd file: /boot/initrd-3.0.76-0.11-default
2018-08-21 11:04:23 Info:begin to modify grub file...
2018-08-21 11:04:23 Info:modify grub file: /boot/grub/menu.lst
2018-08-21 11:04:23 Info:backup file: /boot/grub/menu.lst
2018-08-21 11:04:23 Info:modify grub file success!
2018-08-21 11:04:23 Info:backup file: /boot/grub/menu.lst
2018-08-21 11:04:23 Info:add xen_platform_pci.dev_unplug=all in /boot/grub/menu.lst
2018-08-21 11:04:23 Info:begin to modify fstab file...
2018-08-21 11:04:23 Info:modify fstab file: /etc/fstab
2018-08-21 11:04:23 Info:backup file: /etc/fstab
2018-08-21 11:04:23 Info:modify fstab file success!
2018-08-21 11:04:23 Info:check xen/ide driver is already exist in initrd** or not
```

4. 请耐心等待脚本运行结束。

该脚本在确认驱动安装成功后会自动为云服务器设置标签，标记驱动安装成功，没有设置标签的云服务器无法执行变更规格的操作。

如果回显提示“{镜像名称} already contain xen and virtio driver”，表示检查并安装驱动成功。

- 回显提示“Success to set kvm meta!”或“this server already has kvm meta.”，表示设置标签成功，请继续执行[步骤3：变更规格](#)。
- 回显提示“Failed to set metadata, please try again”，表示设置标签失败，可稍后重试执行脚本。

如果安装失败请参考[XEN实例变更为KVM实例（Linux-手动配置）](#)手动配置或者联系客服寻求技术支持。


图 2-137 运行成功

```
161.548762] device-mapper: uevent: version 1.0.3
161.551753] device-mapper: ioctl: 4.37.1-ioctl (2010-04-03) initialised: dm-devel@redhat.com
generating grub configuration file ...
found linux image: /boot/vmlinuz-3.10.0-1062.12.1.el7.x86_64
found initrd image: /boot/initramfs-3.10.0-1062.12.1.el7.x86_64.img
found linux image: /boot/vmlinuz-3.10.0-957.el7.x86_64
found initrd image: /boot/initramfs-3.10.0-957.el7.x86_64.img
found linux image: /boot/vmlinuz-0-rescue-8f5b018f6eb344909f6cfc5ad0839ef
found initrd image: /boot/initramfs-0-rescue-8f5b018f6eb344909f6cfc5ad0839ef.img
162.148361] SGX: XFS with ACLs, security attributes, no debug enabled
162.189514] xor: automatically using best checksumming function:
162.202066]   avx      : 22448.000 MB/sec
162.233066] raid6: sse2x1   gen() 7382 MB/s
162.277075] raid6: sse2x2   gen() 8589 MB/s
162.297084] raid6: sse2x4   gen() 10273 MB/s
162.318073] raid6: avx2x1   gen() 13410 MB/s
162.337070] raid6: avx2x2   gen() 16503 MB/s
162.356066] raid6: avx2x4   gen() 18976 MB/s
162.358393] raid6: using algorithm avx2x4 gen() (18976 MB/s)
162.361600] raid6: using avx2x2 recovery algorithm
162.431572] Btrfs loaded, crc32c=crc32c-intel
162.446525] fuse init (API version 7.23)
done
2020-09-24 15:12:13 Info:check xen/ide driver is already exist in /boot/initramfs-0-rescue-8f5b018f6eb344909f6cfc5ad0839ef.img
or not
2020-09-24 15:12:23 Info:xen driver:yes
2020-09-24 15:12:23 Info:ide driver:no
2020-09-24 15:12:23 Info:check virtio driver is already exist in /boot/initramfs-0-rescue-8f5b018f6eb344909f6cfc5ad0839ef.img
or not
2020-09-24 15:12:41 Info:virtio driver:yes
2020-09-24 15:12:41 Info:check xen/ide driver is already exist in /boot/initramfs-3.10.0-1062.12.1.el7.x86_64.img or not
2020-09-24 15:12:45 Info:xen driver:yes
2020-09-24 15:12:45 Info:ide driver:no
2020-09-24 15:12:45 Info:check virtio driver is already exist in /boot/initramfs-3.10.0-1062.12.1.el7.x86_64.img or not
2020-09-24 15:12:52 Info:virtio driver:yes
2020-09-24 15:12:52 Info:check xen/ide driver is already exist in /boot/initramfs-3.10.0-957.el7.x86_64.img or not
2020-09-24 15:12:55 Info:xen driver:yes
2020-09-24 15:12:55 Info:ide driver:no
2020-09-24 15:12:55 Info:check virtio driver is already exist in /boot/initramfs-3.10.0-957.el7.x86_64.img or not
2020-09-24 15:13:03 Info:virtio driver:yes
2020-09-24 15:13:03 Info:centos7 already contain xen and virtio driver
% Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left   Speed
  00    30      0     0  100    30      0     0    0     0      0     0      0     0      0     0
2020-09-24 15:13:03 Info:Success to set ksm meta!
```

说明

- 请务必确保云服务器配置成功，否则，可能会导致变更规格后的弹性云服务器不可用。如果运行失败或者，请参考[XEN实例变更为KVM实例（Linux-手动配置）](#)手动配置。
- 脚本安装失败常见问题请参考：
 - [CentOS 5操作系统云服务器执行驱动安装脚本失败怎么办？](#)
 - [Linux云服务器变更规格时执行驱动安装脚本失败怎么办？](#)

步骤 3：变更规格

- 登录控制台。
- 单击管理控制台左上角的 ，选择区域和项目。
- 单击“☰”，选择“计算 > 弹性云服务器”。
- 在弹性云服务器列表，查询待变更弹性云服务器状态。
如果不是关机状态，单击“操作”列下的“更多 > 关机”。
- 单击“操作”列下的“更多 > 变更规格”。
系统进入“云服务器变更规格”页面。
- 根据界面提示，选择变更后的云服务器类型、vCPU和内存。
- （可选）选择“专属主机”。

对于在专属主机上创建的弹性云服务器，系统支持更换云服务器所在的专属主机。

此时，您可以单击下拉列表，选择更换专属主机。如果下拉列表中无可用的专属主机，说明专属主机所剩资源不足，不能用于创建变更规格后的弹性云服务器。

- 勾选复选框“我确认已完成对弹性云服务器的配置”，确认已完成“配置弹性云服务器”操作。

9. 单击“确定”。

📖 说明

- 单击“确定”后云平台自动为您制作系统盘快照。变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。
- 如果变更规格失败后，弹性云服务器无法使用，可能会需要重装操作系统来恢复云服务器，请注意重装操作系统会清除系统盘数据，但不影响数据盘的数据。

（可选）步骤 4：检查磁盘挂载状态

XEN实例变更为KVM实例时，可能会发生磁盘挂载失败的情况，因此，变更规格后，需检查磁盘挂载状态是否正常。如果正常，则变更成功。

- Linux弹性云服务器
详细操作请参考[Linux云服务器变更规格后数据盘脱机怎么办？](#)

后续处理

如果控制台上云服务器列表页，显示弹性云服务器已变更规格成功，但是远程登录云服务器后，操作系统无法启动，此时，请联系客服进行恢复，或重装操作系统进行恢复。重装操作系统的操作指导，请参见[重装操作系统](#)。

📖 说明

重装操作系统会清除系统盘数据，但不影响数据盘的数据。

变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。

2.8.4 XEN 实例变更为 KVM 实例（Linux-手动配置）

操作场景

Linux操作系统XEN实例变更为KVM实例前，必须完成驱动的安装和配置。

本节操作指导您手动安装Linux云服务器驱动、配置磁盘自动挂载等，并将XEN实例变更为KVM实例。

如需使用自动化脚本安装驱动的方法请参考[XEN实例变更为KVM实例（Linux-自动配置）](#)。

📖 说明

- XEN实例：S1、C1、C2、M1型弹性云服务器。
- KVM实例：参考[规格清单](#)，查询对应规格的虚拟化类型。
- 为了同时支持XEN虚拟化和KVM虚拟化，Linux弹性云服务器的正常运行需依赖于xen-pv驱动、virtio驱动等。XEN实例变更为KVM实例前，需要确保Linux弹性云服务器已完成相关配置，包括安装驱动、配置磁盘自动挂载等。

约束与限制

- 变更规格过程中，弹性云服务器需要关机，建议您选择业务空闲时间进行操作。
- Linux操作系统云服务器如果存在由多个物理卷组成的LVM逻辑卷或组建了RAID磁盘阵列，均不支持变更规格，否则可能会导致数据丢失。

- 对于XEN实例，当挂载的VBD磁盘超过24块时，不支持将规格变更为KVM实例。
- 系统支持将“XEN实例”变更为“KVM实例”，不支持将“KVM实例”变更为“XEN实例”。

操作流程

XEN实例变更为KVM实例的操作流程如图2-138所示。

图 2-138 Linux 云服务器变更流程（手动配置）

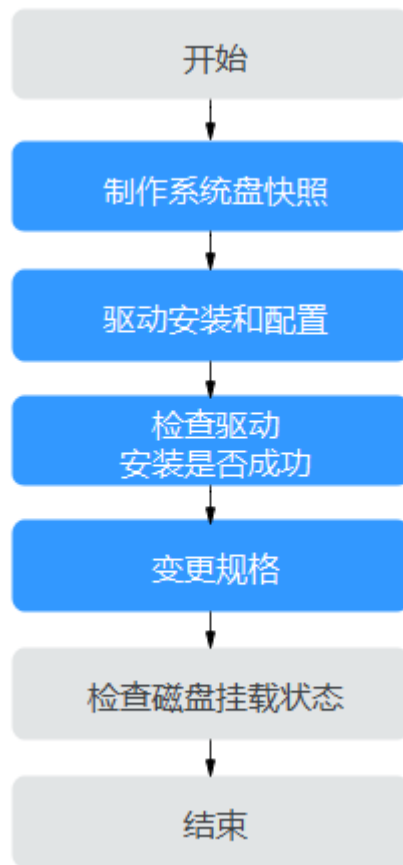


表 2-48 XEN 实例变更为 KVM 实例（手动配置）

序号	任务
步骤1	步骤1：制作系统盘快照
步骤2	步骤2：安装驱动
步骤3	步骤3：检查云服务器配置是否成功
步骤4	步骤4：变更规格
步骤5	（可选）步骤5：检查磁盘挂载状态

步骤 1：制作系统盘快照

如果云服务器未安装驱动就执行变更规格的操作，云服务器可能无法正常使用，造成您的系统盘数据丢失。因此，建议您先制作系统盘快照，防止数据丢失。

1. 制作系统盘快照前请对云服务器完成自检。
对云服务器执行关机、开机操作，确保云服务器重启后业务可以正常运行。再启动制作系统盘快照。
2. 制作系统盘快照的操作，请参见《云硬盘用户指南》的“用户指南 > [创建快照](#)”章节。

说明

变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。

步骤 2：安装驱动

对于不支持使用脚本方式配置的弹性云服务器，请参见本部分内容，手动配置云服务器。

1. 登录弹性云服务器。
2. 卸载云服务器中已安装的Tools。
卸载操作请参见[在Linux系统中卸载Tools](#)。
3. 修改grub的UUID。
具体操作请参见[修改grub文件磁盘标识方式为UUID](#)。
4. 修改fstab的UUID。
具体操作请参见[修改fstab文件磁盘标识方式为UUID](#)。
5. 安装原生的XEN和KVM驱动。
安装方法请参见[安装原生的XEN和KVM驱动](#)。

步骤 3：检查云服务器配置是否成功

根据[步骤2：安装驱动](#)修改grub文件、安装所需驱动后，需检查修改是否生效、驱动是否安装成功。

说明

手动方式配置弹性云服务器时请务必执行以下操作检查云服务器配置是否成功。

1. 登录弹性云服务器。
2. 执行以下命令，检查root分区是否以UUID的形式表示。

```
cat /boot/grub/grub.cfg
```

- 是，表示已修改grub文件的磁盘标识方式为UUID。
- 否，修改失败，请参见[步骤2：安装驱动](#)重新修改。

```
.....menuentry 'Ubuntu Linux, with Linux 3.13.0-24-generic' --class ubuntu --class gnu-linux --class gnu --class os --unrestricted $menuentry_id_option 'gnulinux-3.13.0-24-generic-advanced-ec51d860-34bf-4374-ad46-a0c3e337fd34' {
    recordfail
    load_video
    gfxmode $linux_gfx_mode
    insmod gzio
    insmod part_msdos
    insmod ext2
```

```
if [ x$feature_platform_search_hint = xy ]; then
search --no-floppy --fs-uuid --set=root ec51d860-34bf-4374-ad46-a0c3e337fd34
else
search --no-floppy --fs-uuid --set=root ec51d860-34bf-4374-ad46-a0c3e337fd34
fi
echo 'Loading Linux 3.13.0-24-generic ...'
linux /boot/vmlinuz-3.13.0-24-generic root=UUID=ec51d860-34bf-4374-ad46-a0c3e337fd34 ro
echo 'Loading initial ramdisk ...'
initrd /boot/initrd.img-3.13.0-24-generic
}
```

📖 说明

对于不同操作系统，grub文件的具体路径存在差异，请以具体的操作系统为准。例如：“/boot/grub/menu.lst”、“/boot/grub/grub.cfg”、“/boot/grub2/grub.cfg”、“/boot/grub/grub.conf”。

3. 执行以下命令，检查磁盘的标识方式是否为UUID。

cat /etc/fstab

- 是，表示已修改fstab文件的磁盘标识方式为UUID。
- 否，修改失败，请参见[步骤2：安装驱动重新修改](#)。

```
[root@***** ~]# cat /etc/fstab
UUID=4eb40294-4c6f-4384-bbb6-b8795bbb1130 / xfs defaults 0 0
UUID=2de37c6b-2648-43b4-a4f5-40162154e135 swap swap defaults 0 0
```

4. 检查是否已安装原生的XEN和KVM驱动。

- 如果引导的虚拟文件系统是initramfs，执行以下命令：

```
lsinitrd /boot/initramfs-`uname -r`.img | grep `uname -r` | grep xen
lsinitrd /boot/initramfs-`uname -r`.img | grep `uname -r` | grep virtio
```

- 如果引导的虚拟文件系统是initrd，执行如下命令：

```
lsinitrd /boot/initrd-`uname -r` | grep `uname -r` | grep xen
lsinitrd /boot/initrd-`uname -r` | grep `uname -r` | grep virtio
```

如果安装成功，回显将显示已安装的原生XEN和KVM驱动名称。



```
[root@CTU10000xxxxx home]# lsinitrd /boot/initramfs-`uname -r`.img | grep `uname -r` | grep xen
-rwxr--r-- 1 root root 54888 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/drivers/
block/xen-blkfront.ko
-rwxr--r-- 1 root root 45664 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/
drivers/net/xen-netfront.ko
```

```
[root@CTU10000xxxxx home]# lsinitrd /boot/initramfs-`uname -r`.img | grep `uname -r` | grep virtio
-rwxr--r-- 1 root root 23448 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/drivers/
block/virtio_blk.ko
-rwxr--r-- 1 root root 50704 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/
drivers/net/virtio_net.ko
-rwxr--r-- 1 root root 28424 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/drivers/
scsi/virtio_scsi.ko
drwxr-xr-x 2 root root 0 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/drivers/
virtio
-rwxr--r-- 1 root root 14544 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/drivers/
virtio/virtio.ko
-rwxr--r-- 1 root root 21040 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/drivers/
virtio/virtio_pci.ko
-rwxr--r-- 1 root root 18016 Jul 16 17:53 lib/modules/2.6.32-573.8.1.el6.x86_64/kernel/drivers/
virtio/virtio_ring.ko
```

📖 说明

请务必确保云服务器配置成功，否则，可能会导致变更规格后的弹性云服务器不可用。

步骤 4：变更规格

1. 登录控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表，查询待变更弹性云服务器状态。
如果不是关机状态，单击“操作”列下的“更多 > 关机”。
5. 单击“操作”列下的“更多 > 变更规格”。
系统进入“云服务器变更规格”页面。
6. 根据界面提示，选择变更后的云服务器类型、vCPU和内存。
7. （可选）选择“专属主机”。

对于在专属主机上创建的弹性云服务器，系统支持更换云服务器所在的专属主机。

此时，您可以单击下拉列表，选择更换专属主机。如果下拉列表中无可用的专属主机，说明专属主机所剩资源不足，不能用于创建变更规格后的弹性云服务器。

8. 勾选复选框“我确认已完成对弹性云服务器的配置”，确认已完成“配置弹性云服务器”操作。
9. 单击“确定”。

说明

- 单击“确定”后云平台自动为您制作系统盘快照。变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。
- 如果变更规格失败后，弹性云服务器无法使用，可能会需要重装操作系统来恢复云服务器，请注意重装操作系统会清除系统盘数据，但不影响数据盘的数据。

（可选）步骤 5：检查磁盘挂载状态

XEN实例变更为KVM实例时，可能会发生磁盘挂载失败的情况，因此，变更规格后，需检查磁盘挂载状态是否正常。如果正常，则变更成功。

- Linux弹性云服务器
详细操作请参考[Linux云服务器变更规格后数据盘脱机怎么办?](#)

后续处理

如果控制台上云服务器列表页，显示弹性云服务器已变更规格成功，但是远程登录云服务器后，操作系统无法启动，此时，请联系客服进行恢复，或重装操作系统进行恢复。重装操作系统的操作指导，请参见[重装操作系统](#)。

说明

重装操作系统会清除系统盘数据，但不影响数据盘的数据。

变更规格完成后，如已确认业务恢复正常，请在快照页面手动删除快照。

2.9 变更 ECS 操作系统

2.9.1 重装操作系统

操作场景

弹性云服务器操作系统无法正常启动时，或云服务器系统运行正常，但需要对系统进行优化，使其在最优状态下工作时，用户可以使用重装弹性云服务器的操作系统功能。

重装须知

- 重装操作系统后弹性云服务器IP地址和MAC地址不发生改变。
- 重装操作系统会清除系统盘数据，包括系统盘上的系统分区和所有其他分区，请做好数据备份。
- 重装操作系统不影响数据盘数据。
- 重装操作系统后的几分钟，系统正在注入密码或密钥信息，在此期间请勿对云服务器执行其他操作，避免密码或密钥信息注入失败导致云服务器无法登录。

约束与限制

- 云硬盘的配额需大于0。
- 如果是通过私有镜像创建的弹性云服务器，请确保原有镜像仍存在。
- 如果原有云服务器计费方式为按需模式，请确认账户余额充足。
- 如果原有云服务器计费方式为包年/包月模式，请确保在有效期内。
- H2型弹性云服务器不支持操作系统的重装功能。

前提条件

- 待重装操作系统的弹性云服务器挂载有系统盘。

操作步骤



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在待重装操作系统的弹性云服务器的“操作”列下，单击“更多 > 镜像 > 重装操作系统”。
重装操作系统前请先将云服务器关机，或根据页面提示勾选“立即关机（重装操作系统前需先将云服务器关机）”。
5. 设置登录方式。
如果待重装操作系统的弹性云服务器是使用密钥登录方式创建的，此时可以更换使用新密钥。

图 2-139 重装操作系统

重装操作系统

重装操作系统提供以原镜像进行系统重装的功能，不会收取额外费用。

- 1、重装系统不影响数据盘数据，但是系统盘的所有分区数据和创建的所有快照会被删除，请做好数据备份。
- 2、重装系统成功后，云服务器会自动开机；当前操作系统内的个性化设置（如DNS、主机名等）将被重置，需重新配置。

[收起说明](#)

当前配置

云服务器名称	IP地址	规格	镜像	系统盘
ecs-...	192.168.0.30 (私有IP)	1vCPUs 2GiB ...	CentOS 7.5 64bit(64-bit)	40GiB

立即关机 (重装操作系统前需先将云服务器关机)

系统盘 磁盘加密

登录凭证

密码

可使用云服务器原密码或重新设置密码，请输入并牢记密码。

确认密码

6. 单击“确定”。
7. 在“重装操作系统”页面，确认重装的操作系统规格无误后，阅读并勾选协相关协议或声明，单击“提交申请”。

提交申请后，弹性云服务器的状态变为“重装中”，当该状态消失后，表示重装结束。

说明

重装操作系统过程中，会创建一台临时弹性云服务器，重装操作系统结束后会自动删除。在重装操作系统过程中请不要对该弹性云服务器进行任何操作。

后续处理

如果操作系统重装失败，页面会提示重装操作系统失败。云服务平台支持重试功能，用户可重新执行3-7，重装弹性云服务器的操作系统。

重试后，如果仍未成功，可直接联系客服，客服会在后台进行人工恢复。

2.9.2 切换操作系统

操作场景

切换操作系统是为您的弹性云服务器重新切换一个系统盘。切换完成后弹性云服务器的系统盘ID会发生改变，并删除原有系统盘。

如果弹性云服务器当前使用的操作系统不能满足业务需求（如软件要求的操作系统版本较高），您可以选择切换弹性云服务器的操作系统。

云服务平台支持不同镜像类型（包括公共镜像、私有镜像、共享镜像）与不同操作系统之间的互相切换。您可以将现有的操作系统切换为不同镜像类型的操作系统。

约束与限制

- 切换操作系统操作预计需等待10~20分钟。切换操作系统过程中，弹性云服务器会显示任务状态为“切换操作系统中”。
- 切换操作系统完成后的几分钟，系统正在注入密码或密钥信息，在此期间请勿对云服务器执行其他操作，避免密码或密钥信息注入失败导致云服务器无法登录。
- 待切换操作系统的弹性云服务器需挂载有系统盘。
- “包年/包月”方式购买的弹性云服务器切换操作系统时，由于所选镜像不同，当前云服务器的系统盘容量可能不足，不支持切换后的镜像使用。此时，需先卸载系统盘并进行扩容，然后再重新切换操作系统。
- “包年/包月”方式购买的弹性云服务器切换操作系统时，支持如下场景：
 - 仅支持由免费的操作系统切换至免费的操作系统。
 - 使用源镜像为市场镜像的私有镜像创建而来，且购买方式为“包年/包月”的云服务器不支持切换操作系统。
 - 中国大陆外区域（包括中国港澳台及其他国家、地区）不支持Windows操作系统和Linux操作系统之间的互相切换。
- 云硬盘的配额需大于0。
- 不支持更换系统盘的磁盘类型。
- H2型弹性云服务器不支持操作系统的切换功能。
- 不同操作系统间的切换请参考[不同操作系统切换须知](#)。
- 不支持使用ISO文件创建的ISO镜像切换操作系统，必须完成操作系统和驱动安装将ISO镜像做成系统盘镜像后才能使用。
- 不支持BIOS启动方式与UEFI启动方式的操作系统互相切换。
- 不支持x86架构的弹性云服务器切换为鲲鹏架构类型的操作系统，也不支持鲲鹏架构的弹性云服务器切换到x86架构类型的操作系统。

切换须知

- 切换操作系统后，弹性云服务器将不再保留原操作系统，并删除原有系统盘。
- 切换操作系统会清除系统盘数据，包括系统盘上的系统分区和所有其他分区，请做好数据备份。详细内容，请参考[备份弹性云服务器](#)。
- 切换操作系统不影响数据盘数据。
- 切换操作系统后，您的业务运行环境需要在新的系统中重新部署。
- 切换操作系统成功后弹性云服务器会自动开机。
- 切换操作系统后不支持更换系统盘的云硬盘类型。
- 切换操作系统后弹性云服务器IP地址和MAC地址不发生改变。
- 切换操作系统后，当前操作系统内的个性化设置（如DNS、主机名等）将被重置，需重新配置。
- 切换操作预计需等待10~20分钟完成切换操作系统。切换操作系统过程中，弹性云服务器会显示任务状态为“切换操作系统中”。

不同操作系统切换须知

不同操作系统间的切换是指Windows与Linux操作系统之间的互相切换。

- Windows系统更换为Linux系统：请安装读写Windows系统的NTFS分区工具，例如NTFS-3G等。

- Linux系统更换为Windows系统：请安装可以识别ext3、ext4等分区的识别软件，例如Ext2Read、Ext2Fsd等。

说明

云平台不推荐您将Linux系统更换为Windows系统，当Linux系统中存在LVM分区时，切换为Windows系统后可能会导致LVM逻辑分区无法识别。



计费规则

“按需付费”方式购买的弹性云服务器切换操作系统后，由于所选镜像不同，系统盘的容量可能会增大，由此将带来费用的变更。

前提条件

- 完成云服务器的数据备份。
详细内容，请参见[备份弹性云服务器](#)。
- 若在切换操作系统过程中，要将云服务器的登录鉴权方式由密码改为密钥，请提前创建密钥文件。
详细内容，请参见[（推荐）通过管理控制台创建密钥对](#)。
- 如果您使用私有镜像切换操作系统请参考[《镜像服务用户指南》](#)提前完成私有镜像的制作。
 - 如果需要指定弹性云服务器的镜像，请提前使用指定弹性云服务器创建私有镜像。
 - 如果需要使用本地的镜像文件，请提前将镜像文件导入并注册为云平台的私有镜像。
 - 如果需要使用其他区域的私有镜像，请提前复制镜像。
 - 如果需要使用其他账号的私有镜像，请提前完成镜像共享。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在待切换操作系统的弹性云服务器的“操作”列下，单击“更多 > 镜像 > 切换操作系统”。
切换操作系统前请先将云服务器关机，或根据页面提示勾选“立即关机（切换操作系统前需先将云服务器关机）”。
5. 根据需求选择需要切换的弹性云服务器镜像。

说明

对于“包年/包月”方式购买的弹性云服务器，如果系统盘容量小于您选择的待切换镜像的大小，此时，您需要先卸载系统盘，并进行扩容，然后再挂载至原弹性云服务器执行切换操作。

扩容系统盘的操作指导，请参见[“扩容云硬盘”](#)章节。

图 2-140 切换操作系统



6. 设置登录方式。

如果待切换操作系统的弹性云服务器是使用密钥登录方式创建的，此时可以更换使用新密钥。

7. 单击“确定”。

8. 在“切换云服务器操作系统”页面，确认切换的操作系统规格无误后，阅读并勾选相关协议或声明，单击“提交申请”。

提交切换操作系统的申请后，弹性云服务器的状态变为“切换中”，当该状态消失后，表示切换结束。

说明

切换操作系统过程中，会创建一台临时弹性云服务器，切换操作系统结束后会自动删除。

后续处理

- 如果切换操作系统前后都是Linux系统，且数据盘设置了开机自动挂载分区。切换操作系统后，数据盘分区挂载信息会丢失，请更新/etc/fstab配置。
 - a. 在/etc/fstab写入切换后的分区信息。

建议您先备份/etc/fstab文件。

详细操作请参考[初始化Linux数据盘 \(fdisk\)](#)，设置开机自动挂载磁盘分区。
 - b. 挂载分区。挂载分区后即可开始使用数据盘。

mount 磁盘分区 挂载目录
 - c. 执行以下命令，查看挂载结果。

df -TH
- 如果操作系统切换失败，云服务平台支持重试功能，用户可重新执行3-8，切换弹性云服务器的操作系统。

- 重试后，如果仍未成功，可直接联系客服，客服会在后台进行人工恢复。



2.10 查看弹性云服务器信息

2.10.1 查看创建状态

操作场景

用户申请创建弹性云服务器后，可以查看任务的创建状态。本节介绍如何查看弹性云服务器的创建状态。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 用户执行购买弹性云服务器操作后，在列表的“状态”列，可查看该任务的“创建中”状态。

说明

- 创建云服务器存在以下状态：
 - 创建中：指系统正在处理创建云服务器的请求。
 - 创建失败：指未能成功处理的请求。对于创建失败的任务，系统会自动回退，同时在界面上直观的展示错误码，例如“(Ecs.0013) EIP配额不足”。
 - 运行中：成功处理创建云服务器的请求，是弹性云服务正常运行的状态。在这个状态的云服务器可以运行您的业务。
- 如果用户发现申请状态栏显示创建弹性云服务器的任务失败，而弹性云服务器列表中显示该弹性云服务器已创建成功，关于此问题请参见[为什么创建弹性云服务器的任务失败，但是在云服务器列表中显示创建成功](#)。

2.10.2 查看失败任务

操作场景

“失败信息”栏记录了系统处理中发生异常，导致处理失败的任务，包括该任务的“名称”、“状态”等信息。当有处理失败的操作时，控制台将显示该内容。本节介绍如何查看“失败信息”。

失败信息事件类型

“失败信息”栏记录的失败信息事件类型如[表2-49](#)所示。

表 2-49 失败信息事件类型

事件类型	说明
创建失败	指未能成功处理的请求。对于创建失败的任务，系统会自动回退，同时在界面上直观的展示错误码，例如“(Ecs.0013) EIP配额不足”。
操作失败	<ul style="list-style-type: none">变更规格 用户申请变更规格后，如果规格变更失败，则“失败信息”栏将显示本次变更规格操作。

操作步骤


1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 常用操作“开机/关机/重启/更多”的右侧即为“失败信息”。

图 2-141 失败任务



5. 单击“失败信息”栏对应的数字，即可查看系统处理失败的任务详情。其中：
 - 创建失败：您可以从创建失败的列表中查看创建中和创建失败的任务。
 - 操作失败：您可以从操作失败列表中查看处理异常的任务，包括异常任务的具体操作、错误码等，便于定位系统处理异常的原因，及时予以恢复。


2.10.3 查看弹性云服务器详细信息


操作场景

在您申请了弹性云服务器后，可以通过管理控制台查看和管理您的弹性云服务器。本节介绍如何查看弹性云服务器的详细配置，包括弹性云服务器名称、镜像信息、系统盘、数据盘、虚拟私有云、网卡、安全组、弹性公网IP地址和带宽等信息。

如需查看弹性云服务器的私有IP地址，请直接在弹性云服务器列表页进行查看。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。

3. 单击“”，选择“计算 > 弹性云服务器”。
系统进入弹性云服务器列表页，您可以在本页面查看您已购买的弹性云服务器，以及弹性云服务器的规格、镜像、IP地址等基本信息。
4. 在弹性云服务器列表中的上方，输入弹性云服务器名称、云服务器ID、IP地址等属性类型，并按回车键进行搜索。
5. 单击待查询弹性云服务器的名称。
系统跳转至该弹性云服务器详情页面。
6. 查看弹性云服务器的详细信息。
您可以选择“云硬盘/弹性网卡/安全组/弹性公网IP/监控/标签”页签，为弹性云服务器进行[更改安全组](#)、[绑定弹性网卡](#)、[绑定弹性公网IP](#)、[添加标签](#)等操作。




2.10.4 导出弹性云服务器列表信息

操作场景

您可以将当前账号下拥有的所有弹性云服务器信息，以.xlsx文件的形式导出至本地。该文件记录了弹性云服务器的如下信息：

- 基本信息：名称、区域、可用区、状态、规格、镜像、计费模式等信息。
- 网络信息：私有IP地址、弹性公网IP、VPC、子网、安全组等信息。
- 磁盘信息：磁盘数量、磁盘属性、磁盘容量等信息。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表页，单击右上角的。
系统会将您账号下，当前区域的所有弹性云服务器信息自动导出，并下载至本地。

说明

如需导出部分弹性云服务器的信息，请勾选对应的弹性云服务器，然后再单击右上角的



5. 在本地计算机桌面的默认下载地址文件夹，可以获取到导出的弹性云服务器列表信息。

2.10.5 搜索云服务器

操作场景

购买弹性云服务器后，可以通过管理控制台提供的搜索功能快速检索当前区域的云服务器资源。本节操作介绍搜索云服务器的操作步骤。当前支持使用名称、ID、可用区、状态、规格名称、镜像ID、弹性公网IP地址、私有IP地址、创建时间、计费模式、VPC ID、企业项目、资源标签以上属性进行搜索。

搜索语法

弹性云服务器支持多种类型的搜索，详细的说明和示例请参考表2-50。

说明


- 部分属性支持系统自动匹配属性搜索，但必须输入完整的属性值。
- 选择属性搜索时，以下属性仅支持精确搜索，需输入完整的属性值：ID、镜像ID、弹性公网IP地址、VPC ID、企业项目。
- 多值搜索时私有IP地址也必须使用精确搜索，需输入完整的属性值。
- 私有IP地址的网段需在以下范围内：10.0.0.0/8-24，172.16.0.0/12-24，192.168.0.0/16-24。
- 按标签过滤时可选择键或键值对搜索。支持设置多个标签，如果键不同，则标签按“与”的关系搜索。如果键相同，值不同，则标签按“与”的关系搜索。
- 未选择属性的时候标签不支持多值搜索。
- 私有IP和弹性公网IP不能同时搜索。

表 2-50 搜索语法

搜索类型	支持的属性	格式	示例	说明
属性值 自动匹配 属性	ID 规格名称 弹性公网IP地址 私有IP地址	完整的属性值	ID: 4a79dfec-f0d8-4181-9bef-495b8b7220e1 规格名称: s2.xlarge.4 私有IP地址: 192.168.99.231	使用关键字搜索时，不需要选择属性，仅输入完整的属性值，系统自动匹配属性类型进行搜索。 多个取值使用“,”分隔。否则后输入的取值会替换之前输入的取值。多个属性值之间为“或”的关系。
属性值 模糊搜索	名称 私有IP地址 规格名称	属性：不完整的属性值	名称: ecs-c 规格名称: s7n 私有IP地址: 192.168.0	选择属性后，输入对应的属性信息。
单属性	控制台支持的所有属性	属性：属性值	私有IP地址: 192.168.99.231	选择属性后，输入或选择对应的属性信息。 以下属性仅支持精确搜索，需输入完整的属性值：ID、弹性公网IP地址、镜像ID、VPC ID。

搜索类型	支持的属性	格式	示例	说明
多属性	控制台支持的所有属性	属性1: 取值 属性2: 取值	私有IP地址: 192.168.9 9.231 名称: ecs-c	支持选择多个不同的属性, 搜索时多个属性为“与”的关系。 以下属性仅支持精确搜索, 需输入完整的属性值: ID、弹性公网IP地址、镜像ID、VPC ID。
单属性、多值	ID 规格名称 弹性公网IP地址	属性: 取值 1,取值2	ID: 624eda28 -6bd9-40 2a-934b- 26c8969f 7169,bf6c 0281- f749-42d 7- b732-23a c69d80eb e 规格名称: s2,s3	选择一个属性后, 输入多个取值, 多个取值之间为“或”的关系。 以下属性仅支持精确搜索, 需输入完整的属性值: ID。
	状态 计费模式	属性: 取值 1 属性: 取值2	状态: 运行中 状态: 关闭	选择一个属性后, 勾选多个取值, 多个取值之间为“或”的关系。
多属性、多值	ID、规格名称、状态、弹性公网IP地址、私有IP地址、计费模式	属性1: 取值1,取值2 属性2: 取值1,取值2	ID: 624eda28 -6bd9-40 2a-934b- 26c8969f 7169,bf6c 0281- f749-42d 7- b732-23a c69d80eb e 规格名称: s2,s3	选择多属性、多值搜索时: <ul style="list-style-type: none"> 多个属性之间为“与”的关系。 同一属性的多个取值之间为“或”的关系。 以下属性仅支持精确搜索, 需输入完整的属性值: ID。 以下属性的多个值之间支持使用“ ”分隔或直接勾选: 状态、计费模式。

操作步骤

1. 登录管理控制台。
2. 单击“”，选择“计算 > 弹性云服务器”。
系统进入弹性云服务器列表页。
3. 在搜索的输入框中, 设置搜索条件。

您可以根据需要，选择并设置合适搜索条件。

- 在搜索框直接输入属性值，系统自动识别属性并搜索。
- 设置搜索属性类型进行搜索。
 - i. 单击搜索框，在列表中选择搜索的属性类型。

图 2-142 设置搜索属性类型



- ii. 根据提示信息，在搜索栏中输入或选择搜索条件，例如ID、状态、规格名称等，并单击回车键进行搜索。

示例 1：使用属性值搜索，系统自动匹配属性

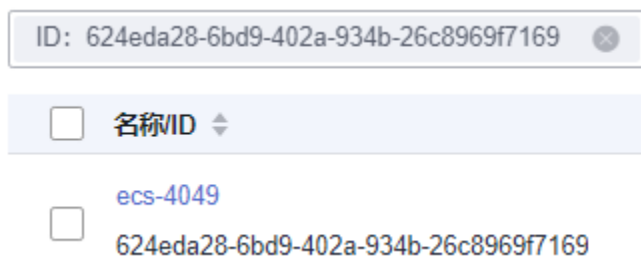
输入完整的属性值后，系统自动识别属性并搜索。多个取值使用“,”分隔，否则后输入的取值会替换之前输入的取值。多个属性值之间为“或”的关系。

- 单值搜索：
在云服务器列表页搜索的输入框中输入完整ID，按回车键进行搜索。

图 2-143 输入完整的 ID



图 2-144 系统自动匹配属性



- 多值搜索：

在云服务器列表页搜索的输入框中输入完整的规格名称，按回车键进行搜索。

图 2-145 输入多个完整的规格名称



图 2-146 系统自动匹配属性并搜索



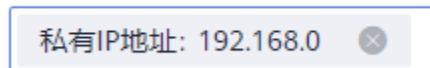
示例 2：单属性搜索

选择属性后，输入/选择对应的属性信息。

以下属性仅支持精确搜索，需输入完整的属性值：ID、弹性公网IP地址、镜像ID、VPC ID。

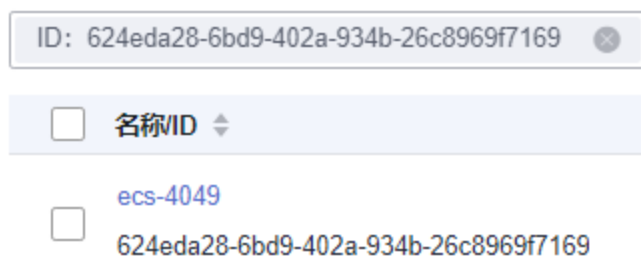
- 模糊搜索：使用私有IP地址搜索
 - a. 在云服务器列表页搜索的输入框中选择“私有IP地址”。
 - b. 根据提示，输入私有IP地址，按回车键进行搜索。“私有IP地址”支持模糊搜索，例如输入“192.168.0”，可以搜索到所有使用192.168.0网段的云服务器。

图 2-147 按私有 IP 地址搜索云服务器



- 精确搜索：使用ID搜索
 - a. 在云服务器列表页搜索的输入框中选择“ID”。
 - b. 根据提示，输入完整的ID，按回车键进行搜索。

图 2-148 按 ID 搜索云服务器



示例 3：多属性搜索

支持选择多个不同的属性，搜索时多个属性为“与”的关系。

以下属性仅支持精确搜索，需输入完整的属性值：ID、弹性公网IP地址、镜像ID、VPC ID。

本例同时使用“名称”和“私有IP地址”进行搜索。

1. 在云服务器列表页搜索的输入框中选择“名称”并输入取值。按回车键进行搜索。本例名称使用模糊搜索。
2. 添加筛选属性条件，选择“私有IP地址”，根据提示输入私有IP地址，按回车键进行搜索。本例私有IP地址使用模糊搜索。

图 2-149 按名称和私有 IP 地址搜索云服务器

名称/ID	监控	可用区	状态	规格/镜像	IP地址
ecs-f 6d8d189a-068b-4f07...		可用区2	关机	1vCPUs 2GiB s2.medium.2	192.168.128.121...
ecs-fb3 3e20a97a-e108-437c...		可用区2	运行中	1vCPUs 2GiB s2.medium.2	192.168.125.168...

示例 4：单属性、多值搜索

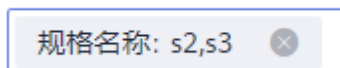
仅以下属性支持单属性、多值搜索：状态、ID、规格名称、私有IP地址、弹性公网IP地址、计费模式、标签。

其中以下属性仅支持精确搜索，需输入完整的属性值：ID、私有IP地址、弹性公网IP地址。

选择一个属性后，输入多个取值，多个取值之间为“或”的关系。

- 模糊搜索
 - a. 在云服务器列表页搜索的输入框中选择“规格名称”。
 - b. 根据提示输入多个规格名称，使用“,”分隔，按回车键进行搜索。

图 2-150 按规格名称搜索云服务器



- 精确搜索
 - a. 在云服务器列表页搜索的输入框中选择“私有IP地址”。
 - b. 根据提示输入多个完整的私有IP地址，使用“,”分隔，按回车键进行搜索。

图 2-151 按私有 IP 地址搜索云服务器

名称/ID	监控	可用区	状态	规格/镜像	IP地址
ecs-407d f7f8462a-111...		可用区2	运行中	1vCPUs 1GiB s3.small.1	192.168.0.94 (...)
ecs-6923 aadf81ad-904...		可用区2	运行中	1vCPUs 1GiB s3.small.1	192.168.0.115 (...)

示例 5：多属性、多值搜索

仅以下属性支持多属性、多值搜索：ID、规格名称、状态、弹性公网IP地址、计费模式。

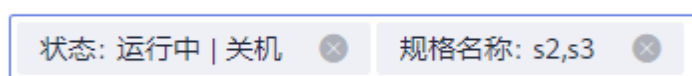
选择多属性、多值搜索：

- 多个属性之间为“与”的关系。
- 同一属性的多个取值之间为“或”的关系。

以下属性仅支持精确搜索，需输入完整的属性值：ID。

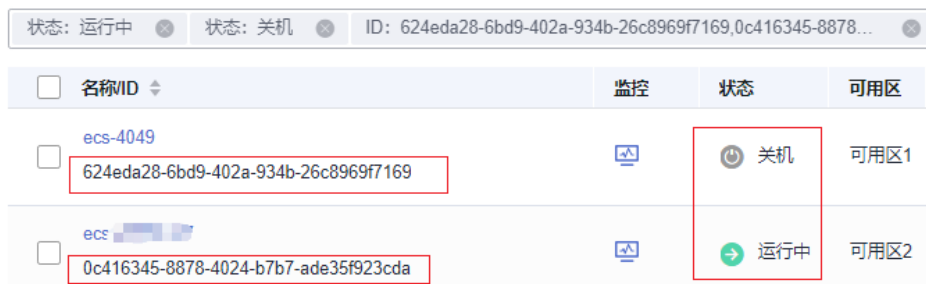
- 模糊搜索
 - a. 在云服务器列表页搜索的输入框中选择“状态”，并选择“运行中”、“关机”，按回车键进行搜索。
 - b. 添加筛选属条件，选择“规格名称”，输入多个规格名称，使用“,”分隔，按回车键进行搜索。

图 2-152 按状态和规格名称搜索云服务器



- 精确搜索
 - a. 在云服务器列表页搜索的输入框中选择“状态”，并选择“运行中”、“关机”，按回车键进行搜索。
 - b. 添加筛选属条件，选择“ID”，根据提示输入多个完整的ID，使用“,”分隔，按回车键进行搜索。

图 2-153 按状态和 ID 搜索云服务器



示例 6：按标签搜索

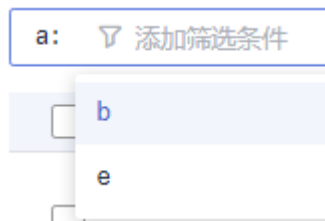
按标签过滤时可选择键或键值对搜索。

支持设置多个标签，如果键不同，则标签按“与”的关系搜索。

如果键相同，值不同，则标签按“与”的关系搜索。

- 单个标签搜索
在云服务器列表页搜索的输入框中选择“标签”，选择标签键值对后自动搜索。

图 2-154 按标签搜索云服务器



- 多个标签搜索

在云服务器列表页搜索的输入框中选择“标签”，选择多个标签键值对后自动搜索。

按“与”的关系搜索，搜索同时具备这两个标签的云服务器。

图 2-155 按标签搜索云服务器



<input type="checkbox"/>	名称ID	监控	状态	可用区	规格/镜像	IP地址	计费模式	企业项目	标签
<input type="checkbox"/>	ecs-0c416345-8878-4024-...		运行中	可用区2	2vCPUs 4GiB s... CentOS_7.5_...	192.168...	包年/包月 31天后到期		<input type="checkbox"/> _type=空值 a=b

3 镜像管理

3.1 镜像概述

什么是镜像

镜像是一个包含了软件及必要配置的云服务器或裸金属服务器模板，包含操作系统或业务数据，还可以包含应用软件（例如，数据库软件）和私有软件。镜像分为公共镜像、私有镜像、共享镜像。

镜像服务（Image Management Service）提供简单方便的镜像自助管理功能。用户可以灵活便捷地使用公共镜像、私有镜像或共享镜像申请云服务器。同时，用户还能通过已有的云服务器或使用外部镜像文件创建私有镜像。

公共镜像

公共镜像是包含常见的标准操作系统的镜像，所有用户可见，包括操作系统以及预装的公共应用，支持Ubuntu，CentOS，Debian等主流操作系统。用户可以从丰富的公共镜像库中选择或创建私有镜像，从而快速创建或批量复制弹性云服务器。请根据您的实际情况自助配置应用环境或相关软件。

更多关于公共镜像的介绍，请参见“[公共镜像概述](#)”。

私有镜像

私有镜像包含操作系统或业务数据、预装的公共应用以及用户的私有应用的镜像，仅用户个人可见。

表 3-1 私有镜像类型

镜像类型	说明
系统盘镜像	包含用户运行业务所需的操作系统、应用软件的镜像。系统盘镜像可以用于创建云服务器，迁移用户业务到云。
数据盘镜像	只包含用户业务数据的镜像。数据镜像可以用于创建云硬盘，将用户的业务数据迁移到云上。

镜像类型	说明
整机镜像	包含用户运行业务所需的操作系统、应用软件和业务数据的镜像。整机镜像包含系统盘和其所挂载的所有数据盘。
ISO镜像	将外部镜像的ISO文件注册到云平台的私有镜像。ISO镜像是特殊的镜像，只能发放用作临时过渡的云服务器。

如果您使用私有镜像切换操作系统请参考《[镜像服务用户指南](#)》提前完成私有镜像的制作。

- 如果需要指定弹性云服务器的镜像，请提前使用指定弹性云服务器创建私有镜像。
- 如果需要使用本地的镜像文件，请提前将镜像文件导入并注册为云平台的私有镜像。
- 如果需要使用其他区域的私有镜像，请提前复制镜像。
- 如果需要使用其他账号的私有镜像，请提前完成镜像共享。

共享镜像

用户将接受云平台其他用户共享的私有镜像，作为自己的镜像进行使用。更多关于共享镜像的使用，请参见[共享镜像](#)。

- 用户只能共享自己没有发布为市场镜像的私有镜像，已经发布为市场镜像的不能共享。
- 镜像共享的范围只能在区域内。
- 每个镜像最多可以共享给128个租户。
- 用户可以随时取消自己共享的镜像，无需通知镜像的接受方。
- 用户可以随时删除自己共享的镜像，无需通知镜像的接受方。
- 加密镜像不能共享。
- 只有通过云备份或云服务器（未通过旧版CSBS服务生成备份）创建的整机镜像，才支持共享。通过其他方式创建的整机镜像，暂不支持共享。

市场镜像

市场镜像提供预装操作系统、应用环境和各类软件的优质第三方镜像。无需配置，可一键部署，满足建站、应用开发、可视化管理等个性化需求。

市场镜像由镜像服务商定价，有免费镜像，也有付费镜像。使用付费镜像创建云服务器时，需要支付市场镜像的费用和云服务器费用。

相关链接

- [了解私有镜像的创建方式](#)
- [镜像源管理类常见问题](#)

3.2 创建镜像

操作场景

您可以使用已有的弹性云服务器创建系统盘镜像、数据盘镜像、整机镜像。

- 系统盘镜像：系统盘镜像包含用户运行业务所需的操作系统、应用软件，可以用于云服务器业务的迁移。
- 数据盘镜像：数据盘镜像是只包含用户业务数据的镜像。用户可以通过创建数据盘镜像对云服务器的数据盘进行导出。数据盘镜像可以用于创建云硬盘，将用户的业务数据迁移到云上。
- 整机镜像：使用弹性云服务器携带其挂载的数据盘一起创建整机镜像，创建的整机镜像包含用户的业务数据，可用于快速发放包含用户业务数据的弹性云服务器。
- ISO 镜像：将外部镜像的ISO文件注册到云平台的私有镜像。ISO镜像是特殊的镜像，只能发放用作临时过渡的云服务器。

如果您使用私有镜像切换操作系统请参考《[镜像服务用户指南](#)》提前完成私有镜像的制作。

前提条件

创建镜像前，请确保弹性云服务器已完成相关初始化配置。

详细内容，请参考[创建私有镜像前云服务器或镜像文件需要完成哪些初始化配置？](#)

操作步骤



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器的“操作”列下，单击“更多 > 镜像 > 创建镜像”。
5. 根据界面要求填写如下信息：
包含“镜像类型和来源”和“配置信息”两个信息块，各参数说明参见[表3-2](#)和[表3-3](#)。

表 3-2 镜像类型和来源

参数	说明
创建方式	选择“创建私有镜像”。
区域	请选择靠近您业务的区域。 如果区域选择错误，可以在私有镜像创建成功后通过“ 跨区域复制镜像 ”将镜像复制到其他区域。
镜像类型	选择“系统盘镜像”。

参数	说明
选择镜像源	选择“云服务器”，然后从列表中选择已完成相关配置的云服务器。

表 3-3 配置信息

参数	说明
加密	镜像的加密属性，不可更改。 <ul style="list-style-type: none">使用未加密的云服务器创建的私有镜像为未加密的私有镜像。使用加密的云服务器创建的私有镜像为加密的私有镜像。
名称	设置一个便于您识别的镜像名称。
企业项目	从下拉列表中选择所在的企业项目。该参数针对企业用户使用，只有开通了企业项目的客户，或者权限为企业主账号的客户才可见。如需使用该功能，请联系您的客户经理申请开通。 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。
标签	可选参数，为镜像设置标签键和标签值，便于识别和管理。
描述	可选参数，对镜像进行描述。

- 单击“立即创建”。

4 磁盘管理

4.1 ECS 磁盘概述

什么是云硬盘

云硬盘（Elastic Volume Service, EVS）可以为弹性云服务器提供高可靠、高性能、规格丰富并且可弹性扩展的块存储服务，满足不同场景的业务需求，适用于分布式文件系统、开发测试、数据仓库以及高性能计算等场景。

云硬盘的类型

不同类型云硬盘的性能各不相同，您可根据应用程序要求选择您所需的云硬盘。

更多关于云硬盘规格、性能等信息，请参见《[云硬盘用户指南](#)》。

相关链接

- [挂载磁盘](#)
- [初始化数据盘](#)
- [登录Windows弹性云服务器，找不到新购买的数据盘？](#)
- [怎样调整系统盘分区？](#)
- [一台弹性云服务器可以挂载多块磁盘吗？](#)
- [弹性云服务器挂载磁盘时有什么限制？](#)

4.2 为 ECS 新增磁盘

操作场景



云服务器的磁盘包括系统盘和数据盘。系统盘在创建云服务器时自动创建并挂载，无需单独购买。数据盘包括如下两种创建方式：

- 在购买云服务器的时候一同购买：该场景下，数据盘由系统自动挂载给云服务器。

- 在购买了云服务器之后，单独购买云硬盘：该场景下，数据盘需要手动挂载给云服务器。

本节操作介绍购买了云服务器之后，新增磁盘的操作步骤。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器的“操作”列下，单击“更多 > 磁盘/备份 > 新增磁盘”。系统跳转至购买磁盘页面。
5. 根据界面提示，设置新添加云硬盘的参数信息。
云硬盘的参数信息配置请参考[购买云硬盘](#)。

说明

- 新增磁盘默认与弹性云服务器计费模式保持一致。
 - 新增磁盘默认与弹性云服务器位于同一区域。
 - 新增磁盘默认与弹性云服务器位于同一可用区，不支持修改磁盘可用区。
 - 购买完成后新增磁盘默认已挂载至选择的弹性云服务器。
 - 新增包年/包月云硬盘与弹性云服务器的到期时间相同。
6. 单击“立即购买”，确认订单详情并完成支付。
返回弹性云服务器控制台在“云硬盘”页签下，查看新增磁盘信息。

后续处理

云硬盘挂载至云服务器后，需要登录云服务器初始化云硬盘，即格式化云硬盘，之后云硬盘才可以正常使用。

初始化数据盘的具体操作，请参见[初始化数据盘](#)。

4.3 为 ECS 挂载磁盘

操作场景

弹性云服务器创建成功后，如果发现磁盘不够用或当前磁盘不满足要求，可以将已有云硬盘挂载给弹性云服务器，或在“存储 > 云硬盘”页面购买新的磁盘，然后再挂载至弹性云服务器。

前提条件

- 已创建可用的云硬盘。
创建云硬盘的操作请参考[“购买云硬盘”](#)。

操作步骤

1. 登录管理控制台。


2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中的右上角，输入弹性云服务器名、IP地址或ID，并进行搜索。
5. 单击待挂载云硬盘的弹性云服务器的名称。
系统跳转至该弹性云服务器详情页面。
6. 选择“云硬盘”页签，并单击“挂载磁盘”。
系统跳转至“挂载磁盘”页面。

图 4-1 挂载磁盘（KVM）



7. 根据界面提示，勾选目标磁盘，并指定该磁盘作为系统盘或数据盘。
 - 对于采用KVM虚拟化类型的弹性云服务器，您可以指定该磁盘作为系统盘或数据盘，但不能指定具体的磁盘挂载点。
 - 对于采用XEN虚拟化类型的弹性云服务器，系统支持指定具体的磁盘挂载点，如/dev/vdb。

📖 说明

- 如果无可用的云硬盘，请单击列表下方的“购买云硬盘”进行购买。
 - 挂载磁盘时的约束限制，请参见[弹性云服务器挂载磁盘时有什么限制](#)。
8. 单击“确定”。
挂载成功后，在弹性云服务器详情页的“云硬盘”页签，即可看到新挂载的磁盘信息。

后续处理

如果挂载的磁盘是新创建的，则云硬盘挂载至云服务器后，需要登录云服务器初始化云硬盘，即格式化云硬盘，之后云硬盘才可以正常使用。

初始化数据盘的具体操作，请参见[初始化数据盘](#)。

4.4 初始化数据盘

4.4.1 初始化概述

一块全新的数据盘挂载到云服务器后，还不能直接存储数据，您需要对数据盘进行分区创建、文件系统挂载等初始化操作后，才可以正常使用。

操作场景

- **系统盘**

创建云服务器时，系统盘会被自动初始化，默认磁盘分区形式为主启动记录分区（MBR, Master boot record）。

- **全新数据盘**

- 随云服务器创建的数据盘，系统已自动将数据挂载至云服务器，需要您对数据盘进行初始化后才能使用。
- 不随云服务器创建的数据盘，需要先挂载至云服务器后，再进行初始化。

全新数据盘初始化操作指导，请参见[表4-1](#)。

- **有数据的数据盘**

有数据的数据盘是指从快照/备份/镜像创建的数据盘，或从其他云服务器上卸载后挂载至另一云服务器上的有数据的数据盘。

- 您可以选择不初始化，直接使用数据盘已有分区：
 - Linux：需要将现有分区挂载至系统指定目录，并设置开机自动挂载至系统指定目录。
挂载操作请参见[初始化Linux数据盘（容量小于等于2TiB）](#)。
 - Windows：无需做任何操作，直接可以使用。
- 也可以选择对数据盘重新初始化：
重新划分分区，数据盘已有数据将会被全部删除，建议您先使用快照为云硬盘备份数据。
 - Linux：需要先卸载目录，并删除现有分区（运行 `fdisk 数据盘名称` 命令，再输入“d”，输入分区编号，输入“w”保存），再重新初始化即可。
 - Windows：需要先删除现有分区（使用“删除卷”工具），再重新初始化即可。

初始化操作请参见[表4-1](#)。

说明

初始化云硬盘不会删除云硬盘快照，初始化后仍可以使用快照回滚数据至原云硬盘。

操作指导

表 4-1 云硬盘初始化操作指导

云硬盘容量	分区格式	分区类型	操作系统	参考文档
容量 ≤ 2TiB	GPT/MBR	<ul style="list-style-type: none">GPT分区不区分主分区、扩展分区以及逻辑分区，且分区个数无限制。MBR分区个数最多支持：<ul style="list-style-type: none">4个主分区3个主分区和1个扩展分区在扩展分区中创建逻辑分区的数量没有限制，可以创建任意多个逻辑分区。 <p>如果您需要划分大于4个分区，只能使用主分区+1个扩展分区，然后在这个扩展分区中划分多个逻辑分区。</p>	Linux	初始化Linux数据盘（容量小于等于2TiB）
			Windows	初始化windows数据盘
容量 > 2TiB	GPT	GPT分区不区分主分区、扩展分区以及逻辑分区，且分区个数无限制。	Linux	初始化Linux数据盘（容量大于2TiB）
			Windows	初始化windows数据盘

须知

- MBR分区支持的云硬盘最大容量为2 TiB，GPT分区支持的最大云硬盘容量为18 EiB，因此当云硬盘容量大于2 TiB或目前小于等于2 TiB但后续可能会扩容至2TiB以上时，请在初始化云硬盘时，分区格式选择GPT分区。
- 切换云硬盘分区格式，则该云硬盘上的数据将会被清除，请您在初始化时谨慎选择云硬盘的分区格式。
- 对于Linux操作系统，支持使用fdisk和parted工具创建MBR分区，仅支持使用parted工具创建GPT分区。

4.4.2 初始化 Linux 数据盘（容量小于等于 2TiB）

操作场景

本文介绍使用方式初始化Linux系统中的数据盘，不同云服务器的操作系统的初始化操作可能不同，请根据您的实际环境进行操作。

前提条件

云硬盘已挂载到云服务器上。

约束与限制

- 从数据源创建的云硬盘无需初始化。该云硬盘在初始状态就具有数据源中的数据，初始化有丢失数据的风险。如果需要重新初始化，建议先备份云硬盘中的数据。
使用CBR备份请参考[备份云硬盘](#)，使用快照备份请参考[管理云硬盘快照](#)。
- 初始化云硬盘不会删除云硬盘快照，初始化后仍可以使用快照回滚数据至原云硬盘。

手动初始化数据盘

📖 说明

MBR分区个数最多支持4个主分区或最多3个主分区+1个扩展分区，在扩展分区中创建任意多个逻辑分区。

例如：您需要划分4个分区，可以使用4个主分区，或者1个主分区+1个扩展分区（3个逻辑分区），或者2个主分区+1个扩展分区（2个逻辑分区），或者3个主分区+1个扩展分区（1个逻辑分区）。

示例说明：使用fdisk工具，针对“/dev/vdb”数据盘划分2个MBR主分区，分别为40GiB（/dev/vdb1）、60GiB（/dev/vdb2）。

步骤1 登录云服务器。

登录弹性云服务器请参见[登录弹性云服务器](#)。

登录裸金属服务器请参见[登录裸金属服务器](#)。

步骤2 针对数据盘“/dev/vdb”划分2个主分区/dev/vdb1、/dev/vdb2。

- 查看/dev/vdb数据盘容量为100GiB。

lsblk

```
[root@ecs-centos76 ~]# lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
vda 253:0 0 40G 0 disk
├vda1 253:1 0 1G 0 part /boot
├vda2 253:2 0 39G 0 part /
vdb 253:16 0 100G 0 disk
```

- 开始新建第一个主分区/dev/vdb1。

fdisk /dev/vdb

```
n
p
1
```

📖 说明

- “Partition type”：“p”表示主分区，“e”表示扩展分区。
- “1”代表主分区编号

```
[root@ecs-test-0001 ~]# fdisk /dev/vdb
Welcome to fdisk (util-linux 2.23.2).
```

```
Changes will remain in memory only, until you decide to write them.
Be careful before using the write command.
```

```
Device does not contain a recognized partition table
Building a new DOS disklabel with disk identifier 0x38717fc1.
```

```
Command (m for help): n
```



```
Partition type:
 p primary (0 primary, 0 extended, 4 free)
 e extended
Select (default p): p
Partition number (1-4, default 1): 1
```

针对第一个分区/dev/vdb1 (40 GiB) 设置**起始磁柱值 (2048)**和**截止磁柱值 (83886079)**。

```
First sector (2048-209715199, default 2048): 2048
Last sector, +sectors or +size{K,M,G} (2048-209715199, default 209715199):83886079
Partition 1 of type Linux and of size 40 GB is set
```

3. 开始新建第二个主分区/dev/vdb2。

```
n
p
2
```

```
Command (m for help): n
Partition type:
 p primary (0 primary, 0 extended, 4 free)
 e extended
Select (default p): p
Partition number (1-4, default 2): 2
```

针对第2个分区/dev/vdb2设置**起始磁柱值 (83886080)**和**截止磁柱值 (209715199)**。

```
First sector (83886080-209715199, default 83886080): 83886080
Last sector, +sectors or +size{K,M,G} (83886080-209715199, default 209715199):209715199
Partition 2 of type Linux and of size 60 GB is set
```

说明

分区的起始磁柱值和截止磁柱值计算方法：

sectors值=容量/512 bytes, 1GiB=1073741824 bytes

- First sector (2048-209715199, default 2048)是数据盘/dev/vdb (100 GiB) 的磁柱范围
起始磁柱=2048
截止磁柱值=sectors值-1=(100 * 1073741824 / 512)-1=209715200-1=209715199
- 数据盘/dev/vdb的第1个分区/dev/vdb1 (40 GiB) :
起始磁柱值=2048 (此处使用数据盘/dev/vdb的起始磁柱)
截止磁柱值=sectors值-1= (40 * 1073741824 / 512) -1=83886079
- 数据盘/dev/vdb的第2个分区/dev/vdb2 (60 GiB) :
起始磁柱值 = /dev/vdb1的截止磁柱值 + 1 = 83886079+1 = 83886080
截止磁柱值 = 起始磁柱值 + sectors - 1 = 83886080+(60 * 1073741824 / 512) -1 = 209715199

步骤3 查看新建分区大小、分区格式信息。

1. 确定之前的分区操作是否正确。

```
p
Command (m for help): p

Disk /dev/vdb: 107.4 GB, 107374182400 bytes, 209715200 sectors
Units = sectors of 1 * 512 = 512 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk label type: dos
Disk identifier: 0x994727e5

Device Boot      Start         End      Blocks   Id  System
/dev/vdb1        2048     83886079    41942016   83  Linux
```

```
/dev/vdb2 83886080 209715199 62914560 83 Linux
```

Command (m for help):

📖 说明

如果之前分区操作有误，请输入“q”，按“Enter”，则会退出fdisk分区工具，之前的分区结果将不会被保留。此时，重新执行创建分区步骤1和步骤2即可。

2. 确认完成后，将分区结果写入分区表中，并变更同步至操作系统。

w

partprobe

📖 说明

如果出现报错“-bash: partprobe: command not found”，则说明系统不识别该命令，需要执行“yum install -y parted”命令来安装该命令。安装成功后再执行上述命令。

3. 再次确认分区格式为MBR。

parted /dev/vdb

p

📖 说明

“Partition Table: msdos”表示磁盘分区格式为MBR

```
[root@ecs-test-0001 ~]# parted /dev/vdb
GNU Parted 3.1
Using /dev/vdb
Welcome to GNU Parted! Type 'help' to view a list of commands.
(parted) p
Model: Virtio Block Device (virtblk)
Disk /dev/vdb: 107GB
Sector size (logical/physical): 512B/512B
Partition Table: msdos
Disk Flags:

Number Start End Size Type File system Flags
 1 1049kB 42.9GB 42.9GB primary
 2 42.9GB 107GB 64.4GB primary

(parted) q
[root@ecs-test-0001 ~]#
```

输入“q”，按“Enter”，退出parted模式。

- 步骤4** 分别对分区/dev/vdb1（40GiB）和/dev/vdb2（60GiB）创建ext4文件系统。

mkfs -t ext4 /dev/vdb1

mkfs -t ext4 /dev/vdb2

📖 说明

创建文件系统格式需要等待一段时间，请确保看到如下回显后，再退出。

```
[root@ecs-test-0001 ~]# mkfs -t ext4 /dev/vdb1
mke2fs 1.42.9 (28-Dec-2013)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
2621440 inodes, 10485504 blocks
524275 blocks (5.00%) reserved for the super user
First data block=0
```

```
Maximum filesystem blocks=2157969408
320 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624

Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done
```

执行以下命令确认文件系统类型是否是ext4

```
parted /dev/vdb
```

```
p
```

```
[root@ecs-test-0001 ~]# parted /dev/vdb
GNU Parted 3.1
Using /dev/vdb
Welcome to GNU Parted! Type 'help' to view a list of commands.
(parted) p
Model: Virtio Block Device (virtblk)
Disk /dev/vdb: 107GB
Sector size (logical/physical): 512B/512B
Partition Table: msdos
Disk Flags:

Number  Start  End    Size  Type  File system  Flags
  1      1049kB 42.9GB 42.9GB primary ext4
  2      42.9GB 107GB  64.4GB primary ext4

(parted) q
[root@ecs-test-0001 ~]#
```

输入“q”，按“Enter”，退出parted模式。

步骤5 新建目录，并将新建分区挂载至新建目录。

```
mkdir -p /mnt/sdc
```

```
mkdir -p /mnt/sdd
```

```
mount /dev/vdb1 /mnt/sdc
```

```
mount /dev/vdb2 /mnt/sdd
```

```
lsblk
```

查看挂载结果

```
[root@ecs-test-0001 ~]# lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
vda   253:0   0 40G  0 disk
├─vda1 253:1   0 40G  0 part /
vdb   253:16  0 100G 0 disk
├─vdb1 253:17  0 40G  0 part /mnt/sdc
└─vdb2 253:18  0 60G  0 part /mnt/sdd
```

表示新建分区“/dev/vdb1”和“/dev/vdb2”已分别挂载至“/mnt/sdc”、“/mnt/sdd”。

步骤6 使用磁盘分区的UUID来设置开机自动挂载磁盘分区。

说明

- 重启后，挂载会失效，因此需要设置开机自动挂载磁盘分区，即在/etc/fstab文件中添加新建磁盘分区信息，启动开机自动挂载磁盘分区。
- 不建议采用在“/etc/fstab”文件中直接指定设备名（比如/dev/vdb1）的方法，因为云中设备的顺序编码在关闭或者开启云服务器过程中可能发生改变（例如：/dev/vdb1可能会变成/dev/vdb2），可能会导致云服务器重启后不能正常运行。
- UUID（universally unique identifier）是Linux系统为磁盘分区提供的唯一的标识字符串。
- 该操作不会影响弹性云服务器中的现有数据。

1. 查询磁盘分区的UUID。

blkid /dev/vdb1**blkid /dev/vdb2**

```
[root@ecs-test-0001 ~]# blkid /dev/vdb1
/dev/vdb1: UUID="0b3040e2-1367-4abb-841d-ddb0b92693df" TYPE="ext4"
/dev/vdb2: UUID="0d6769k2-1745-9dsf-453d-hgd0b34267dj" TYPE="ext4"
```

磁盘分区“/dev/vdb1”和“/dev/vdb2”的UUID分别为0b3040e2-1367-4abb-841d-ddb0b92693df、0d6769k2-1745-9dsf-453d-hgd0b34267dj。

2. 设置开机自动挂载磁盘分区

vi /etc/fstab

按“i”，进入编辑模式，将光标移至文件末尾，按“Enter”，添加下行内容：

```
UUID=0b3040e2-1367-4abb-841d-ddb0b92693df /mnt/sdc ext4 defaults 0 2
UUID=0d6769k2-1745-9dsf-453d-hgd0b34267dj /mnt/sdd ext4 defaults 0 2
```

按“ESC”后，输入“:wq”，按“Enter”，保存设置并退出编辑器。

表 4-2 参数说明

参数示例	说明
UUID=0b3040e2-1367-4abb-841d-ddb0b92693df	磁盘分区的UUID
/mnt/sdc	磁盘分区的挂载目录
ext4	磁盘分区的文件系统格式
defaults	磁盘分区的挂载选项，此处通常设置为defaults即可
0	<ul style="list-style-type: none"> - Linux dump备份选项。 <ul style="list-style-type: none"> ▪ 0表示不使用Linux dump备份。现在通常不使用dump备份，此处设置为0即可。 ▪ 1表示使用Linux dump备份。

参数示例	说明
2	<ul style="list-style-type: none">- fsck选项，即开机时是否使用fsck检查磁盘。<ul style="list-style-type: none">▪ 2表示从挂载点为非根目录 (/) 的分区开始检验。▪ 1表示从挂载点为根目录 (/) 的分区开始检验。▪ 0表示不检验。

步骤7 验证自动挂载功能已生效。

```
umount /dev/vdb1
```

```
umount /dev/vdb2
```

```
mount -a
```

则系统会将“/etc/fstab”文件所有内容重新加载。

查询文件系统挂载信息。

```
mount | grep /mnt/sdc
```

```
mount | grep /mnt/sdd
```

回显类似如下信息，说明自动挂载功能生效。

```
root@ecs-test-0001 ~]# mount | grep /mnt/sdc  
/dev/vdb1 on /mnt/sdc type ext4 (rw,relatime,data=ordered)  
root@ecs-test-0001 ~]# mount | grep /mnt/sdd  
/dev/vdb2 on /mnt/sdd type ext4 (rw,relatime,data=ordered)
```

---结束

4.4.3 初始化 Linux 数据盘（容量大于 2TiB）

操作场景

云硬盘容量大于2TiB时，只能使用parted工具为磁盘新建GPT分区。不同云服务器的操作系统的初始化操作可能不同。

分区格式	操作系统	文件系统	初始化工具	配置示例
GPT	不限	ext* (如ext2、ext3、ext4)、xfs、btrfs	parted	<ul style="list-style-type: none">设备名: /dev/vdb文件系统: ext4挂载目录: /mnt/sdc分区: /dev/vdb1分区格式: GPT容量: 3TiB

前提条件

云硬盘已挂载到云服务器上。

约束与限制

- 从数据源创建的云硬盘无需初始化。该云硬盘在初始状态就具有数据源中的数据，初始化有丢失数据的风险。如果需要重新初始化，建议先备份云硬盘中的数据。
使用CBR备份请参考[备份云硬盘](#)，使用快照备份请参考[管理云硬盘快照](#)。
- 初始化云硬盘不会删除云硬盘快照，初始化后仍可以使用快照回滚数据至原云硬盘。

初始化容量大于 2TiB 的数据盘

示例说明：使用parted工具，针对“/dev/vdb”数据盘划分1个GPT分区。

步骤1 登录云服务器。

登录弹性云服务器请参见[登录弹性云服务器](#)。

登录裸金属服务器请参见[登录裸金属服务器](#)。

步骤2 针对数据盘“/dev/vdb”划分1个分区/dev/vdb1。

- 查看/dev/vdb数据盘容量为3TiB。

lsblk

```
[root@ecs-centos76 ~]# lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
vda 253:0 0 40G 0 disk
└─vda1 253:1 0 40G 0 part /
vdb 253:16 0 3T 0 disk
```

- 开始新建分区/dev/vdb1。

parted /dev/vdb

p

```
[root@ecs-centos74 ~]# parted /dev/vdb
GNU Parted 3.1
Using /dev/vdb
```

```
Welcome to GNU Parted! Type 'help' to view a list of commands.
(parted) p
Error: /dev/vdb: unrecognised disk label
Model: Virtio Block Device (virtblk)
Disk /dev/vdb: 3299GB
Sector size (logical/physical): 512B/512B
Partition Table: unknown
Disk Flags:
(parted)
```

“Partition Table”：“unknown”表示磁盘分区格式未知，新的数据盘还未设置分区格式。

📖 说明

如果出现报错“-bash: parted: command not found”，则说明系统不识别该命令，需要执行“yum install -y parted”命令来安装该命令。安装成功后再执行上述命令。

3. 针对分区/dev/vdb1设置分区格式为GPT。

mklabel gpt

unit s

p

```
(parted) mklabel gpt
(parted) unit s
(parted) p
Model: Virtio Block Device (virtblk)
Disk /dev/vdb: 6442450944s
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:

Number Start End Size File system Name Flags
(parted)
```

📖 说明

- 如果磁盘容量小于等于2 TiB，您想使用parted工具新建MBR分区，则此处命令为：**mklabel msdos**。
- 切换云硬盘分区格式，则该云硬盘上的数据将会被清除，请您在初始化时谨慎选择云硬盘的分区格式。
- 云硬盘初始化时设置的磁盘分区格式（MBR或GPT），后续在该云硬盘下创建的所有分区的格式均和首次初始化时设置的分区格式一致，因此如果您创建第2个及后续分区时，无需执行该步骤。

4. 针对分区/dev/vdb1设置分区名称及大小。

mkpart /dev/vdb1 2048s 100%

p

📖 说明

- “2048s”表示磁盘起始磁柱值，“100%”为磁盘截止磁柱值，表示将磁盘100%容量给到/dev/vdb1分区。
- 如果您需要将数据盘容量划分至两个及以上的分区的起始磁柱值和截止磁柱值计算方法请参见[手动初始化数据盘](#)。

```
(parted) mkpart /dev/vdb1 2048s 100%
(parted) p
Model: Virtio Block Device (virtblk)
Disk /dev/vdb: 6442450944s
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:
```

```
Number Start      End          Size          File system Name      Flags
1      2048s        6442448895s 6442446848s          /dev/vdb1
```

输入“q”，按“Enter”，执行“lsblk”再次查看新建分区“/dev/vdb1”。

```
[root@ecs-centos74 ~]# lsblk
NAME MAJ:MIN RM  SIZE RO TYPE MOUNTPOINT
vda   253:0  0   40G  0 disk
├─vda1 253:1  0   40G  0 part /
vdb   253:16 0    3T  0 disk
├─vdb1 253:17 0    3T  0 part
```

步骤3 为“/dev/vdb1”创建ext4文件系统。

```
mkfs -t ext4 /dev/vdb1
```

📖 说明

创建文件系统格式需要等待一段时间，请观察系统运行状态，不要退出。

```
[root@ecs-test-0001 ~]# mkfs -t ext4 /dev/vdb1
mke2fs 1.42.9 (28-Dec-2013)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
201326592 inodes, 805305856 blocks
40265292 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=2952790016
24576 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
    4096000, 7962624, 11239424, 20480000, 23887872, 71663616, 78675968,
    102400000, 214990848, 512000000, 550731776, 644972544

Allocating group tables: done
Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done
[root@ecs-test-0001 ~]#
```

执行“parted /dev/vdb”命令，再输入“p”，查看分区文件系统类型。

```
[root@ecs-test-0001 ~]# parted /dev/vdb
GNU Parted 3.1
Using /dev/vdb
Welcome to GNU Parted! Type 'help' to view a list of commands.
(parted) p
Model: Virtio Block Device (virtblk)
Disk /dev/vdb: 3299GB
Sector size (logical/physical): 512B/512B
Partition Table: gpt
Disk Flags:

Number Start End      Size  File system Name      Flags
1      1049kB 3299GB 3299GB ext4      /dev/vdb1

(parted) q
[root@ecs-test-0001 ~]#
```

输入“q”，按“Enter”，退出parted模式。

步骤4 新建目录，并将新建分区挂载至新建目录。

```
mkdir -p /mnt/sdc
```



```
mount /dev/vdb1 /mnt/sdc
```

```
lsblk
```

```
[root@ecs-test-0001 ~]# lsblk
NAME MAJ:MIN RM SIZE RO TYPE MOUNTPOINT
vda   253:0   0  40G  0 disk
├─vda1 253:1   0  40G  0 part /
vdb   253:16  0   3T  0 disk
└─vdb1 253:17  0   3T  0 part /mnt/sdc
```

表示新建分区“/dev/vdb1”已挂载至“/mnt/sdc”。

步骤5 使用磁盘分区的UUID来设置开机自动挂载磁盘分区。

📖 说明

- 重启后，挂载会失效，因此需要设置开机自动挂载磁盘分区，即在/etc/fstab文件中添加新建磁盘分区信息，启动开机自动挂载磁盘分区。
- 不建议采用在“/etc/fstab”文件中直接指定设备名（比如/dev/vdb1）的方法，因为云中设备的顺序编码在关闭或者开启云服务器过程中可能发生改变（例如：/dev/vdb1可能会变成/dev/vdb2），可能会导致云服务器重启后不能正常运行。
- UUID（universally unique identifier）是Linux系统为磁盘分区提供的唯一的标识字符串。
- 该操作不会影响弹性云服务器中的现有数据。

1. 查询磁盘分区的UUID。

```
blkid /dev/vdb1
```

```
[root@ecs-test-0001 ~]# blkid /dev/vdb1
/dev/vdb1: UUID="0b3040e2-1367-4abb-841d-ddb0b92693df" TYPE="ext4"
```

磁盘分区“/dev/vdb1”的UUID为0b3040e2-1367-4abb-841d-ddb0b92693df。

2. 设置开机自动挂载磁盘分区。

```
vi /etc/fstab
```

按“i”，进入编辑模式，将光标移至文件末尾，按“Enter”，添加下行内容：

```
UUID=0b3040e2-1367-4abb-841d-ddb0b92693df /mnt/sdc ext4 defaults 0 2
```

按“ESC”后，输入“:wq”，按“Enter”，保存设置并退出编辑器。

表 4-3 参数说明

参数示例	说明
UUID=0b3040e2-1367-4abb-841d-ddb0b92693df	磁盘分区的UUID
/mnt/sdc	磁盘分区的挂载目录
ext4	磁盘分区的文件系统格式
defaults	磁盘分区的挂载选项，此处通常设置为defaults即可

参数示例	说明
0	<ul style="list-style-type: none">- Linux dump备份选项。<ul style="list-style-type: none">▪ 0表示不使用Linux dump备份。现在通常不使用dump备份，此处设置为0即可。▪ 1表示使用Linux dump备份。
2	<ul style="list-style-type: none">- fsck选项，即开机时是否使用fsck检查磁盘。<ul style="list-style-type: none">▪ 2表示从挂载点为非根目录 (/) 的分区开始检验。▪ 1表示从挂载点为根目录 (/) 的分区开始检验。▪ 0表示不检验。

步骤6 验证自动挂载功能已生效。

```
umount /dev/vdb1
```

```
mount -a
```

则系统会将“/etc/fstab”文件所有内容重新加载。

查询文件系统挂载信息。

```
mount | grep /mnt/sdc
```

回显类似如下信息，说明自动挂载功能生效：

```
root@ecs-test-0001 ~]# mount | grep /mnt/sdc  
/dev/vdb1 on /mnt/sdc type ext4 (rw,relatime,data=ordered)
```

----结束

4.4.4 初始化 windows 数据盘

操作场景

本文介绍使用磁盘管理工具和使用脚本来初始化Windows系统中的一块数据盘，下文示例中所使用的配置如下表所示。不同云服务器的操作系统的初始化操作可能不同，请根据您的实际环境进行操作。

前提条件

云硬盘已挂载到云服务器上。

约束与限制

- 从数据源创建的云硬盘无需初始化。该云硬盘在初始状态就具有数据源中的数据，初始化有丢失数据的风险。如果需要重新初始化，建议先备份云硬盘中的数据。
使用CBR备份请参考[备份云硬盘](#)，使用快照备份请参考[管理云硬盘快照](#)。
- 初始化云硬盘不会删除云硬盘快照，初始化后仍可以使用快照回滚数据至原云硬盘。

手动初始化数据盘

示例说明：对Windows 2019操作系统进行初始化操作，新建一个100GiB的GPT分区，设置文件系统格式为NTFS。

步骤1 登录云服务器。

登录弹性云服务器请参见[登录弹性云服务器](#)。

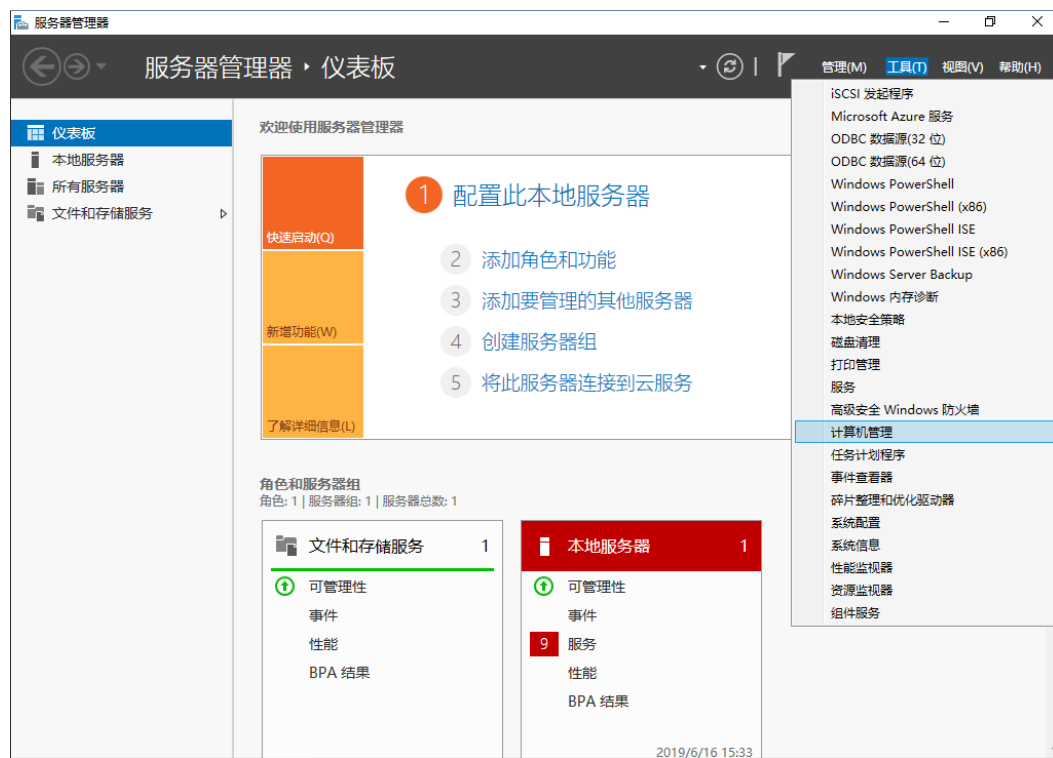
登录裸金属服务器请参见[登录裸金属服务器](#)。

步骤2 在云服务器桌面，单击左下方开始图标。

弹出Windows Server窗口。

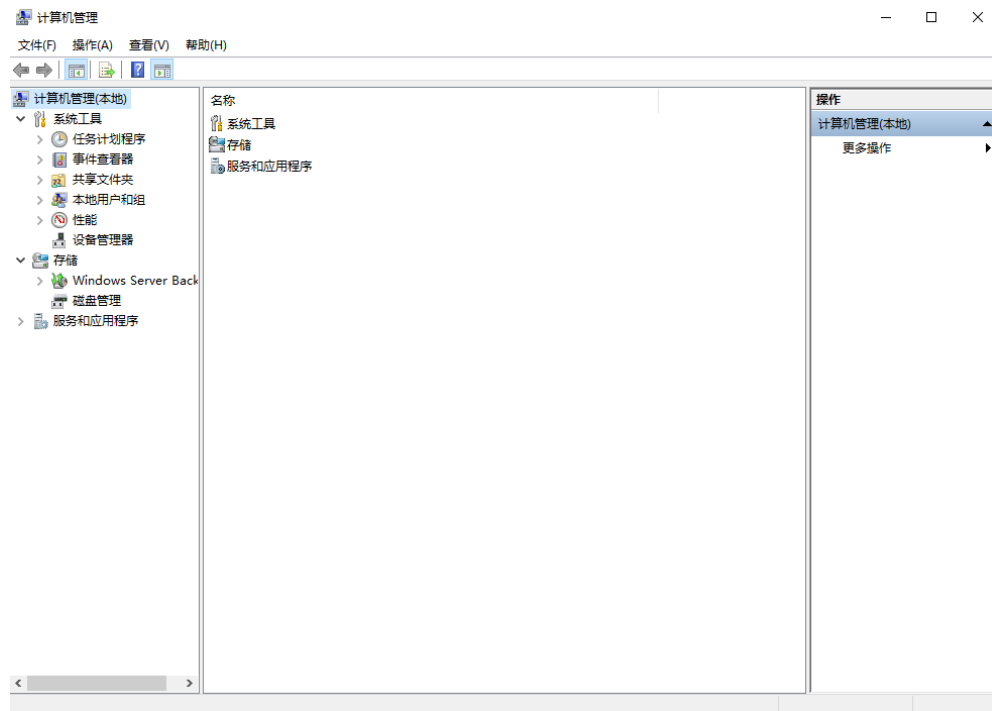
步骤3 单击“服务器管理器”，弹出“服务器管理器”窗口。

图 4-2 服务器管理器



步骤4 “服务器管理器”页面右上方选择“工具 > 计算机管理”，弹出“计算机管理”窗口。

图 4-3 计算机管理



步骤5 选择“存储 > 磁盘管理”。

进入磁盘列表页面，存在未初始化的磁盘时，系统会自动弹出“初始化磁盘”对话框。

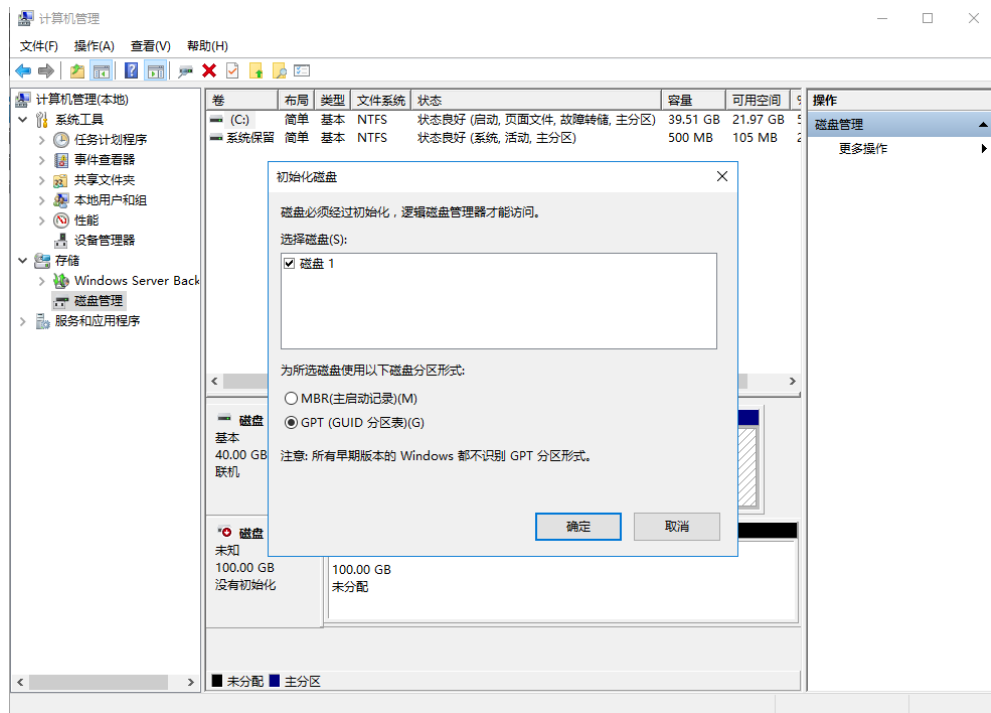
在“初始化磁盘”对话框中显示需要初始化的磁盘，此处以选择“GPT（GUID分区表）”为例，单击“确定”，返回“计算机管理”窗口。

须知

MBR支持的云硬盘最大容量为2 TiB，GPT最大支持的云硬盘容量为18 EiB，当前数据盘支持的最大容量为32 TiB，如果您需要使用大于2 TiB的云硬盘容量，分区形式请采用GPT。

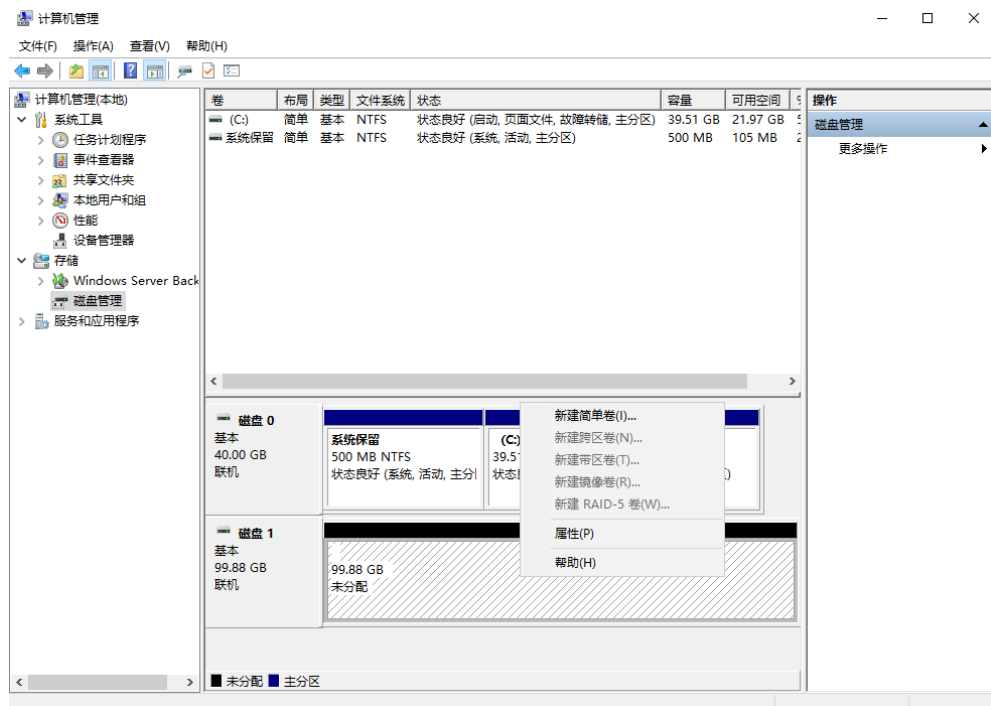
当云硬盘已经投入使用后，此时切换云硬盘分区形式时，云硬盘上的原有数据将会清除，因此请在云硬盘初始化时谨慎选择云硬盘分区形式。切换GPT分区形式前，请对云硬盘数据备份后，再格式化硬盘。

图 4-4 磁盘列表



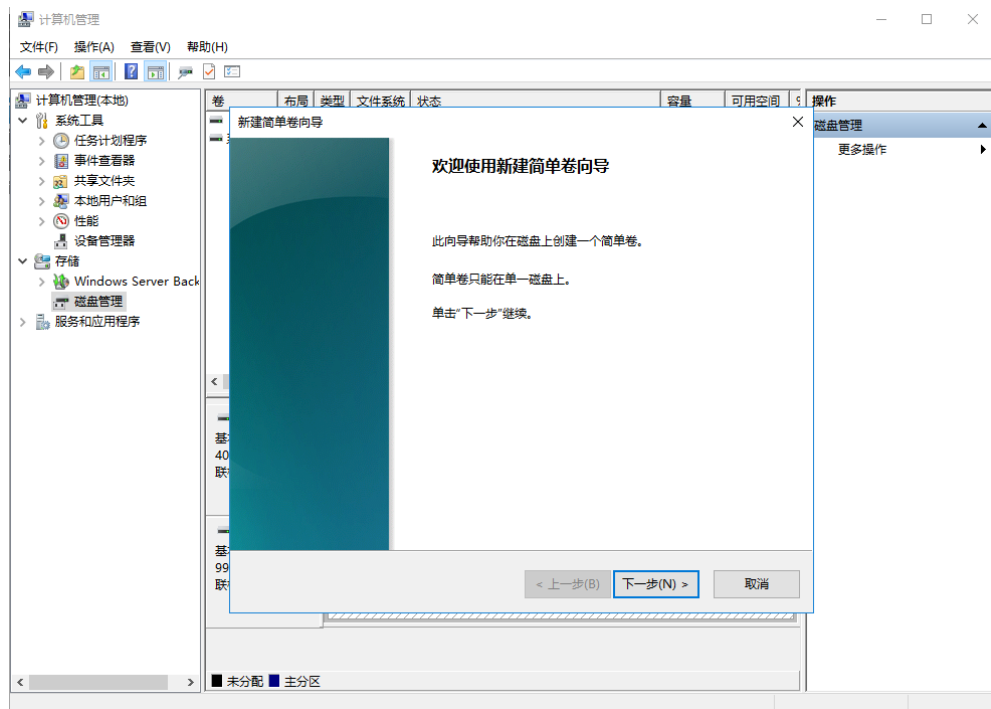
步骤6 在磁盘1右侧的未分配的区域，右键单击选择“新建简单卷”。

图 4-5 计算机管理



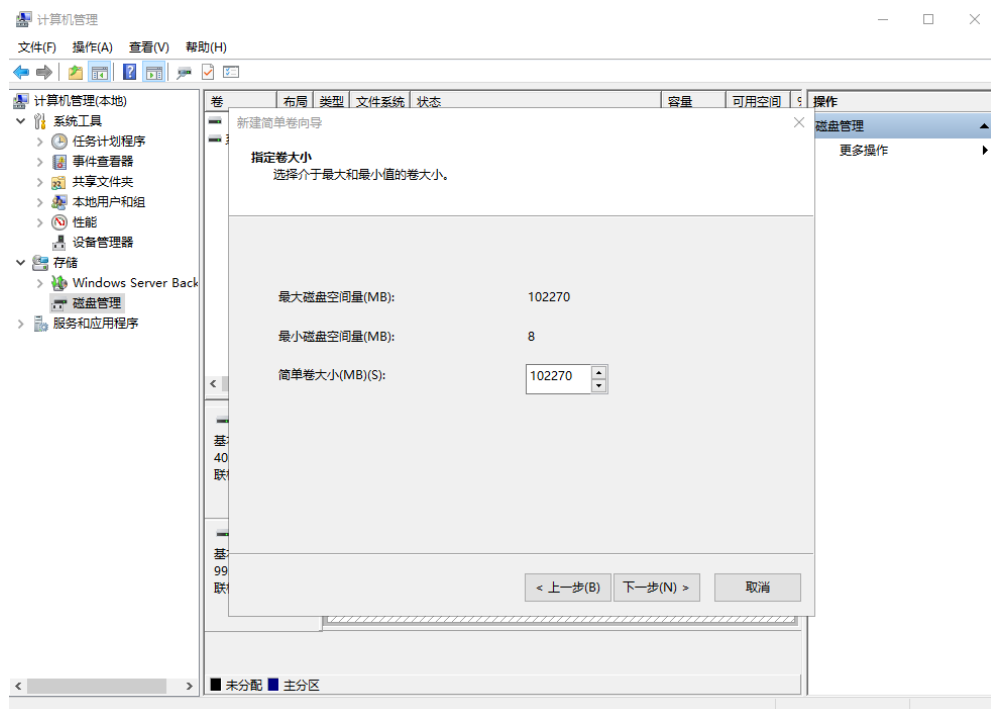
弹出“新建简单卷向导”窗口。

图 4-6 新建简单卷向导



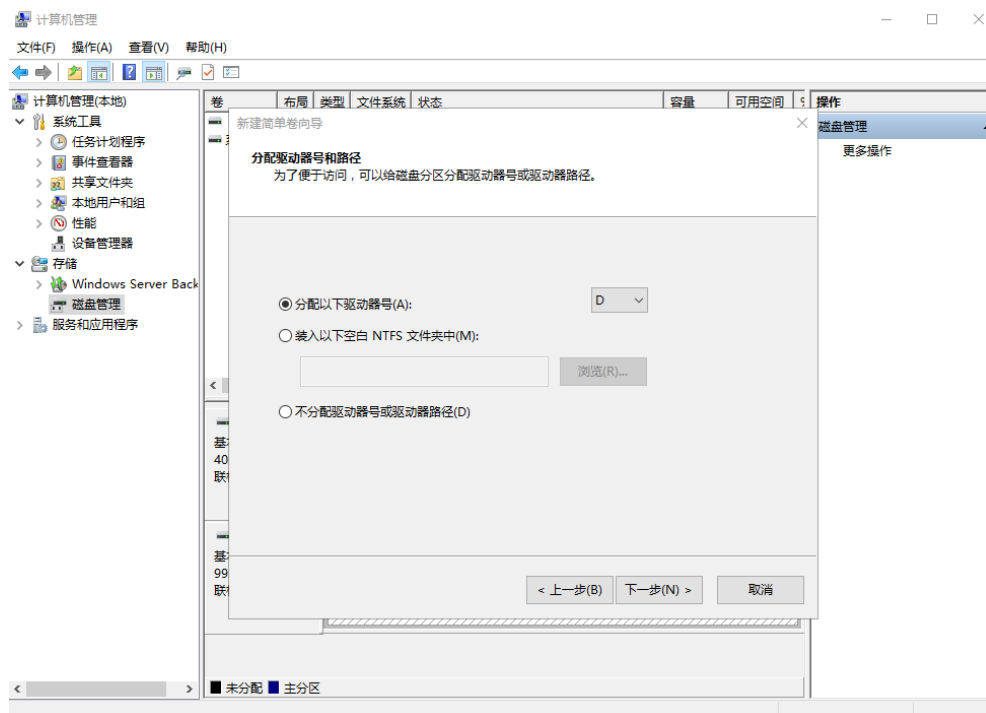
步骤7 根据界面提示，单击“下一步”，进入“指定卷大小”页面。

图 4-7 指定卷大小



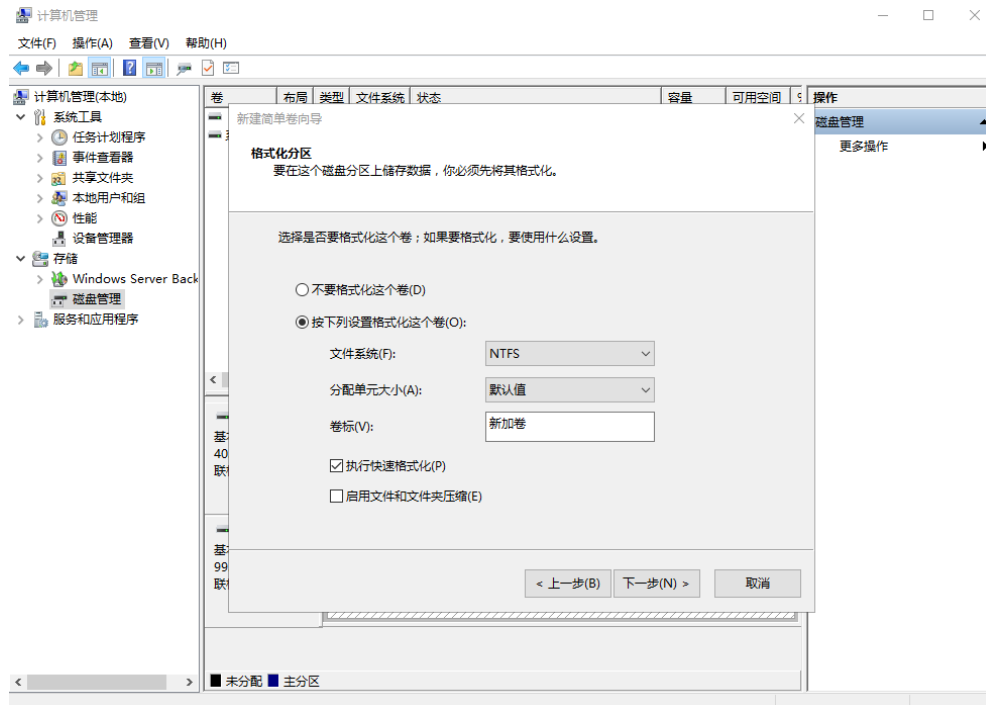
步骤8 指定卷大小，系统默认卷大小为最大值，您还可以根据实际需求指定卷大小，此处以保持系统默认配置为例，单击“下一步”，进入“分配驱动器号和路径”页面。

图 4-8 分配驱动器号和路径



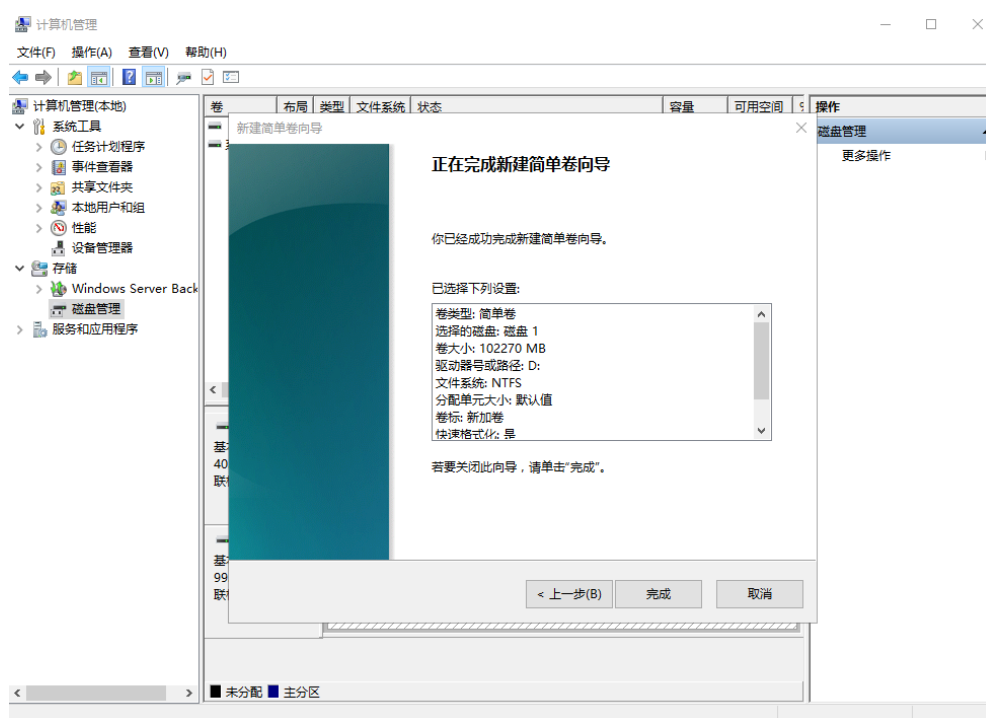
步骤9 分配到驱动器号和路径，系统默认为磁盘分配驱动器号，驱动器号默认为“D”，此处以保持系统默认配置为例，单击“下一步”，进入“格式化分区”页面。

图 4-9 格式化分区



步骤10 格式化分区，系统默认的文件系统为NTFS，并根据实际情况设置其他参数，此处以保持系统默认设置为例，单击“下一步”，进入“完成新建卷”页面。

图 4-10 完成新建卷



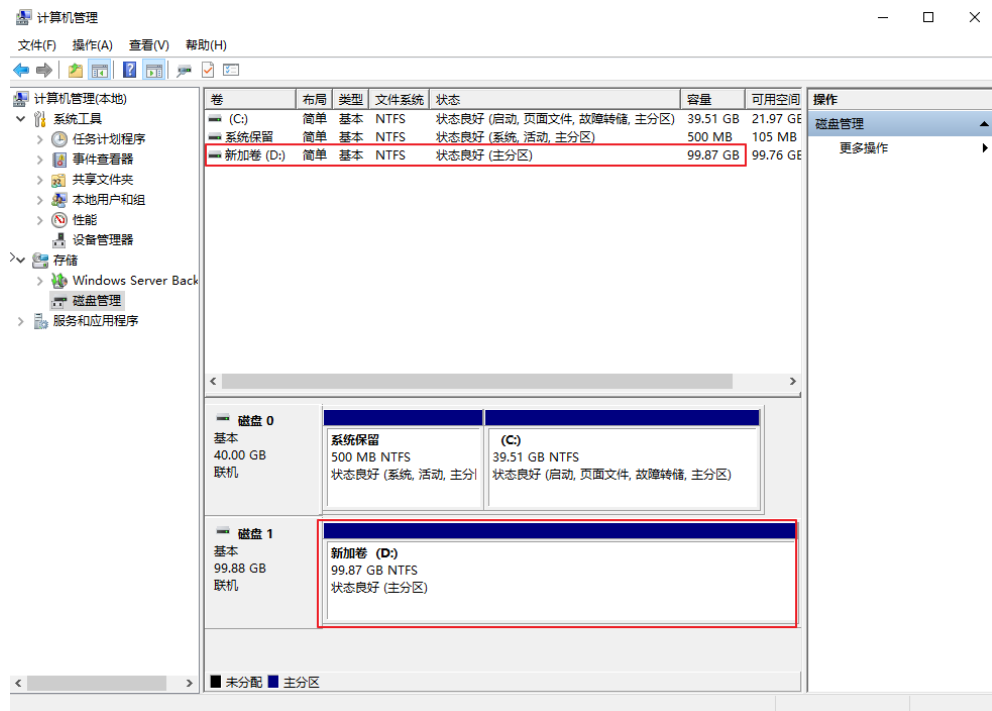
须知

不同文件系统支持的分区大小不同，请根据您的业务需求选择合适的文件系统。

步骤11 单击“完成”。

需要等待片刻让系统完成初始化操作，当卷状态为“状态良好”时，表示初始化磁盘成功。

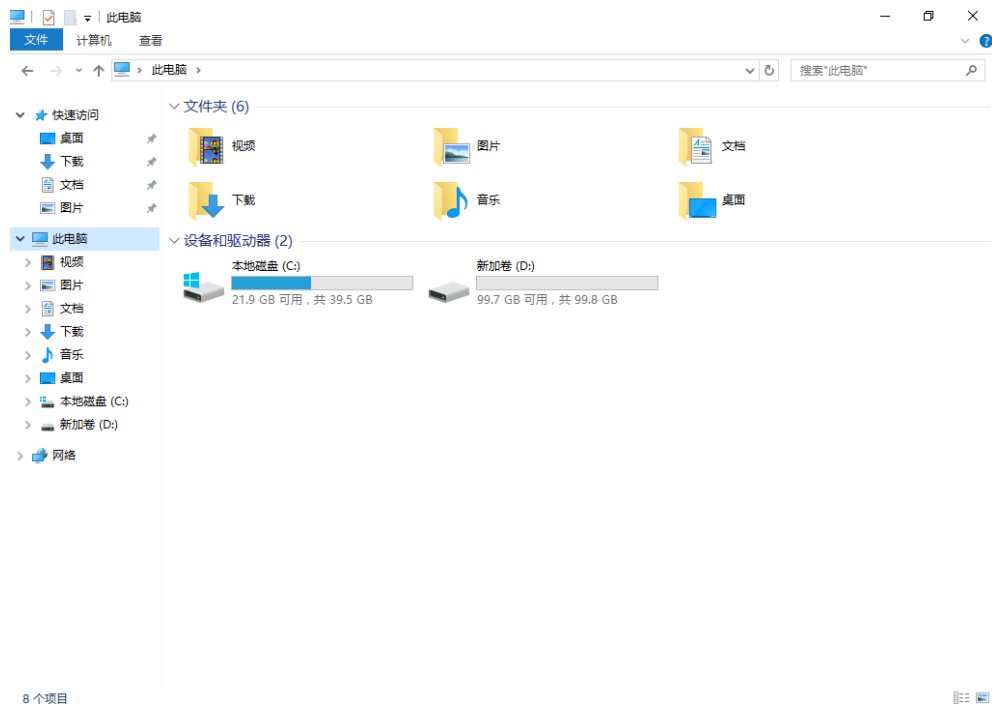
图 4-11 初始化磁盘成功



步骤12 新建卷完成后，单击下方任务栏中，在文件资源管理器中查看是否有新建卷，此处以“新建卷 (D:)”为例。

单击“此电脑”，可以看到“新建卷 (D:)”，表示磁盘初始化成功，任务结束。

图 4-12 文件资源管理器





----结束

4.5 添加“包年/包月”磁盘

操作场景

对于“包年/包月”方式购买的弹性云服务器，支持添加“包年/包月”方式的云硬盘，新添加的云硬盘与弹性云服务器的到期时间相同。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中的右上角，输入弹性云服务器名、IP地址或ID，并进行搜索。
5. 单击待添加云硬盘的“包年/包月”弹性云服务器的名称。
系统跳转至该弹性云服务器详情页面。
6. 选择“云硬盘”页签，并单击“新增磁盘”。
系统跳转至云硬盘购买页面。
7. 根据界面提示，设置新添加云硬盘的参数信息。
8. 单击“立即购买”。
9. 确认云硬盘的配置正确无误后，勾选协议，并单击“提交订单”。
云硬盘购买成功后，会自动添加至指定的“包年/包月”弹性云服务器。

说明

对于9中新增的磁盘，如果卸载后重新挂载，该磁盘只能挂载给原云服务器。

4.6 在线卸载磁盘

操作场景

将挂载在弹性云服务器中的磁盘卸载。

- 对于挂载在系统盘盘位（也就是“/dev/sda”或“/dev/vda”挂载点）上的磁盘，当前仅支持离线卸载。即卸载磁盘时，弹性云服务器需处于“关机”状态。
- 对于挂载在数据盘盘位（非/dev/sda挂载点）上的磁盘，不仅支持离线卸载，在使用部分操作系统时，还支持在线卸载磁盘功能。此时，弹性云服务器处于“运行中”状态。

本节旨在介绍在线卸载弹性云服务器磁盘的使用场景。

约束与限制

- 磁盘需挂载在数据盘盘位，即挂载点为非“/dev/sda”或“/dev/vda”挂载点。如果磁盘挂载在“/dev/sda”或“/dev/vda”挂载点上，则表示该磁盘作为弹性云服务器的系统盘使用，此时不允许在线卸载。

- 在线卸载云硬盘时，对于Windows弹性云服务器，请确认该弹性云服务器已安装UVP VMTools并且正常启用；对于Linux弹性云服务器，可以不安装UVP VMTools。
- 对于Windows弹性云服务器，在线卸载云硬盘前，请确保没有程序正在对该云硬盘进行读写操作。否则，将造成数据丢失。
- 对于Windows弹性云服务器，不支持在线卸载SCSI类型的云硬盘。
- 对于Linux弹性云服务器，在线卸载云硬盘前，客户需要先登录弹性云服务器，执行umount命令，取消待卸载云硬盘与文件系统之间的关联，并确保没有程序正在对该云硬盘进行读写操作。否则，卸载云硬盘将失败。

卸载须知

- 对于Windows弹性云服务器，在线卸载云硬盘时，如果云硬盘处于非“脱机”状态，系统会强制卸载云硬盘。此时，弹性云服务器后台可能会出现xenvbd告警提示，这种情况是正常的。

📖 说明

查看云硬盘状态的方法如下：

1. 选择“开始”，右键单击“计算机”，选择“管理”。
弹出“计算机管理”窗口。
 2. 在左侧导航树中，选择“存储 > 磁盘管理”。
在右侧窗格中出现磁盘列表。
 3. 查看磁盘列表中对应磁盘的“状态”栏。
- 弹性云服务器启动、关机、重启等操作过程中，不建议进行在线卸载云硬盘的操作。
 - 对于[支持在线卸载云硬盘的操作系统](#)以外其他操作系统的弹性云服务器，不建议进行在线卸载云硬盘操作。
 - 对于Linux弹性云服务器，在线卸载云硬盘后重新挂载云硬盘，可能会存在挂载前后盘符发生变化的情况。这是由于Linux系统的盘符分配机制造成的，属于正常情况。
 - 对于Linux弹性云服务器，在线卸载云硬盘后重启弹性云服务器，可能会存在重启前后盘符发生变化的情况。这是由于Linux系统的盘符分配机制造成的，属于正常情况。

支持在线卸载云硬盘的操作系统

支持在线卸载云硬盘的操作系统包括如下两个部分：

- 第一部分请参见[外部镜像文件的镜像格式和操作系统类型](#)。
- 第二部分如[表4-4](#)所示。

表 4-4 支持在线卸载云硬盘的操作系统

操作系统	版本
CentOS	7.3 64bit
	7.2 64bit
	6.8 64bit

操作系统	版本
	6.7 64bit
Debian	8.6.0 64bit
	8.5.0 64bit
Fedora	25 64bit
	24 64bit
SUSE	SUSE Linux Enterprise Server 12 SP2 64bit
	SUSE Linux Enterprise Server 12 SP1 64bit
	SUSE Linux Enterprise Server 11 SP4 64bit
	SUSE Linux Enterprise Server 12 64bit
OpenSUSE	42.2 64bit
	42.1 64bit
Oracle Linux Server release	7.3 64bit
	7.2 64bit
	6.8 64bit
	6.7 64bit
Ubuntu Server	16.04 64bit
	14.04 64bit
	14.04.4 64bit
Windows	Windows Server 2008 R2 Enterprise 64bit
	Windows Server 2012 R2 Standard 64bit
	Windows Server 2016 R2 Standard 64bit
Redhat Linux Enterprise	7.3 64bit
	6.8 64bit

说明

对于其他操作系统的弹性云服务器，请先关机然后再进行卸载磁盘操作，避免由于云硬盘设备与弹性云服务器在线卸载不兼容而产生未知问题。

操作步骤

1. 在弹性云服务器页面，单击待卸载磁盘的弹性云服务器名称，跳转至弹性云服务器详情页。
2. 选择“云硬盘”页签，单击待卸载磁盘所在行的“卸载”，卸载云硬盘。

4.7 扩容云硬盘

操作场景

当您的云硬盘存储容量不足时，您可以通过扩容云硬盘增加弹性云服务器的存储容量。系统盘和数据盘均支持扩容。

操作步骤

扩容云硬盘有如下两种处理方式。

- 申请一块新的云硬盘，并挂载给弹性云服务器。
- 扩容原有云硬盘空间。系统盘和数据盘均支持扩容。
扩容云硬盘的详细步骤，请参见[扩容云硬盘](#)。

扩容云硬盘详细操作请参考[云硬盘扩容概述](#)。

说明

云硬盘扩容成功后，仅扩大了云硬盘的存储容量，还需要登录云服务器自行扩展分区和文件系统。

相关操作

对于Windows弹性云服务器，如果您想通过清理磁盘文件的方式扩大磁盘容量，可以参考[清理 WinSxS 文件夹](#)使用Windows中内置的工具来减小WinSxS文件夹的大小。

4.8 扩容本地盘

操作场景

对于磁盘增强型弹性云服务器，可使用本地盘和云硬盘两类磁盘存储数据。本地盘常用于存放用户的业务数据，磁盘吞吐量更高。

由于磁盘增强型弹性云服务器不支持规格变更，因此，当本地盘容量不足时，需要通过创建更大规格磁盘增强型弹性云服务器的方式进行扩容处理，原磁盘增强型弹性云服务器中保存的数据可以通过云硬盘进行迁移。

操作步骤

1. 根据需要转移的数据量大小，创建云硬盘。
2. 挂载云硬盘。
将步骤1中创建的云硬盘挂载至磁盘增强型弹性云服务器。
3. 备份本地盘数据。
将本地盘中需要转移的数据，备份至当前规格的磁盘增强型弹性云服务器的云硬盘中。
 - 对于Windows系统弹性云服务器，建议您直接复制需要转移的数据至目标云硬盘。

- 对于Linux系统弹性云服务器，建议您使用cp命令复制需要转移的数据至目标云硬盘。
4. 卸载云硬盘。
 - a. 在弹性云服务器页面，选择该规格的磁盘增强型弹性云服务器，确认云服务器为“关机”状态。

如果云服务器为“开机”状态，您可单击“更多 > 关机”将云服务器关机。
 - b. 单击该磁盘增强型弹性云服务器的名称，查看详情。
 - c. 选择“云硬盘”页签，并单击数据盘所在行的“卸载”，卸载该云硬盘。
 5. 准备一台大规格的磁盘增强型弹性云服务器。

该云服务器的规格大于原云服务器规格，且其本地盘容量能够满足用户的需求。
 6. 挂载云硬盘至大规格的磁盘增强型弹性云服务器。

在弹性云服务器页面，单击步骤5中准备的弹性云服务器名称，展开详情。
 7. 选择“云硬盘”页签，并单击“挂载磁盘”。

在“挂载磁盘”对话框中，选择步骤4中卸载的云硬盘，并设置挂载点。
 8. 迁移云硬盘数据。

将步骤7中云硬盘的数据，迁移至大规格磁盘增强型弹性云服务器的本地盘中。



4.9 开启高级磁盘功能

操作场景

- 平台对磁盘功能进行了升级，对于新创建的弹性云服务器，系统支持挂载60块磁盘。但是，对于已有弹性云服务器，磁盘挂载数保持原配额，最多可以挂载24块磁盘（部分类型的弹性云服务器支持挂载40块磁盘），如需支持挂载60块磁盘，需开启高级磁盘功能。
- 开启高级磁盘功能后，支持查询控制台与弹性云服务器内部之间磁盘挂载点的对应关系，查询方法请参见“控制台与弹性云服务器内部之间磁盘挂载点的对应关系”。

本节介绍了弹性云服务器开启高级磁盘功能的方法。

操作步骤

1. 登录控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 单击弹性云服务器名称，进入弹性云服务器详情页。
5. 选择“云硬盘”页签。
6. 查看当前弹性云服务器支持挂载的磁盘数量，并根据页面提示，单击“开启高级磁盘功能，扩大磁盘挂载数量”。

系统弹窗显示“开启高级磁盘功能”页面。
7. 单击“确定”。
8. 关机、并开机弹性云服务器。

使高级磁盘功能生效。

9. 再次进入弹性云服务器详情页，选择“云硬盘”页签，查看界面提示的磁盘挂载数量是否变化。
 - 是，开启成功。
 - 否，开启失败，请重试或联系客服寻求技术支持。

5 弹性网卡管理

5.1 弹性网卡概述

弹性网卡（Elastic Network Interfaces，以下简称ENI）即虚拟网卡，您可以通过创建并配置弹性网卡，并将其附加到您的云服务器实例（包括和裸金属服务器）上，实现灵活、高可用的网络方案配置。

弹性网卡类型

- 主弹性网卡：在创建实例时，随实例默认创建的弹性网卡称为主弹性网卡。无法解除主弹性网卡和实例的绑定关系。
- 扩展弹性网卡：您在弹性网卡控制台创建的是扩展弹性网卡，可以将网卡绑定到实例上，也可以解除网卡和实例的绑定关系。

弹性网卡应用场景

- 灵活迁移
通过将弹性网卡从云服务器实例解绑后再绑定到另外一台服务器实例，保留已绑定私网IP、弹性公网IP和安全组策略，无需重新配置关联关系，将故障实例上的业务流量快速迁移到备用实例，实现服务快速恢复。
- 业务分离管理
可以为服务器实例配置多个分属于同一VPC内不同子网的弹性网卡，特定网卡分别承载云服务器实例的内网、外网、管理网流量。针对子网可独立设置访问安全控制策略与路由策略，弹性网卡也可配置独立安全组策略，从而实现网络隔离与业务流量分离。

弹性网卡的使用限制

- 云服务器可绑定的弹性网卡数量由云服务器实例规格决定，具体请参见[规格清单](#)。

表 5-1 不同类型网卡使用限制

弹性网卡类型	创建网卡	实例绑定	外部通信	网卡与云公共服务区的互通能力
主弹性网卡	随实例默认创建，无法独立创建	无法解除和实例的绑定关系	默认支持和外部正常通信	具备
扩展弹性网卡	可以在弹性网卡控制台单独创建	支持绑定到实例上，也可以解除和实例的绑定关系	需要在弹性云服务器内部配置策略路由来实现和外部正常通信 配置多网卡弹性云服务器的策略路由	不具备

- 扩展弹性网卡不支持直接访问华为云内公共云服务，如内网DNS等，推荐使用VPCEP访问华为云公共云服务，具体参见[购买连接“接口”型终端节点服务的终端节点](#)。

5.2 绑定弹性网卡

操作场景

当您的弹性云服务器需要多个弹性网卡时，可以参考下面步骤为弹性云服务器绑定弹性网卡。

操作步骤



- 登录管理控制台。
- 单击管理控制台左上角的 ，选择区域和项目。
- 单击“”，选择“计算 > 弹性云服务器”。
- 单击待绑定弹性网卡的弹性云服务器名称。
系统跳转至该弹性云服务器详情页面。
- 在“弹性网卡”页签，单击“绑定弹性网卡”。
- 选择不同方式，为弹性云服务器绑定弹性网卡。
 - 选择“选择已有弹性网卡”。
 - （可选）根据“名称”、“ID”或“私有IP”搜索待绑定的弹性网卡。
 - 在弹性网卡列表中，勾选待绑定的弹性网卡。
 - 选择“创建新的弹性网卡”。
根据界面提示，设置待增加弹性网卡的子网和安全组。

图 5-1 选择子网和安全组

绑定弹性网卡

云服务器名称 ecs-...

方式 选择已有弹性网卡 创建新的弹性网卡

所属虚拟私有云 vpc-...

所属子网 subnet-03(192.168.3.0/24)

私有IP地址 自动分配IP地址

选择安全组

<input type="checkbox"/> 安全组名称	描述
<input type="checkbox"/> default	Default security group
<input type="checkbox"/> 通用Web服务器	通用Web服务器，默认放通22、3389、80、443端口和...
<input type="checkbox"/> 通用Web服务器	通用Web服务器，默认放通22、3389、80、443端口和...

- 所属子网：待绑定弹性网卡所属的子网。
 - 私有IP地址：如果需要给弹性云服务器添加一张指定IP地址的网卡，用户需填写“私有IP地址”。
 - 选择安全组：您可以同时勾选多个安全组，此时，弹性云服务器的访问规则遵循几个安全组规则的并集。
7. 单击“确定”。

为保证扩展弹性网卡和外部正常通信，弹性网卡绑定完成后，还需要在云服务器内部为这些弹性网卡配置策略路由。

详细操作，请参见[配置多网卡弹性云服务器的策略路由](#)。

后续任务

部分操作系统无法识别新绑定的弹性网卡，需手动激活弹性网卡。下面以Ubuntu系统为例介绍具体激活弹性网卡的操作步骤，其他操作系统请自行完成相关操作，如有问题，请参见对应操作系统的官网指导或手册来完成操作。

1. 在弹性云服务器所在行的“操作”列下，单击“远程登录”。
登录弹性云服务器。
2. 执行如下命令，查看弹性网卡名称。
ifconfig -a
例如，查询到的弹性网卡名为：eth2。
3. 执行如下命令，进入相应目录。
cd /etc/network
4. 执行如下命令，打开interfaces文件。
vi interfaces
5. 在interfaces文件中，增加类似如下信息。

auto eth2

iface eth2 inet dhcp

6. 执行如下命令，保存并退出interfaces文件。

:wq

7. 执行命令**ifup ethX**或**/etc/init.d/networking restart**，使新增弹性网卡生效。
上述命令中的X为具体的网卡名称序号，例如，**ifup eth2**。
8. 执行如下命令，查看回显信息中是否包括2查询到的弹性网卡。

ifconfig

例如，回显信息中包含弹性网卡eth2。

- 是，表示新增弹性网卡生效，结束。
 - 否，表示新增弹性网卡未生效，执行9。
9. 登录管理控制台，在弹性云服务器所在行的“操作”列下，选择“更多”，并单击“重启”。
 10. 再次执行命令**ifconfig**，查看回显信息中是否包括2查询到的弹性网卡。
 - 是，结束。
 - 否，请联系客服获取技术支持。

5.3 解绑弹性网卡


操作场景

弹性云服务器最多可以有12个弹性网卡，其中包括一个主弹性网卡，且主弹性网卡不可删除。本节操作介绍解绑弹性网卡的操作步骤，当您的弹性云服务器需要解绑主网卡以外的其他弹性网卡时，可以参考以下操作步骤。

注意

解绑弹性网卡前，请谨慎评估该操作可能带来的业务影响，避免由于误操作造成网络中断。

操作步骤

1. 登录管理控制台。
2. 单击“”，选择“计算 > 弹性云服务器”。
3. 在弹性云服务器列表中，单击待解绑弹性网卡的弹性云服务器名称。
系统跳转至该弹性云服务器详情页面。
4. 选择“弹性网卡”页签，并单击待解绑弹性网卡右侧的“解绑”。

说明

对于该弹性云服务器的主弹性网卡（默认为网卡列表中显示的第一个弹性网卡），用户不能执行解绑操作。

5. 在弹出的对话框中单击“是”，解绑弹性网卡。

📖 说明

对于部分弹性云服务器，不支持在线解绑弹性网卡功能，具体以界面显示为准。您需要先关机弹性云服务器，然后再执行解绑弹性网卡操作。

5.4 切换虚拟私有云

操作场景

本节操作介绍为弹性云服务器切换虚拟私有云的操作步骤。

约束限制

- 仅支持处于“运行中”或“关机”状态的弹性云服务器执行此操作。
- 仅支持单网卡切换虚拟私有云。
- 切换虚拟私有云前如果重装/切换过云服务器的操作系统，请先登录云服务器，验证重装/切换时设置的密码或密钥是否注入成功。
 - 如果成功登录云服务器，说明密码或密钥注入完成，可继续执行其他操作。
 - 否则说明系统正在注入密码或密钥信息，在此期间请勿对云服务器执行其他操作。
- 切换虚拟私有云过程中，请勿执行绑定、解绑或更换弹性公网IP等操作，若执行会提示权限不足，无需处理。
- 如果网卡上有IPv6地址，无法切换虚拟私有云。

切换须知

- 切换虚拟私有云支持在开机状态下操作，但是过程中会导致云服务器网络中断。

📖 说明

若弹性云服务器处于开机状态，可能由于网卡有流量访问而导致切换虚拟私有云失败，建议您稍后重试或关机后再进行切换。

- 切换虚拟私有云后，云服务器子网、私有IP地址、MAC地址、操作系统内网卡名称都会发生改变。
- 切换虚拟私有云后，请重新配置源/目的检查和虚拟IP地址。
- 虚拟私有云切换完成后，与网络配置相关的应用软件需要重新配置。与网络相关的服务也需要重新配置，例如ELB、VPN、NAT、流量镜像、DNS等。

前提条件

已准备好待切换的目标VPC、子网、私有IP地址和安全组。

操作步骤

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 弹性云服务器”。
3. 在弹性云服务器列表中，单击“操作”列下的“更多 > 网络/安全组 > 切换VPC”。

系统弹窗显示“切换VPC”页面。

图 5-2 切换虚拟私有云



4. 根据界面提示，在下拉列表中选择可用的虚拟私有云、子网，设置私有IP地址和安全组。

您可以同时勾选多个安全组，此时，弹性云服务器的访问规则遵循几个安全组规则的并集。

说明

使用多个安全组可能会影响弹性云服务器的网络性能，建议您选择安全组的数量不多于5个。

5. 单击“确定”。

5.5 修改私有 IP 地址

操作场景

云平台支持修改主网卡的私有IP地址，具体操作请参见本节内容。如需修改扩展网卡的私有IP地址，请删除网卡，并挂载新网卡。

约束与限制

- 弹性云服务器已关机。
- 如果网卡绑定了虚拟IP或者DNAT规则，需要先解绑。
- 如果网卡上有IPv6地址，无法修改（包括IPv4和IPv6的）私有IP地址。
- 如需修改弹性负载均衡后端服务器的私有IP地址，请先移出后端服务器组后再修改私有IP。

操作步骤

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 弹性云服务器”。
3. 单击待修改私有IP地址的弹性云服务器名称。
系统跳转至该弹性云服务器详情页面。
4. 选择“弹性网卡”页签，并单击主网卡所在行的“修改私有IP地址”。
系统打开“修改私有IP地址”窗口。
5. 请根据需要修改主网卡的“子网”、“私有IP地址”。

📖 说明

只能在同一VPC下更换子网。

如果未填写修改后的“私有IP地址”，系统会自动分配一个新的私有IP地址给主网卡使用。

5.6 管理虚拟 IP 地址

操作场景

虚拟IP地址用于为网卡提供第二个IP地址，同时支持与多个的网卡绑定，从而实现多个弹性云服务器之间的高可用性。

绑定虚拟 IP 地址

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 弹性云服务器”。
3. 在弹性云服务器列表中，单击待绑定虚拟IP地址的弹性云服务器名称。
系统跳转至该弹性云服务器详情页面。
4. 选择“弹性网卡”页签，单击“管理虚拟IP地址”。
5. 选择“IP地址管理”页签，在需要绑定弹性公网IP或者弹性云服务器的虚拟IP地址所在行的操作列下，单击“绑定弹性公网IP”或者“绑定服务器”。
多个主备部署的弹性云服务器可以在绑定虚拟IP地址时选择同一个虚拟IP地址，增强容灾性能。
6. 单击“确定”。

登录 ECS 配置虚拟 IP 地址

当为一个ECS绑定一个虚拟IP或者多个虚拟IP时，在控制台执行完绑定虚拟IP的操作后，您还需要参考以下章节，登录弹性云服务器手工配置虚拟IP地址。

本文提供以下操作系统的配置示例，其他操作系统，请您参考对应官网帮助文档进行配置。

- Linux系统：CentOS 7.2 64bit、Ubuntu 22.04 server 64bit
- Windows系统：Windows Server

Linux 系统（CentOS）

以下操作以“CentOS 7.2 64bit”为例，供您参考。

1. 执行以下命令，查看并记录需要绑定虚拟IP的网卡及对应连接。

nmcli connection

回显类似如下信息：

```
[172.16.0.247_subnet0-ecs-pod6-gaea-dpdk-ipv6 ~]#nmcli connection
NAME                UUID                                TYPE      DEVICE
Wired connection 1  5e72ec5a-6165-3bd6-a34b-ce43981acb27  ethernet  eth0
docker0             cd351a91-c5eb-4b69-83eb-df092a2ccf6b  bridge    docker0
```

本示例的回显信息说明如下：

- **DEVICE**列的eth0为需要绑定虚拟IP的网卡。
- **NAME**列的Wired connection 1为网卡对应的连接。

2. 执行以下命令，在目标网卡连接中添加虚拟IP。

nmcli connection modify "网卡对应的连接名称" +ipv4.addresses 虚拟IP地址

参数说明如下：

- 网卡对应的连接名称：为1中查到的网卡对应的连接，本示例中为Wired connection 1。
- 虚拟IP地址：待添加的虚拟IP地址，如果一次添加多个虚拟IP地址，多个虚拟IP地址之间用“,” 隔开。

命令示例：

- 添加单个虚拟IP：**nmcli connection modify "Wired connection 1" +ipv4.addresses 172.16.0.125**
- 添加多个虚拟IP：**nmcli connection modify "Wired connection 1" +ipv4.addresses 172.16.0.125,172.16.0.126**

3. 执行以下命令，使2的配置生效。

nmcli connection up "网卡对应的连接名称"

命令示例：

nmcli connection up "Wired connection 1"

回显类似如下信息：

```
[172.16.0.247_subnet0-ecs-pod6-gaea-dpdk-ipv6 ~]#nmcli connection up "Wired connection 1"
Connection successfully activated (D-Bus active path: /org/freedesktop/NetworkManager/ActiveConnection/6)
```

4. 执行以下命令，检查虚拟IP配置是否成功。

ip a

回显类似如下信息，可以看到eth0网卡下存在虚拟IP地址，为172.16.0.125。

```
[172.16.0.247_subnet0-ecs-pod6-gaea-dpdk-ipv6 ~]#ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state UP group default qlen 1000
    link/ether fa:16:3e:e5:d5:cd brd ff:ff:ff:ff:ff:ff
    inet 172.16.0.247/24 brd 172.16.0.255 scope global noprefixroute dynamic eth0
        valid_lft 86398sec preferred_lft 86398sec
    inet 172.16.0.125/32 brd 172.16.0.125 scope global noprefixroute eth0
        valid_lft forever preferred_lft forever
    inet6 2001:db8:a583:62c:7d43:a19a:4031:d6fb/128 scope global tentative noprefixroute dynamic
        valid_lft 86400sec preferred_lft 86400sec
    inet6 fe80::5371:9bf9:b652:e35b/64 scope link noprefixroute
        valid_lft forever preferred_lft forever
```

📖 说明

如果您需要删除已添加的虚拟IP，可以使用以下方法：

1. 在目标网卡连接中删除虚拟IP。

nmcli connection modify "网卡对应的连接名称" -ipv4.addresses 虚拟IP地址

一次删除多个虚拟IP地址时，多个IP之间用“,” 隔开，命令示例：

- 删除单个虚拟IP：**nmcli connection modify "Wired connection 1" -ipv4.addresses 172.16.0.125**
- 删除多个虚拟IP：**nmcli connection modify "Wired connection 1" -ipv4.addresses 172.16.0.125,172.16.0.126**

2. 参考3，使删除操作生效。

Linux 系统（Ubuntu）

以下操作以“Ubuntu 22.04 server 64bit”为例，当弹性云服务器的操作系统为Ubuntu 22和Ubuntu 20时，请您参考下述方法。

1. 执行以下命令，查看并记录需要绑定虚拟IP的网卡。

ifconfig

回显类似如下信息，本示例中绑定虚拟IP的网卡名称为eth0。

```
root@ecs-X-ubuntu:~# ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST> mtu 1500
    inet 172.16.0.210 netmask 255.255.255.0 broadcast 172.16.0.255
    inet6 fe80::f816:3eff:fe01:f1c3 prefixlen 64 scopeid 0x20<link>
    ether fa:16:3e:01:f1:c3 txqueuelen 1000 (Ethernet)
    RX packets 43915 bytes 63606486 (63.6 MB)
    RX errors 0 dropped 0 overruns 0 frame 0
    TX packets 3364 bytes 455617 (455.6 KB)
    TX errors 0 dropped 0 overruns 0 carrier 0 collisions 0
...
```

2. 执行以下命令，进入“/etc/netplan”目录。

cd /etc/netplan

3. 执行以下命令，为目标网卡添加虚拟IP地址。

- a. 执行以下命令，打开配置文件“01-netcfg.yaml”。

vim 01-netcfg.yaml

- b. 按i进入编辑模式。

- c. 在对应网卡配置区域内，添加虚拟IP地址。

本示例为eth0添加虚拟IP地址，待添加内容如下：

addresses:

- 172.16.0.26/32

添加后文件内容如下：

```
network:
  version: 2
  renderer: NetworkManager
  ethernets:
    eth0:
      dhcp4: true
      addresses:
        - 172.16.0.26/32
    eth1:
      dhcp4: true
    eth2:
      dhcp4: true
    eth3:
```



```
dhcp4: true
eth4:
  dhcp4: true
```

d. 添加完成后，按“ESC”，并输入“:wq!”，保存后退出文件。

4. 执行以下命令，使3的配置生效。

```
netplan apply
```

5. 执行以下命令，检查虚拟IP配置是否成功。

```
ip a
```

回显类似如下信息，可以看到eth0网卡下存在虚拟IP地址，为**172.16.0.26**。

```
root@ecs-X-ubuntu:/etc/netplan# ip a
...
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc fq_codel state UP group default
qlen 1000
    link/ether fa:16:3e:01:f1:c3 brd ff:ff:ff:ff:ff:ff
    altname enp0s3
    altname ens3
    inet 172.16.0.26/32 scope global noprefixroute eth0
        valid_lft forever preferred_lft forever
    inet 172.16.0.210/24 brd 172.16.0.255 scope global dynamic noprefixroute eth0
        valid_lft 107999971sec preferred_lft 107999971sec
    inet6 fe80::f816:3eff:fe01:f1c3/64 scope link
        valid_lft forever preferred_lft forever
```

说明

如果您需要删除已添加的虚拟IP，可以使用以下方法：

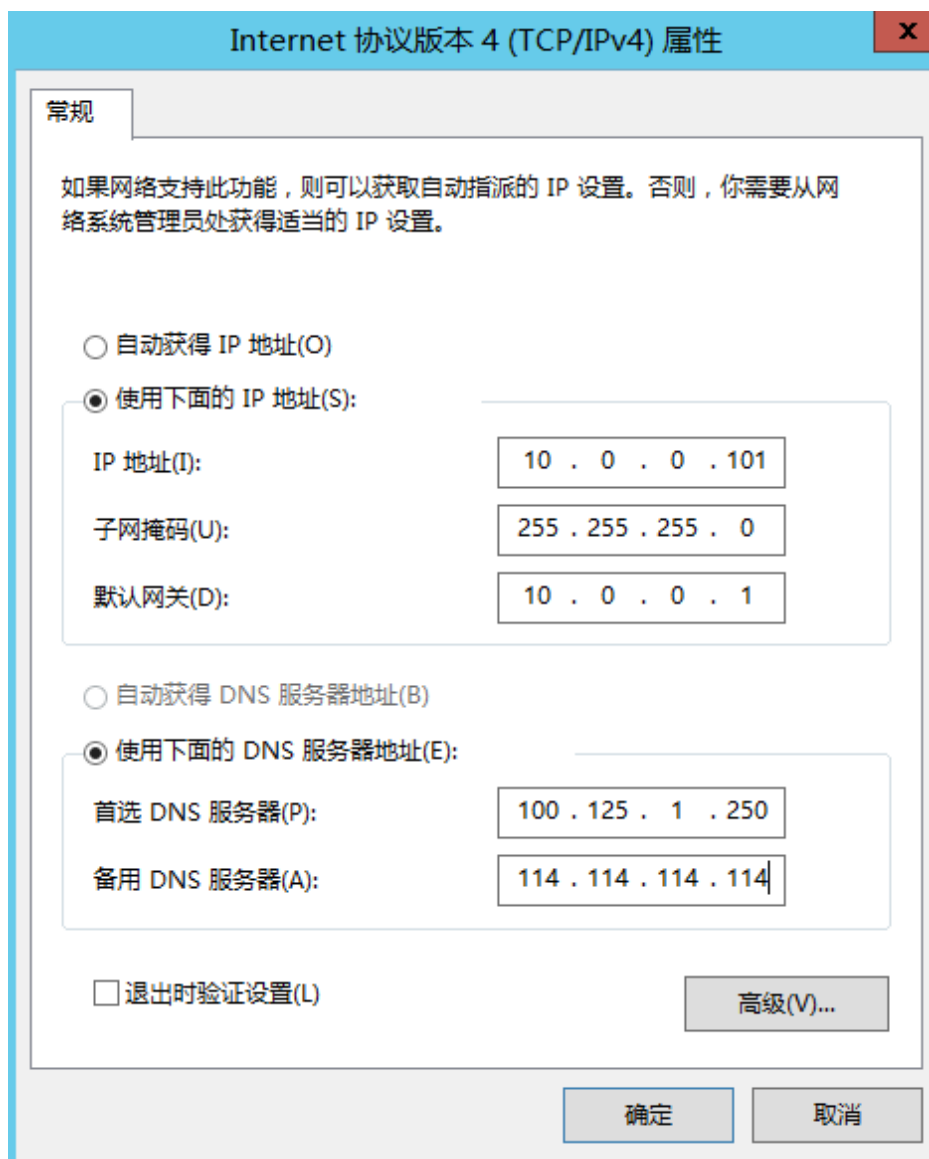
1. 参考3，打开配置文件“01-netcfg.yaml”，并删除对应网卡下虚拟IP的地址。
2. 参考4，使删除操作生效。

Windows 系统（Windows Server）

以下操作以“Windows Server”为例，供您参考。

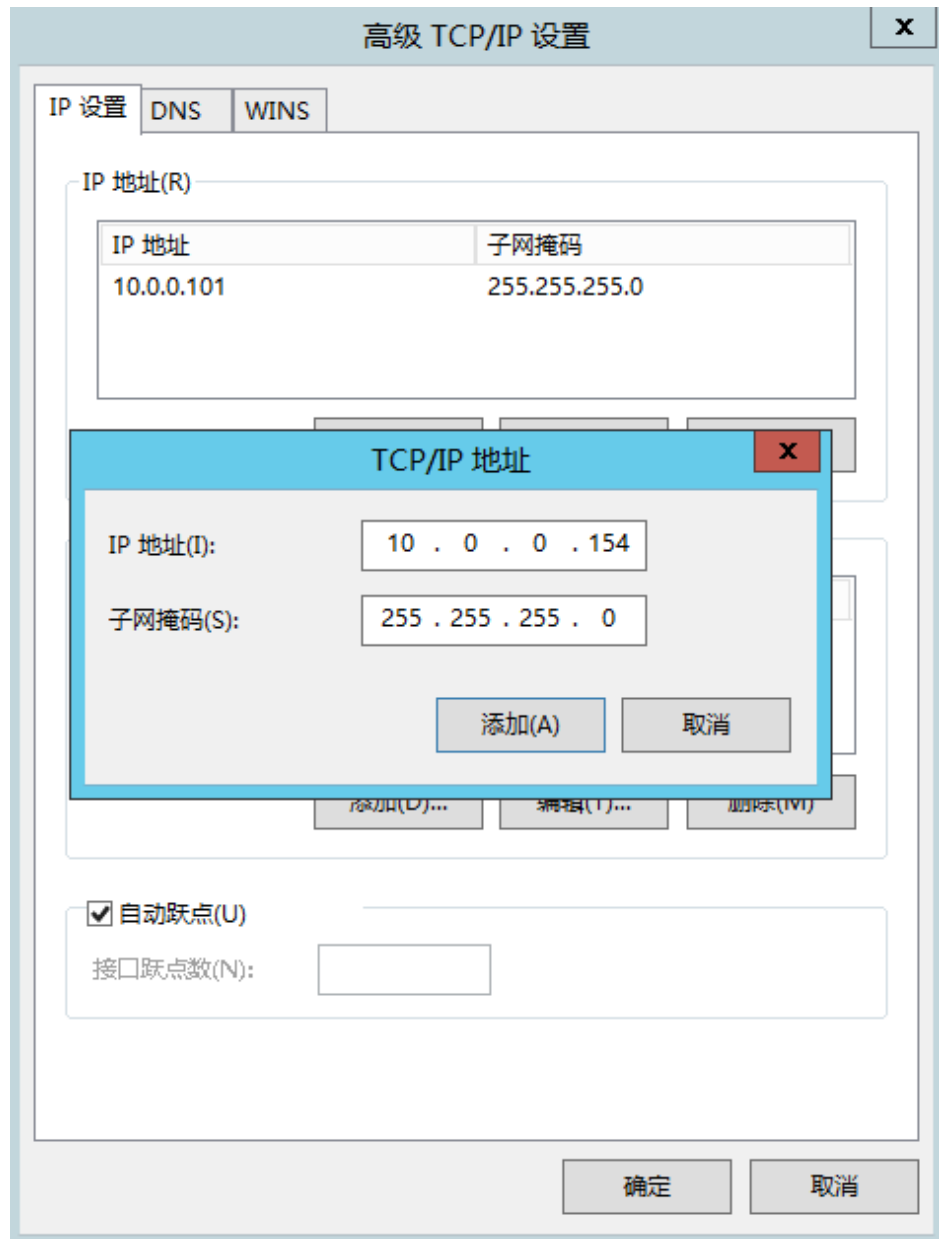
1. 在“控制面板 > 网络和共享中心”路径下，单击对应的本地连接。
2. 在打开的本地连接页面中，单击“属性”。
3. 在“网络”页签中选择“Internet 协议版本 4（TCP/IPv4）”。
4. 单击“属性”。
5. 选择“使用下面的IP地址”，IP地址配置为弹性云服务器的私有IP地址，例如：10.0.0.101。

图 5-3 配置私有 IP 地址



6. 单击“高级”。
7. 在“IP设置”页签内“IP地址”区域，单击“添加”。
添加虚拟IP地址，例如：10.0.0.154。

图 5-4 配置虚拟 IP 地址



- 单击“确定”，保存更改。
- 在“开始”菜单中打开Windows命令行窗口，执行以下命令确认是否配置了虚拟IP地址。

ipconfig /all

回显样例中IPv4 Address包含虚拟IP地址10.0.0.154，表示弹性云服务器内部网卡的虚拟IP地址配置正常。

5.7 开启网卡多队列功能

操作场景

随着网络IO的带宽不断提升，单核CPU处理网络中断存在瓶颈，不能完全满足网卡的需求，通过开启网卡多队列功能，您可以将弹性云服务器中的网卡中断分散给不同的CPU处理，以满足网卡的需求，从而提升网络PPS和带宽性能。

假设以下场景所述的弹性云服务器满足规格和虚拟化类型要求：

- 使用[网卡多队列支持列表](#)中的公共镜像创建的弹性云服务器，默认已开启网卡多队列，无需执行本节操作。
- 对于私有镜像场景，如果您的外部镜像文件的操作系统在[网卡多队列支持列表](#)范围内，需要按照如下流程开启网卡多队列：
 - a. [将外部镜像文件导入镜像服务控制台](#)。
 - b. [为私有镜像添加网卡多队列标签](#)。
 - c. [使用私有镜像创建弹性云服务器](#)。
 - d. [执行网卡多队列的配置脚本](#)。

说明

云服务器开启网卡多队列功能后，如果后续有新增或删除网卡，切换VPC等操作，需要重新对云服务器设置网卡多队列，详细操作请参考[执行网卡多队列的配置脚本](#)。

网卡多队列支持列表

网卡多队列的支持情况和实例规格、虚拟化类型、镜像的操作系统有关，只有同时满足这些要求，弹性云服务器才能开启网卡多队列功能。

- 支持网卡多队列的实例规格请参见“[实例类型](#)”。

说明

网卡多队列数为大于1的值，表示支持网卡多队列。

- 虚拟化类型必须为KVM。
- [表5-3](#)所列的Linux公共镜像，支持网卡多队列。

说明

- Windows操作系统弹性云服务器的驱动pvdriver有一套动态调整网卡队列数的策略，会根据vCPU数计算队列数并进行动态调整，因此无需对windows网卡多队列的队列数进行设置。
- Windows 2008公共镜像已下线，可通过私有镜像支持。
- Linux操作系统弹性云服务器建议将操作系统内核版本升级至2.6.35及以上，否则不支持网卡多队列。
建议您使用命令**uname -r**查询内核版本，如果低于2.6.35请联系技术支持升级内核。

表 5-2 Windows 弹性云服务器网卡多队列支持列表

镜像	是否支持多队列	是否默认开启多队列
Windows Server 2008 R2 Standard/Enterprise/DataCenter 64bit	是	是
Windows Server 2008 Enterprise SP2 64bit	是	是
Windows Server 2008 Web R2 64 bit	是	是
Windows Server 2008 R2 Enterprise 64bit_WithGPUdriver	是	是
Windows Server 2012 R2 Standard 64bit_WithGPUdriver	是	是
Windows Server 2012 R2 Standard/DataCenter 64 bit	是	是
Windows Server 2016 Standard/DataCenter 64 bit	是	是
Windows Server 2019 DataCenter 64 bit	是	是

表 5-3 Linux 弹性云服务器网卡多队列支持列表

镜像	是否支持多队列	是否默认开启多队列
Ubuntu 14.04/16.04/18.04/20.04 server 64bit	是	是
OpenSUSE 42.2/15.* 64bit	是	是
SUSE Enterprise 12 SP1/SP2 64bit	是	是
CentOS 6.8/6.9/7.* /8.* 64bit	是	是
Debian 8.0.0/8.8.0/8.9.0/9.0.0/10.0.0/10.2.0 64bit	是	是
Fedora 24/25/30 64bit	是	是
EulerOS 2.2/2.3/2.5 64bit	是	是

将外部镜像文件导入镜像服务控制台

将外部镜像文件导入镜像服务控制台，详细操作请参见《镜像服务用户指南》中“注册镜像”章节。导入完成后，在镜像详情页查看“网卡多队列”参数取值。

- 如果为“支持”，执行[使用私有镜像创建弹性云服务器](#)。
- 如果为“不支持”，执行[为镜像添加网卡多队列标签](#)。

为镜像添加网卡多队列标签

Windows操作系统暂未商用支持网卡多队列，如果对Windows操作系统镜像添加网卡多队列标签，开启网卡多队列功能，可能会引起操作系统启动速度变慢等问题。

用户可以选择以下任一种方式设置镜像的网卡多队列属性。

方式1:

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 镜像服务”。
3. 单击“私有镜像”页签，在对应镜像所在行的“操作”列下，单击“修改”。
4. 设置镜像的网卡多队列属性。

方式2:

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 镜像服务”。
3. 单击“私有镜像”页签，在镜像列表中，单击镜像名称，进入镜像详情页面。
4. 单击右上角的“修改”，在弹出的“修改镜像”对话框中，设置镜像的网卡多队列属性。

方法3: 通过API为镜像添加网卡多队列标签hw_vif_multiqueue_enabled

1. 获取Token的方法请参考“[认证鉴权](#)”。
2. “更新镜像信息”的API使用方法请参考“[更新镜像信息（OpenStack原生）](#)”。
3. 在请求消息头中增加“X-Auth-Token”。

“X-Auth-Token”的取值为步骤1中获取的Token。

4. 在请求消息头中增加“Content-Type”。

“Content-Type”取值为application/openstack-images-v2.1-json-patch

请求的URI格式为:

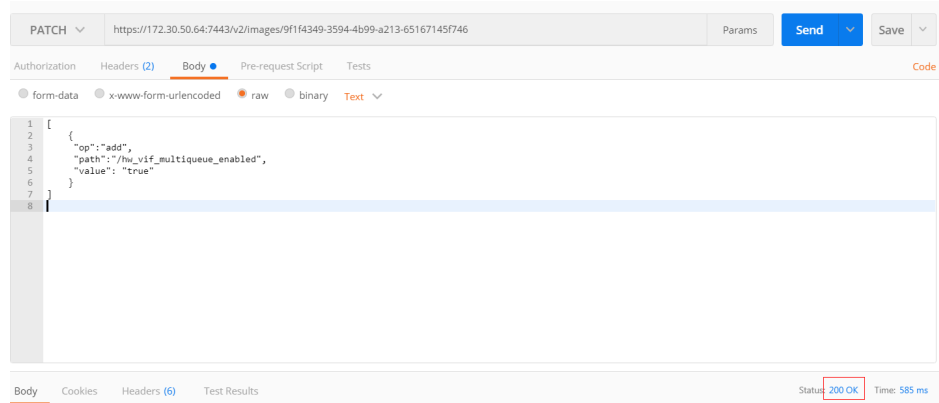
```
PATCH /v2/images/{image_id}
```

请求的body体如下所示。

```
[
  {
    "op": "add",
    "path": "/hw_vif_multiqueue_enabled",
    "value": "true"
  }
]
```

支持网卡多队列修改样例如[图5-5](#)所示。

图 5-5 支持网卡多队列修改样例



使用私有镜像创建弹性云服务器

使用注册好的私有镜像创建云服务器。在配置参数时，需要注意以下两点：

- 区域：必须选择私有镜像所在的区域。
- 镜像：选择“私有镜像”，并在下拉列表中选择需要的镜像。

执行网卡多队列的配置脚本

Windows操作系统使用的半虚拟化驱动pvdriver有一套动态调整网卡队列数的策略，会根据vCPU数计算队列数并进行动态调整，因此无需对windows网卡多队列的队列数进行设置。

Linux操作系统提供了自动开启网卡多队列功能的配置脚本，配置成功后，弹性云服务器支持网卡多队列功能。

说明

自动开启网卡多队列功能的配置脚本当前仅支持处理eth0类型的网卡。

1. 登录弹性云服务器，查看网卡支持和已开启的队列数。

ethtool -l 网卡

示例：

```
[root@localhost ~]# ethtool -l eth0 #查询网卡eth0的队列数
Channel parameters for eth0:
Pre-set maximums:
RX:          0
TX:          0
Other:       0
Combined:    4 #表示此网卡最多支持设置4个队列
Current hardware settings:
RX:          0
TX:          0
Other:       0
Combined:    1 #表示当前已开启的是1个队列
```

如果返回信息中，两个“Combined”字段取值相同，则表示网卡已开启多队列，无需执行以下操作。

2. 执行以下命令，下载配置脚本“multi-queue-hw”。

wget 下载地址URL

下载地址：<https://ecs-instance-driver.obs.cn-north-1.myhuaweicloud.com/multi-queue-hw>

3. 执行以下命令，添加执行权限。
chmod +x multi-queue-hw
4. 执行以下命令，将脚本“multi-queue-hw”放到目录/etc/init.d下。
mv multi-queue-hw /etc/init.d
5. 执行以下命令，运行脚本“multi-queue-hw”。
/etc/init.d/multi-queue-hw start
运行脚本后，立即生效。但关机后，网卡多队列功能将自动失效。
6. 为了使网卡多队列功能开机自动生效，各个OS需要增加开机启动配置：
 - CentOS/Redhat/Fedora/EulerOS/Suse/OpenSuse使用如下命令，增加开机启动项，使网卡多队列配置开机：
chkconfig multi-queue-hw on
 - Ubuntu使用如下命令，增加开机启动项：
update-rc.d multi-queue-hw defaults 90 10
 - Debian使用如下命令，增加开机启动项：
systemctl enable multi-queue-hw

查看网卡队列数

开启网卡多队列功能

1. 登录弹性云服务器。
2. 执行以下命令，查看网卡支持和已开启的队列数。

ethtool -l 网卡

示例：

```
[root@localhost ~]# ethtool -l eth0 #查询网卡eth0的队列数
Channel parameters for eth0:
Pre-set maximums:
RX:          0
TX:          0
Other:       0
Combined: 4 #表示此网卡最多支持设置开启4个队列
Current hardware settings:
RX:          0
TX:          0
Other:       0
Combined: 4 #表示当前开启的是4个队列
```

5.8 开启网卡 IPv6 功能

操作场景

IPv4/IPv6双栈网络，支持为ECS提供两个版本的IP地址，包括IPv4地址和IPv6地址。

您可以在购买ECS时，选择支持IPv6的规格和已开启IPv6的主网卡，这样在ECS购买成功后，即可同时拥有IPv4地址和IPv6地址。该场景下的详细操作，请参见[网络：虚拟私有云、主网卡、扩展网卡、源/目的检查](#)。

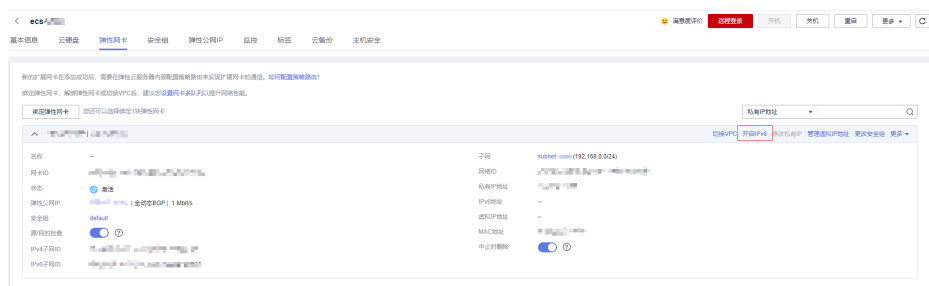
如果购买ECS时未开启IPv6，则ECS不支持在IPv4/IPv6的双栈网络模式下运行，ECS仅拥有IPv4地址。此时，您可以通过“开启IPv6”功能为已创建的ECS开启IPv6功能。

本文介绍如何为已创建ECS开启IPv6功能。

约束与限制

- 请确保ECS的规格支持创建后开启IPv6功能。
当前仅v7及v7以上代系版本的规格支持创建后开启IPv6功能，例如C7、M7等。
您可以在ECS详情页的“弹性网卡”页签进行查看，如果右上角有“开启IPv6”，则表示支持创建后开启IPv6功能。

图 5-6 开启 IPv6



- 同一个网卡上，只能绑定一个IPv6地址。

开启子网的 IPv6 网段

说明

云服务器所属子网的IPv6功能开启后会自动分配IPv6网段，开启后不允许关闭。

- 登录管理控制台，单击**弹性云服务器**列表页。
- 单击待开启IPv6的ECS的名称，进入ECS详情页。
- 在“基本信息”页签的“网卡”区域，单击子网名称，进入子网详情页。
- 在子网详情页的“基本信息”页签，单击“开启IPv6”。

图 5-7 开启子网的 IPv6 网段

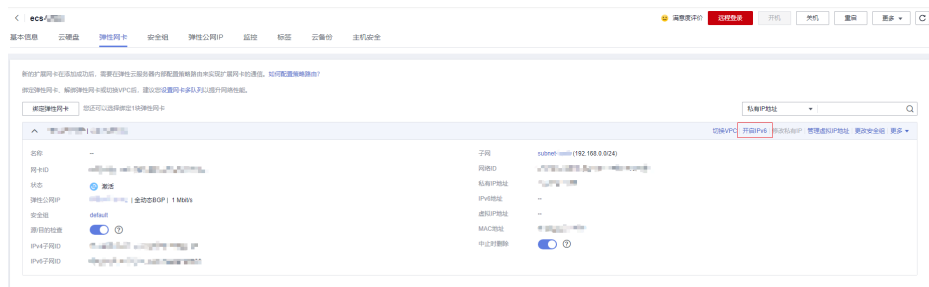


- 单击“确定”，完成子网IPv6网段的开启。

开启 ECS 网卡的 IPv6 功能

- 返回ECS详情页面。
- 在“弹性网卡”页签，单击网卡折叠面板右上角的“开启IPv6”。

图 5-8 开启 IPv6



说明

- 仅当云服务器规格支持创建后开启IPv6功能时，显示“开启IPv6/关闭IPv6”。您可以参考[约束与限制](#)判断云服务器规格是否支持创建后开启IPv6。
 - 开启IPv6后，如果云服务器未获取IPv6地址，建议您重启云服务器后重新查看，或参考[动态获取IPv6地址](#)进行配置。
 - 开启IPv6后，如需实现IPv6公网通信，需要创建EIP共享带宽，并将云服务器的IPv6地址添加至共享带宽。详细内容，请参见[通过VPC快速搭建IPv4/IPv6双栈网络](#)。
 - 若不再使用弹性云服务器的IPv6功能，可以在当前页面单击“关闭IPv6”，关闭网卡的IPv6功能，关闭后该网卡的“IPv6地址”为空。
 - 关闭IPv6后，如果重新开启弹性云服务器的IPv6，在重启云服务器后，需要登录弹性云服务器手动清理IPv6缓存，并重新请求获取IPv6地址。
3. 单击“确定”，完成ECS网卡IPv6功能的开启。

5.9 动态获取 IPv6 地址

操作场景

IPv6的使用，可以有效弥补IPv4网络地址资源有限的问题。如果当前云服务器使用IPv4，那么启用IPv6后，云服务器可在双栈模式下运行，即云服务器可以拥有两个不同版本的IP地址：IPv4地址和IPv6地址，这两个IP地址都可以进行内网/公网访问。

按照[约束与限制](#)中的网络环境要求创建的云服务器，有些不能动态获取到IPv6地址，需要进行相关配置才行。如果云服务器使用的是公共镜像，则支持情况如下：

- Windows公共镜像默认已开启IPv6动态获取功能，无需配置，文中的[Windows 2012操作系统](#)和[Windows 2008操作系统](#)部分供您验证、参考。
- Linux公共镜像开启动态获取IPv6功能时，需要先判断是否支持IPv6协议栈，再判断是否已开启动态获取IPv6。目前，所有Linux公共镜像均已支持IPv6协议栈。

约束与限制

- 请确保云服务器所在的子网已开启IPv6功能。
若云服务器所在子网未开启IPv6功能，需参考[开启云服务器的IPv6功能](#)进行开启，开启后不允许关闭。
- 请确保云服务器规格支持IPv6功能。
不同区域、不同可用区支持IPv6双栈的云服务器规格不同。ECS是否支持IPv6双栈，请选择区域、可用区后，以控制台的显示为准，查询方法如下图所示。

图 5-9 查询支持 IPv6 的 ECS 规格



当ECS规格列表中包含“IPv6”参数，且取值为“是”时，表示该ECS规格支持IPv6。

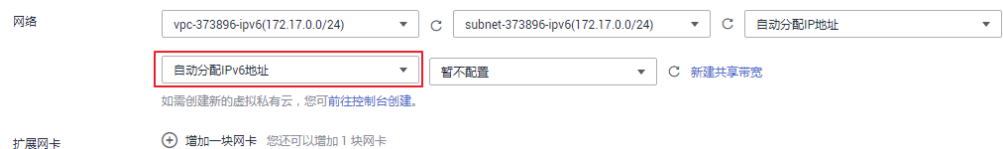
说明

规格是否支持IPv6由“可用区”和“规格”两个参数决定。

如果设置“可用区”后，规格列表中不显示“IPv6”参数或参数值为“否”，表示当前规格不支持IPv6。

- 请确保创建云服务器时已选择“自动分配IPv6地址”。

图 5-10 选择“自动分配 IPv6 地址”



- 云服务器启动之后动态插拔的网卡不支持IPv6地址动态获取功能。
- 仅弹性云服务器支持IPv6双栈，裸金属服务器不支持。
- 同一个网卡上，只能绑定一个IPv6地址。
- 云服务器网络配置需处于正常状态。

检查网络配置是否正常的方法，请参见[检查云服务器网络配置状态](#)。

如果网络配置不正常，请[提交工单](#)获取技术支持。

操作导航

- Windows系统：本文以Windows 2012版本、Windows 2008版本为例，介绍Windows操作系统启用IPv6的方法，如[表5-4](#)所示。
- Linux系统：本文提供了自动配置、手动配置两种方式启用IPv6，推荐您使用自动配置方法，如[表5-4](#)所示。

对于CentOS 6.x和Debian操作系统，云服务器内部配置IPv6自动获取功能之后，将该云服务器制作为私有镜像，使用该镜像在非IPv6网络环境中创建云服务器时，由于等待获取IPv6地址超时，导致云服务器启动较慢，此时您可以参考[设置云服务器获取IPv6地址超时时间](#)操作。


表 5-4 不同操作系统启用 IPv6 操作指导

操作系统	方式	操作指导
Windows 2012	自动配置启用IPv6	Windows 2012操作系统
Windows 2008	自动配置启用IPv6	Windows 2008操作系统
Linux	自动配置启用IPv6（推荐）	Linux操作系统（自动配置启用IPv6）
Linux	手动配置启用IPv6	Linux操作系统（手动配置启用IPv6）

开启云服务器的 IPv6 功能

说明

云服务器所属子网的IPv6功能开启后会自动分配IPv6网段，开启后不允许关闭。

1. 登录管理控制台。
2. 在管理控制台左上角单击 ，选择区域和项目。
3. 单击“≡”，选择“计算 > 弹性云服务器”。
4. 单击待开启IPv6功能的弹性云服务器，进入详情页面。
5. 单击“虚拟私有云”名称，进入弹性云服务器所属虚拟私有云列表。
6. 在虚拟私有云列表中，单击“子网个数”列对应的数字超链接。
进入子网列表页面。
7. 在子网列表中，单击待修改的子网名称超链接。
进入子网详情页面。
8. 在子网详情页，单击“开启IPv6”。
9. 单击“是”，完成子网IPv6网段的开启。

Windows 2012 操作系统

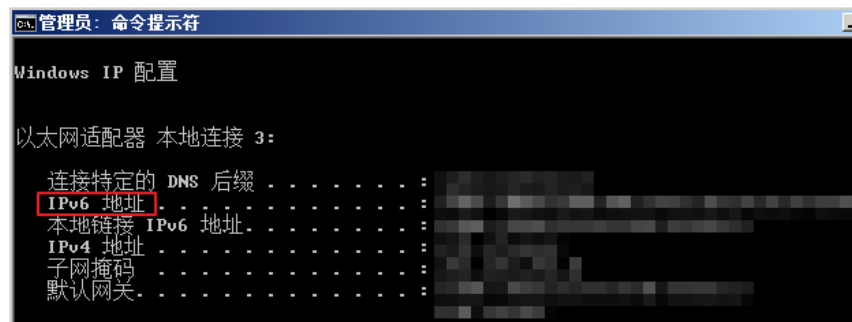
步骤1 检查是否启用IPv6。

打开cmd窗口，执行如下命令，查看当前云服务器是否启用IPv6。

ipconfig

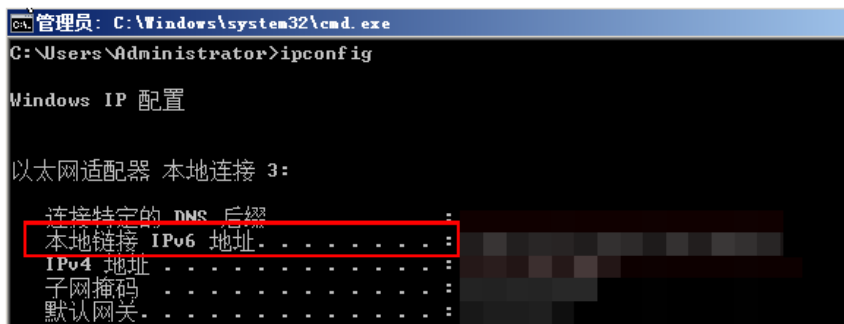
- 如果已启用IPv6，则会显示IPv6的地址。

图 5-11 显示 IPv6 的地址



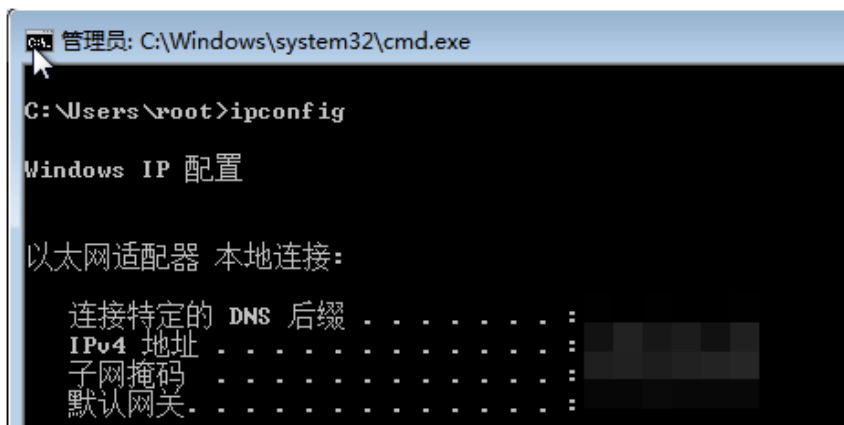
- 如果显示只有本地链接IPv6地址，则表示无法动态获取到IPv6地址。请执行**步骤2**。

图 5-12 本地链接 IPv6 地址



- 如果未启用IPv6，则不会显示IPv6的地址。请执行**步骤3**。

图 5-13 未启用 IPv6



说明

Windows公共镜像默认已经配置了IPv6动态获取功能，即回显如图5-11所示，无需特殊配置。

步骤2 配置动态获取IPv6。

1. 单击“开始 > 控制面板”。
2. 单击“网络和共享中心”。
3. 单击以太网连接。

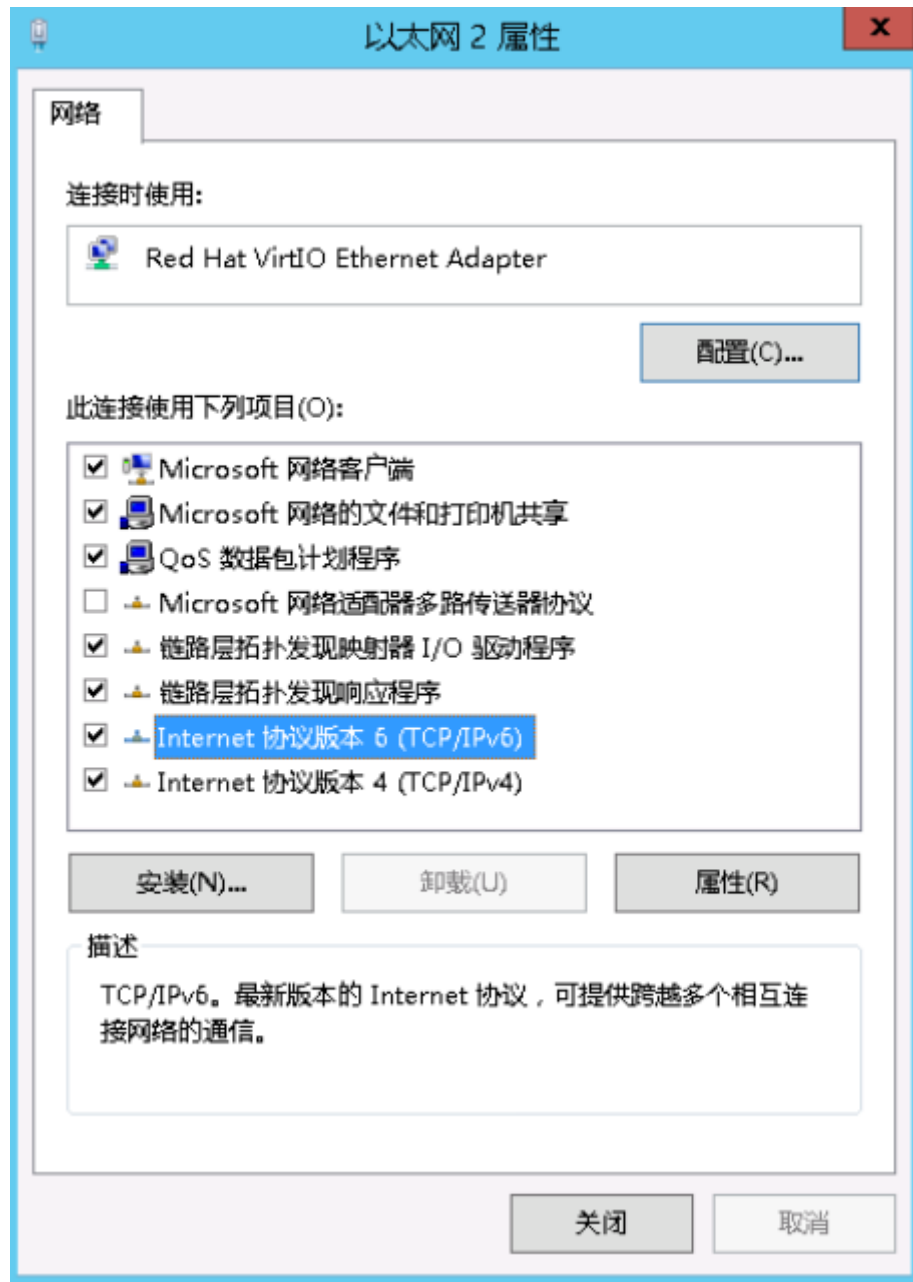
图 5-14 以太网连接



4. 在以太网状态的弹窗中单击左下角的“属性”。

- 勾选“Internet协议版本 6 (TCP/IPv6)”，然后单击“确定”。

图 5-15 配置动态获取 IPv6

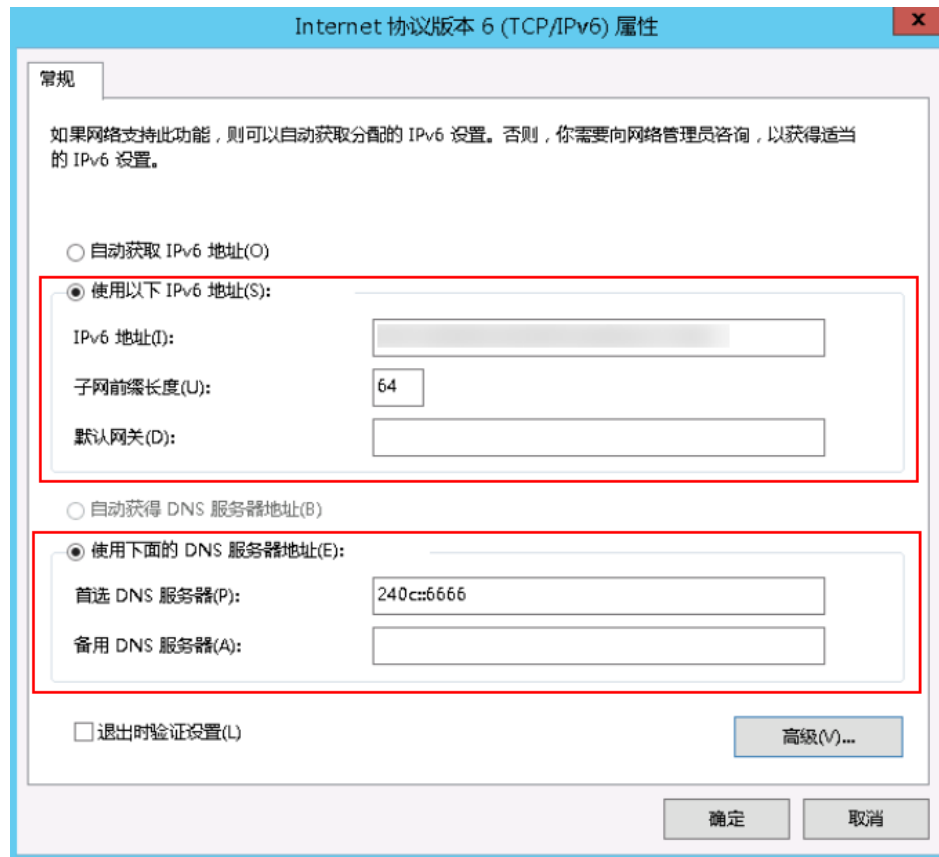


- 执行**步骤1**检查是否已开启动态获取IPv6。

步骤3 启用和配置IPv6。

- 在“Internet 协议版本 6 (TCP/IPv6) 属性”弹窗中，配置云服务器的IPv6地址和DNS服务器地址。
 - IPv6地址：创建云服务器时分配的IPv6地址，请从控制台云服务器的列表页面获取。
 - 子网前缀长度：64
 - DNS服务器：推荐使用240c::6666

图 5-16 在控制台获取 IPv6 地址



2. （可选配置）根据操作系统不同请分别执行以下命令。
Windows Server 2012操作系统云服务器请在PowerShell或者cmd中执行如下命令：

Set-NetIPv6Protocol -RandomizeIdentifiers disabled

3. 执行步骤1检查是否已开启动态获取IPv6。

----结束

Windows 2008 操作系统

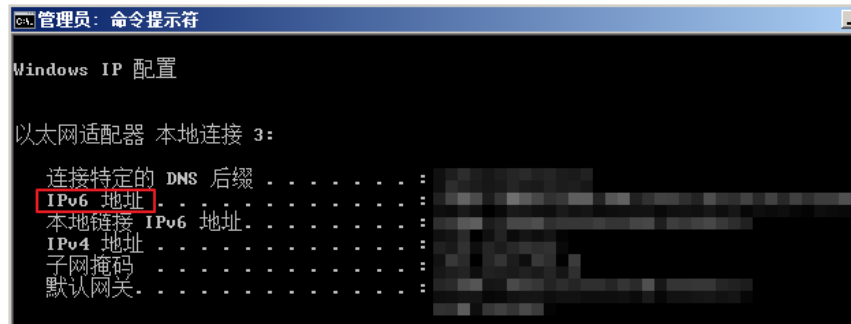
步骤1 检查是否启用IPv6。

打开cmd窗口，执行如下命令，查看当前云服务器是否启用IPv6。

ipconfig

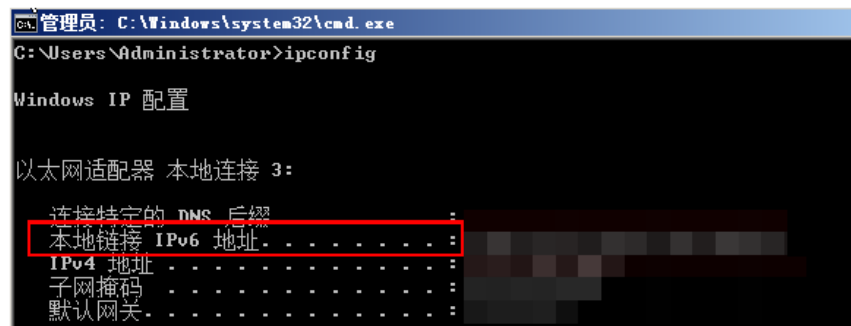
- 如果已启用IPv6，则会显示IPv6的地址。

图 5-17 显示 IPv6 的地址



- 如果显示只有本地链接IPv6地址，则表示无法动态获取到IPv6地址。请执行[步骤2](#)。

图 5-18 本地链接 IPv6 地址



- 如果未启用IPv6，则不会显示IPv6的地址。请执行[步骤3](#)。

图 5-19 未启用 IPv6



说明

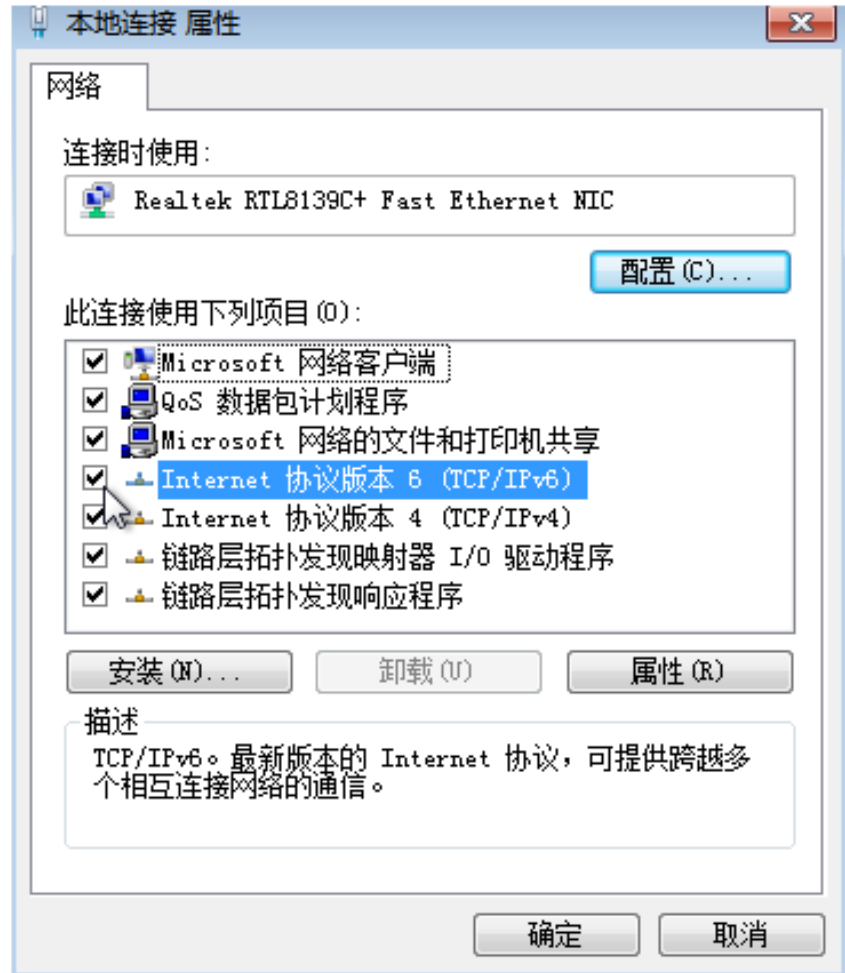
Windows公共镜像默认已经配置了IPv6动态获取功能，即回显如[图5-17](#)所示，无需特殊配置。

步骤2 配置动态获取IPv6。

1. 单击“开始 > 控制面板”。
2. 单击“网络和共享中心”。

3. 左键单击“更改适配器设置”。
4. 右键单击网络连接并选择“属性”。
5. 勾选“Internet协议版本 6 (TCP/IPv6)”，然后单击“确定”。

图 5-20 配置动态获取 IPv6

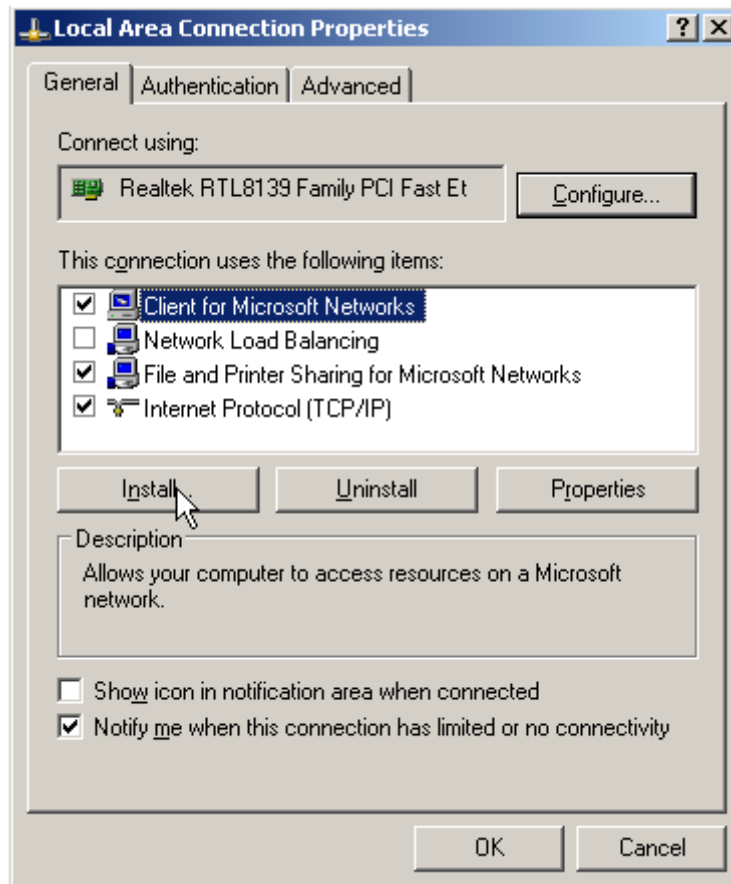


6. 执行**步骤1**检查是否已开启动态获取IPv6。

步骤3 启用和配置IPv6。

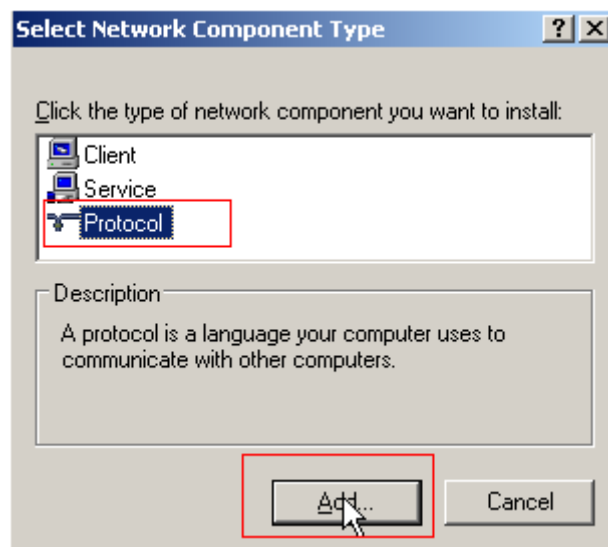
1. 选择“开始 > 控制面板 > 网络连接 > 本地连接”。
2. 选择“属性”，确认勾选以下选项后单击“安装”。

图 5-21 启用和配置 IPv6



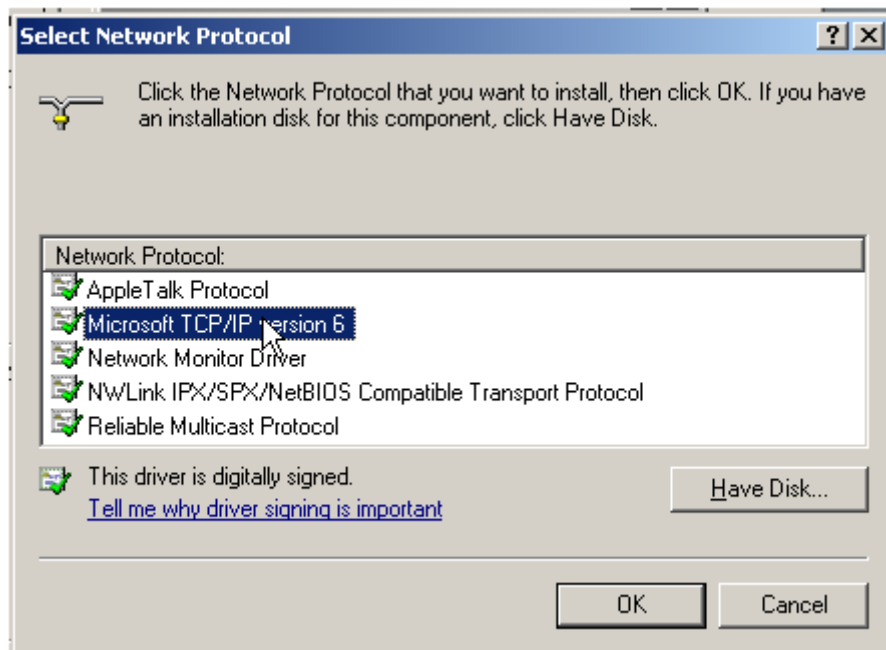
3. 选择“协议”，然后单击“添加”。

图 5-22 添加协议



4. 在网络协议列表中选择“Microsoft TCP/IP版本 6”，然后单击“确定”。

图 5-23 网络协议列表



5. （可选配置）根据操作系统不同请分别执行以下命令。

Windows Server 2008操作系统云服务器请在PowerShell或者cmd中执行如下命令：

```
netsh interface ipv6 set global randomizeidentifiers=disable
```

设置云服务器先禁用本地连接，再重启本地连接。

禁用本地连接：单击“开始 > 控制面板 > 网络和共享中心 > 更改适配器配置”，选择本地连接，单击右键选择“禁用”。

重启本地连接：单击“开始 > 控制面板 > 网络和共享中心 > 更改适配器配置”，选择本地连接，单击右键选择“启用”。

6. 执行[步骤1](#)检查是否已开启动态获取IPv6。

----结束

Linux 操作系统（自动配置启用 IPv6）

ipv6-setup-xxx工具能为开启IPv6协议栈的Linux操作系统自动配置动态获取IPv6地址。其中，xxx表示工具系列：rhel或debian。

您也可以参考[Linux操作系统（手动配置启用IPv6）](#)手动配置启用IPv6。

⚠ 注意

- ipv6-setup-xxx工具运行时会自动重启网络服务，导致网络短暂不可用。
- CentOS 6.x和Debian操作系统的云服务器内部配置IPv6自动获取功能之后，将该云服务器制作作为私有镜像，使用该镜像在非IPv6网络环境中创建云服务器时，由于等待获取IPv6地址超时，导致云服务器启动较慢，您可以参考[设置云服务器获取IPv6地址超时时间](#)设置获取IPv6地址超时时间为30s，然后再重新制作私有镜像。

步骤1 执行如下命令，查看当前云服务器是否启用IPv6。

ip addr

- 如果没有开启IPv6协议栈，则只能看到IPv4地址，如下图所示，请参考[设置云服务器获取IPv6地址超时时间](#)先开启IPv6协议栈。

图 5-24 云服务器未开启 IPv6 协议栈

```
eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP qlen 1000
link/ether fa:16:3e: b rd ff:ff:ff:ff:ff:ff
inet b rd scope global noprefixroute dynamic eth0
valid_lft 1193sec preferred_lft 1193sec
```

- 如果已开启IPv6协议栈，则可以看到LLA地址（fe80开头）。

图 5-25 云服务器已开启 IPv6 协议栈

```
eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP group default qlen 1000
link/ether fa:16:3e: b rd ff:ff:ff:ff:ff:ff
inet b rd scope global noprefixroute dynamic eth0
valid_lft 76391sec preferred_lft 76391sec
inet6 fe80::f816: /64 scope link
valid_lft forever preferred_lft forever
```

- 如果已开启IPv6协议栈并且已获取到IPv6地址，则会看到如下地址：

图 5-26 云服务器已开启 IPv6 协议栈并且已获取到 IPv6 地址

```
eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP group default qlen 1000
link/ether fa:16:3e:75:af:4c b rd ff:ff:ff:ff:ff:ff
inet b rd scope global noprefixroute dynamic eth0
valid_lft 86395sec preferred_lft 86395sec
inet6 2407:c080:802: /128 scope global dynamic
valid_lft 7496sec preferred_lft 7196sec
inet6 fe80::f816:3eff: /64 scope link noprefixroute
valid_lft forever preferred_lft forever
```

说明

Linux公共镜像均已开启IPv6协议栈，如[图5-25](#)所示；

步骤2 开启Linux云服务器IPv6协议栈。

1. 执行如下命令，确认内核是否支持IPv6协议栈。

sysctl -a | grep ipv6

- 如果有输出信息，表示内核支持IPv6协议栈。
- 如果没有任何输出，说明内核不支持IPv6协议栈，需要执行**2.b**加载IPv6模块。

2. 执行以下命令，加载IPv6模块。

modprobe ipv6

3. 修改“/etc/sysctl.conf”配置文件，增加如下配置：

net.ipv6.conf.all.disable_ipv6=0

4. 保存配置并退出，然后执行如下命令，加载配置。

sysctl -p

步骤3 自动配置启用IPv6。

1. 下载对应系统版本的工具ipv6-setup-rhel或ipv6-setup-debian，并上传至待操作的云服务器。

ipv6-setup-xxx工具会添加或者修改网卡设备的配置文件，添加IPv6动态获取的配置信息，然后重启网卡或者网络服务。ipv6-setup-rhel和ipv6-setup-debian的工具下载地址如表5-5所示。

表 5-5 工具下载地址

系列	发行版	下载地址
RHEL	<ul style="list-style-type: none">- CentOS 6/7- EulerOS 2.2/2.3- Fedora 25	https://ecs-instance-driver.obs.cn-north-1.myhuaweicloud.com/ipv6/ipv6-setup-rhel
Debian	<ul style="list-style-type: none">- Ubuntu 16/18- Debian 8/9/10	https://ecs-instance-driver.obs.cn-north-1.myhuaweicloud.com/ipv6/ipv6-setup-debian

2. 执行以下命令，添加执行权限。
chmod +x ipv6-setup-xxx
3. 执行以下命令，指定一个网卡设备，配置动态获取IPv6地址。
./ipv6-setup-xxx --dev [dev]
示例：
./ipv6-setup-xxx --dev eth0

📖 说明

- 如需对所有网卡配置动态获取IPv6地址，命令为**./ipv6-setup-xxx**，即不带参数。
- 如需查询工具的用法，请执行命令**./ipv6-setup-xxx --help**。

----结束

Linux 操作系统（手动配置启用 IPv6）

⚠ 注意

CentOS 6.x和Debian操作系统的云服务器内部配置IPv6自动获取功能之后，将该云服务器制作作为私有镜像，使用该镜像在非IPv6网络环境中创建云服务器时，由于等待获取IPv6地址超时，导致云服务器启动较慢，您可以参考[设置云服务器获取IPv6地址超时时间](#)设置获取IPv6地址超时时间为30s，然后再重新制作私有镜像。

步骤1 执行如下命令，查看当前云服务器是否启用IPv6。

ip addr

- 如果没有开启IPv6协议栈，则只能看到IPv4地址，如下图所示，请参考[步骤2](#)先开启IPv6协议栈。

图 5-27 未开启 IPv6 协议栈

```
eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP qlen 1000
link/ether fa:16:3e:..... brd ff:ff:ff:ff:ff:ff
inet ..... brd ..... scope global noprefixroute dynamic eth0
    valid_lft 1193sec preferred_lft 1193sec
```

- 如果已开启IPv6协议栈，则可以看到LLA地址（fe80开头）。

图 5-28 已开启 IPv6 协议栈

```
eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP group default qlen 1000
link/ether fa:16:3e: b8: ff:ff:ff:ff:ff:ff
inet b8:ff:ff:ff:ff:ff:ff scope global noprefixroute dynamic eth0
    valid_lft 76391sec preferred_lft 76391sec
inet6 fe80::f816: b8: /64 scope link
    valid_lft forever preferred_lft forever
```

- 如果已开启IPv6协议栈并且已获取到IPv6地址，则会看到如下地址：

图 5-29 已开启 IPv6 协议栈并且已获取到 IPv6 地址

```
eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP group default qlen 1000
link/ether fa:16:3e:75:af:4c b8: ff:ff:ff:ff:ff:ff
inet b8:ff:ff:ff:ff:ff:ff scope global noprefixroute dynamic eth0
    valid_lft 86395sec preferred_lft 86395sec
inet6 2407:c080:802: b8: /128 scope global dynamic
    valid_lft 7496sec preferred_lft 7196sec
inet6 fe80::f816:3eff: b8: /64 scope link noprefixroute
    valid_lft forever preferred_lft forever
```

说明

Linux公共镜像均已开启IPv6协议栈，如图5-28所示；

步骤2 开启Linux云服务器IPv6协议栈。

1. 执行如下命令，确认内核是否支持IPv6协议栈。

```
sysctl -a | grep ipv6
```

- 如果有输出信息，表示内核支持IPv6协议栈。
- 如果没有任何输出，说明内核不支持IPv6协议栈，需要执行步骤2.2加载IPv6模块。

2. 执行以下命令，加载IPv6模块。

```
modprobe ipv6
```

3. 修改“/etc/sysctl.conf”配置文件，增加如下配置：

```
net.ipv6.conf.all.disable_ipv6=0
```

4. 保存配置并退出，然后执行如下命令，加载配置。

```
sysctl -p
```

步骤3 手动配置启用IPv6。操作系统不同，步骤有所差别。

- Ubuntu 18.04/20.04操作系统云服务器配置动态获取IPv6。

- a. 执行以下命令，进入“/etc/netplan/”。

```
cd /etc/netplan
```

- b. 执行以下命令，查询配置文件名。

```
ls
```

图 5-30 查询配置文件名

```
root@ecs-: /etc/netplan# ls
01-netcfg.yaml  01-network-manager-all.yaml
```

- c. 执行以下命令，编辑“01-network-manager-all.yaml”配置文件。

```
vi 01-network-manager-all.yaml
```

- d. 在“01-network-manager-all.yaml”下增加如下内容，注意yaml文件格式及缩进：

```
ethernets:  
  eth0:  
    dhcp6: true
```

图 5-31 修改结果

```
# Let NetworkManager manage all devices on this system  
network:  
  version: 2  
  renderer: NetworkManager  
  ethernets:  
    eth0:  
      dhcp6: true
```

修改完成后保存退出。

- e. 执行以下命令，使更改生效。

sudo netplan apply

- Ubuntu 22.04操作系统云服务器配置动态获取IPv6。

- a. 执行以下命令，进入“/etc/netplan/”。

cd /etc/netplan

- b. 执行以下命令，查询配置文件名。

ls

图 5-32 查询配置文件名

```
root@ecs-485b:/etc/netplan# ls  
01-netcfg.yaml
```

- c. 执行以下命令，编辑“01-netcfg.yaml”配置文件。

vi 01-netcfg.yaml

- d. 在“01-netcfg.yaml”中增加如下内容，注意yaml文件格式及缩进：

```
ethernets:  
  eth0:  
    dhcp6: true
```

图 5-33 修改结果

```
network:  
  version: 2  
  renderer: NetworkManager  
  ethernets:  
    eth0:  
      dhcp4: true  
      dhcp6: true  
    eth1:  
      dhcp4: true  
    eth2:  
      dhcp4: true  
    eth3:  
      dhcp4: true  
    eth4:  
      dhcp4: true
```

修改完成后保存退出。

- e. 执行以下命令，使更改生效。

```
sudo netplan apply
```

- f. 执行以下命令，编辑“/etc/NetworkManager/NetworkManager.conf”文件。

```
vi /etc/NetworkManager/NetworkManager.conf
```

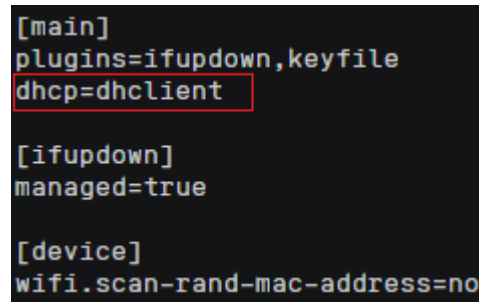
- g. 在“NetworkManager.conf”中增加如下内容，注意文件格式及缩进：

```
[main]
plugins=ifupdown,keyfile
dhcp=dhclient

[ifupdown]
managed=true

[device]
wifi.scan-rand-mac-address=no
```

图 5-34 修改结果



```
[main]
plugins=ifupdown,keyfile
dhcp=dhclient

[ifupdown]
managed=true

[device]
wifi.scan-rand-mac-address=no
```

- h. 执行以下命令，使配置生效。

```
systemctl restart NetworkManager
```

- Debian操作系统云服务器配置动态获取IPv6。

- a. 编辑“/etc/network/interfaces”文件，使之包含以下内容：

```
auto lo
iface lo inet loopback
auto eth0
iface eth0 inet dhcp
iface eth0 inet6 dhcp
pre-up sleep 3
```

- b. 如果有多个网卡，则在“/etc/network/interfaces”文件中，增加对应网卡的配置，以eth1为例，需要增加：

```
auto eth1
iface eth1 inet dhcp
iface eth1 inet6 dhcp
pre-up sleep 3
```

- c. 执行如下命令重启网络服务。

```
service networking restart
```

📖 说明

如果将网卡进行down/up操作之后无法获取IPv6地址，也可以通过此命令重启网络服务。

- d. 执行步骤[步骤1](#)检查是否已开启动态IPv6。

- CentOS/EulerOS/Fedora操作系统云服务器配置动态获取IPv6。

- a. 编辑主网卡配置文件“/etc/sysconfig/network-scripts/ifcfg-eth0”。
补充如下配置项：
IPV6INIT=yes
DHCPV6C=yes
- b. 编辑“/etc/sysconfig/network”，按如下所示添加或修改以下行。
NETWORKING_IPV6=yes
- c. CentOS 6系列从网卡需要编辑对应的配置文件，以eth1为例，编辑“/etc/sysconfig/network-scripts/ifcfg-eth1”。
补充如下配置项：
IPV6INIT=yes
DHCPV6C=yes

CentOS 6.3系统中默认ip6tables会过滤dhcpv6-client请求，所以CentOS 6.3除了需要编辑“ifcfg-eth*”文件外，还需要额外添加一条允许dhcpv6-client请求的ip6tables规则。操作如下：

- i. 执行以下命令，添加ip6tables规则。
ip6tables -A INPUT -m state --state NEW -m udp -p udp --dport 546 -d fe80::/64 -j ACCEPT
- ii. 执行以下命令，保存ip6tables规则。
service ip6tables save

图 5-35 命令示例

```
root@ecs-cd02 log]# ip6tables -A INPUT -m state --state NEW -m udp -p udp --dport 546 -d fe80::/64 -j ACCEPT
nf_conntrack version 0.5.0 (7964 buckets, 31856 max)
root@ecs-cd02 log]# service ip6tables save
ip6tables: Saving firewall rules to /etc/sysconfig/ip6tablef OK ]
```

- d. （可选配置）CentOS 7/CentOS 8系列需要将扩展网卡的IPv6 LLA地址模式修改为EUI64。
 - i. 执行如下命令查看网卡信息。
nmcli con

图 5-36 查看网卡信息

```
[root@ecs-166b ~]# nmcli con
NAME                UUID                                  TYPE      DEVICE
System eth0         5fb06bd0-0bb0-7ffb-45f1-d6edd65f3e03  ethernet  eth0
Wired connection 1  9c92fad9-6ecb-3e6c-eb4d-8a47c6f50c04  ethernet  eth1
Wired connection 1  3a73717e-65ab-93e8-b518-24f5af32dc0d  ethernet  eth2
```

- ii. 将eth1的IPv6 LLA地址模式按以下命令修改为EUI64：
nmcli con modify "Wired connection 1" ipv6.addr-gen-mode eui64

📖 说明

CentOS不同系列，网卡信息存在差异，命令中的“Wired connection 1”需要根据实际查询的网卡信息的“NAME”列进行替换。

- iii. 通过ifconfig命令将eth1进行down/up操作。
ifdown eth1
ifup eth1
- e. 重启网络服务。
 - i. CentOS 6系列执行以下命令，重启网络服务。
service network restart

- ii. CentOS 7/EulerOS/Fedora系列执行以下命令，重启网络服务。
systemctl restart NetworkManager
- f. 执行步骤**步骤1**检查是否已开启动态IPv6。
- SUSE/openSUSE/CoreOS操作系统云服务器配置动态获取IPv6。
SUSE 11 SP4不支持IPv6自动获取。
SUSE 12 SP1、SUSE 12 SP2无需特殊配置。
openSUSE 13.2、openSUSE 42.2无需特殊配置。
CoreOS 10.10.5无需特殊配置。

---结束

检查云服务器网络配置状态

1. 执行以下命令，检查云服务器网络服务状态是否正常。

systemctl status NetworkManager

如果网络服务状态正常，则执行结果中服务处于active (running)状态且是enable的。

图 5-37 网络服务器状态

```
Welcome to Huawei Cloud Service
[root@ecs-~]# systemctl status NetworkManager
* NetworkManager.service - Network Manager
   Loaded: loaded (/usr/lib/systemd/system/NetworkManager.service; enabled; vendor preset: enabled)
   Active: active (running) since Tue 2024-05-21 18:32:01 CST; 24h ago
     Docs: man:NetworkManager(8)
   Main PID: 581 (NetworkManager)
  CGroup: /system.slice/NetworkManager.service
          └─581 /usr/sbin/NetworkManager --no-daemon
              └─555 /sbin/dhclient -d -q -sf /usr/libexec/nm-dhcp-helper -pf /var/run/dhclient-eth0.pid -lf /var/lib/NetworkManager/dhclient-5fb82b08-8bb8-7ffb-45f1-d6dd512e83-eth0...
```

2. 执行以下命令，检查云服务器网卡获取IP地址的方式。

cat /etc/sysconfig/network-scripts/ifcfg-ethx

📖 说明

- “ethx”需替换为具体的网卡，例如eth0。
- 该命令以CentOS 7为例进行介绍。

图 5-38 网卡获取 IP 地址的方式

```
[root@ecs-~]# cat /etc/sysconfig/network-scripts/ifcfg-eth0
DEVICE="eth0"
BOOTPROTO="dhcp"
ONBOOT="yes"
TYPE="Ethernet"
PERSISTENT_DHCLIENT="yes"
```

- 如果BOOTPROTO字段的值为dhcp：表示云服务器网卡通过动态DHCP获取IP地址，继续执行**3**。
 - 如果BOOTPROTO字段的值为static：表示云服务器网卡通过静态配置的方式设置网络，继续执行**4**。
3. 如果云服务器网卡通过动态DHCP获取IP地址，则需要执行以下命令，确认DHCP相关进程是否正常。

systemctl status NetworkManager

回显中包含相关网卡的dhclient进程则代表正常。

图 5-39 检查 DHCP 进程

```
[root@ecs-~]# systemctl status NetworkManager
NetworkManager.service - Network Manager
Loaded: loaded (/usr/lib/systemd/system/NetworkManager.service; enabled; vendor preset: enabled)
Active: active (running) since Tue 2024-05-21 18:32:01 CST; 1 day 4h ago
Docs: man:NetworkManager(8)
Main PID: 581 (NetworkManager)
CGroup: /system.slice/NetworkManager.service
└─581 /usr/sbin/NetworkManager --no-daemon
    └─585 /sbin/dhclient -d -q -sf /usr/libexec/nm-dhcp-helper -pf /var/run/dhclient-eth0.pid -lf /var/lib/NetworkManager/dhclient-57b6bd9-8bb8-7ffb-45f1-d6dd65f3e83-eth0...
```

4. 如果云服务器网卡通过静态配置的方式设置网络，则需要执行以下命令，确认IP地址配置是否生效。

```
ip a
```

回显中IP配置字段是forever则代表配置已经生效。

图 5-40 检查静态配置

```
[root@ecs-~]# ip a
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN group default qlen 1000
    link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
    inet 127.0.0.1/8 scope host lo
        valid_lft forever preferred_lft forever
    inet6 ::1/128 scope host
        valid_lft forever preferred_lft forever
2: eth0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc mq state UP group default qlen 1000
    link/ether fa:16:3e:fc:fc:14 brd ff:ff:ff:ff:ff:ff
    inet 10.0.0.24/24 brd 10.0.0.255 scope global noprefixroute eth0
        valid_lft forever preferred_lft forever
    inet6 fe80::f816:3eff:fcfc:14/64 scope link
        valid_lft forever preferred_lft forever
```

设置云服务器获取 IPv6 地址超时时间

CentOS 6.x和Debian操作系统的云服务器内部配置IPv6自动获取功能之后，将该云服务器制作作为私有镜像，使用该镜像在非IPv6网络环境中创建云服务器时，由于等待获取IPv6地址超时，导致云服务器启动较慢，您可以参考本节操作设置获取IPv6地址超时时间为30s，然后再重新制作私有镜像。

- CentOS 6.x:
 - a. 执行以下命令编辑“dhclient.conf”文件。

```
vi /etc/dhcp/dhclient.conf
```
 - b. 按“i”进入编辑模式，在文件中增加timeout属性。

```
timeout 30;
```
 - c. 输入:wq保存后退出。
- Debian 7.5:
 - a. 执行以下命令编辑“networking”文件。

```
vi /etc/init.d/networking
```
 - b. 按“i”进入编辑模式，增加延迟命令timeout，修改点如下图所示。

图 5-41 修改点 1

```
115 case "$1" in
116 start)
117     if init_is_upstart; then
118         exit 1
119     fi
120     process_options
121     check_ifstate
122
123     if [ "$CONFIGURE_INTERFACES" = no ]
124     then
125         log_action_msg "Not configuring network interfaces, see /etc/default/networking"
126         exit 0
127     fi
128     set -f
129     exclusions=$(process_exclusions)
130     log_action_begin_msg "Configuring network interfaces"
131     if /usr/bin/timeout 30 ifup -a $exclusions $verbose && ifup_hotplug $exclusions $verbose
132     then
133         log_action_end_msg $?
134     else
135         log_action_end_msg $?
136     fi
137     ;;
138
139 stop)
140     if init_is_upstart; then
141         exit 0
142     fi
143     check_network_file_systems
144     check_network_swap
145
146     log_action_begin_msg "Deconfiguring network interfaces"
147     if /usr/bin/timeout 30 ifdown -a --exclude=lo $verbose; then
148         log_action_end_msg $?
```

图 5-42 修改点 2

```
154 reload)
155     process_options
156     log_action_begin_msg "Reloading network interfaces configuration"
157     state=$(cat /run/network/ifstate)
158     if /usr/bin/timeout 30 ifdown -a --exclude=lo $verbose || true
159     then
160         if /usr/bin/timeout 30 ifup --exclude=lo $state $verbose ; then
161             log_action_end_msg $?
162         else
163             log_action_end_msg $?
164         fi
165     ;;
166
167 force-reload|restart)
168     if init_is_upstart; then
169         exit 1
170     fi
171     process_options
172
173     log_warning_msg "Running $0 $1 is deprecated because it may not re-enable some interfaces"
174     log_action_begin_msg "Reconfiguring network interfaces"
175     if /usr/bin/timeout 30 ifdown -a --exclude=lo $verbose || true
176     then
177         set -f
178         exclusions=$(process_exclusions)
179         if /usr/bin/timeout 30 ifup -a --exclude=lo $exclusions $verbose && ifup_hotplug $exclusions $verbose
180     then
181         log_action_end_msg $?
182     else
183         log_action_end_msg $?
```

- Debian 8.2.0/8.8.0
 - a. 执行以下命令编辑“network-pre.conf”文件。
vi /lib/systemd/system/networking.service.d/network-pre.conf
 - b. 按“i”进入编辑模式，在文件中增加timeout属性。
[Service]
TimeoutStartSec=30
- Debian 9.0
 - a. 执行以下命令编辑“networking.service”文件。
vi /etc/systemd/system/network-online.target.wants/networking.service
 - b. 按“i”进入编辑模式，将TimeoutStartSec=5min改为
TimeoutStartSec=30。

6 弹性公网 IP 管理

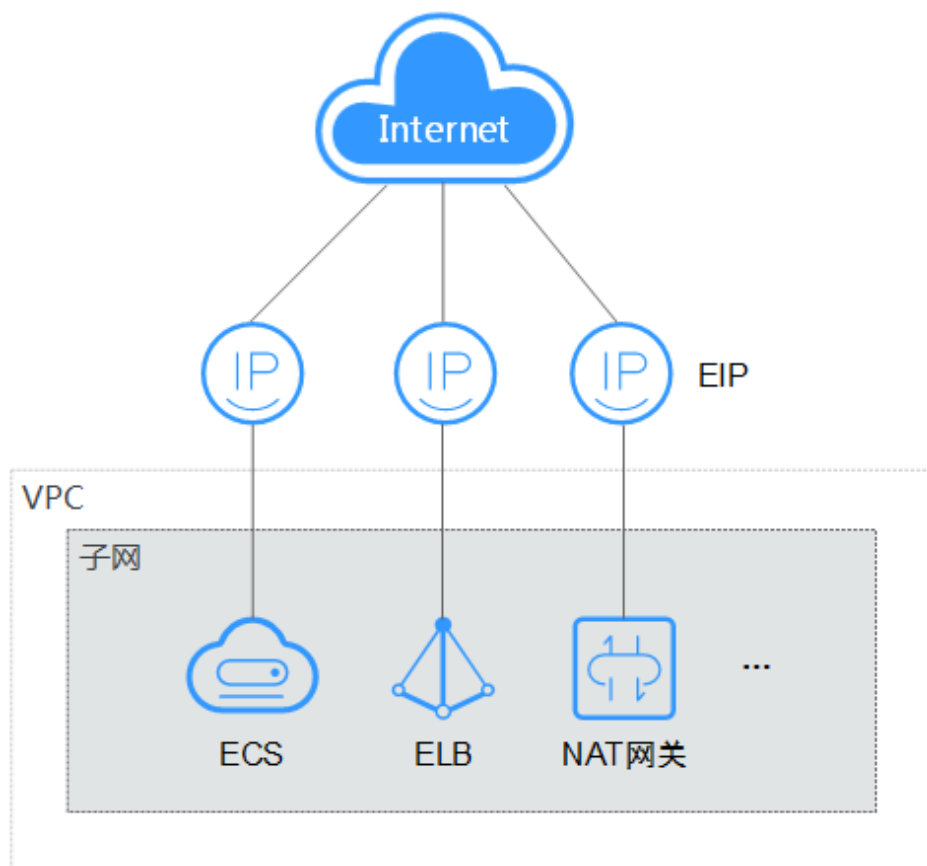
6.1 弹性公网 IP 概述

弹性公网 IP

弹性公网IP（Elastic IP，简称EIP）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。拥有多种灵活的计费方式，可以满足各种业务场景的需要。

一个弹性公网IP只能绑定一个云资源使用。

图 6-1 通过 EIP 访问公网



相关链接



- [绑定弹性公网IP](#)
- [解绑弹性公网IP](#)
- [修改弹性公网IP的带宽](#)
- [弹性公网IP类常见问题](#)
- [访问网站或应用故障类常见问题](#)

6.2 绑定弹性公网 IP

操作场景

可以通过申请弹性公网IP并将弹性公网IP绑定到弹性云服务器上，实现弹性云服务器访问公网的目的。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。

- 在弹性云服务器列表中，找到待绑定弹性公网IP的弹性云服务器，单击“操作”列下的“更多 > 网络/安全组 > 绑定弹性公网IP”。
- 选择弹性公网IP。

说明

如果当前区域没有可用的弹性公网IP，则弹性公网IP列表为空，请购买弹性公网IP后重新执行绑定操作。

图 6-2 绑定弹性公网 IP



- 单击“确定”。
- 完成绑定后，可以在云服务器列表页查看已绑定的弹性公网IP。

6.3 解绑弹性公网 IP



操作场景

本节操作介绍从云服务器解绑弹性公网IP的操作步骤。

注意

解绑弹性公网IP会导致弹性云服务器实例无法访问公网，在解绑前请确保不再需要访问公网或者有替代的网络连接方式。

操作步骤

- 登录管理控制台。
- 单击管理控制台左上角的 ，选择区域和项目。
- 单击“”，选择“计算 > 弹性云服务器”。
- 在弹性云服务器列表中，在待调整带宽的弹性云服务器操作列下，单击“操作”列下的“网络/安全组 > 解绑弹性公网IP”。

5. 确认弹性公网IP信息，单击“确定”。

说明

未绑定弹性云服务器的弹性公网IP会继续计费，如果不再使用请根据控制台的引导释放弹性公网IP。

详细操作，请参见[释放/退订弹性公网IP](#)。

6.4 更换弹性公网 IP

操作场景

如果您的弹性云服务器绑定了弹性公网IP，您可以参考本节操作更换弹性公网IP。

暂不支持直接更换弹性云服务器的弹性公网IP。您需要先解绑当前的弹性公网IP，再重新为弹性云服务器绑定新的弹性公网IP。

如果无可更换的弹性公网IP，您需要先申请新的弹性公网IP，再进行更换操作。

注意

更换弹性公网IP前，请谨慎评估该操作可能带来的业务影响，避免由于误操作造成网络中断。

约束限制

为防止误删除操作，弹性公网IP存在24小时缓存机制，对于已释放过弹性公网IP的用户，24小时内会优先分配之前使用过的弹性公网IP。

如需申请新的弹性公网IP地址，建议您先购买新的弹性公网IP后再释放旧的弹性公网IP。


详细内容请参考[弹性公网IP的分配策略](#)。

前提条件

已购买新的弹性公网IP。

详细内容，请参见[申请弹性公网IP](#)。


解绑弹性公网 IP

1. 登录管理控制台。
2. 单击管理控制台左上角的，选择区域和项目。
3. 选择指定云服务器，单击“操作”列下的“更多 > 网络/安全组 > 解绑弹性公网IP”。
4. 确认弹性公网IP信息，解绑EIP。

说明

未绑定弹性云服务器的弹性公网IP会继续计费，如果不再使用请根据控制台的引导释放弹性公网IP。

绑定新的弹性公网 IP

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择指定云服务器，单击“操作”列下的“更多 > 网络/安全组 > 绑定弹性公网 IP”。
4. 选择弹性公网IP，单击“确定”完成新的弹性公网IP的绑定。

说明

如果当前区域没有可用的弹性公网IP，则弹性公网IP列表为空，请购买弹性公网IP后重新执行绑定操作。

图 6-3 绑定新的弹性公网 IP



6.5 修改弹性公网 IP 的带宽

操作场景

弹性公网IP的带宽可以为弹性云服务器提供公网和弹性云服务器间的访问流量。如果当前弹性公网IP的带宽不满足您的业务需要，可以参考本节操作调整弹性云服务器带宽。


约束与限制

降低带宽大小，可能会影响业务流量造成丢包，请确认对业务产生的影响，谨慎操作。

前提条件

弹性云服务器已绑定弹性公网IP，详细操作，请参见[绑定弹性公网IP](#)。

操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中，在待调整带宽的弹性云服务器操作列下，单击“操作”列下的“更多 > 网络/安全组 > 修改带宽”。
5. 根据页面提示，修改“带宽名称”、“计费方式”和“带宽大小”。

6.6 无公网 IP 的弹性云服务器访问 Internet

操作场景

为保证安全和节省公网IP资源，通常只为特定的弹性云服务器配置公网IP，可直接访问Internet，其他弹性云服务器只配置私网IP，无法直接访问Internet。

因此，当只配置了私网IP的弹性云服务器需要访问Internet，执行软件升级、给系统打补丁或者其他需求时，可选择以下方案：

- **（推荐）通过NAT（NAT Gateway）网关服务**
NAT网关能够为VPC内的弹性云服务器提供SNAT和DNAT功能，通过灵活简易的配置，即可轻松构建VPC的公网出入口，实现多个弹性云服务器共享弹性公网IP访问Internet或使弹性云服务器对互联网提供服务。
- **通过Linux操作系统的代理主机**
如果您已拥有一台绑定了EIP的弹性云服务器，可将该弹性云服务器作为代理，为同一网段且同一安全组内的其他无公网IP的云服务器提供访问通道，正常访问Internet。

（推荐）通过 NAT（NAT Gateway）网关服务

📖 说明

本方案需要：

- 已购买弹性公网IP和公网NAT网关。
- 公网NAT网关和其他需要访问Internet的弹性云服务器处于同一VPC内。

表 6-1 NAT 网关操作指引

操作场景	网关规则	相关链接
无公网IP的弹性云服务器访问Internet	SNAT	通过公网NAT网关的SNAT规则访问公网
无公网IP的弹性云服务器对互联网提供服务	DNAT	通过公网NAT网关的DNAT规则面向公网提供服务

通过 Linux 操作系统的代理主机

说明

本方案需要：

- 已拥有一台绑定了公网IP的弹性云服务器作为代理弹性云服务器。
- 代理弹性云服务器和其他需要访问Internet的弹性云服务器均处于同一网段，并且在同一安全组内。

本节操作中，以代理弹性云服务器的操作系统是CentOS 7.9为例进行介绍，适用于CentOS 7.9及以下版本、Huawei Cloud EulerOS 2.0版本。

其他操作系统及版本，请您参考对应官网帮助文档进行配置，或采用（推荐）通过 [NAT（NAT Gateway）网关服务](#)。



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中的右上角，输入代理云服务器名称进行搜索。
5. 单击代理弹性云服务器的名称，查看详情。
6. 在代理弹性云服务器详情页面，选择“弹性网卡”页签，并展开 ，将“源/目的检查”选项设置为“OFF”。

图 6-4 关闭“源/目的检查”



默认情况下，“源/目的检查”状态为“启用”，系统会检查弹性云服务器发送的报文中源IP地址是否正确，否则不允许弹性云服务器发送该报文。这有助于防止伪装报文攻击，提升安全性。但在该场景中，这种保护机制会导致报文的发送者无法接收到返回的报文。因此，需设置“源/目的检查”状态为禁用。

7. 登录代理弹性云服务器。
详细操作方法请参见[Linux ECS登录方式概述](#)。

8. 执行以下命令，检测代理弹性云服务器是否可以正常连接Internet。
ping www.huaweicloud.com
回显包含类似如下信息时，表示代理弹性云服务器可正常连接Internet。

图 6-5 检测是否可以正常连接 Internet

```
[root@ecs-f4f0 ~]# ping www.baidu.com
PING www.a.shifen.com (61.135.169.121) 56(84) bytes of data:
64 bytes from 61.135.169.121 (61.135.169.121): icmp_seq=1 ttl=47 time=2.77 ms
64 bytes from 61.135.169.121 (61.135.169.121): icmp_seq=2 ttl=47 time=2.65 ms
64 bytes from 61.135.169.121 (61.135.169.121): icmp_seq=3 ttl=47 time=2.61 ms
64 bytes from 61.135.169.121 (61.135.169.121): icmp_seq=4 ttl=47 time=2.83 ms
64 bytes from 61.135.169.121 (61.135.169.121): icmp_seq=5 ttl=47 time=2.69 ms
64 bytes from 61.135.169.121 (61.135.169.121): icmp_seq=6 ttl=47 time=2.63 ms
```

9. （可选）执行以下命令，安装iptables服务并配置iptables服务开机自启动。
仅CentOS 7.x版本的弹性云服务器，需要执行此步骤。

```
yum install iptables-services -y  
systemctl start iptables  
systemctl enable iptables
```

10. 执行以下命令，查看代理弹性云服务器的IP转发功能是否开启。

```
cat /proc/sys/net/ipv4/ip_forward
```

- 回显为“0”表示关闭，请执行11。
- 回显为“1”表示开启，请执行16。

11. 执行以下命令，打开IP转发功能配置文件。

```
vi /etc/sysctl.conf
```

12. 按“i”，进入编辑模式。

13. 修改如下参数的值。

将参数“net.ipv4.ip_forward”的值修改为“1”。

📖 说明

如果“sysctl.conf”文件中不存在参数“net.ipv4.ip_forward”，执行以下命令进行添加：

```
echo net.ipv4.ip_forward=1 >> /etc/sysctl.conf
```

14. 按“Esc”，输入:wq，按“Enter”。

保存设置并退出vi编辑器。

15. 执行以下命令，使配置文件修改生效。

```
sysctl -p /etc/sysctl.conf
```

16. 执行以下命令，清除原有iptables规则。

```
iptables -F
```

17. 执行以下命令，配置SNAT，使代理弹性云服务器所在的网段内其他弹性云服务器可通过代理弹性云服务器访问Internet。

```
iptables -t nat -A POSTROUTING -o eth0 -s subnet/netmask-bits -j SNAT --to nat-instance-ip
```

假设代理弹性云服务器所在的网段为192.168.125.0，子网掩码为24位，私网IP地址为192.168.125.4，则执行如下命令。

```
iptables -t nat -A POSTROUTING -o eth0 -s 192.168.125.0/24 -j SNAT --to 192.168.125.4
```

📖 说明

为了确保重启后上述规则不丢失，可以执行 `vi /etc/rc.local` 编辑 `rc.local` 文件，将 17 中的规则复制到 `rc.local` 文件，按“ESC”退出编辑模式，输入“:wq”保存并退出。

18. 执行以下命令，保存 iptables 的配置并设置开机自启动。

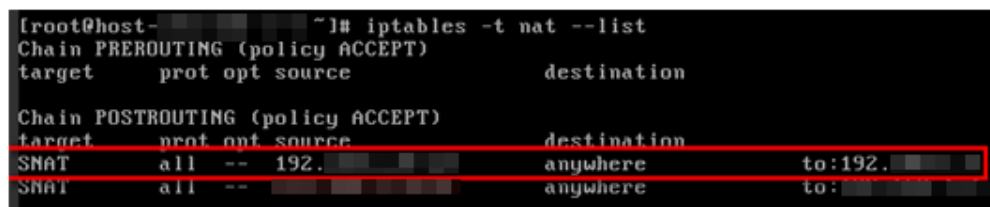
```
service iptables save  
chkconfig iptables on
```

19. 执行以下命令，查看 SNAT 配置是否成功。


```
iptables -t nat --list
```

回显类似如图 6-6 所示时，表示 SNAT 配置成功。

图 6-6 SNAT 配置成功



```
[root@host- ~]# iptables -t nat --list  
Chain PREROUTING (policy ACCEPT)  
target prot opt source destination  
  
Chain POSTROUTING (policy ACCEPT)  
target prot opt source destination  
SNAT all -- 192.168.125.0/24 anywhere to:192.168.125.0/24  
SNAT all -- anywhere anywhere to:
```

20. 添加自定义路由。
 - a. 登录管理控制台。
 - b. 单击管理控制台左上角的 ，选择区域和项目。
 - c. 选择“网络 > 虚拟私有云”。
 - d. 选择需要添加路由表的虚拟私有云，在“路由表”页面，单击“添加路由信息”。
 - e. 根据界面提示，填写路由信息。

- 目的地址：是目的网段，默认是 0.0.0.0/0。

- 下一跳地址：是代理弹性云服务器的私有 IP 地址。

您可以在弹性云服务器页面，查看该弹性云服务器的私有 IP 地址。

21. 如需删除添加的 iptables 规则，需执行以下命令：

```
iptables -t nat -D POSTROUTING -o eth0 -s subnet/netmask-bits -j SNAT  
--to nat-instance-ip
```

假设代理弹性云服务器所在的网段为 192.168.125.0，子网掩码为 24 位，私网 IP 地址为 192.168.125.4，则执行如下命令。

```
iptables -t nat -D POSTROUTING -o eth0 -s 192.168.125.0/24 -j SNAT --to  
192.168.125.4
```

7 安全管理

7.1 提升云服务器安全性方法概述

操作场景

如果云服务器没有设置相关的安全防护，可能受到病毒入侵或外部攻击，导致数据泄露或丢失，影响业务的正常运行。

怎样保护云服务器免受攻击或病毒入侵？本节操作介绍常见的提升云服务器安全的措施。

防护类型简介

提升云服务器的安全性，分为云服务器“外部安全防护”和“内部安全防护”两方面。

表 7-1 提升云服务器安全的方法

类型	说明	防护方法
外部安全防护	常见的DDoS攻击、木马或病毒的入侵都是常见的外部安全问题。针对这类问题有多种常见的防护方案，例如开启主机安全防护您可以根据您的实际业务选择合适的防护方案。	<ul style="list-style-type: none">● 开启主机安全防护● 配置云堡垒机● 监控云服务器● 开启防DDoS攻击● 定期备份数据
内部安全防护	弱密码、开放错误的端口可能引起内部安全防护问题，不提升云服务器的内部安全防护，外部安全防护方案就无法有效的拦截和阻断各种外部攻击。	<ul style="list-style-type: none">● 增加登录密码的强度● 提升云服务器的端口安全● 定期升级操作系统

开启主机安全防护

主机安全服务（Host Security Service，HSS）是提升服务器整体安全性的服务，通过主机管理、风险防御、入侵检测、安全运营、网页防篡改功能，可全面识别并管理云服务器中的信息资产，实时监测云服务器中的风险，降低服务器被入侵的风险。

使用主机安全需要在云服务器中安装Agent。安装Agent后，您的云服务器将受到HSS云端防护中心全方位的安全保障，在安全控制台可视化界面上，您可以统一查看并管理同一区域内所有主机的防护状态和主机安全风险。

您在使用主机安全服务前，需要先在弹性云服务器上安装Agent。针对新创建的云服务器和已有的云服务器，提供不同的安装方式：

● 场景一：新创建云服务器

购买弹性云服务器，选择部分操作系统的公共镜像时，系统推荐您配套使用主机安全服务（Host Security Service，HSS）。

开启“主机安全”需要设置“安全防护”参数：

- 免费试用一个月主机安全基础防护：开启后，免费体验HSS基础版30天，为您的主机提供操作系统漏洞检测、弱口令检测、暴力破解检测等功能。

📖 说明

主机安全基础防护免费使用期限结束后，该防护配额将自动释放，停止相应的实时防护能力。

如您需要保留或升级原有安全防护能力，建议您购买高阶防护。详细情况，请参见主机安全的[免费试用HSS基础版到期后怎么办？](#)。

购买弹性云服务器时，默认设置该选项。

- 购买高阶防护：高阶防护支持基础版、企业版、旗舰版和网页防篡改版，需付费使用。

购买后您可以在主机安全服务控制台切换不同版本。各版本之间的差异请参考[服务版本差异](#)。

- 不使用安全防护：若您不需要进行安全防护，可选择此选项。

选择主机安全后系统自动安装主机安全Agent，开启账号防御，启用主机安全服务的功能。

若基础版或企业版不满足要求，您可以[购买其他版本配额](#)，在主机安全服务控制台切换不同版本，获取更高级的防护，且不需要重新安装Agent。

图 7-1 开通主机安全



● 场景二：未配置主机安全的云服务器

对于已经创建完成的弹性云服务器，可能由于创建时尚未支持主机安全服务或未设置“安全防护”参数。如需使用主机安全，您需要手动安装Agent。

具体操作请参见[手动安装Agent](#)、[手动开启防护](#)。

配置云堡垒机

云堡垒机（Cloud Bastion Host, CBH）可以实时收集和监控网络环境中每个组成部分的系统状态、安全事件和网络活动，保障网络和数据不受来自外部或内部用户的入侵和破坏，便于集中报警、及时处理及审计定责。

登录云堡垒机后，您需要依次创建用户、资源、访问策略，依次创建成功后才能对资源进行管理和运维。

- 云堡垒机支持Web浏览器、SSH客户端和MSTSC客户端三种登录方式。
Web浏览器登录：支持系统管理和资源运维功能。建议系统管理员admin或管理人员使用Web浏览器登录进行系统管理和授权审计。
SSH客户端登录：在不改变用户原来使用SSH客户端习惯的前提下，可对授权资源进行运维管理。运维人员可选择使用SSH客户端直接登录运维资源。
MSTSC客户端登录：在不改变用户原来使用MSTSC客户端习惯的前提下，可对授权资源进行运维管理。运维人员可选择使用MSTSC客户端直接登录运维资源。
- 在使用云堡垒机进行系统管理和运维前，管理人员需要在CBH系统中**创建系统用户**，为用户分配不同系统角色。
根据角色系统权限的不同，用户拥有不同的系统操作和访问权限，新创建的用户登录系统，即可访问角色权限内模块。

📖 说明

- 仅admin拥有管理系统角色的权限。
- 云堡垒机系统集中管理云资源，主要包括管理资源账户和运维权限管理。为实现统一管理资源，需**添加资源到系统**。
一个主机或应用资源可能有多个登录主机或应用的账户。
CBH系统纳管主机或应用的账户（资源账户）后，无需反复输入账户和密码，通过登录资源账户，自动登录资源进行运维管控。
系统默认资源账户Empty，登录Empty资源账户时需手动输入主机账户和对应密码。
- 若需通过云堡垒机运维资源，还需**配置访问控制策略**，关联用户和资源，赋予用户相应资源访问控制权限。
- 用户获取资源访问控制权限后，通过系统登录资源**进行运维**，运维过程被全程监控记录。
运维用户可根据资源类型选择不同登录方式。
- 用户获取相应系统权限和运维权限后，可通过云堡垒机登录已授权的资源进行运维操作，以及在系统进行系统数据管理操作。
管理员可在系统Web页面审计用户系统登录和操作，以及**审计用户运维会话**。

监控云服务器

监控是保持弹性云服务器可靠性、可用性和性能的重要部分，通过监控，用户可以观察弹性云服务器资源。为使用户更好地掌握自己的弹性云服务器运行状态，云平台提供了云监控。您可以使用该服务监控您的弹性云服务器，执行自动实时监控、告警和通知操作，帮助您更好地了解弹性云服务器的各项性能指标。

主机监控分为基础监控、操作系统监控和进程监控。

- 基础监控

基础监控无需安装Agent，是ECS自动上报的监控指标。基础监控指标的监控周期为5分钟（KVM实例）。

- 操作系统监控

操作系统监控需要在弹性云服务器中安装Agent插件，为用户提供服务器的系统级、主动式、细颗粒度监控服务。操作系统监控的监控周期为1分钟（KVM实例）。

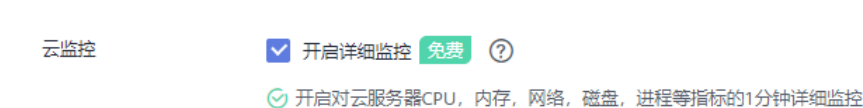
购买云服务器时的开启操作系统监控的方法：

您可以在购买时勾选“开启详细监控”，勾选后云平台将自动安装操作系统监控所需的Agent插件。

说明

当前仅部分区域的部分操作系统支持在购买时开启详细监控，实际情况请以控制台显示为准。

图 7-2 购买云服务器时开启操作系统监控



为已创建完成的云服务器开启操作系统监控的方法：

如果创建时未勾选“开启详细监控”，如需使用操作系统监控，您需要手动安装Agent。

安装配置Agent相关操作请参考云监控服务[“Agent安装配置方式说明”](#)。

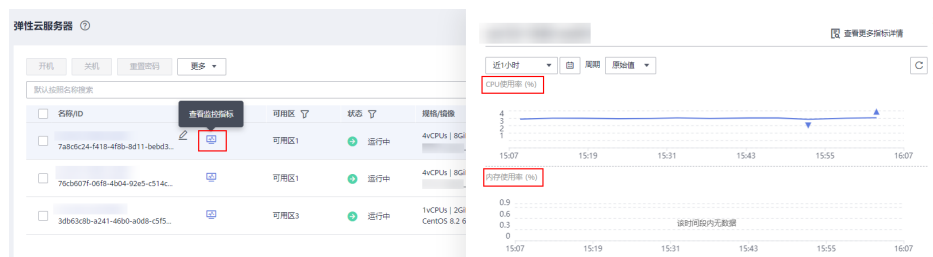
- 进程监控

进程监控需要在弹性云服务器中安装Agent插件，对主机内活跃进程进行监控，进程监控的监控周期为1分钟（KVM实例）。

开启主机监控后您可以通过设置弹性云服务器告警规则，自定义监控目标与通知策略，及时了解弹性云服务器运行状况，从而起到预警作用。

在ECS的控制台单击即可查看监控指标。

图 7-3 查看云服务器监控指标



开启防 DDoS 攻击

针对DDoS攻击，华为云提供多种安全防护方案，您可以根据您的实际业务选择合适的防护方案。华为云DDoS防护服务（Anti-DDoS Service，简称AAD）提供了DDoS原生基础防护（Anti-DDoS流量清洗）、DDoS原生专业防护和DDoS高防三个子服务。

其中，DDoS原生基础防护为免费服务，DDoS原生专业防护和DDoS高防为收费服务。关于DDoS原生专业防护和DDoS高防请单击[DDoS防护](#)了解更多。

购买弹性云服务器时，如果选择了购买弹性公网IP，那么控制台会提示您已免费开启DDoS基础防护。

图 7-4 免费开启 DDoS 基础防护



DDoS原生基础防护（Anti-DDoS流量清洗）服务（以下简称Anti-DDoS）为云服务器提供DDoS攻击防护和攻击实时告警通知。同时，Anti-DDoS可以提升用户带宽利用率，确保用户业务稳定运行。

Anti-DDoS通过对互联网访问公网IP的业务流量进行实时监测，及时发现异常DDoS攻击流量。在不影响正常业务的前提下，根据用户配置的防护策略，清洗掉攻击流量。同时，Anti-DDoS为用户生成监控报表，清晰展示网络流量的安全状况。

定期备份数据

数据备份是防止系统出现数据丢失，将全部或部分数据以其他方式保留的过程。本节以云备份方法为例，了解更多备份方法请参考[云备份概述](#)。

云备份可以为云服务器、云硬盘提供简单易用的备份服务，当发生病毒入侵、人为误删除、软硬件故障等事件时，可将数据恢复到任意备份点。保障用户数据的安全性和正确性，确保业务安全。

购买云服务器时启用云备份的方法：

购买云服务器时，设置开启云备份完成后，系统会将云服务器绑定至云备份存储库并绑定所选备份策略，定期备份弹性云服务器。

- 现在购买：
 - a. 输入云备份存储库的名称：只能由中文字符、英文字母、数字、下划线、中划线组成，且长度小于等于64个字符。例如：vault-f61e。默认的命名规则为“vault_xxxx”。
 - b. 输入存储库的容量：此容量为备份云服务器所需的容量。存储库的空间不能小于云服务器的空间。取值范围为[云服务器总容量，10485760]GB。
 - c. 设置备份策略：在下拉列表中选择备份策略，或进入云备份控制台查看或编辑备份策略。
- 使用已有：
 - a. 选择云备份存储库的名称：在下拉列表中选择已有的云备份存储库。
 - b. 设置备份策略：在下拉列表中选择备份策略，或进入云备份控制台查看或编辑备份策略。

- 暂不购买：跳过云备份的配置步骤。如云服务器购买成功后仍需设置备份保护，请进入云备份控制台找到目标存储库，绑定服务器。

图 7-5 设置云备份



已创建完成的云服务器创建云备份的方法：

备份云服务器可以通过“云服务器备份”和“云硬盘备份”功能实现：

- 云服务器备份（推荐）：如果是对弹性云服务器中的所有云硬盘（系统盘和数据盘）进行备份，推荐使用云服务器备份功能，同时对所有云硬盘进行备份，避免因备份创建时间差带来的数据不一致问题。
- 云硬盘备份：如果对指定的单个或多个云硬盘（系统盘或数据盘）进行备份，推荐使用云硬盘备份功能，在保证数据安全的同时降低备份成本。

增加登录密码的强度

“密钥对”方式创建的弹性云服务器安全性更高，建议选择“密钥对”方式。

如果您习惯使用“密码”方式，请增强密码的复杂度，保证密码符合要求，防止恶意攻击。详细内容，请参见[密码使用场景介绍](#)。

系统不会定期自动修改弹性云服务器密码。为安全起见，建议您定期修改密码。

密码设置建议：

- 密码应该长度不少于10位。
- 建议不要使用有一定特征和规律容易被破解的常用口令的密码（如：在常用彩虹表中的密码、滚键盘密码等），且密码复杂度至少包含大写字母、小写字符、数字、特殊字符四类中的三种。
- 密码尽量不要包含账户名如：adminstrator/administrator，test/test，root/root，oracle/oracle，mysql/mysql。
- 建议至少每90天更改一次密码。
- 建议不要重复使用最近5次（含5次）内已使用的密码。
- 建议根据不同应用设置不同的账号密码，不建议多个应用使用同一密码。

提升云服务器的端口安全

安全组是云服务器的守卫，是重要的网络安全隔离手段，可以保护云服务器的网络安全。安全组可以控制进出云服务器的网络流量。网络流量分为出方向和入方向，出方向是指您想访问别人，入方向就是别人想访问您。如果把云服务器比作一个宫殿，那安全组就像是一个守卫者，谁能进出，都由安全组规则控制。

通过配置安全组规则，限定云服务器出方向和入方向的访问端口，通常建议您关闭高危端口，仅开启必要的云服务器端口。

常见的高危端口如表7-2所示，建议您修改敏感端口为其它非高危端口来承载业务。请参考[云服务器常用端口](#)。

表 7-2 常见的高危端口

协议	端口
TCP	42 135 137 138 139 444 445 593 1025 1068 1434 3127 3128 3129 3130 4444 4789 5554 5800 5900 9996
UDP	135~139 1026 1027 1028 1068 1433 1434 4789 5554 9996

定期升级操作系统

云服务器申请完成后，系统内的所有配置都是需要您自行维护，云平台不负责客户系统补丁的升级，对于官方发布的一些漏洞预警，会有[安全公告](#)，需要您自行升级维护。

7.2 安全组

7.2.1 安全组概述

安全组

安全组是一个逻辑上的分组，为具有相同安全保护需求并相互信任的云服务器提供访问策略。安全组创建后，用户可以在安全组中定义各种访问规则，当云服务器加入该安全组后，即受到这些访问规则的保护。

您也可以根据需要创建自定义的安全组，或使用默认安全组，系统会为每个用户默认创建一个默认安全组，默认安全组的规则是在出方向上的数据报文全部放行，入方向访问受限，安全组内的云服务器无需添加规则即可互相访问。默认安全组您可以直接使用，详情请参见[默认安全组和规则](#)。

说明

安全组需在网络互通的情况下生效。若实例属于不同VPC，但同属于一个安全组，则此安全组不生效，您可以使用对等连接等产品建立VPC连接互通。VPC连接请参见[VPC连接](#)。

安全组规则

安全组创建后，您可以在安全组中设置出方向、入方向规则，这些规则会对安全组内部的云服务器出入方向网络流量进行访问控制，当云服务器加入该安全组后，即受到这些访问规则的保护。

每个安全组都自带默认安全组规则，详情请参见[默认安全组和规则](#)。您也可以自定义添加安全组规则，请参见[配置安全组规则](#)。

安全组的使用限制

- 为了确保良好的网络性能体验，建议一个实例最多关联5个安全组。
- 建议一个安全组关联的实例数量不应超过6000个，否则会引起安全组性能下降。
- 在一个安全组中，对于入方向规则来说，源地址是安全组的规则数量+源地址是IP地址组的规则数量+端口是不连续端口号的规则数量 ≤ 120条，否则超过数量的安全组规则将不生效。当同时存在IPv4和IPv6类型的安全组规则时，两种类型的安全组规则单独计算，即IPv4规则和IPv6规则可以各有120条。

对于安全组出方向规则来说，目的地址和端口存在一样的限制。

以安全组Sg-A的入方向IPv4规则为例，表7-3中提供了部分符合限制条件的规则供您参考。其中，当一条安全组规则同时符合多个限制时，比如规则A02即使用了不连续端口，又使用了安全组作为源地址，此时只占用一条配额。

表 7-3 入方向安全组规则示例

规则编号	策略	类型	协议端口	源地址
规则A01	允许	IPv4	全部	当前安全组: Sg-A
规则A02	允许	IPv4	TCP: 22,25,27	其他安全组: Sg-B
规则A03	允许	IPv4	TCP: 80-82	IP地址组: ipGroup-A
规则A04	允许	IPv4	TCP: 22-24,25	IP地址: 192.168.0.0/16

- 如果您添加安全组规则时，使用IP地址组或者不连续端口，那么该安全组规则对不同规格云服务器的生效情况存在差异，为了避免您的安全组规则不生效，请您查看表7-4了解详情。

表 7-4 安全组规则限制

安全组规则	云服务器类型
添加安全组规则时，“源地址”和“目的地址”可选择“IP地址组”	不支持的X86云服务器规格如下： <ul style="list-style-type: none"> • 通用计算型（S1型、C1型、C2型） • 内存优化型（M1型） • 高性能计算型（H1型） • 磁盘增强型（D1型） • GPU加速型（G1型、G2型） • 超大内存型（E1型、E2型、ET2型）

安全组规则	云服务器类型
添加安全组规则时，“协议端口”可配置为不连续端口号	不支持的X86云服务器规格如下： <ul style="list-style-type: none">● 通用计算型（S1型、C1型、C2型）● 内存优化型（M1型）● 高性能计算型（H1型）● 磁盘增强型（D1型）● GPU加速型（G1型、G2型）● 超大内存型（E1型、E2型、ET2型）
	所有鲲鹏云服务器规格不支持配置不连续端口。 如果您在鲲鹏云服务器中添加安全组规则时，使用了不连续端口号，那么除了该条规则不会生效，该规则后的其他规则也不会生效。比如： 您先配置了安全组规则A（不连续端口号22,24），再配置了下一条安全组规则B（独立端口号9096），则安全组规则A和B均不会生效。

📖 说明

- X86云服务器规格详情，请参见[规格清单（X86）](#)。
- 鲲鹏云服务器规格详情，请参见[规格清单（鲲鹏）](#)。
- 当您的组网中存在以下情况时，来自ELB流量不受网络ACL和安全组规则的限制。
ELB实例的监听器开启“获取客户端IP”功能时，不受限制。
比如规则已明确拒绝来自ELB实例的流量进入后端云服务器，此时该规则无法拦截ELB的流量，流量依然会抵达后端云服务器。

实践建议

- 请您遵循白名单原则配置安全组规则，即安全组内实例默认拒绝所有外部的访问请求，通过添加允许规则放通指定的网络流量。
- 添加安全组规则时，请遵循最小授权原则。例如，放通22端口用于远程登录云服务器时，建议仅允许指定的IP地址登录，谨慎使用0.0.0.0/0（所有IP地址）。
- 请您尽量保持单个安全组内规则的简洁，通过不同的安全组来管理不同用途的实例。如果您使用一个安全组管理您的所有业务实例，可能会导致单个安全组内的规则过于冗余复杂，增加维护管理成本。
- 您可以将实例按照用途加入到不同的安全组内。例如，当您具有面向公网提供网站访问的业务时，建议您将运行公网业务的Web服务器加入到同一个安全组，此时仅需要放通对外部提供服务的特定端口，例如80、443等，默认拒绝外部其他的访问请求。同时，请避免在运行公网业务的Web服务器上运行内部业务，例如MySQL、Redis等，建议您将内部业务部署在不需要连通公网的服务器上，并将这些服务器关联至其他安全组内。
- 对于安全策略相同的多个IP地址，您可以将其添加到一个IP地址组内统一管理，并在安全组内添加针对该IP地址组的授权规则。当IP地址发生变化时，您只需要在IP地址组内修改IP地址，那么IP地址组对应的安全组规则将会随之变更，无需逐次修

改安全组内的规则，降低了安全组管理的难度，提升效率。具体方法，请参见[使用IP地址组提升安全组规则管理效率](#)。

- 请您尽量避免直接修改已运行业务的安全组规则。如果您需要修改使用中的安全组规则，建议您先克隆一个测试安全组，然后在测试安全组上进行调试，确保测试安全组内实例网络正常后，再修改使用中的安全组规则，减少对业务的影响。具体方法，请参见[克隆安全组](#)。
- 您在安全组内新添加实例，或者修改安全组的规则后，此时不需要重启实例，安全组规则会自动生效。

如果您的安全组规则配置完未生效，请参考[为什么配置的安全组规则不生效?](#)。

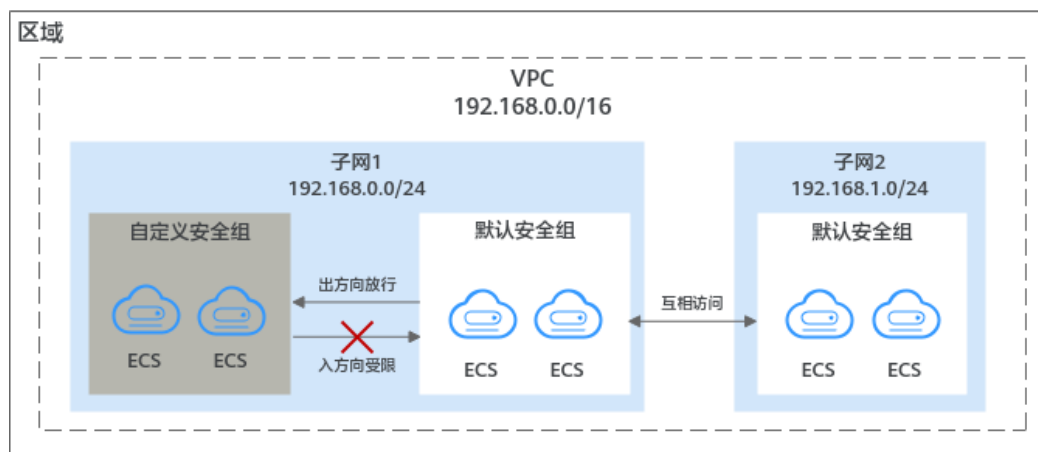
7.2.2 默认安全组和规则

默认安全组规则说明如下：

- 入方向规则：入方向流量受限，只允许安全组内实例互通，拒绝来自安全组外部的所有请求进入实例。
- 出方向规则：出方向流量放行，允许所有请求从安全组内实例流出。

如[图7-6](#)所示。

图 7-6 默认安全组



说明

- 默认安全组名称为default，默认安全组和您创建的自定义安全组均不收取费用。
- 您无法删除默认安全组，可以在默认安全组内修改已有规则或者添加新的规则。
- 默认安全组是为了简化您初次创建实例的流程，系统为您自动创建的。默认安全组拒绝所有外部请求，如果您需要登录实例，请参见[通过本地服务器远程登录云服务器](#)添加安全组规则。

默认安全组规则如[表7-5](#)所示：

表 7-5 默认安全组规则

方向	策略	类型	协议端口	源地址/目的地址	描述
入方向	允许	IPv4	全部	源地址：默认安全组 (default)	针对全部IPv4协议，允许安全组内的实例可使用任何协议和端口互相通信，确保安全组内实例网络互通。
入方向	允许	IPv6	全部	源地址：默认安全组 (default)	针对全部IPv6协议，允许安全组内的实例可使用任何协议和端口互相通信，确保安全组内实例网络互通。
出方向	允许	IPv4	全部	目的地址：0.0.0.0/0	针对全部IPv4协议，允许所有流量从安全组内实例流出，即实例可访问外部任意IP和端口。
出方向	允许	IPv6	全部	目的地址：::/0	针对全部IPv6协议，允许所有流量从安全组内实例流出，即实例可访问外部任意IP和端口。

首次创建弹性云服务器时，系统会自动新建虚拟私有云vpc-default，同时对安全组进行如下操作：

- 新增Sys-WebServer安全组
- 新增Sys-FullAccess安全组
- 在默认安全组default中新增安全组规则

表 7-6 默认 default 安全组规则

规则方向	策略	类型	协议端口	源地址/目的地址	描述
入方向规则	允许	IPv4	TCP: 3389	源地址：0.0.0.0/0	针对全部IPv4协议，允许所有IP地址通过默认Windows远程桌面连接到Windows云服务器。
入方向规则	允许	IPv4	TCP: 22	源地址：0.0.0.0/0	针对全部IPv4协议，允许所有IP地址通过SSH远程连接到Linux云服务器。
入方向规则	允许	IPv4	全部	源地址：默认安全组 (default)	针对全部IPv4协议，允许本安全组内实例的请求进入，即该条规则确保安全组内的实例网络互通。
入方向规则	允许	IPv6	全部	源地址：默认安全组 (default)	针对全部IPv6协议，允许本安全组内实例的请求进入，即该条规则确保安全组内的实例网络互通。
出方向规则	允许	IPv4	全部	目的地址：0.0.0.0/0	针对全部IPv4协议，允许安全组内的实例可访问外部IP的所有端口。
出方向规则	允许	IPv6	全部	目的地址：::/0	针对全部IPv6协议，允许安全组内的实例可访问外部IP的所有端口。

表 7-7 Sys-WebServer 安全组规则

规则方向	策略	类型	协议端口	目的地址/源地址	描述
入方向规则	允许	IPv4	ICMP: 全部	源地址: 0.0.0.0/0	针对全部IPv4协议, 允许在云服务器上使用ping命令验证网络的连通性。
入方向规则	允许	IPv4	全部	源地址: 当前安全组(Sys-WebServer)	针对全部IPv4协议, 允许本安全组内实例的请求进入, 即该条规则确保安全组内的实例网络互通。
入方向规则	允许	IPv4	TCP: 443	源地址: 0.0.0.0/0	针对全部IPv4协议, 允许所有IP地址通过HTTPS协议访问云服务器上搭建的网站。
入方向规则	允许	IPv4	TCP: 80	源地址: 0.0.0.0/0	针对全部IPv4协议, 允许所有IP地址通过HTTP协议访问云服务器上搭建的网站。
入方向规则	允许	IPv4	TCP: 22	源地址: 0.0.0.0/0	针对全部IPv4协议, 允许所有IP地址通过SSH远程连接到Linux云服务器。
入方向规则	允许	IPv4	TCP: 3389	源地址: 0.0.0.0/0	针对全部IPv4协议, 允许所有IP地址通过默认Windows远程桌面连接到Windows云服务器。
入方向规则	允许	IPv6	全部	源地址: 当前安全组(Sys-WebServer)	针对全部IPv6协议, 允许本安全组内实例的请求进入, 即该条规则确保安全组内的实例网络互通。
出方向规则	允许	IPv4	全部	目的地址: 0.0.0.0/0	针对全部IPv4协议, 允许安全组内的实例可访问外部IP的所有端口。
出方向规则	允许	IPv6	全部	目的地址: ::/0	针对全部IPv6协议, 允许安全组内的实例可访问外部IP的所有端口。

表 7-8 Sys-FullAccess 安全组规则

规则方向	策略	类型	协议端口	目的地址/源地址	说明
入方向规则	允许	IPv4	全部	源地址: 当前安全组(Sys-FullAccess)	针对全部IPv4协议, 允许本安全组内实例的请求进入, 即该条规则确保安全组内的实例网络互通。
入方向规则	允许	IPv6	全部	源地址: 当前安全组(Sys-FullAccess)	针对全部IPv6协议, 允许本安全组内实例的请求进入, 即该条规则确保安全组内的实例网络互通。
入方向规则	允许	IPv4	全部	源地址: 0.0.0.0/0	针对全部IPv4协议, 允许所有入站流量的数据报文通过。
入方向规则	允许	IPv6	全部	源地址: ::/0	针对全部IPv6协议, 允许所有入站流量的数据报文通过。

规则方向	策略	类型	协议端口	目的地址/源地址	说明
出方向规则	允许	IPv4	全部	目的地址：0.0.0.0/0	针对全部IPv4协议，允许安全组内的实例可访问外部IP的所有端口。
出方向规则	允许	IPv6	全部	目的地址：::/0	针对全部IPv6协议，允许安全组内的实例可访问外部IP的所有端口。

7.2.3 安全组配置示例

当您在VPC子网内创建实例（云服务器、云容器、云数据库等）时，您可以使用系统提供的默认安全组default，您也可以创建其他安全组。无论是默认安全组，还是您创建的安全组，您均可以在安全组内设置出方向和入方向规则，以此控制出入实例的流量。以下为您介绍一些常用的安全组的配置示例：

- [通过本地服务器远程登录云服务器](#)
- [在本地服务器远程连接云服务器上传或者下载文件（FTP）](#)
- [在云服务器上搭建网站对外提供Web服务](#)
- [使用ping命令验证网络连通性](#)
- [实现不同安全组的实例内网网络互通](#)
- [云服务器提供数据库访问服务](#)
- [限制云服务器访问外部网站](#)

须知

如果您的安全组规则配置完成后不生效，请您[提交工单](#)联系客服处理。

使用须知

在配置安全组规则之前，您需要先了解以下信息：

- 不同安全组之间的实例默认网络隔离，无法互相访问。
- 安全组默认拒绝所有来自外部的请求，即本安全组内的实例网络互通，外部无法访问安全组内的实例。
您需要遵循白名单原则添加安全组入方向规则，允许来自外部的特定请求访问安全组内的实例。
- 安全组入方向规则的源地址设置为0.0.0.0/0或::/0，表示允许或拒绝所有外部IP地址访问您的实例，如果将“22、3389、8848”等[高危端口](#)暴露到公网，可能导致网络入侵，造成业务中断、数据泄露或数据勒索等严重后果。建议您将安全组规则设置为仅允许已知的IP地址访问。
- 安全组的出方向规则一般默认全部放通，即允许安全组内的实例访问外部。
如果出方向规则被删除，将会导致安全组内实例无法正常访问外部，您可以参考[表7-9](#)重新添加规则。

表 7-9 安全组默认出方向规则

方向	优先级	策略	类型	协议端口	目的地址	描述
出方向	1	允许	IPv4	全部	0.0.0.0/0	针对全部IPv4协议，允许安全组内的实例可访问外部任意IP和端口。
出方向	1	允许	IPv6	全部	::/0	针对全部IPv6协议，允许安全组内的实例可访问外部任意IP和端口。

通过本地服务器远程登录云服务器

安全组默认拒绝所有来自外部的请求，如果您需要通过本地服务器远程登录安全组内的云服务器，那么需要根据您的云服务器操作系统类型，在安全组入方向添加对应的规则。

- 通过SSH远程登录Linux云服务器，需要放通SSH(22)端口，请参见[表7-10](#)。
- 通过RDP远程登录Windows云服务器，需要放通RDP(3389)端口，请参见[表7-11](#)。

表 7-10 通过 SSH 远程登录 Linux 云服务器

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	自定义TCP: 22	IP地址: 0.0.0.0/0

表 7-11 通过 RDP 远程登录 Windows 云服务器

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	自定义TCP: 3389	IP地址: 0.0.0.0/0

须知

源地址设置为0.0.0.0/0表示允许所有外部IP地址远程登录云服务器，如果将当前端口暴露到公网，可能存在网络安全风险，建议您将源IP设置为已知的IP地址，配置示例请参见[表7-12](#)。

表 7-12 通过已知 IP 地址远程登录云服务器

云服务器类型	方向	优先级	策略	类型	协议端口	源地址
Linux云服务器	入方向	1	允许	IPv4	自定义 TCP: 22	IP地址: 192.168.0.0/24
Windows云服务器	入方向	1	允许	IPv4	自定义 TCP: 3389	IP地址: 10.10.0.0/24

在本地服务器远程连接云服务器上传或者下载文件（FTP）

安全组默认拒绝所有来自外部的请求，如果您需要在本地服务器远程连接云服务器上传或者下载文件，那么您需要开通FTP(20、21)端口。

表 7-13 通过任意服务器远程连接云服务器上传或者下载文件

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	自定义TCP: 20-21	IP地址: 0.0.0.0/0

须知

- 源地址设置为0.0.0.0/0表示允许所有外部IP地址远程连接云服务器上传或者下载文件，如果将当前端口暴露到公网，可能存在网络安全风险，建议您将源IP设置为已知的IP地址，配置示例请参见表7-14。
- 您需要在弹性云服务器上先安装FTP服务器程序，再查看20、21端口是否正常工作。

表 7-14 通过已知的服务器远程连接云服务器上传或者下载文件

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	自定义TCP: 20-21	IP地址: 192.168.0.0/24

在云服务器上搭建网站对外提供 Web 服务

安全组默认拒绝所有来自外部的请求，如果您在云服务器上搭建了可供外部访问的网站，则您需要在安全组入方向添加对应的规则，放通对应的端口，例如HTTP(80)、HTTPS(443)。

表 7-15 在云服务器上搭建网站对外提供 Web 服务

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	自定义TCP: 80	IP地址: 0.0.0.0/0
入方向	1	允许	IPv4	自定义TCP: 443	IP地址: 0.0.0.0/0

使用 ping 命令验证网络连通性

ICMP协议用于网络消息的控制和传递，因此在进行一些基本测试操作之前，需要开通ICMP协议访问端口。比如，您需要在某个个人PC上使用ping命令来验证云服务器的网络连通性，则您需要在云服务器所在安全组的入方向添加以下规则，放通ICMP端口。

表 7-16 使用 ping 命令验证网络连通性

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	ICMP: 全部	IP地址: 0.0.0.0/0
入方向	1	允许	IPv6	ICMP: 全部	IP地址: ::/0

实现不同安全组的实例内网网络互通

同一个VPC内，位于不同安全组内的实例网络不通。如果您需要在同一个VPC内的实例之间共享数据，比如安全组sg-A内的云服务器访问安全组sg-B内的MySQL数据库，您需要通过在安全组sg-B中添加一条入方向规则，允许来自安全组sg-A内云服务器的内网请求进入。

表 7-17 实现不同安全组的实例网络互通

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	自定义TCP: 3306	安全组: sg-A

须知

如果您通过中间网络实例在不同子网的实例之间转发流量，如[安全组应用示例](#)中的“控制虚拟IP访问安全组内实例”，子网Subnet-A的ECS通过虚拟IP和子网Subnet-B的ECS互相通信。由于存在中间网络实例，此时安全组规则的源地址选择实例所在的安全组时，无法放通中间网络实例转发的流量，源地址必须设置成中间网络实例的私有IP地址或者子网网段。

云服务器提供数据库访问服务

安全组默认拒绝所有来自外部的请求，如果您在云服务器上部署了数据库服务，允许其他实例通过内网访问数据库服务，则您需要在部署数据库服务器所在的安全组内，添加入方向规则，放通对应的端口，实现其他实例通过内网获取数据库数据的请求。常见的数据库类型及其对应的端口如下：

- MySQL(3306)
- Oracle(1521)
- MS SQL(1433)
- PostgreSQL(5432)
- Redis(6379)

表 7-18 云服务器提供数据库访问服务

方向	优先级	策略	类型	协议端口	源地址	描述
入方向	1	允许	IPv4	自定义 TCP: 3306	安全组: sg-A	允许安全组sg-A内云服务器访问MySQL数据库服务。
入方向	1	允许	IPv4	自定义 TCP: 1521	安全组: sg-B	允许安全组sg-B内云服务器访问Oracle数据库服务。
入方向	1	允许	IPv4	自定义 TCP: 1433	IP地址: 172.16.3.2 1/32	允许私网IP地址为172.16.3.21的云服务器访问MS SQL数据库服务。
入方向	1	允许	IPv4	自定义 TCP: 5432	IP地址: 192.168.0. 0/24	允许私网IP地址属于192.168.0.0/24网段的云服务器访问PostgreSQL数据库服务。
入方向	1	允许	IPv4	自定义 TCP: 6379	IP地址组: ipGroup-A	允许私网IP地址属于IP地址组ipGroup-A范围内的云服务器访问Redis数据库服务。

须知

本示例中源地址提供的配置仅供参考，请您根据实际需求设置源地址。

限制云服务器访问外部网站

安全组的出方向规则一般默认全部放通，默认规则如表7-20所示。如果您需要限制服务器只能访问特定网站，则按照如下要求配置：

1. 首先，您需要遵循白名单规则，在安全组出方向规则中添加指定的端口和IP地址。

表 7-19 限制云服务器访问外部网站

方向	优先级	策略	类型	协议端口	目的地址	描述
出方向	1	允许	IP v4	自定义 TCP: 80	IP地址: 132.15.XX.XX	允许安全组内云服务器访问指定的外部网站，网站地址为 http://132.15.XX.XX:80。
出方向	1	允许	IP v4	自定义 TCP: 443	IP地址: 145.117.XX.XX	允许安全组内云服务器访问指定的外部网站，网站地址为 https://145.117.XX.XX:443。

2. 其次，删除安全组出方向中原有放通全部流量的规则，如表7-20所示。

表 7-20 安全组默认出方向规则

方向	优先级	策略	类型	协议端口	目的地址	描述
出方向	1	允许	IPv 4	全部	0.0.0.0/0	针对全部IPv4协议，允许安全组内的实例可访问外部任意IP和端口。
出方向	1	允许	IPv 6	全部	::/0	针对全部IPv6协议，允许安全组内的实例可访问外部任意IP和端口。

7.2.4 配置安全组规则

操作场景

安全组类似防火墙功能，是一个逻辑上的分组，用于设置网络访问控制。用户可以在安全组中定义各种访问规则，当弹性云服务器加入该安全组后，即受到这些访问规则的保护。

- 入方向规则：放通入方向网络流量，指从外部访问安全组规则下的云服务器。
- 出方向规则：放通出方向网络流量。指安全组规则下的云服务器访问安全组外的实例。

默认安全组规则请参见[默认安全组和规则](#)。常用的安全组规则配置示例请参见[安全组配置示例](#)。

操作步骤

1. 进入[弹性云服务器列表](#)页面。
2. 在弹性云服务器列表，单击待变更安全组规则的弹性云服务器名称。系统跳转至该弹性云服务器详情页面。

- 选择“安全组”页签，展开安全组，查看安全组规则。
- 单击安全组ID。
系统自动跳转至安全组页面。
- 根据界面提示，设置入方向规则参数。
单击 \oplus ，可以依次增加多条入方向规则。

图 7-7 添加安全组入方向规则

表 7-21 入方向规则参数说明

参数	说明	取值样例
优先级	安全组规则优先级。 优先级可选范围为1-100，默认值为1，即最高优先级。优先级数字越小，规则优先级级别越高。	1
策略	安全组规则策略，支持的策略如下： <ul style="list-style-type: none"> 如果“策略”设置为允许，表示允许源地址访问安全组内云服务器的指定端口。 如果“策略”设置为拒绝，表示拒绝源地址访问安全组内云服务器的指定端口。 安全组规则匹配流量时，首先按照优先级进行排序，其次按照策略排序，拒绝策略高于允许策略，更多信息请参见 流量匹配安全组规则的顺序 。	允许
类型	源地址支持的IP地址类型，如下： <ul style="list-style-type: none"> IPv4 IPv6 	IPv4
协议端口	安全组规则中用来匹配流量的网络协议类型，目前支持TCP、UDP、ICMP和GRE协议。	TCP

参数	说明	取值样例
	<p>安全组规则中用来匹配流量的目的端口，取值范围为：1~65535。</p> <p>在入方向规则中，表示外部访问安全组内实例的指定端口。</p> <p>端口填写支持下格式：</p> <ul style="list-style-type: none"> ● 单个端口：例如22 ● 连续端口：例如22-30 ● 多个端口：例如22,23-30，一次最多支持20个不连续端口组，端口组之间不能重复。 ● 全部端口：为空或1-65535 	22或22-30
源地址	<p>在入方向规则中，用来匹配外部请求的源地址，支持以下格式：</p> <ul style="list-style-type: none"> ● IP地址：表示源地址为某个固定的IP地址。当源地址选择IP地址时，您可以在一个框内同时输入或者粘贴多个IP地址，不同IP地址以“,”隔开。一个IP地址生成一条安全组规则。 <ul style="list-style-type: none"> - 单个IP地址：IP地址/掩码。 单个IPv4地址示例为192.168.10.10/32。 单个IPv6地址示例为2002:50::44/128。 - IP网段：IP地址/掩码。 IPv4网段示例为192.168.52.0/24。 IPv6网段示例为2407:c080:802:469::/64。 - 所有IP地址： 0.0.0.0/0表示匹配所有IPv4地址。 ::/0表示匹配所有IPv6地址。 ● 安全组：表示源地址为另外一个安全组，您可以在下拉列表中，选择同一个区域内的其他安全组。当安全组A内有实例a，安全组B内有实例b，在安全组A的入方向规则中，放通源地址为安全组B的流量，则来自实例b的内网访问请求被允许进入实例a。 ● IP地址组：表示源地址为一个IP地址组，IP地址组是一个或者多个IP地址的集合。您可以在下拉列表中，选择可用的IP地址组。对于安全策略相同的IP网段和IP地址，此处建议您使用IP地址组简化管理。 如果没有可用的IP地址组，请您参见创建IP地址组进行创建。 	IP地址： 192.168.52.0/24,10.0.0.0/24
描述	<p>安全组规则的描述信息，非必填项。</p> <p>描述信息内容不能超过255个字符，且不能包含“<”和“>”。</p>	-

6. 根据界面提示，设置出方向规则参数。


单击 ，可以依次增加多条出方向规则。

图 7-8 添加安全组出方向规则



表 7-22 出方向规则参数说明

参数	说明	取值样例
优先级	安全组规则优先级。 优先级可选范围为1-100，默认值为1，即最高优先级。优先级数字越小，规则优先级级别越高。	1
策略	安全组规则策略，支持的策略如下： <ul style="list-style-type: none"> 如果“策略”设置为允许，表示允许安全组内的云服务器访问目的地址的指定端口。 如果“策略”设置为拒绝，表示拒绝安全组内的云服务器访问目的地址的指定端口。 安全组规则匹配流量时，首先按照优先级进行排序，其次按照策略排序，拒绝策略高于允许策略，更多信息请参见 流量匹配安全组规则的顺序 。	允许
类型	目的地址支持的IP地址类型，如下： <ul style="list-style-type: none"> IPv4 IPv6 	IPv4
协议端口	安全组规则中用来匹配流量的网络协议类型，目前支持TCP、UDP、ICMP和GRE协议。	TCP

参数	说明	取值样例
	<p>安全组规则中用来匹配流量的目的端口，取值范围为：1~65535。</p> <p>在出方向规则中，表示安全组内实例访问外部地址的指定端口。</p> <p>端口填写支持下格式：</p> <ul style="list-style-type: none"> • 单个端口：例如22 • 连续端口：例如22-30 • 多个端口：例如22,23-30，一次最多支持20个不连续端口组，端口组之间不能重复。 • 全部端口：为空或1-65535 	22或22-30
目的地址	<p>在出方向规则中，用来匹配内部请求的目的地址。支持以下格式：</p> <ul style="list-style-type: none"> • IP地址：表示目的地址为某个固定的IP地址。当目的地址选择IP地址时，您可以在一个框内同时输入或者粘贴多个IP地址，不同IP地址以“,”隔开。一个IP地址生成一条安全组规则。 <ul style="list-style-type: none"> - 单个IP地址：IP地址/掩码。 单个IPv4地址示例为192.168.10.10/32。 单个IPv6地址示例为2002:50::44/128。 - IP网段：IP地址/掩码。 IPv4网段示例为192.168.52.0/24。 IPv6网段示例为2407:c080:802:469::/64。 - 所有IP地址： 0.0.0.0/0表示匹配所有IPv4地址。 ::/0表示匹配所有IPv6地址。 • 安全组：表示目的地址为另外一个安全组，您可以在下拉列表中，选择当前账号下，同一个区域内的其他安全组。当安全组A内有实例a，安全组B内有实例b，在安全组A的出方向规则放通目的地址为安全组B的流量，则实例a访问实例b的内网请求被允许流出。 • IP地址组：表示目的地址为一个IP地址组，IP地址组是一个或者多个IP地址的集合。您可以在下拉列表中，选择可用的IP地址组。对于安全策略相同的IP网段和IP地址，此处建议您使用IP地址组简化管理。 如果没有可用的IP地址组，请您参见创建IP地址组进行创建。 	IP地址： 192.168.52.0/24,10.0.0.0/24
描述	<p>安全组规则的描述信息，非必填项。</p> <p>描述信息内容不能超过255个字符，且不能包含“<”和“>”。</p>	-

7. 单击“确定”，完成安全组规则配置。

检查安全组规则是否生效

在安全组规则中放开某个端口后，您还需要确保实例内对应的端口也已经放通，安全组规则才会对实例生效。

假设您在某台ECS上部署了网站，希望用户能通过HTTP(80)端口访问到您的网站，则您需要在ECS所在安全组的入方向中，添加表7-23中的规则，放通HTTP(80)端口。

表 7-23 安全组规则示例

方向	优先级	策略	类型	协议端口	源地址
入方向	1	允许	IPv4	自定义TCP: 80	IP地址: 0.0.0.0/0

安全组规则添加完成后，您需要执行以下操作，检查云服务器内端口开放情况，并验证配置是否生效。

1. 登录云服务器，检查云服务器端口开放情况。

- **检查Linux云服务器端口**

执行以下命令，查看TCP 80端口是否被监听。

```
netstat -an | grep 80
```

若回显类似图7-9，说明80端口已开通。

图 7-9 Linux TCP 80 端口验证结果

```
tcp      0      0 0.0.0.0:80          0.0.0.0:*          LISTEN
```

- **检查Windows云服务器端口**

- i. 通过“开始菜单 > 运行 > cmd”，打开命令执行窗口。
- ii. 执行以下命令，查看TCP 80端口是否被监听。

```
netstat -an | findstr 80
```

若回显类似图7-10，说明TCP 80端口已开通。

图 7-10 Windows TCP 80 端口验证结果

```
TCP      0.0.0.0:80          0.0.0.0:0          LISTENING
```

2. 打开浏览器，在地址栏里输入“http://云服务器的弹性公网IP地址”。
如果访问成功，说明安全组规则已经生效。

7.2.5 更改安全组

操作场景

本节操作介绍当弹性云服务器的网卡需要变更所属安全组时的操作步骤。

约束限制

- 更改安全组操作会覆盖弹性云服务器原有的安全组设置。
- 使用多个安全组可能会影响弹性云服务器的网络性能，建议您选择安全组的数量不多于5个。

操作步骤

1. 登录管理控制台。
 2. 单击“☰”，选择“计算 > 弹性云服务器”。
 3. 在弹性云服务器列表中，单击“操作”列下的“更多 > 网络/安全组 > 更改安全组”。
- 系统弹窗显示“更改安全组”页面。

图 7-11 更改安全组



4. 根据界面提示，在下拉列表中选择待更改安全组的网卡，并重新选择安全组。您可以同时勾选多个安全组，弹性云服务器的访问规则遵循几个安全组规则的并集。如需创建新的安全组，请单击“新建安全组”。

说明

使用多个安全组可能会影响弹性云服务器的网络性能，建议您选择安全组的数量不多于5个。

5. 单击“确定”。

7.3 主机安全

什么是主机安全

主机安全服务（Host Security Service，HSS）是提升服务器整体安全性的服务，通过主机管理、风险防御、入侵检测、安全运营、网页防篡改功能，可全面识别并管理云服务器中的信息资产，实时监测云服务器中的风险，降低服务器被入侵的风险。

使用主机安全需要在云服务器中安装Agent。安装Agent后，您的云服务器将受到HSS云端防护中心全方位的安全保障，在安全控制台可视化界面上，您可以统一查看并管理同一区域内所有主机的防护状态和主机安全风险。

怎样使用主机安全服务

您在使用主机安全服务前，需要先在弹性云服务器上安装Agent。针对新创建的云服务器和已有的云服务器，提供不同的安装方式：

● 场景一：新创建云服务器

购买弹性云服务器，选择部分操作系统的公共镜像时，系统推荐您配套使用主机安全服务（Host Security Service, HSS）。

开启“主机安全”需要设置“安全防护”参数：

- 免费试用一个月主机安全基础防护：开启后，免费体验HSS基础版30天，为您的主机提供操作系统漏洞检测、弱口令检测、暴力破解检测等功能。

📖 说明

主机安全基础防护免费使用期限结束后，该防护配额将自动释放，停止相应的实时防护能力。

如您需要保留或升级原有安全防护能力，建议您购买高阶防护。详细情况，请参见主机安全的[免费试用HSS基础版到期后怎么办？](#)。

购买弹性云服务器时，默认设置该选项。

- 购买高阶防护：高阶防护支持基础版、企业版、旗舰版和网页防篡改版，需付费使用。

购买后您可以在主机安全服务控制台切换不同版本。各版本之间的差异请参考[服务版本差异](#)。

- 不使用安全防护：若您不需要进行安全防护，可选择此选项。

选择主机安全后系统自动安装主机安全Agent，开启账号防御，启用主机安全服务的功能。

若基础版或企业版不满足要求，您可以[购买其他版本配额](#)，在主机安全服务控制台切换不同版本，获取更高级的防护，且不需要重新安装Agent。

图 7-12 开通主机安全



● 场景二：未配置主机安全的云服务器

对于已经创建完成的弹性云服务器，可能由于创建时尚未支持主机安全服务或未设置“安全防护”参数。如需使用主机安全，您需要手动安装Agent。

具体操作请参见[手动安装Agent](#)、[手动开启防护](#)。

怎样查看主机安全状态

在云服务器的列表页，您可以查看当前区域下的云服务器的主机安全状态。

1. 登录管理控制台。
2. 单击“☰”，选择“安全与合规 > 主机安全服务”。

3. 在“资产管理 > 主机管理”的“云服务器”页签，查看服务器的防护状态。

图 7-13 查看主机安全状态



表 7-24 状态说明

参数	说明
Agent状态	<ul style="list-style-type: none"> ● 未安装：未安装Agent，或Agent已安装但未成功启动。 ● 在线：Agent运行正常。 ● 离线：Agent与HSS服务器通信异常，HSS无法提供安全防护功能。单击“离线”，您可以查看Agent不在线的华为云主机列表，并查看“离线原因”。
防护状态	<ul style="list-style-type: none"> ● 开启：HSS为该服务器提供全面的主机安全防护。 ● 关闭：单击“关闭防护”可以暂停HSS对服务器的防护，降低该服务器的资源消耗。
检测结果	<ul style="list-style-type: none"> ● 有风险：主机存在风险。 ● 无风险：主机暂未发现风险。 ● 未检测：主机未开启防护。

更多关于主机安全服务的使用，请参见[主机安全服务](#)。

7.4 云堡垒机

什么是云堡垒机

云堡垒机（Cloud Bastion Host, CBH）是华为云的一款4A统一安全管控平台，为企业提供集中的账号（Account）、授权（Authorization）、认证（Authentication）和审计（Audit）管理服务。

云堡垒机提供云计算安全管控的系统 and 组件，包含部门、用户、资源、策略、运维、审计等功能模块，集单点登录、统一资产管理、多终端访问协议、文件传输、会话协同等功能于一体。通过统一运维登录入口，基于协议正向代理技术和远程访问隔离技术，实现对服务器、云主机、数据库、应用系统等云上资源的集中管理和运维审计。

快速配置云堡垒机

配置后可以实时收集和监控网络环境中每个组成部分的系统状态、安全事件和网络活动，保障网络和数据不受来自外部或内部用户的入侵和破坏，便于集中报警、及时处理及审计定责。

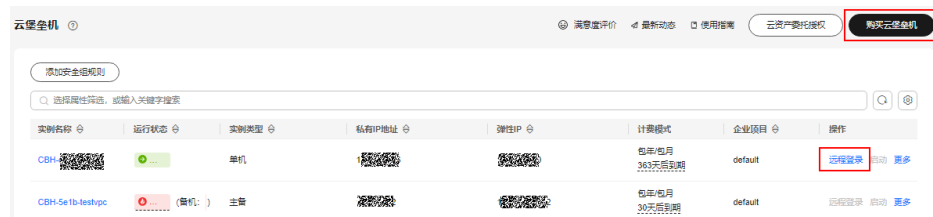
步骤1 登录管理控制台。

步骤2 在控制台页面中选择“安全与合规 > 云堡垒机 CBH”，进入服务在右上角单击**购买云堡垒机**，购买后在实例“操作”列单击“远程登录”进入云堡垒机登录页面。

说明

- 首次以系统管理员admin登录系统，登录密码为购买实例时配置密码。系统管理员admin是系统默认用户，即第一个可登录用户，拥有系统最高操作权限，且无法删除和更改权限配置。
- 系统所有用户首次登录系统后，请按照提示修改密码和绑定手机号码。

图 7-14 登录云堡垒机



步骤3 登录云堡垒机系统后，选择“用户 > 用户管理”，右上角单击“新建”，根据实际需求和界面提示**创建运维用户**。

说明

- 角色划分用户系统操作权限，默认有系统管理员、策略管理员、审计管理员和运维员四种角色。admin可以**自定义角色**，分配系统操作权限。
- “登录名”在整个CBH系统是唯一的，不能重复。
- 创建用户成功后，可通过修改用户配置，配置用户登录云堡垒机系统的**多因子认证**。

图 7-15 新建运维用户

新建用户

* 登录名
长度为1-64个字符，以字母或者数字开头，不支持的字符:\|:;|=,*?<>@*以及空格

* 认证类型

* 密码

* 确认密码
长度为8-32个字符，密码只能包含大写字母、小写字母、数字和特殊字符(!@\$%^_-=+[];./?~#*)且至少包含四种字符中的三种，不能包含用户名或倒序用户名

* 姓名
长度为1-255个汉字或字符，允许输入汉字、字母、数字、“@”、“.”、“_”或“-”

* 手机
手机号十分重要，请输入正确的手机号码。若是国际号码，请输入：“+”+国家代码+手机号码

步骤4 运维用户创建成功后，选择“资源 > 主机管理”，右上角单击“新建”，[配置主机](#)基本信息和网络参数。

说明

- “主机地址”为主机与云堡垒机网络通畅的IP地址，可选择主机的EIP地址或私有IP地址，建议优先选择可用的私有IP地址。
- 增强版支持以主机方式运维数据库，支持四种数据库类型MySQL、SQL Server、Oracle和DB2。
- 应用发布资源是通过Windows远程访问应用实现运维，需先[配置应用发布服务器](#)。
- 创建资源成功后，需添加资源账户才能登录资源运维。资源账户的登录方式可选择以下几种：
 - 自动登录：系统纳管资源账户名和密码，再次登录资源时无需输入账户名和密码。
 - 手动登录：自动生成一个Empty账户。登录资源时，需手工输入资源账户名和密码进行登录。
 - 提权登录：当用户登录提权资源账户时，自动切换身份到特权帐号。

步骤7 单击“下一步”，关联运维用户和主机资源，单击“确认”完成配置。

步骤8 访问控制策略配置成功后，使用运维用户登录云堡垒机系统，选择“运维 > 主机运维”。

步骤9 选择目标主机资源，单击“登录”，[登录主机资源](#)进行运维操作。

📖 说明

- SSH、TELNET和Rlogin协议类型主机资源，可使用SSH客户端运维资源。
- FTP、SFTP和SCP协议类型主机资源，需使用FTP/SFTP/SCP客户端运维资源。
- MySQL、SQL Server、Oracle和DB2协议类型的主机资源，需提前配置SSO单点登录工具和数据库管理工具。通过SSO单点登录工具调用数据库客户端，实现SSO单点客户端运维资源。
- SSH、RDP、VNC和TELNET协议类型主机资源，支持直接通过Web浏览器运维；应用发布资源仅能通过Web浏览器远程访问进行运维。

----结束

7.5 管理项目和企业项目

创建项目并授权

- **创建项目**

进入管理控制台页面，单击右上方的用户名，在下拉列表中选择“统一身份认证”，进入统一身份认证服务页面。选择左侧导航中的“项目”，单击“创建项目”，选择区域并输入项目名称。

- **授权**

通过为用户组授予权限（包括资源集和操作集），实现项目和用户组的关联。将用户加入到用户组，使用户具有用户组中的权限，从而精确地控制用户所能访问的项目，以及所能操作的资源。具体步骤如下：

- a. 在“用户组”页面，选择目标用户组，单击操作列的“授权”，进入“选择策略”环节。
- b. 在列表中勾选策略，完成用户组的“策略”设置。
- c. 单击“下一步”，将“选择授权范围方案”设置为“指定区域项目资源”。
- d. 在区域项目列表中，选择具体的项目，单击“确定”。完成对用户组的授权。
- e. 在“用户”页面，选择目标用户，单击操作列的“授权”，进入“选择授权方式”环节。
- f. 勾选“继承所选用户组的策略”，并在“已有用户组”列表中勾选[步骤1](#)中进行了策略授权的用户组。
- g. 单击“确定”，将用户加入到用户组。

创建企业项目并授权

- **创建企业项目**

进入管理控制台页面，单击右上方的“企业 > 企业管理”，进入企业项目管理页面。单击“创建企业项目”，完成企业项目的创建。

📖 说明

开通了企业项目的客户，或者权限为企业主账号的客户才可以看到控制台页面上方的“企业”入口。如需使用该功能，请联系客服申请开通。

- **授权**

通过为企业项目添加用户组，并设置策略，实现企业项目和用户组的关联。将用户加入到用户组，使用户具有用户组中的权限，从而精确地控制用户所能访问的项目，以及所能操作的资源。具体步骤如下：

- a. 在企业项目管理页面，单击企业项目的名称，进入企业项目详情页面。
- b. 在“权限管理”页签，单击“用户组授权”，系统跳转至IAM的用户组页面，在“用户组”页签中为企业项目关联用户组并授权。

详细操作，请参见[创建用户组并授权](#)。

- **关联资源与企业项目**

企业项目可以将云资源按企业项目统一管理。

- 购买弹性云服务器时选择企业项目

在购买页面，“企业项目”下拉列表中选择目标企业项目，实现资源与企业项目关联。

- 资源迁入

对于账号下的存量弹性云服务器，您可以在“企业项目管理”页面将资源迁入目标企业项目。

“default”为默认企业项目，账号下原有资源和未选择企业项目的资源均在默认企业项目内。

更多信息，请参阅《[企业管理用户指南](#)》。

7.6 敏感操作保护

操作场景

弹性云服务器支持敏感操作保护，在控制台进行敏感操作时，需要输入一种能证明身份的凭证，身份验证通过后方可进行相关操作。为了账号安全，建议开启操作保护功能，该功能对账号以及账号下的用户都生效。

目前支持：关机、重启、删除、重置密码、卸载磁盘、解绑弹性公网IP。

开启操作保护

操作保护默认关闭，您可以参考以下步骤开启操作保护。

1. 登录管理控制台。
2. 在“控制台”页面，鼠标移动至右上方的用户名，在下拉列表中选择“安全设置”。

图 7-18 安全设置



3. 在“安全设置”页面中，选择“敏感操作 > 操作保护 > 立即启用”。

图 7-19 敏感操作



4. 在“操作保护设置”页面中，选择“开启”，单击“确定”后，开启操作保护。开启后，您以及账号中的IAM用户进行敏感操作时，例如删除弹性云服务器资源，需要输入验证码进行验证，避免误操作带来的风险和损失。

说明

- 用户如果进行敏感操作，将进入“操作保护”页面，选择认证方式，包括邮箱、手机和虚拟MFA三种认证方式。
 - 如果用户只绑定了手机，则认证方式只能选择手机。
 - 如果用户只绑定了邮箱，则认证方式只能选择邮件。
 - 如果用户未绑定邮箱、手机和虚拟MFA，进行敏感操作时，华为云将提示用户绑定邮箱、手机或虚拟MFA。
- 如需修改验证手机、邮箱、虚拟MFA设备，请在[基本信息](#)中修改。

操作保护验证

当您已经开启操作保护，在进行敏感操作时，系统会先进行操作保护验证：

- 若您绑定了邮箱，需输入邮箱验证码。
- 若您绑定了手机，需输入手机验证码。
- 若您绑定了虚拟MFA，需输入MFA设备上的6位动态验证码。

当尝试关机弹性云服务器时，弹出以下验证框，选择一种验证方式：

图 7-20 操作保护身份验证



关闭操作保护

如需关闭操作保护，请参考以下步骤操作。

1. 登录管理控制台。
2. 在“控制台”页面，鼠标移动至右上方的用户名，在下拉列表中选择“安全设置”。

图 7-21 安全设置



3. 在“安全设置”页面中，选择“敏感操作 > 操作保护 > 立即修改”。

图 7-22 修改敏感操作



4. 在“操作保护设置”页面中，选择“关闭”，单击“确定”后，关闭操作保护。

相关链接

- [如何绑定虚拟MFA设备？](#)
- [如何获取MFA验证码？](#)

8 备份管理

8.1 云备份概述

什么是云备份

云备份（Cloud Backup and Recovery, CBR）可以为云服务器、云硬盘提供简单易用的备份服务，当发生病毒入侵、人为误删除、软硬件故障等事件时，可将数据恢复到任意备份点。

云备份保障用户数据的安全性和正确性，确保业务安全。

备份、快照、镜像有什么区别？

备份分为云服务器备份和云硬盘备份。

镜像分为系统盘镜像、数据盘镜像、整机镜像。

备份类型	备份对象	适用场景	区别和优势	备份方法	恢复方法
云服务器备份	弹性云服务器中的所有云硬盘（系统盘和数据盘）	<ul style="list-style-type: none"> ● 云服务器受到攻击或病毒入侵 通过云服务器备份，可立即恢复到最近一次没有受黑客攻击或病毒入侵的备份时间点。 ● 数据被误删 通过云服务器备份，可立即恢复到删除前的备份时间点，找回被删除的数据。 ● 应用程序更新出错 通过云服务器备份，可立即恢复到应用程序更新前的备份时间点，使系统正常运行。 ● 云服务器宕机 通过云服务器备份，可立即恢复到宕机之前的备份时间点，使云服务器能再次正常启动。 	<p>备份的同一个云服务器下的所有云硬盘数据具有一致性，即同时对所有云硬盘进行备份，不存在因备份创建时间差带来的数据不一致问题。</p> <p>且云备份支持根据备份策略自动备份。</p>	创建云服务器备份	<ul style="list-style-type: none"> ● 使用云服务器备份恢复数据（恢复至原服务器） ● 使用云服务器备份恢复数据（创建新的云服务器）
云硬盘备份	指定的单个或多个云硬盘（系统盘或数据盘）	<ul style="list-style-type: none"> ● 系统盘没有个人数据，因而只需要对部分的数据盘进行备份 当云硬盘出现故障或云硬盘中的数据发生逻辑错误时（如误删数据、遭遇黑客攻击或病毒危害等），可快速恢复数据。 ● 备份作为基线数据 设置备份策略，根据策略自动对云硬盘进行数据备份，通过定期创建的备份作为基线数据，用来创建新的云硬盘或者恢复数据到云硬盘。 	<p>备份数据则存储在对象存储(OBS)中，可以实现在云硬盘存储损坏情况下的数据恢复</p> <p>保证数据安全的同时降低备份成本</p>	创建云硬盘备份	<ul style="list-style-type: none"> ● 使用云硬盘备份恢复数据（恢复至原磁盘） ● 使用云硬盘备份恢复数据（创建新的磁盘）

备份类型	备份对象	适用场景	区别和优势	备份方法	恢复方法
快照	指定的单个或多个云硬盘（系统盘或数据盘）	<ul style="list-style-type: none"> ● 日常备份数据 通过对云硬盘定期创建快照，实现数据的日常备份，可以应对由于误操作、病毒以及黑客攻击等导致数据丢失或不一致的情况。 ● 快速恢复数据 应用软件升级或业务数据迁移等重大操作前，您可以创建一份或多份快照，一旦升级或迁移过程中出现问题，可以通过快照及时将业务恢复到快照创建点的数据状态。 例如，当由于云服务器A的系统盘A发生故障而无法正常开机时，此时您可以使用系统盘A已有的快照新建一块云硬盘B并挂载至正常运行的云服务器B上，从而云服务器B能够通过云硬盘B读取原系统盘A的数据。 ● 快速部署多个业务 通过同一个快照可以快速创建出多个具有相同数据的云硬盘，从而可以同时为多种业务提供数据资源。例如数据挖掘、报表查询和开发测试等业务。 这种方式既保护了原始数据，又能通过快照创建的新云硬盘快速部署其他业务，满足企业对业务数据的多元化需求。 <p>说明</p> <ul style="list-style-type: none"> ● 只支持回滚快照数据至原云硬盘，不支持快照回滚到其他云硬盘。 ● 重装操作系统或切换操作系统后，系统盘快照会自动删除；数据盘快照不受影响，可以照常使用。 	<ul style="list-style-type: none"> ● 快照数据与云硬盘数据存储在一起，可以支持快速备份和恢复 ● 快速保存指定时刻云硬盘的数据，同时还可以通过快照创建新的云硬盘，这样云硬盘在初始状态就具有快照中的数据 	创建快照	使用快照回滚数据

备份类型	备份对象	适用场景	区别和优势	备份方法	恢复方法
系统盘镜像	系统盘	<ul style="list-style-type: none"> ● 快速恢复系统 更换操作系统、应用软件升级或业务数据迁移等重大操作前，将系统盘创建成系统盘镜像，一旦迁移过程中出现问题，可以通过系统盘镜像切换操作系统，或重新创建新的云服务器。 ● 快速部署多个业务 通过同一个系统盘镜像可以快速创建出多个具有相同操作系统的云服务器，从而快速部署多个业务。 	系统盘镜像可以实现在云服务器操作系统损坏情况下快速切换至损坏前的操作系统。	创建系统盘镜像	<ul style="list-style-type: none"> ● 使用系统盘镜像切换故障云服务器的操作系统 ● 使用系统盘镜像创建新的云服务器
数据盘镜像	指定的数据盘	<p>快速复制数据</p> <p>通过同一个数据盘镜像可以快速创建出多个具有相同数据的云硬盘，再将新创建的云硬盘挂载到其他服务器上，从而可以同时为多种业务提供数据资源。</p>	数据盘镜像可以复制全盘的数据并创建新的云硬盘挂载到其他服务器上，实现云服务器数据盘的复制和共享。	创建数据盘镜像	使用数据盘镜像创建数据盘
整机镜像	弹性云服务器中的所有云硬盘（系统盘和数据盘）	<ul style="list-style-type: none"> ● 快速恢复系统 更换操作系统、应用软件升级或业务数据迁移等重大操作前，将云服务器的系统盘和数据盘创建成整机镜像，一旦迁移过程中出现问题，可以通过整机镜像切换操作系统，或重新创建新的云服务器。 ● 快速部署多个业务 通过同一个整机镜像可以快速创建出多个具有相同操作系统和数据的云服务器，从而快速部署多个业务。 	通过整机镜像实现业务的整迁移。	创建整机镜像	使用整机镜像创建新的云服务器

云备份产品架构

云备份由备份、存储库和策略组成。

- **备份：**

备份即一个备份对象执行一次备份任务产生的备份数据，包括备份对象恢复所需要的全部数据。

- 云服务器备份：云服务器备份提供对弹性云服务器和裸金属服务器的基于多云硬盘一致性快照技术的数据保护。同时，未部署数据库等应用的服务器产生的备份为服务器备份，部署数据库等应用的服务器产生的备份为数据库服务器备份。
- 云硬盘备份：云硬盘备份提供对云硬盘的基于快照技术的数据保护。

- **存储库**

云备份使用存储库来存放备份。创建备份前，需要先创建至少一个存储库，并将服务器或磁盘绑定至存储库。服务器或磁盘产生的备份则会存放至绑定的存储库中。

存储库分为备份存储库和复制存储库两种。备份存储库用于存放备份对象产生的备份，复制存储库用于存放复制操作产生的备份。

不同类型的备份对象产生的备份需要存放在不同类型的存储库中。

- **策略**

策略分为备份策略和复制策略。

- 备份策略：需要对备份对象执行自动备份操作时，可以设置备份策略。通过在策略中设置备份任务执行的时间、周期以及备份数据的保留规则，将备份存储库绑定到备份策略，可以为存储库执行自动备份。
- 复制策略：需要对备份或存储库执行自动复制操作时，可以设置复制策略。通过在策略中设置复制任务执行的时间、周期以及备份数据的保留规则，将备份存储库绑定到复制策略，可以为存储库执行自动复制。复制产生的备份需要存放在复制存储库中。

备份机制

首次备份为全量备份，备份云服务器/磁盘已使用空间。

示例：某磁盘大小为100GB，已使用空间为40GB，则备份的是40GB的已使用空间。

后续备份均为增量备份，备份上次备份后变化的数据，缩短备份时长、节约备份空间。

删除备份时，仅删除不被其他备份依赖的数据块，不影响使用其他备份进行恢复。无论是全量还是增量备份，都可以快速、方便地将数据恢复至备份所在时刻的状态。

云备份会在备份过程中自动创建快照并且为每个磁盘保留最新的快照。如果该磁盘已备份，再次备份后会自动将旧快照删除，保留最新的快照。

云备份通过云服务器/磁盘与对象存储服务的结合，将数据备份到对象存储中，高度保障用户的备份数据安全。

备份的方式及适用场景

云备份提供两种配置方式，一次性备份和周期性备份。一次性备份是指用户手动创建的一次性备份任务。周期性备份是指用户通过创建备份策略并绑定存储库的方式创建的周期性备份任务。

表 8-1 备份的方式及适用场景

对比项	一次性备份	周期性备份
备份策略	不需要	需要
备份次数	手动执行一次性备份	根据备份策略进行周期性备份
备份名称	支持自定义，默认为“manualbk_xxxx”	系统自动生成，默认为“autobk_xxxx”
备份方式	默认首次全量备份，后续增量备份	默认首次全量备份，后续增量备份
适用场景	资源进行操作系统补丁安装、升级，应用升级等操作之前，以便安装或者升级失败之后，能够快速恢复到变更之前的状态。	资源的日常备份保护，以便发生不可预见的故障而造成数据丢失时，能够使用邻近的备份进行恢复。

8.2 备份弹性云服务器

操作场景



云备份提供申请即用的备份服务，使您的数据更加安全可靠。例如，当云服务器或磁盘出现故障或者人为错误导致数据误删时，可以自助快速恢复数据。本节操作介绍备份云服务器和云硬盘的操作步骤。

您可以通过[云备份产品架构](#)、[备份机制（全量备份和增量备份）](#)、[备份的方式及适用场景](#)了解更多产品详情。

目前弹性云服务器备份可以通过“云服务器备份”和“云硬盘备份”功能实现：



- 云服务器备份（推荐）：如果是对弹性云服务器中的所有云硬盘（系统盘和数据盘）进行备份，推荐使用云服务器备份功能，同时对所有云硬盘进行备份，避免因备份创建时间差带来的数据不一致问题。
- 云硬盘备份：如果对指定的单个或多个云硬盘（系统盘或数据盘）进行备份，推荐使用云硬盘备份功能，在保证数据安全的同时降低备份成本。

云服务器备份操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器的“操作”列下，单击“更多 > 磁盘/备份 > 创建云服务器备份”。
 - 如果云服务器已绑定存储库，则根据界面提示，配置备份信息。
 - 云服务器列表：默认已勾选待备份的云服务器
 - 名称：用户自定义云备份名称。

- 描述：输入待创建的备份的描述。
 - 执行全量备份：勾选后，系统会为绑定的云服务器执行全量备份，备份所占存储容量也会相应增加。
 - 如果云服务器未绑定存储库，请先购买存储库，配置备份信息。
详细操作请参考[购买云服务器备份存储库](#)。
5. 单击“确定”。系统会自动为云服务器创建备份。
- 在“备份副本”页签，产生的备份的“备份状态”为“可用”时，表示备份任务执行成功。
- 云服务器备份在执行备份过程中，若备份进度已经超过10%可重启云服务器。但为了保证数据完整性，建议备份完成后再执行重启。
- 执行备份成功后，后续可以使用云服务器备份恢复服务器数据或创建镜像，详情请参见[使用云服务器备份恢复数据](#)和[使用备份创建镜像](#)。

云硬盘备份操作步骤

1. 登录管理控制台。
 2. 单击管理控制台左上角的，选择区域和项目。
 3. 单击“≡”，选择“计算 > 弹性云服务器”。
 4. 在弹性云服务器的“操作”列下，单击“更多 > 磁盘/备份 > 创建云硬盘备份”。
 - 如果云服务器已绑定存储库，则根据界面提示，配置备份信息。
 - 服务器列表：默认已勾选待备份的云服务器，单击查看云服务器挂载的磁盘。并选择待备份的磁盘。
 - 名称：用户自定义云备份名称。
 - 描述：输入待创建的备份的描述。
 - 执行全量备份：勾选后，系统会为绑定的磁盘执行全量备份，备份所占存储容量也会相应增加。
 - 如果云服务器未绑定存储库，请先购买存储库，再配置备份信息。
详细操作请参考[购买云硬盘备份存储库](#)。
5. 单击“确定”。系统会自动为磁盘创建备份。
- 根据界面提示，跳转至云备份控制台的“备份副本”页签，当产生的备份的“备份状态”为“可用”时，表示备份任务执行成功。
- 如果在备份过程中对磁盘的数据进行删除等操作，被删除的文件可能不会被备份成功。为了保证数据完整性，建议备份完成后再对数据进行操作。
- 执行备份成功后，后续可以使用云硬盘备份恢复磁盘数据，详情请参见[使用云硬盘备份恢复数据](#)。

后续操作

备份成功后，您可以查看云服务器的云备份信息以及备份详情。

1. 登录管理控制台。


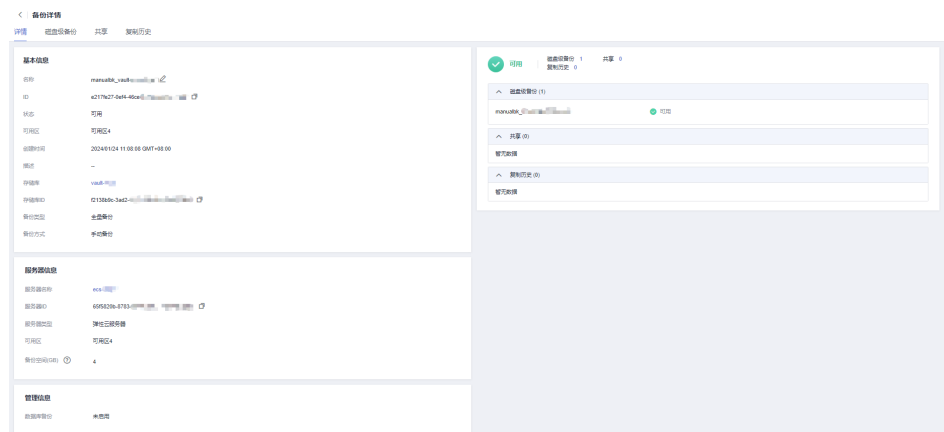
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在云服务器列表中，单击待查看的弹性云服务器名称，进入详情页。
5. 查看备份详情。
 - a. 在云服务器详情页“基本信息”页签右下角的“云备份”区域，可以看到具体的备份及时间信息。

图 8-1 云备份区域



- b. 单击具体的备份，跳转至云备份控制台的“备份详情”页面。

图 8-2 备份详情



通过备份详情，可以查看备份的名称、ID、状态、可用区、创建时间、存储库等信息，还可以查看备份相关的云服务器信息。

更多关于备份详情的详细操作，请参见[备份管理](#)。

6. 在云服务器详情页，单击“云备份”页签，查看云备份的基本信息和付费信息。

9 密码和密钥对管理

9.1 重置 ECS 密码

9.1.1 密码使用场景介绍

弹性云服务器系统密码涉及到客户重要的私人信息，提醒您妥善保管密码。如果您忘记密码或密码过期，可以重置密码。

弹性云服务器的重置密码场景如表9-1所示。

表 9-1 重置密码场景

操作场景	前提条件
在控制台重置ECS密码	已安装重置密码插件。 说明 <ul style="list-style-type: none">使用公共镜像的弹性云服务器默认已安装一键式重置密码插件。Windows/Linux操作系统的云服务器
未安装重置密码插件且忘记密码时，如何重置密码(Windows)?	未安装重置密码插件。
未安装重置密码插件且忘记密码时，如何重置密码(Linux)?	未安装重置密码插件。

背景知识

弹性云服务器的密码规则如表9-2所示。

表 9-2 密码设置规则

参数	规则
密码	<ul style="list-style-type: none">● 密码长度范围为8到26位。● 密码至少包含以下4种字符中的3种：<ul style="list-style-type: none">- 大写字母- 小写字母- 数字- Windows操作系统云服务器支持的特殊字符：!@\$%^-_=+[{()}];,./?~#*- Linux操作系统云服务器支持的特殊字符：!@\$%^-_=+[{}];,./?~#*● 密码不能包含用户名或用户名的逆序。● Windows操作系统的云服务器，不能包含用户名中超过两个连续字符的部分。

9.1.2 在控制台重置 ECS 密码

操作场景

如果在创建弹性云服务器时未设置密码，或密码丢失、过期，可以参考本节操作重置密码。

前提条件

- 密码丢失或过期前，已安装密码重置插件。
 - 公共镜像创建的弹性云服务器默认已安装一键式重置密码插件。
 - 私有镜像创建的云服务器且未安装密码重置插件，可参考[未安装重置密码插件且忘记密码时，如何重置密码\(Windows\)?](#)和[未安装重置密码插件且忘记密码时，如何重置密码\(Linux\)?](#)。
- 请勿删除重置密码进程CloudResetPwdAgent和CloudResetPwdUpdateAgent，否则，会导致一键式重置密码功能不可用。
- 使用SUSE 11 SP4镜像创建的弹性云服务器，内存需要大于等于4GiB时才能支持一键式重置密码功能。
- 弹性云服务器使用的VPC网络DHCP不能禁用。
- 弹性云服务器网络正常通行。
- 请确保一键式重置密码插件未被安全软件阻止运行，否则重置密码功能无法使用。
- 重置密码之后，必须重启弹性云服务器方可生效。

操作步骤

您可参考以下步骤在控制台上修改一台或多台弹性云服务器的登录密码。

📖 说明

开机状态下重置密码后需重启云服务器，新密码才会生效。您可以重置密码后手动重启，或在重置密码时候勾选“自动重启”。

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 弹性云服务器”。
3. 选中待重置密码的弹性云服务器，并选择“操作”列下的“更多 > 重置密码”。

图 9-1 重置密码



4. 根据界面提示，设置弹性服务器的新密码，并确认新密码。

📖 说明

如果提示无法进行该操作，请根据提示信息，参考[未安装重置密码插件且忘记密码时，如何重置密码\(Windows\)?](#)和[未安装重置密码插件且忘记密码时，如何重置密码\(Linux\)?](#)完成重置密码操作。

新密码的复杂度应满足[表9-3](#)。

表 9-3 密码设置规则

参数	规则
密码	<ul style="list-style-type: none">• 密码长度范围为8到26位。• 密码至少包含以下4种字符中的3种：<ul style="list-style-type: none">- 大写字母- 小写字母- 数字- Windows操作系统云服务器支持的特殊字符：!@\$%^_-=+[{()}];,./?~#*- Linux操作系统云服务器支持的特殊字符：!@\$%^_-=+[{()}];,./?~#*• 密码不能包含用户名或用户名的逆序。• Windows操作系统的云服务器，不能包含用户名中超过两个连续字符的部分。

5. 单击“确认”。

系统执行重置密码操作，该操作预计需要10分钟，请勿频繁执行。

- 如果在开机状态下重置密码，需手动重启使新密码生效。
- 如果在关机状态下重置密码，待重新开机后新密码生效。

9.2 管理 ECS 一键式重置密码插件

9.2.1 获取 ECS 一键式重置密码插件

操作场景

若弹性云服务器“重置密码”失败，有可能是由于未安装一键式重置密码插件或插件失效，您可以参考本章节获取一键式重置密码插件并校验其完整性。

获取并校验一键式重置密码插件完整性（Linux）

1. 以root用户登录弹性云服务器。
2. 执行以下命令，下载一键式重置密码插件及sha256校验码。

您可以根据弹性云服务器所在区域、操作系统是32位还是64位，参考[表9-4](#)选择对应的下载地址。

📖 说明

为了便于您通过内网下载一键式重置密码插件，我们提供了不同区域的下载地址，不同区域、相同操作系统的插件相同。

如果[表9-4](#)中不包含您的弹性云服务器所在区域，请您为弹性云服务器绑定EIP，然后选择就近区域进行下载。

绑定EIP的详细操作，请参见[绑定弹性公网IP](#)。

以“华北-北京一”区域的32位x86操作系统为例：

```
wget https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip
```

```
wget https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
```

3. 执行以下命令，获取本地一键式重置密码插件的Hash值。

```
sha256sum {软件包本地目录}/CloudResetPwdAgent.zip
```

{软件包本地目录}：请根据实际下载目录进行替换，例如/root。

4. 将步骤2获取的sha256 Hash值和步骤3获取的Hash值进行比较。
 - 一致：通过校验。
 - 不一致：请重新下载对应版本的一键式重置密码插件，重复步骤2~步骤4进行校验。

获取并校验一键式重置密码插件完整性（Windows）

1. 登录弹性云服务器。
2. 下载一键式重置密码插件及sha256校验码。
您可以根据弹性云服务器所在区域，参考表9-4选择对应的下载地址。

说明

如果表9-4中不包含您的弹性云服务器所在区域，请您为弹性云服务器绑定EIP，然后选择就近区域进行下载。

绑定EIP的详细操作，请参见[绑定弹性公网IP](#)。

3. 以管理员权限打开cmd，执行以下命令，获取本地一键式重置密码插件的Hash值。
certutil -hashfile {软件包本地目录}\CloudResetPwdAgent.zip SHA256
{软件包本地目录}：请根据实际下载目录进行替换。
4. 将步骤2获取的sha256 Hash值和步骤3获取的Hash值进行比较。
 - 一致：通过校验。
 - 不一致：请重新下载对应版本的一键式重置密码插件，重复步骤2~步骤4进行校验。

一键式重置密码插件及 sha256 校验码下载地址

表 9-4 一键式重置密码插件下载地址

区域	操作系统	文件名	下载地址
华北-北京一	Linux(x86_32)	CloudResetPwdAgent.zip	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip.sha256

区域	操作系统	文件名	下载地址
	Linux(x86_64)	CloudResetPwdAgent.zip	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(arm64)	CloudResetPwdAgent.zip	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Windows	CloudResetPwdAgent.zip	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
华北-北京四	Linux(x86_64)	CloudResetPwdAgent.zip	https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Windows	CloudResetPwdAgent.zip	https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
华东-上海二	Linux(x86_32)	CloudResetPwdAgent.zip	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip.sha256

区域	操作系统	文件名	下载地址
	Linux(x86_64)	CloudResetPwdAgent.zip	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(armv7h)	CloudResetPwdAgent.zip	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Windows	CloudResetPwdAgent.zip	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
华南-广州	Linux(x86_32)	CloudResetPwdAgent.zip	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(x86_64)	CloudResetPwdAgent.zip	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(armv7h)	CloudResetPwdAgent.zip	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256

区域	操作系统	文件名	下载地址
	Windows	CloudResetPwdAgent.zip	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
中国-香港	Linux(x86_32)	CloudResetPwdAgent.zip	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(x86_64)	CloudResetPwdAgent.zip	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(arm64)	CloudResetPwdAgent.zip	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Windows	CloudResetPwdAgent.zip	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip

区域	操作系统	文件名	下载地址
		sha256校验码	https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
亚太-曼谷	Linux(x86_32)	CloudResetPwdAgent.zip	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/linux/32/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(x86_64)	CloudResetPwdAgent.zip	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Linux(arm64)	CloudResetPwdAgent.zip	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/arm/linux/64/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
	Windows	CloudResetPwdAgent.zip	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip
		sha256校验码	https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip.sha256

相关操作

获取完成后，您可以安装或更新一键式重置密码插件：

- [安装单台ECS的一键式重置密码插件（单台操作）](#)
- [更新单台ECS的一键式重置密码插件（单台操作）](#)

9.2.2 安装单台 ECS 的一键式重置密码插件（单台操作）

弹性云服务器提供一键式重置密码功能。当弹性云服务器的密码丢失或过期时，如果提前安装了一键式重置密码插件，则可以应用一键式重置密码功能，给弹性云服务器设置新密码。

该方法方便、有效，建议您成功创建、登录弹性云服务器后，安装密码重置插件，具体操作请参见本节内容。

说明

使用公共镜像创建的弹性云服务器默认已安装一键式重置密码插件。检查是否已安装一键式重置密码插件，请参见[步骤1](#)。

安装须知

1. 用户自行决定是否安装一键式重置密码插件，使弹性云服务器具备一键式重置密码功能。
2. 安装完成后，请勿自行卸载插件，否则可能导致管理控制台判断失误，从而无法完成密码重置。
3. 重装/切换弹性云服务器操作系统后，一键式重置密码功能失效。如需继续使用该功能，请重新安装一键式重置密码插件。
4. 更换弹性云服务器系统盘后，一键式重置密码功能失效。如需继续使用该功能，请重新安装一键式重置密码插件。
5. CoreOS和KylinOS操作系统的弹性云服务器，不支持安装一键式重置密码插件。
6. 如果没有提前安装一键式重置密码插件，密码丢失或过期后，不能使用一键式重置密码功能。
7. 弹性云服务器需绑定弹性公网IP，才能安装一键式重置密码插件。

前提条件

- 对于Windows弹性云服务器，需保证C盘可写入，且剩余空间大于300MB。
- 对于Linux弹性云服务器，需保证根目录可写入，且剩余空间大于300MB。
- 对于Linux弹性云服务器，若开启了selinux，请参考[关闭selinux](#)禁用selinux。
- 使用SUSE 11 SP4镜像创建的弹性云服务器，内存需要大于等于4GiB时才能支持一键式重置密码功能。
- 弹性云服务器使用的VPC网络DHCP不能禁用。
- 弹性云服务器网络正常通行。
- 设置弹性云服务器的网卡属性为DHCP，使弹性云服务器可以动态获取IP地址。

说明

Linux弹性云服务器的设置方法，请参见[设置网卡属性为DHCP（Linux）](#)。

Windows弹性云服务器的设置方法，请参见[设置网卡属性为DHCP（Windows）](#)。

- 弹性云服务器安全组出方向规则满足如下要求：
 - 协议端口：TCP
 - 端口范围：80
 - 目的地址：169.254.0.0/16

如果您使用的是默认安全组出方向规则，则已经包括了如上要求，可以正常初始化。默认安全组出方向规则为：

- 协议端口：全部
- 端口范围：全部协议端口
- 目的地址：0.0.0.0/0

Linux 弹性云服务器插件安装方法

步骤1 检查弹性云服务器是否已安装一键式重置密码插件。

1. 以root用户登录弹性云服务器。
2. 执行以下命令，查询是否已安装CloudResetPwdAgent。

```
ls -lh /Cloud*
```

图 9-2 查询是否已安装一键式重置密码插件

```
[root@ecs-test ~]# ls -lh /Cloud*
total 20K
drwx----- 2 root root 4.0K Jun 13 14:13 bin
drwxr-xr-x 2 root root 4.0K Jun 13 11:53 conf
drwx----- 3 root root 4.0K Jun 13 11:53 depend
drwx----- 2 root root 4.0K Jun 13 11:53 lib
drwx----- 2 root root 4.0K Jun 13 14:13 logs
[root@ecs-test ~]#
[root@ecs-test ~]#
```

检查结果是否如图9-2所示。

- 是，表示已安装一键式重置密码插件，结束。
- 否，表示未安装一键式重置密码插件，请继续执行如下操作进行安装。

步骤2 请参考[获取并校验一键式重置密码插件完整性（Linux）](#)，下载对应的一键式重置密码插件CloudResetPwdAgent.zip并完成完整性校验。

安装一键式重置密码插件对插件的具体放置目录无特殊要求，请您自定义。

步骤3 执行以下命令，解压软件包CloudResetPwdAgent.zip。

安装一键式重置密码插件对插件的解压目录无特殊要求，请您自定义。

```
unzip -o -d 插件解压目录 CloudResetPwdAgent.zip
```

示例：

假设插件解压的目录为/home/linux/test，则命令行如下：

```
unzip -o -d /home/linux/test CloudResetPwdAgent.zip
```

步骤4 安装一键式重置密码插件。

1. 执行以下命令，进入文件CloudResetPwdAgent.Linux。

```
cd {插件解压目录}/CloudResetPwdAgent/CloudResetPwdAgent.Linux
```

示例：

假设插件解压的目录为/home/linux/test，则命令行如下：

```
cd /home/linux/test/CloudResetPwdAgent/CloudResetPwdAgent.Linux
```

2. 执行以下命令，添加文件setup.sh的运行权限。

```
chmod +x setup.sh
```

3. 执行以下命令，安装插件。

```
sudo sh setup.sh
```

如果脚本执行成功打印“cloudResetPwdAgent install successfully.”，且未打印“Failed to start service cloudResetPwdAgent”，表示安装成功。

说明

- 您也可以根据[步骤1](#)，检查密码重置插件是否安装成功。
- 如果密码重置插件安装失败，请检查安装环境是否符合要求，并重试安装操作。

- 步骤5** 修改重置密码插件的文件权限。

```
chmod 700 /CloudrResetPwdAgent/bin/cloudResetPwdAgent.script
```

```
chmod 700 /CloudrResetPwdAgent/bin/wrapper
```

```
chmod 600 /CloudrResetPwdAgent/lib/*
```

----结束

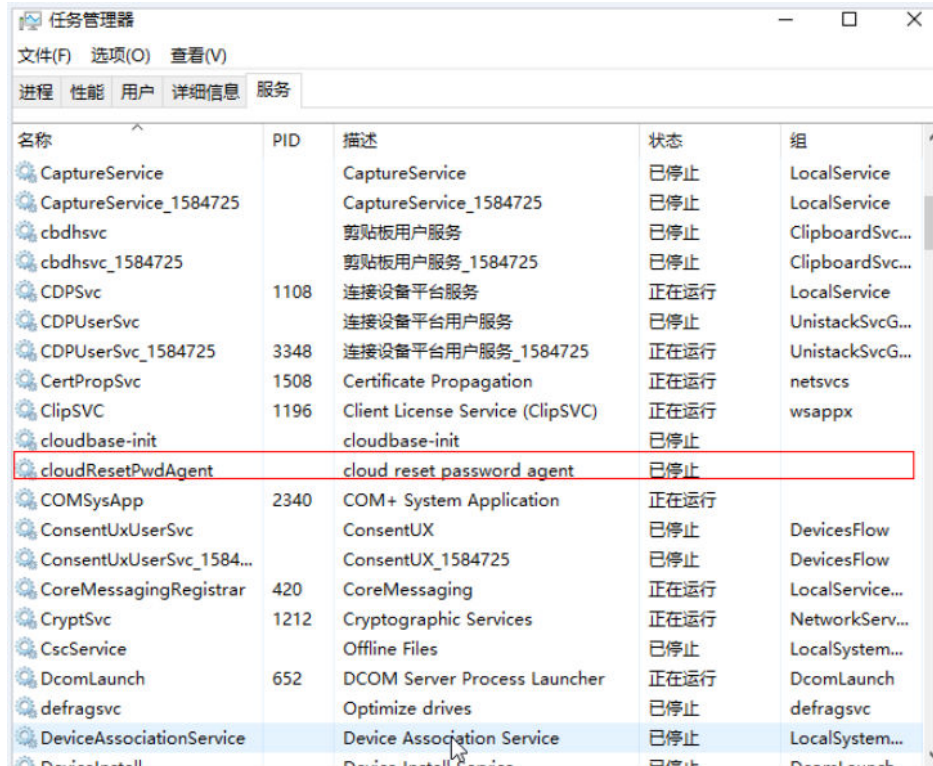
Windows 弹性云服务器插件安装方法

- 步骤1** 登录弹性云服务器。

- 步骤2** 检查弹性云服务器是否已安装密码重置插件CloudResetPwdAgent。检查方法如下：

查看任务管理器，如果找到cloudResetPwdAgent服务，如[图9-3](#)所示，表示弹性云服务器已安装密码重置插件。

图 9-3 安装插件成功



- 是，结束。
- 否，执行步骤3。

步骤3 请参考[获取并校验一键式重置密码插件完整性（Windows）](#)，下载对应的一键式重置密码插件CloudResetPwdAgent.zip并完成完整性校验。

安装一键式重置密码插件对插件的具体放置目录无特殊要求，请您自定义。

步骤4 解压软件包CloudResetPwdAgent.zip。

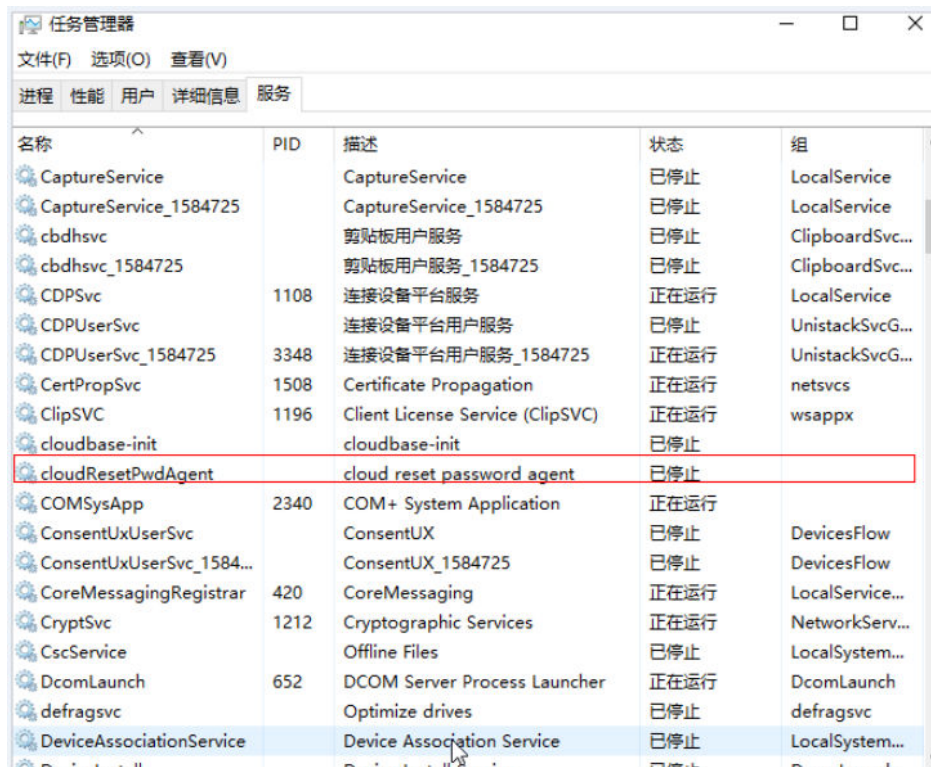
安装一键式重置密码插件对插件的解压目录无特殊要求，请您自定义。

步骤5 安装一键式重置密码插件。

1. 双击“CloudResetPwdAgent.Windows”文件夹下的“setup.bat”文件。
安装密码重置插件。
2. 查看任务管理器，检查密码重置插件是否安装成功。

如果在任务管理器中查找到了cloudResetPwdAgent服务，如图9-4所示，表示安装成功，否则安装失败。

图 9-4 安装插件成功



说明

如果密码重置插件安装失败，请检查安装环境是否符合要求，并重试安装操作。

---结束

后续处理

- 安装一键式重置密码插件后，如果无法通过弹性云服务器开机自动启动该插件，可以将其添加至开机启动项。具体方法请参见[一键式密码重置插件启动失败时如何操作？](#)。
- 一键式重置密码插件安装成功后，请勿删除重置密码进程CloudResetPwdAgent，否则，会导致一键式重置密码功能不可用。
- 云平台对一键式重置密码插件进行了升级，对于新创建的弹性云服务器，默认采用PIPE模式，不会占用端口。对于已创建云服务器，仍采用AUTO模式，随机占用31000~32999中的一个端口。占用端口的原则是：在该范围内，系统会按照自小到大的顺序，占用当前空闲的端口。

相关操作

- **卸载插件：**如果不再继续使用一键重置密码功能，您可以根据如下指导卸载插件。
 - Linux弹性云服务器
 - i. 登录弹性云服务器。
 - ii. 执行以下命令，进入bin目录，删除服务cloudResetPwdAgent。
cd /CloudrResetPwdAgent/bin

- ii. (可选) 停止**CloudResetPwdUpdateAgent**。
不同Windows版本的弹性云服务器插件有差异, 请检查**CloudResetPwdUpdateAgent**是否存在, 如果存在, 请执行本步骤停止; 如果不存在, 则跳过本步骤。
 - 1) 登录弹性云服务器。
 - 2) 打开“任务管理器”。
 - 3) 在“服务”页签, 找到“**CloudResetPwdUpdateAgent**”服务, 右键单击菜单中的“停止”。
- **重启插件**: 如果一键重置密码功能异常, 您可以根据如下指导重启插件。
 - Linux弹性云服务器
 - i. 登录弹性云服务器。
 - ii. 执行以下命令, 进入bin目录, 重启服务**cloudResetPwdAgent**。

```
cd /CloudResetPwdAgent/bin
sudo ./cloudResetPwdAgent.script reboot
```
 - iii. 检查**CloudResetPwdUpdateAgent**是否存在, 如果存在, 进入bin目录, 重启服务**cloudResetPwdUpdateAgent**。

```
cd /CloudResetPwdUpdateAgent/bin
sudo ./cloudResetPwdUpdateAgent.script reboot
```
 - Windows弹性云服务器
 - i. 重启**CloudResetPwdAgent**。
 - 1) 登录弹性云服务器。
 - 2) 打开“任务管理器”。
 - 3) 在“服务”页签, 找到“**CloudResetPwdAgent**”服务, 右键单击菜单中的“重启”。
 - ii. (可选) 重启**CloudResetPwdUpdateAgent**。
不同Windows版本的弹性云服务器插件有差异, 请检查**CloudResetPwdUpdateAgent**是否存在, 如果存在, 请执行本步骤重启; 如果不存在, 则跳过本步骤。
 - 1) 登录弹性云服务器。
 - 2) 打开“任务管理器”。
 - 3) 在“服务”页签, 找到“**CloudResetPwdUpdateAgent**”服务, 右键单击菜单中的“重启”。

9.2.3 更新单台 ECS 的一键式重置密码插件（单台操作）

弹性云服务器提供一键式重置密码功能。当弹性云服务器的密码丢失或过期时, 如果提前安装了一键式重置密码插件, 则可以应用一键式重置密码功能, 给弹性云服务器设置新密码。

更新一键式重置密码插件的操作请参考本节内容。

安装须知

1. 弹性云服务器需绑定弹性公网IP, 才能更新一键式重置密码插件。
2. 使用公共镜像创建的弹性云服务器默认已安装一键式重置密码插件, 在更新前请先卸载一键式重置密码插件。

前提条件

- 对于Windows弹性云服务器，需保证C盘可写入，且剩余空间大于300MB。
- 对于Linux弹性云服务器，需保证根目录可写入，且剩余空间大于300MB。
- 对于Linux弹性云服务器，若开启了selinux，请参考[关闭selinux](#)禁用selinux。
- 使用SUSE 11 SP4镜像创建的弹性云服务器，内存需要大于等于4GiB时才能支持一键式重置密码功能。
- 弹性云服务器使用的VPC网络DHCP不能禁用。
- 弹性云服务器网络正常通行。
- 设置弹性云服务器的网卡属性为DHCP，使弹性云服务器可以动态获取IP地址。

📖 说明

Linux弹性云服务器的设置方法，请参见[设置网卡属性为DHCP（Linux）](#)。

Windows弹性云服务器的设置方法，请参见[设置网卡属性为DHCP（Windows）](#)。

- 弹性云服务器安全组出方向规则满足如下要求：
 - 协议端口：TCP
 - 端口范围：80
 - 目的地址：169.254.0.0/16

如果您使用的是默认安全组出方向规则，则已经包括了如上要求，可以正常初始化。默认安全组出方向规则为：

- 协议端口：全部
- 端口范围：全部协议端口
- 目的地址：0.0.0.0/0

Linux 弹性云服务器插件更新方法

步骤1 卸载插件。

1. 登录弹性云服务器。
2. 执行以下命令，进入bin目录，删除服务cloudResetPwdAgent。
cd /CloudrResetPwdAgent/bin
sudo ./cloudResetPwdAgent.script remove
3. 执行以下命令，删除插件。
sudo rm -rf /CloudrResetPwdAgent
4. 检查CloudResetPwdUpdateAgent是否存在，如果存在，进入bin目录，删除服务cloudResetPwdUpdateAgent。
cd /CloudResetPwdUpdateAgent/bin
sudo ./cloudResetPwdUpdateAgent.script stop
sudo ./cloudResetPwdUpdateAgent.script remove
5. 执行以下命令，删除插件。
sudo rm -rf /CloudResetPwdUpdateAgent

步骤2 请参考[获取并校验一键式重置密码插件完整性（Linux）](#)，下载对应的一键式重置密码插件CloudResetPwdAgent.zip并完成完整性校验。

安装一键式重置密码插件对插件的具体放置目录无特殊要求，请您自定义。

步骤3 执行以下命令，解压软件包CloudResetPwdAgent.zip。

安装一键式重置密码插件对插件的解压目录无特殊要求，请您自定义。

```
unzip -o -d 插件解压目录 CloudResetPwdAgent.zip
```

示例：

假设插件解压的目录为/home/linux/test，则命令行如下：

```
unzip -o -d /home/linux/test CloudResetPwdAgent.zip
```

步骤4 安装一键式重置密码插件。

1. 执行以下命令，进入文件CloudResetPwdAgent.Linux。

```
cd {插件解压目录}/CloudResetPwdAgent/CloudResetPwdAgent.Linux
```

示例：

假设插件解压的目录为/home/linux/test，则命令行如下：

```
cd /home/linux/test/CloudResetPwdAgent/CloudResetPwdAgent.Linux
```

2. 执行以下命令，添加文件setup.sh的运行权限。

```
chmod +x setup.sh
```

3. 执行以下命令，安装插件。

```
sudo sh setup.sh
```

如果脚本执行成功打印“cloudResetPwdAgent install successfully.”，且未打印“Failed to start service cloudResetPwdAgent”，表示安装成功。

说明

- 您也可以根据**步骤1**，检查密码重置插件是否安装成功。
- 如果密码重置插件安装失败，请检查安装环境是否符合要求，并重试安装操作。

步骤5 修改重置密码插件的文件权限。

```
chmod 700 /CloudrResetPwdAgent/bin/cloudResetPwdAgent.script
```

```
chmod 700 /CloudrResetPwdAgent/bin/wrapper
```

```
chmod 600 /CloudrResetPwdAgent/lib/*
```

----结束

Windows 弹性云服务器插件更新方法

步骤1 卸载插件。

1. 卸载并删除CloudResetPwdAgent。

- 进入C:\CloudResetPwdAgent\bin文件夹。
- 双击“UninstallApp-NT.bat”。
- 删除C:\CloudResetPwdAgent的文件。

2. （可选）卸载并删除CloudResetPwdUpdateAgent。

不同Windows版本的弹性云服务器插件有差异，请检查

CloudResetPwdUpdateAgent是否存在，如果存在，请执行本步骤卸载并删除；如果不存在，则跳过本步骤。

- a. 进入C:\CloudResetPwdUpdateAgent文件夹。
- b. 双击“UninstallApp-NT.bat”。
- c. 删除C:\CloudResetPwdUpdateAgent的文件。

如果删除失败，请先通过“任务管理器”删除CloudResetPwdUpdateAgent进程后，再次删除C:\CloudResetPwdUpdateAgent文件夹。

步骤2 请参考[获取并校验一键式重置密码插件完整性（Windows）](#)，下载对应的一键式重置密码插件CloudResetPwdAgent.zip并完成完整性校验。

安装一键式重置密码插件对插件的具体放置目录无特殊要求，请您自定义。

步骤3 解压软件包CloudResetPwdAgent.zip。

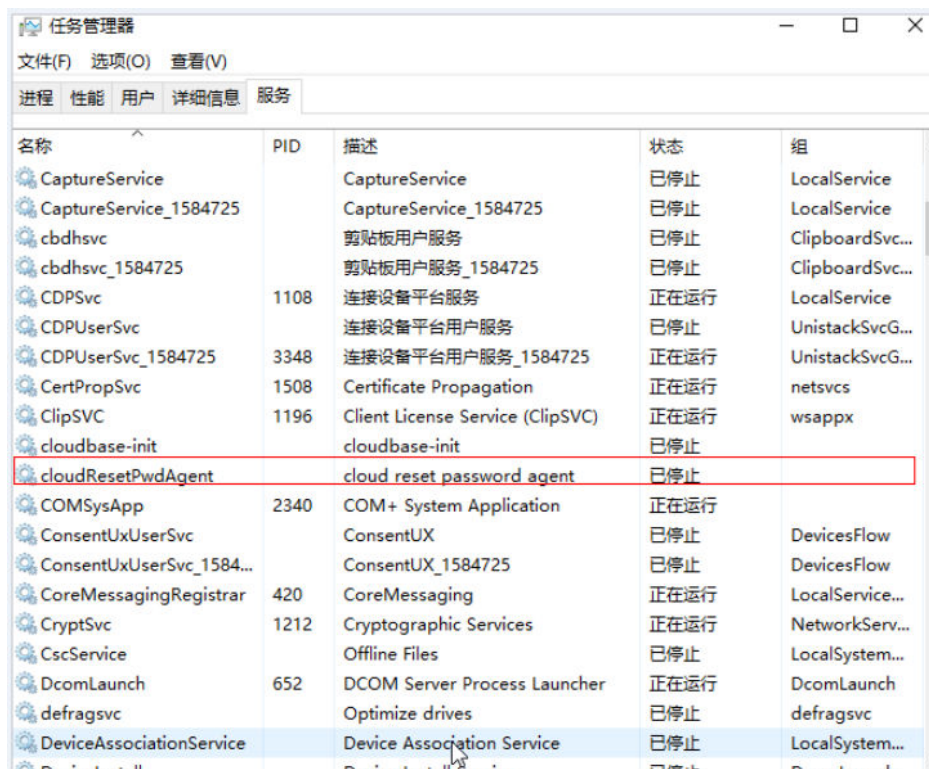
安装一键式重置密码插件对插件的解压目录无特殊要求，请您自定义。

步骤4 安装一键式重置密码插件。

1. 双击“CloudResetPwdAgent.Windows”文件夹下的“setup.bat”文件。
安装密码重置插件。
2. 查看任务管理器，检查密码重置插件是否安装成功。

如果在任务管理器中查找到了cloudResetPwdAgent服务，如图9-5所示，表示安装成功，否则安装失败。

图 9-5 安装插件成功



说明

如果密码重置插件安装失败，请检查安装环境是否符合要求，并重试安装操作。

---结束

后续处理

- 更新一键式重置密码插件后，如果无法通过弹性云服务器开机自动启动该插件，可以将其添加至开机启动项。具体方法请参见[一键式密码重置插件启动失败时如何操作？](#)。
- 一键式重置密码插件更新成功后，请勿删除重置密码进程CloudResetPwdAgent，否则，会导致一键式重装密码功能不可用。
- 云平台对一键式重置密码插件进行了升级，对于新创建的弹性云服务器，默认采用PIPE模式，不会占用端口。对于已创建云服务器，仍采用AUTO模式，随机占用31000~32999中的一个端口。占用端口的原则是：在该范围内，系统会按照自小到大的顺序，占用当前空闲的端口。

9.2.4 通过脚本批量更新 ECS 一键式重置密码插件（Linux）

操作场景

当您需要对多台Linux系统的云服务器批量更新一键式重置密码插件时，可参考本文档操作。

前提条件

- 准备好执行机，执行机需满足的条件请参见[约束与限制](#)。
- 需要提前准备待批量安装插件的云服务器的IP地址、root用户的密码信息或者私钥文件。
- 执行机应该与待更新机器在同一VPC下。
- 在执行完步骤6之后可以解绑eip。

约束与限制

- 需要选取一台操作系统为CentOS 7（公共镜像）且已绑定弹性公网IP的云服务器作为执行机，且与待批量安装插件的弹性云服务器之间网络需要互通。

说明

- 若已配置yum内部源，可不需要绑定弹性公网IP。
- 仅支持对使用同一密钥对的弹性云服务器执行批量安装插件的操作。

操作步骤

1. 以root用户登录执行机。
2. 执行以下命令，安装批量脚本运行所需要的依赖。
yum install ansible -y
3. 请参考[获取并校验一键式重置密码插件完整性（Linux）](#)，下载对应的一键式重置密码插件CloudResetPwdAgent.zip并完成完整性校验。
安装一键式重置密码插件对插件的具体放置目录无特殊要求，请您自定义。
4. 执行以下命令，将批量操作脚本下载到root目录下。
curl URL > ~/batch_update_log4j_version.py
其中，URL为批量操作的执行脚本。
请根据云服务器所在区域选择脚本的下载地址：

- 华北-北京一: https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/batch_update_resetpwd/batch_update_log4j_version.py
 - 华北-北京四: https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/linux/batch_update_resetpwd/batch_update_log4j_version.py
 - 华东-上海二: https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/linux/batch_update_resetpwd/batch_update_log4j_version.py
 - 华南-广州: https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/linux/batch_update_resetpwd/batch_update_log4j_version.py
 - 中国-香港: https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/linux/batch_update_resetpwd/batch_update_log4j_version.py
 - 亚太-曼谷: https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/linux/batch_update_resetpwd/batch_update_log4j_version.py
5. 执行以下命令，将更新插件脚本下载到root目录下。
- ```
curl URL > ~/update_log4j_version_for_resetpwdagent.sh
```
- 其中，URL为更新插件脚本的下载地址。
- 请根据云服务器所在区域选择脚本的下载地址：
- 华北-北京一: [https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/batch\\_update\\_resetpwd/update\\_log4j\\_version\\_for\\_resetpwdagent.sh](https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/linux/batch_update_resetpwd/update_log4j_version_for_resetpwdagent.sh)
  - 华北-北京四: [https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/linux/batch\\_update\\_resetpwd/update\\_log4j\\_version\\_for\\_resetpwdagent.sh](https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/linux/batch_update_resetpwd/update_log4j_version_for_resetpwdagent.sh)
  - 华东-上海二: [https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/linux/batch\\_update\\_resetpwd/update\\_log4j\\_version\\_for\\_resetpwdagent.sh](https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/linux/batch_update_resetpwd/update_log4j_version_for_resetpwdagent.sh)
  - 华南-广州: [https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/linux/batch\\_update\\_resetpwd/update\\_log4j\\_version\\_for\\_resetpwdagent.sh](https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/linux/batch_update_resetpwd/update_log4j_version_for_resetpwdagent.sh)
  - 中国-香港: [https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/linux/batch\\_update\\_resetpwd/update\\_log4j\\_version\\_for\\_resetpwdagent.sh](https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/linux/batch_update_resetpwd/update_log4j_version_for_resetpwdagent.sh)
  - 亚太-曼谷: [https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/linux/batch\\_update\\_resetpwd/update\\_log4j\\_version\\_for\\_resetpwdagent.sh](https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/linux/batch_update_resetpwd/update_log4j_version_for_resetpwdagent.sh)
6. 检查如下脚本是否在root目录下：
- batch\_update\_log4j\_version.py
  - update\_log4j\_version\_for\_resetpwdagent.sh
  - CloudResetPwdAgent.zip
7. 执行以下命令，新建并编辑host\_list.txt，按i进入编辑模式。
- ```
vi host_list.txt
```
- 将需要自动安装驱动的云服务器的相关信息填写到host_list.txt文件中。

文件填写格式与登录待切换弹性云服务器的方式需要匹配。

- 使用密钥对方式鉴权的云服务器，填写方式如下：

⚠ 注意

- 使用密钥对方式鉴权的云服务器，请将创建弹性云服务器时保存的私钥文件，上传至与host_list.txt同一个文件夹下。
- 使用密钥对方式鉴权的云服务器，请确保私钥文件的权限为400。

chmod 400 私钥文件

每行仅需填写云服务器IP地址。

示例：

```
192.168.1.10  
192.168.1.11
```

- 使用密码方式登录的云服务器，填写方式如下：

请严格按照每行ip,root用户密码，中间以英文逗号隔开的格式填写。

示例：

```
192.168.1.10,'*****'  
192.168.1.11,'*****'
```

8. 运行批量执行操作脚本“batch_update_log4j_version.py”。

- 密钥对鉴权方式的服务器

如果私钥文件与批量操作执行脚本在一个目录下，则直接指定私钥文件名称即可。

python batch_update_log4j_version.py {私钥文件路径/私钥文件名称}

图 9-6 运行成功

```
2024-04-11 16:59:04 Start copying the file /root/CloudResetPwdAgent.zip to all hosts.  
2024-04-11 16:59:09 Start copying the file /root/update_log4j_version_for_resetpwdagent.sh to all hosts.  
2024-04-11 16:59:10 Start executing scripts on all hosts, it will take a while..  
2024-04-11 17:00:03 Please check the execution result.  
status ip msg  
[SUCCESS] ██████████ install resetpwd successful  
Total: 1 Success: 1 Failed: 0  
You can check the logs/exec_origin.log for details.
```

若如图9-6所示，则表示脚本运行成功。

- 密码方式鉴权的服务器。

python batch_update_log4j_version.py

图 9-7 运行成功

```
2024-04-11 16:22:27 Start copying the file /root/CloudResetPwdAgent.zip to all hosts.  
2024-04-11 16:22:32 Start copying the file /root/update_log4j_version_for_resetpwdagent.sh to all hosts.  
2024-04-11 16:22:33 Start executing scripts on all hosts, it will take a while..  
2024-04-11 16:23:47 Please check the execution result.  
status ip msg  
[SUCCESS] ██████████ install resetpwd successful  
Total: 1 Success: 1 Failed: 0  
You can check the logs/exec_origin.log for details.
```

若如图9-7所示，则表示脚本运行成功。

9. 执行如下命令，在“/root/logs/exec_origin.log”的最后一行查看运行结果日志。

vim /root/logs/exec_origin.log

若如图9-8所示，则表示批量更新一键式重置密码插件成功。

图 9-8 运行结果日志

```
begin install CloudResetPwdAgent
=====cp LinuxCloudResetPwdAgent=====
=====y1 /etc/hosts=====
=====begin install CloudResetPwdAgent=====
Detected RHEL or Fedora:
Installing the cloudResetPwdAgent daemon using systemd...
creating default service file...
Reading file /etc/systemd/system/cloudResetPwdAgent.service

/home/CloudResetPwdAgent/CloudResetPwdAgent.Linux
cloudResetPwdAgent install successfully.
2024-04-11 17:08:03 Info:sh setup.sh successful
2024-04-11 17:08:03 Info:install resetpwd successful
2024-04-11 17:08:03 Info:update file permission successfulRemoved symlink /etc/systemd/system/multi-user.target.wants/cloudResetPwdAgent.service.
ls: cannot access /CloudResetPwdUpdateAgent: No such file or directory
Created symlink from /etc/systemd/system/multi-user.target.wants/cloudResetPwdAgent.service to /etc/systemd/system/cloudResetPwdAgent.service.
Redirecting to /bin/systemctl status cloudResetPwdAgent.service
```

后续处理

- 更新一键式重置密码插件后，如果无法通过弹性云服务器开机自动启动该插件，可以将其添加至开机启动项。具体方法请参见[一键式密码重置插件启动失败时如何操作？](#)。
- 一键式重置密码插件更新成功后，请勿删除重置密码进程CloudResetPwdAgent，否则，会导致一键式重装密码功能不可用。
- 云平台对一键式重置密码插件进行了升级，对于新创建的弹性云服务器，默认采用PIPE模式，不会占用端口。对于已创建云服务器，仍采用AUTO模式，随机占用31000~32999中的一个端口。占用端口的原则是：在该范围内，系统会按照自小到大的顺序，占用当前空闲的端口。

9.2.5 通过脚本批量更新 ECS 一键式重置密码插件（Windows）

操作场景

当您需要对多台Windows系统的云服务器批量更新一键式重置密码插件时，可参考本文档操作。

前提条件

- 准备好执行机，执行机需满足的条件请参见[约束与限制](#)。
- 需要提前准备待批量安装插件的云服务器的IP地址、Administrator用户的密码信息。
- 执行机应该与待更新机器在同一VPC下。
- 在执行完步骤7之后可以解绑eip。

约束与限制

需要选取一台操作系统为CentOS 8.2 (公共镜像)且已绑定弹性公网IP的云服务器作为执行机，且与待批量安装插件的弹性云服务器之间网络需要互通。

📖 说明

若已配置yum内部源，可不需要绑定弹性公网IP。

操作步骤

1. 以root用户登录执行机。
2. 执行以下命令，安装批量脚本运行所需要的依赖。

- a. 执行以下命令，安装epel。
yum install epel-release -y
- b. 执行以下命令，安装ansible。
yum install ansible -y --skip-broken
执行**ansible --version**命令检查ansible是否安装成功。

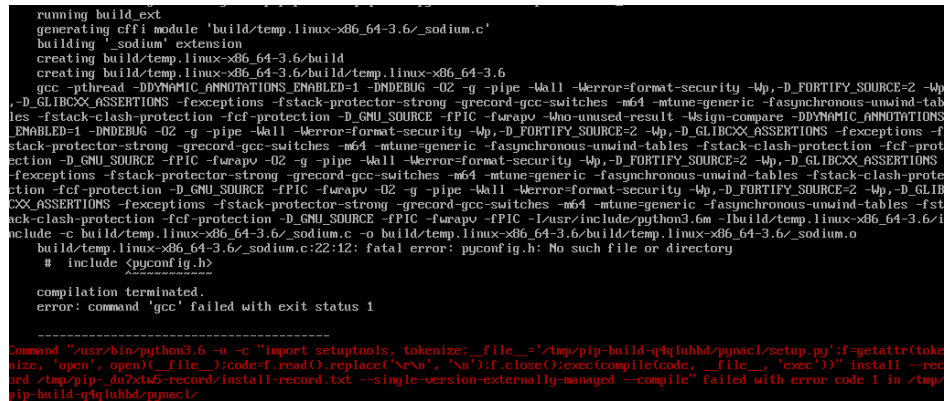
说明

如果因为yum源配置问题导致无法安装ansible，可以使用如下命令安装ansible：

```
yum install python3 python3-pip
pip3 install --upgrade pip
pip3 install ansible
```

- c. 执行以下命令，安装pip。
python3.6 -m pip install bcrypt==3.2.0 paramiko==3.3.1 cryptography==3.3.0 pywinrm PyYAML Jinja2 httpplib2 six
若出现如图9-9所示报错信息，请执行以下操作。
 - i. 执行以下命令，安装依赖：
dnf install python3-devel
 - ii. 然后再次执行以下命令：
python3.6 -m pip install bcrypt==3.2.0 paramiko==3.3.1 cryptography==3.3.0 pywinrm PyYAML Jinja2 httpplib2 six

图 9-9 报错信息



```
running build_ext
generating ffi module 'build/temp.linux-x86_64-3.6/_sodium.c'
building '_sodium' extension
creating build/temp.linux-x86_64-3.6/build
creating build/temp.linux-x86_64-3.6/build/temp.linux-x86_64-3.6
gcc -pthread -DDYNAMIC_ANNOTATIONS_ENABLED=1 -DNDEBUG -O2 -g -pipe -Wall -Werror=format-security -Wp,-D_FORTIFY_SOURCE=2 -Wp,-D_GLIBCXX_ASSERTIONS -fexceptions -fstack-protector-strong -grecord-gcc-switches -m64 -mtune=generic -fasynchronous-unwind-tables -fstack-clash-protection -fcf-protection -D_GNU_SOURCE -fPIC -fwrapv -Wno-unused-result -Wsign-compare -DDYNAMIC_ANNOTATIONS_ENABLED=1 -DNDEBUG -O2 -g -pipe -Wall -Werror=format-security -Wp,-D_FORTIFY_SOURCE=2 -Wp,-D_GLIBCXX_ASSERTIONS -fexceptions -fstack-protector-strong -grecord-gcc-switches -m64 -mtune=generic -fasynchronous-unwind-tables -fstack-clash-protection -fcf-protection -D_GNU_SOURCE -fPIC -fwrapv -O2 -g -pipe -Wall -Werror=format-security -Wp,-D_FORTIFY_SOURCE=2 -Wp,-D_GLIBCXX_ASSERTIONS -fexceptions -fstack-protector-strong -grecord-gcc-switches -m64 -mtune=generic -fasynchronous-unwind-tables -fstack-clash-protection -fcf-protection -D_GNU_SOURCE -fPIC -fwrapv -fPIC -I/usr/include/python3.6m -Ibuild/temp.linux-x86_64-3.6/include -c build/temp.linux-x86_64-3.6/_sodium.c -o build/temp.linux-x86_64-3.6/build/temp.linux-x86_64-3.6/_sodium.o
# include <pyconfig.h>
~~~~~
compilation terminated.
error: command 'gcc' failed with exit status 1

-----
Command "/usr/bin/python3.6 -u -c "import setuptools, tokenize; file_='/tmp/pip-build-gqgluhbd/pynacl/setup.py';f=getattr(tokenize, 'open', open)(__file__);code=f.read().replace('\r\n', '\n');f.close();exec(compile(code, __file__, 'exec'))" install --record /tmp/pip-4u7xtu6-record/install-record.txt --single-version-externally-managed --compile" failed with error code 1 in /tmp/pip-build-gqgluhbd/pynacl/
```

3. 执行以下命令，参考表9-4，下载对应区域的Windows操作系统的一键式重置密码插件CloudResetPwdAgent.zip并完成完整性校验。

以“华北-北京四”区域的Windows操作系统为例：

```
wget https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip
```

```
wget https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/windows/reset_pwd_agent/CloudResetPwdAgent.zip.sha256
```

完整性校验步骤，请参见[获取并校验一键式重置密码插件完整性（Linux）](#)的步骤4。

安装一键式重置密码插件对插件的具体放置目录无特殊要求，请您自定义。

4. 执行以下命令，将对应OS架构的Windows版本的安装包下载到root目录下：
执行命令pwd确认当前路径是否为/root，若不是请切换至/root路径下。

- 32位操作系统，x86架构：

```
wget https://www.7-zip.org/a/7z2107.exe '--no-check-certificate'
```

- 64位操作系统，x86架构：

```
wget https://www.7-zip.org/a/7z2107-x64.exe '--no-check-certificate'
```

- 64位操作系统，鲲鹏架构：

```
wget https://www.7-zip.org/a/7z2107-x64.exe '--no-check-certificate'
```

5. 执行以下命令，将批量操作脚本下载到root目录下。

```
curl URL > ~/batch_update_log4j_version_for_windows.py
```

其中，URL为批量操作的执行脚本。

请根据云服务器所在区域选择脚本的下载地址：

- 华北-北京一：https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/windows/batch_update_resetpwd/batch_update_log4j_version_for_windows.py

- 华北-北京四：https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/windows/batch_update_resetpwd/batch_update_log4j_version_for_windows.py

- 华东-上海二：https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/windows/batch_update_resetpwd/batch_update_log4j_version_for_windows.py

- 华南-广州：https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/windows/batch_update_resetpwd/batch_update_log4j_version_for_windows.py

- 中国-香港：https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/windows/batch_update_resetpwd/batch_update_log4j_version_for_windows.py

- 亚太-曼谷：https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/windows/batch_update_resetpwd/batch_update_log4j_version_for_windows.py

6. 执行以下命令，将更新插件脚本下载到root目录下。

```
curl URL > ~/update_log4j_version_for_resetpwdagent_windows.bat
```

其中，URL为更新插件脚本的下载地址。

请根据云服务器所在区域选择脚本的下载地址：

- 华北-北京一：https://cn-north-1-cloud-reset-pwd.obs.cn-north-1.myhuaweicloud.com/windows/batch_update_resetpwd/update_log4j_version_for_resetpwdagent_windows.bat

- 华北-北京四：https://cn-north-4-cloud-reset-pwd.obs.cn-north-4.myhuaweicloud.com/windows/batch_update_resetpwd/update_log4j_version_for_resetpwdagent_windows.bat

- 华东-上海二：https://cn-east-2-cloud-reset-pwd.obs.cn-east-2.myhuaweicloud.com/windows/batch_update_resetpwd/update_log4j_version_for_resetpwdagent_windows.bat

- 华南-广州：https://cn-south-1-cloud-reset-pwd.obs.cn-south-1.myhuaweicloud.com/windows/batch_update_resetpwd/update_log4j_version_for_resetpwdagent_windows.bat

- 中国-香港: https://ap-southeast-1-cloud-reset-pwd.obs.ap-southeast-1.myhuaweicloud.com/windows/batch_update_resetpwd/update_log4j_version_for_resetpwdagent_windows.bat
 - 亚太-曼谷: https://ap-southeast-2-cloud-reset-pwd.obs.ap-southeast-2.myhuaweicloud.com/windows/batch_update_resetpwd/update_log4j_version_for_resetpwdagent_windows.bat
7. 检查如下文件是否在root目录下。
 - batch_update_log4j_version_for_windows.py
 - update_log4j_version_for_resetpwdagent_windows.bat
 - CloudResetPwdAgent.zip
 - 7z*.exe
 8. 执行以下命令, 新建并编辑host_list.txt, 按i进入编辑模式。

vi host_list.txt

将需要自动安装驱动的云服务器器的相关信息填写到host_list.txt文件中。

填写Administrator用户的IP和密码, 请严格按照每行“IP,用户密码”的格式填写, 中间以英文逗号隔开。

示例:

```
192.168.1.10,'*****'  
192.168.1.11,'*****'
```

9. 执行以下命令, 添加ansible配置文件。

```
mkdir -p /etc/ansible
```

```
touch /etc/ansible/ansible.cfg
```

10. 运行批量执行操作脚本“batch_update_log4j_version_for_windows.py”。
- ```
python3.6 batch_update_log4j_version_for_windows.py
```

图 9-10 运行脚本

```
2022-01-08 17:00:36 Start copying the file /root/CloudResetPwdAgent.zip to all hosts.
2022-01-08 17:00:43 Start copying the file /root/update_log4j_version_for_resetpwdagent_windows.bat to all hosts.
2022-01-08 17:00:45 Start copying the file /root/7z.exe to all hosts.
2022-01-08 17:00:46 Start executing scripts on all hosts, it will take a while..
2022-01-08 17:00:57 Please check the execution result.
status ip msg
[SUCCESS] 192.168.96.118 started successfully
Total: 1 Success: 1 Failed: 0
You can check the logs/exec_origin.log for details.
```

11. 执行如下命令, 在“/root/logs/exec\_origin.log”的最后一行查看运行结果日志。

```
vim /root/logs/exec_origin.log
```

若如下图所示, 则表示运行成功。

图 9-11 运行成功

```
"C:\temp\CloudResetPwdAgent\CloudResetPwdUpdateAgent.Windows\CloudResetPwdUpdateAgent\lib\json-20160810.jar",
"C:\temp\CloudResetPwdAgent\CloudResetPwdUpdateAgent.Windows\CloudResetPwdUpdateAgent\lib\log4j-api-2.17.0.jar",
"C:\temp\CloudResetPwdAgent\CloudResetPwdUpdateAgent.Windows\CloudResetPwdUpdateAgent\lib\log4j-core-2.17.0.jar",
"C:\temp\CloudResetPwdAgent\CloudResetPwdUpdateAgent.Windows\CloudResetPwdUpdateAgent\lib\resetpwdupdateagent.jar",
"C:\temp\CloudResetPwdAgent\CloudResetPwdUpdateAgent.Windows\CloudResetPwdUpdateAgent\lib\wrapper.dll",
"C:\temp\CloudResetPwdAgent\CloudResetPwdUpdateAgent.Windows\CloudResetPwdUpdateAgent\lib\wrapper.jar",
"205 File(s) copied",
"The cloud reset password update agent service is starting.",
"The cloud reset password update agent service was started successfully.",
"
2022/01/08 17:00:56.63 \Info:\install ResetPwdUpdateAgent success\\"",
2022/01/08 17:00:56.83 \Info:*****Update Success*****\\"",
"Press any key to continue . . ."
```

## 后续处理

- 更新一键式重置密码插件后，如果无法通过弹性云服务器开机自动启动该插件，可以将其添加至开机启动项。具体方法请参见[一键式密码重置插件启动失败时如何操作？](#)。
- 一键式重置密码插件更新成功后，请勿删除重置密码进程CloudResetPwdAgent，否则，会导致一键式重装密码功能不可用。
- 云平台对一键式重置密码插件进行了升级，对于新创建的弹性云服务器，默认采用PIPE模式，不会占用端口。对于已创建云服务器，仍采用AUTO模式，随机占用31000~32999中的一个端口。占用端口的原则是：在该范围内，系统会按照自小到大的顺序，占用当前空闲的端口。

## 9.3 管理 ECS 密钥对

### 9.3.1 密钥对使用场景介绍

#### 密钥对

密钥对，即SSH密钥对，是为用户提供远程登录云服务器的认证方式，是一种区别于传统的用户名和密码登录的认证方式。

密钥对包含一个公钥和一个私钥，公钥自动保存在KPS（Key Pair Service）中，私钥由用户保存在本地。若用户将公钥配置在Linux云服务器中，则可以使用私钥登录Linux云服务器，而不需要输入密码。由于密钥对可以让用户无需输入密码登录到Linux云服务器，因此，可以防止由于密码被拦截、破解造成的账户密码泄露，从而提高Linux云服务器的安全性。

您可以通过[数据加密服务](#)（Data Encryption Workshop）管理密钥对，包括创建、导入、绑定、查看、重置、替换、解绑、删除密钥对等。

本章节主要介绍如何创建和导入密钥对，其余操作，请参见“[管理密钥对](#)”。

#### 使用场景

用户在购买弹性云服务器时，建议选择密钥对进行用户身份认证，或者通过提供的密钥对获取Windows操作系统弹性云服务器的登录密码。

- 登录Linux操作系统的弹性云服务器  
若用户购买的是Linux操作系统的弹性云服务器，可以直接使用密钥对远程登录云服务器。
  - 创建弹性云服务器时，选择“密钥对方式”登录，详细操作，请参见[步骤三：高级配置](#)的“设置‘登录凭证’”。
  - 创建弹性云服务器完成后，通过“[绑定密钥对](#)”的方式为云服务器绑定密钥对。
- 获取Windows操作系统弹性云服务器的登录密码  
若用户购买的是Windows操作系统的弹性云服务器，可以通过密钥对的私钥获取登录密码，该密码为随机密码，安全性高。  
详细内容，请参见[获取Windows弹性云服务器的密码](#)

## 创建密钥对操作指引

您可以使用已有密钥对或新建一个密钥对，用于远程登录身份验证。

- 新建密钥对

如果没有可用的密钥对，需新建一个密钥对，生成公钥和私钥，并在登录弹性云服务器时提供私钥进行鉴权。创建密钥对的方法如下：

- **（推荐）通过管理控制台创建密钥对**：公钥自动保存在系统中，私钥由用户保存在本地。
- **通过PuTTY Key Generator工具创建密钥对**：公钥和私钥均保存在用户本地。  
创建成功的密钥对，还需要执行**导入密钥对**，导入系统才能正常使用密钥对。

- 使用已有密钥对

如果本地已有密钥对（例如，使用PuTTYgen工具生成的密钥对），可以在管理控制台导入密钥对公钥，由系统维护您的公钥文件。具体操作请参见**导入密钥对**。

### 说明

如果已有密钥对的公钥文件是通过puttygen.exe工具的“Save public key”按钮保存的，该公钥文件不能直接导入管理控制台。

如需继续使用该密钥对作远程身份验证，请参见**通过puttygen.exe工具创建的密钥对，导入管理控制台失败怎么办？**

## 约束与限制


- 仅支持远程登录Linux云服务器。
- 通过管理控制台创建的SSH-2密钥对仅支持“RSA-2048”加解密算法。
- 密钥对仅适用于本区域的云服务器。
- 通过外部导入的密钥对支持的加解密算法为：
  - RSA-1024
  - RSA-2048
  - RSA-4096
- 私钥是保证您的弹性云服务器安全的重要手段之一，用于远程登录身份认证，为保证弹性云服务器安全，只能下载一次，请妥善保管。

### 9.3.2 （推荐）通过管理控制台创建密钥对

#### 操作场景

您可以通过管理控制台创建密钥对，创建完成后，公钥自动保存在系统中，私钥由用户保存在本地。

#### 操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的  ，选择区域和项目。

- 单击“☰”，选择“计算 > 弹性云服务器”。
- 在左侧导航树中，选择“密钥对”。
- 在“密钥对管理”页面，单击“创建密钥对”。

#### 📖 说明

密钥对包括私有密钥对和账户密钥对。私有密钥对仅限本用户使用，账户密钥对账户下其他用户也可使用。

您可以根据业务需要选择创建合适的密钥对。

- 根据界面提示输入参数。
  - 输入密钥对“名称”。
  - 输入“密钥对类型”。
  - 选择“KMS加密”。

如果同意托管私钥并勾选“我同意将密钥对私钥托管。”则显示该参数，用于对私钥进行加密。

#### 📖 说明

- 若未托管私钥，为保证安全，私钥只能下载一次，请妥善保管。  
若不慎遗失，您可以通过[重置密钥对](#)的方式，重新为弹性云服务器绑定密钥对。
  - 若已授权托管私钥，可根据需要将托管的私钥导出使用，详细内容请参见[导出私钥](#)。
- 勾选“我已经阅读并同意《密钥对管理服务免责声明》”
- 单击“确定”

## 相关操作

- 如果私钥文件丢失，您可以通过“[重置密钥对](#)”功能重新配置弹性云服务器的密钥对。
- 如果私钥文件泄露，您可以通过“[替换密钥对](#)”功能使用新的密钥对替换弹性云服务器内的公钥。

## 9.3.3 通过 PuTTY Key Generator 工具创建密钥对

### 操作场景

您可以通过puttygen.exe工具创建密钥对，创建完成后，公钥和私钥均保存在用户本地。

#### 📖 说明

通过puttygen.exe工具创建成功的密钥对，还需要执行[导入密钥对](#)，导入系统才能正常使用密钥对。

### 操作步骤

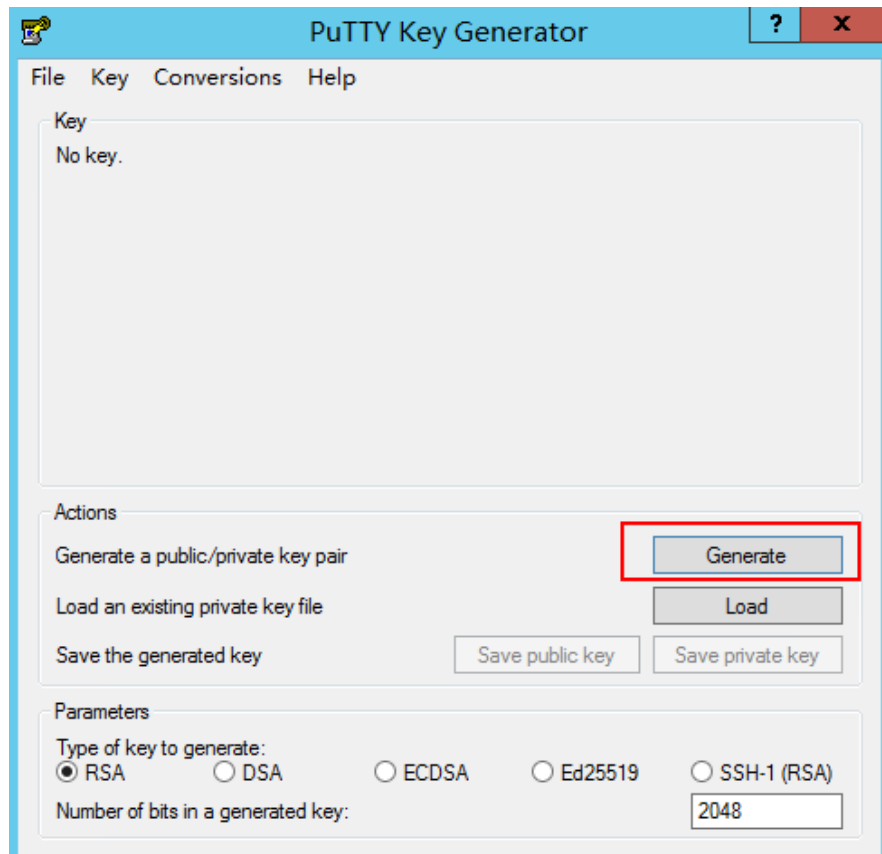
- 在以下路径中下载并安装PuTTY和PuTTYgen。  
<https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html>

 说明

PuTTYgen是密钥生成器，用于创建密钥对，生成一对公钥和私钥供PuTTY使用。

2. 生成公钥和私钥文件。
  - a. 双击“puttygen.exe”，打开“PuTTY Key Generator”。

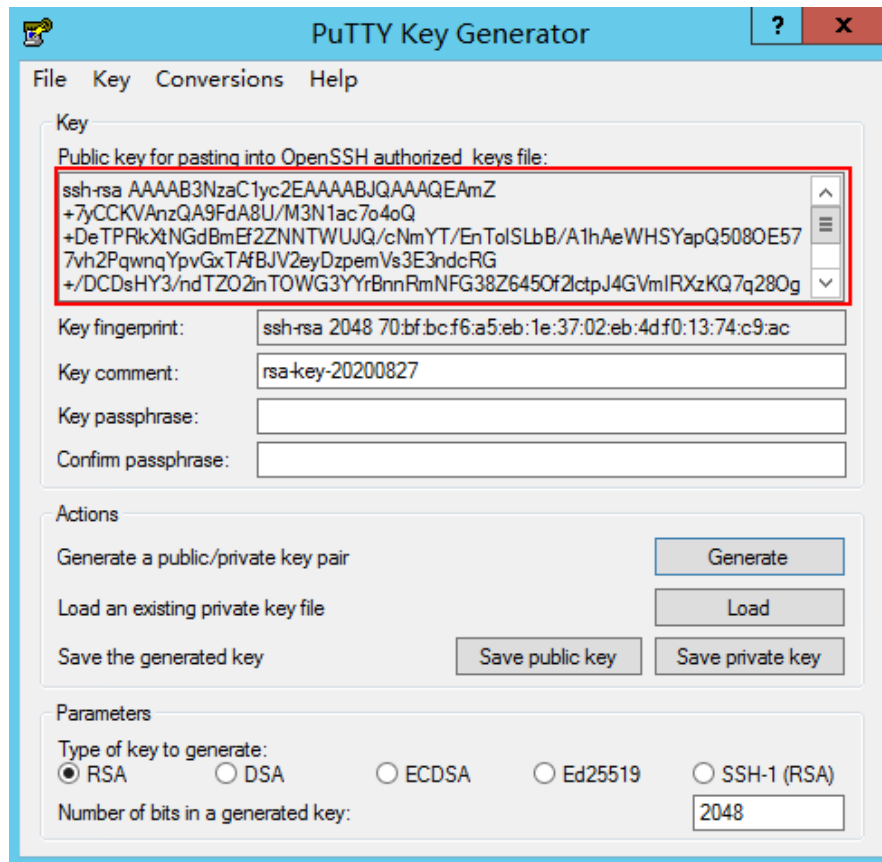
图 9-12 PuTTY Key Generator



- b. 单击“Generate”。

密钥生成器将自动生成一对公钥和私钥，其中，图9-13的红框内容为生成的公钥文件。

图 9-13 生成公钥和私钥文件



3. 复制红框中的公钥内容，并将其粘贴在文本文档中，以.txt格式保存在本地，保存公钥文件。

#### 说明

请勿直接单击“Save public key”保存公钥文件。因为通过puttygen.exe工具的“Save public key”按钮保存在本地的公钥文件，公钥内容的格式会发生变化，不能直接导入管理控制台使用。

4. 保存私钥文件。

根据不同用途，需将私钥文件保存为不同格式。为保证安全，私钥只能下载一次，请妥善保管。

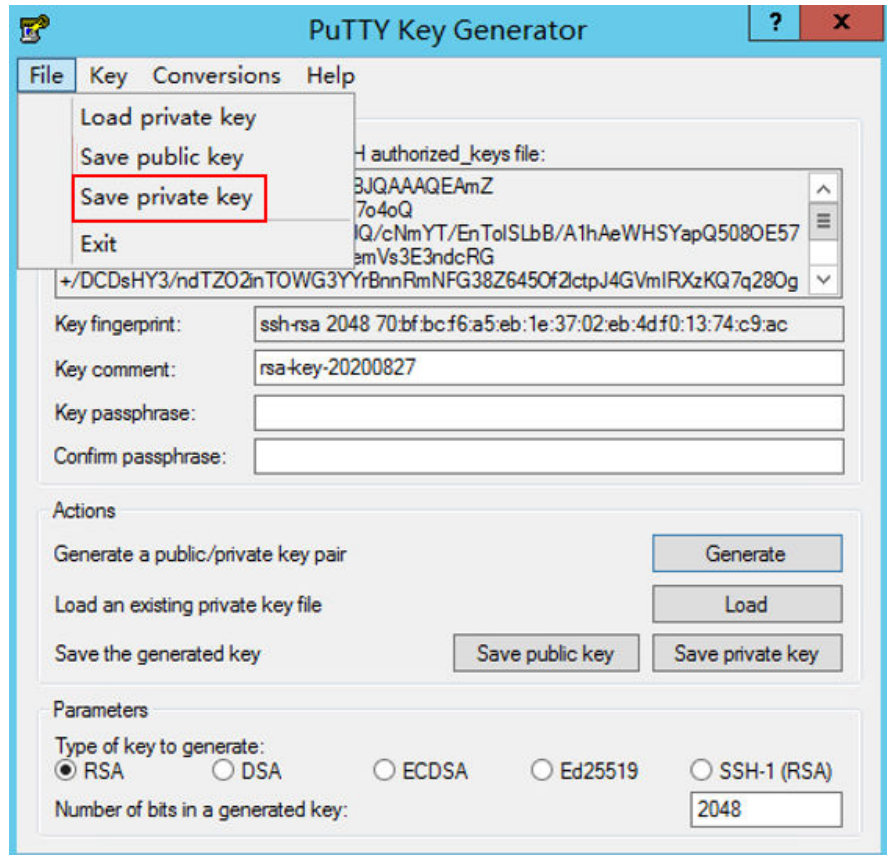
- 保存私钥文件为“.ppk”格式。

当用户使用PuTTY工具登录Linux弹性云服务器时，需使用“.ppk”格式的私钥文件。保存方法如下：

- i. 在“PuTTY Key Generator”界面，选择“File > Save private key”。



图 9-14 保存私钥文件



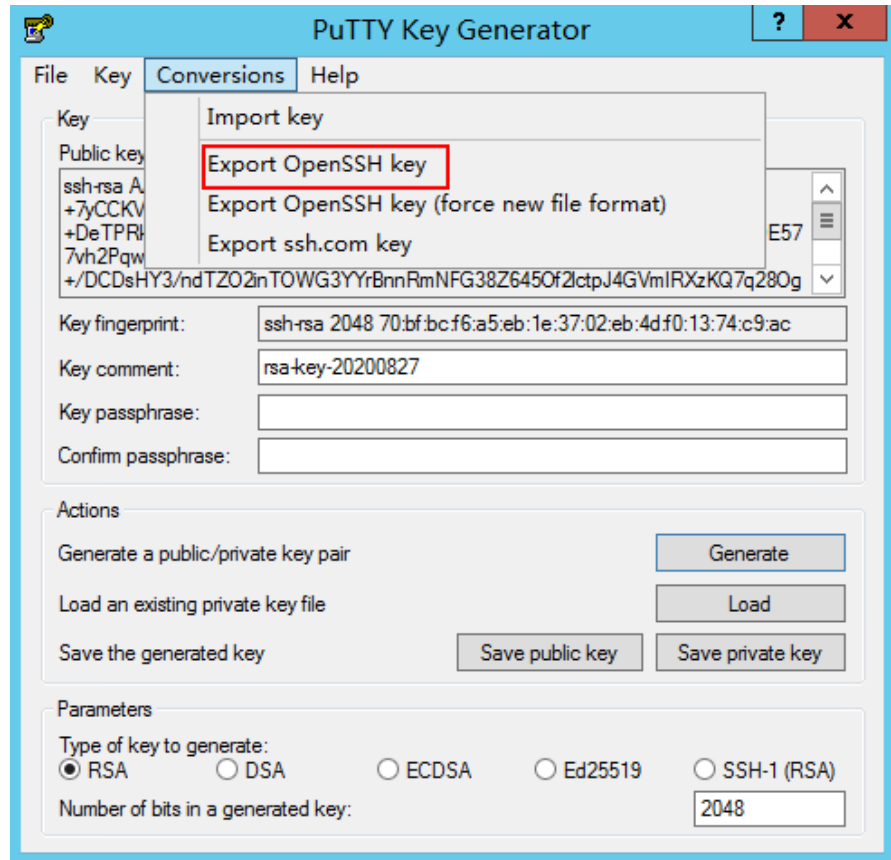
- ii. 保存转化后的私钥到本地。例如：kp-123.ppk
- 保存私钥文件为“.pem”格式。  
当用户使用Xshell工具登录Linux弹性云服务器，或者获取Windows弹性云服务器的密码时，需使用“.pem”格式的私钥文件。保存方法如下：
  - i. 选择“Conversions > Export OpenSSH key”。

**说明**

如果该私钥文件用于Windows弹性云服务器的获取密码操作，在选择“Export OpenSSH key”时，请勿填写“Key passphrase”信息，否则会导致获取密码失败。



图 9-15 保存私钥文件



- ii. 保存私钥到本地。例如：kp-123.pem
5. 公钥和私钥文件按照需要正确保存后，请参见[导入密钥对](#)的“复制公钥文件内容方式”，将公钥导入系统中。

## 相关操作

- 如果私钥文件丢失，您可以通过“[重置密钥对](#)”功能重新配置弹性云服务器的密钥对。
- 如果私钥文件泄露，您可以通过“[替换密钥对](#)”功能使用新的密钥对替换弹性云服务器内的公钥。

## 9.3.4 导入密钥对

### 操作场景

以下两种场景，需要执行导入密钥对的操作：

- 通过puttygen.exe工具新建密钥对，需要将密钥对导入系统才能正常使用。
- 如果本地已有密钥对（例如，使用PuTTYgen工具生成的密钥对），可以在管理控制台导入密钥对公钥，由系统维护您的公钥文件。

## 说明

如果已有密钥对的公钥文件是通过puttygen.exe工具的“Save public key”按钮保存的，该公钥文件不能直接导入管理控制台。

如需继续使用该密钥对作远程身份验证，请参见[通过puttygen.exe工具创建的密钥对，导入管理控制台失败怎么办？](#)

## 操作步骤

1. 登录管理控制台。
2. 单击“☰”，选择“计算 > 弹性云服务器”。
3. 在左侧导航树中，选择“密钥对”。
4. 在“密钥对管理”页面，单击“导入密钥对”。

图 9-16 导入密钥对



5. 在“导入密钥对”页面，导入公钥内容、私钥内容。  
通过单击“选择文件”，可以导入本地保存的公钥文件（例如，[通过PuTTY Key Generator工具创建密钥对](#)的步骤3中保存的“.txt”格式文件）。  
更多参数设置，请参见[导入密钥对](#)。

图 9-17 参数设置

### 导入密钥对

密钥对当前免费使用，但有配额限制。首次创建账号密钥对时，需要具有Tenant Administrator系统角色的用户完成一次账号密钥对创建。

导入公钥有两种方式：

- 1、选择文件导入，导入成功后，可以修改公钥名称。
- 2、复制内容到“公钥内容”文本框，手工输入公钥名称。

说明：仅支持导入使用RSA算法的密钥，且密钥长度为1024、2048、4096 (bit)。

名称

公钥   
未选择任何文件

公钥内容

私钥内容

KMS加密

密钥ID a0ff3eb-4804-

我同意将密钥对私钥托管。[了解详情](#)

我已经阅读并同意 [《密钥对管理服务免责声明》](#)

6. 单击“确定”，完成密钥对的导入。

## 相关链接

- [无法导入密钥对，怎么办？](#)
- [通过puttygen.exe工具创建的密钥对，导入管理控制台失败怎么办？](#)

## 9.3.5 获取和清除云服务器密码（Windows）

### 9.3.5.1 获取 Windows 弹性云服务器的密码

#### 操作场景

登录Windows操作系统的弹性云服务器时，需使用密码方式登录。因此，用户需先根据创建弹性云服务器时使用的密钥文件，获取该弹性云服务器初始安装时系统生成的管理员密码（Administrator账户或Cloudbase-init设置的账户）。该密码为随机密码，安全性高，请放心使用。

请根据您的个人需求，通过管理控制台或API方式获取Windows弹性云服务器的初始密码，获取方法请参见本节内容。

## 前提条件

已获取创建弹性云服务器时使用的私钥文件（.pem文件）。

私钥文件（.pem文件）与创建弹性云服务器时选择的密钥对有关，在创建密钥对时生成。密钥对的创建和使用方法，请参见[密钥对使用场景介绍](#)。

若私钥文件不慎遗失，可以通过[重置密钥对](#)的方式，重新为弹性云服务器绑定密钥对。若已授权托管私钥，可根据需要将托管的私钥导出使用，详细内容请参见[导出私钥](#)。

## 通过管理控制台获取密码



1. 获取创建弹性云服务器时使用的私钥文件（.pem文件）。
2. 登录管理控制台。
3. 单击管理控制台左上角的 ，选择区域和项目。
4. 单击“”，选择“计算 > 弹性云服务器”。
5. 在弹性云服务器列表，选择待获取密码的弹性云服务器。
6. 选择“操作 > 更多”，单击“获取密码”。

图 9-18 获取密码



### 说明

如果不显示“获取密码”，可能是由于弹性云服务器未安装一键式重置密码插件。

您可以参考[重置Windows云服务器密码（未安装重置密码插件）](#)为弹性云服务器重置密码。

7. 通过密钥文件获取密码，获取方法有如下两种：
  - 单击“选择文件”，从本地上传密钥文件。
  - 将密钥文件内容复制粘贴在空白框中。
8. 单击“获取密码”，获取随机密码。

## 通过 API 方式获取密码

1. 获取创建弹性云服务器时使用的私钥文件（.pem文件）。
2. 设置调用接口的环境。
3. 具体方法请参见《弹性云服务器接口参考》的“接口调用方法”。
4. 获取密文密码。

调用获取密码的接口，获取密钥的公钥加密后的密文密码。该密文密码是通过RSA算法加密的。其中，URI格式为“GET /v2/{project\_id}/servers/{server\_id}/os-server-password”。

### 📖 说明

调用接口的具体方法请参见《弹性云服务器接口参考》的“获取云服务器密码”章节。

5. 解密密文密码。  
通过创建云服务器时使用的密钥的私钥文件对4中获取的密文密码进行解密。
  - a. 使用OpenSSL工具，执行以下命令，将密文密码转换为.key -nocrypt格式。  
**openssl pkcs8 -topk8 -inform PEM -outform DER -in *rsa\_pem.key* -out *pkcs8\_der.key* -nocrypt**
  - b. 调用Java类库org.bouncycastle.jce.provider.BouncyCastleProvider，使用密钥文件，编写代码解密密文。



### 9.3.5.2 清除 Windows 弹性云服务器的密码

#### 操作场景

为安全起见，建议用户获取初始密码后，执行清除密码操作，清除系统中记录的初始密码信息。

该操作不会影响弹性云服务器的正常登录与运行。清除密码后，系统不能恢复获取密码功能，因此，请在执行清除密码操作前，记录弹性云服务器密码信息。

#### 操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表，选择待获取密码的弹性云服务器。
5. 选择“操作 > 更多”，单击“清除密码”。  
系统弹窗提示用户是否确认清除密码。
6. 单击“确定”，执行清除密码操作。

# 10 启动模板管理

## 10.1 启动模板概述

### 什么是启动模板？

启动模板是可用于快速创建实例的模板。模板中包含了创建实例所需的配置信息，包括实例类型、网络配置、密钥对（不包含密码）等信息。

启动模板不支持修改，但可以创建多个版本，每个版本可以配置不同的参数，通过版本管理实现模板配置信息的不断更新，您可以使用模板的任意一个版本创建实例。

### 创建启动模板的方法

在控制台创建启动模板

具体操作请参考[创建启动模板](#)。

## 10.2 创建启动模板



### 操作场景

本节操作介绍在管理控制台创建启动模板的操作步骤。

### 约束限制



- 每个账号在每个区域最多可创建30个启动模板。
- 创建启动模板时，所有配置项均为可选。  
但如果缺失了创建实例的必要参数，例如规格、镜像类型，那么在使用该模板创建实例时需重新设置这些缺失的配置项。
- 启动模板创建成功后，不支持修改，如需修改模板中包含的配置信息，可以通过创建新版本的方式更新配置项。
- 支持的区域包括：亚太-新加坡、中国-香港。

## 通过控制台创建启动模板

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“启动模板”。
5. 在“启动模板”页面，单击“创建模板”。
6. 在启动模板页面，完成基础配置、网络配置、高级配置。  
关于基础配置、网络配置、高级配置的参数含义可以参考[购买弹性云服务器的操作步骤](#)。
7. 在确认配置页面，输入模板名称和模板版本描述，并单击“立即创建”。  
在启动模板列表页即可查看已创建的模板。

## 通过 ECS 的自定义购买页面创建启动模板

您还可以选择在创建云服务器时将实例的配置保存为启动模板。

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 单击“购买弹性云服务器”，系统进入购买页面。
5. 在自定义购买页面完成云服务器的基础配置、网络配置、高级配置。  
关于基础配置、网络配置、高级配置的参数含义可以参考[购买弹性云服务器的操作步骤](#)。
6. 在确认配置页面，选择“保存为启动模板”，并输入模板名称和模板版本描述，并单击“立即创建”。  
在启动模板列表页即可查看已创建的模板。



# 10.3 管理启动模板

## 操作场景

启动模板支持以下操作：

- [查看启动模板详情](#)
- [删除启动模板](#)

## 查看启动模板详情


1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。

4. 在左侧导航树中，选择“启动模板”。
5. 在“启动模板”页面，单击启动模板名称，查看启动模板的详情。

表 10-1 启动模板详情

| 参数   | 说明                                     |
|------|----------------------------------------|
| 模板名称 | 启动模板名称。                                |
| 模板ID | 启动模板ID。                                |
| 创建时间 | 启动模板的创建时间。                             |
| 描述信息 | 自定义的启动模板的描述信息。                         |
| 版本信息 | 版本信息中包含了当前版本模板的配置信息，例如区域、规格、镜像类型等配置信息。 |

## 删除启动模板

1. 登录管理控制台。
2. 单击“”，选择“计算 > 弹性云服务器”。
3. 在左侧导航树中，选择“启动模板”。
4. 找到要删除的启动模板，在操作列下单击“删除”。
5. 单击“是”。



# 11 智能购买组管理

## 11.1 智能购买组概述

### 什么是智能购买组？

智能购买组是一种新的算力获取方式，通过自定义配置，可以自动化获得不同可用区、不同实例类型以及不同计费模式的算力。极大程度降低投入成本，同时简化算力获取方式。

### 应用场景

智能购买组适用于图像渲染、无状态Web服务、基因测序、离线分析、函数计算、批量计算、样本分析、CI/CD与测试等场景。

### 使用须知

- 智能购买组可以跨可用区部署实例，但是不可以跨区域部署实例。
- 每个智能购买组的目标容量限制：
  - 如果以实例个数为目标容量最多创建500台实例。
  - 如果以vCPUs为目标容量最大创建40000vCPUs。
- 每个智能购买组最多指定一个启动模板。

### 智能购买组的优势

- 支持使用跨计费模式的实例部署集群  
华为云提供了多种计费模式：其中竞价计费型实例成本低，但存在中断回收机制，当市场价格高于报价或库存资源不足就会被中断回收。按需实例可以随时创建和删除，库存充足稳定，但价格高于竞价实例。  
智能购买组实现了同时搭配使用竞价计费型实例和按需实例，既可保证最低算力，同时又兼具性价比和实例灵活性。
- 支持使用跨可用区的实例部署集群  
智能购买组支持跨可用区部署实例集群，提高了集群的容灾能力。
- 支持使用多种类型实例部署集群

智能购买组支持选择多种类型的实例规格，可满足不同的使用场景。

- 灵活的实例分配策略

智能购买组设置多种实例分配策略以满足更多形态的业务需求。

支持设置集群目标容量、竞价计费型实例目标容量和按需实例目标容量。支持设置智能购买组的目标容量保持策略，即支持在与目标容量有差异时自动补齐对应容量的实例或释放超过容量的实例。

- 高性价比

智能购买组支持设置实例的分配策略为“最低价格策略”，确保智能购买组购买的所有实例的价格总和最低。

## 计费说明

智能购买组免费，但由智能购买组创建的实例正常计费。

具体请参考[弹性云服务器产品价格详情](#)。

## 11.2 创建智能购买组


### 操作场景

本节操作介绍在管理控制台创建智能购买组的操作步骤。

### 约束与限制

支持的区域包括：亚太-新加坡、中国-香港。

### 操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“≡”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“智能购买组”。
5. 在“智能购买组”页面，单击“创建智能购买组”。

6. 设置智能购买组名称。

智能购买组的名称，长度为2~64个字符，只能包含中文、字母、数字、下划线（\_）和中划线（-）。

7. 设置目标容量。

计划使用智能购买组交付的算力，您可以指定实例个数或vCPU数量。

勾选“使用按需实例提供部分算力”后，需设置按需实例容量，按需实例可以保证智能购买组的最低算力需求。

目标容量决定智能购买组交付的算力总和，以及竞价实例和按需实例的算力占比。

每个智能购买组的目标容量限制：

- 如果以实例个数为目标容量最多创建500台实例。

- 如果以vCPUs为目标容量最大创建40000vCPUs。
8. 选择启动模板。

智能购买组需要选择一个实例启动模板和对应的版本作为配置来源，您还可以基于该启动模板自定义多个可选的实例配置。
  9. 设置实例分配策略。
    - 最低价格策略：智能购买组购买的所有实例的价格总和最低。
    - 多可用区均衡策略：智能购买组购买的实例在多个可用区之间均匀分配算力。
    - 容量优化策略：智能购买组购买的实例按照大规格优先进行购买。

即目标容量为实例个数时智能购买组的vCPU容量最大；目标容量为vCPU大小时智能购买组的实例个数最小。
  10. 选择智能购买组类型。
    - 一次性交付：智能购买组仅在启动时尝试一次性创建目标容量的实例，即使未达到目标容量也不再重试。
    - 持续保持：智能购买组实时监控当前集群的容量与目标容量，如未达到目标容量继续创建实例，直至满足目标容量。如果超出目标容量则自动缩容至目标容量。
  11. 选择启动时间。

智能购买组开始创建实例的时间，与结束时间结合使用，确定智能购买组的有效时段。

    - 立即启动：智能购买组创建成功后立即启动。
    - 指定启动时间：自定义智能购买组的启动时间。
  12. 选择结束时间。

与启动时间结合使用，确定智能购买组的有效时段。

    - 无限期：智能购买组长期有效。
    - 指定结束时间：自定义智能购买组的过期时间。
  13. 设置全局竞价实例价格上限。

智能购买组内竞价实例的价格上限。当智能购买组内某一竞价实例的市场价格高于该价格上限时，该竞价实例将被中断回收。

如果同时设置了单个竞价实例的价格上限和全局竞价实例的价格上限，则以单个竞价实例的价格上限为准。

设置的竞价实例的价格上限不得小于0，如果高于按需实例价格，代表对竞价实例价格上限无约束限制。
  14. 智能购买组内实例释放设置。
    - 智能购买组过期时释放实例：当设置了智能购买组的结束时间时，如果智能购买组超过有效时间，即删除组内实例。
    - 超过容量时释放实例：当智能购买组的实例个数/vCPU之和超过目标容量时，即删除超出容量的实例。
- 说明**
- 如果不勾选超过容量时释放实例，那么当智能购买组超过目标容量时，超过容量的实例将被移出智能购买组，但并不释放。在ECS列表页面可以查询到被移出的实例，如需停止计费，请手动删除。
15. 单击“立即创建”。

## 执行结果

智能购买组创建成功后，会在您指定的时间启动并交付集群。如果您选择了持续保持模式，智能购买组会实时监控当前集群的容量与目标容量，在竞价实例被回收时自动创建新的实例。


## 11.3 管理智能购买组

### 操作场景

智能购买组支持以下操作：

- [查看智能购买组详情](#)
- [修改智能购买组](#)
- [删除智能购买组](#)

### 查看智能购买组详情


1. 登录管理控制台。
2. 单击“”，选择“计算 > 弹性云服务器”。
3. 在左侧导航树中，选择“智能购买组”。
4. 在智能购买组列表页面，单击智能购买组名称，查看智能购买组的详情。

在详情页面可以查看智能购买组的基本信息和容量配置信息。


基本信息包括智能购买组名称、使用的启动模板、智能购买组类型、实例分配策略等信息。

在容量配置信息中可以查看当前容量与目标容量，竞价实例/按需实例的当前容量和目标容量等信息。

### 修改智能购买组

1. 登录管理控制台。
  2. 单击“”，选择“计算 > 弹性云服务器”。
  3. 在左侧导航树中，选择“智能购买组”。
  4. 找到智能购买组，在操作列下单击“修改”。
- 支持修改智能购买组名称、目标容量、按需实例容量、全局竞价实例的价格上限、智能购买组过期时释放实例、超过容量时释放实例。
5. 修改完后单击“是”。

### 删除智能购买组

1. 登录管理控制台。
2. 单击“”，选择“计算 > 弹性云服务器”。
3. 在左侧导航树中，选择“智能购买组”。
4. 找到智能购买组，在操作列下单击“删除”。

5. 选择是否删除智能购买组内实例。  
如果选择不删除实例，您可以在ECS列表页查看保留的实例，如需停止实例计费请手动删除实例。
6. 单击“是”。

# 12 事件管理

## 12.1 事件概述

在日常运维中，华为云会对ECS实例所在底层宿主机的软硬件故障进行预测和主动规避。

当宿主机上的故障风险无法规避时，为避免因ECS实例的资源可用性或性能受损对您的业务造成的更大影响，系统会对受影响的实例生成事件并进行上报，例如实例重部署、本地盘换盘等，事件详细内容请参见[事件类型](#)。系统上报事件不会频繁发生。

您可通过云服务器控制台查看事件详情，包括事件类型，实例ID、事件状态等信息。也可以通过云监控服务的“事件监控”查看ECS实例的事件，详细内容请参见[查看事件监控数据](#)。

### 事件类型

系统支持上报的事件如[表12-1](#)所示。

表 12-1 支持系统上报的事件

| 事件类型  | 事件说明                                              | 事件影响                                                                                                | 用户侧处理建议                                                                                                           |
|-------|---------------------------------------------------|-----------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------|
| 实例重部署 | 当系统检测到ECS实例的底层宿主机异常，计划将ECS实例部署到新主机时，会自动上报实例重部署事件。 | 实例重部署过程中，云服务器将会有短暂时间不可用。<br>系统会在计划事件执行时间前24~72小时发送该系统事件通知。<br><b>须知</b><br>对于包含本地盘的实例，会丢失所有本地盘上的数据。 | 您可以根据业务需要选择如下处理方式，并在事件处理完成后，及时验证业务受损情况，如遇问题，请联系技术支持。<br><b>响应实例重部署事件</b><br>建议您在授权时选择业务低谷期为开始时间，如不指定将会以当前时间为开始时间。 |

| 事件类型  | 事件说明                                                              | 事件影响                                            | 用户侧处理建议                                                                                                                                                                                                                                                                                                                    |
|-------|-------------------------------------------------------------------|-------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 本地盘换盘 | 系统检测到ECS实例（含裸金属类型实例）的底层宿主主机存在磁盘故障风险，会对受影响的ECS实例自动生成本地盘换盘事件。       | 本地盘换盘会丢失本地盘上的数据。                                | <p>您可以根据业务需要选择如下处理方式，并在事件处理完成后，及时验证业务受损情况，如遇问题，请联系技术支持。</p> <p><b>须知</b></p> <p>本地盘换盘操作会丢失本地盘上的数据，如果无需保留本地盘数据，可根据业务需要选择如下处理方式。</p> <ul style="list-style-type: none"> <li>● <b>重部署</b>：会丢失所有本地盘数据</li> <li>● <b>授权换盘</b>：会丢失故障本地盘数据<br/>建议您在授权时选择业务低谷期为开始时间，如不指定将会以当前时间为开始时间。<br/>通常会在开始时间后5个工作日内完成本地盘换盘，请耐心等待。</li> </ul> |
| 实例迁移  | 当系统检测到ECS实例的底层宿主主机异常，需要进行重启、关机、下线等系统维护时，计划对ECS实例进行迁移，会自动上报实例迁移事件。 | 系统会先尝试对云服务器进行热迁移，如遇异常，则会触发HA机制（云服务器将会有短暂时间不可用）。 | 建议您在事件处理完成后，及时验证业务受损情况，如遇问题，请联系技术支持。                                                                                                                                                                                                                                                                                       |

| 事件类型 | 事件说明                                                           | 事件影响                          | 用户侧处理建议                                                                                                                                                                                        |
|------|----------------------------------------------------------------|-------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 系统维护 | 系统检测到ECS实例（含裸金属类型实例）的宿主机存在软硬件故障风险，计划对受影响的实例进行维护操作，会自动生成系统维护事件。 | 系统维护过程中，宿主机可能会进入下电状态，云服务器不可用。 | <p>您可以根据业务需要选择如下处理方式，并在事件处理完成后，及时验证业务受损情况，如遇问题，请联系技术支持。</p> <p><b>响应系统维护事件</b></p> <p>建议您在授权时确保实例的业务已离线并选择业务低谷期为开始时间，如不指定将会以当前时间为开始时间。</p> <p>不同故障系统维护的耗时不同。通常会在授权开始时间后5个工作日内完成系统维护，请耐心等待。</p> |

## 事件状态

系统上报的事件状态如表12-2所示，您可以根据状态判断系统上报事件的进展，也可以通过状态进行事件筛选。

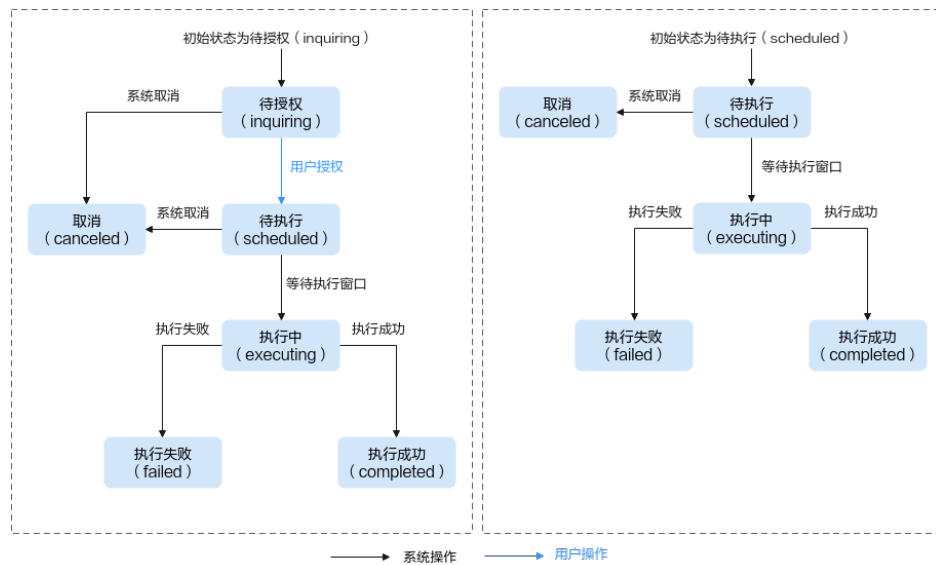
表 12-2 事件状态

| 类型   | 描述                                                                      |
|------|-------------------------------------------------------------------------|
| 待授权  | 需要用户对事件操作进行授权，授权时支持指定开始时间，系统将会在一定时间内完成操作，详细内容请参见 <a href="#">响应事件</a> 。 |
| 待执行  | 等待系统资源调度中。                                                              |
| 执行中  | 系统已完成资源调度，正在修复中。                                                        |
| 执行成功 | 系统已完成事件执行。请验证业务受损情况，如遇异常，请联系技术支持。                                       |
| 执行失败 | 系统自动修复失败。                                                               |
| 取消   | 该事件已被系统取消。                                                              |

事件状态会随着用户和系统的操作进行变化。



图 12-1 事件状态示意图



## 12.2 查询事件

### 操作场景

本文介绍如何通过弹性云服务器的控制台查询系统上报的事件。

您还可以通过云监控服务的“事件监控”查看ECS实例的事件，详细内容请参见[查看事件监控数据](#)。

### 前提条件



如果您通过IAM用户进行本操作，则需要在操作前对IAM用户进行授权。

查询及响应事件操作依赖以下策略：

- 查询事件：ecs:instanceScheduledEvents:list
- 接受并授权执行事件：ecs:instanceScheduledEvents:accept
- 修改预约时间：ecs:instanceScheduledEvents:update

为IAM用户授权的操作，请参见[ECS自定义策略](#)。

### 操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。

在“事件”页面的列表中，可以查看系统上报的事件ID、实例名称、事件类型等信息。

在列表的“操作”列，您可以进行事件的响应操作，详细内容请参见[响应事件](#)。

## 12.3 响应事件

### 12.3.1 响应实例重部署事件

#### 操作场景

当系统上报“实例重部署”事件时，您需要对ECS实例进行“授权重部署”操作。

本文介绍如何响应“待授权”状态的“实例重部署”事件，指导您完成实例重部署的授权操作。

授权重部署包括两种模式：

- 立即部署：即系统收到请求后，会立即进行实例重部署。
- 预约部署：即系统收到请求后，会在设置的重部署时间进行实例重部署。设置重部署时间后，您还可以修改预约时间，详细内容，请参见[修改预约时间](#)。

#### 须知

- 实例重部署时，会重启实例，请选择合适的时间授权并提前进行业务切流。
- 实例重部署后，实例的系统盘和云硬盘类型的数据盘不受影响。
- 对于包含本地盘的实例，实例重部署后，实例的本地盘数据会被清空。为了保障数据安全，请在授权重部署前进行本地盘数据备份。
- 对于包含本地盘的实例，授权重部署前需要对本地盘进行[预处理操作](#)，防止重部署后启动云服务器进入Emergency Mode（紧急模式）。

请根据ECS的实例类型，选择授权重部署操作指导：

- [授权重部署（仅包含云硬盘的实例）](#)
- [授权重部署（包含本地盘的实例）](#)

#### 前提条件


如果您通过IAM用户进行本操作，则需要在操作前对IAM用户进行授权。

查询及响应事件操作依赖以下策略：

- 查询事件：ecs:instanceScheduledEvents:list
- 接受并授权执行事件：ecs:instanceScheduledEvents:accept
- 修改预约时间：ecs:instanceScheduledEvents:update

为IAM用户授权的操作，请参见[ECS自定义策略](#)。

#### 授权重部署（仅包含云硬盘的实例）

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。

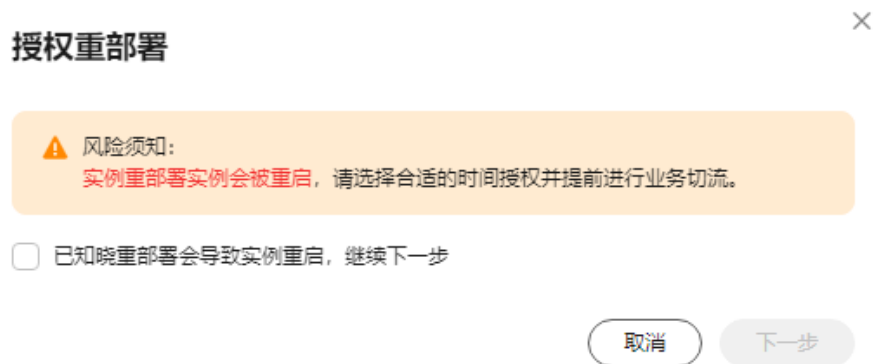
- 单击“☰”，选择“计算 > 弹性云服务器”。
- 在左侧导航树中，选择“事件”。

图 12-2 事件列表



- 在“事件”页面，单击“🔍”，通过以下条件过滤出待授权的“实例重部署”事件。
  - 选择“事件类型 > 实例重部署”。
  - 选择“事件状态 > 待授权”。
- 在待授权的事件“操作”列，单击“授权重部署”。


图 12-3 授权重部署



### 📖 说明

- 在进行“实例重部署”操作之前，请先了解风险须知。
- 了解风险须知后，勾选已知晓风险的提示，并单击“下一步”。
  - 进行“重部署设置”。
    - 选择“立即部署”。  
即系统收到请求后，会立即进行实例重部署。

图 12-4 立即部署



The screenshot shows a dialog box titled "授权重部署" (Authorization for Instance Redeployment) with a close button (X) in the top right corner. The main text asks: "确定并授权对云服务器ecs-进行实例重部署操作吗?" (Are you sure and authorize the instance redeployment operation for ECS instance?). Below this, there are two radio buttons under the label "重部署设置" (Redeployment Settings): "立即部署" (Immediate Deployment) is selected, and "预约部署" (Scheduled Deployment) is unselected. A text prompt says "同意授权重部署, 请输入 AGREE" (I agree to authorize instance redeployment, please enter AGREE). There is a text input field containing "AGREE". At the bottom right, there are two buttons: "取消" (Cancel) and "确定" (Confirm).

- 选择“预约部署”，并设置重部署时间。  
即系统在收到请求后，会在设置的预约部署时间进行实例重部署。

图 12-5 预约部署



The screenshot shows a dialog box titled "授权重部署" (Authorization for Instance Redeployment) with a close button (X) in the top right corner. The main text asks: "确定并授权对云服务器ecs-进行实例重部署操作吗?" (Are you sure and authorize the instance redeployment operation for ECS instance?). Below this, there are two radio buttons under the label "重部署设置" (Redeployment Settings): "立即部署" (Immediate Deployment) is unselected, and "预约部署" (Scheduled Deployment) is selected. A text prompt says "设置重部署时间" (Set redeployment time) followed by a date/time picker field containing "请选择日期时间" (Please select date and time). Below this, there is a text input field containing "AGREE". At the bottom right, there are two buttons: "取消" (Cancel) and "确定" (Confirm).

设置重部署时间后，您还可以修改预约时间，详细内容，请参见[修改预约时间](#)。

9. 输入“AGREE”确定授权实例重部署。
10. 单击“确定”，完成授权重部署。

授权重部署后，“实例重部署”事件更新为“待执行”状态。系统收到授权请求后，会根据授权设置将事件更新为执行中，开始进行实例重部署。

待“实例重部署”事件更新为“执行成功”状态，实例重部署完成，请检查云服务器业务运行情况。

## 授权重部署（包含本地盘的实例）


1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。

图 12-6 事件列表





5. 在“事件”页面，单击“🔍”，通过以下条件过滤出待授权的“实例重部署”事件。
  - a. 选择“事件类型 > 实例重部署”。
  - b. 选择“事件状态 > 待授权”。
6. 在待授权的事件“操作”列，单击“授权重部署”。

图 12-7 授权重部署

### 授权重部署

#### ⚠️ 风险须知：

1. 实例重部署实例会被重启，请选择合适的时间授权并提前进行业务切流。
2. 实例重部署会导致所有本地盘（）数据被清空，为了保障数据安全，请在计划事件授权前进行数据备份。
3. 授权重部署前需要对本地盘进行预处理操作，防止重部署后启动云服务器进入 Emergency Mode（紧急模式）。

- 已知晓重部署会导致实例重启
- 已知晓本地盘（）数据会被清空
- 已完成本地盘预处理，继续下一步

取消

下一步

#### 📖 说明

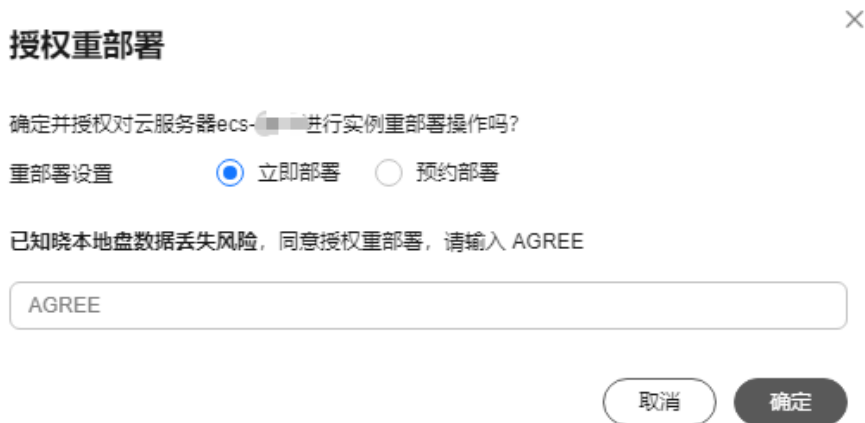
在进行“实例重部署”操作之前，请先了解风险须知，并完成本地盘预处理操作，为所有数据盘增加nofail参数。

详细内容，请参见[实例重部署预处理](#)。

7. 了解风险须知后，勾选已知晓风险的提示，并单击“下一步”。
8. 进行“重部署设置”。

- 选择“立即部署”。  
即系统收到请求后，会立即进行实例重部署。

图 12-8 立即部署



The screenshot shows a dialog box titled "授权重部署" (Authorize Re-deployment) with a close button (X) in the top right corner. The main text asks: "确定并授权对云服务器ecs-[redacted]进行实例重部署操作吗?" (Are you sure and authorize the instance re-deployment operation for ECS instance [redacted]?). Below this, there are two radio buttons: "重部署设置" (Re-deployment Settings), "立即部署" (Immediate Deployment) which is selected, and "预约部署" (Scheduled Deployment). A warning message states: "已知晓本地盘数据丢失风险，同意授权重部署，请输入 AGREE" (I am aware of the risk of local disk data loss, I agree to authorize re-deployment, please enter AGREE). There is a text input field containing "AGREE". At the bottom right, there are two buttons: "取消" (Cancel) and "确定" (Confirm).

- 选择“预约部署”，并设置重部署时间。  
即系统在收到请求后，会在设置的预约部署时间进行实例重部署。

图 12-9 预约部署



The screenshot shows a dialog box titled "授权重部署" (Authorize Re-deployment) with a close button (X) in the top right corner. The main text asks: "确定并授权对云服务器ecs-[redacted]进行实例重部署操作吗?" (Are you sure and authorize the instance re-deployment operation for ECS instance [redacted]?). Below this, there are two radio buttons: "重部署设置" (Re-deployment Settings), "立即部署" (Immediate Deployment), and "预约部署" (Scheduled Deployment) which is selected. A "设置重部署时间" (Set Re-deployment Time) field is present, showing the time "2024/05/21 10:43:44" and a calendar icon. A warning message states: "已知晓本地盘数据丢失风险，同意授权重部署，请输入 AGREE" (I am aware of the risk of local disk data loss, I agree to authorize re-deployment, please enter AGREE). There is a text input field containing "AGREE". At the bottom right, there are two buttons: "取消" (Cancel) and "确定" (Confirm).

设置重部署时间后，您还可以修改预约时间，详细内容，请参见[修改预约时间](#)。




9. 输入“AGREE”确定授权实例重部署。
10. 单击“确定”，完成授权重部署。

授权重部署后，“实例重部署”事件更新为“待执行”状态。系统收到授权请求后，会根据授权设置将事件更新为执行中，开始进行实例重部署。

待“实例重部署”事件更新为“执行成功”状态，实例重部署完成，请检查云服务器业务运行情况。

## 修改预约时间

完成“预约部署”方式的授权重部署操作后，“实例重部署”事件状态变为“待执行”，此时，您可以修改预约时间。

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。
5. 在“事件”页面，单击“”，通过以下条件过滤出待执行的“实例重部署”事件。
  - a. 选择“事件类型 > 实例重部署”。
  - b. 选择“事件状态 > 待执行”。
6. 在待执行的事件“操作”列，单击“修改预约时间”。
7. 设置“修改预约执行时间”，并单击“确定”，完成预约时间的修改。

### 说明

您需要在“可预约最晚时间”前至少30分钟修改预约时间，否则无法设置。

## 12.3.2 响应本地盘换盘事件

### 操作场景

当系统检测到ECS实例（含裸金属类型实例）的本地盘出现软硬件故障时，会对受影响的ECS实例自动生成“本地盘换盘”事件。

本文介绍响应“待授权”状态的“本地盘换盘”事件的方案：

- **授权换盘**：您可以选择通过“授权换盘”操作响应“本地盘换盘”事件，授权更换故障的本地盘。  
授权换盘包括两种模式：
  - 立即换盘：即系统收到换盘请求后，会在5个工作日内完成换盘操作。
  - 预约换盘：即系统收到换盘请求后，会在预约时间后5个工作日内完成换盘操作。  
设置预约换盘时间后，您还可以修改预约时间，详细内容，请参见[修改预约时间](#)。

### 须知

- 本地盘换盘期间，实例的系统盘、云硬盘类型的数据盘以及正常的本地盘不受影响。
  - 本地盘换盘期间，故障本地盘会被卸载并清空所有数据，为了保障数据安全，请提前进行本地盘数据备份。
  - 超高I/O型（I系列）的ECS，本地盘换盘期间，实例会被关机且不可用，关机后仍然继续计费。
  - 磁盘增强型（D系列）的ECS，本地盘换盘期间，实例不会关机，仅故障本地盘不可用。
  - 裸金属类型的ECS，本地盘换盘期间，可能伴随宿主机的下电与重启，在授权前，请确保云服务器中部署的业务离线，或云服务器停止对业务无影响。如果云服务器上的业务不可停止，请勿授权，并联系技术支持进行处理。
  - 授权换盘前需执行[换盘预处理](#)操作，避免因换盘引起的异常风险。
- 
- **重部署**：您可以选择通过“重部署”操作响应“本地盘换盘”事件，授权立即对实例进行重部署。

### 须知

- 实例重部署操作会重启实例，请选择合适的时间触发并提前进行业务切流。
- 实例重部署操作不会影响实例的系统盘和云硬盘类型的数据盘。
- 实例重部署会导致所有本地盘数据被清空，为了保障数据安全，请在重部署前进行数据备份。
- 实例重部署前需要对本地盘进行[预处理](#)操作，防止重部署后启动云服务器进入Emergency Mode（紧急模式）。

## 前提条件

如果您通过IAM用户进行本操作，则需要在操作前对IAM用户进行授权。

查询及响应事件操作依赖以下策略：

- 查询事件：ecs:instanceScheduledEvents:list
- 接受并授权执行事件：ecs:instanceScheduledEvents:accept
- 修改预约时间：ecs:instanceScheduledEvents:update

为IAM用户授权的操作，请参见[ECS自定义策略](#)。

## 授权换盘



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。



图 12-10 事件列表




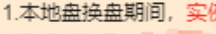


5. 在“事件”页面，单击“”，通过以下条件过滤出待授权的“本地盘换盘”事件。
  - a. 选择“事件类型 > 本地盘换盘”。
  - b. 选择“事件状态 > 待授权”。
6. 在待授权的事件“操作”列，单击“授权换盘”。

图 12-11 授权换盘（超高 I/O 型实例）

### 授权换盘

- 风险须知：**
- 1.本地盘换盘期间，实例会被关机且不可用。故障本地盘（Serial number: ）会被卸载，云盘、非故障本地盘不受影响。 [查询对应盘符](#)
  - 2.本地盘换盘期间，故障本地盘（Serial number: ）上的数据会被清空，为了保障数据安全，请提前进行数据备份。
  - 3.授权换盘前需执行[换盘预处理](#)操作，避免因换盘引起的异常风险。

- 已知晓换盘期间实例关机不可用
- 已知晓故障本地盘（Serial number: ）上的数据会被清空
- 已完成本地盘预处理，继续下一步

取消

下一步

图 12-12 授权换盘（磁盘增强型实例）

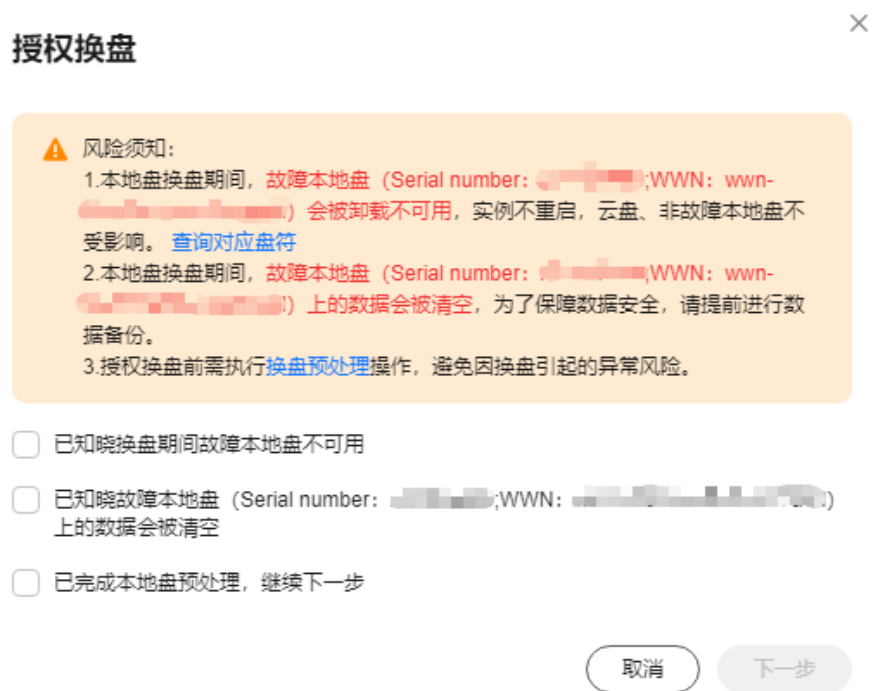


图 12-13 授权换盘（裸金属类型实例）



### 说明

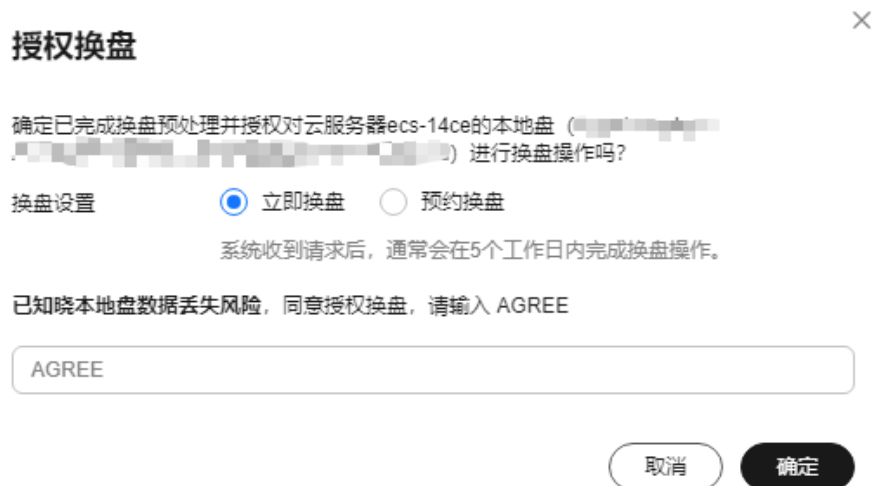
在进行“授权换盘”操作之前，请先了解风险须知，并完成换盘预处理操作：

- 如需保留本地盘上的数据，请先备份故障本地盘的数据。
- 在云服务器的/etc/fstab中，注释掉故障本地盘对应的挂载点，防止换盘后启动云服务器时进入维护模式。
- 在云服务器操作系统内部，使用umount命令，卸载故障本地盘对应的设备或者文件系统。

请根据实例类型，选择对应的预处理操作指导，详细内容，请参见[换盘预处理](#)。

7. 了解风险须知后，勾选已知晓风险的提示，并单击“下一步”。
8. 进行“授权设置”。
  - 选择“立即换盘”。  
即系统收到换盘请求后，会在5个工作日内完成换盘操作。

图 12-14 立即换盘



The screenshot shows a dialog box titled "授权换盘" (Authorization for Disk Replacement) with a close button (X) in the top right corner. The main text asks for confirmation to perform a disk replacement on a specific ECS instance. Under "换盘设置" (Disk Replacement Settings), the "立即换盘" (Immediate Replacement) radio button is selected, while "预约换盘" (Scheduled Replacement) is unselected. A note states that the operation will be completed within 5 working days. A risk warning is displayed, and a text input field contains the word "AGREE". At the bottom, there are "取消" (Cancel) and "确定" (Confirm) buttons.

- 选择“预约换盘”，并设置换盘时间。  
即系统收到换盘请求后，会在预约时间后5个工作日内完成换盘操作。

图 12-15 预约换盘



The screenshot shows the same "授权换盘" dialog box. In this view, the "预约换盘" (Scheduled Replacement) radio button is selected, and "立即换盘" is unselected. A "设置换盘时间" (Set Disk Replacement Time) section is visible, featuring a date-time picker with the placeholder text "请选择日期时间" (Please select date and time). The rest of the dialog, including the risk warning, the "AGREE" input field, and the "取消" (Cancel) and "确定" (Confirm) buttons, remains the same as in the previous figure.

设置预约换盘时间后，您还可以修改预约时间，详细内容，请参见[修改预约时间](#)。

9. 输入“AGREE”确定授权本地盘换盘。

10. 单击“确定”，完成授权换盘。

授权换盘后，“本地盘换盘”事件更新为“待执行”状态。系统收到换盘请求后，会根据授权设置将事件更新为执行中，开始进行换盘。

待“本地盘换盘”事件更新为“执行成功”状态，本地盘换盘完成，请检查云服务器业务运行情况。

## 重部署


1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“≡”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。

图 12-16 事件列表






5. 在“事件”页面，单击“”，通过以下条件过滤出待授权的“本地盘换盘”事件。
  - a. 选择“事件类型 > 本地盘换盘”。
  - b. 选择“事件状态 > 待授权”。
6. 在事件“操作”列，单击“重部署”。

图 12-17 重部署

## 重部署

- 风险须知：**
1. 实例重部署会重启实例，请选择合适的时间触发并提前进行业务切流。
  2. 实例重部署会导致所有本地盘（）数据被清空，为了保障数据安全，请在重部署前进行数据备份。
  3. 实例重部署前需要对本地盘进行预处理操作，防止重部署后启动云服务器进入 Emergency Mode（紧急模式）。

- 已知晓重部署会导致实例重启
- 已知晓本地盘（）数据会被清空
- 已完成本地盘预处理，继续下一步

取消

下一步

### 说明

在进行“重部署”操作之前，请先了解本地盘数据丢失风险，完成实例重部署预处理操作，为所有数据盘增加nofail参数。

详细内容，请参见[实例重部署预处理](#)。

- 了解风险须知后，勾选已知晓风险提示，并单击“下一步”。
- 输入“AGREE”确定进行实例重部署操作。

图 12-18 确定实例重部署

### 重部署

确定对云服务器ecs-进行实例重部署操作吗？  
重部署任务下发后，实例会进入重建中，重部署完成后实例状态恢复运行中，您可以在云服务器列表页查看重部署是否完成。

已知晓本地盘数据丢失风险，同意重部署，请输入 AGREE




取消

确定

- 单击“确定”。  
确定实例重部署后，“本地盘换盘”事件更新为“待执行”状态。系统收到重部署请求后，会将事件更新为执行中，开始进行实例重部署。  
待“本地盘换盘”事件更新为“执行成功”状态，实例重部署完成，请检查云服务器业务运行情况。

## 修改预约时间

完成“预约换盘”方式的授权换盘操作后，“本地盘换盘”事件状态变为“待执行”，此时，您可以修改预约换盘时间。

- 登录管理控制台。
- 单击管理控制台左上角的，选择区域和项目。
- 单击“”，选择“计算 > 弹性云服务器”。
- 在左侧导航树中，选择“事件”。
- 在“事件”页面，单击“”，通过以下条件过滤出待执行的“本地盘换盘”事件。
  - 选择“事件类型 > 本地盘换盘”。
  - 选择“事件状态 > 待执行”。
- 在待执行的事件“操作”列，单击“修改预约时间”。
- 设置“修改预约执行时间”，并单击“确定”，完成预约换盘时间的修改。

## 说明

您需要在“可预约最晚时间”前至少30分钟修改预约时间，否则无法设置。

### 12.3.3 响应系统维护事件

#### 操作场景

当系统上报“系统维护”事件时，您需要对ECS实例（含裸金属类型实例）进行“授权维护”操作。

本文介绍如何响应“待授权”状态的“系统维护”事件，指导您完成系统维护的授权操作。

授权维护包括两种模式：

- 立即授权维护：即系统收到请求后，会在5个工作日内完成实例宿主机的系统维护。
- 预约授权维护：即系统收到请求后，会在设置的预约维护时间后5个工作日内完成实例宿主机的系统维护。

设置预约维护时间后，您还可以修改预约时间，详细内容，请参见[修改预约时间](#)。

#### 须知

系统维护操作需要对主机进行下电，在此期间云服务器将停止无法工作。在授权维护前，请确保云服务器中部署的业务离线或云服务器停止对业务无影响。

如果云服务器上的业务不可停止，请勿授权，并联系技术支持进行处理。

#### 前提条件

如果您通过IAM用户进行本操作，则需要在操作前对IAM用户进行授权。

查询及响应事件操作依赖以下策略：

- 查询事件：ecs:instanceScheduledEvents:list
- 接受并授权执行事件：ecs:instanceScheduledEvents:accept
- 修改预约时间：ecs:instanceScheduledEvents:update

为IAM用户授权的操作，请参见[ECS自定义策略](#)。

#### 授权维护



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。

图 12-19 事件列表




5. 在“事件”页面，单击“”，通过以下条件过滤出待授权的“系统维护”事件。
  - a. 选择“事件类型 > 系统维护”。
  - b. 选择“事件状态 > 待授权”。
6. 在待授权的事件“操作”列，单击“授权维护”。

图 12-20 授权维护

### 授权维护

**风险须知：**  
系统维护操作需要对主机进行下电，期间实例将停止无法工作，因此授权前您需确保该实例的业务已经离线。


取消

下一步

7. 了解风险须知后，单击“下一步”。
8. 进行“维护设置”。
  - 选择“立即授权维护”。  
即系统收到请求后，会在5个工作日内完成实例宿主机的系统维护。

图 12-21 立即授权维护

### 授权维护

确定业务已离线并授权对云服务器ecs-进行系统维护操作吗？

维护设置  立即授权维护  预约授权维护

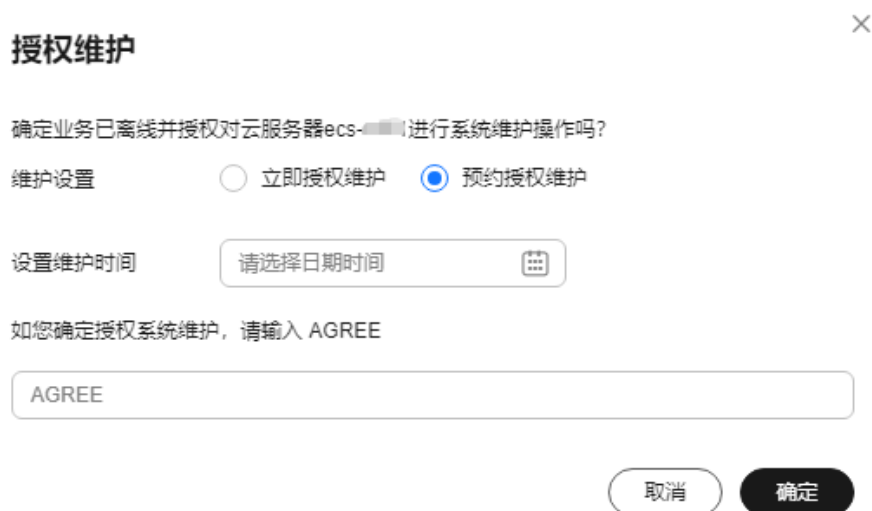
如您确定授权系统维护，请输入 AGREE

取消

确定

- 选择“预约授权维护”，并设置维护时间。  
即系统在收到请求后，会在设置的预约维护时间后5个工作日内完成系统维护操作。

图 12-22 预约授权维护



授权维护

确定业务已高线并授权对云服务器ecs-进行系统维护操作吗?

维护设置  立即授权维护  预约授权维护

设置维护时间

如您确定授权系统维护, 请输入 AGREE

设置预约维护时间后, 您还可以修改预约时间, 详细内容, 请参见[修改预约时间](#)。

9. 输入“AGREE”确定授权系统维护。
10. 单击“确定”, 完成授权维护。

授权维护后, “系统维护”事件更新为“待执行”状态。系统收到授权请求后, 会根据授权设置将事件更新为执行中, 开始进行系统维护。

待“系统维护”事件更新为“执行成功”状态, 系统维护完成, 请检查云服务器业务运行情况。

## 修改预约时间

完成“预约授权维护”方式的授权维护操作后, “系统维护”事件状态变为“待执行”, 此时, 您可以修改预约时间。




1. 登录管理控制台。
2. 单击管理控制台左上角的 , 选择区域和项目。
3. 单击“”, 选择“计算 > 弹性云服务器”。
4. 在左侧导航树中, 选择“事件”。
5. 在“事件”页面, 单击“”, 通过以下条件过滤出待执行的“系统维护”事件。
  - a. 选择“事件类型 > 系统维护”。
  - b. 选择“事件状态 > 待执行”。
6. 在待执行的事件“操作”列, 单击“修改预约时间”。



图 12-23 修改预约时间

×

### 修改预约时间

当前预约执行时间 2024/05/22 10:18:26 GMT+08:00

修改预约执行时间  📅

可预约最晚时间 2024/05/29 10:18:26 GMT+08:00

取消 确定

7. 设置“修改预约执行时间”，并单击“确定”，完成预约时间的修改。

#### 📖 说明

您需要在“可预约最晚时间”前至少30分钟修改预约时间，否则无法设置。

## 12.4 换盘预处理

### 12.4.1 本地盘换盘预处理（磁盘增强型实例）

#### 操作场景

当进行“本地盘换盘”事件的“授权换盘”操作时，需要先对本地盘进行换盘预处理。

本文介绍如何对磁盘增强型（D系列）的ECS实例进行本地盘的换盘预处理操作。


#### 须知

本地盘换盘操作会丢失故障本地盘上的数据，如果无需保留本地盘上的数据，可通过本操作换盘预处理。

如果需要保留本地盘上的数据，请结束操作并联系技术支持进行处理。

#### 操作步骤（Linux 实例）

##### 获取故障本地盘信息

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。

- 在事件列表中，查看“本地盘换盘”事件的“事件描述”列。通过“事件描述”信息获取故障本地盘的WWN号，本文以“wwn-0x5000c500e01a4930”为例进行介绍。

### 获取故障盘挂载点

- [远程登录](#)待换盘预处理的弹性云服务器。
- 执行以下命令，查看故障本地盘的WWN号对应的盘符信息。

```
ll /dev/disk/by-id/ | grep WWN号
```

以WWN号“wwn-0x5000c500e01a4930”为例，命令如下：

```
ll /dev/disk/by-id/ | grep wwn-0x5000c500e01a4930
```

回显信息如下：

```
lrwxrwxrwx 1 root root 9 May 13 14:05 wwn-0x5000c500e01a4930 -> ../../sdb
```

- 执行以下命令，查询故障本地盘的挂载点。

```
df -Th | grep 盘符信息
```

以盘符信息“/dev/sdb”为例，命令如下：

```
df -Th | grep /dev/sdb
```

回显信息如下：

```
/dev/sdb ext4 3.6T 28K 3.4T 1% /data
```

### 须知

本地盘的换盘预处理会导致本地盘数据丢失。如需保留本地盘数据，请结束操作并联系技术支持进行处理。

### 卸载故障的本地磁盘

- 执行以下命令，卸载故障本地盘。

```
umount 挂载点
```

以挂载点“/data”为例，命令如下：

```
umount /data
```

### 检查报障云服务器的/etc/fstab中是否配置了故障盘对应的自动挂载信息

#### 说明

如果报障云服务器的/etc/fstab中配置了故障盘对应的自动挂载信息，则需要删除或注释掉，防止本地盘换盘后启动云服务器进入Emergency Mode（紧急模式）。

- 执行以下命令，查询磁盘分区的UUID。

```
blkid 磁盘分区
```

以查询磁盘分区“/dev/sdb”的UUID为例，命令如下：

```
blkid /dev/sdb
```

回显信息如下：

```
/dev/sdb: UUID="626c4774-e60e-4d86-bbe6-031bac126e4c" TYPE="ext4"
```

- 执行以下命令，查询/etc/fstab中是否存在磁盘分区的自动挂载信息。

```
cat /etc/fstab | grep 磁盘分区的UUID
```

以查询磁盘分区的UUID “626c4774-e60e-4d86-bbe6-031bac126e4c” 为例，命令如下：

```
cat /etc/fstab | grep 626c4774-e60e-4d86-bbe6-031bac126e4c
```

回显信息如下，表示/etc/fstab中存在磁盘分区的自动挂载信息：

```
UUID=626c4774-e60e-4d86-bbe6-031bac126e4c /mnt ext4 defaults 0 0
```

3. 如果存在自动挂载信息，执行以下命令，删除磁盘分区的自动挂载信息，避免在更换本地盘后，云服务器开机时进入Emergency Mode（紧急模式）。

- a. 执行以下命令，编辑/etc/fstab文件。

```
vi /etc/fstab
```

- b. 单击“i”进入编辑模式。
- c. 删除或注释掉步骤2中查询的磁盘分区的自动挂载信息。



在自动挂载信息前面添加#，表示注释掉自动挂载信息，示例如下：

```
UUID=626c4774-e60e-4d86-bbe6-031bac126e4c /mnt ext4 defaults 0 0
```

- d. 单击Esc键并输入:wq后回车，保存并退出编辑模式。

## 操作步骤（Windows 实例）

### 获取故障盘标识符

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“

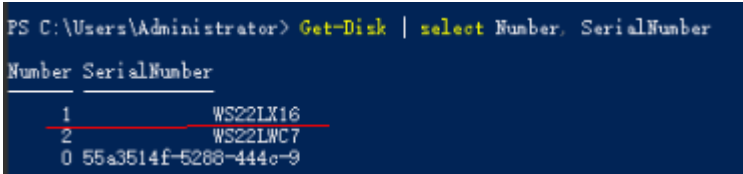
通过“事件描述”信息获取故障本地盘的SN号，本文以“WS22LX16”为例进行介绍。

### 获取磁盘编号，确认影响范围

1. [远程登录](#)待换盘预处理的弹性云服务器。
2. 以管理员身份打开“Windows PowerShell”，查询磁盘编号和SN号的对应关系，获取磁盘编号。

```
Get-Disk | select Number, SerialNumber
```

图 12-24 查询磁盘编号和 SN 号的对应关系



```
PS C:\Users\Administrator> Get-Disk | select Number, SerialNumber
Number SerialNumber

1 WS22LX16
2 WS22LWC7
0 55a3514f-5288-444e-9
```

### 须知

本地盘换盘操作会导致本地盘数据丢失。如需保留本地盘数据，请结束操作并联系技术支持进行处理。

## 12.4.2 本地盘换盘预处理（超高 I/O 型实例）

### 操作场景

当进行“本地盘换盘”事件的“授权换盘”操作时，需要先对本地盘进行换盘预处理。

本文介绍如何对超高I/O型（I系列）的ECS实例进行本地盘的换盘预处理操作。



#### 须知

本地盘换盘操作会丢失故障本地盘上的数据，如果无需保留本地盘上的数据，可通过本操作换盘预处理。

如果需要保留本地盘上的数据，请结束操作并联系技术支持进行处理。

### 操作步骤（Linux 实例）

#### 获取故障本地盘标识符

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。
5. 在事件列表中，查看“本地盘换盘”事件的“事件描述”列。

通过“事件描述”信息获取故障本地盘的SN号，本文以“1003004212241800EY”为例进行介绍。

#### 获取故障本地盘挂载点

1. [远程登录](#)待换盘预处理的弹性云服务器。
2. 执行以下命令，查看故障盘的SN号对应的盘符信息。

```
ll /dev/disk/by-id/ | grep SN号
```

以SN号“1003004212241800EY”为例，命令如下：

```
ll /dev/disk/by-id/ | grep 1003004212241800EY
```

回显信息如下：

```
lrwxrwxrwx 1 root root 13 Dec 14 18:55 nvme-RP2001T6RK004VX_1003004212241800EY -> ../../nvme0n1
```

3. 执行以下命令，查询故障盘的挂载点。

```
df -Th | grep 盘符信息
```

以盘符信息“/dev/nvme0n1”为例，命令如下：

```
df -Th | grep /dev/nvme0n1
```

回显信息如下：

```
/dev/nvme0n1 ext4 1.5T 77M 1.4T 1% /data
```

**须知**

本地盘的换盘预处理会导致本地盘数据丢失。如需保留本地盘数据，请结束操作并联系技术支持进行处理。

**卸载故障的本地磁盘**

1. 执行以下命令，卸载故障的本地磁盘。

**umount 挂载点**

以挂载点“/data”为例，命令如下：

**umount /data**

**检查报障云服务器的/etc/fstab中是否配置了故障盘对应的自动挂载信息****说明**

如果报障云服务器的/etc/fstab中配置了故障盘对应的自动挂载信息，则需要删除或注释掉，防止本地盘换盘后启动云服务器进入Emergency Mode（紧急模式）。

1. 执行以下命令，查询磁盘分区的UUID。

**blkid 磁盘分区**

以查询磁盘分区“/dev/nvme0n1”的UUID为例：

**blkid /dev/nvme0n1**

回显信息如下：

```
/dev/nvme0n1: UUID="11e39b6b-07fc-4675-8302-04c7ce7e30bd" TYPE="ext4"
```

2. 执行以下命令，查询/etc/fstab中是否存在磁盘分区的自动挂载信息。

**cat /etc/fstab | grep 磁盘分区的UUID**

以查询磁盘分区的UUID“11e39b6b-07fc-4675-8302-04c7ce7e30bd”为例，命令如下：

**cat /etc/fstab | grep 11e39b6b-07fc-4675-8302-04c7ce7e30bd**

回显信息如下，表示/etc/fstab中存在磁盘分区的自动挂载信息：

```
UUID=11e39b6b-07fc-4675-8302-04c7ce7e30bd /mnt ext4 defaults 0 0
```

3. 如果存在挂载信息，执行以下命令，删除磁盘分区的自动挂载信息，避免在更换本地盘后，云服务器开机时进入Emergency Mode（紧急模式）。

- a. 执行以下命令，编辑/etc/fstab文件。

**vi /etc/fstab**

- b. 单击“i”进入编辑模式。

- c. 删除或注释掉步骤2中查询的磁盘分区的自动挂载信息。

在自动挂载信息前面添加#，表示注释掉自动挂载信息，示例如下：

```
UUID=11e39b6b-07fc-4675-8302-04c7ce7e30bd /mnt ext4 defaults 0 0
```

- d. 单击Esc键并输入:wq后回车，保存并退出编辑模式。

**云服务器关机**


1. 在事件列表中，查看“本地盘换盘”事件对应的实例ID。
2. 在左侧导航选择“弹性云服务器”，在报障的云服务器“操作”列，单击“更多 > 关机”。

### 📖 说明

本地盘的换盘预处理需要通过将云服务器关机再开机的过程，将本地盘信息同步至虚拟化层。

## 操作步骤（Windows 实例）

### 获取故障本地盘标识符

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。
5. 在事件列表中，查看“本地盘换盘”事件的“事件描述”列。

通过“事件描述”信息获取故障本地盘的SN号，本文以“WS22LX16”为例进行介绍。

### 获取磁盘编号，确认影响范围

1. [远程登录](#)待换盘预处理的弹性云服务器。
2. 以管理员身份打开“Windows PowerShell”，查询磁盘编号和SN号的对应关系，获取磁盘编号。

```
Get-Disk | select Number, SerialNumber
```

图 12-25 查询磁盘编号和 SN 号的对应关系

```
PS C:\Users\Administrator> Get-Disk | select Number, SerialNumber
Number SerialNumber

1 WS22LX16
2 WS22LWC7
0 55a3514f-5288-444e-9
```

### 须知

本地盘换盘操作会导致本地盘数据丢失。如需保留本地盘数据，请结束操作并联系技术支持进行处理。

### 云服务器关机

1. 在事件列表中，查看“本地盘换盘”事件对应的实例ID。
2. 在左侧导航选择“弹性云服务器”，在报障的云服务器“操作”列，单击“更多 > 关机”。

### 📖 说明

本地盘的换盘预处理需要通过将云服务器关机再开机的过程，将本地盘信息同步至虚拟化层。

## 12.4.3 本地盘换盘预处理（裸金属类型实例）

### 操作场景

当进行“本地盘换盘”事件的“授权换盘”操作时，需要先对本地盘进行换盘预处理。

本文介绍如何对裸金属类型的ECS实例进行本地盘的换盘预处理操作。

#### 须知



本地盘换盘操作会丢失故障本地盘上的数据，如果无需保留本地盘上的数据，可通过本操作换盘预处理。

本地盘换盘操作可能伴随宿主机的下电与重启，在预处理前，请确保云服务器中部署的业务离线，或云服务器停止对业务无影响。

如果需要保留本地盘上的数据或如果云服务器上的业务不可停止，请结束操作并联系技术支持进行处理。

### 操作步骤（Linux 实例）

#### 获取故障本地盘标识符

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。
5. 在事件列表中，查看“本地盘换盘”事件的“事件描述”列。  
通过“事件描述”信息获取故障本地盘的SN号，本文以“99K0A046FL3G”为例进行介绍。

#### 获取故障本地盘挂载点

1. [远程登录](#)待换盘预处理的裸金属类型实例。
2. 执行以下命令，查看故障盘的SN号对应的盘符信息。

```
ll /dev/disk/by-id/ | grep SN号
```

以SN号“99K0A046FL3G”为例，命令如下：

```
ll /dev/disk/by-id/ | grep 99K0A046FL3G
```

回显信息如下：

```
/lrxxrxxrxx 1 root root 9 Sep 18 19: 20 ata-MG06ACA10TE_99K0A046FL3G ->././sda
lrwxrwxrwx 1 root root 10 Sep 18 19: 20 ata-MG06ACA10TE_99K0A046FL3G-part1->././sda1
```

3. 执行以下命令，查询故障盘的挂载点。

```
df -Th | grep 盘符信息
```

以盘符信息“/dev/sda1”为例，命令如下：

```
df -Th | grep /dev/sda1
```

回显信息如下：

```
/dev/sda1 ext4 9.1T 7.5T 1.6T 83% /srv/data
```



**须知**

本地盘的换盘预处理会导致本地盘数据丢失。如需保留本地盘数据，请结束操作并联系技术支持进行处理。

**卸载故障的本地磁盘**

1. 执行以下命令，卸载故障的本地磁盘。

**umount 挂载点**

以挂载点“/data”为例，命令如下：

**umount /data**

**检查报障云服务器的/etc/fstab中是否配置了故障盘对应的自动挂载信息****说明**

如果报障云服务器的/etc/fstab中配置了故障盘对应的自动挂载信息，则需要删除或注释掉，防止本地盘换盘后启动云服务器进入Emergency Mode（紧急模式）。

1. 执行以下命令，查询磁盘分区的UUID。

**blkid 磁盘分区**

以查询磁盘分区“/dev/sda1”的UUID为例：

**blkid /dev/sda1**

回显信息如下：

```
/dev/sda1: UUID="e7100f3e-af65-49da-a013-c4ace8e5aba7" TYPE
="ext4"RTLABEL="logical"PARTUUID="cd358d32-e02b-4b23-bbd9-8a8bdae0e070"
```

2. 执行以下命令，查询/etc/fstab中是否存在磁盘分区的自动挂载信息。

**cat /etc/fstab | grep 磁盘分区的UUID**

以查询磁盘分区的UUID“e7100f3e-af65-49da-a013-c4ace8e5aba7”为例，命令如下：

**cat /etc/fstab | grep e7100f3e-af65-49da-a013-c4ace8e5aba7**

回显信息如下，表示/etc/fstab中存在磁盘分区的自动挂载信息：

```
UUID=e7100f3e-af65-49da-a013-c4ace8e5aba7 /srv/data ext4 defaults,noatime, nodiratime 1 0
```

3. 如果存在挂载信息，执行以下命令，删除磁盘分区的自动挂载信息，避免在更换本地盘后，云服务器开机时进入Emergency Mode（紧急模式）。

- a. 执行以下命令，编辑/etc/fstab文件。

**vi /etc/fstab**

- b. 单击“i”进入编辑模式。

- c. 删除或注释掉步骤2中查询的磁盘分区的自动挂载信息。

在自动挂载信息前面添加#，表示注释掉自动挂载信息，示例如下：

```
UUID=e7100f3e-af65-49da-a013-c4ace8e5aba7/srv/data ext4 defaults,noatime, nodiratime 1
0
```

- d. 单击Esc键并输入:wq后回车，保存并退出编辑模式。

**云服务器关机**

1. 在事件列表中，查看“本地盘换盘”事件对应的实例ID。
2. 在左侧导航选择“弹性云服务器”，在报障的云服务器“操作”列，单击“更多 > 关机”。




### 📖 说明

本地盘的换盘预处理需要通过将云服务器关机再开机的过程，将本地盘信息同步至虚拟化层。

## 操作步骤（Windows 实例）

### 获取故障本地盘标识符

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在左侧导航树中，选择“事件”。
5. 在事件列表中，查看“本地盘换盘”事件的“事件描述”列。

通过“事件描述”信息获取故障本地盘的SN号，本文以“WS22LX16”为例进行介绍。

### 获取磁盘编号，确认影响范围

1. [远程登录](#)待换盘预处理的裸金属类型实例。
2. 以管理员身份打开“Windows PowerShell”，查询磁盘编号和SN号的对应关系，获取磁盘编号。

```
Get-Disk | select Number, SerialNumber
```

图 12-26 查询磁盘编号和 SN 号的对应关系

```
PS C:\Users\Administrator> Get-Disk | select Number, SerialNumber
Number SerialNumber

1 WS22LX16
2 WS22LWC7
0 55a3514f-5288-444e-9
```

### 须知

本地盘换盘操作会导致本地盘数据丢失。如需保留本地盘数据，请结束操作并联系技术支持进行处理。

### 云服务器关机

1. 在事件列表中，查看“本地盘换盘”事件对应的实例ID。
2. 在左侧导航选择“弹性云服务器”，在报障的云服务器“操作”列，单击“更多 > 关机”。

### 📖 说明

本地盘的换盘预处理需要通过将云服务器关机再开机的过程，将本地盘信息同步至虚拟化层。

## 12.5 实例重部署预处理

### 操作场景

对于包含本地盘的云服务器，当进行“本地盘换盘”和“实例重部署”事件的“授权重部署”操作时，需要先对本地盘进行实例重部署预处理，避免云服务器异常风险。

本文介绍如何对Linux操作系统的磁盘增强型（D系列）和超高I/O型（I系列）的ECS实例进行实例重部署的预处理操作。

#### 须知

对于包含本地盘的实例，实例重部署操作会丢失所有本地盘上的数据，如果无需保留本地盘上的数据，可通过本操作实例重部署预处理。

如果需要保留本地盘上的数据，请勿授权，结束操作并联系技术支持进行处理。

### 操作步骤

实例重部署预处理，即在实例的/etc/fstab配置文件中为所有数据盘增加nofail参数。

1. [远程登录](#)待预处理的弹性云服务器。
2. 执行以下命令，打开并编辑/etc/fstab配置文件。  
**vim /etc/fstab**
3. 单击“i”进入编辑模式，在/etc/fstab配置文件中为所有数据盘加入nofail参数。

示例如下：

```
UUID-8232fee7-f20a-416c-a2e0-cbc8c85a01a2 /mnt/nvme0n1 ext4 defaults,nofail 0 2
```

- UUID-8232fee7-f20a-416c-a2e0-cbc8c85a01a2：数据盘的UUID。
  - /mnt/nvme0n1：数据盘挂载点，可通过`mount | grep "盘符信息"`命令进行查询。
  - ext4：数据盘的文件系统类型，可通过`blkid 盘符信息`命令进行查询。
  - nofail：用于设置当文件系统中包含本地盘但实际缺失时，不中断云服务器的启动流程，避免云服务器异常风险。
4. 单击:wq退出编辑模式。
  5. 执行以下命令，使配置生效。

```
systemctl daemon-reload
```

# 13 操作系统 dump 管理

## 13.1 配置操作系统 dump

### 操作场景

云服务器的操作系统在运行过程中如果出现故障，可能导致系统崩溃，例如操作系统蓝屏或死机。在这种情况下，如果使用操作系统dump，系统会将内存中的数据以存储文件的形式进行保存，以便于进行云服务器的故障定位和问题分析。

本文介绍如何配置操作系统dump的方法。

- [配置操作系统dump（Windows）](#)
- [配置操作系统dump（Linux）](#)

### 背景知识

操作系统dump，即内核转储或崩溃转储，是当云服务器系统运行出现异常时，对操作系统内存数据的实时完整备份。

操作系统dump功能主要利用操作系统内核对硬件异常的处理能力。操作系统内核可以监测到硬件异常，并根据异常的类型来决定处理方式。当系统遇到无法处理的异常或错误时，系统的部分或全部运行状态信息将会转储到磁盘上，这些状态信息包括：CPU的寄存器、物理内存、进程状态、文件系统状态、硬件设备状态等。

操作系统dump可以由以下几种方式触发：

- 系统崩溃：当操作系统遇到无法处理的异常或错误时，例如非法的内存访问、内核panic等，系统会自动触发dump。
- 手动触发：系统管理员也可以通过一些命令或者操作手动触发系统dump，例如在Linux系统中可以通过向/proc/sysrq-trigger写入c字符来触发系统dump。
- 计划触发：某些情况下，系统管理员为了获取系统的某一时刻的运行状态信息，或者模拟系统崩溃的情况，也可以设置在某一特定的时间触发系统dump。

在多数情况下，为了准确获得系统dump，需要对系统进行相关的配置，例如在Linux系统中，需要配置kdump服务等。本文主要介绍手动触发dump的操作指导。

## 约束与限制

- 执行操作系统dump会重启云服务器，请提前做好业务备份，避免数据丢失。
- 鲲鹏架构的云服务器不支持操作系统dump。
- 支持运行中、关机和故障状态的云服务器进行操作系统dump。
- 仅支持对4vCPUs，8GiB内存以上规格的ECS使用操作系统dump功能。
- 预留内存需大于ECS规格的内存。

## 配置操作系统 dump（Windows）

本文以Windows Server 2016版本操作系统为例进行介绍，更多版本的操作，请参考对应操作系统官网帮助文档。

1. [远程登录云服务器](#)。
2. 在“控制面板中，打开“系统 > 高级系统配置”。

图 13-1 高级系统配置



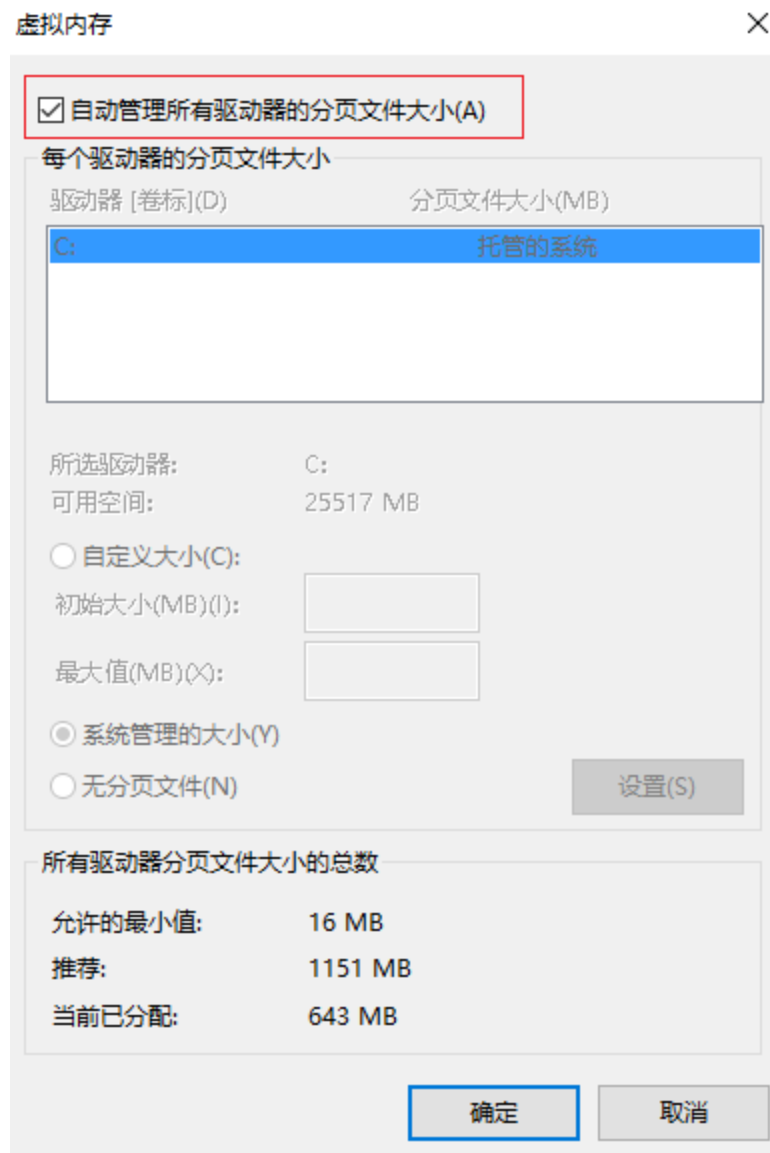
3. 在“系统属性”的“性能”区域，单击“设置”。
4. 在“高级”页签的“虚拟内存”区域，单击“更改”。

图 13-2 性能选项设置



5. 在“虚拟内存”页面，勾选“自动管理所有驱动器的分页文件大小”。

图 13-3 虚拟内存设置



6. 单击“确定”，完成虚拟内存设置。
7. 在“系统属性”的“启动和故障恢复”区域，单击“设置”。
8. 设置参数。
  - 写入调试信息：设置为“完全内存转储”。
  - 转储文件：可以将路径从“%SystemRoot%\Memory.dmp”更改为指向具有足够磁盘空间的本地驱动器，例如“E: |Memory.dmp”。

图 13-4 启动和故障恢复设置



9. 单击“确定”。
10. 重启云服务器，完成配置。

## 配置操作系统 dump (Linux)

本文以CentOS 7.5版本操作系统为例进行介绍，更多版本的操作，请参考对应操作系统官网帮助文档。

1. [远程登录云服务器](#)。
2. 执行以下命令，安装kexec-tools。  
**yum install kexec-tools**
3. 执行以下命令，查看crashkernel预留内存大小。  
**cat /etc/default/grub**



```
GRUB_TIMEOUT=5
GRUB_DISTRIBUTOR="$(sed 's, release .*$,g' /etc/system-release)"
GRUB_DEFAULT=saved
GRUB_DISABLE_SUBMENU=true
GRUB_TERMINAL_OUTPUT="console"
GRUB_CMDLINE_LINUX="net.ifnames=0 consoleblank=600 console=tty0 console=ttyS0,115200n8
spectre_v2=off nopti crashkernel=auto "
GRUB_DISABLE_RECOVERY="true"
```

4. 执行以下命令，设置crashkernel的大小。

#### **vim /etc/default/grub**

当前crashkernel预留内存大小为“auto”，您可以在“GRUB\_CMDLINE\_LINUX”字段，修改“**crashkernel**”的值，设置crashkernel的大小。

```
GRUB_TIMEOUT=5
GRUB_DISTRIBUTOR="$(sed 's, release .*$,g' /etc/system-release)"
GRUB_DEFAULT=saved
GRUB_DISABLE_SUBMENU=true
GRUB_TERMINAL_OUTPUT="console"
GRUB_CMDLINE_LINUX="net.ifnames=0 consoleblank=600 console=tty0 console=ttyS0,115200n8
spectre_v2=off nopti crashkernel=256M"
GRUB_DISABLE_RECOVERY="true"
```

#### 说明

您可以根据系统架构决定crashkernel预留内存大小，过小会导致生成coredump文件失败。

当前示例的系统内存是1GiB，预留内存设置为256M。

5. 执行以下命令，重新生成grub配置文件。

#### **grub2-mkconfig -o /boot/grub2/grub.cfg**

6. 执行以下命令，编辑内核参数文件/etc/sysctl.conf。

#### **vim /etc/sysctl.conf**

在/etc/sysctl.conf配置文件中增加以下参数，以便于在收到NMI时将操作系统切换到应急模式，并生成内存转储。

```
kernel.panic_on_unrecovered_nmi=1
kernel.unknown_nmi_panic=1
```

7. 重启系统。

#### **reboot**

8. 执行以下命令，修改kdump默认配置/etc/kdump.conf。

#### **vim /etc/kdump.conf**

```
#指定coredump文件存储位置
path /var/crash
#增加-c参数，代表压缩coredump文件
core_collector makedumpfile -c -l --message-level 1 -d 31
#生成coredump后，重启系统
default reboot
```

9. 执行以下命令，开启kdump服务并设置开机自启动。

#### **systemctl start kdump.service**

#### **systemctl enable kdump.service**

10. 执行以下命令，检查kdump是否开启成功。

#### **service kdump status**



```
Redirecting to /bin/systemctl status kdump.service
● kdump.service - Crash recovery kernel arming
 Loaded: loaded (/usr/lib/systemd/system/kdump.service; enabled; vendor preset: enabled)
 Active: active (exited) since Wed 2024-05-22 21:19:47 CST; 11min ago
```

```
Main PID: 591 (code=exited, status=0/SUCCESS)
CGroup: /system.slice/kdump.service
```

```
systemctl is-active kdump.service
```

```
active
```

## 触发操作系统 dump

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表的“操作”列，单击“更多 > 故障排查 > 触发操作系统 dump”。
5. 了解风险须知后，勾选已知晓风险的提示。
6. 单击“确定”，完成触发操作系统dump。

## 13.2 查看操作系统 dump

### 操作场景

触发操作系统dump后，您可以在云服务器上查看对应的内存dump文件。



本文介绍如何触发和查看操作系统dump的内核转储文件的方法。

- [触发操作系统dump](#)
- [查看操作系统dump \( Windows \)](#)
- [查看操作系统dump \( Linux \)](#)

### 前提条件

已经完成[配置操作系统dump](#)。

### 触发操作系统 dump

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表的“操作”列，单击“更多 > 故障排查 > 触发操作系统 dump”。
5. 了解风险须知后，勾选已知晓风险的提示。
6. 单击“确定”，完成触发操作系统dump。

### 查看操作系统 dump ( Windows )

本文以Windows Server 2016版本操作系统为例进行介绍，更多版本的操作，请参考对应操作系统官网帮助文档。

1. [远程登录云服务器](#)。
2. 根据[配置操作系统dump \( Windows \)](#)中设置的转储文件路径，可以查看到操作系统dump的内核转储文件。  
例如，在“C:\Windows”路径下，获取到文件“memory.dmp”。  
您可以通过工具检查和分析内核转储文件，详细操作，请参见[分析内核模式转储文件](#)。

## 查看操作系统 dump ( Linux )

本文以CentOS 7.5版本操作系统为例进行介绍，更多版本的操作，请参考对应操作系统官网帮助文档。

1. [远程登录云服务器](#)。
2. 执行以下命令，查看是否生成操作系统dump的内核转储文件。

```
ls /var/crash/
```

```
127.0.0.1-2024-05-22-21:35:26
```

您可以通过Linux平台的crash工具来分析操作系统内核的转储文件，详细操作，请参见[crash](#)。

# 14 自助运维管理

## 14.1 配置 ECS 自助运维自定义策略

### 操作场景

ECS自助运维功能依赖云运维中心（Cloud Operations Center，简称COC），需开通并授权COC服务。开通并授权后，系统自动创建名为ServiceLinkedAgencyForCOC和ServiceAgencyForCOC的委托，并授予必要的权限，用于通过COC代理操作对应的云服务器资源。详细内容，请参见[查看委托](#)。

当您想要使用IAM用户进行ECS自助运维操作时，需联系IAM用户所属华为账号为该IAM用户授予COC的资源操作类权限。

本文为您介绍为IAM用户授权的方法。

### 前提条件

本文默认IAM用户已具备ECS相关操作权限。

#### 说明

若操作过程中出现ECS权限相关问题，请为IAM用户授权，详细操作，请参见[创建用户并授权使用ECS](#)。

### 操作步骤

1. 使用IAM用户所属华为账号登录[IAM权限管理](#)页面。
2. 在页面右上角，单击“创建自定义策略”，并增加COC自定义权限策略。  
详细操作步骤，请参见[创建自定义策略](#)。

以添加开通COC以及操作COC的自定义策略为例，进行介绍

- a. （可选）创建名为“开通COC”的自定义策略，并增加如下权限。

#### 说明

如果IAM用户所属华为账号已开通COC，则无需为IAM用户添加“开通COC”的权限。

```
{
 "Version": "1.1",
```

```
"Statement": [
 {
 "Effect": "Allow",
 "Action": [
 "iam:agencies:list*",
 "iam:agencies:createAgency",
 "iam:agencies:createServiceLinkedAgencyV5",
 "coc:agency:get",
 "coc:agency:create",
 "iam:permissions:grantRoleToAgency",
 "iam:permissions:grantRoleToAgencyOnDomain",
 "iam:roles:listRoles"
]
 }
]
```

- b. 创建名为“COC操作权限”的自定义策略，并增加如下权限。

```
{
 "Version": "1.1",
 "Statement": [
 {
 "Effect": "Allow",
 "Action": [
 "coc:instance:listResources",
 "coc:application:listResources",
 "coc:schedule:list",
 "coc:schedule:enable",
 "coc:schedule:update",
 "coc:schedule:disable",
 "coc:schedule:approve",
 "coc:schedule:create",
 "coc:schedule:delete",
 "coc:schedule:count",
 "coc:schedule:get",
 "coc:schedule:getHistories",
 "coc:application:GetDiagnosisTaskDetails",
 "coc:application:CreateDiagnosisTask",
 "coc:document:create",
 "coc:document:listRunbookAtomics",
 "coc:document:getRunbookAtomicDetails",
 "coc:document:list",
 "coc:document:delete",
 "coc:document:update",
 "coc:document:get",
 "coc:document:analyzeRisk",
 "coc:instance:autoBatchInstances",
 "coc:instance:executeDocument",
 "coc:instance:start",
 "coc:instance:reboot",
 "coc:instance:stop",
 "coc:job:action",
 "coc:instance:reinstallOS",
 "coc:instance:changeOS",
 "coc:hostAccount:describe",
 "coc:instance:syncResources"
]
 }
]
}
```

3. 创建用户组并授予COC的资源操作类权限。

详细操作步骤，请参见[创建用户组并授权](#)。

需要为用户组授予如下权限：

- COC ReadOnlyAccess
- 步骤2中创建的“开通COC”以及“COC操作权限”自定义策略。

图 14-1 为用户组授权



4. 在IAM控制台的“用户”页面，单击待授权的IAM用户“操作”列的“授权”，为IAM用户授予COC权限。

详细操作步骤，请参见[给IAM用户授权](#)。

在“授权”页面，将待授权的IAM用户加入到步骤3创建的用户组，完成IAM用户的授权。

## 14.2 为 ECS 配置定时运维任务

### 操作场景

ECS支持定时运维，您可以通过配置定时运维任务，对ECS进行定时开机、定时关机等操作。该功能适用于需要在特定时间对多台ECS进行开机、关机等运维操作的场景。

例如，若您的业务有比较规律的波峰和波谷，可以选择在业务波峰时开机、波谷时关机，以便于降低使用成本。如果采用手动方式对多台ECS进行开机、关机操作，耗时长且容易出错，此时，可以通过设置定时运维任务，自动化的进行多台ECS的定时运维操作，节省人力和时间成本。

本文介绍如何设置ECS的定时运维任务。

### 约束与限制

- ECS的定时运维依赖云运维中心（Cloud Operations Center，简称COC），需开通并授权COC服务。  
对于IAM用户，需额外配置COC资源操作类权限，详细内容，请参见[配置ECS自助运维自定义策略](#)。
- 单个租户最多支持创建100个定时任务。
- 定时运维任务仅支持应用于同一区域的目标ECS实例。

### 操作步骤

您可以为单台或多台ECS创建定时开机、定时关机等运维任务。

#### 为单台 ECS 创建定时任务

1. 登录管理控制台，进入[弹性云服务器](#)列表页面。
2. 在待创建定时任务ECS的“操作”列，单击“更多 > 定时任务 > 定时开机”或“更多 > 定时任务 > 定时关机”。
3. （可选）在“开通云运维中心并添加权限”页面，阅读服务声明并勾选后，单击“开通并授权”。

若当前账号未开通并授权COC服务，则会显示该页面。

4. 在“创建定时任务”页面，根据界面提示信息，完成相关参数配置。

表 14-1 定时任务参数配置

| 参数      | 说明                                                                                                                                                                                                                                                                              |
|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 任务名称    | 定时任务的名称。<br>由中文、字母、数字、中划线、下划线组成，长度在3~100之间。                                                                                                                                                                                                                                     |
| 企业项目    | 定时任务所属的企业项目，用于提供项目内的资源管理、成员管理。<br>默认项目为default。                                                                                                                                                                                                                                 |
| 时区      | 执行定时任务的时间所属时区。                                                                                                                                                                                                                                                                  |
| 定时类型    | 定时任务的执行类型，包括： <ul style="list-style-type: none"><li>● 单次执行：在指定时间执行一次定时任务。</li><li>● 周期执行：在指定的周期和时间循环执行定时任务。</li></ul>                                                                                                                                                           |
| 执行时间    | 定时任务的执行时间。 <ul style="list-style-type: none"><li>● 当“定时类型”选择“单次执行”时，设置定时任务的开始执行时间。</li><li>● 当“定时类型”选择“周期执行”时，设置定时任务的执行周期和执行时间。<ul style="list-style-type: none"><li>- 简单周期：设置每周自动执行定时任务的具体日期和时间。</li><li>- Cron表达式：通过设置秒、分、时、日、月、周的Cron表达式，设置定时任务的执行周期和时间。</li></ul></li></ul> |
| 规则结束时间  | 周期执行的定时任务的结束时间。<br>仅当“定时类型”选择“周期执行”时，会显示该参数。                                                                                                                                                                                                                                    |
| 任务类型    | 定时任务的类型，包括： <ul style="list-style-type: none"><li>● 当步骤2选择“定时开机”时，此参数默认为“ECS开机”，用于设置ECS定时开机任务。</li><li>● 当步骤2选择“定时关机”时，此参数默认为“ECS关机”，用于设置ECS定时关机任务。</li></ul>                                                                                                                   |
| IAM权限委托 | 执行定时任务所需的IAM委托权限，以便于通过COC代理操作对应的ECS资源。                                                                                                                                                                                                                                          |
| 目标实例    | 待执行定时任务的目标实例。默认已选择一台目标实例。                                                                                                                                                                                                                                                       |

| 参数   | 说明                                                                                                                                                                                                       |
|------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 分批策略 | 目标实例执行定时任务时的分批策略，包括： <ul style="list-style-type: none"> <li>• 自动分批：系统自动设置目标实例的分批批次，按不同批次执行定时任务。</li> <li>• 手动分批：手动设置目标实例的分批批次，按不同批次执行定时任务。</li> <li>• 不分批：目标实例不分批，同时执行定时任务。</li> </ul> 每批次中目标实例数的上限为100。 |

5. 单击“提交”，完成定时任务的创建。

## 为多台 ECS 创建定时任务

1. 登录管理控制台，进入[弹性云服务器](#)页面。
2. 在ECS列表中，勾选待创建定时任务的多台ECS。
3. 在列表上方，单击“更多 > 定时任务 > 定时开机”或“更多 > 定时任务 > 定时关机”。
4. （可选）在“开通云运维中心并添加权限”页面，阅读服务声明并勾选后，单击“开通并授权”。

若当前账号未开通并授权COC服务，则会显示该页面。

5. 在“创建定时任务”页面，根据界面提示信息，完成相关参数配置。

表 14-2 定时任务参数配置

| 参数   | 说明                                                                                                                       |
|------|--------------------------------------------------------------------------------------------------------------------------|
| 任务名称 | 定时任务的名称。<br>由中文、字母、数字、中划线、下划线组成，长度在3~100之间。                                                                              |
| 企业项目 | 定时任务所属的企业项目，用于提供项目内的资源管理、成员管理。<br>默认项目为default。                                                                          |
| 时区   | 执行定时任务的时间所属时区。                                                                                                           |
| 定时类型 | 定时任务的执行类型，包括： <ul style="list-style-type: none"> <li>• 单次执行：在指定时间执行一次定时任务。</li> <li>• 周期执行：在指定的周期和时间循环执行定时任务。</li> </ul> |



| 参数      | 说明                                                                                                                                                                                                                                                                      |
|---------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 执行时间    | 定时任务的执行时间。 <ul style="list-style-type: none"><li>当“定时类型”选择“单次执行”时，设置定时任务的开始执行时间。</li><li>当“定时类型”选择“周期执行”时，设置定时任务的执行周期和执行时间。<ul style="list-style-type: none"><li>简单周期：设置每周自动执行定时任务的具体日期和时间。</li><li>Cron表达式：通过设置秒、分、时、日、月、周的Cron表达式，设置定时任务的执行周期和时间。</li></ul></li></ul> |
| 规则结束时间  | 周期执行的定时任务的结束时间。<br>仅当“定时类型”选择“周期执行”时，会显示该参数。                                                                                                                                                                                                                            |
| 任务类型    | 定时任务的类型，包括： <ul style="list-style-type: none"><li>当步骤3选择“定时开机”时，此参数默认为“ECS开机”，用于设置ECS定时开机任务。</li><li>当步骤3选择“定时关机”时，此参数默认为“ECS关机”，用于设置ECS定时关机任务。</li></ul>                                                                                                               |
| IAM权限委托 | 执行定时任务所需的IAM委托权限，以便于通过COC代理操作对应的ECS资源。                                                                                                                                                                                                                                  |
| 目标实例    | 待执行定时任务的目标实例。默认已选择步骤2的目标实例。                                                                                                                                                                                                                                             |
| 分批策略    | 目标实例执行定时任务时的分批策略，包括： <ul style="list-style-type: none"><li>自动分批：系统自动设置目标实例的分批批次，按不同批次执行定时任务。</li><li>手动分批：手动设置目标实例的分批批次，按不同批次执行定时任务。</li><li>不分批：目标实例不分批，同时执行定时任务。</li></ul> 每批次中目标实例数的上限为100。                                                                          |

6. 单击“提交”，完成定时任务的创建。

## 后续操作

定时任务创建成功后，您可以通过[定时运维](#)页面启用、禁用、编辑、删除定时任务，还可以设置定时任务的审批配置和消息通知。

详细内容，请参见[定时任务管理](#)。

# 15 资源与标签

## 15.1 标签管理

### 15.1.1 标签概述

#### 操作场景

标签是弹性云服务器的标识。为弹性云服务器添加标签，可以方便用户识别和管理拥有的弹性云服务器资源。

您可以在创建弹性云服务器时添加标签，也可以在弹性云服务器创建完成后，在云服务器的详情页添加标签，您最多可以给弹性云服务器添加10个标签。

#### 说明

创建弹性云服务器时添加的标签，将同步添加至一同创建的EIP和云硬盘（包括系统盘、数据盘）上。如果云服务器使用的是已有EIP，则该标签不会在EIP上标识。

如您的组织已经设定弹性云服务器的相关标签策略，则需按照标签策略规则为弹性云服务器添加标签。如果添加的标签不符合标签策略规则，则可能会导致弹性云服务器创建失败，请联系组织管理员了解标签策略详情。

云服务器创建成功后，您可以在弹性云服务器、EIP和云硬盘详情页，查看到对应的标签。

#### 有关标签的基本知识

标签用于标识资源，当您拥有相同类型的许多云资源时，可以使用标签按各种维度（例如用途、所有者或环境）对云资源进行分类。

图 15-1 标签示例

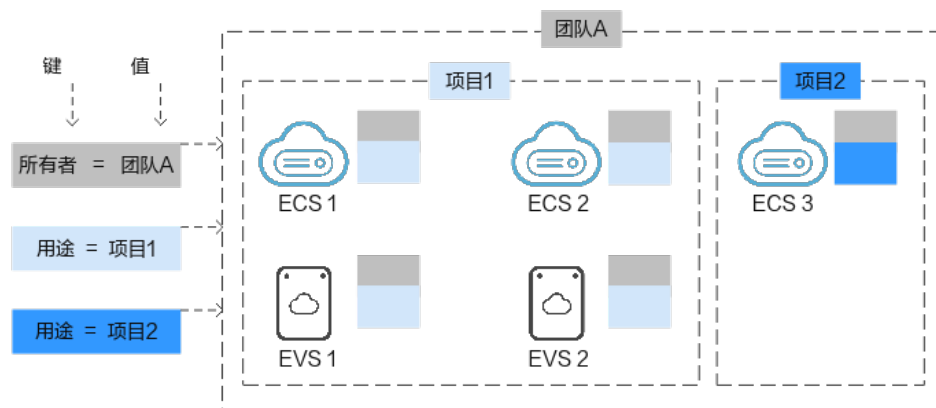


图15-1说明了标签的工作方式。在此示例中，您为每个云资源分配了两个标签，每个标签都包含您定义的一个“键”和一个“值”，一个标签使用键为“所有者”，另一个使用键为“用途”，每个标签都拥有相关的值。

您可以根据为云资源添加的标签快速搜索和筛选特定的云资源。例如，您可以为账户中的资源定义一组标签，以跟踪每个云资源的所有者和用途，使资源管理变得更加轻松。

## 标签命名规则

- 每个标签由一对键值对（Key-Value）组成。
- 每个弹性云服务器最多可以添加10个标签。
- 对于每个资源，每个标签键（Key）都必须是唯一的，每个标签键（Key）只能有一个值（Value）。
- 标签共由两部分组成：“标签键”和“标签值”，其中，“标签键”和“标签值”的命名规则如表15-1所示。

表 15-1 标签命名规则

| 参数  | 规则                                                                                                        | 样例           |
|-----|-----------------------------------------------------------------------------------------------------------|--------------|
| 标签键 | <ul style="list-style-type: none"><li>• 不能为空。</li><li>• 对于同一台云服务器，Key值唯一。</li><li>• 长度不超过36个字符。</li></ul> | Organization |
| 标签值 | <ul style="list-style-type: none"><li>• 长度不超过43个字符。</li></ul>                                             | Apache       |

### 15.1.2 添加标签

标签用于标记云资源，如实例、镜像和磁盘等。如果您的账户下有多种云资源，并且不同云资源之间有多种关联，您可以为云资源添加标签，实现云资源的分类和统一管理。更多信息，请参见[标签概述](#)。

有三种途径为弹性云服务器添加标签：

- [在创建弹性云服务器时添加标签](#)



- [在弹性云服务器详情页添加标签](#)
- [在标签管理页面添加标签](#)

预定义标签的使方法请参考[预定义标签的使用方法](#)。

## 约束与限制

如您的组织已经设定云资源的相关标签策略，则需按照标签策略规则为云资源添加标签。如果添加的标签不符合标签策略规则，则可能会导致云资源创建失败，请联系组织管理员了解标签策略详情。

## 在创建弹性云服务器时添加标签

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”，进入弹性云服务器页面。
4. 单击“购买弹性云服务器”。
5. 配置弹性云服务器的规格参数。

在“高级配置 > 标签”中，添加标签键和标签值。标签命名规则如[表15-1](#)所示。

### 说明

其他参数说明请参考[创建弹性云服务器](#)。

图 15-2 添加标签



## 在弹性云服务器详情页添加标签



1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“”，选择“计算 > 弹性云服务器”，进入弹性云服务器页面。
4. 在弹性云服务器列表中，单击待管理标签的弹性云服务器名称。  
系统跳转至该弹性云服务器详情页面。
5. 选择“标签”页签，单击“添加标签”，在弹出的“添加标签”窗口，输入标签的键和值。标签命名规则如[表15-1](#)所示。

图 15-3 在云服务器详情页添加标签



- 单击“确定”，完成标签的添加。  
标签添加成功后，您可以单击标签“操作”列的“编辑”修改标签值。

## 在标签管理页面添加标签

### 说明

此方法适用于为多个资源统一添加拥有同样标签键的标签。

- 登录管理控制台。
- 在右上角的用户名下选择“标签管理”，进入标签管理服务页面。

图 15-4 标签管理




- 在“资源标签”页面，勾选资源所在的区域，“资源类型”请选择“ECS-ECS”，单击“搜索”。  
系统列出所选区域下的所有云服务器资源。
- 在“搜索结果”区域，单击“创建标签键”，输入键（例如：项目），单击“确定”。  
创建完成后，该标签键会添加至资源标签列表。如果列表中没有显示该标签，单击 ，在下拉列表中勾选创建的标签键。  
默认该标签键的值为“无标签”，您还需要为每一个资源对应的标签值赋值，完成标签与资源的关联，继续下一步。

图 15-5 资源列表

搜索结果 上次更新: 2019/08/01 14:08:33 GMT+08:00 (0分钟前) API非法标签

| 资源类型    | 资源名称                           | 区域(项目)            | 标签数量 | 项目  |
|---------|--------------------------------|-------------------|------|-----|
| ECS-ECS | 03215acd-eb0c-49a9-b26c-f2c... | 华北北京一(cn-north-1) | 0    | 无标签 |
| ECS-ECS | 03215acd-eb0c-49a9-b26c-f2c... | 华北北京一(cn-north-1) | 0    | 无标签 |
| ECS-ECS | 03215acd-eb0c-49a9-b26c-f2c... | 华北北京一(cn-north-1) | 0    | 无标签 |
| ECS-ECS | 03215acd-eb0c-49a9-b26c-f2c... | 华北北京一(cn-north-1) | 0    | 无标签 |


5. 单击“编辑”，切换资源标签列表为可编辑状态。
6. 在云服务器资源所在行，单击 ，输入标签“值”（例如：A）。为标签键赋值后，“标签数量”将加1。按照同样方法依次为其他云服务器添加标签值。

图 15-6 输入标签值

搜索结果 上次更新: 2019/08/01 14:08:33 GMT+08:00 (4分钟前)

| 资源类型    | 资源名称                           | 区域(项目)            | 标签数量 | 项目  |
|---------|--------------------------------|-------------------|------|-----|
| ECS-ECS | 03215acd-eb0c-49a9-b26c-f2c... | 华北北京一(cn-north-1) | 0    | A   |
| ECS-ECS | 03215acd-eb0c-49a9-b26c-f2c... | 华北北京一(cn-north-1) | 0    | 无标签 |

## 预定义标签的使用方法

如果有多台云服务器或其他云资源需要添加同一标签，为了避免重复输入标签键和值，您可以在标签管理服务中预定义标签，然后在添加标签时直接选择键和值。具体步骤如下：


1. 登录管理控制台。
2. 在右上角的用户名下选择“标签管理”，进入标签管理服务页面。
3. 在左侧导航中选择“预定义标签”，单击“创建标签”，输入标签键和值（例如，项目-A）。
4. 选择“服务列表 > 计算 > 弹性云服务器”，按照上述添加标签的方法，在标签键和标签值输入框中下拉选择预定义的标签。

### 15.1.3 使用标签检索资源

为云资源添加标签后，您可以通过本文所述的两种方法使用标签检索资源。

#### 使用标签搜索云服务器

在弹性云服务器列表页，按标签键或键值对搜索目标云服务器。

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“≡”，选择“计算 > 弹性云服务器”。
4. 在云服务器列表页搜索的输入框中选择“资源标签”，选择标签键值对后自动搜索。

支持多个标签搜索，按“与”的关系搜索目标云服务器。

图 15-7 按标签搜索云服务器

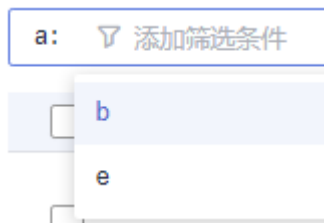


图 15-8 按标签搜索云服务器



## 通过标签管理筛选资源


1. 登录管理控制台。
  2. 在右上角的用户名下选择“标签管理”，进入标签管理服务页面。
  3. 在“资源标签”页面，设置搜索条件（包括区域、资源类型、资源标签）。
  4. 单击“搜索”。
- 搜索结果区域将列出所有符合搜索条件的资源。

### 15.1.4 删除标签

如果某个标签已经不再适用于您的资源管理，您可以删除资源标签。有三种途径删除资源标签：

- [在弹性云服务器详情页删除](#)
- [在标签管理页面单个删除](#)
- [在标签管理页面批量删除](#)

#### 在弹性云服务器详情页删除

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”，进入弹性云服务器页面。
4. 在弹性云服务器列表中，单击待删除标签的弹性云服务器名称。  
系统跳转至该弹性云服务器详情页面。
5. 选择“标签”页签，单击标签所在行“操作”列下的“删除”。
6. 单击“确定”，完成标签的删除。

#### 在标签管理页面单个删除

1. 登录管理控制台。






2. 在右上角的用户名下选择“标签管理”，进入标签管理服务页面。

图 15-9 标签管理



3. 在“资源标签”页面，设置弹性云服务器资源搜索条件，单击“搜索”。
4. 页面下方展示的搜索结果包含“编辑”与“只读”两种状态，单击“编辑”，切换资源标签列表为可编辑状态。

若需要删除的标签的“键”没有展示在列表中，单击 ，在下拉列表中勾选需要删除的标签的“键”。勾选需要展示的标签键建议不超过10个。


5. 单击待删除标签的弹性云服务器资源所在行的 ，资源标签删除完成。
6. （可选）单击搜索结果区域右侧的  按钮。  
资源标签列表刷新为最新状态，并更新列表刷新时间。

## 在标签管理页面批量删除

### 须知

在批量删除标签时，请谨慎操作。执行删除操作后，其所标识的所有弹性云服务器资源对应的该标签均会被删除，且不可恢复。

1. 登录管理控制台。
2. 在右上角的用户名下选择“标签管理”，进入标签管理服务页面。
3. 在“资源标签”页面，设置弹性云服务器资源搜索条件，单击“搜索”。
4. 勾选待删除标签的弹性云服务器资源。
5. 单击列表上方的“管理标签”，进入管理标签页面。

- 单击待删除标签所在行的“删除”。单击“确认”，资源标签删除完成。
- (可选)单击搜索结果区域右侧的  按钮。  
资源标签列表刷新为最新状态，并更新列表刷新时间。

## 15.2 配额调整

### 什么是配额？

为防止资源滥用，平台限制了各服务资源的配额，对用户的资源数量和容量做了限制。如您最多可以创建多少台弹性云服务器、多少块云硬盘。

如果当前资源配额限制无法满足使用需要，您可以申请扩大配额。

### 怎样查看我的配额？


- 登录管理控制台。
- 单击管理控制台左上角的 ，选择区域和项目。
- 在页面右上角，选择“资源 > 我的配额”。  
系统进入“服务配额”页面。

图 15-10 我的配额



- 您可以在“服务配额”页面，查看各项资源的总配额及使用情况。  
如果当前配额不能满足业务要求，请参考后续操作，申请扩大配额。

### 如何申请扩大配额？

- 登录管理控制台。
- 在页面右上角，选择“资源 > 我的配额”。  
系统进入“服务配额”页面。

图 15-11 我的配额



3. 在页面右上角，单击“申请扩大配额”。

图 15-12 申请扩大配额

| 服务                  | 资源名称               | 已用配额 | 总配额 |
|---------------------|--------------------|------|-----|
| 弹性网络 AS             | 带宽                 | 0    |     |
|                     | 中间件                | 0    |     |
| 云容器引擎 CCE           | 集群                 | 0    |     |
|                     | 镜像                 | 0    |     |
| 弹性伸缩 AS             | 弹性伸缩               | 0    |     |
|                     | 弹性伸缩 (AS)          | 0    |     |
| 函数工作流 FunctionGraph | 函数                 | 3    |     |
|                     | 函数 (FunctionGraph) | 0    |     |
| 云硬盘 EVS             | 云硬盘 (EVS)          | 120  |     |
|                     | 快照                 | 4    |     |
| 弹性公网 IP             | 弹性公网 IP            | 0    |     |
|                     | 弹性公网 IP            | 0    |     |
| 云堡垒机堡垒机             | 弹性公网 IP            | 0    |     |
|                     | 弹性公网 IP            | 0    |     |
| 弹性文件存储 SFS          | 文件存储               | 0    |     |
|                     | 文件存储 (SFS)         | 0    |     |
| CDN                 | CDN                | 0    |     |
|                     | CDN                | 0    |     |

4. 在“新建工单”页面，根据您的需求，填写相关参数。  
其中，“问题描述”项请填写需要调整的内容和申请原因。
5. 填写完毕后，勾选协议并单击“提交”。

# 16 使用 CES 监控 ECS

## 16.1 监控弹性云服务器

监控是保持弹性云服务器可靠性、可用性和性能的重要部分，通过监控，用户可以观察弹性云服务器资源。为使用户更好地掌握自己的弹性云服务器运行状态，云服务平台提供了云监控。您可以使用该服务监控您的弹性云服务器，执行自动实时监控、告警和通知操作，帮助您更好地了解弹性云服务器的各项性能指标。

主机监控分为基础监控和操作系统监控。

- 基础监控为ECS自动上报的监控指标。
- 操作系统监控通过在弹性云服务器中安装Agent插件，为用户提供服务器的系统级、主动式、细颗粒度监控服务。

安装配置Agent相关操作请参考“云监控服务 > 用户指南 > 主机监控”。

通过后续章节，您可以了解以下内容：

- 弹性云服务器当前支持的基础监控指标
- 弹性云服务器操作系统监控的监控指标（安装Agent）
- 弹性云服务器进程监控的监控指标（安装Agent）
- 如何自定义弹性云服务器告警规则
- 如何查看弹性云服务器运行状态进行日常监控

### 一键告警

弹性云服务器运行在物理机上，虽然提供了多种机制来保证系统的可靠性、容错能力和高可用性，但是，服务器的硬件、电源等部件仍有较小概率的损坏。云平台默认提供了自动恢复功能，当弹性云服务器所在的硬件出现故障时，系统会自动将弹性云服务器迁移至正常的物理机，保障您受到的影响最小，该过程会导致云服务器重启。[了解更多](#)

您可以在云监控服务控制台为弹性云服务器开启“一键告警”功能，以便在HA发生（弹性云服务器所在的物理机出现故障，系统自动迁移弹性云服务器至正常的物理机）时，及时获得通知。配置方法请参见[一键告警](#)。

## 相关链接

- [Windows云服务器带宽和CPU利用率高问题排查方法](#)
- [Linux云服务器带宽和CPU占用率高问题排查方法](#)

## 16.2 弹性云服务器支持的基础监控指标

### 功能说明

本节定义了弹性云服务器上报云监控的基础监控指标的命名空间，监控指标列表，各项监控指标的具体含义与使用说明，用户可以通过云监控检索弹性云服务器服务产生的监控指标和告警信息。

### 命名空间

SYS.ECS

### 基础监控指标支持列表

对于不同的操作系统、不同的弹性云服务器类型，支持的基础监控指标有所差异，具体如表16-1所示。其中，√表示支持，×表示不支持。

#### 说明

- 对于部分监控指标，需弹性云服务器使用的镜像安装UVP VMTools，UVP VMTools安装的具体操作，请参见<https://github.com/UVP-Tools/UVP-Tools/>。
- 对于部分监控指标，需弹性云服务器安装Agent。安装成功后，您可以登录控制台，打开“云监控”服务，在“主机监控 > 弹性云服务器 > 操作系统监控”页面查看各项监控指标，如“AGT.用户空间CPU使用率”，具体请参见[弹性云服务器支持的操作系统监控指标（安装Agent）](#)。
- Windows云服务器安装Agent的具体操作，请参见《云监控用户指南》的“安装配置Agent（Windows）”章节。
- Linux云服务器安装Agent的具体操作，请参见《云监控用户指南》的“安装配置Agent（Linux）”章节。

表 16-1 弹性云服务器基础监控指标支持列表

| 指标ID     | 监控指标   | Windows弹性云服务器    |                  | Linux弹性云服务器                      |                  |
|----------|--------|------------------|------------------|----------------------------------|------------------|
|          |        | 虚拟化类型为XEN的弹性云服务器 | 虚拟化类型为KVM的弹性云服务器 | 虚拟化类型为XEN的弹性云服务器                 | 虚拟化类型为KVM的弹性云服务器 |
| -        | -      |                  |                  |                                  |                  |
| cpu_util | CPU使用率 | √                | √                | √                                | √                |
| mem_util | 内存使用率  | √                | √                | √（镜像需安装UVP VMTools，否则无法获取该监控指标。） | ×（暂不支持）          |

| 指标ID                                  | 监控指标     | Windows弹性云服务器                              |   | Linux弹性云服务器                                |         |
|---------------------------------------|----------|--------------------------------------------|---|--------------------------------------------|---------|
| disk_util_inband                      | 磁盘使用率    | √                                          | √ | √（镜像需安装UVP VMTools，否则无法获取该监控指标。）           | ×（暂不支持） |
| disk_read_bytes_rate                  | 磁盘读带宽    | √                                          | √ | √                                          | √       |
| disk_write_bytes_rate                 | 磁盘写带宽    | √                                          | √ | √                                          | √       |
| disk_read_requests_rate               | 磁盘读IOPS  | √                                          | √ | √                                          | √       |
| disk_write_requests_rate              | 磁盘写IOPS  | √                                          | √ | √                                          | √       |
| network_incoming_bytes_rate_inband    | 带内网络流入速率 | √（镜像需安装UVP VMTools，否则无法获取该监控指标。）           | √ | √（镜像需安装UVP VMTools，否则无法获取该监控指标。）           | ×（暂不支持） |
| network_outgoing_bytes_rate_inband    | 带内网络流出速率 | √（镜像需安装UVP VMTools，否则无法获取该监控指标。）           | √ | √（镜像需安装UVP VMTools，否则无法获取该监控指标。）           | ×（暂不支持） |
| network_incoming_bytes_aggregate_rate | 带外网络流入速率 | √（镜像如果安装了UVP VMTools，则不提供该指标，请使用带内网络流出速率。） | √ | √（镜像如果安装了UVP VMTools，则不提供该指标，请使用带内网络流出速率。） | √       |
| network_outgoing_bytes_aggregate_rate | 带外网络流出速率 | √（镜像如果安装了UVP VMTools，则不提供该指标，请使用带内网络流出速率。） | √ | √（镜像如果安装了UVP VMTools，则不提供该指标，请使用带内网络流出速率。） | √       |

| 指标ID                      | 监控指标       | Windows弹性云服务器 |            | Linux弹性云服务器 |            |
|---------------------------|------------|---------------|------------|-------------|------------|
| cpu_credit_usage          | CPU积分使用量   | ×             | √ (仅T6型支持) | ×           | √ (仅T6型支持) |
| cpu_credit_balance        | CPU积分累积量   | ×             | √ (仅T6型支持) | ×           | √ (仅T6型支持) |
| network_vm_connections    | 网络连接数      | ×             | √          | ×           | √          |
| network_vm_bandwidth_in   | 虚拟机入方向带宽   | ×             | √          | ×           | √          |
| network_vm_bandwidth_out  | 虚拟机出方向带宽   | ×             | √          | ×           | √          |
| network_vm_pps_in         | 虚拟机入方向PPS  | ×             | √          | ×           | √          |
| network_vm_pps_out        | 虚拟机出方向PPS  | ×             | √          | ×           | √          |
| network_vm_newconnections | 虚拟机整机新建连接数 | ×             | √          | ×           | √          |

### 说明

“√”表示支持该监控指标，“×”表示不支持该监控指标。

各项基础监控指标的具体含义与使用说明如表16-2所示。

以下监控指标（原始指标）的监控周期如下：

- 虚拟化类型为XEN的弹性云服务器：4分钟。
- 虚拟化类型为KVM的弹性云服务器：5分钟。

表 16-2 弹性云服务器支持的基础监控指标

| 指标ID             | 指标名称   | 指标含义                                                                                                                                                  | 取值范围 | 测量对象<br>(维度) | 监控周期<br>(原始指标, 本列监控周期值适用于KVM实例) |
|------------------|--------|-------------------------------------------------------------------------------------------------------------------------------------------------------|------|--------------|---------------------------------|
| cpu_util         | CPU使用率 | 该指标用于统计弹性云服务器的CPU使用率。<br>该指标为从物理机层面采集的CPU使用率, 数据准确性低于从弹性云服务器内部采集的数据, 查看 <a href="#">详情</a> 。<br>单位: 百分比。<br>计算公式: 单个弹性云服务器CPU使用率 / 单个弹性云服务器的CPU总核数。   | ≥ 0% | 云服务器         | 5分钟                             |
| mem_util         | 内存使用率  | 该指标用于统计弹性云服务器的内存使用率。<br>如果用户使用的镜像未安装UVP VMTools, 则无法获取该监控指标。<br>单位: 百分比。<br>计算公式: 该弹性云服务器内存使用量 / 该弹性云服务器内存总量。<br><b>说明</b><br>内存使用率监控指标不支持QingTian实例。 | ≥ 0% | 云服务器         | 5分钟                             |
| disk_util_inband | 磁盘使用率  | 该指标用于统计弹性云服务器的磁盘使用情况。<br>如果用户使用的镜像未安装UVP VMTools, 则无法获取该监控指标。<br>单位: 百分比。<br>计算公式: 该弹性云服务器磁盘使用容量 / 该弹性云服务器磁盘总容量。                                      | ≥ 0% | 云服务器         | 5分钟                             |



| 指标ID                               | 指标名称     | 指标含义                                                                                                                                  | 取值范围                  | 测量对象<br>(维度) | 监控周期<br>(原始指标, 本列监控周期值适用于KVM实例) |
|------------------------------------|----------|---------------------------------------------------------------------------------------------------------------------------------------|-----------------------|--------------|---------------------------------|
| disk_read_bytes_rate               | 磁盘读带宽    | 该指标用于统计每秒从弹性云服务器读出数据量。<br>单位: 字节/秒。<br>计算公式: 该弹性云服务器的磁盘读出的字节数之和 / 测量周期。<br>$byte\_out = (rd\_bytes - last\_rd\_bytes) / \text{时间差}$ 。 | $\geq 0$<br>Byte/s    | 云服务器         | 5分钟                             |
| disk_write_bytes_rate              | 磁盘写带宽    | 该指标用于统计每秒写到弹性云服务器的数据量。<br>单位: 字节/秒。<br>计算公式: 该弹性云服务器的磁盘写入的字节数之和 / 测量周期。                                                               | $\geq 0$<br>Byte/s    | 云服务器         | 5分钟                             |
| disk_read_requests_rate            | 磁盘读 IOPS | 该指标用于统计每秒从弹性云服务器读取数据的请求次数。<br>单位: 请求/秒。<br>计算公式: 请求读取该弹性云服务器磁盘的次数之和 / 测量周期。<br>$req\_out = (rd\_req - last\_rd\_req) / \text{时间差}$ 。  | $\geq 0$<br>request/s | 云服务器         | 5分钟                             |
| disk_write_requests_rate           | 磁盘写 IOPS | 该指标用于统计每秒从弹性云服务器写数据的请求次数。<br>单位: 请求/秒。<br>计算公式: 请求写入该弹性云服务器磁盘的次数之和 / 测量周期。<br>$req\_in = (wr\_req - last\_wr\_req) / \text{时间差}$ 。    | $\geq 0$<br>request/s | 云服务器         | 5分钟                             |
| network_incoming_bytes_rate_inband | 带内网络流入速率 | 该指标用于在弹性云服务器内统计每秒流入弹性云服务器的网络流量。<br>单位: 字节/秒。<br>计算公式: 该弹性云服务器的带内网络流入字节数之和/测量周期。                                                       | $\geq 0$<br>Byte/s    | 云服务器         | 5分钟                             |

| 指标ID                                  | 指标名称     | 指标含义                                                                                                    | 取值范围          | 测量对象<br>(维度) | 监控周期<br>(原始指标, 本列监控周期值适用于KVM实例) |
|---------------------------------------|----------|---------------------------------------------------------------------------------------------------------|---------------|--------------|---------------------------------|
| network_outgoing_bytes_rate_inband    | 带内网络流出速率 | 该指标用于在弹性云服务器内统计每秒流出弹性云服务器的网络流量。<br>单位: 字节/秒。<br>计算公式: 该弹性云服务器的带内网络流出字节数之和 / 测量周期。                       | ≥ 0<br>Byte/s | 云服务器         | 5分钟                             |
| network_incoming_bytes_aggregate_rate | 带外网络流入速率 | 该指标用于在虚拟化层统计每秒流入弹性云服务器的网络流量。<br>单位: 字节/秒。<br>计算公式: 该弹性云服务器的带外网络流入字节数之和 / 测量周期。<br>当使用SRIOV时, 无法获取该监控指标。 | ≥ 0<br>Byte/s | 云服务器         | 5分钟                             |
| network_outgoing_bytes_aggregate_rate | 带外网络流出速率 | 该指标用于在虚拟化层统计每秒流出弹性云服务器的网络流量。<br>单位: 字节/秒。<br>计算公式: 该弹性云服务器的带外网络流出字节数之和 / 测量周期。<br>当使用SRIOV时, 无法获取该监控指标。 | ≥ 0<br>Byte/s | 云服务器         | 5分钟                             |
| network_vm_connections                | 网络连接数    | 该指标表示弹性云服务器已经使用的TCP和UDP的连接数总和。<br>单位: 个<br><b>说明</b><br>该指标通过带外采集, 因此可能会出现该指标值大于OS中查询到的网络连接数的情况。        | ≥ 0           | 云服务器         | 5分钟                             |
| network_vm_bandwidth_in               | 虚拟机入方向带宽 | 虚拟机整机每秒接收的比特数, 此处为公网和内网流量总和。<br>单位: 字节/秒                                                                | ≥ 0           | 云服务器         | 5分钟                             |

| 指标ID                      | 指标名称       | 指标含义                                         | 取值范围     | 测量对象(维度) | 监控周期(原始指标, 本列监控周期值适用于KVM实例) |
|---------------------------|------------|----------------------------------------------|----------|----------|-----------------------------|
| network_vm_bandwidth_out  | 虚拟机出方向带宽   | 虚拟机整机每秒发送的比特数, 此处为公网和内网流量总和。<br>单位: 字节/秒     | $\geq 0$ | 云服务器     | 5分钟                         |
| network_vm_pps_in         | 虚拟机入方向PPS  | 虚拟机整机每秒接收的数据包数, 此处为公网和内网数据包总和。<br>单位: 包/秒    | $\geq 0$ | 云服务器     | 5分钟                         |
| network_vm_pps_out        | 虚拟机出方向PPS  | 虚拟机整机每秒发送的数据包数, 此处为公网和内网数据包总和。<br>单位: 包/秒    | $\geq 0$ | 云服务器     | 5分钟                         |
| network_vm_newconnections | 虚拟机整机新建连接数 | 虚拟机整机新建连接数, 包括TCP协议、UDP协议以及ICMP协议等。<br>单位: 个 | $\geq 0$ | 云服务器     | 5分钟                         |

## 维度

| 维度   | Key         | Value   |
|------|-------------|---------|
| 云服务器 | instance_id | 云服务器ID。 |

## 16.3 弹性云服务器支持的操作系统监控指标 (安装 Agent)

### 功能说明

通过在弹性云服务器中安装Agent插件, 可以为用户提供服务器的系统级、主动式、细颗粒度监控服务。本节定义了弹性云服务器上报云监控的操作系统监控指标。

操作系统监控目前支持的监控指标有: CPU相关监控项、CPU负载类相关监控项、内存相关监控项、磁盘相关监控项、磁盘I/O相关监控项、文件系统类相关监控项、网卡类相关监控项、NTP类相关监控项、TCP连接数类相关监控、GPU相关监控项。

安装Agent后, 对于不同的操作系统、不同的弹性云服务器类型, 您可以查看不同类型的操作系统监控指标。指标采集周期是1分钟。

## 命名空间

AGT.ECS

## 操作系统监控指标：CPU

表 16-3 CPU 相关监控指标说明

| 指标             | 指标名称              | 指标含义                                                                                                                                                                                                      | 取值范围   | 测量对象（维度） | 监控周期（原始指标） |
|----------------|-------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|----------|------------|
| cpu_usage      | (Agent) CPU使用率    | 该指标用于统计测量对象当前CPU使用率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/stat中的变化得出cpu使用率。用户可以通过top命令查看%Cpu(s)值。</li><li>采集方式（Windows）：通过WindowsAPI GetSystemTimes获取。</li></ul> | 0-100% | 云服务器     | 1分钟        |
| cpu_usage_idle | (Agent) CPU空闲时间占比 | 该指标用于统计测量对象当前CPU空闲时间占比。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/stat中的变化得出CPU空闲时间占比。</li><li>采集方式（Windows）：通过WindowsAPI GetSystemTimes获取。</li></ul>                 | 0-100% | 云服务器     | 1分钟        |

| 指标               | 指标名称                | 指标含义                                                                                                                                                                                                                              | 取值范围   | 测量对象(维度) | 监控周期(原始指标) |
|------------------|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|----------|------------|
| cpu_usage_user   | (Agent) 用户空间 CPU使用率 | 该指标用于统计测量对象当前用户空间占用CPU使用率。<br>单位: 百分比<br><ul style="list-style-type: none"> <li>采集方式 (Linux): 通过计算采集周期内/proc/stat中的变化得出cpu使用率。用户可以通过top命令查看 %Cpu(s) us值。</li> <li>采集方式 (Windows): 通过 WindowsAPI GetSystemTimes获取。</li> </ul>      | 0-100% | 云服务器     | 1分钟        |
| cpu_usage_system | (Agent) 内核空间 CPU使用率 | 该指标用于统计测量对象当前内核空间占用CPU使用率。<br>单位: 百分比<br><ul style="list-style-type: none"> <li>采集方式 (Linux): 通过计算采集周期内/proc/stat中的变化得出内核空间CPU使用率。用户可以通过top命令查看 %Cpu(s) sy值。</li> <li>采集方式 (Windows): 通过 WindowsAPI GetSystemTimes获取。</li> </ul>  | 0-100% | 云服务器     | 1分钟        |
| cpu_usage_other  | (Agent) 其他CPU使用率    | 该指标用于统计测量对象其他占用CPU使用率。<br>单位: 百分比<br><ul style="list-style-type: none"> <li>采集方式 (Linux): 其他CPU使用率=1- 空闲CPU使用率(%) - 内核空间CPU使用率- 用户空间CPU使用率。</li> <li>采集方式 (Windows): 其他CPU使用率=1- 空闲CPU使用率(%) - 内核空间CPU使用率- 用户空间CPU使用率。</li> </ul> | 0-100% | 云服务器     | 1分钟        |

| 指标               | 指标名称                            | 指标含义                                                                                                                                                                                                     | 取值范围   | 测量对象(维度) | 监控周期(原始指标) |
|------------------|---------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|----------|------------|
| cpu_usage_nice   | (Agent)<br>Nice进程<br>CPU使用<br>率 | 该指标用于统计测量对象当前Nice进程CPU使用率。<br>单位: 百分比<br><ul style="list-style-type: none"><li>采集方式 (Linux): 通过计算采集周期内/proc/stat中的变化得出Nice进程CPU使用率。用户可以通过top命令查看 %Cpu(s) ni值。</li><li>采集方式 (Windows): 暂不支持。</li></ul>    | 0-100% | 云服务器     | 1分钟        |
| cpu_usage_iowait | (Agent)<br>iowait状态占比           | 该指标用于统计测量对象当前iowait状态占用CPU的比率。<br>单位: 百分比<br><ul style="list-style-type: none"><li>采集方式 (Linux): 通过计算采集周期内/proc/stat中的变化得出iowait状态占比。用户可以通过top命令查看 %Cpu(s) wa值。</li><li>采集方式 (Windows): 暂不支持。</li></ul>  | 0-100% | 云服务器     | 1分钟        |
| cpu_usage_irq    | (Agent)<br>CPU中断<br>时间占比        | 该指标用于统计测量对象当前CPU处理中断时占用CPU时间的比率。<br>单位: 百分比<br><ul style="list-style-type: none"><li>采集方式 (Linux): 通过计算采集周期内/proc/stat中的变化得出CPU中断时间占比。用户可以通过top命令查看 %Cpu(s) hi值。</li><li>采集方式 (Windows): 暂不支持。</li></ul> | 0-100% | 云服务器     | 1分钟        |

| 指标                | 指标名称               | 指标含义                                                                                                                                                                                                           | 取值范围   | 测量对象（维度） | 监控周期（原始指标） |
|-------------------|--------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|----------|------------|
| cpu_usage_softirq | (Agent) CPU软中断时间占比 | 该指标用于统计测量对象当前CPU处理软中断时间占用CPU时间的比率。<br>单位：百分比<br><ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>通过计算采集周期内/proc/stat中的变化得出CPU软中断时间占比。用户可以通过top命令查看 %Cpu(s) si值。</li> <li>采集方式（Windows）：暂不支持。</li> </ul> | 0-100% | 云服务器     | 1分钟        |

## 操作系统监控指标：CPU 负载

表 16-4 CPU 负载指标说明

| 指标             | 指标名称            | 指标含义                                                                                          | 取值范围 | 测量对象（维度） | 监控周期（原始指标） |
|----------------|-----------------|-----------------------------------------------------------------------------------------------|------|----------|------------|
| load_averager1 | (Agent) 1分钟平均负载 | 该指标用于统计测量对象过去1分钟的CPU平均负载。<br>采集方式（Linux）：通过/proc/loadavg中load1/逻辑CPU个数得到。用户可以通过top命令查看load1值。 | ≥0   | 云服务器     | 1分钟        |
| load_averager5 | (Agent) 5分钟平均负载 | 该指标用于统计测量对象过去5分钟的CPU平均负载。<br>采集方式（Linux）：通过/proc/loadavg中load5/逻辑CPU个数得到。用户可以通过top命令查看load5值。 | ≥0   | 云服务器     | 1分钟        |

| 指标             | 指标名称             | 指标含义                                                                                             | 取值范围 | 测量对象（维度） | 监控周期（原始指标） |
|----------------|------------------|--------------------------------------------------------------------------------------------------|------|----------|------------|
| load_average15 | (Agent) 15分钟平均负载 | 该指标用于统计测量对象过去15分钟的CPU平均负载。<br>采集方式（Linux）：通过/proc/loadavg中load15/逻辑CPU个数得到。用户可以通过top命令查看load15值。 | ≥0   | 云服务器     | 1分钟        |

### 说明

Windows系统暂不支持CPU负载指标。



## 操作系统监控指标：内存

表 16-5 内存相关监控指标说明

| 指标            | 指标名称            | 指标含义                                                                                                                                                                                                                                                                                                                                                            | 取值范围  | 测量对象<br>(维度) | 监控周期<br>(原始指标) |
|---------------|-----------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|--------------|----------------|
| mem_available | (Agent)<br>可用内存 | 该指标用于统计测量对象的可用内存。<br>单位：GB<br><ul style="list-style-type: none"><li>采集方式（Linux）：<br/>通过/proc/meminfo文件获取，<ul style="list-style-type: none"><li>若/proc/meminfo中显示MemAvailable，则直接可得</li><li>若/proc/meminfo中不显示MemAvailable，则MemAvailable=MemFree+Buffers+Cached</li></ul></li><li>采集方式（Windows）：计算方法为（内存总量-已用内存量）。通过WindowsAPI GlobalMemoryStatusEx获取。</li></ul> | ≥0 GB | 云服务器         | 1分钟            |

| 指标              | 指标名称          | 指标含义                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                  | 取值范围   | 测量对象<br>(维度) | 监控周期<br>(原始指标) |
|-----------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------------|----------------|
| mem_usedPercent | (Agent) 内存使用率 | <p>该指标用于统计测量对象的内存使用率。</p> <p>单位：百分比</p> <ul style="list-style-type: none"><li>采集方式（Linux）：<br/>通过/proc/meminfo文件获取，<math>(\text{MemTotal} - \text{MemAvailable}) / \text{MemTotal}</math><ul style="list-style-type: none"><li>若/proc/meminfo中显示 MemAvailable，则 <math>\text{MemUsedPercent} = (\text{MemTotal} - \text{MemAvailable}) / \text{MemTotal}</math></li><li>若/proc/meminfo中不显示 MemAvailable，则 <math>\text{MemUsedPercent} = (\text{MemTotal} - \text{MemFree} - \text{Buffers} - \text{Cached}) / \text{MemTotal}</math></li></ul></li><li>采集方式（Windows）：计算方法为（已用内存量/内存总量*100%）。</li></ul> | 0-100% | 云服务器         | 1分钟            |
| mem_free        | (Agent) 空闲内存量 | <p>该指标用于统计测量对象的空闲内存量。</p> <p>单位：GB</p> <ul style="list-style-type: none"><li>采集方式（Linux）：<br/>通过/proc/meminfo获取。</li><li>采集方式（Windows）：暂不支持。</li></ul>                                                                                                                                                                                                                                                                                                                                                                                                                                                | ≥0 GB  | 云服务器         | 1分钟            |

| 指标               | 指标名称               | 指标含义                                                                                                                                                                                    | 取值范围  | 测量对象<br>(维度) | 监控周期<br>(原始指标) |
|------------------|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|--------------|----------------|
| mem_buffers      | (Agent) Buffers占用量 | 该指标用于统计测量对象的Buffers内存量。<br>单位: GB<br><ul style="list-style-type: none"> <li>采集方式 (Linux): 通过/proc/meminfo获取。用户可以通过top命令查看 KiB Mem:buffers值。</li> <li>采集方式 (Windows): 暂不支持。</li> </ul>   | ≥0 GB | 云服务器         | 1分钟            |
| mem_cached       | (Agent) Cache占用量   | 该指标用于统计测量对象Cache内存量。<br>单位: GB<br><ul style="list-style-type: none"> <li>采集方式 (Linux): 通过/proc/meminfo获取。用户可以通过top命令查看 KiB Swap:cached Mem 值。</li> <li>采集方式 (Windows): 暂不支持。</li> </ul> | ≥0 GB | 云服务器         | 1分钟            |
| total_open_files | (Agent) 文件句柄总数     | 该指标用于统计测量对象的所有进程使用的句柄总和。<br>单位: 个<br><ul style="list-style-type: none"> <li>采集方式 (Linux): 通过/proc/{pid}/fd文件汇总所有进程使用的句柄数。</li> <li>采集方式 (Windows): 暂不支持。</li> </ul>                     | ≥0    | 云服务器         | 1分钟            |

## 操作系统监控指标：磁盘

### 📖 说明

- 目前仅支持物理磁盘指标的采集，不支持通过网络文件系统协议挂载的磁盘。
- 会默认屏蔽docker相关的挂载点。挂载点前缀如下：  
/var/lib/docker;/mnt/paas/kubernetes;/var/lib/mesos

表 16-6 磁盘相关监控指标说明

| 指标         | 指标名称                   | 指标含义                                                                                                                                                                                                                                                                                                                                                      | 取值范围  | 测量对象<br>(维度)  | 监控周期<br>(原始指标) |
|------------|------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------|---------------|----------------|
| disk_free  | (Agent)<br>磁盘剩余<br>存储量 | 该指标用于统计测量对象<br>磁盘的剩余存储空间。<br>单位：GB<br><ul style="list-style-type: none"><li>采集方式（Linux）：<br/>执行df -h命令，查看<br/>Avail列数据。挂载点前<br/>缀路径长度不能超过64<br/>个字符，必须以字母开<br/>头，只能包含0-9/a-<br/>z/A-Z/-/./~。</li><li>采集方式<br/>（Windows）：使用<br/>WMI接口<br/>GetDiskFreeSpaceExW<br/>获取磁盘空间数据。挂<br/>载点前缀路径长度不能<br/>超过64个字符，必须以<br/>字母开头，只能包含<br/>0-9/a-z/A-Z/-/./~。</li></ul> | ≥0 GB | 云服务器 -<br>挂载点 | 1分钟            |
| disk_total | (Agent)<br>磁盘存储<br>总量  | 该指标用于统计测量对象<br>磁盘存储总量。<br>单位：GB<br><ul style="list-style-type: none"><li>采集方式（Linux）：<br/>执行df -h命令，查看<br/>Size列数据。<br/>挂载点前缀路径长度不<br/>能超过64个字符，必须<br/>以字母开头，只能包含<br/>0-9/a-z/A-Z/-/./~。</li><li>采集方式<br/>（Windows）：使用<br/>WMI接口<br/>GetDiskFreeSpaceExW<br/>获取磁盘空间数据。挂<br/>载点前缀路径长度不能<br/>超过64个字符，必须以<br/>字母开头，只能包含<br/>0-9/a-z/A-Z/-/./~。</li></ul>     | ≥0 GB | 云服务器 -<br>挂载点 | 1分钟            |

| 指标                       | 指标名称                  | 指标含义                                                                                                                                                                                                                                                                                                                                                                                   | 取值范围   | 测量对象<br>(维度)      | 监控周期<br>(原始指标) |
|--------------------------|-----------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|-------------------|----------------|
| disk_use<br>d            | (Agent)<br>磁盘已用<br>存量 | 该指标用于统计测量对象<br>磁盘的已用存储空间。<br>单位：GB <ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>执行df -h命令，查看<br/>Used列数据。挂载点前<br/>缀路径长度不能超过64<br/>个字符，必须以字母开<br/>头，只能包含0-9/a-<br/>z/A-Z/-/./~。</li> <li>采集方式<br/>（Windows）：使用<br/>WMI接口<br/>GetDiskFreeSpaceExW<br/>获取磁盘空间数据。挂<br/>载点前缀路径长度不能<br/>超过64个字符，必须以<br/>字母开头，只能包含<br/>0-9/a-z/A-Z/-/./~。</li> </ul>                               | ≥0 GB  | 云服务器 -<br>挂载<br>点 | 1分钟            |
| disk_use<br>dPercen<br>t | (Agent)<br>磁盘使用<br>率  | 该指标用于统计测量对象<br>磁盘使用率，以百分比为<br>单位。计算方式为：磁盘已<br>用存储量/磁盘存储总量。<br>单位：百分比 <ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>通过计算Used/Size得<br/>出。挂载点前缀路径长<br/>度不能超过64个字符，<br/>必须以字母开头，只能<br/>包含0-9/a-z/A-<br/>Z/-/./~。</li> <li>采集方式<br/>（Windows）：使用<br/>WMI接口<br/>GetDiskFreeSpaceExW<br/>获取磁盘空间数据。挂<br/>载点前缀路径长度不能<br/>超过64个字符，必须以<br/>字母开头，只能包含<br/>0-9/a-z/A-Z/-/./~。</li> </ul> | 0-100% | 云服务器 -<br>挂载<br>点 | 1分钟            |

## 操作系统监控指标：磁盘 I/O

表 16-7 磁盘 I/O 相关监控指标说明

| 指标                       | 指标名称          | 指标含义                                                                                                                                                                                                                                                                                                                                                                                                                                                         | 取值范围          | 测量对象<br>(维度)                                                                    | 监控周期<br>(原始指标) |
|--------------------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------------------------------------------------------------------------|----------------|
| disk_agt_read_bytes_rate | (Agent) 磁盘读速率 | <p>该指标用于统计每秒从测量对象读出数据量。<br/>单位：Byte/s</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第六列数据的变化得出磁盘读速率。<br/>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>采集方式（Windows）： <ul style="list-style-type: none"> <li>使用WMI中Win32_PerfFormattedData_PerfDisk_LogicalDisk对象获取磁盘I/O数据。</li> <li>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>高CPU情况下存在获取超时的现象，会导致无法获取监控数据。</li> </ul> </li> </ul> | ≥ 0<br>Byte/s | <ul style="list-style-type: none"> <li>云服务器 - 磁盘</li> <li>云服务器 - 挂载点</li> </ul> | 1分钟            |

| 指标                          | 指标名称            | 指标含义                                                                                                                                                                                                                                                                                                                                                                                                                                                                | 取值范围     | 测量对象<br>(维度)                                                                    | 监控周期<br>(原始指标) |
|-----------------------------|-----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|---------------------------------------------------------------------------------|----------------|
| disk_agt_read_requests_rate | (Agent) 磁盘读操作速率 | <p>该指标用于统计每秒从测量对象读取数据的请求次数。</p> <p>单位：请求/秒</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第四列数据的变化得出磁盘读操作速率。<br/>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>采集方式（Windows）： <ul style="list-style-type: none"> <li>使用WMI中Win32_PerfFormattedData_PerfDisk_LogicalDisk对象获取磁盘I/O数据。</li> <li>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>高CPU情况下存在获取超时的现象，会导致无法获取监控数据。</li> </ul> </li> </ul> | ≥ 0 请求/秒 | <ul style="list-style-type: none"> <li>云服务器 - 磁盘</li> <li>云服务器 - 挂载点</li> </ul> | 1分钟            |

| 指标                        | 指标名称          | 指标含义                                                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 取值范围          | 测量对象<br>(维度)                                                                    | 监控周期<br>(原始指标) |
|---------------------------|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------|---------------------------------------------------------------------------------|----------------|
| disk_agt_write_bytes_rate | (Agent) 磁盘写速率 | <p>该指标用于统计每秒写到测量对象的数据量。<br/>单位: Byte/s</p> <ul style="list-style-type: none"> <li>采集方式 (Linux) :<br/>通过计算采集周期内/proc/diskstats中对应设备第十列数据的变化得出磁盘写速率。<br/>挂载点前缀路径长度不能超过64个字符, 必须以字母开头, 只能包含0-9/a-z/A-Z/-/./~。</li> <li>采集方式 (Windows) :                             <ul style="list-style-type: none"> <li>使用WMI中Win32_PerfFormattedData_PerfDisk_LogicalDisk对象获取磁盘I/O数据。</li> <li>挂载点前缀路径长度不能超过64个字符, 必须以字母开头, 只能包含0-9/a-z/A-Z/-/./~。</li> <li>高CPU情况下存在获取超时的现象, 会导致无法获取监控数据。</li> </ul> </li> </ul> | ≥ 0<br>Byte/s | <ul style="list-style-type: none"> <li>云服务器 - 磁盘</li> <li>云服务器 - 挂载点</li> </ul> | 1分钟            |



| 指标                           | 指标名称            | 指标含义                                                                                                                                                                                                                                                                                                                                                                                                                                                               | 取值范围     | 测量对象<br>(维度)                                                                    | 监控周期<br>(原始指标) |
|------------------------------|-----------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|---------------------------------------------------------------------------------|----------------|
| disk_agt_write_requests_rate | (Agent) 磁盘写操作速率 | <p>该指标用于统计每秒向测量对象写数据的请求次数。</p> <p>单位：请求/秒</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第八列数据的变化得出磁盘写操作速率。<br/>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>采集方式（Windows）： <ul style="list-style-type: none"> <li>使用WMI中Win32_PerfFormattedData_PerfDisk_LogicalDisk对象获取磁盘I/O数据。</li> <li>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>高CPU情况下存在获取超时的现象，会导致无法获取监控数据。</li> </ul> </li> </ul> | ≥ 0 请求/秒 | <ul style="list-style-type: none"> <li>云服务器 - 磁盘</li> <li>云服务器 - 挂载点</li> </ul> | 1分钟            |

| 指标             | 指标名称               | 指标含义                                                                                                                                                                                                                                                                                             | 取值范围             | 测量对象<br>(维度)                                                                 | 监控周期<br>(原始指标) |
|----------------|--------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|------------------------------------------------------------------------------|----------------|
| disk_readTime  | (Agent)<br>读操作平均耗时 | <p>该指标用于统计测量对象磁盘读操作平均耗时。</p> <p>单位：ms/Count</p> <ul style="list-style-type: none"><li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第七列数据的变化得出磁盘读操作平均耗时。</li></ul> <p>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</p> <ul style="list-style-type: none"><li>采集方式（Windows）：暂不支持。</li></ul>  | ≥ 0 ms/<br>Count | <ul style="list-style-type: none"><li>云服务器 - 磁盘</li><li>云服务器 - 挂载点</li></ul> | 1分钟            |
| disk_writeTime | (Agent)<br>写操作平均耗时 | <p>该指标用于统计测量对象磁盘写操作平均耗时。</p> <p>单位：ms/Count</p> <ul style="list-style-type: none"><li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第十一列数据的变化得出磁盘写操作平均耗时。</li></ul> <p>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</p> <ul style="list-style-type: none"><li>采集方式（Windows）：暂不支持。</li></ul> | ≥ 0 ms/<br>Count | <ul style="list-style-type: none"><li>云服务器 - 磁盘</li><li>云服务器 - 挂载点</li></ul> | 1分钟            |

| 指标                | 指标名称                | 指标含义                                                                                                                                                                                                                                                                         | 取值范围     | 测量对象<br>(维度)                                                                 | 监控周期<br>(原始指标) |
|-------------------|---------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|------------------------------------------------------------------------------|----------------|
| disk_ioUtils      | (Agent)<br>磁盘I/O使用率 | 该指标用于统计测量对象磁盘I/O使用率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第十三列数据的变化得出磁盘I/O使用率。<br/>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li><li>采集方式（Windows）：暂不支持。</li></ul>                                       | 0-100%   | <ul style="list-style-type: none"><li>云服务器 - 磁盘</li><li>云服务器 - 挂载点</li></ul> | 1分钟            |
| disk_queue_length | (Agent)<br>平均队列长度   | 该指标反映指定时间段内磁盘的繁忙程度，可用于评估磁盘I/O性能，指标值越大代表磁盘越繁忙，I/O性能越差。<br>单位：个<br><ul style="list-style-type: none"><li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第十四列数据的变化与指标采集周期时间相除得出指标的值。<br/>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li><li>采集方式（Windows）：暂不支持。</li></ul> | $\geq 0$ | <ul style="list-style-type: none"><li>云服务器 - 磁盘</li><li>云服务器 - 挂载点</li></ul> | 1分钟            |

| 指标                             | 指标名称            | 指标含义                                                                                                                                                                                                                                                                             | 取值范围           | 测量对象<br>(维度)                                                                    | 监控周期<br>(原始指标) |
|--------------------------------|-----------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------|---------------------------------------------------------------------------------|----------------|
| disk_write_bytes_per_operation | (Agent) 平均写操作大小 | <p>该指标用于统计指定时间段内，平均每个写I/O操作传输的字节数。</p> <p>单位：Byte/op</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第十列数据的变化与第八列数据的变化相除得出磁盘平均写操作大小。</li> <li>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>采集方式（Windows）：暂不支持。</li> </ul> | ≥ 0<br>Byte/op | <ul style="list-style-type: none"> <li>云服务器 - 磁盘</li> <li>云服务器 - 挂载点</li> </ul> | 1分钟            |
| disk_read_bytes_per_operation  | (Agent) 平均读操作大小 | <p>该指标用于统计指定时间段内，平均每个读I/O操作传输的字节数。</p> <p>单位：Byte/op</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：<br/>通过计算采集周期内/proc/diskstats中对应设备第六列数据的变化与第四列数据的变化相除得出磁盘平均读操作大小。</li> <li>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>采集方式（Windows）：暂不支持。</li> </ul> | ≥ 0<br>Byte/op | <ul style="list-style-type: none"> <li>云服务器 - 磁盘</li> <li>云服务器 - 挂载点</li> </ul> | 1分钟            |

| 指标                       | 指标名称              | 指标含义                                                                                                                                                                                                                                                                        | 取值范围         | 测量对象<br>(维度)                                                                    | 监控周期<br>(原始指标) |
|--------------------------|-------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------|---------------------------------------------------------------------------------|----------------|
| disk_io_svctm            | (Agent) 平均I/O服务时长 | <p>该指标用于统计指定时间段内，平均每个读或写I/O的操作时长。</p> <p>单位：ms/op</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：通过计算采集周期内/proc/diskstats中对应设备第十三列数据的变化与第四列数据和第八列数据和的变化相除得出磁盘平均I/O时长。<br/>挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。</li> <li>采集方式（Windows）：暂不支持。</li> </ul> | ≥ 0<br>ms/op | <ul style="list-style-type: none"> <li>云服务器 - 磁盘</li> <li>云服务器 - 挂载点</li> </ul> | 1分钟            |
| disk_device_used_percent | 块设备使用率            | <p>该指标用于统计测量对象物理磁盘使用率，以百分比为单位。计算方式为：所有已挂载磁盘分区已用存储量/磁盘存储总量。</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：通过汇总每个挂载点的磁盘使用量，再通过磁盘扇区大小和扇区数量计算出磁盘总大小，计算出整体磁盘使用率</li> <li>采集方式（Windows）：暂不支持。</li> </ul>                                                                 | 0-100%       | 云服务器 - 磁盘                                                                       | 1分钟            |

## 操作系统监控指标：文件系统

表 16-8 文件系统类监控指标说明

| 指标                     | 指标名称               | 指标含义                                                                                                                        | 取值范围                                                                    | 测量对象<br>(维度) | 监控周期<br>(原始指标) |
|------------------------|--------------------|-----------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|--------------|----------------|
| disk_fs_rwstate        | (Agent) 文件系统读写状态   | 该指标用于统计测量对象挂载文件系统的读写状态。状态分为：可读写（0）/只读（1）。<br>采集方式（Linux）：通过读取/proc/mounts中第四列文件系统挂载参数获得。                                    | <ul style="list-style-type: none"> <li>0: 可读写</li> <li>1: 只读</li> </ul> | 云服务器 - 挂载点   | 1分钟            |
| disk_inodesTotal       | (Agent) inode空间大小  | 该指标用于统计测量对象当前磁盘的inode空间量。<br>采集方式（Linux）：执行df -i命令，查看Inodes列数据。挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。            | $\geq 0$                                                                | 云服务器 - 挂载点   | 1分钟            |
| disk_inodesUsed        | (Agent) inode已使用空间 | 该指标用于统计测量对象当前磁盘已使用的inode空间量。<br>采集方式（Linux）：执行df -i命令，查看IUsed列数据。挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。          | $\geq 0$                                                                | 云服务器 - 挂载点   | 1分钟            |
| disk_inodesUsedPercent | (Agent) inode已使用占比 | 该指标用于统计测量对象当前磁盘已使用的inode占比。<br>单位：百分比<br>采集方式（Linux）：执行df -i命令，查看IUse%列数据。挂载点前缀路径长度不能超过64个字符，必须以字母开头，只能包含0-9/a-z/A-Z/-/./~。 | 0-100%                                                                  | 云服务器 - 挂载点   | 1分钟            |

 说明

Windows系统暂不支持文件系统类监控指标。

## 操作系统监控指标：网卡

表 16-9 网卡相关监控指标说明

| 指标             | 指标名称            | 指标含义                                                                                                                                                                                | 取值范围              | 测量对象（维度） | 监控周期（原始指标） |
|----------------|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------|------------|
| net_bitRecv    | (Agent) 出网带宽    | 该指标用于统计测量对象网卡每秒发送的比特数。<br>单位：bit/s<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li><li>采集方式（Windows）：使用WMI中MibIfRow对象获取网络指标数据。</li></ul>     | $\geq 0$ bit/s    | 云服务器     | 1分钟        |
| net_bitSent    | (Agent) 入网带宽    | 该指标用于统计测量对象网卡每秒接收的比特数。<br>单位：bit/s<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li><li>采集方式（Windows）：使用WMI中MibIfRow对象获取网络指标数据。</li></ul>     | $\geq 0$ bit/s    | 云服务器     | 1分钟        |
| net_packetRecv | (Agent) 网卡包接收速率 | 该指标用于统计测量对象网卡每秒接收的数据包数。<br>单位：Counts/s<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li><li>采集方式（Windows）：使用WMI中MibIfRow对象获取网络指标数据。</li></ul> | $\geq 0$ Counts/s | 云服务器     | 1分钟        |

| 指标             | 指标名称            | 指标含义                                                                                                                                                                                | 取值范围            | 测量对象（维度） | 监控周期（原始指标） |
|----------------|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------|----------|------------|
| net_packetSent | (Agent) 网卡包发送速率 | 该指标用于统计测量对象网卡每秒发送的数据包数。<br>单位：Counts/s<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li><li>采集方式（Windows）：使用WMI中MibIfRow对象获取网络指标数据。</li></ul> | ≥ 0<br>Counts/s | 云服务器     | 1分钟        |
| net_errin      | (Agent) 接收误包率   | 该指标用于统计测量对象网卡每秒接收的错误数据包数量占所接收的数据包的比率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li><li>采集方式（Windows）：暂不支持。</li></ul>            | 0-100%          | 云服务器     | 1分钟        |
| net_errout     | (Agent) 发送误包率   | 该指标用于统计测量对象网卡每秒发送的错误数据包数量占所发送的数据包的比率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li><li>采集方式（Windows）：暂不支持。</li></ul>            | 0-100%          | 云服务器     | 1分钟        |



| 指标          | 指标名称          | 指标含义                                                                                                                                                                          | 取值范围   | 测量对象（维度） | 监控周期（原始指标） |
|-------------|---------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|----------|------------|
| net_dropin  | (Agent) 接收丢包率 | 该指标用于统计测量对象网卡每秒接收并已丢弃的数据包数量占所接收的数据包的比率。<br>单位：百分比<br><ul style="list-style-type: none"> <li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li> <li>采集方式（Windows）：暂不支持。</li> </ul> | 0-100% | 云服务器     | 1分钟        |
| net_dropout | (Agent) 发送丢包率 | 该指标用于统计测量对象网卡每秒发送并已丢弃的数据包数量占所发送的数据包的比率。<br>单位：百分比<br><ul style="list-style-type: none"> <li>采集方式（Linux）：通过计算采集周期内/proc/net/dev中的变化得出。</li> <li>采集方式（Windows）：暂不支持。</li> </ul> | 0-100% | 云服务器     | 1分钟        |

## 操作系统监控指标：NTP

表 16-10 NTP 类监控指标说明

| 指标         | 指标名称           | 指标含义                                                                       | 取值范围   | 测量对象（维度） | 监控周期（原始指标） |
|------------|----------------|----------------------------------------------------------------------------|--------|----------|------------|
| ntp_offset | (Agent) NTP偏移量 | 该指标用于统计测量对象当前NTP偏移量。<br>单位：ms<br>采集方式（Linux）：执行chronyc sources -v命令，获取偏移量。 | ≥ 0 ms | 云服务器     | 1分钟        |

## 操作系统监控指标：TCP

表 16-11 TCP 类监控指标说明

| 指标                  | 指标名称                    | 指标含义                                                                                                                                                                                                           | 取值范围 | 测量对象（维度） | 监控周期（原始指标） |
|---------------------|-------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|----------|------------|
| net_tcp_total       | (Agent) TCP TOTAL       | 该指标用于统计测量对象所有状态的TCP连接数总和。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>           | ≥ 0  | 云服务器     | 1分钟        |
| net_tcp_established | (Agent) TCP ESTABLISHED | 该指标用于统计测量对象处于ESTABLISHED状态的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul> | ≥ 0  | 云服务器     | 1分钟        |
| net_tcp_sys_sent    | (Agent) TCP SYS_SENT    | 该指标用于统计测量对象处于请求连接状态的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>        | ≥ 0  | 云服务器     | 1分钟        |

| 指标                | 指标名称                            | 指标含义                                                                                                                                                                                                              | 取值范围     | 测量对象（维度） | 监控周期（原始指标） |
|-------------------|---------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|----------|------------|
| net_tcp_sys_recv  | (Agent)<br>TCP<br>SYS_RECV      | 该指标用于统计测量对象服务器端收到的请求连接的TCP数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>          | $\geq 0$ | 云服务器     | 1分钟        |
| net_tcp_fin_wait1 | (Agent)<br>TCP<br>FIN_WAIT<br>1 | 该指标用于统计测量对象客户端主动关闭且没有收到服务端ACK的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul> | $\geq 0$ | 云服务器     | 1分钟        |
| net_tcp_fin_wait2 | (Agent)<br>TCP<br>FIN_WAIT<br>2 | 该指标用于统计测量对象处于FIN_WAIT2状态的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>      | $\geq 0$ | 云服务器     | 1分钟        |

| 指标                 | 指标名称                         | 指标含义                                                                                                                                                                                                          | 取值范围 | 测量对象（维度） | 监控周期（原始指标） |
|--------------------|------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|----------|------------|
| net_tcp_time_wait  | (Agent)<br>TCP<br>TIME_WAIT  | 该指标用于统计测量对象处于TIME_WAIT状态的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>  | ≥ 0  | 云服务器     | 1分钟        |
| net_tcp_close      | (Agent)<br>TCP<br>CLOSE      | 该指标用于统计测量对象关闭的或未打开的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>        | ≥ 0  | 云服务器     | 1分钟        |
| net_tcp_close_wait | (Agent)<br>TCP<br>CLOSE_WAIT | 该指标用于统计测量对象处于CLOSE_WAIT状态的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul> | ≥ 0  | 云服务器     | 1分钟        |

| 指标               | 指标名称                       | 指标含义                                                                                                                                                                                                             | 取值范围 | 测量对象（维度） | 监控周期（原始指标） |
|------------------|----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|----------|------------|
| net_tcp_last_ack | (Agent)<br>TCP<br>LAST_ACK | 该指标用于统计测量对象被动关闭等待ACK报文的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>       | ≥ 0  | 云服务器     | 1分钟        |
| net_tcp_listen   | (Agent)<br>TCP<br>LISTEN   | 该指标用于统计测量对象处于LISTEN状态的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul>        | ≥ 0  | 云服务器     | 1分钟        |
| net_tcp_closing  | (Agent)<br>TCP<br>CLOSING  | 该指标用于统计测量对象处于服务端和客户端同时主动关闭状态的TCP连接数量。<br>单位：Count<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过/proc/net/tcp文件获取到所有状态的TCP连接，再统计每个状态的连接数量。</li><li>采集方式（Windows）：通过WindowsAPI GetTcpTable2获取。</li></ul> | ≥ 0  | 云服务器     | 1分钟        |

| 指标              | 指标名称           | 指标含义                                                                                                                                                                                                                      | 取值范围   | 测量对象（维度） | 监控周期（原始指标） |
|-----------------|----------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|----------|------------|
| net_tcp_retrans | (Agent) TCP重传率 | 该指标用于统计测量对象重新发送的报文数与总发送的报文数之间的比值。<br>单位：百分比<br><ul style="list-style-type: none"> <li>采集方式（Linux）：通过从/proc/net/snmp文件中获取对应的数据，计算采集周期内发送包数和重传包数的比值得出。</li> <li>采集方式（Windows）：重传率通过WindowsAPI GetTcpStatistics获取</li> </ul> | 0-100% | 云服务器     | 1分钟        |

## 操作系统监控指标：GPU

表 16-12 GPU 类监控指标说明

| 指标         | 指标名称    | 指标含义                                                                                                                                                                                        | 取值范围                                                                                      | 测量对象（维度）                                                                   | 监控周期（原始指标） |
|------------|---------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------|----------------------------------------------------------------------------|------------|
| gpu_status | gpu健康状态 | 该指标用于统计虚拟机上GPU健康状态，是一个综合指标。<br>该指标无单位。<br><ul style="list-style-type: none"> <li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li> <li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li> </ul> | <ul style="list-style-type: none"> <li>0：代表健康</li> <li>1：代表亚健康</li> <li>2：代表故障</li> </ul> | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - GPU</li> </ul> | 1分钟        |

| 指标                       | 指标名称       | 指标含义                                                                                                                                                                                           | 取值范围   | 测量对象（维度）                                                                | 监控周期（原始指标） |
|--------------------------|------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|-------------------------------------------------------------------------|------------|
| gpu_usage_encoder        | 编码使用率      | 该指标用于统计该GPU的编码能力使用率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>               | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_usage_decoder        | 解码使用率      | 该指标用于统计该GPU的解码能力使用率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>               | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_volatile_correctable | 可纠正ECC错误数量 | 该指标用于统计该GPU重置以来可纠正的ECC错误数量，每次重置后归0。<br>单位：个。<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul> | ≥ 0    | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |

| 指标                          | 指标名称          | 指标含义                                                                                                                                                                                           | 取值范围     | 测量对象（维度）                                                                | 监控周期（原始指标） |
|-----------------------------|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|-------------------------------------------------------------------------|------------|
| gpu_volatile_uncorrectable  | 不可纠正ECC错误数量   | 该指标用于统计该GPU重置以来不可纠正的ECC错误数量，每次重置后归0。<br>单位：个<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul> | $\geq 0$ | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_aggregate_correctable   | 累计可纠正ECC错误数量  | 该指标用于统计该GPU累计的可纠正ECC错误数量。<br>单位：个<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>            | $\geq 0$ | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_aggregate_uncorrectable | 累计不可纠正ECC错误数量 | 该指标用于统计该GPU累计的不可纠正ECC错误数量。<br>单位：个<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>           | $\geq 0$ | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |



| 指标                          | 指标名称                         | 指标含义                                                                                                                                                                                   | 取值范围                                                                                                                   | 测量对象（维度）                                                     | 监控周期（原始指标） |
|-----------------------------|------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------|------------|
| gpu_retired_page_single_bit | retired page single bit 错误数量 | 该指标用于统计该GPU当前卡隔离的单比特页的数量。<br>单位：个<br><ul style="list-style-type: none"> <li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li> <li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li> </ul> | ≥ 0                                                                                                                    | <ul style="list-style-type: none"> <li>云服务器 - GPU</li> </ul> | 1分钟        |
| gpu_retired_page_double_bit | retired page double bit 错误数量 | 该指标用于统计该GPU当前卡隔离的双比特页的数量。<br>单位：个<br><ul style="list-style-type: none"> <li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li> <li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li> </ul> | ≥ 0                                                                                                                    | <ul style="list-style-type: none"> <li>云服务器 - GPU</li> </ul> | 1分钟        |
| gpu_performance_state       | (Agent) 性能状态                 | 该指标用于统计测量对象当前的GPU性能状态。<br>该指标无单位。<br><ul style="list-style-type: none"> <li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li> <li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li> </ul> | P0-P15、P32,<br><ul style="list-style-type: none"> <li>P0：表示最大性能状态</li> <li>P15：表示最小性能状态</li> <li>P32：表示状态未知</li> </ul> | <ul style="list-style-type: none"> <li>云服务器 - GPU</li> </ul> | 1分钟        |

| 指标            | 指标名称           | 指标含义                                                                                                                                                                               | 取值范围   | 测量对象（维度）                                                                | 监控周期（原始指标） |
|---------------|----------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|-------------------------------------------------------------------------|------------|
| gpu_usage_mem | (Agent) 显存使用率  | 该指标用于统计测量对象当前的显存使用率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>   | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_usage_gpu | (Agent) GPU使用率 | 该指标用于统计测量对象当前的GPU使用率。<br>单位：百分比<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>  | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_free_mem  | GPU显存剩余量       | 该指标用于统计测量对象当前的GPU显存剩余量。<br>单位：MB<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul> | ≥ 0 MB | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |

| 指标                  | 指标名称      | 指标含义                                                                                                                                                                                      | 取值范围    | 测量对象（维度）                                                                | 监控周期（原始指标） |
|---------------------|-----------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------|-------------------------------------------------------------------------|------------|
| gpu_graphics_clocks | GPU显卡时钟频率 | 该指标用于统计测量对象当前的GPU显卡（着色器）时钟频率。<br>单位：MHz<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul> | ≥ 0 MHz | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_memory_clocks   | GPU内存时钟频率 | 该指标用于统计测量对象当前的GPU内存时钟频率。<br>单位：MHz<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>      | ≥ 0 MHz | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_power_draw      | GPU功率     | 该指标用于统计测量对象当前的GPU功率。<br>单位：W<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>            | NA      | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |

| 指标                    | 指标名称          | 指标含义                                                                                                                                                                                        | 取值范围             | 测量对象（维度）                                                                | 监控周期（原始指标） |
|-----------------------|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------|-------------------------------------------------------------------------|------------|
| gpu_rx_throughput_pci | GPU PCI 入方向带宽 | 该指标用于统计测量对象当前的GPU PCI入方向带宽。<br>单位：MByte/s<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul> | $\geq 0$ MByte/s | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_sm_clocks         | GPU流式处理器时钟频率  | 该指标用于统计测量对象当前的GPU流式处理器时钟频率。<br>单位：MHz<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>     | $\geq 0$ MHz     | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_temperature       | GPU温度         | 该指标用于统计测量对象当前的GPU温度。<br>单位：℃<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>              | $\geq 0$ °C      | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |

| 指标                    | 指标名称          | 指标含义                                                                                                                                                                                        | 取值范围        | 测量对象（维度）                                                                | 监控周期（原始指标） |
|-----------------------|---------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-------------------------------------------------------------------------|------------|
| gpu_tx_throughput_pci | GPU PCI 出方向带宽 | 该指标用于统计测量对象当前的GPU PCI出方向带宽。<br>单位：MByte/s<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul> | ≥ 0 MByte/s | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_used_mem          | GPU显存使用量      | 该指标用于统计测量对象当前的GPU显存使用量。<br>单位：MB<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul>          | ≥ 0 MB      | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |
| gpu_video_clocks      | GPU视频时钟频率     | 该指标用于统计测量对象当前的GPU视频（包含编解码）时钟频率。<br>单位：MHz<br><ul style="list-style-type: none"><li>采集方式（Linux）：通过调用GPU卡的libnvidia-ml.so.1库文件获取。</li><li>采集方式（Windows）：通过调用GPU卡的nvidia-smi.dll库获取。</li></ul> | ≥ 0 MHz     | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - GPU</li></ul> | 1分钟        |

## 操作系统监控指标：NPU

表 16-13 NPU 类监控指标说明

| 指标                    | 指标名称        | 指标含义                                                                              | 取值范围                                                                                                            | 测量对象（维度）                                                                | 监控周期（原始指标） |
|-----------------------|-------------|-----------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------|------------|
| npu_device_health     | NPU健康状况     | 该指标用于统计虚拟机上NPU卡的健康状态，是一个综合指标。<br>该指标无单位。<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。 | <ul style="list-style-type: none"><li>0：代表健康</li><li>1：代表存在一般告警</li><li>2：代表存在重要告警</li><li>3：代表存在紧急告警</li></ul> | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_util_rate_mem     | NPU显存使用率    | 该指标用于统计该NPU的编码能力使用率。<br>单位：百分比<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。           | 0-100%                                                                                                          | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_util_rate_ai_core | NPU卡AI核心使用率 | 该指标用于统计该NPU的AI核心使用率。<br>单位：百分比<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。           | 0-100%                                                                                                          | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |

| 指标                          | 指标名称          | 指标含义                                                                          | 取值范围   | 测量对象（维度）                                                                | 监控周期（原始指标） |
|-----------------------------|---------------|-------------------------------------------------------------------------------|--------|-------------------------------------------------------------------------|------------|
| npu_util_rate_ai_cpu        | NPU卡AI CPU使用率 | 该指标用于统计该NPU的AI CPU的使用率。<br>单位：百分比。<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。   | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_util_rate_ctrl_cpu      | NPU控制 CPU使用率  | 该指标用于统计该NPU的控制CPU的使用率。<br>单位：百分比。<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。    | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_util_rate_mem_bandwidth | NPU显存带宽使用率    | 该指标用于统计该NPU的显存的带宽使用率。<br>单位：百分比。<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。     | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_freq_mem                | NPU显存频率       | 该指标用于统计该NPU的显存的时钟频率。<br>单位：兆赫兹（MHz）。<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。 | ≥ 0    | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |

| 指标               | 指标名称        | 指标含义                                                                            | 取值范围 | 测量对象（维度）                                                                | 监控周期（原始指标） |
|------------------|-------------|---------------------------------------------------------------------------------|------|-------------------------------------------------------------------------|------------|
| npu_freq_ai_core | NPU卡AI核心频率  | 该指标用于统计该NPU AI核心的时钟频率。<br>单位：兆赫兹（MHz）。<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。 | ≥ 0  | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_usage_memory | NPU显存使用量    | 该指标用于统计该NPU 显存的使用量。<br>单位：兆Byte（MB）。<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。   | ≥ 0  | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_sbe          | NPU单bit错误数量 | 该指标用于统计该NPU卡当前的单比特页错误的数量。<br>单位：个<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。      | ≥ 0  | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |
| npu_dbe          | NPU双bit错误数量 | 该指标用于统计该NPU卡当前的多比特页错误的数量。<br>单位：个<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。      | ≥ 0  | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - NPU</li></ul> | 1分钟        |



| 指标              | 指标名称  | 指标含义                                                                                           | 取值范围 | 测量对象（维度）                                                                   | 监控周期（原始指标） |
|-----------------|-------|------------------------------------------------------------------------------------------------|------|----------------------------------------------------------------------------|------------|
| npu_power       | NPU功率 | 该指标用于统计该NPU卡的功率。其中，310卡仅支持显示额定功率，其余卡显示实际功率<br>单位：瓦（W）<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。 | ≥ 0  | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - NPU</li> </ul> | 1分钟        |
| npu_temperature | NPU温度 | 该指标用于统计该NPU卡当前的温度<br>单位：摄氏度（℃）<br>采集方式（Linux）：通过调用NPU卡的libdcmi.so库文件获取。                        | ≥ 0  | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - NPU</li> </ul> | 1分钟        |

### 说明

Windows系统暂不支持NPU类监控指标。

## 操作系统监控指标：DAVP

表 16-14 DAVP 类监控指标说明

| 指标                 | 指标名称     | 指标含义                                                                                           | 取值范围                                                                     | 测量对象（维度）                                                                    | 监控周期（原始指标） |
|--------------------|----------|------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------|-----------------------------------------------------------------------------|------------|
| davp_device_health | DAVP健康状态 | 该指标用于统计虚拟机上DAVP卡的健康状态，是一个综合指标。<br>该指标无单位。<br>采集方式（Linux）：通过调用DAVP卡的vatools工具下的libvaml.so库文件获取。 | <ul style="list-style-type: none"> <li>0：代表健康</li> <li>1：代表异常</li> </ul> | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - DAVP</li> </ul> | 1分钟        |

| 指标                       | 指标名称           | 指标含义                                                                                    | 取值范围   | 测量对象（维度）                                                                 | 监控周期（原始指标） |
|--------------------------|----------------|-----------------------------------------------------------------------------------------|--------|--------------------------------------------------------------------------|------------|
| davp_util_rate_mem       | DAVP显存使用率      | 该指标用于统计该DAVP卡的显存使用率。<br>单位：百分比<br>采集方式（Linux）：通过调用DAVP卡的vtools工具下的libvaml.so库文件获取。      | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - DAVP</li></ul> | 1分钟        |
| davp_usage_mem           | DAVP显存使用量      | 该指标用于统计该DAVP卡的显存使用量<br>单位：MB<br>采集方式（Linux）：通过调用DAVP卡的vtools工具下的libvaml.so库文件获取。        | ≥ 0    | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - DAVP</li></ul> | 1分钟        |
| davp_util_rate_ai_core   | DAVP卡AI核心使用率   | 该指标用于统计该DAVP卡的AI核心的使用率。<br>单位：百分比。<br>采集方式（Linux）：通过调用DAVP卡的vtools工具下的libvaml.so库文件获取。  | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - DAVP</li></ul> | 1分钟        |
| davp_util_rate_vdsp_core | DAVP卡vdsp核心使用率 | 该指标用于统计该DAVP卡的vdsp核心使用率。<br>单位：百分比。<br>采集方式（Linux）：通过调用DAVP卡的vtools工具下的libvaml.so库文件获取。 | 0-100% | <ul style="list-style-type: none"><li>云服务器</li><li>云服务器 - DAVP</li></ul> | 1分钟        |

| 指标                         | 指标名称         | 指标含义                                                                                    | 取值范围   | 测量对象（维度）                                                                    | 监控周期（原始指标） |
|----------------------------|--------------|-----------------------------------------------------------------------------------------|--------|-----------------------------------------------------------------------------|------------|
| davp_util_rate_encode_core | DAVP卡编码核心使用率 | 该指标用于统计该DAVP卡的编码核心使用率。<br>单位：百分比。<br>采集方式（Linux）：通过调用DAVP卡的vtools工具下的libvaml.so库文件获取。   | 0-100% | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - DAVP</li> </ul> | 1分钟        |
| davp_util_rate_decode_core | DAVP卡解码核心使用率 | 该指标用于统计该DAVP卡的解码核心使用率。<br>单位：百分比。<br>采集方式（Linux）：通过调用DAVP卡的vtools工具下的libvaml.so库文件获取。   | 0-100% | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - DAVP</li> </ul> | 1分钟        |
| davp_sysc_temperature      | DAVP卡系统模块温度  | 该指标用于统计该DAVP卡的系统模块温度。<br>单位：摄氏度（℃）。<br>采集方式（Linux）：通过调用DAVP卡的vtools工具下的libvaml.so库文件获取。 | ≥ 0    | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - DAVP</li> </ul> | 1分钟        |

### 说明

Windows系统暂不支持DAVP类监控指标。

### 维度

| 维度         | Key         | Value       |
|------------|-------------|-------------|
| 云服务器       | instance_id | 云服务器ID。     |
| 云服务器 - 磁盘  | disk        | 云服务器磁盘。     |
| 云服务器 - 挂载点 | mount_point | 云服务器磁盘的挂载点。 |

| 维度          | Key  | Value                                 |
|-------------|------|---------------------------------------|
| 云服务器 - GPU  | gpu  | GPU类型云服务器中显卡。                         |
| 云服务器 - NPU  | npu  | NPU类型云服务器中显卡。                         |
| 云服务器 - DAVP | davp | DAVP类型云服务器，其中搭载了DaoCloud DAVP1 视频加速卡。 |

## 16.4 弹性云服务器支持的进程监控指标（安装 Agent）

### 功能说明

通过在弹性云服务器中安装Agent插件，可以对主机内的活跃进程进行监控，默认采集活跃进程消耗的CPU、内存，以及打开的文件数量等信息。

本节定义了弹性云服务器上报云监控的进程监控指标。

### 命名空间

AGT.ECS

### 进程监控指标说明

对于不同的操作系统、不同的弹性云服务器类型，在安装Agent后均默认支持查看以下进程监控指标。

表 16-15 进程监控指标说明

| 指标               | 指标名称   | 指标含义                                                                                                                                                                                                             | 取值范围       | 测量对象（维度） | 监控周期（原始指标） |
|------------------|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------|----------|------------|
| proc_pHashId_cpu | CPU使用率 | <p>进程消耗的CPU百分比，pHashId是（进程名+进程ID）的md5值。</p> <p>单位：百分比</p> <ul style="list-style-type: none"> <li>采集方式（Linux）：通过计算/proc/pid/stat的变化得出。</li> <li>采集方式（Windows）：通过Windows API GetProcessTimes获取进程CPU使用率。</li> </ul> | 0-1*CPU核心数 | 云服务器     | 1分钟        |

| 指标                  | 指标名称          | 指标含义                                                                                                                                                                                                                                                                                                                                                                                 | 取值范围   | 测量对象<br>(维度) | 监控周期(原始指标) |
|---------------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------------|------------|
| proc_pHashId_memory | 内存使用率         | <p>进程消耗的内存百分比, pHashId是(进程名+进程ID)的md5值。</p> <p>单位: 百分比</p> <ul style="list-style-type: none"> <li>采集方式(Linux):<br/>RSS*PAGESIZE/<br/>MemTotal<br/>RSS: 通过获取/proc/pid/statm第二列得到<br/>PAGESIZE: 通过命令getconf PAGESIZE获取<br/>MemTotal: 通过/proc/meminfo获取</li> <li>采集方式(Windows): 使用Windows API procGlobalMemoryStatusEx获取内存总量, 通过GetProcessMemoryInfo获取内存已使用量, 计算两者比值得到内存使用率。</li> </ul> | 0-100% | 云服务器         | 1分钟        |
| proc_pHashId_file   | 打开文件数         | <p>进程打开文件数, pHashId是(进程名+进程ID)的md5值。</p> <ul style="list-style-type: none"> <li>采集方式(Linux): 通过执行ls -l /proc/pid/fd可以查看数量。</li> <li>采集方式(Windows): 暂不支持。</li> </ul>                                                                                                                                                                                                                  | ≥0     | 云服务器         | 1分钟        |
| proc_running_count  | (Agent)运行中进程数 | <p>该指标用于统计测量对象处于运行状态的进程数。</p> <ul style="list-style-type: none"> <li>采集方式(Linux): 通过统计 /proc/pid/status 中Status值获取每个进程的状态, 进而统计各个状态进程总数。</li> <li>采集方式(Windows): 暂不支持。</li> </ul>                                                                                                                                                                                                    | ≥0     | 云服务器         | 1分钟        |

| 指标                  | 指标名称          | 指标含义                                                                                                                                                                   | 取值范围 | 测量对象<br>(维度) | 监控周期(原始指标) |
|---------------------|---------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|--------------|------------|
| proc_idle_count     | (Agent) 空闲进程数 | 该指标用于统计测量对象处于空闲状态的进程数。 <ul style="list-style-type: none"><li>采集方式 (Linux)：通过统计 /proc/pid/status 中Status值获取每个进程的状态，进而统计各个状态进程总数。</li><li>采集方式 (Windows)：暂不支持。</li></ul> | ≥0   | 云服务器         | 1分钟        |
| proc_zombie_count   | (Agent) 僵死进程数 | 该指标用于统计测量对象处于僵死状态的进程数。 <ul style="list-style-type: none"><li>采集方式 (Linux)：通过统计 /proc/pid/status 中Status值获取每个进程的状态，进而统计各个状态进程总数。</li><li>采集方式 (Windows)：暂不支持。</li></ul> | ≥0   | 云服务器         | 1分钟        |
| proc_blocked_count  | (Agent) 阻塞进程数 | 该指标用于统计测量对象被阻塞的进程数。 <ul style="list-style-type: none"><li>采集方式 (Linux)：通过统计 /proc/pid/status 中Status值获取每个进程的状态，进而统计各个状态进程总数。</li><li>采集方式 (Windows)：暂不支持。</li></ul>    | ≥0   | 云服务器         | 1分钟        |
| proc_sleeping_count | (Agent) 睡眠进程数 | 该指标用于统计测量对象处于睡眠状态的进程数。 <ul style="list-style-type: none"><li>采集方式 (Linux)：通过统计 /proc/pid/status 中Status值获取每个进程的状态，进而统计各个状态进程总数。</li><li>采集方式 (Windows)：暂不支持。</li></ul> | ≥0   | 云服务器         | 1分钟        |

| 指标                   | 指标名称          | 指标含义                                                                                                                                                                                             | 取值范围 | 测量对象 (维度)                                                                 | 监控周期 (原始指标) |
|----------------------|---------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------|---------------------------------------------------------------------------|-------------|
| proc_total_count     | (Agent) 系统进程数 | 该指标用于统计测量对象的总进程数。<br><ul style="list-style-type: none"> <li>采集方式 (Linux)：通过统计 /proc/pid/status 中Status值获取每个进程的状态，进而统计各个状态进程总数。</li> <li>采集方式 (Windows)：通过 psapi.dll系统进程状态支持模块得到进程总数。</li> </ul>  | ≥0   | 云服务器                                                                      | 1分钟         |
| proc_specified_count | (Agent) 指定进程数 | 该指标用于统计测量对象指定的进程数。<br><ul style="list-style-type: none"> <li>采集方式 (Linux)：通过统计 /proc/pid/status 中Status值获取每个进程的状态，进而统计各个状态进程总数。</li> <li>采集方式 (Windows)：通过 psapi.dll系统进程状态支持模块得到进程总数。</li> </ul> | ≥0   | <ul style="list-style-type: none"> <li>云服务器</li> <li>云服务器 - 进程</li> </ul> | 1分钟         |

## 维度

| 维度        | Key         | Value    |
|-----------|-------------|----------|
| 云服务器      | instance_id | 云服务器ID。  |
| 云服务器 - 进程 | proc        | 云服务器的进程。 |


## 16.5 设置告警规则

### 操作场景

通过设置弹性云服务器告警规则，用户可自定义监控目标与通知策略，及时了解弹性云服务器运行状况，从而起到预警作用。

设置弹性云服务器的告警规则包括设置告警规则名称、监控对象、监控指标、告警阈值、监控周期和是否发送通知等参数。本节介绍了设置弹性云服务器告警规则的具体方法。

## 操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 选择“管理与监管 > 云监控服务”。
4. 在左侧导航树栏，选择“告警 > 告警规则”。
5. 在“告警规则”界面，单击“创建告警规则”进行添加，或者选择已有的告警规则进行修改，设置弹性云服务器的告警规则。

以修改弹性云服务器的告警规则为例，介绍如何设置告警规则：

- a. 单击待修改的告警规则名称，进入详情页。
- b. 单击右上角的“修改”。
- c. 在“修改告警规则”界面，根据界面提示配置参数。
- d. 单击“确定”。

弹性云服务器告警规则设置完成后，当符合规则的告警产生时，系统会自动进行通知。

### 说明

更多关于弹性云服务器监控规则的信息，请参见《[云监控用户指南](#)》。

## 16.6 查看监控指标

### 操作场景

云服务平台提供的云监控，可以对弹性云服务器的运行状态进行日常监控。您可以通过管理控制台，直观地查看弹性云服务器的各项监控指标。

由于监控数据的获取与传输会花费一定时间，因此，云监控显示的是当前时间5~10分钟前的弹性云服务器状态。如果您的弹性云服务器刚刚创建完成，请等待5~10分钟后查看监控数据。

### 前提条件

- 弹性云服务器正常运行。  
关机、故障、删除状态的弹性云服务器，无法在云监控中查看其监控指标。当弹性云服务器再次启动或恢复后，即可正常查看。


### 说明

关机、故障24小时的弹性云服务器，云监控将默认该弹性云服务器不存在，并在监控列表中删除，不再对其进行监控，但告警规则需要用户手动清理。

- 弹性云服务器已对接云监控，即已在云监控页面设置告警规则。  
对接云监控之前，用户无法查看到未对接资源的监控数据。具体操作，请参见[设置告警规则](#)。
- 弹性云服务器已正常运行一段时间（约10分钟）。  
对于新创建的弹性云服务器，需要等待一段时间，才能查看上报的监控数据和监控视图。



## 操作步骤

1. 登录管理控制台。
2. 单击管理控制台左上角的 ，选择区域和项目。
3. 单击“☰”，选择“计算 > 弹性云服务器”。
4. 在弹性云服务器列表中的右上角，输入弹性云服务器名称、IP地址或ID，并进行搜索。
5. 单击弹性云服务器的名称，查看详情。
6. 在弹性云服务器详情页面，选择“监控”页签，查看监控数据。
7. 在弹性云服务器监控区域，您可以通过选择时长，查看对应时间的监控数据。  
当前支持查看弹性云服务器“近1小时”、“近3小时”、“近12小时”、“近24小时”、“近7天”的监控数据。

# 17 使用 CTS 审计 ECS 操作事件

## 17.1 ECS 支持云审计的关键操作

### 操作场景

平台提供了云审计服务。通过云审计服务，您可以记录与云服务器相关的操作事件，便于日后的查询、审计和回溯。

### 前提条件

已开通云审计服务。

### 支持审计的关键操作列表

表 17-1 云审计服务支持的云服务器操作列表

| 操作名称     | 资源类型 | 事件名称                                                                  |
|----------|------|-----------------------------------------------------------------------|
| 创建云服务器   | ecs  | createServer<br>createServerV2<br>createServerV21                     |
| 删除云服务器   | ecs  | deleteServer<br>deleteServerV2<br>deleteServerV21<br>deleteVmByNative |
| 启动云服务器   | ecs  | startServer                                                           |
| 重启云服务器   | ecs  | rebootServer                                                          |
| 关闭云服务器   | ecs  | stopServer                                                            |
| 添加云服务器网卡 | ecs  | addNic                                                                |

| 操作名称                          | 资源类型   | 事件名称                           |
|-------------------------------|--------|--------------------------------|
| 删除云服务器网卡                      | ecs    | deleteNic<br>delNic            |
| 挂载磁盘                          | ecs    | attachVolume<br>attachVolumeV2 |
| 挂载磁盘（EVS页面触发）                 | ecs    | attachVolume                   |
| 卸载磁盘                          | ecs    | detachVolume                   |
| 重装操作系统                        | ecs    | reinstallOs<br>reinitOs        |
| 切换操作系统                        | ecs    | changeOs<br>reinitOs           |
| 变更规格                          | ecs    | resizeServer                   |
| 配置虚拟机自动恢复标签                   | ecs    | addAutoRecovery                |
| 删除虚拟机自动恢复标签                   | ecs    | deleteAutoRecovery             |
| 更新元数据/设置指定key的元数据（不包含委托名称）    | ecs    | updateMetadata                 |
| 更新元数据/设置指定key的元数据（包含委托名称）     | ecs    | updateMetadataAndServerAgency  |
| 更新委托                          | ecs    | updateServerAgency             |
| 获取VNC远程登录                     | ecs    | remoteConsole                  |
| 修改云服务器信息                      | ecs    | updateServer                   |
| 虚拟机状态管理操作（OpenStack Nova API） | server | operateServer                  |

## 17.2 查询 ECS 审计事件

### 操作场景

用户进入云审计服务创建管理类追踪器后，系统开始记录云服务资源的操作。在创建数据类追踪器后，系统开始记录用户对OBS桶中数据的操作。云审计服务管理控制台会保存最近7天的操作记录。


本节介绍如何在云审计服务管理控制台查看或导出最近7天的操作记录。

- [在新版事件列表查看审计事件](#)
- [在旧版事件列表查看审计事件](#)

## 使用限制

- 单账号跟踪的事件可以通过云审计控制台查询。多账号的事件只能在账号自己的事件列表页面去查看，或者到组织追踪器配置的OBS桶中查看，也可以到组织追踪器配置的CTS/system日志流下面去查看。
- 用户通过云审计控制台只能查询最近7天的操作记录。如果需要查询超过7天的操作记录，您必须配置转储到对象存储服务(OBS)或云日志服务(LTS)，才可在OBS桶或LTS日志组里面查看历史事件信息。否则，您将无法追溯7天以前的操作记录。
- 云上操作后，1分钟内可以通过云审计控制台查询管理类事件操作记录，5分钟后才可通过云审计控制台查询数据类事件操作记录。
- 云审计控制台对用户的操作事件日志保留7天，过期自动删除，不支持人工删除。

## 在新版事件列表查看审计事件

1. 登录管理控制台。
2. 单击左上角 ，选择“管理与监管 > 云审计服务 CTS”，进入云审计服务页面。
3. 单击左侧导航树的“事件列表”，进入事件列表信息页面。
4. 事件列表支持通过高级搜索来查询对应的操作事件，您可以在筛选器组合一个或多个筛选条件：
  - 事件名称：输入事件的名称。
  - 事件ID：输入事件ID。
  - 资源名称：输入资源的名称，当该事件所涉及的云资源无资源名称或对应的API接口操作不涉及资源名称参数时，该字段为空。
  - 资源ID：输入资源ID，当该资源类型无资源ID或资源创建失败时，该字段为空。
  - 云服务：在下拉框中选择对应的云服务名称。
  - 资源类型：在下拉框中选择对应的资源类型。
  - 操作用户：在下拉框中选择一个或多个具体的操作用户。
  - 事件级别：可选项为“normal”、“warning”、“incident”，只可选择其中一项。
    - normal：表示操作成功。
    - warning：表示操作失败。
    - incident：表示比操作失败更严重的情况，例如引起其他故障等。
  - 企业项目ID：输入企业项目ID。
  - 访问密钥ID：输入访问密钥ID（包含临时访问凭证和永久访问密钥）。
  - 时间范围：可选择查询最近1小时、最近1天、最近1周的操作事件，也可以自定义最近7天内任意时间段的操作事件。
5. 在事件列表页面，您还可以导出操作记录文件、刷新列表、设置列表展示信息等。
  - 在搜索框中输入任意关键字，按下Enter键，可以在事件列表搜索符合条件的数据。



9. 在需要查看的记录右侧，单击“查看事件”，会弹出一个窗口显示该操作事件结构的详细信息。

×

**查看事件**

```
{
 "request": "",
 "trace_id": "676d4ae3-842b-11ee-9299-9159eee6a3ac",
 "code": "200",
 "trace_name": "createDockerConfig",
 "resource_type": "dockerlogincmd",
 "trace_rating": "normal",
 "api_version": "",
 "message": "createDockerConfig, Method: POST Url=/v2/manage/utills/secret, Reason:",
 "source_ip": " ",
 "domain_id": " ",
 "trace_type": "ApiCall",
 "service_type": "SWR",
 "event_type": "system",
 "project_id": " ",
 "response": "",
 "resource_id": "",
 "tracker_name": "system",
 "time": "2023/11/16 10:54:04 GMT+08:00",
 "resource_name": "dockerlogincmd",
 "user": {
 "domain": {
 "name": " ",
 "id": " "
 }
 }
}
```

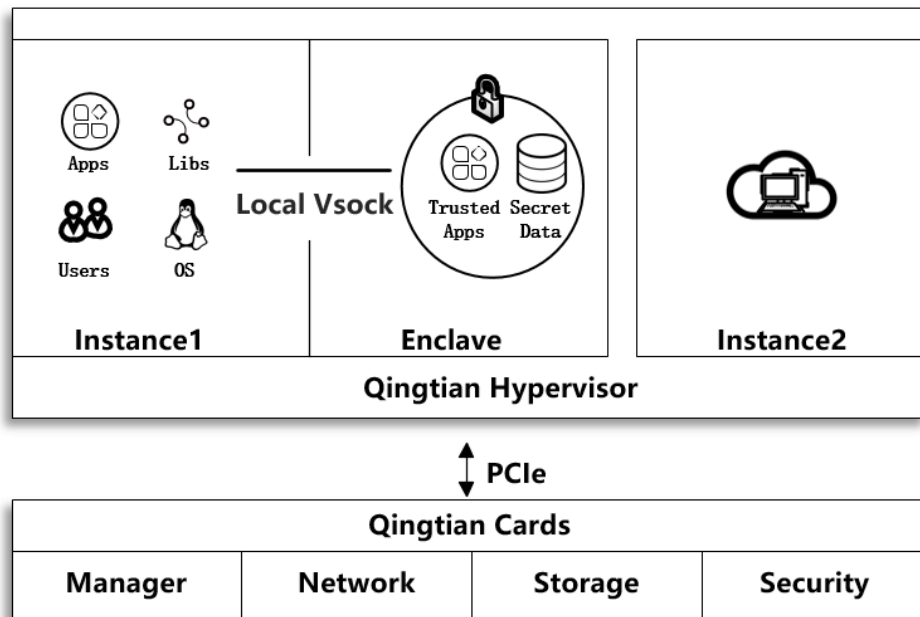
10. 关于事件结构的关键字段详解，请参见《云审计服务用户指南》中的[事件结构](#)和[事件样例](#)。
11. （可选）在旧版事件列表页面，单击右上方的“体验新版”按钮，可切换至新版事件列表页面。

# 18 QingTian Enclave 管理

## 18.1 QingTian Enclave 简介

### 18.1.1 什么是 QingTian Enclave

- 在现有的基于QingTian架构的虚拟机产品中，我们增加了一个置于其内部的，安全的、完全隔离的虚拟机，这个虚拟机我们称之为QingTian Enclave虚拟机，外部对这个QingTian Enclave虚拟机具备所有权的虚拟机，我们称之为父虚拟机。QingTian Enclave是完全独立的虚拟机，无持久化存储、交互式访问或外部网络连接。父虚拟机与QingTian Enclave之间通过安全的本地通道进行通信。即使是父虚拟机上的root用户，也不能访问QingTian Enclave或通过 SSH 连接到QingTian Enclave。
- QingTianHypervisor可以将QingTian Enclave的CPU 和内存与父虚拟机的CPU和内存资源隔离开，提供给您一个完全隔离的可执行环境，显著减少了攻击面。因此，使用QingTian Enclave，您能保护敏感的核心数据和应用程序，为您在QingTian Enclave中运行的服务增加安全保障。
- QingTian Enclave还提供了证明（Attestation）的功能，您可以通过该功能验证QingTian Enclave实例的可信度量值。[华为云密钥管理服务（Key management Service, 简称KMS）](#)为QingTian证明功能提供了内在支持，您能限制应用程序必须在预期的QingTian Enclave运行环境中才能调用KMS API处理敏感数据。



## 约束条件

Qingtian Enclave有以下约束：

| 虚拟机名称                  | 限制要求                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                       |
|------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 父虚拟机（主虚拟机）             | <ol style="list-style-type: none"> <li>1. 至少预留2个vCPU，至少预留512M内存空间。</li> <li>2. 必须为Linux操作系统。</li> </ol>                                                                                                                                                                                                                                                                                                                                                                                                                                                                                    |
| Qingtian Enclave（从虚拟机） | <ol style="list-style-type: none"> <li>1. 裸金属实例不支持Qingtian Enclave。</li> <li>2. 必须为Linux操作系统。</li> <li>3. 启动Qingtian Enclave最低内存规格为128M，并且不能小于Enclave镜像文件（Enclave Image File，简称EIF）文件大小的4倍。</li> <li>4. 当在配置文件中，为Qingtian Enclave配置2M大页的内存管理方式时，Qingtian Enclave启动支持的最大内存为512M。</li> <li>5. 当在配置文件中，为Qingtian Enclave配置1G大页的内存管理方式时，Qingtian Enclave最大内存为256G。</li> <li>6. QingTian Enclave所使用的内存和CPU都必须隔离自同一个NUMA节点。</li> <li>7. vcpu数量配置需要为偶数，最高不超过父虚拟机单NUMA node cpu数量减 2；总数量最高不超过62。</li> <li>8. 在Qingtian Enclave中运行的应用程序需要和OS（内核，ramdisk，init程序）一起被打包成Qingtian Enclave镜像。</li> </ol> |



## 📖 说明

内存和CPU资源隔离的详细内容，请参见[资源隔离](#)。

父虚拟机和QingTian Enclave关系：

1. 每个父虚拟机可以创建最多两个QingTian Enclave。
2. 不支持与父虚拟机共物理内核。
3. 只有在父虚拟机处于运行状态时，QingTian Enclave才处于运行状态。如果父虚拟机被停止或终止，则QingTian Enclave被终止。
4. QingTian Enclave分配的资源（内存和CPU等），都是从父虚拟机中分割出来的，内存区间要求是2M/1G对齐的连续物理区间。

另外还需要注意：

1. 支持QingTian Enclave特性的父虚拟机规格：c7t
2. 支持QingTian Enclave特性的局点：华北-北京四、华东-上海一、华南-广州、亚太-新加坡、土耳其-伊斯坦布尔
3. 如果您在QingTian Enclave中的业务被意外终止，您需要手动重新运行该业务
4. QingTian Enclave的默认配置为使用1G大页，具有1G内存，2vcpu

## 计费标准

在公测期间，使用QingTian Enclave并不会收取额外费用，您只需要支付ECS的购买费用。

## 相关服务

QingTian Enclave与以下华为云服务集成：

1. 密钥管理服务

密钥管理服务（KMS）是华为云数据加密服务族中的一个核心服务。KMS提供可用性高的密钥生成、存储、管理和审计解决方案。KMS密钥由硬件安全模块HSM保护，并与许多华为云数据存储服务集成。您可以借此服务开发自己的安全的数据应用。

2. 华为统一身份认证服务

统一身份认证（Identity and Access Management，简称IAM）是华为云提供权限管理的基础服务，可以帮助您安全地控制华为云服务和资源的访问权限。

## 18.1.2 QingTian Enclave 概念

- QingTian Enclave

QingTian Enclave是完全隔离的虚拟机，它的内存和CPU来自于其父虚拟机被预先隔离出来的资源。QingTian Enclave既没有外部网络，也没有持久化存储。父虚拟机内的进程、应用程序、内核或者用户都无法访问QingTian Enclave中的资源。

- 父虚拟机

父虚拟机是一个能够将其CPU及内存资源隔离分配给QingTian Enclave的ECS实例。这些资源能够在QingTian Enclave的生存周期内被其使用。启动QingTian Enclave成功后，该QingTian Enclave只能与其父虚拟机进行安全通信。

- QingTian Enclave镜像文件  
QingTian Enclave镜像文件 (.eif) 提供了启动QingTian Enclave并在其中运行QingTian Enclave应用程序所需的系统信息，包括Linux操作系统、其他第三方库和QingTian Enclave应用程序。镜像创建详情见[Linux系统上QingTian Enclave应用的开发](#)。
- QingTian命令行工具  
作为华为云QingTian命令行工具 (qt CLI)，在QingTian Enclave使用场景中，qt CLI可以用于创建、关闭和查询QingTian Enclave信息。qt CLI必须在父虚拟机上安装和使用。详情见[QingTian CLI \(qt CLI\)](#)。
- Enclave SDK  
Enclave SDK由一系列开源库组成，以使用户开发自己的QingTian Enclave应用程序。它集成了一些与华为云KMS交互的接口，例如加解密和产生随机数等，并为远程证明提供了内在支持。
- QingTian密码学证明  
QingTian密码学证明是QingTian Enclave在与KMS服务交互时证明自己合法性的过程。它依赖QingTianHypervisor产生的具有数字签名的证明文档。一个QingTian Enclave证明文档内包含的具体信息可以作为第三方服务认证及鉴权的条件。您可以在IAM服务中使用kms:RecipientAttestation相关条件键值 (condition key) 来控制对KMS服务特定接口操作的访问权限，例如生成随机数或者加解密操作。
- 证明文档  
证明文档 (Attestation Document) 由QingTianHypervisor产生并签名，其文件内容为QingTian Enclave的信息，包括PCR、密码摘要以及用户声明。外部服务可以通过证明文件来验证QingTian Enclave的身份是否可信。用户可以利用证明文档构建自己的可信系统，也可以与KMS交互时使用。详情可见[证明文档](#)。
- qt-proxy  
qt-proxy是一个运行在父虚拟机上的网络代理服务。用户可以使用这个服务，使父虚拟机转发来自于QingTian Enclave的网络包，实现QingTian Enclave与外界进行通信。这是QingTian Enclave用来与外部服务交互的唯一通信渠道。
- PCR  
平台配置寄存器 (Platform configuration registers, PCRs) 是QingTian Enclave独有的可信度量值。一部分在QingTian Enclave创建时自动产生，用于验证QingTian Enclave自创建以来的完整性，另一部分可以由用户自己定义以确保QingTian Enclave能够运行在他所希望执行的平台上。另外，Attestation Document包含了相关的PCR，用户可以使用PCR作为IAM访问控制策略的条件键 (condition key)，以实现更为严格的访问控制。详情可见[PCR简介](#)。
- 本地连接通道  
本地连接通道 (Local Vsock Connection) 是QingTian Enclave实例与父虚拟机之间唯一的，安全的本地通信通道。
- QingTian安全模块  
QingTian安全模块 (QingTian Security Module, QTSM) 整体由qtsm-lib函数库和qtsm-server服务组成。您可以在您的QingTian Enclave 应用程序中调用qtsm-lib用户态接口，qtsm-server会处理具体的QTSM请求并将返回相应结果。qtsm-lib提供的用户态接口包括查询指定index对应的PCR值 (qtsm\_describe\_pcr)，扩展指定index的PCR值 (qtsm\_extend\_pcr)，锁定指定index的PCR值 (qtsm\_lock\_pcr)，批量锁定指定index的PCR值 (qtsm\_lock\_pcrs)，获取QTSM信息 (qtsm\_get\_describe) 和获取已签名的证明文档 (qtsm\_get\_attestation)。

## 18.2 快速入门

以下教程将引导您了解并使用QingTian Enclave特性，包括如何启动QingTian Enclave父虚拟机，如何构建QingTian Enclave镜像文件，如何查询正在运行的QingTian Enclave，以及停止QingTian Enclave。

### 1. 准备一台支持QingTian Enclave特性的虚拟机实例

- 在购买虚拟机时，您需要在高级选项中勾选Enclave，并选择Linux作为系统镜像。建议使用HCE2.0镜像。
- 连接到父虚拟机，请参考华为云ECS用户指南。
- 在父虚拟机中安装qt CLI工具，并在配置文件中按需设置资源隔离参数，然后启动资源隔离服务。使用qt CLI工具需要安装必要的python库，详情请见 [QingTian CLI \(qt CLI\)](#)。

- 您可以安装qt CLI工具和其他必要rpm包：

```
yum install qt-enclave-bootstrap
yum install virtio-qtbox
yum install qingtian-tool
```

- 在配置文件中按需配置隔离参数，本教程内使用默认1G内存，2个vCPU。然后启动隔离服务：

```
systemctl start qt-enclave-env
```

### 2. 在父虚拟机中安装Docker程序，我们推荐使用二进制的方式进行Docker安装，可以在Docker官方网站下载需要版本。这里以27.0.1为例：

```
wget https://download.docker.com/linux/static/stable/x86_64/docker-27.0.1.tgz
```

对下载的压缩包进行解压操作：

```
tar xzf docker-27.0.1.tgz
```

解压完成后将docker目录下所有文件复制到/usr/bin目录下：

```
cp docker/* /usr/bin
```

启动docker服务，并将日志等级设置为error等级：

```
dockerd -l error &
```

验证Docker版本：

```
docker version
```

运行hello-world容器，查看Docker是否安装正确：

```
docker run hello-world
```

### 3. 构建QingTian Enclave镜像

本教程中，使用以下**hello\_enclave.sh**脚本作为QingTian Enclave应用程序：

```
#!/bin/bash
while true
do
 echo "hello enclave!"
 sleep 2
done
```

Dockerfile内容如下：

```
FROM ubuntu:latest
COPY hello_enclave.sh /root/hello_enclave.sh
CMD ["/root/hello_enclave.sh"]
```

确认脚本有可执行权限：

```
chmod +x hello_enclave.sh
```

构建一个名为hello-enclave的Docker镜像：



您将观察到终端每隔2秒打印一次hello enclave!。

- 停止正在运行的QingTian Enclave实例

如果您想停止运行以上的QingTian Enclave实例，可以使用以下命令：

```
qt enclave stop --enclave-id 0
stop enclave 0 successfully
{
 "EnclaveID": 0
}
```

## 18.3 典型使用案例

在这个章节中，我们将介绍一个QingTian Enclave的使用场景。其中包含了一些角色和基本的工作流程，同时使用了华为云数据加密服务（DEW）的KMS子服务，华为云统一身份认证服务（IAM）和华为云对象存储服务（OBS）。

### 18.3.1 工作流程

#### 相关角色

在这个QingTian Enclave使用场景里，涉及到了以下几个角色：

- 安全员：拥有对机密数据和华为云KMS密钥的控制权限。作为华为云账号的所有者，它具有最高权限，例如他可以创建IAM用户并为其设置最小权限，创建加密密钥和加密敏感数据等。在本场景中，我们还假设安全员同时承担了Enclave镜像构建的工作，他获得预期的Enclave镜像度量值PCR0和PCR8，并在IAM授权策略中使用这些度量值作为条件键。
- 父虚拟机管理员：作为被安全员授权的运维人员，他拥有父虚拟机实例的访问权限，并能管理QingTian Enclave实例的生命周期。他使用安全员构建好的QingTian Enclave镜像文件来启动QingTian Enclave实例。
- QingTian Enclave应用程序开发人员：开发在QingTian Enclave中运行的应用程序。在该案例中，该应用程序需要从华为OBS服务的Bucket1获取密文对象，然后调用kms-decrypt接口对密文进行解密，进行数据处理后，最后将结果输出到Bucket2。

#### 数据和环境准备

接下来，我们将概述一下关于数据加密、证明设置和QingTian Enclave实例创建的相关流程。

1. 安全员在华为云数据加密服务的密钥管理子服务中创建KMS密钥。
2. 安全员使用KMS密钥加密一段敏感数据（例如银行卡信息）。详细信息可以参考华为云数据加密服务帮助文档中API参考章节中的应用示例：[加解密小量数据](#)。
3. 安全员使用obsutil命令行工具将加密后的密文上传到华为云对象存储服务（OBS）的一个桶中。详细信息可以参考华为云对象存储服务帮助文档工具指南章节：[上传对象](#)。
4. 安全员编写并打包QingTian Enclave开发程序：制作Docker镜像，然后使用qt命令行工具将Docker镜像转换成QingTian Enclave镜像文件。可以参考QingTian Enclave镜像文件制作。并记录创建QingTian Enclave镜像文件时生成的PCR0和PCR8。
5. 安全员将PCR0和PCR8设置为IAM访问控制策略的条件键（控制kms-decrypt API）。



- 父虚拟机管理员：启动父虚拟机并设置启动qt-proxy服务，然后使用QingTian Enclave镜像启动QingTian Enclave。

## 远程证明和数据解密

下面我们将概述在QingTian Enclave应用程序的执行流程。

- 通过qt-proxy服务，QingTian Enclave应用程序将密文从华为OBS桶中下载到QingTian Enclave实例中。
- Enclave应用程序产生一对RSA公私钥（pubKey, priKey），用于建立与KMS服务之间的端到端数据加密（它并不需要依赖HTTPS）。然后使用QingTian Enclave SDK调用KMS提供的支持AttestationDocument（证明文档）作为入参的kms-decrypt API。证明文档中包括QingTian Enclave的PCR值和应用程序产生的加密公钥（pubKey）。
- 华为云KMS接收并验证证明文档是否由QingTianAttestation PKI签发。在kms-decrypt API的访问控制检查中，会将证明文档中的PCR值与IAM授权策略中的PCR值进行比较，如果匹配成功则允许调用该接口，否则拒绝访问。
- KMS将数据解密后，然后用AttestationDocument提供的pubKey对数据进行重新加密后发送回QingTian Enclave应用程序。QingTian Enclave应用程序使用priKey对接收到的密文数据进行解密。

### 18.3.2 构建 QingTian Enclave 镜像

在开发人员开发完成一个QingTian Enclave应用程序后，还需要在一个可信赖的环境中构建QingTian Enclave镜像文件（.eif）。该镜像文件提供了启动QingTian Enclave实例所需要的所有信息，包括应用程序代码、运行时依赖、操作系统和文件系统等。在本节我们将说明如何创建QingTian Enclave镜像文件。

- 制作docker源镜像

用户将开发好的enclave应用程序及其相关的执行环境打包成docker镜像。详情可见[Linux系统上QingTian Enclave应用的开发](#)。

- 获取镜像库中的镜像

本章将使用docker仓中提供的ubuntu镜像为例。从docker中获取镜像源（虚拟机内需要配置网络才可查询）。查询镜像源命令：

```
docker search ubuntu
```

将ubuntu镜像pull到本地：

```
docker pull ubuntu
```

镜像pull到本地后可以通过以下命令查询到：

```
docker image ls
```

如果您使用本地docker镜像，可以直接进行步骤3镜像转换。

- 镜像转换

接下来，您需要将docker镜像转换为QingTian Enclave镜像，转换前可以通过（openssl或其他工具）创建私钥（private-key.pem）和证书（server.pem）。生成私钥和证书为可选项。**qt make-img**命令中的必要参数为docker源镜像和创建后生成的QingTian Enclave目标镜像。

```
qt enclave make-img --docker-uri ubuntu --eif /home/docker/ubuntu.eif --
private-key /home/docker/private-key.pem--signing-certificate
/home/docker/server.pem
{
 "digest": "SHA384",
 "PCR0":
 "b8c59692da8a5bcb739a83d15a0ceca670bd78da06cb2250ec70548f72254e674419e9888db9c0364a9b
```

```
88dd58017a62"
 "PCR8":
 "dbf4a7f9fab7f18619b5899c407081981ad6762fb9a809da78548821b5021965423181584acd7b2017033
 76f1133a546"
 }
}
```

至此，QingTian Enclave可用的EIF镜像已制作完成。您将获得一组散列PCR0和PCR8，这些散列值用作预期的Enclave度量值，它们可用于IAM授权策略中的条件键，以实现KMS API的条件访问控制。详情可见[PCR简介](#)。

### 18.3.3 启动 QingTian Enclave

#### 资源隔离

启动QingTian Enclave虚拟机之前，您首先要在父虚拟机内隔离资源供QingTian Enclave虚拟机使用。隔离的资源包括cpu个数和内存大小，可以通过在虚拟机内访问修改`*/etc/qingtian/enclave/qt-enclave-env.conf`配置文件指定需要隔离的资源：

```
#1G大页
hugepage_size:1024
#内存1G.
memory_mib:1024
cpu个数
cpu_count:2
cpu列表
cpu_list:2,3
```

我们建议您最好不要反复启动资源隔离服务，否则将会出现大页内存不足，导致QingTian Enclave无法正常启动或者隔离服务无法正常启动。本教程中保持默认值不变，使用1G大页，隔离2个vCPU和1G内存。确认配置文件参数之后执行以下命令：

```
systemctl restart qt-enclave-env.service
```

其中`*/etc/qingtian/enclave/qt-enclave-env.conf`配置文件中的配置项之间有约束条件，详情可见[qt-enclave-env服务介绍](#)。

#### 启动 QingTian Enclave

在父虚拟机内，使用`qt enclave start`命令创建QingTian Enclave虚拟机，该命令中需要指定QingTian Enclave虚拟机镜像文件。在QingTian Enclave虚拟机启动后，QingTian Enclave应用程序及其依赖项会被从该镜像文件中引导到QingTian Enclave虚拟机中。这里以创建一个拥有2个vCPU，1G内存，CID为4的QingTian Enclave虚拟机为例：

```
[root@localhost ~]# qt enclave start --cpus 2 --mem 1024 --eif /home/docker/ubuntu.eif --cid 4
Started enclave with EnclaveID : 0, EnclaveCID : 4, NumberOfCPUs : 2, MemoryMiB : 1024
{
 "EnclaveID": 0,
 "EnclaveCID": 4,
 "NumberOfCPUs": 2,
 "MemoryMiB": 1024,
 "LaunchMode": "debug"
}
```

在该实例中，原镜像ubuntu中的CMD语句为`/bin/bash`，因此QingTian Enclave启动之后将会执行该语句，执行完成后将退出并关闭。

## 18.4 密码学证明

QingTian Enclave支持密码学证明。QingTian Enclave支持通过密码学证明过程来证明其身份并与外部服务建立信任。证明过程需要使用证明文档，它包含对QingTian

Enclave运行环境的可信度量值。您可以使用这些度量值在外部服务中创建访问控制策略，以控制特定QingTian Enclave实例对指定操作的访问权限。

您可以通过使用QingTian Enclave SDK从QingTianHypervisor获取证明文档，该证明文档包含相关的度量值及数字签名。QingTian Enclave应用可以在向外部服务发送请求时携带该证明文档。当外部服务收到该请求时，会验证证明文档中包含的度量值是否匹配预期的度量值，以决定是否允许Enclave应用的访问请求或建立安全。

## 18.4.1 PCR 简介

QingTian Enclave的度量值是由标准可信度量运算得到的一组哈希值（Hashes）组成，这组哈希值保存在QingTian安全模块（QingTian Security Module, QTSM）的平台配置寄存器（Platform configuration registers, PCR）。

### 说明

QingTian Enclave度量值最多可支持32个PCR。QingTian系统占用index 0~15的PCR，用户的Enclave应用可使用index 16~31的PCR。

使用debug-mode启动Enclave时，不进行镜像校验。QingTian系统的index 0~15的PCR默认全为零避免打印泄漏。用户的Enclave应用可继续使用index 16~31的PCR。

## 系统 PCR

| PCR  | 度量内容                     | 备注                                                                                                   |
|------|--------------------------|------------------------------------------------------------------------------------------------------|
| PCR0 | QingTian Enclave镜像文件     | QingTian Enclave镜像文件本身，不包括证书与签名信息。                                                                   |
| PCR3 | IAM Agency               | 分配给父虚拟机实例的IAM Agency的连续度量值。确保父虚拟机具有正确的IAM Agency时，证明流程才能成功。<br>用户通过页面设置，当前仅在父虚拟机启动时下发一次，重置后需重启虚拟机生效。 |
| PCR4 | 父虚拟机Instance UUID        | 父虚拟机UUID的连续度量值。确保只有当父虚拟机具有特定的UUID时，证明过程才会成功。                                                         |
| PCR8 | QingTian Enclave镜像文件签名证书 | QingTian Enclave镜像文件的签名证书。                                                                           |

当前QingTian Enclave提供的度量值支持PCR0和PCR8，后续会进行扩展。

1. PCR0是对QingTian Enclave镜像文件的度量值，在镜像制作完成时，PCR0的值就是确定的。以下是PCR0实例：

```
EXTEND_PCR: index: 0
```

```
EXTEND_PCR: data:
```

```
0d1ae7330f437ee563178df30a7c7b7634125d31cac14f6784933db5e90080008438b38fdbb39c886ffe058
```



```
6ab099b56
EXTEND_PCR res: data:
b8c59692da8a5bcb739a83d15a0ceca670bd78da06cb2250ec70548f72254e674419e9888db9c0364a9b8
8dd58017a62
```

2. 为了进一步加强QingTian Enclave的安全策略，您可以创建IAM Agency并将其附加到父虚拟机。在KMS密钥策略的条件密钥中使用IAM Agency的SHA384哈希值作为PCR3。这样做可确保只有在具有正确IAM Agency的实例上运行的QingTian Enclave才能针对KMS密钥执行特定的KMS操作。您可以使用任何能够将字符串转换为SHA384哈希的工具生成哈希。以下是PCR3实例：

```
$IAM_AGENCY="iam: :6c031a4leefc480bb60f20c003891fcd:agency:cdd"; \
python -c"import hashlib, sys; \
h=hashlib.sha384(); h.update(b'\0'*48); \
h.update("$IAM_AGENCY".encode('utf-8')); \
print(h.hexdigest())"
```

3. PCR4基于父虚拟机UUID的SHA384计算值。因此，您可以在启动父虚拟机后生成PCR。您可以使用任何能够将字符串转换为SHA384哈希的工具生成此哈希。以下是PCR4实例：

```
$INSTANCE_ID="ecb23eec- 51d4-462f-8dbd-63bfbae7869b"; \
python -c"import hashlib, sys; \
h=hashlib.sha384(); h.update(b'\0'*48); \
h.update("$INSTANCE_ID".encode('utf-8')); \
print(h.hexdigest())"
```

4. PCR8是对QingTian Enclave镜像文件的签名证书的度量值，用户可以选择用自己的证书和私钥对镜像文件进行签名。只有当镜像文件使用了签名证书和私钥进行签名，才会有对应的PCR8。使用PCR8可以确认是镜像是通过特定的签名证书来进行签名的，即使镜像文件改变，只要指定的签名证书不变，PCR8就不会变化。以下是PCR8实例：

```
EXTEND_PCR: index: 8
EXTEND_PCR: data:
c5b3e075e00c261e7fc364f1541067b2a42d4b793225ab10e5cfb8eaca31b3d598af9dd2e491828c2569a9
953401abcb
EXTEND_PCR res: data:
4f8b066ce5ac24150612ba9a55bbb9211f626152ada40ede160f4d7ecbfa214c2a549181f6611a3d16a12e
c88a577a01
```

## 18.4.2 证明文档

证明文档用于证明QingTian Enclave实例的可信度量结果。证明文档由QingTianHypervisor生成，文档内容包括PCR列表、QingTianPKI证书链、密码算法声明以及Enclave应用自定义数据。证明文档由华为云QingTian Attestation PKI（Public Key Infrastructure）签署。

QingTianHypervisor生成的证明文档以简明二进制对象表示（CBOR）进行编码，并用CBOR Object Signing and Encryption（COSE）进行签名。关于CBOR，详情可见[RFC 8949: Concise Binary Object Representation（CBOR）](#)。

关于QingTian Enclave证明文档结构，我们遵循CDDL规范（RFC 8610）描述如下：

```
AttestationDocument = {
 module_id: text, ; 安全模块编号
 timestamp: uint .size 8, ; 时间戳
 digest: digest, ; 摘要算法
 pcrs: { + index => pcr }, ; PCR列表
 certificate: cert, ; Enclave证明文档的签名证书
 cabundle: [* cert], ; QingTianPKI证书链
 ? user_data: user_data, ; 可选，应用相关的数据
 ? nonce: user_data, ; 可选，不重复使用的数值
 ? pubkey: user_data, ; 可选，应用程序相关的公钥
}

cert = bytes .size (1..4096) ; DER编码证书
```

```
user_data = bytes .size (0..4096)
pcr = bytes .size (48) ; PCR内容
index = 0..31
digest = "SHA384"
```

证明文档中的可选参数（pubkey，user\_data和nonce）可以支持QingTian Enclave与外部实体之间运行自定义（或自协商）的应用级安全协议。比如，QingTian Enclave应用可以创建一个非对称密钥对（PriKey, PubKey），并通过QingTian Enclave证明文档来提供对PubKey的可信证明，那么外部实体就可以基于该PubKey证明来运行一些自定义的可信密钥分发、可信密钥协商等应用级安全协议。

### 18.4.3 文档签名验证

接下来，我们将介绍证明文档的验证过程。当您从QingTianHypervisor请求证明文档时，您将会收到一个包含已签名证明文档的二进制blob。已签名的证明文档是一个由CBOR编码，COSE签名的对象。整个验证过程包括以下几个步骤：

1. 解码CBOR对象并将其映射到COSE\_Sign1结构中
2. 从COSE\_Sign1结构中提取证明文档
3. 验证证书文档中CA证书链的有效性
4. 验证证明文档的数字签名的有效性

证明文档由华为云QingTianAttestation PKI签署。华为云QingTian根证书可以从[https://qingtian-enclave.obs.myhuaweicloud.com/huawei\\_qingtian-enclaves\\_root-G1.zip](https://qingtian-enclave.obs.myhuaweicloud.com/huawei_qingtian-enclaves_root-G1.zip) 下载获取，该压缩文件的SHA256的摘要值为：

```
99e9203a64cfb0c6495afd815051e97bea8a37895dc083d715674af64adeadfe
```

QingTianAttestation PKI根证书的有效期为30年，根证书主体（Subject）的格式如下：

```
CN=huaweicloud.qingtian-enclaves, C=CN, O=Huawei Technologies, OU=Huawei Cloud
```

## COSE 与 CBOR

COSE\_Sign1签名结构通常被使用在为消息进行单一签名的场景，其内容和签名的相关参数都放在受保护的header中。COSE\_Sign1数据结构是一个CBOR数组，该数组包括如下字段：

```
[
 protected header; // 被保护的header信息
 unprotected header; // 不被保护的header信息
 payload; // 被签名的数据， attestation doc的CBOR封装数据
 signature; // 签名
]
```

在证明文档的上下文中，该数组示例为：

```
18(// COSE_Sign1 CBOR标签
 {1: -35}, // 算法: ECDS 384
 {}, // 空
 attestation doc, // 证明文档
 signature, // 签名
)
```

## 验证证书

验证证书链是证书有效性验证过程中不可或缺的一部分。证明文档中的CA Bundle包含了根证书和中间证书的证书列表，并按以下顺序提供：

```
[ROOT_CERT - INTERM_1 - INTERM_2 ... -INTERM_N]
 0 1 2 N
```

为了验证目标证书TARGET\_CERT（即证明文档中的Certificate）的有效性，对于某些证书验证工具而言，您可能需要按如下顺序构造证书验证列表：

```
[TARGET_CERT - INTERM_N - INTERM_N-1 ... - ROOT_CERT]
```

## 18.4.4 结合华为云 KMS 服务

华为云KMS服务内置了对QingTian Enclave证明的支持。通过使用QingTian Enclave SDK中包含的华为云KMS API，您可以在QingTian Enclave实例中基于QingTian Enclave证明来执行华为云KMS操作，比如解密、生成随机数和加密等操作。华为云KMS服务提取来自QingTian Enclave的证明文档并根据预设的IAM授权策略对其进行访问权限控制。

比如，如下是一个IAM授权策略的举例。该授权策略允许调用KMS的解密数据或解密数据密钥功能API，但限制条件是要求请求者必须在QingTian Enclave环境中运行，且Enclave的度量值PCR0和PCR8都必须和指定的PCR值相同。

```
{
 "Version": "1.1",
 "Statement": [
 {
 "Effect": "Allow",
 "Action": [
 "kms:cmk:decrypt",
 "kms:dek:decrypt"
],
 "Resource": "*",
 "Condition": {
 "StringEqualsIgnoreCase": {
 "kms:RecipientAttestation/PCR0": [
 "c5158cb6ee9dbb0ead648c3dc80e472c85e0d67f19fb53fbd3fb94c3371aec63cdb93b80d727a7084248873b1d8e8b41"
],
 "kms:RecipientAttestation/PCR8": [
 "705afb1012d27f4e07a25e674e6a17dec57305e29cd412184b7bcb78d9e67f16a0cc26d8706a4fab418a5da5788bc949"
]
 }
 }
 }
]
}
```

## 18.5 QingTian Enclave 应用开发

一个功能完备的QingTian Enclave应用程序至少包括如下两个部分：

1. 一个运行在父虚拟机内安全不敏感的应用程序。
2. 一个运行在Enclave环境中安全敏感的应用程序。

由于QingTian Enclave环境的独立性，QingTian Enclave环境中的应用程序和运行在父虚拟机中的应用程序间只能通过唯一的vsock通道进行通信。

## 18.5.1 Linux 系统上 QingTian Enclave 应用的开发

### QingTian Enclave SDK

QingTian Enclave SDK由一系列开源库组成，以便您开发自己的QingTian Enclave应用程序。其中包括QingTian安全模块（QingTian Security Module，QTSM）提供的qtsm-lib函数库。此外，SDK集成了KMS接口，该接口内置了获取证明文档及调用华为云KMS相关服务的功能。在典型使用案例里，我们描述了在QingTian Enclave调用KMS解密接口的实例。

表 18-1 接口介绍

| 类型            | 接口                                       | 接口描述                |
|---------------|------------------------------------------|---------------------|
| libqtsm<br>接口 | qtsm_describe_pcr                        | 查询指定index的PCR数据信息   |
|               | qtsm_extend_pcr                          | 扩展指定index对应的PCR值    |
|               | qtsm_lock_pcr                            | 锁定指定index的PCR数据信息   |
|               | qtsm_lock_pcrs                           | 批量锁定指定index对应的PCR值  |
|               | qtsm_get_describe                        | 获取qtsm信息            |
|               | qtsm_get_attestation                     | 获取Attestation Doc   |
|               | qtsm_get_random                          | 获取硬件随机数             |
| KMS接口         | kms_generate_datakey_blocking            | 生成新的密钥对，获取公钥与私钥     |
|               | kms_generate_datakey_blocking_with_proxy | 集成qtproxy代理的密钥对获取接口 |
|               | kms_gen_random_blocking                  | 获取随机数               |
|               | kms_gen_random_blocking_with_proxy       | 集成qtproxy代理的随机数获取接口 |
|               | kms_decrypt_data_blocking                | 解密数据                |
|               | kms_decrypt_data_blocking_with_proxy     | 集成qtproxy代理的解密数据接口  |

源码可以在开源仓库<https://gitee.com/HuaweiCloudDeveloper/huawei-qingtian/tree/master/enclave>免费获取，您可以基于测试示例开发自己的QingTian Enclave应用程序。

### Vsock 通信实例

本节主要通过vsock示例来介绍如何在Linux环境下开发QingTian Enclave应用程序。本节中提供的vsock程序只支持在Linux环境下运行。

通过该vsock程序，可以帮助开发者了解到父虚拟机和QingTian Enclave间如何进行消息传递从而实现双方交互。该vsock程序带有Server和Client参数，通过指定该参数可

以指定父虚拟机或者QingTian Enclave虚拟机扮演相应角色。在该vsock程序中，客户端程序会发送一段简单的文本信息通过vsock通道传递给服务端程序，同时服务端程序会监听相应的vsock通道，一旦收到来自客户端的消息后，会将其收到的消息打印在终端界面上。

在下面介绍中，我们将说明如何让QingTian Enclave扮演服务端等待并接收来自客户端父虚拟机的hello world信息：

#### 1. 编写一个SocketCommunication.py程序：

```
#!/usr/local/env python3
import argparse
import socket
import sys

CID_DEFAULT = 3
PORT_DEFAULT = 9999
TIMEOUT = 5
BLACKLOG_DEFAULT = 5

class Client:
 def __init__(self, cid, port):
 self.clientAddr = (cid, port)
 self.connect()

 def connect(self):
 self.socket = socket.socket(socket.AF_VSOCK, socket.SOCK_STREAM)
 self.socket.settimeout(TIMEOUT)
 print("connecting to the server")
 try:
 self.socket.connect(self.clientAddr)
 except socket.error:
 print("client's socket connection err")
 sys.exit(1)

 def send(self, msg):
 print("client sends hello to the server")
 self.socket.sendall(msg)

 def disconnect(self):
 self.socket.close()

 def receiveData(self):
 while True:
 try:
 message = self.socket.recv().decode()
 except (socket.error, UnicodeDecodeError):
 break
 if message:
 print(message, end = " ", flush = True)
 print()

def clientHandler(args):
 client = Client(args.cid, args.port)
 message = "Hello world"
 client.send(message.encode())
 client.disconnect()

class Server:
 def __init__(self, port):
 self.socket = socket.socket(socket.AF_VSOCK, socket.SOCK_STREAM)
 self.serverAddr = (socket.VMADDR_CID_ANY, port)
 self.socket.bind(self.serverAddr)
 self.socket.listen(BLACKLOG_DEFAULT)

 def receiveData(self):
 while True:
 print("waiting for a connection")
 (conn, clientAddr) = self.socket.accept()
```

```
try:
 print("connection from ", clientAddr)
 while True:
 try:
 data = conn.recv(256).decode()
 except (socket.error, UnicodeDecodeError):
 break
 if data:
 print("data: ", data)
 else:
 print("connection close")
 break
 finally:
 conn.close()

def serverHandler(args):
 server = Server(args.port)
 server.receiveData()

def main():
 parser = argparse.ArgumentParser(description = "Hello world demo", prog='SocketCommunication')
 subparsers = parser.add_subparsers(description = "Communication roles")
 parserClient = subparsers.add_parser("Client", description = "Client",
 help = "Communicate with server using a given cid and port.")
 parserClient.add_argument("-c", "--cid", default = CID_DEFAULT, type = int, help = "Client's Cid")
 parserClient.add_argument("-p", "--port", default = PORT_DEFAULT, type = int, help = "Client's
port")
 parserClient.set_defaults(func = clientHandler)
 parserServer = subparsers.add_parser("Server", description = "Server", help = "Listen on a given
port")
 parserServer.add_argument("-p", "--port", default = PORT_DEFAULT, type = int, help = "Server's
Port")
 parserServer.set_defaults(func = serverHandler)
 if len(sys.argv) < 2:
 parser.print_usage()
 sys.exit(1)
 args = parser.parse_args()
 args.func(args)

if __name__ == "__main__":
 main()
```

## 2. 创建一个Dockerfile文件，内容如下：

```
#start the Docker image from ubuntu
FROM ubuntu AS base-img
WORKDIR /home/builder
COPY vsocket example
COPY . vsocket
install relative dependencies
RUN apt-get update && \
 apt-get install python3 -y && \
 apt-get install gcc -y && \
 apt-get install gawk -y
Launch a client
CMD ["python3", "/home/builder/vsocket/SocketCommunication.py", "Server", "-p 9999"]
```

## 3. 构建docker镜像：

```
sudo docker build -t vsock-sample-client -f Dockerfile .
```

## 4. 将docker镜像转化为QingTian Enclave镜像：

```
qt enclave make-img --docker-uri vsock-sample-client --eif vsock_sample.eif
```

## 5. 使用vsock\_sample.eif以debug模式启动QingTian Enclave：

```
qt enclave start --cpus 2 --mem 4096 --eif vsock_sample.eif --debug-mode --cid 4
```

然后使用qt enclave console命令查看QingTian Enclave内只读终端输出：

```
qt enclave console --enclave-id 0
waiting for a connection
```

6. 然后另起一个父虚拟机终端，启动客户端程序：

```
python3 SocketCommunication.py Client -c 4 -p 9999
```
7. 当服务端程序收到来自vsock的消息后，会打印如下消息到终端上：

```
connection from (3, 4180219645)
data: Hello world
connection close
waiting for a connection
```

## libqtsm 与 SDK 使用示例

本节主要基于开源示例代码来介绍如何在QingTian Enclave应用程序中使用libqtsm与SDK接口。本节中提供的示例程序只支持在Linux环境下运行。

1. 安装libqtsm开发包：

```
yum install libqtsm-devel
```

2. 获取开源示例代码，拷贝至enclave镜像创建环境中，地址如下：

<https://gitee.com/HuaweiCloudDeveloper/huawei-qingtian/tree/master/enclave/qtsm>

3. 创建一个Dockerfile文件，内容如下：

```
start the Docker image from ubuntu
FROM ubuntu AS base-img
WORKDIR /home/builder
COPY libqtsm example
COPY ./qtsm qtsm_tests/
install relative dependencies
RUN apt-get update && \
 apt-get install gcc -y && \
 apt-get install make -y && \
 apt-get install libssl-dev -y && \
 apt-get install libglib2.0-dev -y && \
 apt-get install curl -y && \
 apt-get install libcurl4-openssl-dev -y && \
 apt-get install -y libcbor-dev && \
 apt-get install -y libjson-c-dev
build a test demo
RUN cd qtsm_tests/tests/ && \
 make
RUN cp /home/builder/qtsm_tests/tests/gtest_libqtsm /root/
Launch a client
CMD "/root/gtest_libqtsm"
```

4. 构建docker镜像，将docker镜像转换为enclave镜像并启动enclave。

5. 获取SDK接口的开源示例代码，地址如下：

<https://gitee.com/HuaweiCloudDeveloper/huawei-qingtian/tree/master/enclave/qtsm-sdk-c/samples>

## 18.5.2 QingTian Enclave 网络代理工具

### QingTian Enclave 网络代理工具（qproxy）概述

qproxy（QingTian Enclave proxy）是QingTian Enclave的网络代理工具。借助此工具，您可以无需任何改动，直接将部署在基于QingTian架构虚拟机中的网络业务移植到QingTian Enclave中，实现传统网络业务的无感迁移。

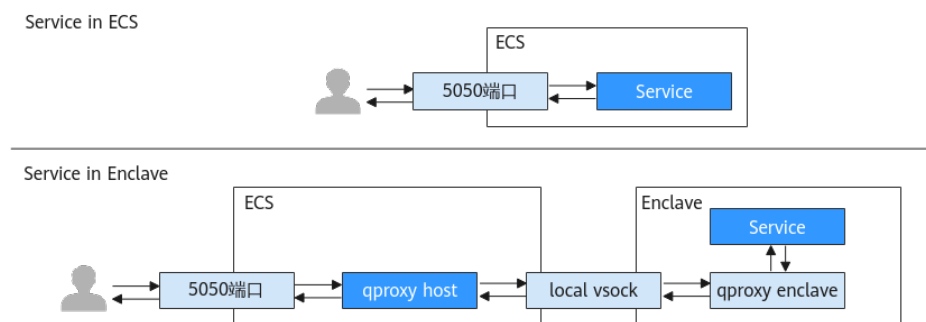
qproxy是一个可执行的二进制文件，需要在父虚拟机和QingTian Enclave中通过不同的命令运行：

- 在父虚拟机中，通过执行“`/path/to/qproxy host --config=/path/to/config_qproxy.toml <cid>`”命令启用qproxy。
- 在QingTian Enclave中，通过执行“`/path/to/qproxy enclave --config=/path/to/config_qproxy.toml`”命令启用qproxy。

以一个简单的网络服务为例，网络服务直接部署在虚拟机中，负责处理来自最终用户的请求。最终用户通过该服务暴露的端口，发起网络请求，等待服务响应。在将该网络服务迁移至QingTian Enclave之后，通过qproxy代理，可以提供和原本网络服务相同的服务。

使用同一个qproxy二进制文件，通过`qlog host`和`qlog enclave`命令，使其分别运行在父虚拟机和QingTian Enclave中，两者会建立基于local vsock的通信链路。父虚拟机中的qproxy组件会监听5050端口，接收用户请求，并通过local vsock将收到的请求转发给QingTian Enclave中的qproxy组件。该组件会将请求发送到QingTian Enclave中监听的5050端口。网络服务处理完成之后，响应数据会原路返回。

图 18-1 网络服务无感迁移至 QingTian Enclave



本文介绍qproxy的操作步骤。

## 前提条件

1. 获取qproxy源码。  
克隆QingTian Enclave代码仓。  

```
git clone https://gitee.com/HuaweiCloudDeveloper/huawei-qingtian.git
```
2. 获取cargo工具链。
  - a. 执行以下命令安装rustup。  

```
curl --proto 'https' --tls1.2 -sSf https://sh.rustup.rs | sh
```
  - b. 安装完成后执行以下命令加载rustup。  

```
source $HOME/.cargo/env
```
  - c. 执行以下命令验证rustc和cargo安装成功。  

```
rustc -V
cargo -V
```
3. qproxy的前置依赖。

表 18-2 前置依赖

| 依赖项     | 最低测试版本 |
|---------|--------|
| cargo   | 1.77.0 |
| libcbor | 0.10.2 |



| 依赖项                | 最低测试版本 |
|--------------------|--------|
| libssl (libcrypto) | 3.0.0  |
| libcurl            | 4.0.0  |

4. QingTian Enclave的环境准备。
  - a. 安装qt CLI工具和其他必要rpm包。
  - b. 安装Docker。
  - c. 安装python3以及几个必要的python module : docker和knack。详细内容，请参考[快速入门](#)和[安装qt CLI](#)。

## 操作步骤

### 步骤1 构建qproxy。

生成的qproxy程序在目录qingtian-tools/qproxy/target/release下。

### 步骤2 创建工作目录。

新建工作目录workspace，将qproxy程序拷贝到workspace下，后续创建的文件都放在workspace中。

### 步骤3 配置config\_qproxy.toml。

在workspace目录中创建config\_qproxy.toml文件，内容如下：

```
[[inbound_connections]]
host_port = 5050
enclave_port = 5050
vssock_port = 9995

[[inbound_connections]]
host_port = 5443
enclave_port = 5443
vssock_port = 9994

[log_location]
host_log = "host.log" # qproxy host log name, e.g./var/log/qproxy/host.log
enclave_log = "enclave.log" # qproxy enclave log name, e.g./var/log/qproxy/enclave.log
log_level = "info" # qproxy logger level, e.g. "off", "info", "warn", "error", "debug", "trace"
host_log_dir = "/var/log/qproxy" # qproxy host log dir, and its default value is "/var/log/qproxy"
enclave_log_dir = "/var/log/qproxy" # qproxy enclave log dir, and its default value is "/var/log/qproxy"
```

### 步骤4 制作含qproxy的QingTian Enclave镜像。

1. 在workspace目录中创建http请求测试脚本app.sh：

```
#!/bin/bash

PORT=5050

while true; do
 echo -e "HTTP/1.1 200 OK\nContent-Type: text/plain\n\nHello world!" | nc -l 127.0.0.1 $PORT
done
```

2. 赋予app.sh执行权限，在workspace目录中执行以下命令：

```
chmod +x app.sh
```

3. 在workspace目录中创建start.sh脚本：

```
#!/bin/bash
ip link set lo up
```

```
/root/qproxy enclave --config=/root/config_qproxy.toml &
/root/app.sh
```

4. 赋予start.sh执行权限，在workspace目录中执行以下命令：  
chmod +x start.sh
5. 在workspace目录中创建Dockerfile文件：  
FROM ubuntu:latest  
COPY ./qproxy /root/qproxy  
COPY ./config\_qproxy.toml /root/config\_qproxy.toml  
COPY ./start.sh /root/start.sh  
COPY ./app.sh /root/app.sh  
RUN apt-get update && \  
 apt-get install -y netcat-openbsd && \  
 apt-get install -y iproute2  
CMD "/root/start.sh"
6. 制作docker镜像，在workspace目录中运行以下命令：  
docker build -f Dockerfile -t test\_qproxy\_enclave .
7. 制作QingTian Enclave镜像，在workspace目录中运行以下命令：  
qt enclave make-img --docker-uri test\_qproxy\_enclave --eif test\_qproxy\_enclave.eif

### 步骤5 启动qproxy。

1. 启动一台QingTian Enclave，在workspace目录中运行以下命令：  
qt enclave start --cpus 2 --mem 1024 --cid 4 --eif test\_qproxy\_enclave.eif
2. 启动父虚拟机中的qproxy，在workspace目录中运行以下命令：  
./qproxy host --config=./config\_qproxy.toml 4 &
3. 父虚拟机中执行curl命令：  
curl localhost:5050  
可以看到终端输出"Hello world!"。

### 步骤6 qproxy环境变量设置。

qproxy包含两分子命令，一部分在QingTian Enclave ( qproxy enclave ) 中运行，另一部分在父虚拟机 ( qproxy host ) 中运行。

通过设置环境变量RUST\_LOG可以控制二进制文件的日志输出信息等级。

- 设置RUST\_LOG=OFF，日志信息会全部关闭不显示。
- 设置RUST\_LOG=info，日志信息会显示info，warn和error三个级别的信息。
- 默认情况下，只会显示warn和error两个级别的日志信息。
- 可以通过查看[EnvFilter文档](#)获取更多信息。

----结束

## qproxy 帮助信息

### qproxy help

```
$ qproxy --help
Usage: qproxy <COMMAND>

Commands:
 host The part of qproxy that runs outside the enclave
 enclave The part of qproxy that runs inside the enclave
 check-config Check the qproxy configuration file
 help Print this message or the help of the given subcommand(s)

Options:
 -h, --help Print help
 -V, --version Print version
```

## qproxy enclave help

```
$ qproxy enclave --help
The part of qproxy that runs inside the enclave

Usage: qproxy enclave [OPTIONS]

Options:
 --parent-cid <PARENT_CID>
 The CID of the parent VM of this enclave

 [env: QPROXY_PARENT_CID=]
 [default: 3]

 --config <CONFIG>
 Path to the configuration file

-t, --threads <THREADS>
 Number of threads the async runtime is allowed to use

 [env: TOKIO_WORKER_THREADS=]

--control-port <CONTROL_PORT>
 The port where to listen for control messages from the enclave

 Leave at default value unless you know what you are doing

 [env: QPROXY_CONTROL_PORT=]
 [default: 6666]
 [0..=65535]

-h, --help
 Print help (see a summary with '-h')
```

## qproxy host help

```
$ qproxy host --help
The part of qproxy that runs outside the enclave

Usage: qproxy host [OPTIONS] <CID>

Arguments:
 <CID>
 The CID of the enclave

 [env: QPROXY_LISTEN_CID=]

Options:
 --config <CONFIG>
 Path to the configuration file

-t, --threads <THREADS>
 Number of threads the async runtime is allowed to use

 [env: TOKIO_WORKER_THREADS=]

--ipv4
 Only resolve IPv4 addresses

--ipv6
 Only resolve IPv6 addresses

--control-port <CONTROL_PORT>
 The port where to listen for control messages from the enclave

 Leave at default value unless you know what you are doing

 [env: QPROXY_CONTROL_PORT=]
 [default: 6666]
 [0..=65535]
```

```
-h, --help
Print help (see a summary with '-h')
```

## 配置信息 ( Config )

### 配置参数解释

- `outbound_connections`: 配置转发QingTian Enclave外发流量到外部特定服务 ( hostname/IP ) 及其端口。
- `inbound_connections`: 配置转发父虚拟机指定端口收到的流量到QingTian Enclave中。

表 18-3 配置参数解释

| Bound                | Variable     | Type   | Description                                                                                   |
|----------------------|--------------|--------|-----------------------------------------------------------------------------------------------|
| outbound_connections | hostname     | String | 用于转发流量的主机名 ( hostname ) , 可以是IP地址, 例如: hostname = "api.myservice.com"。                        |
|                      | vsock_port   | u32    | QingTian Enclave内vsock要使用的端口, 必须唯一, 并且不能和qproxy端口冲突。                                          |
|                      | tcp_port     | u32    | 服务器转发流量的端口 ( 将连接到 hostname:port ) , 不能和qproxy端口冲突 ( 默认是8080 ) 。                               |
| inbound_connections  | host_port    | u32    | 主机端 ( host ) qproxy将会监听的端口, 如 host_port = 80, 主机端qproxy将会监听 0.0.0.0:80。                       |
|                      | enclave_port | u32    | QingTian Enclave端qproxy将会监听的端口, 如 enclave_port = 80, QingTian Enclave端qproxy将会监听127.0.0.1:80。 |
|                      | vsock_port   | u32    | QingTian Enclave内vsock要使用的端口, 必须唯一, 并且不能和qproxy端口冲突。                                          |

### 日志文件配置

表 18-4 日志文件配置

| Variable    | Type   | Description                                                                                                |
|-------------|--------|------------------------------------------------------------------------------------------------------------|
| host_log    | String | qproxy在主机端 ( host ) 日志文件名称, 例如: host_log = "host.log"表示主机端 ( host ) qproxy日志文件名为host.log。                  |
| enclave_log | String | qproxy在QingTian Enclave端日志文件名称, 例如: enclave_log = "enclave.log"表示QingTian Enclave中qproxy日志文件名为enclave.log。 |

| Variable        | Type   | Description                                                                                             |
|-----------------|--------|---------------------------------------------------------------------------------------------------------|
| log_level       | String | qproxy日志级别, 例如: log_level = "info"表示显示信息, 警告和错误。全部级别有 "off", "info", "warn", "error", "debug", "trace"。 |
| host_log_dir    | String | qproxy在主机端 ( host ) 日志目录, 默认是 "/var/log/qproxy"                                                         |
| enclave_log_dir | String | qproxy在QingTian Enclave端日志目录, 默认是 "/var/log/qproxy"                                                     |

### 配置文件参考

```
[[outbound_connections]]
hostname = "api.myservice.com"
vsock_port = 7777
tcp_port = 443

[[outbound_connections]]
hostname = "another.api.com"
vsock_port = 7778
tcp_port = 5555

[[inbound_connections]]
host_port = 80
enclave_port = 80
vsock_port = 9000

[[inbound_connections]]
host_port = 443
enclave_port = 443
vsock_port = 9001

[log_location]
host_log = "host.log"
enclave_log = "enclave.log"
log_level = "info"
host_log_dir = "/var/log/qproxy"
enclave_log_dir = "/var/log/qproxy"
```

## 18.5.3 QingTian Enclave 日志转发工具

### QingTian Enclave 日志转发工具 ( qlog ) 概述

qlog ( QingTian Enclave log ) 是QingTian Enclave的运维工具。QingTian Enclave是运行在QingTian架构虚拟机中完全隔离的子虚拟机, 即使是root用户也不能直接通过ssh登入。因此, 为了方便运维人员监控运行在QingTian Enclave中的业务和定位问题, 我们推出了qlog工具。qlog可以收集QingTian Enclave虚拟机中指定的日志文件和资源使用情况, 并将结果发送到父虚拟机中。目前qlog支持收集QingTian Enclave中的资源使用情况包括CPU使用率和内存使用率。

qlog是一个可执行二进制文件, 需要在父虚拟机中和QingTian Enclave中通过不同的命令运行:

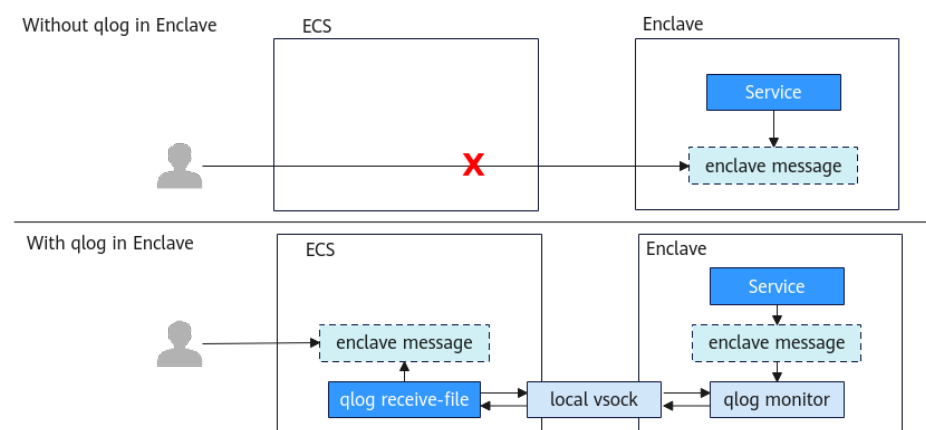
- 在父虚拟机中, 通过执行 “/path/to/qlog receive-file <cid>/path/to/config\_qlog.toml” 命令启用qlog。
- 在QingTian Enclave中, 通过执行 “/path/to/qlog monitor /path/to/config\_qlog.toml” 命令启用qlog。

传统虚拟机中业务的日志会直接存储在类似“/var/log/service.log”的目录文件当中。

在将业务移植到QingTian Enclave中之后，通过qlog工具，可以将上述日志文件（“/var/log/service.log”）导出到父虚拟机当中。

使用同一个qlog二进制文件，通过**qlog receive-file**和**qlog monitor**命令，使其分别运行在父虚拟机和QingTian Enclave中，两者会建立基于local vsock的通信链路。运行在QingTian Enclave中的qlog组件会收集指定的业务日志或QingTian Enclave的资源使用情况，将收集到的数据发送给父虚拟机中的qlog组件。父虚拟机中的qlog组件再将接收到的数据储存到指定的文件中，如“/var/log/qlog/service.log”。

图 18-2 网络服务无感迁移至 QingTian Enclave



本文介绍qlog的操作步骤。

## 前提条件

1. 获取qlog程序。  
克隆QingTian Enclave代码仓。  
`git clone https://gitee.com/HuaweiCloudDeveloper/huawei-qingtian.git`
2. 获取cargo工具链。
  - a. 执行下述命令安装rustup  
`curl --proto '=https' --tlsv1.2 -sSf https://sh.rustup.rs | sh`
  - b. 安装完成后执行下述命令加载rustup  
`source $HOME/.cargo/env`
  - c. 执行下述命令验证rustc和cargo安装成功  
`rustc -V`  
`cargo -V`
3. qlog的前置依赖。

表 18-5 前置依赖

| 依赖项   | 最低测试版本 |
|-------|--------|
| glibc | 2.34   |
| cargo | 1.77.0 |

4. QingTian Enclave的环境准备。
  - a. 安装qt CLI工具和其他必要rpm包。
  - b. 安装Docker。
  - c. 安装python3以及几个必要的python module : docker和knack。详细内容, 请参考[快速入门](#)和[安装qt CLI](#)。

## 操作步骤

### 步骤1 构建qlog。

进入qingtian-tools/qlog目录中, 执行cargo命令:

```
cargo build --release
```

生成的qlog程序在目录qingtian-tools/qlog/target/release下。

### 步骤2 创建工作目录。

新建工作目录workspace, 将qlog程序拷贝到workspace下, 后续创建的文件都放在workspace中。

### 步骤3 配置config\_qlog.toml。

在workspace目录中创建config\_qlog.toml文件, 内容如下:

```
port: 6000
workspace: /var/log/qlog
server_logfile: server.log
client_logfile: client.log
server_threads: 4
client_threads: 1
log_level: info
rotate_size: 65536
rotate_num: 10
monitor_items:
 - name: service #业务名
 monitor_type: file
 monitor_path: /var/log/service.log #QingTian Enclave中业务日志的路径
 outputfile: service.log #同步到父虚拟机中的日志文件名
 - name: resource
 monitor_type: resource
 monitor_interval: 15
 outputfile: resource.log
```

### 步骤4 制作含qlog的QingTian Enclave镜像。

1. 在workspace目录中创建start.sh脚本:

```
#!/bin/bash
/root/qlog monitor /root/config_qlog.toml &

LOG_FILE="/var/log/service.log"
LOG_MESSAGE="Hello, service."

while true; do
 TIMESTAMP=$(date '+%Y-%m-%d %H:%M:%S')
 echo "$TIMESTAMP - $LOG_MESSAGE" >> "$LOG_FILE"
 sleep 3
done
```

2. 赋予start.sh脚本执行权限, 在workspace目录中执行下述命令:

```
chmod +x start.sh
```

3. 在workspace目录中创建Dockerfile文件, 内容如下:

```
FROM ubuntu:latest
COPY ./qlog /root/qlog
```

```
COPY ./config_qlog.toml /root/config_qlog.toml
COPY ./start.sh /root/start.sh
CMD /root/start.sh
```

4. 赋予start.sh执行权限，在workspace目录中执行以下命令：  
chmod +x start.sh
5. 制作docker镜像，在workspace目录中运行以下命令：  
docker build -f Dockerfile -t test\_qlog\_enclave .
6. 制作QingTian Enclave镜像，在workspace目录中运行以下命令：  
qt enclave make-img --docker-uri test\_qlog\_enclave --eif test\_qlog\_enclave.eif

### 步骤5 启动qlog。

1. 启动一台QingTian Enclave，在workspace目录中运行下述命令：  
qt enclave start --cpus 2 --mem 1024 --cid 4 --eif test\_qlog\_enclave.eif
2. 启动父虚拟机中的qlog，在workspace目录中运行下述命令：  
./qlog receive-file 4 ./config\_qlog.toml &
3. 父虚拟机中执行下述命令，查看service日志：  
tail -F /var/log/qlog/service.log  
可以看到每3s，打印一行日志"Hello, service."。
4. 父虚拟机中执行下述命令，查看QingTian Enclave中资源使用情况：  
tail -F /var/log/qlog/resource.log  
可以看到每15s，打印一次CPU使用率和内存使用率。

----结束

## qlog 帮助信息

### qlog help

```
$ qlog --help
A tool to monitor logs and resource usage over a Vsock connection

Usage: qlog <COMMAND>

Commands:
 monitor Monitor resource usage
 receive-file Receive data from qlog monitor
 help Print this message or the help of the given subcommand(s)

Options:
 -h, --help Print help
```

### qlog monitor help

```
$ qlog help monitor
Monitor logs and/or resource usage

Usage: qlog monitor [OPTIONS] <CONFIG>

Arguments:
 <CONFIG> A configuration file in yaml format, which defines items to be monitored

Options:
 -c, --cid <CID> CID to listen on (defaults to VMADDR_CID_ANY) [default: 4294967295]
 -h, --help Print help
```

### qlog receive-file help

```
$ qlog help receive-file
Receive data from qlog monitor

Usage: qlog receive-file [OPTIONS] <CID> <CONFIG>
```



Arguments:  
 <CID> Enclave VM's CID  
 <CONFIG> A configuration file in yaml format, which defines items to be monitored

Options:  
 -w, --workspace <WORKSPACE> Set to workspace if specified, prior to configuration file  
 -h, --help Print help

## 配置信息 ( Config )

表 18-6 配置参数解释

| Variable       | Type   | Description                                                    |
|----------------|--------|----------------------------------------------------------------|
| port           | u32    | vsoc端口                                                         |
| workspace      | String | 写入进程日志和监控数据的工作空间                                               |
| server_logfile | String | 将服务器端 ( qlog monitor ) 进程日志写入 <workspace>/<server_logfile>     |
| client_logfile | String | 将客户端 ( qlog receive-file ) 进程日志写入 <workspace>/<client_logfile> |
| server_threads | u32    | 服务器端 ( qlog monitor ) 线程数量                                     |
| client_threads | u32    | 客户端 ( qlog receive-file ) 线程数量                                 |
| log_level      | String | 日志等级 ( TRACE/DEBUG/INFO/WARN/ERROR/OFF )                       |
| rotate_size    | u32    | 可选参数: 日志文件转储大小, 默认是2MB                                         |
| rotate_num     | u32    | 可选参数: 日志文件转储时, 保留旧日志文件数量, 默认是10                                |
| monitor_items  | Vector | 监控事物列表                                                         |

表 18-7 监控项参数解释

| Variable | Type   | Description |
|----------|--------|-------------|
| name     | String | 事物名称        |

| Variable         | Type   | Description                                          |
|------------------|--------|------------------------------------------------------|
| monitor_type     | String | 监控类型：file表示日志文件，resource表示资源使用                       |
| monitor_path     | String | 可选参数，QingTian Enclave中要监控的日志文件路径，monitor_type只能是file |
| monitor_interval | u32    | 可选参数，资源监视间隔时间，单位是秒，默认值是15，monitor_type只能是resource    |
| outputfile       | String | 监控数据文件名称，写入监控数据的路径是 <workspace>/<outputfile>         |

### 配置文件参考

```
port: 6000
workspace: /var/log/qlog
server_logfile: server.log
client_logfile: client.log
server_threads: 4
client_threads: 1
log_level: info
rotate_size: 65536
rotate_num: 10
monitor_items:
- name: item1
 monitor_type: file
 monitor_path: /var/log/item1.log
 outputfile: output1.log
- name: item2
 monitor_type: file
 monitor_path: /var/log/item2.log
 outputfile: output2.log
- name: item3
 monitor_type: file
 monitor_path: /var/log/item3.log
 outputfile: output3.log
- name: item4
 monitor_type: file
 monitor_path: /var/log/item4.log
 outputfile: output4.log
- name: item5
 monitor_type: resource
 monitor_interval: 15
 outputfile: output5.log
```

## 18.6 QingTian CLI ( qt CLI )

## 18.6.1 安装 qt CLI

如果您使用其他Linux镜像，请通过华为云QingTian开源仓库进行编译安装。如果您使用Huawei Cloud EulerOS系统镜像时，通过以下命令直接安装：

```
yum install qingtian-tool
```

该rpm包中包含两个工具：

qt-enclave-env：提供资源隔离功能，QingTian Enclave虚拟机创建之前需要隔离指定内存供enclave虚拟机使用，以保证QingTian Enclave虚拟机空间的绝对安全。

qt CLI：是一个QingTian命令行工具。您可以使用qt CLI制作QingTian Enclave虚拟机启动所需要的EIF镜像，和管理QingTian Enclave虚拟机生命周期。

注意：在使用qt CLI之前，您需要预先安装python3以及几个必要的python module：docker 和knack。可以参考以下命令进行安装：

```
pip3 install docker knack
```

## 18.6.2 qt-enclave-env 工具介绍

qt-enclave-env是一个service，service启动后从qt-enclave-env.conf配置文件中读取需要隔离的资源信息，并执行资源隔离操作。隔离操作需要在创建QingTian Enclave虚拟机之前执行。

接下来我们将介绍该服务的配置文件/etc/qingtian/enclave/qt-enclave-env.conf：

```
#enclave虚拟机隔离大页内存类型，可配置2或1024，分别表示2M大页或1G大页
hugepage_size:1024
配置需要隔离的内存大小，数值需要为hugepage_size的整数倍。
memory_mib:1024
配置需要隔离的cpu个数，该配置项与cpu_list互斥，二者只能配置一个，否则service启动失败
cpu_count:2
配置需要隔离的cpu列表，可填入0以外的其他cpu id，该配置项与cpu_count互斥，二者只能配置一个，否则
service启动失败
cpu_list:2,3
```

需要注意的是：qt-enclave-env服务预留大页内存是否成功受父虚拟机本身内存碎片化影响，系统长时间运行或者反复重启qt-enclave-env可能导致无法预留出足够的大页内存。推荐的使用方式是在系统刚启动时，启动一次qt-enclave-env预留足够的内存即可。

## 18.6.3 qt enclave 子命令介绍

qt为一级命令，其下目前包含enclave一个二级子命令：

```
[root@localhost ~]# qt
 /_ \() |_ _()
 || ||_--_|| | _--_
 || ||_ _ /_ || ||_ _ /_
 || || || | | || || | | ||
 _ _ || | | | | _ _ ||
 _/ |
 |_/

Welcome to the cool QingTian new CLI!

enclave : Enclave life-circle management.
```

qt enclave中提供enclave相关的镜像制作，拉起、销毁、查询QingTian Enclave等子命令：

```
[root@localhost ~]# qt enclave
usage: qt enclave [-h] {make-img,start,stop,query,console} ...
qt enclave: error: the following arguments are required: _subcommand
enclave command line interface
[root@localhost ~]# qt enclave -h

Group
 qt enclave : Enclave life-circle management.

Commands:
 console : Console an enclave via the enclave-id while debugging.
 make-img : Make an eif image from a docker image.
 query : Query an enclave via the enclave-id or query all enclaves.
 start : Start an enclave via an eif image.
 stop : Stop an enclave via the enclave-id.
```

## qt enclave make-img

该命令用于将用户制作好的docker镜像转换为QingTian Enclave虚拟机可用的镜像，命令格式如下：

```
[root@localhost ~]# qt enclave make-img -h

Command
 qt enclave make-img : Make an eif image from a docker image.

Arguments
 --docker-uri [Required]
 --eif [Required]
 --private-key
 --signing-certificate

Global Arguments
 --debug : Increase logging verbosity to show all debug logs.
 --help -h : Show this help message and exit.
 --only-show-errors : Only show errors, suppressing warnings.
 --output -o : Output format. Allowed values: json, jsonc, none, table, tsv, yaml,
 yamlc. Default: json.
 --query : JMESPath query string. See http://jmespath.org/ for more information
 and examples.
 --verbose : Increase logging verbosity. Use --debug for full debug logs.

Examples
 Given docker-uri and eif to make an eif image
 qt enclave make-img --docker-uri [DOCKER-URI] --eif [EIF]

 Make an eif image with private-key and signing-certificate
 qt enclave make-img --docker-uri [DOCKER-URI] --eif [EIF] --private-key [PRIVATE-KEY]
 --signing-certificate [SIGNING-CERTIFICATE]
```

必选项：--docker-uri，指定Docker存储库中Docker映像的统一资源标识符（URI）。可以通过docker image ls命令查询当前本地镜像uri。

必选项：--eif，提供路径名称用于存放生成后的EIF镜像

可选项：--private-key，提供对QingTian Enclave镜像进行签名的私钥绝对路径，如果指定了PRIVATE-KEY，要求必须同时指定签名密钥SIGNING-CERTIFICATE；

可选项：--signing-certificate，提供对QingTian Enclave镜像进行签名的证书绝对路径，如果指定了SIGNING-CERTIFICATE，如要同时指定签名密钥PRIVATE-KEY；

命令行返回值：如果配置了以上两项证书可选项，必须保证证书有效。证书有效情况下命令输出将包括额外的PCR0，PCR8散列值用于度量Enclave镜像与签名证书。证书无效，镜像创建失败。

创建镜像命令示例：

```
[root@localhost docker]# qt enclave make-img --docker-uri ubuntu --eif /home/docker/ubuntu.eif --private-key /home/docker/private-key.pem --signing-certificate /home/docker/server.pem
{
 "digest": "SHA384",
 "PCR0":
"b8c59692da8a5bcb739a83d15a0ceca670bd78da06cb2250ec70548f72254e674419e9888db9c0364a9b88dd58017a62"
 "PCR8":
"dbf4a7f9fab7f18619b5899c407081981ad6762fb9a809da78548821b5021965423181584acd7b201703376f1133a546"
}
```

## qt enclave start

该命令用于创建QingTian Enclave虚拟机，命令格式如下：

```
[root@localhost ~]# qt enclave start -h
```

### Command

qt enclave start : Start an enclave via an eif image.

### Arguments

--cid : Default: 4.  
--eif [Required]  
--cpus : Default: 2.  
--debug-mode  
--mem : Default: 1024.

### Global Arguments

--debug : Increase logging verbosity to show all debug logs.  
--help -h : Show this help message and exit.  
--only-show-errors : Only show errors, suppressing warnings.  
--output -o : Output format. Allowed values: json, jsonc, none, table, tsv, yaml, yamlc.  
Default: json.  
--query : JMESPath query string. See <http://jmespath.org/> for more information and examples.  
--verbose : Increase logging verbosity. Use --debug for full debug logs.

### Examples

Given an eif image, an unused cid, the number of cpus and memory needed  
qt enclave start [--cpus CPUS] [--mem MEM] --eif EIF [--cid CID]

可选项：--cpus，指定要分配给enclave虚拟机的vCPU数量，不能大于隔离的cpu数目，未配置使用默认值2；

可选项：--mem，指定分配给QingTian Enclave虚拟机的内存大小（MB），不能大于隔离内存大小，需大于QingTian Enclave镜像大小，未配置使用默认值1024MB；

必选项：--eif，指定EIF镜像路径；

可选项：--cid，设置QingTian Enclave虚拟机cid，用于指定父虚拟机与QingTian Enclave虚拟机间vsock通信的socket IP。可用的cid范围为：4-4294967294，未配置使用默认值4；

可选项：--debug-mode，指定是否在调试模式下启动QingTian Enclave实例，该模式下使用全部为0的PCR散列值，可以收集打印QingTian Enclave虚拟机内部日志；

命令行返回值：创建成功后输出创建的QingTian Enclave虚拟机详细信息。

Enclave虚拟机启动命令示例：

```
qt enclave start --cpus 2 --mem 1024 --eif /home/docker/ubuntu.eif --cid 4
```

## qt enclave query

该命令用于在父虚拟机内查询当前已创建的QingTian Enclave虚拟机信息，命令格式如下：

```
[root@localhost ~]# qt enclave query -h

Command
qt enclave query : Query an enclave via the enclave-id or query all enclaves.

Arguments
--enclave-id

Global Arguments
--debug : Increase logging verbosity to show all debug logs.
--help -h : Show this help message and exit.
--only-show-errors : Only show errors, suppressing warnings.
--output -o : Output format. Allowed values: json, jsonc, none, table, tsv, yaml, yamlc.
 Default: json.
--query : JMESPath query string. See http://jmespath.org/ for more information and
 examples.
--verbose : Increase logging verbosity. Use --debug for full debug logs.

Examples
Given an enclave-id to query an enclave
qt enclave query --enclave-id [ENCLAVE-ID]

Query all enclaves without enclave-id
qt enclave query
```

可选项：--enclave-id，query时带此参数则查询指定QingTian Enclave虚拟机信息，否则查询全部当前已存在QingTian Enclave虚拟机信息；

命令行返回值，查询到的QingTian Enclave虚拟机信息：

- EnclaveID：QingTian Enclave虚拟机的id号；
- ProcessID：父虚拟机中持有QingTian Enclave资源的进程号PID；
- EnclaveCID：QingTian Enclave虚拟机与父虚拟机通信使用的vsock socket id；
- NumberOfCPUs：从父虚拟机分配给QingTian Enclave虚拟机的vCPU个数；
- MemoryMiB：从父虚拟机分配给QingTian Enclave虚拟机的内存大小(MB)；

QingTian Enclave虚拟机查询命令示例：

```
[root@localhost ~]# qt enclave query
[{
 "EnclaveID": 0,
 "ProcessID": 29990,
 "EnclaveCID": 4,
 "NumberOfCPUs": 2,
 "MemoryMiB": 1024,
 "LaunchMode": "debug"
}]
```

若当前无QingTian Enclave虚拟机存在，则该命令查询返回空；

带--enclave-id参数查询场景，若指定enclave-id的QingTian Enclave虚拟机不存在，查询命令返回空。

## qt enclave stop

该命令用于在父虚拟机内销毁已创建的QingTian Enclave虚拟机，命令格式如下：

```
[root@localhost ~]# qt enclave stop -h
```

```
Command
qt enclave stop : Stop an enclave via the enclave-id.

Arguments
--enclave-id [Required]

Global Arguments
--debug : Increase logging verbosity to show all debug logs.
--help -h : Show this help message and exit.
--only-show-errors : Only show errors, suppressing warnings.
--output -o : Output format. Allowed values: json, jsonc, none, table, tsv, yaml,
 yamlc. Default: json.
--query : JMESPath query string. See http://jmespath.org/ for more information
 and examples.
--verbose : Increase logging verbosity. Use --debug for full debug logs.

Examples
Given an enclave-id to stop an enclave
qt enclave stop --enclave-id [ENCLAVE-ID]
```

必选项: --enclave-id, 指定需要销毁的QingTian Enclave虚拟机的enclave-id

命令行返回值: 成功返回销毁成功, 销毁失败无返回。

QingTian Enclave虚拟机销毁命令示例:

```
[root@localhost ~]# qt enclave stop --enclave-id 1
stop 1 success
```

## qt enclave console

在启动QingTian Enclave时, 指定为debug-mode时, 用于在父虚拟机中查看QingTian Enclave中的只读控制台输出, 命令格式如下:

```
[root@localhost ~]# qt enclave console -h

Command
qt enclave console : Console an enclave via the enclave-id while debugging.

Arguments
--enclave-id [Required]

Global Arguments
--debug : Increase logging verbosity to show all debug logs.
--help -h : Show this help message and exit.
--only-show-errors : Only show errors, suppressing warnings.
--output -o : Output format. Allowed values: json, jsonc, none, table, tsv, yaml,
 yamlc. Default: json.
--query : JMESPath query string. See http://jmespath.org/ for more information
 and examples.
--verbose : Increase logging verbosity. Use --debug for full debug logs.

Examples
Given an enclave-id to console an enclave
qt enclave console --enclave-id [ENCLAVE-ID]
```

必选项: --enclave-id, 指定需要获取只读控制台输出的QingTian Enclave虚拟机的enclave-id。

命令行执行成功后, 会打印QingTian Enclave虚拟机的只读控制台输出, 如下所示:

```
hello enclave!
hello enclave!
hello enclave!
hello enclave!
```

您可以使用ctrl+c的方式退出该命令。需要注意的是, 在同一时间我们只允许一个qt enclave console命令作用于一个指定的QingTian Enclave实例。

## 18.7 错误码

| 错误码 | 错误信息                                                     | 描述                         | 处理措施                             |
|-----|----------------------------------------------------------|----------------------------|----------------------------------|
| 01  | Missing necessary argument.                              | 必要参数缺失                     | 请您对命令参数进行检查                      |
| 02  | Invalid argument provided.                               | 无效参数                       | 请您对命令参数进行检查                      |
| 03  | File operation failure.                                  | 文件操作错误                     | 请您检查目标文件或目录是否存在                  |
| 04  | loctl get sandbox capacity failure.                      | loctl获取 sandbox capacity错误 | 请您联系华为云技术支撑人员帮您处理                |
| 05  | loctl define sandbox failure.                            | loctl定义 sandbox错误          | 请您联系华为云技术支撑人员帮您处理                |
| 06  | Invalid parameters provided in configuration file.       | 配置文件中存在无效参数                | 请您检查对应配置文件                       |
| 07  | Missing necessary parameters in configuration file.      | 配置文件中缺少必要参数                | 请您检查对应配置文件                       |
| 08  | Mmap memory failure.                                     | Mmap内存错误                   | 请您联系华为云技术支撑人员帮您处理                |
| 09  | loctl add memory failure.                                | loctl增加内存错误                | 请您联系华为云技术支撑人员帮您处理                |
| 10  | Load image failure because provided memory is too small. | 加载镜像错误, 可能是因为设置的内存大小不够     | 请您在启动QingTian Enclave时, 增加内存参数设置 |
| 11  | loctl add cpu failure.                                   | loctl增加vcpu错误              | 请您联系华为云技术支撑人员帮您处理                |
| 12  | Lock acquire failure.                                    | 请求锁失败                      | 请您检查qt CLI日志信息, 查看对应锁文件权限是否正常    |
| 13  | Socket initialization failure.                           | Socket初始化失败                | 请您联系华为云技术支撑人员帮您处理                |
| 14  | Socket binding failure.                                  | Socket绑定错误                 | 请您联系华为云技术支撑人员帮您处理                |
| 15  | Socket listen failure.                                   | Socket监听错误                 | 请您联系华为云技术支撑人员帮您处理                |



| 错误码 | 错误信息                                          | 描述                           | 处理措施                                                    |
|-----|-----------------------------------------------|------------------------------|---------------------------------------------------------|
| 16  | Socket accept failure.                        | Socket执行接收错误                 | 请您联系华为云技术支撑人员帮您处理                                       |
| 17  | Write heartbeat to the enclave failure.       | 给QingTian Enclave写入心跳信息出错    | 请您联系华为云技术支撑人员帮您处理                                       |
| 18  | Read heartbeat from the enclave failure.      | 读取QingTian Enclave心跳信息出错     | 请您联系华为云技术支撑人员帮您处理                                       |
| 19  | loctl start an enclave failure.               | loctl启动QingTian Enclave失败    | 请您联系华为云技术支撑人员帮您处理                                       |
| 20  | Wait heartbeat timeout.                       | 等待心跳信息超时                     | 请您在启动QingTian Enclave时，增加内存参数配置；如未能解决，请您联系华为云技术支撑人员帮您处理 |
| 21  | Get json print object failure.                | 获取json打印对象失败                 | 请您检查cjson库是否正常                                          |
| 22  | Write enclave's configuration file failure.   | 生成QingTian Enclave配置文件失败     | 请您查看qt CLI日志信息，排查相关文件权限；如未能解决，请您联系华为云技术支撑人员帮您处理         |
| 23  | Socket connection failure                     | Socket连接错误                   | 请您联系华为云技术支撑人员帮您处理                                       |
| 24  | Write cmd to the enclave server failure.      | 向QingTian Enclave后台服务写入命令失败  | 请您联系华为云技术支撑人员帮您处理                                       |
| 25  | Read message from the enclave server failure. | 获取QingTian Enclave后台服务信息失败   | 请您联系华为云技术支撑人员帮您处理                                       |
| 26  | Create cJSON object failure.                  | 创建cjson对象失败                  | 请您检查cjson库是否正常                                          |
| 27  | Create cJSON array failure.                   | 创建cjson数组失败                  | 请您检查cjson库是否正常                                          |
| 28  | The required enclave is not running.          | 被请求的QingTian Enclave实例并没有在运行 | 请您使用qt enclave query命令查看目前正在运行的QingTian Enclave         |
| 29  | Invalid enclave pid.                          | 无效的Enclave PID               | 请您联系华为云技术支撑人员帮您处理                                       |

| 错误码 | 错误信息                                                         | 描述                                  | 处理措施                                         |
|-----|--------------------------------------------------------------|-------------------------------------|----------------------------------------------|
| 30  | Add number into cJSON object failure.                        | 向cJSON打印对象中增加数字失败                   | 请您检查cJSON库是否正常                               |
| 31  | Add string into cJSON object failure.                        | 向cJSON打印对象中增加字符串失败                  | 请您检查cJSON库是否正常                               |
| 32  | The required enclave is not running in the debug mode.       | 被请求的QingTian Enclave实例没有运行在debug模式下 | 请您使用qt enclave query 查询该QingTian Enclave运行模式 |
| 33  | Enclave console read failure.                                | QingTian Enclave console 命令读取失败     | 请您联系华为云技术支撑人员帮您处理                            |
| 34  | Write img header failure.                                    | 创建镜像时，写入镜像头文件失败                     | 请您联系华为云技术支撑人员帮您处理                            |
| 35  | Write cmdline failure.                                       | 创建镜像时，写入cmdline失败                   | 请您联系华为云技术支撑人员帮您处理                            |
| 36  | Write kernel failure.                                        | 创建镜像时，写入kernel失败                    | 请您联系华为云技术支撑人员帮您处理                            |
| 37  | Write initrd failure.                                        | 创建镜像时，写入initrd失败                    | 请您联系华为云技术支撑人员帮您处理                            |
| 38  | Write certificate failure.                                   | 创建镜像时，写入证书失败                        | 请您联系华为云技术支撑人员帮您处理                            |
| 39  | Get pcr failure.                                             | 获取PCR值失败                            | 请您联系华为云技术支撑人员帮您处理                            |
| 40  | Add signature failure.                                       | 创建镜像时，对镜像签名出错                       | 请您联系华为云技术支撑人员帮您处理                            |
| 41  | Check enclave image info failure while building an eif file. | 构建eif镜像时检查到enclave镜像信息异常            | 请您联系华为云技术支撑人员帮您处理                            |
| 42  | The required enclave is in maintenance state.                | 被请求的QingTian Enclave实例正处于维护阶段       | 请您联系华为云技术支撑人员帮您处理                            |
| 43  | The cid has already been used.                               | cid已经在使用中                           | 请您重新指定未在使用中的cid                              |

## 18.8 常见问题

### 18.8.1 通用类

#### 什么是 QingTian Enclave?

QingTian Enclave是基于QingTian架构虚拟机产品所提供的的一个安全特性。借助QingTian Enclave提供的隔离、高度受限的环境，客户可以部署其安全敏感型应用程序，缩小其攻击面。

#### 为什么需要使用 QingTian Enclave?

QingTian Enclave在普通的ECS实例中提供了隔离的计算环境，以处理高度敏感的数据。

QingTian Enclave是完全独立的虚拟机，无持久化存储、交互式访问或外部网络连接，您的ECS实例通过安全的本地通道与QingTian Enclave之间进行通信。

#### 什么时候需要使用 QingTian Enclave?

当您需要处理安全敏感型数据，并希望处理该数据的过程与有权访问您实例的用户、应用程序或者三方库隔离时，您应该使用QingTian Enclave。

您可以利用QingTian Enclave开发和运行各种程序，以应对不同的使用场景，比如个人隐私信息处理，专有代码和算法运算，多方计算等。

#### 如何开始使用 QingTian Enclave?

您可以参考QingTian Enclave[快速入门](#)，来开启您的QingTian Enclave开发之旅。

#### 什么是内存和 CPU 隔离?

内存和CPU隔离技术可以防止ECS实例上的用户、应用程序和第三方库直接访问已被隔离给QingTian Enclave使用的内存和CPU。您可以通过华为云QingTian命令行工具（qt CLI），使用这部分被隔离的内存和CPU启动QingTian Enclave虚拟机，详细操作请参考[QingTian CLI \(qt CLI\)](#)。

#### QingTian Enclave 是如何保证内存和 CPU 隔离的?

QingTian Enclave采用经过验证的基于CPU的隔离技术，并结合QingTian架构的独特设计和基于华为自研智能网卡的可信根证书，共同实现隔离。由华为云开发和设计的极简QingTian Hypervisor可以将服务器的物理资源进行分区。相比于其他虚拟化技术，QingTian Hypervisor裁减了所有不必要的功能。QingTian Enclave扩展了QingTian Hypervisor的隔离功能，保护了分配给QingTian Enclave的CPU和内存，保证其与ECS实例的CPU和内存资源相隔离，为您提供一个完全隔离的可执行环境。

#### 哪些实例规格支持 QingTian Enclave?

目前c7t实例已支持QingTian Enclave特性。

## 什么是证明（Attestation）文档？

证明文档用于证明QingTian Enclave实例的可信度量结果。证明文档由QingTian Hypervisor生成，文档内容包括PCR（Platform configuration registers，平台配置寄存器）列表、QingTian Attestation PKI（Public Key Infrastructure，公钥基础设施）证书链、密码算法声明以及Enclave应用自定义数据。

证明文档由华为云QingTian Attestation PKI签署。华为云KMS（Key management Service，密钥管理服务）内置了对QingTian Enclave证明的支持。通过使用QingTian Enclave SDK中包含的华为云KMS API，您可以在QingTian Enclave实例中基于QingTian Enclave证明来执行华为云KMS操作，比如解密、生成随机数和加密等操作。华为云KMS服务提取来自QingTian Enclave的证明文档并根据预设的华为云IAM（Identity and Access Management，统一身份认证服务）授权策略对其进行访问权限控制。

## QingTian Enclave 证明文件的信任根源是什么？该如何验证它？

证明文档由华为云QingTian Attestation PKI（Public Key Infrastructure）签署。华为云QingTian根证书可以从[https://qingtian-enclave.obs.myhuaweicloud.com/huawei\\_qingtian-enclaves\\_root-G1.zip](https://qingtian-enclave.obs.myhuaweicloud.com/huawei_qingtian-enclaves_root-G1.zip)下载获取。文档签名验证方法请参考[文档签名验证](#)。

## QingTian Enclave 如何进行收费？

目前使用QingTian Enclave不会收取额外费用，您只需要支付ECS实例的购买费用。

## 为什么执行隔离服务会失败（systemctl start qt-enclave-env）？

在系统运行过程中，不可避免地会产生碎片化内存，使得隔离服务时无法获取连续内存大页。

执行**systemctl status qt-enclave-env**命令，查看错误日志内是否包含**allocating hugepages error**。

如果包含该错误日志，则说明此时系统内所提供的可用空闲连续大页资源少于您所期望的内存大页数量（\$wanted\_mem\_num）。

您可以采用如下两种方案进行处理。

- 方案一：
  - a. 执行以下命令，查看当前系统所能提供的最大可用连续内存大页数量（\$free\_mem\_num）。

```
cat /sys/devices/system/node/node0/hugepages/hugepages-1048576kB/free_hugepages
```
  - b. 执行以下命令，修改/etc/qingtian/enclave/qt-enclave-env.conf配置文件，使得memory\_mib的值低于(\$free\_mem\_num \* 1024)。

```
vim /etc/qingtian/enclave/qt-enclave-env.conf
```

建议您在系统启动后立即执行资源隔离服务，避免反复执行隔离命令造成内存碎片化。
- 方案二：
  - a. 执行以下命令，修改grub文件（/etc/default/grub）。

```
vim /etc/default/grub
```

在grub文件的Linux命令行参数 ( GRUB\_CMDLINE\_LINUX ) 中, 添加 ( default\_hugepagesz=1G hugepagesz=1G hugepages=\$wanted\_mem\_num ) 。

- b. 执行以下命令, 使得配置生效。  
grub2-mkconfig -o /boot/efi/EFI/hce/grub.cfg
- c. 重启虚拟机, 重启后可再次查看可用大页数量。  
**reboot**

## 18.8.2 开发部署类

### 一个 ECS 实例可以部署多少个 QingTian Enclave?

QingTian Enclave支持为每个ECS实例创建最多2个Enclave虚拟机。

### 什么是 Vsock? 我该如何使用 Vsock 与 QingTian Enclave 进行通信?

Vsock是一种套接字接口, 由上下文 ID (CID) 和端口号定义, 上下文ID与TCP/IP连接中的IP地址相同。

Vsock使用标准、定义明确的POSIX套接字API ( 例如连接、侦听、接受 ) 与QingTian Enclave进行通信。您可以开发基于vsock与QingTian Enclave通信的程序, 详细操作请参考[Linux系统上QingTian Enclave应用的开发](#)。另外应用程序也可以借助代理, 通过vsock发送HTTP请求。

### 为什么使用 kc2 实例启动 QingTian Enclave 后, 出现 Vsock 性能下降?

对于kc2实例, 如果使用隔离的全部CPU启动QingTian Enclave, 会导致Vsock性能下降。

建议您在启动隔离服务时, 比预期QingTian Enclave使用的CPU数量, 多隔离2个CPU。这样可以使被隔离服务的一部分CPU用于启动QingTian Enclave, 一部分CPU用于Vsock转发, 避免出现Vsock性能下降。

1. 执行以下命令, 修改qt-enclave-env服务的配置文件/etc/qingtian/enclave/qt-enclave-env.conf。  
**vim /etc/qingtian/enclave/qt-enclave-env.conf**
  - 方式一: 修改cpu\_count数量为QingTian Enclave的CPU数量加2。
  - 方式二: 修改cpu\_list中的CPU数量为QingTian Enclave的CPU数量加2。
2. 执行以下命令, 重启qt-enclave-env服务。  
**systemctl restart qt-enclave-env**
3. 执行以下命令, 重启QingTian Enclave。  
**qt enclave start --cpus \${isolated\_cpus\_count-2} --mem \${wanted\_mem} --eif \${eif\_file\_location} --cid \${wanted\_cid}**

其中:

  - isolated\_cpus\_count-2: 隔离出来的CPU数量减少2
  - wanted\_mem: 期望内存大小
  - eif\_file\_location: eif文件位置
  - wanted\_cid: 期望cid值重启QingTian Enclave后观察vsock性能。

## ECS 实例开启 SELinux 后，为什么启动 qt-enclave-env 服务会失败？

问题现象：ECS实例开启SELinux后，启动qt-enclave-env服务失败，查看qt-enclave-env服务日志，失败的日志显示“insmod virtio-qtbox.ko Permission denied”。

问题原因：SELinux通过强制访问控制、细粒度的访问控制、策略强制、类型强制、安全上下文和审计功能等特性，提供了强大的安全机制，可以帮助保护Linux 系统免受恶意攻击和数据泄露的威胁。因此，qt-enclave-env服务中无法直接调用insmod virtio-qtbox.ko命令插入内核模块。

解决方案：请您通过手动执行命令insmod /opt/qingtian/enclave/virtio-qtbox.ko，或者关闭SELinux特性，并重启qt-enclave-env服务规避该问题。