

UCLLOUD 优刻得

中国第一家公有云科创板上市公司
股票代码：688158

UCloudStack 私有云产品白皮书

优刻得私有云
构建下一代可持续云基础设施
赋能企业未来

版权信息

版权所有©2024 优刻得科技股份有限公司保留一切权利。

本档中出现的任何文字叙述、文档格式、图片、方法及过程等内容，除另有特别注明外，其著作权或其它相关权利均属于优刻得科技股份有限公司。非经优刻得科技股份有限公司书面许可，任何单位和个人不得以任何方式和形式对本档内的任何部分擅自进行摘抄、复制、备份、修改、传播、翻译成其它语言、将其全部或部分用于商业用途。

注意

您购买的产品、服务或特性等应受优刻得科技股份有限公司商业合同和条款约束，本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权利范围之内。除非合同另有约定，优刻得科技股份有限公司对本档内容不做任何明示或暗示的声明或保证。

关于文档

优得刻科技股份有限公司在编写本档时已尽最大努力保证其内容准确可靠，但优得刻科技股份有限公司不对本文本中的遗漏、不准确或错误导致的损失和损害承担责任。

由于产品版本升级或其它原因，本档内容会不定期更新，除非另有约定，本档仅作为使用指导，本档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

目录

前言	13
1 产品简介	15
1.1 产品概述	15
1.2 核心优势	15
1.3 产品架构	17
1.3.1 基础设施	17
1.3.2 虚拟核心引擎	18
1.3.3 统一资源调度	18
1.3.4 统一计算平台	19
1.3.5 统一存储服务	20
1.3.6 分布式虚拟网络	21
1.3.7 PaaS 服务	23
1.3.8 统一运营服务	24
1.3.9 统一运维服务	25
1.3.10 统一管理平台	28
1.4 技术架构特性	29
1.4.1 API 幂等性	29
1.4.2 全异步架构	29
1.4.3 分布式	30
1.4.4 高可用	33
1.4.5 业务实现分离	34
1.4.6 组件化	34
1.5 客户痛点	35
1.5.1 自建私有云的痛点	35
1.5.2 解决之道	35
1.6 应用场景	36
1.6.1 虚拟化&云化	36
1.6.2 业务快速交付	36

1.6.3 超融合一体机.....	36
1.6.4 政企专有云.....	36
1.7 交付和服务.....	37
2 平台物理架构.....	40
2.1 物理集群节点.....	40
2.1.1 管理节点.....	40
2.1.2 超融合节点.....	40
2.1.3 独立计算节点.....	42
2.1.4 独立存储节点.....	42
2.1.5 商业存储节点.....	42
2.1.6 推荐节点方案.....	43
2.2 物理网络架构.....	45
2.2.1 架构规模.....	45
2.2.2 网络区域.....	46
2.2.3 服务器区域.....	47
2.2.4 标准架构扩展.....	48
2.3 硬件选型.....	50
2.3.1 最低硬件配置.....	50
2.3.2 推荐硬件配置.....	52
2.4 平台资源占用.....	55
2.5 机柜空间规划.....	56
3 平台技术架构.....	58
3.1 计算虚拟化.....	58
3.1.1 CPU 超分.....	60
3.1.2 镜像文件.....	61
3.1.3 GPU 透传.....	62
3.1.4 USB 透传.....	63
3.1.5 物理机纳管.....	64
3.1.6 集群平滑扩容.....	64
3.2 智能调度.....	65

3.2.1 均衡调度.....	66
3.2.2 亲和策略.....	66
3.2.3 在线迁移.....	66
3.2.4 离线迁移.....	69
3.2.5 宕机迁移.....	69
3.3 存储虚拟化.....	71
3.3.1 分布式存储.....	72
3.3.2 智能存储集群.....	74
3.3.3 超大规模扩展.....	76
3.3.4 高可用和高可靠.....	77
3.3.5 多副本冗余机制.....	78
3.3.6 数据重均衡.....	81
3.3.7 数据故障重建.....	84
3.3.8 数据清洗.....	85
3.3.9 自动精简配置.....	85
3.3.10 存储功能简介.....	86
3.4 网络虚拟化.....	89
3.4.1 分布式网络.....	93
3.4.2 分布式架构.....	93
3.4.3 通信机制.....	96
3.4.4 SDN 控制器.....	98
3.4.5 网络功能简介.....	99
3.5 复用公有云.....	100
3.6 一云多芯架构.....	101
3.7 混合云架构.....	103
4 核心产品服务.....	108
4.1 基本概念.....	108
4.1.1 地域.....	108
4.1.2 集群.....	108
4.1.3 存储集群.....	111

4.2 虚拟机.....	112
4.2.1 概述.....	112
4.2.2 实例规格.....	113
4.2.3 生命周期管理.....	114
4.2.4 镜像服务.....	115
4.2.5 虚拟机存储.....	118
4.2.6 存储热迁移.....	119
4.2.7 虚拟机网络.....	122
4.2.8 安全组.....	127
4.2.9 隔离组.....	132
4.2.10 USB 透传.....	134
4.2.11 VNC 登录.....	135
4.2.12 自定义启动源.....	135
4.2.13 自定义主机名称.....	136
4.2.14 自定义 DNS.....	136
4.2.15 自定义 MAC.....	136
4.2.16 自定义引导方式.....	136
4.2.17 自定义 CPU 启动模式.....	137
4.2.18 自定义高可用模式.....	137
4.3 GPU 虚拟机.....	138
4.3.1 概述.....	138
4.3.2 应用场景.....	138
4.4 云硬盘.....	139
4.4.1 概述.....	139
4.4.2 功能与特性.....	140
4.4.3 应用场景.....	142
4.5 共享云盘.....	142
4.6 快照服务.....	143
4.7 商业存储服务.....	144
4.7.1 概述.....	144

4.7.2 功能与特性.....	145
4.7.3 使用流程.....	146
4.8 私有网络.....	148
4.8.1 VPC 概述.....	148
4.8.2 VPC 逻辑结构.....	149
4.8.3 VPC 连接.....	151
4.8.4 功能与特性.....	152
4.8.5 自定义路由.....	154
4.8.6 网络拓扑.....	154
4.8.7 VPC 互通.....	155
4.9 组播.....	155
4.9.1 概述.....	155
4.9.2 组播组成员.....	156
4.9.3 组播路由器.....	156
4.9.4 组播地址.....	156
4.9.5 组播转发机制.....	157
4.10 外网 IP.....	157
4.10.1 物理架构.....	158
4.10.2 逻辑架构.....	158
4.10.3 EIP 通信模式.....	160
4.10.4 功能特性.....	161
4.11 高可用 VIP.....	163
4.11.1 概述.....	163
4.11.2 工作机制.....	163
4.12 NAT 网关.....	164
4.12.1 产品概述.....	164
4.12.2 应用场景.....	165
4.12.3 架构原理.....	165
4.12.4 功能特性.....	167
4.13 负载均衡.....	173

4.13.1 产品概述.....	173
4.13.2 应用场景.....	173
4.13.3 架构原理.....	174
4.13.4 功能特性.....	176
4.13.5 负载均衡高可用.....	180
4.13.6 SSL 证书.....	180
4.13.7 LB 安全性.....	181
4.14 IPSECVPN 服务.....	182
4.14.1 背景.....	182
4.14.2 概述.....	183
4.14.3 逻辑架构.....	183
4.14.4 VPN 隧道建立.....	186
4.14.5 VPN 隧道参数.....	187
4.14.6 应用场景.....	190
4.14.7 使用流程.....	190
4.15 裸金属.....	191
4.15.1 概述.....	191
4.15.2 使用流程.....	192
4.16 弹性伸缩.....	193
4.16.1 概述.....	193
4.16.2 水平伸缩.....	193
4.16.3 垂直伸缩.....	198
4.17 备份服务.....	199
4.17.1 概述.....	199
4.17.2 存储池.....	199
4.17.3 备份任务.....	199
5 PaaS 产品服务.....	202
5.1 MySQL 服务.....	202
5.1.1 产品概述.....	202
5.1.2 实例管理.....	203

5.1.3 从库管理.....	205
5.1.4 参数配置.....	205
5.1.5 备份管理.....	205
5.1.6 监控告警.....	206
5.1.7 日志事件.....	206
5.1.8 参数模板.....	206
5.2 REDIS 服务	207
5.2.1 产品概述.....	207
5.2.2 实例管理.....	208
5.2.3 从库管理.....	209
5.2.4 参数配置.....	210
5.2.5 备份管理.....	210
5.2.6 监控告警.....	211
5.2.7 日志事件.....	211
5.2.8 参数模板.....	211
5.3 文件存储.....	212
5.3.1 概述.....	212
5.3.2 实例管理.....	213
5.3.3 实例扩容.....	214
5.3.4 监控告警.....	215
5.3.5 日志事件.....	215
5.3.6 备份管理.....	215
5.3.7 文件管理.....	216
5.4 对象存储.....	216
5.4.1 概述.....	216
5.4.2 实例管理.....	217
5.4.3 实例扩容.....	219
5.4.4 监控告警.....	219
5.4.5 日志事件.....	220
5.4.6 备份管理.....	220

5.4.7 桶管理.....	220
5.4.8 桶文件管理.....	221
6 运维运营管理.....	222
6.1 统一管理服务.....	222
6.2 平台管理账号.....	223
6.2.1 管理员概述.....	223
6.2.2 管理员账号安全.....	224
6.2.3 管理员账号管理.....	225
6.2.4 管理员权限管理.....	226
6.3 多租户管理.....	227
6.3.1 概述.....	227
6.3.2 租户管理.....	229
6.3.3 租户自服务.....	231
6.3.4 账号权限管理.....	232
6.4 多地域管理.....	236
6.4.1 概述.....	236
6.4.2 多地域特性.....	237
6.4.3 多地域管理能力.....	238
6.5 全局资源视图.....	238
6.6 物理资源管理.....	239
6.6.1 物理节点管理.....	239
6.6.2 镜像管理.....	241
6.6.3 外网网段管理.....	242
6.6.4 专线接入管理.....	244
6.7 虚拟资源管理.....	244
6.8 QoS 配置管理.....	245
6.9 资源模板.....	246
6.10 标签管理.....	246
6.10.1 概述.....	246
6.10.2 资源类型.....	247

6.10.3 使用限制.....	247
6.11 监报告警.....	248
6.11.1 概述.....	248
6.11.2 监控图表.....	249
6.11.3 告警模板.....	250
6.11.4 告警记录.....	252
6.12 通知组.....	253
6.13 操作日志.....	253
6.13.1 操作日志.....	253
6.13.2 通知规则.....	255
6.14 资源事件.....	255
6.14.1 资源事件.....	255
6.14.2 通知规则.....	256
6.15 回收站.....	257
6.15.1 概述.....	257
6.15.2 恢复资源.....	258
6.15.3 续费资源.....	258
6.15.4 销毁资源.....	259
6.16 计量计费.....	259
6.16.1 概述.....	259
6.16.2 资源计价格.....	260
6.16.3 资金管理.....	260
6.16.4 价格配置.....	262
6.16.5 订单管理.....	264
6.16.6 交易管理.....	265
6.16.7 账单管理.....	266
6.17 审批流程.....	272
6.17.1 概述.....	272
6.17.2 使用流程.....	273
6.17.3 开启审批.....	274

6.17.4 审批管理.....	275
6.18 报表统计.....	276
6.18.1 资源用量统计.....	276
6.18.2 资源统计报表.....	277
6.19 大屏监控.....	279
6.20 API 控制台.....	280
7 云平台管理.....	281
7.1 客制化能力.....	281
7.1.1 自定义网站展示 UI.....	281
7.1.2 自定义监控大屏 UI.....	281
7.1.3 自定义登录页 UI.....	282
7.2 平台系统配置.....	283
7.2.1 邮箱配置.....	283
7.2.2 磁盘设置.....	284
7.2.3 网络设置.....	285
7.2.4 计费配置.....	285
7.2.5 回收策略.....	285
7.3 平台数据备份.....	286
7.4 服务目录.....	287
7.5 自定义规格.....	288
7.6 配额管理.....	290
7.7 巡检服务.....	291
7.8 统一授权.....	292
7.8.1 授权管理.....	293
7.8.2 节点管理.....	293
8 双活数据中心.....	294
8.1 概述.....	294
8.2 部署结构.....	295
8.3 双活机制.....	296
8.4 双活收益.....	299

8.5 方案场景	299
9 平台安全性	301
9.1 控制台安全性	301
9.2 账号认证授权	302
9.3 网络安全控制	303
9.4 数据存储安全	304
9.5 日志审计体系	305
10 平台可靠性	306
10.1 数据中心	306
10.2 硬件设施	306
10.3 云平台软件	307
10.4 云平台服务	309
11 灾备服务	313
11.1 本地灾备	313
11.2 异地灾备	314
11.3 公有云灾备服务	317
11.4 灾备网络架构	319
11.5 灾备切换	321
11.5.1 计划内切换	321
11.5.2 计划外切换	322
11.5.3 灾备回切	322
11.6 表格样式	323
11.7 代码样式	324

前言

UCloud（优刻得科技股份有限公司）是中立、安全的云计算服务平台，坚持中立，不涉足客户业务领域。公司自主研发 IaaS、PaaS、大数据流通平台、AI 服务平台等一系列云计算产品，并深入了解互联网、传统企业在不同场景下的业务需求，提供公有云、私有云、混合云、专有云在内的综合性行业解决方案。

依托公司在莫斯科、圣保罗、拉各斯、伦敦等全球部署的 32 大高效节能绿色数据中心，以及国内北、上、广、深、杭等 11 地线下服务站，UCloud 已为全球上万家企业级客户提供云服务支持，间接服务终端用户数量达到数亿人。

UCloud 深耕用户需求，秉持产品快速定制、贴身按需服务的理念，推出适合行业特性的产品与服务，业务已覆盖包含互联网、金融、新零售、制造、教育、政府等在内的诸多行业。

公司核心团队来自腾讯、阿里、百度、华为、VMware 等国内外知名互联网和 IT 企业，同时引进传统金融、医疗、零售、制造业等行业精英人才，目前员工总数超过 1000 人。

在云计算之前，企业业务应用上线，需要经历组网规划、容量规划、设备选型、下单、付款、发货、运输、安装、部署、调试的整个完整过程。随着云计算、大数据、人工智能等新技术对各行各业的赋能，以虚拟化和软件定义技术方向构建新一代数据中心为基础，实现业务的集中管理、资源动态调配、业务快速部署及统一运营运维，满足企业关键应用向 x86 及国产化系统迁移时，对资源高性能、高可靠、安全性、高可用及易用性上的要求，同时提高基础架构的自动化管理水平及业务快速交付能力，继而推动企业的数字化转型与业务创新。

UCloud 提供计算、存储、网络、数据库、中间件、云分发、多媒体音视频、大数据、人工智能及云安全防护等产品服务，为多种业务场景提供全方位的公有云计算产品和服务。为帮助企业级客户实现数字化转型，快速构建新一代基于虚拟化的私有数据中心，优刻得私有化解决方案，构建下一代可持续云基础设施，赋能企业未来。通过构建基于公有云且自主可控的下一代云基础设施，提供私有云、分布式存储及智能大数据平台等“底座”，凭借多年公有云运营经验和解决方案能力，助力企业数字化转型。同时依托客户实践，提供超融合和信创云架构满足敏捷交付、降低成本和信创转型等需求。

在私有云层面，UCloud 基于多年运营的公有云架构，复用服务器操作系统内核及虚拟化核心组件（KVM、OVS、Qemu、Libvirt），同时优化核心调度系统及管理平台，对公有云平台进行定制重构，缩减部署规模及业务开销，提高云平台的可靠性、可用性及可运营性，提供一套可私有化交付且与公有云用户体验一致的企业私有云平台——UCloudStack。

通过 UCloudStack 私有云平台，企业可在现有数据中心及设备上快速构建一套成熟且完整的私有云及混合云解决方案。为企业简化业务上云过程，提升组织管理和业务管理效率的同时，降低业务转型及信息系统的总体拥有成本，助力企业数字化转型。

1 产品简介

1.1 产品概述

UCloudStack 企业私有云平台，提供虚拟化、SDN 网络、分布式存储、数据库缓存、对象存储、文件存储、容器服务及云管等核心服务的统一管理、资源调度、监控日志及运营运维等一整套云资源管理能力，助力企业数字化转型。



计算虚拟化



网络虚拟化



分布式存储



云管平台



数据库服务



缓存服务



对象存储服务



文件存储服务

平台基于 UCloud 公有云基础架构，复用内核及核心虚拟化组件，将公有云架构私有化部署，具有自主可控、稳定可靠、持续进化及开放兼容等特点，企业可通过控制台或 APIs 快速构建资源及业务，支持与公有云无缝打通，灵活调用公有云能力，帮助企业快速构建安全可靠的业务架构。

UCloudStack 定位为轻量级交付，3 节点即可构建生产环境且可平滑扩容，不强行绑定硬件及品牌，兼容 X86 和 ARM 架构，并提供统一资源调度和管理，支持纯软件、超融合一体机及一体机柜多种交付模式，有效降低用户管理维护成本，为用户提供一套安全可靠且自主可控的云服务平台。

1.2 核心优势

- 超十年云计算技术沉淀

超 10 年公有云运营成熟经验和技術积累，从海量项目实践中积累了私有云

产品及解决方案能力。核心组件历经上万家企业级客户大规模的磨炼和验证，确保产品稳如磐石。

- **中立安全自主可控**

中立、安全的云计算服务平台，基于公有云架构，复用核心虚拟化组件自主研发，可控性高且可靠性经上万家企业验证。专注于 IaaS、PaaS 的全栈云产品自主研发与服务。

- **技术架构轻量灵活**

轻量灵活且先进的技术架构，最小 3 个节点即可构建生产环境，并可平滑扩容至数千、数万节点。资源规划尽在掌控，轻松应对业务增长挑战。

- **稳定可靠开放兼容**

平台服务高可用，虚拟资源智能调度，数据存储多副本，自愈型分布式网络，为业务保驾护航。不绑定硬件品牌，兼容 X86、ARM、龙芯、申威等架构及生态适配，设备异构搭建统一管理。

- **全方位服务保障**

1 对 1 金牌服务，精准匹配客户需求，量身打造全栈式私有化方案，以最佳实践协助企业轻松上云，为私有化项目提供全生命周期服务保障，助力企业数字化建设。

- **众多用户可信赖选择**

适配兼容 100+ 信创生态软硬件，全面满足企业上云、数字化改造等多场景业务需求。累计服务政府、金融、教育、制造等 14+ 行业、400+ 上市企业，成为全球 50000+ 用户及伙伴可信赖的数字转型服务商。

1.3 产品架构



UCloudStack 平台整体产品架构由基础硬件设施、虚拟核心引擎、智能调度系统、统一计算平台、分布式存储、分布式虚拟网络、PaaS 服务、统一运营服务、统一运维服务组成, 为平台管理者和租户人员提供统一云管理和自服务门户。

1.3.1 基础设施

用于承载 UCloudStack 平台的服务器、交换机及存储设备等。

- 平台支持并兼容通用 X86、ARM 及 MIPS 架构硬件服务器, 不限制服务器和硬件品牌;
- 支持 SSD、SATA、SAS 等磁盘存储, 同时支持计算存储超融合节点及对接磁盘阵列设备, 无厂商锁定;
- 支持华为、思科、H3C 等通用交换机、路由器网络设备接入, 所有网络功能均通过 SDN 软件定义, 仅需物理交换机支持 Vlan、Trunk、IPv6、端口聚合、堆叠等特性;
- 支持混合云接入并适配客户现有硬件资源, 充分利用资源的同时, 无缝对接现有资源服务。

1.3.2 虚拟核心引擎

承载平台核心的操作系统内核、虚拟化计算、存储、网络的实现和逻辑。

- **内核模块**: 承载云平台运行的服务器操作系统及内核模块, 复用公有云深度优化的 Linux 内核; 同时兼容 ARM 生态的 UOS、银河麒麟等服务器操作系统及内核。
- **虚拟化计算**: 通过 KVM、Libvirt 及 Qemu 实现计算虚拟化, 支持标准虚拟化架构, 提供虚拟机全生命周期管理, 兼容 X86 和 ARM 架构体系, 支持热升级、重装系统、CPU 超分、GPU 透传、在线迁移、宕机迁移、反亲和部署等特性, 并支持导入导出虚拟机镜像满足业务迁移需求。
- **分布式网络 SDN**: 基于主机 Overlay 提供软件定义 SDN 网络, 实现与硬件设备解绑, 支持分布式网络转发和 SDN 控制器高可用, 无需单独的网络节点和专有 SDN 网络设备; 同时可实现 VPC 网络、外网 IP、VIP、弹性网卡、安全组、组播及虚拟负载均衡、NAT 网关、IPSec VPN 及专线等 NFV 网络服务的开通及管理, 并支持 IPv4&IPv6 双栈。
- **分布式存储 SDS**: 采用分布式存储作为虚拟化后端存储池, 为平台提供多存储介质的高性能块存储服务。提供多副本、多级故障域、智能故障自恢复、磁盘 QoS 等数据保护机制及自动精简配置, 保障数据完整性。支持云盘在线扩容、克隆、快照及回滚功能, 同时支持缓存分层机制, 通过 SSD/NVMe 作为缓存提高存储性能。

1.3.3 统一资源调度

平台统一资源智能调度体系将平台计算、存储、网络资源统一池化, 为虚拟资源提供均衡部署、在线迁移、离线迁移、宕机迁移、亲和/反亲和、存储热迁移及网络流量分发等全面资源调度能力, 保证资源可用性和业务可靠性。

- **均衡部署**: 支持实时监测节点负载信息, 创建云资源时, 优先选择低负荷节点进行部署, 同时默认支持分散部署策略, 将云资源尽可能分散部署于集群内的空闲节点上, 保证集群资源均衡。

- **在线迁移**: 支持在线迁移能力, 管理员手动将一台虚拟机热迁移到另一台计算服务器节点或者随机迁移到任意满足条件的计算服务器节点上, 迁移高负载资源支持查看进度。
- **离线迁移**: 支持离线迁移能力, 关机状态下的虚拟机可迁移至其它计算集群, 启动时在新集群自动调度至最优节点。
- **宕机迁移**: 提供云资源故障转移机制, 当云资源所在物理机由于硬件故障出现宕机时, 支持将资源自动切换至集群内最适合的健康物理机上。
- **亲和/反亲和**: 支持隔离组进行虚拟机资源的自动调度机制, 可通过隔离组自定义虚拟机与其他虚拟机或宿主机之间的亲和关系。支持亲和性、反亲和性及强制亲和、强制反亲和多种调度类型, 保障业务的高可用。
- **存储热迁移**: 支持虚拟硬盘跨存储集群迁移。支持将虚拟机的系统盘和数据盘迁移至不同类型的存储集群和设备上, 支持分布式存储集群之间的热迁移、支持分布式存储和商业存储之间的热迁移。
- **网络流量调度**: 支持平台虚拟资源的网络流表控制及下发, 保证分布式网络架构的性能及可用性。

1.3.4 统一计算平台

统一虚拟化、容器、GPU 计算层, 共享底层物理服务器资源, 直接采用物理机运行并提供虚拟化、安全容器、GPU 及裸金属计算服务。异构资源统一管理平滑扩展, 并可通过智能调度和资源超分进一步提升资源利用率和可用性, 为云平台提供安全、可靠、稳定的基础计算环境。

- **虚拟机**: 运行在物理主机上的虚拟机, 提供从镜像或磁盘创建能力, 支持重启/关机/断电/启动/暂存、删除、VNC 登陆、重装系统、重置密码、热升级、ISO 镜像加载、USB 透传、绑定外网 IP 及安全组、挂载数据盘及反亲和策略部署等虚拟机全生命周期功能。支持用户自定义虚拟机的引导方式、Cloud-init、CPU 启动模式、DNS、MAC 地址, 同时支持用户对虚拟机磁盘进行快照备份管理。

- **GPU 计算**: 平台提供 GPU 设备透传能力, 为用户提供 GPU 虚拟机能力。支持用户在平台上创建并运行 GPU 虚拟机, 让虚拟机拥有高性能计算和图形处理能力。
- **虚拟机镜像**: 虚拟机运行时所需的操作系统镜像, 支持 CentOS、Windows、Ubuntu 等常用基础操作系统。提供基础镜像和自制镜像管理能力, 支持将虚拟机导出为镜像, 通过镜像重建虚拟机; 同时支持用户自定义导入 QCOW2、ISO 格式的虚拟机镜像。
- **安全容器**: 提供安全容器计算引擎服务, 采用 MicroVM 轻量虚拟机作为容器沙箱环境, 提供更高的隔离级别, 进一步提升容器安全。
- **裸金属**: 裸金属服务为用户提供物理服务器统一管理能力, 用户在控制台即可对物理服务器进行无缝纳管, 同时可将其转化为裸金属资源成为物理云资源。支持进行电源管理、访问远程控制台和查看硬件监控、重新部署操作系统等物理服务器的生命周期管理。
- **弹性伸缩**: 支持弹性伸缩功能, 用户可通过定义弹性伸缩策略, 在业务需求增长时自动增加计算资源 (虚拟机) 以保证计算能力; 在业务需求下降时自动减少计算资源以节省成本。基于负载均衡和健康检查机制, 可同时适用于请求量波动和业务量稳定的业务场景。

1.3.5 统一存储服务

基于分布式存储构建统一存储层, 为应用提供块、对象及文件存储服务, 兼容多存储协议的同时, 可结合多副本、快照、QoS 及加密等方式保证资源及数据的安全性, 采用 SSD 缓存加速降低成本同时提升存储性能, 同时支持商业存储作为后端存储, 使用户可快速构建领先的存储解决方案。

- **块存储**: 一种基于分布式存储系统为虚拟机提供持久化存储空间的块设备。具有独立的生命周期, 支持随意绑定/解绑至多个虚拟机使用, 基于网络分布式访问, 并支持容量扩容、克隆、快照、QoS、加密等特性, 为虚拟资源提供高安全、高可靠、高性能及可扩展的磁盘。

- **外置存储**: 平台通过 iSCSI 协议, FC SAN 协议对接商业存储, 将商业存储作为虚拟化后端存储池, 提供存储池管理及逻辑卷分配, 可直接作为虚拟机的系统盘及数据盘进行使用, 丰富平台存储解决方案的同时, 利旧商业存储降低成本。
- **共享盘**: 共享云硬盘是一种支持多个云服务器并发读写访问的数据块级存储设备, 支持将云硬盘、SAN 存储 LUN 设备设备为共享盘, 并作为虚拟机的数据盘, 使多个虚拟机同时对共享盘进行数据读写操作。
- **快照服务**: 提供磁盘快照及快照回滚能力, 可应用于容灾备份及版本回退等业务场景, 降低因误操作、版本升级等导致的数据丢失风险;
- **文件存储**: 提供基于 NFS 协议的文件存储能力, 支持基于云租户隔离的文件存储实例, 提供 NFSv4 协议共享能力, 支持文件存储实例的创建、扩容及快照备份管理, 同时支持目录文件的生命周期管理。
- **对象存储**: 对象存储服务兼容标准 S3 访问协议, 支持基于云租户隔离的对象存储。提供私有、公共读、公共读写三种类型的存储桶, 支持存储桶、对象文件、访问令牌等管理功能, 并提供多版本、对象锁定、生命周期及快照备份能力。

1.3.6 分布式虚拟网络

纯软件定义分布式虚拟网络, 支持用户自定义构建安全隔离的 VPC 网络, 提供 IPv4/IPv6 双栈单播和组播通信, 并结合安全组和 QoS 保证网络流量的安全性。同时基于 NFV 能力, 提供外网 IP、高可用 VIP、弹性网卡、负载均衡、NAT 网关、VPN 网关、专线接入等服务全面覆盖网络接入场景, 并结合 DPDK 实现云资源网络性能的全面增强。

- **VPC 网络**: 软件定义虚拟私有网络 (VPC—Virtual Private Cloud), 一个属于用户的、逻辑隔离的二层网络广播域环境。在一个私有网络内, 用户可以构建并管理多个三层网络, 即子网 (Subnet), 包括 IP 网段、IP 地址、网关等虚拟资源作为租户虚拟机业务的网络通信载体。

- **外网 IP**: 外网弹性 IP 服务, 提供直通和 NAT 两种通信模式, 为虚拟机、负载均衡、NAT 网关、VPN 网关及 PaaS 等资源提供外网 IP 地址, 使虚拟资源可与平台以外的网络进行高性能连接, 并提供 IPv4/IPv6 双栈网络连接服务。
- **高可用 VIP**: 提供高可用虚拟 VIP 服务, 归属于 VPC 子网或外网 IP 网段内可漂移 IP 地址, 用户可将 VIP 与高可用服务结合, 以便在服务出现故障时进行服务入口的漂移, 以实现服务的高可用。
- **安全组**: 虚拟防火墙, 提供出入双方向流量访问控制规则, 定义哪些网络或协议能访问资源, 用于限制虚拟资源的网络访问流量, 支持 IPv4/IPv6 双栈能力的 TCP、UDP、ICMPv4、ICMPv6 及多种应用协议, 为云平台提供必要的安全保障。
- **弹性网卡**: 弹性网卡 (Elastic Network Interface, ENI) 是一种可随时附加到虚拟机的弹性网络接口, 支持绑定和解绑, 可在多个虚拟机间灵活迁移, 为虚拟机提供高可用集群搭建能力, 同时可实现精细化网络管理及廉价故障转移方案。
- **NAT 网关**: 企业级 NAT 网关服务, 为云平台资源提供 SNAT 和 DNAT 能力。支持 VPC 级、子网级及虚拟机实例级的 SNAT 规则, 使不同维度的资源通过 NAT 网关访问外网, 同时支持 TCP、UDP 多协议的 DNAT 端口映射服务。
- **负载均衡**: 软件定义负载均衡服务, 提供 4 层和 7 层协议转发能力, 基于 TCP、UDP、HTTP、HTTPS 网络监听协议, 通过负载均衡算法、健康检查、会话保持等均衡能力, 将访问流量自动分配至多台虚拟机, 用于实现流量负载及高可用。支持 HTTP 内容转发、HTTP 重定向 HTTPS、获取客户端真实 IP、监听器协议、附加 HTTP Host 字段等服务, 并提供 SSL Offloading 及 SNI 多域名证书能力。
- **IPSecVPN**: 提供 IPSecVPN 网关服务, 通过 IPSec 协议加密的隧道技术, 将 UCloudStack 与 UCloud 公有云、IDC 数据中心、第三方公有云

的内网打通，在互联网上为两个私有网络提供安全通道，通过加密保证连接的安全。

- **VPC 互通**：提供 VPC 网络互通能力，支持同租户不同 VPC 之间的网络互通管理。
- **专线接入**：提供专线接入的管理和配置，为用户本地数据中心与私有云 VPC 之间建立高速、低时延、稳定安全的专属连接通道。
- **组播服务**：提供 VPC 网络的组播通信服务，实现 VPC 网络中点到多点的高效数据传送能力，支持组播组 IP、组播组端口、发送方、接收方的管理和配置。
- **DPDK 能力**：支持高性能网络特性，通过引入 OVS DPDK 能力，增强并提高云资源网络性能。

1.3.7 PaaS 服务

平台基于稳定可靠、弹性伸缩的计算、存储、网络等 IaaS 资源为用户提供高效部署便捷管理的 MySQL 数据库、Redis 缓存及 Kubernetes 容器服务。

- **MySQL 数据库**：基于 MySQL 提供的关系型数据库 PaaS 服务，支持单机版和主备版两种架构，并提供数据库的一键部署、从库管理、参数配置、平滑扩展、监报告警、日志事件及备份恢复等自动化高效部署及运维管理能力，支持 MySQL5.7 和 MySQL8.0 版本。
- **Redis 缓存**：基于 Redis 提供的缓存 PaaS 服务，支持单机版和主备版两种架构，并提供缓存实例的一键部署、从库管理、参数配置、平滑扩展、监报告警、日志事件及备份恢复等自动化高效部署及运维管理能力，支持 Redis4.0 和 Redis7.0 版本。
- **Kubernetes 容器服务**：基于 Kubernetes 提供容器集群管理服务，为用户提供开箱即用、全托管免运维的生产级容器集群，采用轻量级虚拟化为容器提供强隔离运行环境，与虚拟机共用物理节点、VPC 网络、负载均衡及块存储等基础云资源。完全兼容原生 Kubernetes API，并为容

器化应用提供高效部署、资源调度、服务发现及监控日志等完整功能。

1.3.8 统一运营服务

提供对云服务的统一运营能力，包括多租户、权限管理、OAuth 认证、云服务目录、计量计费、配额管理、规格配置、报表统计、资源模板、资源标签、回收站、审批流程及自定义 UI 等运营管理功能，并可通过管理控制台进行简易运营操作。

- **多租户**：平台支持多租户模式，提供租户间资源隔离能力，不同租户之间资源完全隔离互不影响。支持租户生命周期管理和租户自服务模式，并支持租户的充值、提现、价格、配额等配置管理。
- **多级账号体系**：支持管理员、主账号、子账号多级账号管理体系，管理员为平台管理账号，主账号为租户管理员，子账号为租户子人员，其权限由项目授权范围和角色权限决定。不同的账号登录平台可管理的功能不同，所有登录账号均可进行账号的信息查看及账号安全配置。
- **资源级权限**：支持对租户下的子账号进行细粒度权限管理（如三权分立配置），将云资源以项目组进行分类，以项目组和自定义角色划分权限范围，并将权限授权给子账号，对子账号进行资源授权管理，实现平台资源级权限控制。
- **OAuth 认证**：平台账号支持对接 OAuth 2.0，并可使用 OAuth 第三方平台帐号登录平台，同时提供 OAuth2.0 账户认证接口，使第三方平台直接通过 OAuth 对接平台的账户系统。
- **云服务目录**：提供服务目录管理能力，统一展示平台支持的所有云服务的整体授权和启用情况，可统一管控全局和租户在各地域的服务开通。
- **计量计费**：提供计量计费能力，支持按时、按年、按月三种计费方式，并支持资源的计费、扣费、自动续费、手动续费、退费及过期回收等计费策略。提供产品全局定价能力，支持针对特定租户设置价格折扣，并可基于账户提供充值、提现、订单、交易及账单查询及管理。

- **配额管理**: 平台提供全局配额管理的能力, 支持用户自定义各地域的产品服务数量及容量默认配额。支持为租户自定义分配资源配额, 限制租户的资源使用量。
- **规格配置**: 提供自定义规格能力, 支持用户自定义各地域、各集群的虚拟机、虚拟硬盘、外网 IP、MySQL、Redis、对象存储及文件存储的等产品服务的规格。
- **报表统计**: 提供报表统计能力, 支持资源用量统计及资源统计两种报表, 将各种数据整理成易于理解和分析形式, 提高整体运营和管理的效率。
- **资源模板**: 提供虚拟机资源模板能力, 支持用户自定义虚拟机规格及参数配置为模板, 通过资源模板快速创建虚拟机, 并可结合水平弹性伸缩完成业务节点的快速伸缩。
- **资源标签**: 提供资源标签能力, 支持用户对云资源添加一个或多个标签, 并通过标签分类、筛选、搜索云资源。
- **回收站**: 提供回收站功能, 支持虚拟机、虚拟硬盘、外网 IP 及自制镜像资源的回收, 当资源被删除或过期未续费时均支持自动进入回收站。提供资源回收全局策略配置, 并支持对回收站中资源进行手动续费、资源恢复及销毁操作。
- **审批流程**: 平台提供审批流程功能, 管理员和租户可自定义创建审批流程, 为租户或子账号的指定云服务启动审批流程。提供审批管理能力, 支持手动审批、自动审批及代理审批, 并支持审批通过后的自动下发资源操作能力。
- **自定义 UI**: 提供平台 UI 风格自定义能力, 支持用户自定义平台登录页、平台网站及监控大屏的 UI 配置, 如 Logo、Favicon、背景图、标题文字等, 同时支持用户自定义平台默认语言环境和默认币种配置。

1.3.9 统一运维服务

提供对虚拟资源和物理资源的统一运维能力, 包括多数据中心管理、集群、

节点、镜像、外网网络及云服务等物理和虚拟资源的管理及监报告警、操作日志、资源事件、一键巡检、平台配置、部署升级及迁移服务等运维管理功能，对云平台及云产品的运维进行统一管理。

- **多数据中心管理：**支持多数据中心统一运维、统一运营、统一计费的管理能力，可通过一套管理平台管理多数据中心云平台，用户可在统一控制台使用多个数据中心的资源。
- **物理资源管理：**提供统一物理设备管理能力，包括地域、计算集群、存储集群、物理节点、外网网络、裸金属、外置存储等物理资源，支持地域、集群、节点的 CPU、内存、存储容量的资源用量统计和用量预警告警通知。
 - **集群管理：**提供多计算集群和存储集群管理，一个数据中心可以部署多个集群。提供自定义集群类型能力，支持获取计算集群中计算实例信息，并支持获取存储集群的冗余策略和磁盘架构等信息。同时支持集群权限控制，将集群资源独享给一个或部分租户使用。
 - **节点管理：**提供物理宿主机节点管理和维护能力，支持对节点进行锁定、解锁、进入维护模式、退出维护模式等管理，并支持对磁盘进行点灯和关灯操作。
 - **外网网络：**提供平台对外通信的网络网段管理能力，支持管理员将物理网络的空闲地址分配至云平台作为 EIP 资源池，不同的外网网段支持分配给指定的云租户，支持 IPv4 和 IPv6 两种类型外网网段管理，并支持外网网段路由配置。
 - **裸金属管理：**提供物理服务器的生命周期管理，包含添加物理机、批量导入物理机、远程开机、远程关机、远程强制关机、远程重启、转换为裸金属、远程一键重装操作系统、更新物理服务器参数信息。
 - **外置存储管理：**提供 iSCSI 和 FC 集中式存储接入能力，支持集中式存储 LUN 设备扫描发现功能，并可对发现的 LUN 设备进行分配和绑定管理。

- **云资源管理**: 平台提供云资源统一管理服务能力, 支持管理者统一管理平台所有用户创建的计算、存储、网络、数据库、缓存、对象存储、文件存储、MySQL、Redis、回收站、监控告警、资源模板、资源标签等云服务资源, 并支持管理者为指定租户创建并管理云服务资源。
- **监控告警**: 平台具备全线产品的运维监控及告警能力, 为用户提供实时监控图表和历史监控图表。支持对云平台计算、存储、网络、数据库、云缓存、容器、对象存储、文件存储等产品及平台地域、集群、节点等资源的监控指标自定义设置告警规则, 并在资源故障或监控指标超过阈值时, 以邮件和 Webhook 的方式进行告警通知, 同时可支持用户获取并查看历史告警记录信息。
- **操作日志**: 提供平台全面操作审计日志, 支持收集并展示用户在控制台或 API 对资源的操作行为日志及登录登出平台的审计日志。支持云产品服务模块的操作日志收集和展示, 并支持对操作日志事件进行监控, 并进行告警通知。
- **资源事件**: 提供资源事件日志能力, 对于云平台核心资源的部分操作进行记录及通知, 如资源生命周期状态的变化, 操作运维执行情况等。支持对资源事件日志进行监控并进行告警通知。
- **一键巡检**: 支持对管理节点、计算节点以及平台自身的服务进行多维度的健康扫描检测, 检查平台节点 CPU、内存、磁盘等资源的使用情况, 便于管理员了解平台健康状况及问题评估。支持用户通过控制台一键创建自动巡检任务, 并支持用户查看并下载一键巡检报告。
- **平台配置**: 提供平台运维管理配置, 支持对账号登录安全、回收策略、告警通知邮箱、产品全局策略、平台自身备份等能力进行细化配置, 使平台策略更加符合管理者预期。
- **部署升级**: 平台软件及组件具备自动化部署能力, 并可提供 PXE 网络分发模式, 提升批量部署的效率。同时平台自身具备在线平滑升级力, 升级过程中无需手工干预, 并保证不对平台业务和应用造成影响, 可正

常不间断的对外提供服务。

- **迁移服务**：提供异构环境至云平台的整机在线迁移能力，支持无代理和有代理在线迁移，助力业务高效上云。

1.3.10 统一管理平台

平台为用户提供 Web 控制台和 API 接口两种方式接入和管理云平台。Web 控制台为平台管理者和租户分别提供管理控制台和租户控制台，并提供监控大屏服务；开发者可通过 APIs 自定义构建云平台资源，支持无缝迁移上云。

- **租户自服务门户**：提供租户自服务门户，源于超 1 年公有云使用体验。租户和子账号可通过控制台完成虚拟机、容器、虚拟硬盘、VPC 等 IaaS 和 PaaS 资源的申请、使用、修改和销毁等全生命周期的自服务操作，无需关心底层资源，即刻获得优质云资源。
- **管理自服务门户**：管理员可通过管理控制台进行平台所有资源的统一管理和运维，平台资源全局掌控，租户资源全面接管。并支持多租户、计量计费、报表统计、流程审批、监报告警、日志事件等管理；同时控制台可提供中英文站点、站内信、收藏夹等能力，并支持展示产品版本信息和在线帮助文档。
- **监控大屏**：提供平台大屏监控，可使平台运营者清晰了解平台整体软硬件运行情况下和状态，支持查看平台整体通知告警、资源分配、资源使用量、宿主机 TOP5、虚拟机 TOP5，可自定义监控大屏的标题，并支持通过数据中心切换不同地或的监控大屏信息。
- **全局概览**：提供基于管理视角的关键信息概览的柱状及饼状图，包括全地域资源用量统计、地域资源用量统计 Top 5 排行、云服务资源数量使用排行榜、虚拟机 CPU 总数量、虚拟机内存总数量、云盘磁盘空间信息等。
- **租户概览**：提供基于租户视角的关键信息概览展示页，包括告警记录、虚拟机总量信息、虚拟硬盘总量信息、外网 I 总量信息、系统信息、虚

拟机 TOP 排名、资源配额使用情况。

- **开放标准 API:** 云平台可对外开放完善易用的 API 接口及文档，允许第三方平台对接；可提供 Python、Golang、Java 等多种语言的 SDK 开发工具，便于第三方接入和使用。同时支持 API 控制台，通过界面提供对云平台 API 的测试调用，并提供参数的解释和文档说明。

1.4 技术架构特性

1.4.1 API 幂等性

幂等性是指一次和多次请求某一个资源应该具有同样的作用，保证资源请求无论调用多少次得到的结果始终一致。如多次调用更新虚拟机的 API 请求，返回的结果都是一致的。

UCloudStack 通过分布式锁、业务字段唯一约束及 Token 唯一约束等技术手段保证平台资源 API 幂等性。对虚拟机、云硬盘、VPC、负载均衡等资源的操作请求（除创建请求）均支持重复提交，并保证多次调用同一个 API 请求返回结果的一致性，同时避免网络中断导致 API 未能获取确切结果，从而导致重复操作的问题。

1.4.2 全异步架构

云平台使用消息总线进行服务通信连接，在调用服务 API 时，源服务发消息给目的服务，并注册一个回调函数，然后立即返回；一旦目的服务完成任务，即触发回调函数回复任务结果。

云平台服务之间和服务内部均采用异步调用方法，通过异步消息进行通信，并结合异步 HTTP 调用机制，保证平台所有组件均实现异步操作。

基于异步架构机制，云平台可同时管理数十万以上的虚拟机及虚拟资源，后端系统每秒可并发处理上万条 API 请求。

UCloudStack 采用的插件机制，每个插件设置相应的代理程序，同时在 HTTP 包头为每个请求设置回调 URL，插件任务结束后，代理程序发送应答给调

用者的 URL。

1.4.3 分布式

(1) 分布式底层系统

UCloudStack 核心模块提供计算、存储及调度等分布式底层支持，用于智能调度、资源管理、安全管理、集群部署及集群监控等功能模块。

- **智能调度**

基于分布式服务调用和远程服务调用为租户提供智能调度模块。智能调度模块实时监测集群和所有服务节点的状态和负载，当某集群扩容、服务器故障、网络故障及配置发生变更时，智能调度模块将自动迁移被变更集群的虚拟资源到健康的服务器节点，保证云平台的高可靠性和高可用性；

- **资源管理**

通过分布式资源管理模块，负责集群计算、存储、网络等资源的分配及管理，为云平台租户提供资源配额、资源申请、资源调度、资源占用及访问控制，提升整个集群的资源利用率；

- **安全管理**

分布式底层系统提供安全管理模块，为租户提供身份认证、授权机制、访问控制等功能。通过 API 密钥对和用户名密码等多种方式进行服务间调用及用户身份认证；通过角色权限机制进行用户对资源访问的控制；通过 VPC 隔离机制和安全组对资源网络进行访问控制，保证平台的安全性；

- **集群部署**

分布式底层系统为云平台提供自动化部署集群节点的模块，为运维人员提供集群部署、配置管理、集群管理、集群扩容、在线迁移及服务节点下线等功能，为平台管理者提供自动化部署通道；

- **集群监控**

监控模块主要负责平台物理资源和虚拟资源信息收集、监控及告警。监控模

块在物理机及虚拟资源上部署 Agent，获取资源的运行状态信息，并将信息指标化展示给用户；同时监控模块提供监报告警规则，通过配置告警规则，对集群的状态事件进行监控及报警，并有效存储监控报警历史记录。

(2) 分布式存储系统

平台可采用高可靠、高安全、高扩展、高性能的分布式存储系统，提供块存储服务，保证本地数据的安全性和可靠性。

- 软件定义分布式存储，将大量通用机器的磁盘存储资源聚合在一起，采用通用的存储系统标准，对数据中心的所有存储进行统一管理。
- 分布式存储系统采用多副本数据备份机制，写入数据时先向主副本写入数据，由主副本负责向其他副本同步数据，并将每一份数据的副本跨磁盘、跨服务器、跨机柜、跨数据中心分别存储于不同磁盘上，多维度保证数据安全。
- 多副本机制存储数据，将自动屏蔽软硬件故障，磁盘损坏和软件故障，系统自动检测到并自动进行副本数据备份和迁移，保证数据安全性，不会影响业务数据存储和使用。
- 分布式存储服务支持水平扩展、增量扩容及数据自动平衡性，保证存储系统的高扩展性。
- 支持 PB 级存储容量，总文件数量可支持亿量级。
- 支持不间断数据存储和访问服务，保证存储系统的高可用性。
- 支持高性能云硬盘，IOPS 和吞吐量随存储容量规模线性增长，保证响应时延。

在部署上，计算节点自带 SSD 磁盘构建为高性能的存储池，计算节点自带的 SATA/SAS 磁盘构建为普通性能存储池。分布式存储系统将块设备内建为弹性块存储，可供虚拟机直接挂载使用，在数据写入时通过三副本、写入确认机制及副本分布策略等措施，最大限度保障数据安全性和可用性。在本地可通过快照技术，将本地数据定时备份，在数据丢失或损坏时，可通过快照快速恢复本地业

务的数据。

(3) 分布式网络架构

采用分布式 Overlay 网络，提供 VPC、NAT 网关、负载均衡、安全组、外网 IP 等网络功能。

- 云平台 Overlay 网络分布式运行在所有计算节点。
 - 管理服务仅作为管理角色，不承担网络组件部署及生产网络传输；
 - 虚拟网络流表分发服务为高可用架构，仅做流表分发不透传生产网络传输；
 - 所有生产网络仅在计算节点上传输，无需通过管理服务或流表分发服务进行转发；
 - 管理服务和流表分发服务故障，不影响已部署好的虚拟资源运行及通信。
- 超融合计算节点或独立存储节点根据磁盘和业务分不同的集群 (Set) 。
 - 每个物理集群 45 台节点，控制集群规模。
 - 业务数据网络仅在单集群中进行传输，即在单组接入交换机中进行传输。
- 分布式存储直接通过物理网络进行挂载，无需通过 overlay 网络进行挂载和传输。
 - 通过 libvirt 融合分布式存储 rbd 和 qemu，qemu 通过 librbd 操作分布式存储；
 - 虚拟化进程与分布式存储进程通过本机&跨物理机内网进行通信；
 - 云平台内网至少使用万兆交换机并做端口聚合，可满足虚拟机和分布式存储的性能需求；

分布式网络架构将业务数据传输分散至各个计算节点，除业务逻辑等北向流

量需要管理服务外，所有虚拟化资源的业务实现等南向流量均分布在计算节点或存储节点上，即平台业务扩展并不受管理节点数量限制。

1.4.4 高可用

UCloudStack 私有云平台架构，从硬件设施、网络设备、服务器节点、虚拟化组件、分布式存储均提供高可用技术方案，保证整个云平台业务不间断运行。

- 数据中心机柜级别冗余性设计，所有设备均对称部署于机柜，单机柜掉电或故障不影响业务；
- 网络服务区域隔离设计，内网业务和外网业务在物理设备上完全隔离，避免内外网业务相互影响；
- 网络设备扩展性设计，所有网络设备分为核心和接入两层架构，一套核心可水平扩展几十套接入设备；
- 网络设备冗余性设计，所有网络设备均为一组两台堆叠，避免交换机单点故障；
- 交换机下联接入冗余性设计，所有服务器双上联交换机的接口均做 LACP 端口聚合，避免单点故障；
- 服务器网络接入冗余性设计，所有服务器节点均做双网卡绑定，分别接入内网和外网，避免单点故障；
- 管理节点冗余性和扩展性设计，多台管理节点均为 HA 部署，并支持横向扩展，避免管理节点单点故障；
- 通过智能调度系统将虚拟机均衡部署于计算节点，可水平扩展计算节点数量；
- 分布式存储冗余性设计，将数据均衡存储于所有磁盘，并多副本、写确认机制及副本分布策略保证数据安全；
- 进行服务器节点及存储扩展时，只需增加相应数量的硬件设备，并相应的配置资源调度管理系统；

- 云平台内各组件均采用高可用架构设计，如管理服务、调度服务、网络流表分发服务等，保证平台高可用；
- 云平台提供的产品服务，如负载均衡、NAT 网关、数据库服务及缓存服务均采用高可用架构构建，保证云平台提供服务的可靠性。

1.4.5 业务实现分离

云平台架构从业务逻辑上分为北向接口和南向接口，将云平台的业务逻辑和业务实现进行分离，业务管理逻辑不可用时，不影响虚拟资源的正常运行，整体提升云平台业务可用性和可靠性。

- 北向接口：仅定义业务逻辑，提供业务接口，负责北向数据落地。业务接口包括帐户认证、资源调度、监控、计费、API 网关及 WEB 控制台等业务服务接口。
- 南向接口：仅定义业务实现，负责将北向接口的业务转换为实现，如虚拟机运行、VPC 网络构建、分布式存储数据存储等。

业务实现分离后，当云平台业务端（如 WEB 控制台）发生故障时，并不影响已运行在云平台上的虚拟机及运行在虚拟机中的业务，是实现业务系统的连续性保障机制之一。

1.4.6 组件化

云平台所有虚拟资源组件化，支持热插拔、编排组合及横向扩展。

- 组件化包括虚拟机、磁盘、网卡、IP、路由器、交换机、安全组等。
- 每种组件均支持热插拔，如将一个 IP 绑定至一个在运行中的虚拟机。
- 每种组件均支持横向扩展，如横向增加虚拟机的磁盘，提升整体云平台的健壮性。

1.5 客户痛点

1.5.1 自建私有云的痛点

- **可控性差**: 业界基于开源架构封装的私有云核心组件和服务源自社区, 可控性差且可靠性未经验证, 平台特性升级受限于社区且需专精运维人员, 同时开源框架构造繁杂, 部署实施环节复杂, 实施难度大。
- **投入成本大**: OEM 公有云直接部署的专有云平台, 所有服务均需独立的服务器集群, 起始部署规模较大且通常限制硬件架构及品牌, 部署实施需要投入大量基础设施和人力资源; 在运维方面通常需要托管运维, 建设成本较高。
- **运维复杂**: 自建数据中心及通用虚拟化系统, 对于业务构建所需的数据库、缓存、负载均衡等一系列应用, 需自己通过虚拟机进行搭建并维护, 同时还需考虑服务的集群部署、监控、日志、备份、容灾及可靠性和可用性等。
- **兼容性差**: 通用数据中心及虚拟化系统, 对国产化硬件、操作系统、中间件的适配及兼容性较差。

1.5.2 解决之道

- **可控可靠**: 采用非开源架构, 基于公有云自主研发, 复用内核及核心虚拟化组件, 将公有云部署规模重构为可运营、可运维、可快速交付且可私有化交付的云平台, 可控性高且可靠性经上万家企业验证。
- **轻量构建**: 平台所有产品服务使用统一底层资源池, 所有产品无需准备专用服务器集群, 3 节点即可轻量构建生产环境, 规模轻量且可水平扩展, 支持自动化一键部署并提供平滑升级能力, 一个运维人员即可轻松管理。
- **丰富组件**: 公有云一致用户体验的自服务平台, 除基础 IaaS 产品外, 为政企用户提供高可用、高可靠且可自服务的负载均衡、NAT 网关、

IPSecVPN、数据库服务及缓存服务等 PaaS 类产品服务。

- **兼容性高**：平台本身不绑定硬件架构及品牌，兼容 X86、ARM、MIPS 等主流架构，可异构搭建并进行统一资源管理；同时已适配信创体系硬件、操作系统及中件间，如华为泰山加鲲鹏、飞腾加银河麒麟、UOS 及南大通用数据库等。

1.6 应用场景

1.6.1 虚拟化&云化

通过将业务系统和内部应用部署至 UCloudStack 平台，可为用户提供一套集虚拟化、分布式存储、SDN 网络为一体的私有云平台。平台支持多数据中心管理，可将业务部署至多个数据中心构建灾备云或边缘计算，同时支持与公有云无缝打通，灵活调用公有云能力，帮助政企快速构建安全可靠的业务架构。

1.6.2 业务快速交付

平台服务所见即所得，可通过自服务云管理平台一键部署并管理业务交付所需的基础设施和中件间，包括在线扩容、负载分发、数据库缓存及监控日志等应用基础环境服务能力；同时平台支持镜像导入导出，可方便快捷将业务系统迁移至云平台，并可对所有业务系统的资源进行统一管理。

1.6.3 超融合一体机

平台提供一体机交付模式，多款机型应用不同业务场景，集成 UCloudStack 私有云平台，出厂预装开箱即用，服务模块热插拔可按需部署，提供虚拟化、网络、存储、数据库、缓存及云管等一系列云服务能力；同时可通过与 IDC 数据中心互联，构建混合云解决方案。

1.6.4 政企专有云

UCloudStack 提供租户控制台和管理员控制台，支持多租户、账户注册、计量计费等功能特性，同时为云平台管理者提供运营运维管理功能，包括资源管理、

租户管理、价格配置、资源规格配置、部署升级及监控日志等服务，为政企提供行业专有云解决方案。



UCloudStack 轻量级私有云属于 IaaS+PaaS 复合型产品，并可按需搭载大数据、安全屋、AI 等公有云产品，适用于全行业客户需要云化且私有部署的业务应用上云场景，典型行业如下：

- **政府、央企、军工、交通、制造型企业**

对外承担公共服务职责，内外部业务应用系统和商用软件需要快速交付、资源共享、智能调度及统一管理为上云需求的行业客户。

- **泛互联网行业，如 B2B 电商、大数据、教育等企业**

需要构建行业专属云，结合自有 SaaS 业务为其用户提供整体解决方案的行业客户。

- **人工智能和科研实验室行业**

需要大量可快速交付且私有化部署的虚拟化环境，用于科研项目和训练系统的快速部署和管理的行业客户。

1.7 交付和服务

UCloudStack 定位为轻量级交付，3 节点即可构建生产环境且可平滑扩容，并提供统一资源调度和管理，支持纯软件、超融合一体机及超融合机柜多种交付模式，有效降低用户管理维护成本，为用户提供一套安全可靠且自主可控的云服务平台。



- 纯软件交付

客户提供承载云平台运行的硬件服务器、网络设备及相关基础设施，UCloud 优刻得提供 UCloudStack 轻量级私有云软件；通常在基础网络设施环境完备的情况下，UCloudStack 软件可在 2 小时内完成部署并交付。

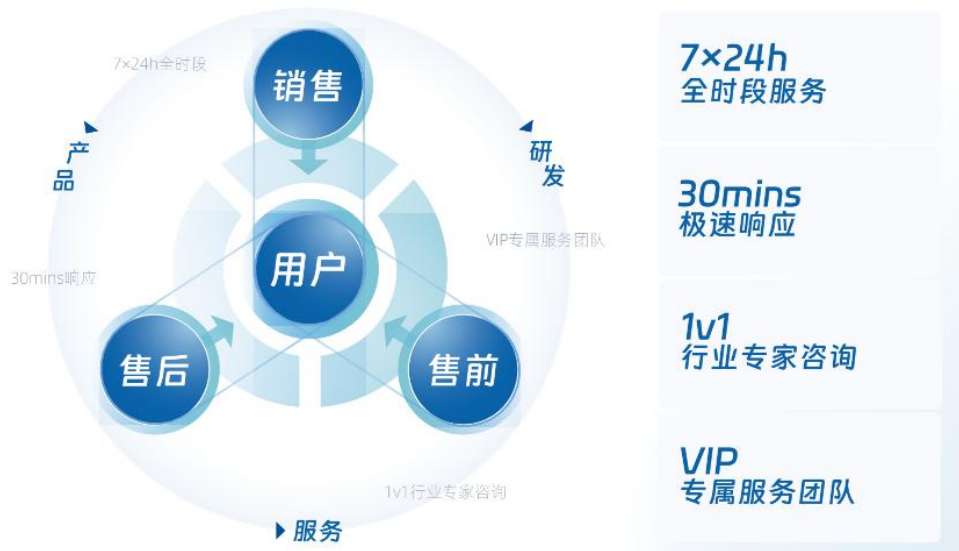
- 超融合一体机

客户仅需提供数据中心基础设施，UCloud 优刻得提供超融合一体机（出厂预装 UCloudStack），通常在基础网络设施环境完备的情况下，可在小时内完成初始化并交付。



- 超融合机柜

客户仅需提供数据中心即可，UCloud 优刻得提供超融合一体机柜（包含网络设备、服务器节点&一体机、PDU、线缆及 UCloudStack 软件），通常以一个机柜的形式进行交付。



在服务方面，提供全面服务保障体系，7x24 小时全时段服务，30 秒极速响应，1v1 行业专家咨询，VIP 专属服务团队。用户可根据业务场景和需求，自主选择适合的服务，如基础质保服务、高级质保服务、金牌质保服务、续保服务、培训服务及驻场服务；并可提供一对一专家的方案咨询、架构设计、迁移上云、巡检调优等。

2 平台物理架构

2.1 物理集群节点

云平台系统常见集群节点角色有 4 种, 分别是管理节点、计算存储融合节点、独立计算节点以及独立存储节点。

2.1.1 管理节点

集群内部署的核心管理服务, 承载私有云平台的北向接口服务模块, 包括帐户认证、计量计费、资源管理、网关及服务监控等服务, 提供标准 API 和 WEB 控制台两种接入和管理方式。

- 管理节点负责虚拟资源全生命周期的管理, 由于北向业务与南向实现接口分离及分布式网络机制, 网络流量通过所在计算节点直接转发, 平台业务扩展并不受管理节点数量限制;
- 云平台支持统一的底层资源, 基于管理服务仅转发和透传管理流量, 平台支持并推荐将管理服务部署于计算节点的虚拟机中, 通过平台虚拟机的智能调度提供管理能力的高可用。

管理服务与计算服务间通过 TCP/IP 协议进行通信, 提供管理服务通过内网或外网与计算节点通信的能力, 支持管理服务与计算服务分离部署, 如管理服务部署至公有云或其中一个数据中心, 计算节点分布在各个数据中心, 通过全局云管平台跨机房、跨数据中心及跨地域统一管理。

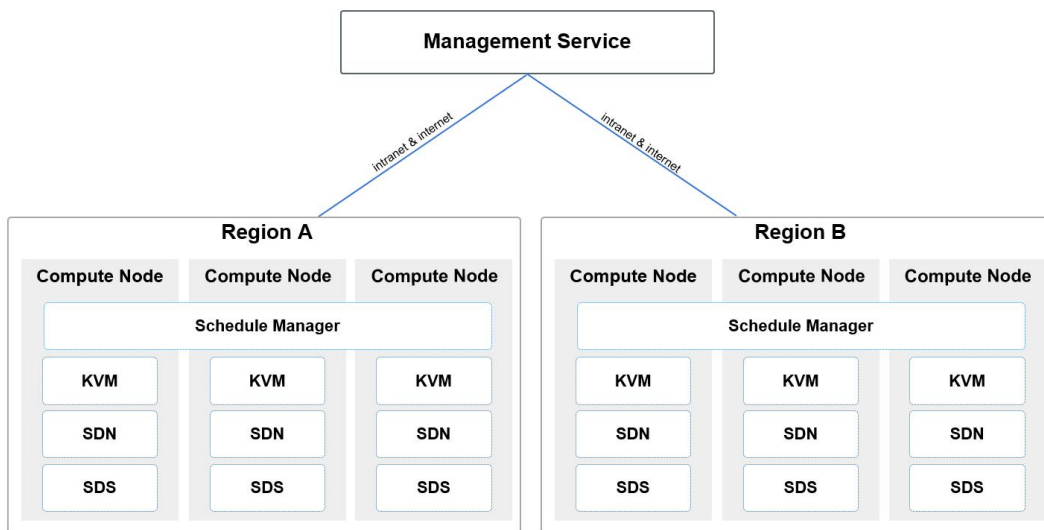
2.1.2 超融合节点

计算存储融合节点, 同时包含计算资源和存储资源, 用于运行虚拟机、虚拟网络、分布式存储、数据库服务、缓存服务等资源, 同时承载智能调度控制和监控服务。

云平台分布式存储使用所有计算节点的数据磁盘, 每个节点仅支持部署一种类型的数据磁盘, 如 SATA (使用 SSD 缓存加速)、SSD 等。

生产环境至少部署 3 台以上，保证分布式系统的正常部署和运行。采用 **SATA+SSD** 缓存的方案构建超融合节点，必须保证 **SSD** 缓存盘和 **HDD** 数据盘的容量比不高于 **1:20**，数量比不高于 **1:5**。

在部署上，每台计算节点均会部署用于运行计算存储网络的 KVM、Qemu、Libvirt、OVS、SDS 分布式存储等核心组件，同时在每个地域中至少有 3 台计算节点会部署核心调度及管理模块，如下图所示：



其中【Schedule Manager】即为 UCloudStack 云平台的核心调度及管理模块，用于虚拟资源的运行调度及虚拟网络的流表下发管理，每一个地域仅需部署一套高可用的 Schedule Manager。一般为主备模式，可在 3 台或多台计算节点上进行部署，当部署调度模块的主计算节点服务器物理故障时，部署调度模块的备计算节点将自动接替调度服务，保证核心调度及流表控制服务的可用性。

每个地域或数据中心的部署的 Schedule Manager 均会开放一个 API 端点，作为管理服务连接并管理数据中心计算资源的统一入口。API 端点支持通过内网和互联网的连接模式，在 TCP/IP 网络通信可达的情况下，管理服务（Management Service）支持部署于相同数据中心，也可部署于公有云或其它数据中心，并可为多数据中心计算资源提供统一调度和管理，满足云平台多应用场景部署。

2.1.3 独立计算节点

集群内宿主机节点，用于独立运行所有计算和网络资源，通过挂载独立存储节点的磁盘作为云平台的存储资源。

一般由几台到几千台服务器组成，生产环境至少部署 2 台以上，保证虚拟机的调度及稳定迁移。

通常建议将相同配置的计算节点服务器放置在一个集群内进行虚拟资源的调度。

2.1.4 独立存储节点

独立存储节点，用于独立承载分布式存储的节点，构建独立存储区域。适合将计算和存储分离，搭建独立存储网络的场景。独立存储节点，使用独立的存储网络接入设备，与计算业务物理或逻辑隔离。

- 部署独立存储节点，可节省计算节点的 CPU、内存等资源；
- 一般由几台和几千台服务器组成，生产环境至少部署 3 台以上，保证分布式系统的正常部署和运行；
- 独立存储节点为【可选】节点，如果采用融合节点，可使用计算存储超融合节点上的数据磁盘作为分布式存储的存储池。

部署存储节点时，每个节点需配置相同介质类型的数据磁盘，如全 SSD 存储节点、SATA+SSD 缓存存储节点等（适用于采用 SSD 缓存加速场景），将相同磁盘类型的节点组成一个存储集群，分别作为普通存储和高性能存储资源池。

注意 采用 SATA+SSD 缓存的方案构建独立存储节点，须保证 SSD 缓存盘和 HDD 数据盘的容量比不高于 1:20；盘数量比不高于 1:5。

2.1.5 商业存储节点

平台支持采用独立的商业存储（如 FCSAN、IP SAN 等）设备作为存储节点，构建存储区域，云平台的虚拟机镜像及云服务的数据均存储于商业存储设备，适

合已有商业存储设备的利旧场景，整体节省信息化转型的总拥有成本。

云平台基于 ISCSI、FC 协议对接商业存储，作为云平台的后端存储；并支持将存储设备中的 LUN 分配给租户，由租户将 LUN 分配或挂载至虚拟机的系统盘或数据盘，进行数据的读写。

商业存储设备采用设备已有的网络接入平台物理网络，与计算业务网络物理或逻辑隔离。

- 商业存储设备节点，可代替分布式存储节点，或与超融合、独立存储节点共同构建存储服务，构建不同的存储资源。
- 一般由几台至上百台商业存储设备组成，生产环境至少部署 2 台以上，保证存储系统的正常部署和运行；
- 商业存储节点为【可选】节点，可在生产环境中根据实际需要进行部署。

注意 平台仅将商业存储的 LUN 作为存储卷进行使用，不对存储卷本身进行管理，如 LUN 的创建、映射、扩容、快照、备份、回滚、克隆等。

2.1.6 推荐节点方案

UCloudStack 云平台轻量且架构灵活，物理节点方案可根据企业业务需求及应用场景进行灵活调整，可部署的推荐方案举例如下：

(1) 计算存储融合节点

如 3+3 的节点方案，即 3 台 SATA+SSD 缓存超融合计算节点、3 台 SSD 超融合计算节点，管理服务部署于计算节点的虚拟机中，后续可根据业务规模水平扩展，如将 SATA+SSD 缓存超融合计算节点扩容为 9 台。

(2) 独立计算节点+独立存储节点

适用于存算分离的大规模场景，如 N ($N \geq 2$) 台计算节点、3 台 SSD 存储节点、3 台 SATA+SSD 缓存节点，分别构建计算集群、SSD 存储集群及 SATA 存储集群，共同构建一套云平台，后续可根据业务规模分别对集群的计算节点或存储节点进行水平扩展，满足业务发展规划。

! 注意 SSD 和 SATA 节点的配比取决于业务需求，如高存储容量需求较大，则需配置较多的 SATA+SSD 缓存节点；若高性能业务需求较多，则需配置较多的 SSD 全闪节点。

(3) 独立计算节点+商业存储设备

适用于利旧商业存储设备的场景，如 N ($N \geq 2$) 台计算节点、 N 台商业存储设备，分别构建计算集群和商业存储集群，共同构建一套云平台，后续可根据业务规模分虽对计算集群和商业存储设备进行扩展。

(4) 计算存储融合节点+独立存储节点

适用于超融合集群扩展大容量存储资源的场景，如 N ($N \geq 3$) 台计算存储融合节点、 N ($N \geq 3$) 台独立存储节点，分别构建计算集群和商业存储集群，共同构建一套云平台，后续可根据业务规模分虽对计算集群和存储集群进行扩展。

(5) 计算存储融合节点+独立计算节点

适用于超融合集群扩展算力资源的场景，如 N ($N \geq 2$) 台计算节点、 N 台商业存储设备，分别构建计算集群和商业存储集群，共同构建一套云平台，后续可根据业务规模分虽对计算集群和商业存储设备进行扩展。

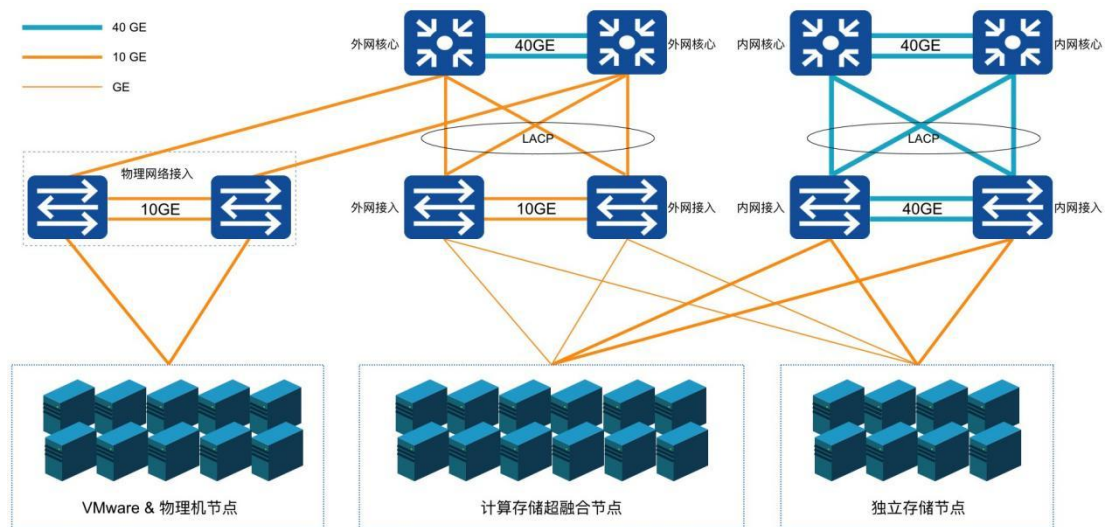
(6) 计算存储融合节点+独立计算节点+独立存储节点+商业存储设备

适用于复杂型算力架构场景，将即有的设备利旧的同时，满足大规模算力和存储空间的需求。如 N ($N \geq 2$) 台计算节点、 N 台商业存储设备，分别构建计算集群和商业存储集群，共同构建一套云平台，后续可根据业务规模分虽对计算集群和商业存储设备进行扩展。

最佳实践中，生产环境至少需要 3 台 SATA/SSD 超融合节点部署搭建 UCloudStack 平台，即 UCloudStack 最小生产规模为 3 台服务器。具体服务器配置要求详见【硬件选型】章节的【最低配置】。

2.2 物理网络架构

为构建高可用、高可靠、高安全的企业专有云平台，UCloudStack 平台均采用高可用冗余性设计。本文以标准网络拓扑图为基础进行物理网络架构描述，本架构设计至少需要 6 台万兆交换机、2 台千兆交换机、多台计算&存储节点服务器。若有 IPMI 管理及网络设备管理等需求，可根据需求增加 IPMI 和 Management 交换机并接入网络。



平台网络设计为核心、接入二层架构，接入交换机双上联到核心，且按计算业务分集群划分。本架构设计从业务场景上提供公网服务，因此整体业务架构分为内网区域和外网区域两张网络，分别承载云平台内网通信和外网通信，两张网络在网络设备层面物理隔离。

2.2.1 架构规模

标准的网络架构为单数据中心网络架构，以下述配置为例，单数据中心可支撑 900~1000 台规模的节点数量：

- 两台交换机堆叠在一起，称为一组交换机，如一组内网接入交换机或一组外网接入交换机；
- 通常一组接入交换机为 96 个业务接口（每台交换机 48 个接口），堆叠检测及备用占用 3*2 个接口，可用业务端口为 90 个；

- 每个服务器节点使用两个网卡占用一组接入交换机的 2 个接口，即一组接入交换机可接入 45 台服务器；
- 每增加一组交换机即可扩展 45 个节点，一组核心交换机至少可接入 20 组接入交换机，即至少可支撑 900 个节点服务器。

2.2.2 网络区域

网络区域的设备通常包括内网核心交换机、外网核心交换机、内网接入交换机、外网接入交换机。若服务器节点规模较小且暂不考虑扩容，可仅采用内/外网接入交换机。

- **内网核心交换机：**采用 2 台 40GE 的三层交换机堆叠作为一组内网核心，用于承载内网接入交换机的汇聚和管理；
- **外网核心交换机：**采用 2 台万兆三层交换机堆叠作为一组外网核心，用于承载外网接入交换机的汇聚和管理；
- **内网接入交换机：**采用 2 台万兆交换机堆叠作为一组内网接入，用于承载 45 台服务器内网接入；
- **外网接入交换机：**采用 2 台千兆交换机堆叠作为一组外网接入，用于承载 45 台服务器外网接入；

除 Internet 连接外，网络均为大二层环境，采用 LLDP 协议获取网络拓扑信息，所有网络接入均为端口聚合，保证高可用；同时通过控制接口广播报文流量，抑制网络广播风暴。

外网核心交换机与 Internet 之间可以为二层聚合、三层聚合、L3 ECMP、L3 A/S 等互连模式，同时支持串联或旁挂防火墙、IDS、IPS 及防 DDOS 等安全设备。

云平台提供的网络功能均采用软件定义的方式实现，物理交换机仅作为网络流量转发设备，即仅使用交换机部分通用能力，如堆叠、Vlan、Trunk、LACP 及 IPV6 等，无需采用 SDN 交换机实现虚拟网络的通信。

标准网络架构中，通常推荐至少采用万兆及以上级别的交换机，保证平台节点内网接入、虚拟资源通信及分布式存储的性能及可用性。由于外网接入带宽一般较小，通常推荐采用千兆交换机作为外网接入设备。

2.2.3 服务器区域

服务器区域的设备通常包括计算存储超融合节点、独立计算节点、独立存储节点、以及管理节点。若直接使用计算节点的虚拟机作为管理节点，即可省去物理管理节点服务器。

(1) 计算节点【必选】

采用 x86/ARM 架构服务器作为计算节点或计算存储超融合节点，用于运行虚拟机、虚拟网络、分布式存储及数据库缓存等服务，承载整个云平台的资源核心实现及运行。

- 采用 2 个 GE 网卡分别上联到两台外网接入交换机，并做双网卡 bond，作为计算节点外网接入。
- 采用 2 个 10GE 网卡分别上联到两台内网接入交换机，并做双网卡 bond，作为计算节点内网接入。
- 若为超融合节点，则分布式存储使用所有计算节点上的数据磁盘，所有计算节点上的数据磁盘组成统一分存储资源池，用于构建分布式存储。
- 若为独立计算节点，则分布式存储使用存储节点上的数据磁盘作为统一存储资源池，通过网络跨集群挂载。



警告 为提升平台虚拟化算力资源的稳定性和可用性，若计算节点的 CPU 型号不一致，必须将不同型号 CPU 节点划分至不同集群。

(2) 独立存储节点【可选】

若计算存储需要分离部署，可采用 x86/ARM 架构且磁盘较多服务器作为独立存储节点，用于承载独立的分布式存储服务。

- 存储节点与计算服务通过内网进行通信，仅需 2 个 10GE 网卡分别上联

到两台内网接入交换机，并做双网卡 bond，作为存储节点的内网接入。

- 如需将计算存储网络物理隔离，可采用独立存储接入交换机，存储节点的网卡上联至存储接入交换机。

分布式存储使用存储节点及超融合节点上的所有数据磁盘，三副本保证数据安全。若采用商业存储（SAN），则云平台服务可使用商业存储划分的 LUN 存储空间作为云平台的后端存储。

- 每台分布式存储节点的数据盘仅支持单种介质类型，即单节点不支持 SSD 和 HDD 混插，分别组建不同的存储集群。
- 为保证存储性能，若分布式存储节点采用 SATA 机械盘组建大容量存储集群，必须配置 SSD/NVME 高性能磁盘作为缓存盘，并须保证 SSD 缓存盘和 HDD 数据盘的容量比不高于 1:20；盘数量比不高于 1:5。

注意 为保证分布式存储的性能及可用性，存储节点必须采用万兆以上速率的网卡。

(3) 管理节点【可选】

平台默认推荐使用平台虚拟机部署管理服务，如需物理服务器承载并运行管理服务，可采用 x86/ARM 服务器作为云平台管理节点，用于承载云平台管理模块及服务。

- 采用 2 个 GE 网卡分别上联到两台外网接入交换机，并做双网卡 bond，作为管理节点外网接入；
- 采用 2 个 10GE 网卡分别上联到两台内网接入交换机，并做双网卡 bond，作为管理节点内网接入。

注意 以上网卡 bond 均采用“mode=4”模式，即 IEEE 802.3ad 动态链路聚合。

2.2.4 标准架构扩展

在实际项目中，根据用户需求和所提供的环境，可对标准网络架构进行调整，如项目较小规模（45 节点内）或仅需一个简单的测试环境或等场景。

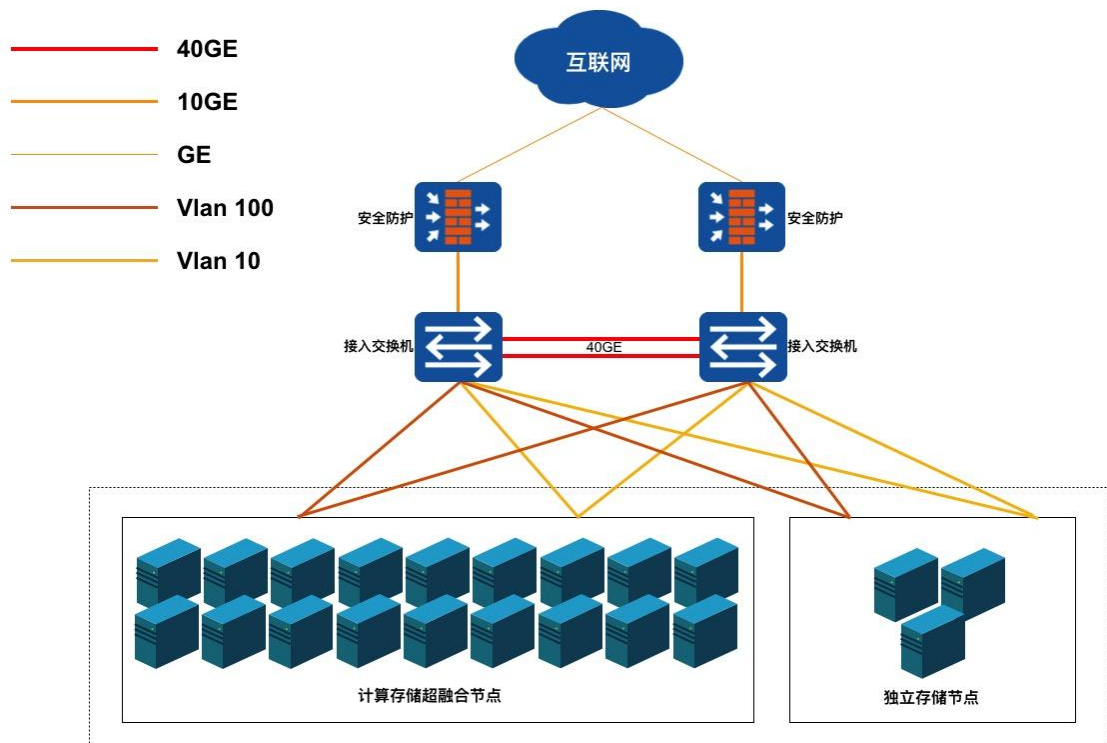
(1) 需内外网物理隔离且考虑接入冗余，可采用 2 组共 4 台接入交换机部署：

- 2 台堆叠用于服务器内网接入，2 台堆叠用于服务器外网接入；
- 每台服务器内外网分别使用 2 个接口绑定接入内外网接入交换机，可支持 45 台服务器节点冗余接入（每台交换机 48 个接口，考虑堆叠检测和备用的端口占用）。

(2) 需内外网物理隔离且不考虑接入冗余，可用 2 台接入交换机进行业务部署：

- 1 台用于服务器内网接入，1 台用于服务器外网接入；
- 每台服务器分别使用 1 个接口接入内网接入交换机及外网接入交换机，支持 45 台服务器节点接入；

(3) 若内外网无需物理隔离且考虑接入冗余，可采用 2 台交换机堆叠，通过 Vlan 隔离内外网，如下图所示：



- **方案一：**通过在交换机上划分 Vlan，服务器分别使用 2 个接口绑定接入交换机内外网 Vlan 接口，即每台服务器需 2 组 bond(4 个接口) 实现内外网业务通信，可支持 22 节点；

- **方案二：**通过在服务器操作系统内划分 Vlan（即子接口），服务器分别使用 2 个接口绑定接入交换机 Trunk 接口，即每台服务器仅需 2 个接口绑定实现内外网业务通信，可支持 45 节点；

(4) 若内外网无需物理隔离且不考虑接入冗余，可采用 1 台交换机，通过交换机 Vlan 或服务器内划分 Vlan 进行内外网隔离及接入。

(5) 若实际环境中需要采用独立的计算节点和独立的存储节点，并需要将计算网络和存储网络进行物理隔离。

- 可以为独立存储节点单独划分一对接入交换机上联至内网核心交换机，实现计算和存储网络进行分离。
- 平台计算虚拟机可通过物理网络挂载多个存储网络的存储集群，采用独立的存储网络设计可将存储节点及分布式存储系统内部同步流量与虚拟机计算读写存储的流量进行分离，提高平台整体的稳定性和性能。

2.3 硬件选型

2.3.1 最低硬件配置

用户在部署云平台时可选择超融合节点或存算分离节点进行部署，超融合节点和独立节点最低硬件配置要求如下：

(1) 超融合节点

用于生产环境的最低主机和网络硬件配置，一般生产环境至少需要 3 台超融合节点和 2 台万兆接入交换机。针对测试环境和生产环境最低配置要求如下：

配置项	测试环境	生产环境	备注
CPU	CPU 不低于 10 核	CPU 不低于 16 核	每增加一块数据盘，需增加 2 核
内存	内存不低于 32GB	内存不低于 32GB	每增加一块数据盘，需增加 4GB

网卡	1 个 10GB 网口	2 个 10GB 网口	若考虑网卡冗余，推荐 2 张 10GB 网卡
系统盘	单块硬盘不小于 240GB	2 个 SSD 480GB RAID1	
数据盘	SSD 盘：最少 1 块。	SSD 盘：最少 1 块	1、整个分布式存储集群，需要 2TB 可用容量，作为云平台自身的容量预留。 2、如节点只配置 1 块数据盘，单块 HDD 不低于 2TB，单块 SSD 不低于 1.92TB。
	HDD 盘：最少 1 块，且必须配置 SSD/NVME 缓存盘，存储盘的容量比不高于 1:20；数量比不高于 1:5	HDD 盘：最少 1 块，且必须配置 SSD/NVME 缓存盘，存储盘的容量比不高于 1:20；数量比不高于 1:5	
节点数量	3 台	3 台	
接入交换机	1 台万兆以太网交换机	2 台万兆以太网交换机	

(2) 存储分离节点

用于生产环境的最低主机和网络硬件配置，一般生产环境至少需要 2 台计算节点、3 台存储节点及 2 台万兆接入交换机。针对测试环境和生产环境最低配置要求如下：

硬件类型	配置项	测试环境	生产环境
计算节点	CPU	CPU 不低于 10 核	CPU 不低于 16 核
	内存	内存不低于 32GB	内存不低于 32GB
	网卡	1 个 10GB 网口	2 个 10GB 网口做网卡冗余
	系统盘	1 个 SSD 480GB	2 个 SSD 480GB RAID1
	节点数量	1	2
存储节点	CPU	CPU 不低于 10 核	CPU 不低于 16 核
	内存	内存不低于 32GB	内存不低于 32GB

	网卡	1 个 10GB 网口	2 个 10GB 网口做网卡冗余	
	系统盘	1 个 SSD 480GB	2 个 SSD 480GB RAID1	
	数据盘	SSD 盘: 最少 1 块。	SSD 盘: 最少 1 块	
		HDD 盘: 最少 1 块, 且必须配置 SSD/NVME 缓存盘, 存储盘的容量比不高于 1:20; 数量比不高于 1:5	HDD 盘: 最少 1 块, 且必须配置 SSD/NVME 缓存盘, 存储盘的容量比不高于 1:20; 数量比不高于 1:5	
		1、整个分布式存储集群, 需要 2TB 可用容量, 作为云平台自身的容量预留。 2、如节点只配置 1 块数据盘, 单块 HDD 不低于 2TB, 单块 SSD 不低于 1.92TB。		
	商业存储	商业存储模式无需存储节点, 商业存储上需预留 1~2TB 可用容量, 作为云平台自身容量预留。		
节点数量	3	3		
网络设备	接入交换机	1 台万兆以太网交换机	2 台万兆以太网交换机	

! 注意 最低配置建议, 只保证云平台的正常部署和稳定运行, 不考虑租户或用户的资源预留。生产环境下服务器硬件和架构层必须保证冗余机制。

2.3.2 推荐硬件配置

(1) 网络设备推荐

业务	配置描述	构建方案
内网核心交换机	CE88-40G 板卡(16 口)*4, 64*40GE	可选 (48 节点以上扩容)
外网核心交换机	48*10GE + 6*100GE	可选 (48 节点以上扩容)
内网接入交换机	48*10GE + 6*100GE	必选
存储接入交换机	48*10GE + 6*100GE	可选 (独立存储区域)

外网接入交换机	48*GE + 4*10GE + 2*40GE	可选 (内外网物理隔离)
管理汇报交换机	48*GE + 4*10GE + 2*40GE	可选 (构建运维管理网)
IPMI 接入交换机	48*GE + 4*10GE + 2*40GE	可选 (构建 IPMI 管理网)
网络设备 MGT 接入	48*GE + 4*10GE + 2*40GE	可选 (构建 MGT 管理网)

(2) 服务器推荐配置

机型	配置描述
融合型——低配	<p>Factor Form 2U</p> <p>CPU Intel Xeon Silver 4310 Processor(12CORES_2.1GHz_120W_X86) *2</p> <p>DDR4_32GB_RDIMM_3200MHz *4</p> <p>OS HDD 480G_SSD_SATA3_512E_2.5"_6Gb/s *2</p> <p>Cache HDD 960G_SSD_U.2_N/A_512E_2.5"_32Gb/s*2</p> <p>Data HDD SATA3_HDD_8TB *4</p> <p>LSI-9311-8I*1</p> <p>双口万兆光口网卡(不含光模块)*2</p> <p>PSU=800W*2/导轨</p>
融合型——中配	<p>Factor Form 2U</p> <p>CPU Intel Xeon Silver 4310 Processor(12CORES_2.1GHz_120W_X86) *2</p> <p>DDR4_32GB_RDIMM_3200MHz *8</p> <p>OS HDD 480G_SSD_SATA3_512E_2.5"_6Gb/s *2</p> <p>Cache HDD 1.92T_SSD_U.2_N/A_512E_2.5"_32Gb/s*2</p> <p>Data HDD SATA3_HDD_8TB *10</p> <p>LSI-9311-8I</p>

	<p>双口万兆光口网卡(不含光模块)*2</p> <p>PSU=800W*2/导轨</p>
融合型——高配	<p>Factor Form 2U</p> <p>CPU Intel Xeon Silver 4310 Processor(12CORES_2.1GHz_120W_X86) *2</p> <p>DDR4_32GB_RDIMM_3200MHz *16</p> <p>OS HDD 480G_SSD_SATA3_512E_2.5" _6Gb/s *2</p> <p>Cache HDD 3.84T_NVME_U.2_N/A_512E_2.5" _32Gb/s*2</p> <p>Data HDD SATA3_HDD_16TB *10</p> <p>LSI-9311-8I</p> <p>双口万兆光口网卡(不含光模块)*2</p> <p>PSU=800W*2/导轨</p>
存储型——低配	<p>"Factor Form 2U</p> <p>CPU Intel Xeon Silver 4310 Processor(12CORES_2.1GHz_120W_X86) *2</p> <p>DDR4_32GB_RDIMM_3200MHz *4</p> <p>OS HDD 480G_SSD_SATA3_512E_2.5"" _6Gb/s *2</p> <p>Cache HDD 3.84T_NVME_U.2_N/A_512E_2.5"" _32Gb/s*2</p> <p>Data HDD SATA3_HDD_16TB *10</p> <p>LSI-9311-8I</p> <p>双口万兆光口网卡(不含光模块)*2</p> <p>PSU=800W*2/导轨"</p>
存储型——高配	<p>"Factor Form 4U</p> <p>CPU Intel® Xeon® Silver 4314 Processor(16CORES_2.4GHz_135W_X86) *2</p> <p>DDR4_32GB_RDIMM_3200MH*8</p> <p>OS HDD</p>

	480G_SSD_SATA3_512E_2.5""_6Gb/s *2 Data HDD SATA3_HDD_16TB *36 LSI-9311-8I 双口万兆光口网卡(不含光模块)*2 PSU=1300W*2/导轨"
计算型	"Factor Form 2U CPU Intel Xeon Silver 4310 Processor(12CORES_2.1GHz_120W_X86) *2 DDR4_32GB_RDIMM_3200MHz *8 OS HDD 480G_SSD_SATA3_512E_2.5""_6Gb/s *2 LSI-9311-8I 双口万兆光口网卡(不含光模块)*2 PSU=800W*2/导轨"

2.4 平台资源占用

云平台运行本身需要占用服务器的 CPU、存储及存储资源，其中存储服务的 CPU 和内存为预估值，实际生产环境中，根据使用负载等情况，可能会有变动，具体如下：

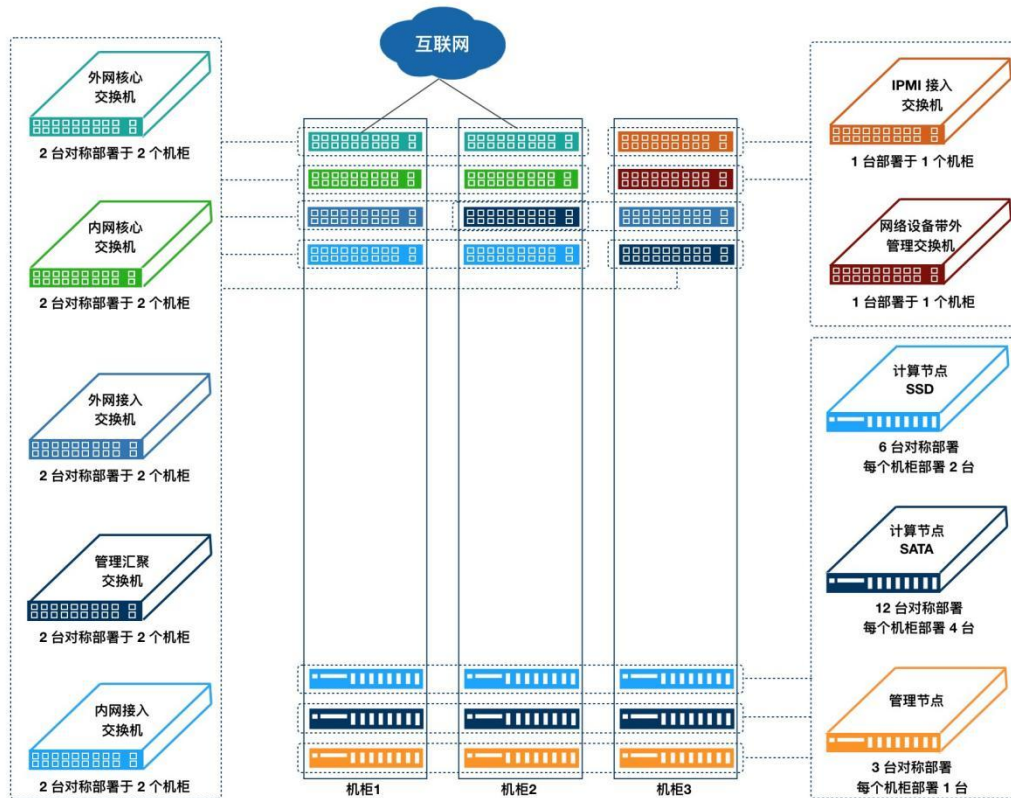
模块	角色	数量	CPU	内存	存储	说明
调度服务	调度管理服务	3	4C	8GB	400GB	CPU 内存占用管理集群物理资源, 存储占用管理节点本地存储资源
	每计算节点——计算服务	N	4C	4GB	400GB	CPU 内存占用计算集群物理资源, 存储占用计算节点本地存储资源
	存储服务	3	4C	4GB	400GB	均占用存储节点本地物理资源

存储服务	每存储节点——N 块硬盘存储服务	N	2C	4GB		均占用存储节点 本地物理资源
	缓存加速模式	每 TB		4GB		每 TB 缓存容量需 要消耗 4GB 内存
云管理服务	管理服务	1	4C	8GB	240GB	CPU 内存占用计 算集群资源, 存储 数据占用分布式 存储资源
	公共服务	1	4C	8GB	640GB	CPU 内存占用计 算集群资源, 存储 数据占用分布式 存储资源

2.5 机柜空间规划

网络设备和服务器的物理机柜空间规划如下图所示:

示例架构机柜空间规划方案



所有设备在机柜中对称部署, 实现机柜级冗余, 单机柜掉电或故障不影响云

平台业务。一个机柜可支撑 15 个节点，根据网络架构设计一组接入交换机支撑 45 个节点，即一组接入交换机支撑 3 个机柜。3 个机柜为 1 组，平均 1 组机柜支撑 45 个节点、1 组内网接入交换机、1 组外网接入交换机、1 台 IPMI 接入交换机。

如上图项目案例中的设备包括 8 台业务交换机、4 台运维管理交换机、21 台服务器设备及 3 个机柜：

- 一组内网核心交换机对称部署于 2 个机柜，即其中两个机柜各部署 1 台；
- 一组内网接入交换机对称部署于 2 个机柜，即其中两个机柜各部署 1 台；
- 一组外网核心交换机对称部署于 2 个机柜，即其中两个机柜各部署 1 台；
- 一组外网接入交换机对称部署于 2 个机柜，即其中两个机柜各部署 1 台；
- 一组管理汇聚交换机对称部署于 2 个机柜，即其中两个机柜各部署 1 台；
- 1 台 IPMI 接入交换机和 1 台网络设备带外管理交换机部署于 1 个机柜；
- 3 台管理节点对称部署于 3 个机柜，即每个机柜各部署 1 台；
- 12 台计算&SATA 节点对称部署于 3 个机柜，即每个机柜各部署 4 台；
- 6 台计算&SSD 节点对称部署于 3 个机柜，即每个机柜各部署 2 台。

若服务器分集群部署云平台，建议不同集群的服务器对称部署于多个机柜。

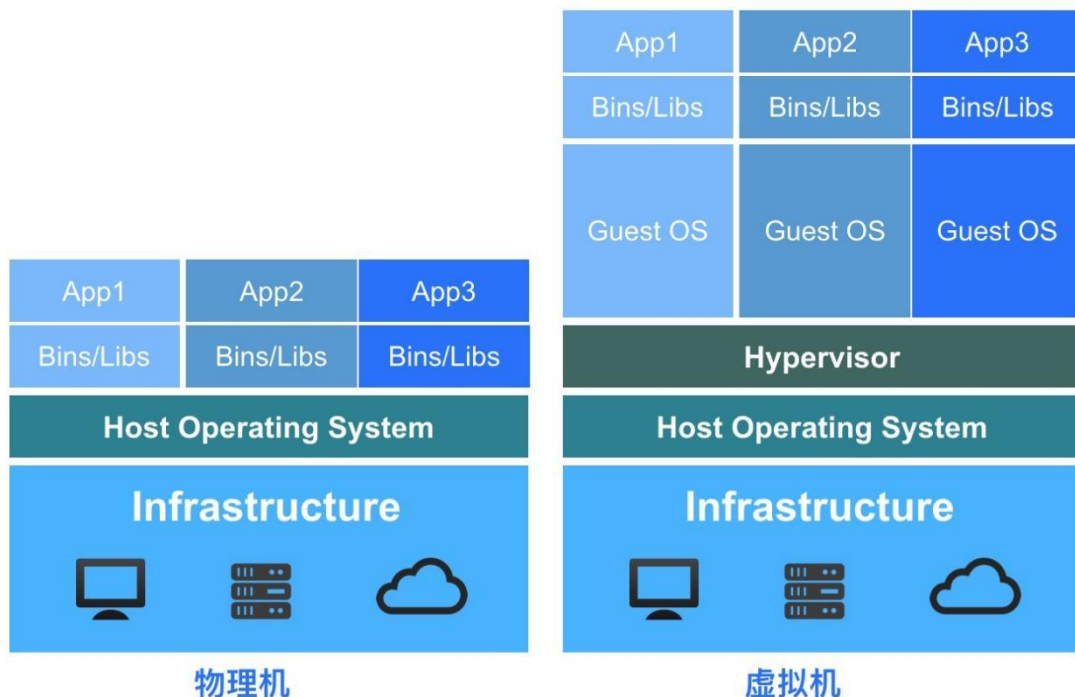
3 平台技术架构

UCloudStack 平台基于 UCloud 公有云平台，复用公有云核心组件，具备计算虚拟化、智能调度、存储虚拟化、网络虚拟化的基础能力，为用户提供软件定义的计算、存储、网络及资源管理等服务，在保证资源服务性能、可用性及安全性的同时，提供统一资源调度及管理服务，适应企业基础设施服务的多种应用场景。

在标准化通用产品服务外，平台针对国产化场景提供一云多芯的信创云平台，信创私有云在标准版特性基础之上，面向国产化场景，提供通过信创互认证的 IaaS 和 PaaS 功能，兼容硬件、操作系统、中间件到上层应用的全信创生态，打造新基建下的安全、可信的云环境。

3.1 计算虚拟化

云计算技术是虚拟化技术的延伸，计算虚拟化是在硬件之上增加一个 Hypervisor，通过它虚拟出多个完全隔离的主机并可安装不同的操作系统，承载不同的应用程序运行，最大程度上解决了一台物理机被一个系统或一个应用占用的问题，有效的提高资源使用率。



如上图所示，物理机和虚拟机在应用部署及资源占用上有本质区别：

(1) 物理机环境

- 操作系统是直接安装在物理机上，通常一台物理机只支持安装一个操作系统；
- 所有的应用程序和服务均需部署在物理机操作系统上，共享底层硬件资源；
- 多个应用程序对底层操作系统的及组件要求不一致时，可能会导致应用无法正常运行，需要将两个应用程序分属部署至一台物理机上，在非业务高峰时资源利用率较低。

(2) 虚拟机环境

- 在硬件底层及操作系统上增加 Hypervisor 层，作为计算虚拟化的引擎；
- 虚拟化引擎支持将底层硬件虚拟为多个主机，即虚拟机；
- 每个虚拟机都拥有独立的硬件设施，如 CPU、内存、磁盘、网卡等；
- 每个虚拟机可以独立安装并运行不同的操作系统 (GuestOS)，相互完全隔离，彼此不受影响；
- 每个虚拟机操作系统与物理机的操作系统一致，拥有独立的组件及库文件，可运行专属应用服务；
- 多个应用程序的虚拟机在完全隔离且彼此不影响的情况下运行在一台物理机上，并共享物理机的资源，提高物理机的资源使用率及管理效率。

优刻得私有云平台计算虚拟化采用 KVM 和 Qemu 等 Hypervisor 组件及技术，将通用裸金属架构的 x86/ARM 服务器资源进行抽象，以虚拟机的方式呈现给用户。虚拟机将 CPU、内存、I/O、磁盘等服务器物理资源转化为一组可统一管理、调度和分配的逻辑资源，并基于虚拟机在物理机上构建多个同时运行、相互隔离的虚拟机执行环境，可充分利用硬件辅助的完全虚拟化技术，实现高资源利用率的同时满足应用更加灵活的资源动态分配需求，如快速部署、资源均衡部

署、重置系统、在线变更配置及热迁移等特性，降低应用业务的运营成本，提升部署运维的灵活性及业务响应的速度。

平台计算虚拟化通过 KVM 硬件辅助的全虚拟化技术实现，因此需要 CPU 虚拟化特性的支持，即要求计算节点 CPU 支持虚拟化技术，如 Intel VT 和 AMD V 技术。KVM 属于 Linux Kernel 的一个模块，虚拟化平台可通过加载内核模块的方式启动 KVM，管理虚拟硬件的设备驱动，用于模拟 CPU 和内存资源，同时需要加载 QEMU 模块模拟 I/O 设备。KVM 虚拟机包括虚拟内存、虚拟 CPU 和虚拟机 I/O 设备，其中 KVM 用于 CPU 和内存的虚拟化，QEMU 用于 I/O 设备的虚拟化。

虚拟机不直接感知物理 CPU，它的计算单元会通过 Hypervisor 抽象的 vCPU 和内存进行呈现，通过与 GuestOS 的结合共同构建虚拟机系统。I/O 设备的虚拟化是 Hypervisor 复用外设资源，通过软件模拟真实硬件进行呈现，为 GuestOS 提供诸如网卡、磁盘、USB 设备等外设。

计算虚拟化是 UCloudStack 企业专有云平台的服务器虚拟化组件，是整个云平台架构的核心组件。在提供基础计算资源的同时，支持 CPU 超分、QCOW2 镜像文件、GPU 透传、USB 透传、物理机纳管及集群平滑扩容等特性。

3.1.1 CPU 超分

云平台支持平台物理 CPU 超分，即平台可虚拟化的 vCPU 数量可大于 pCPU 数量，在分配给虚拟机的 CPU 资源未全部使用时，共享未使用的部分给其它虚拟机使用，进一步提高平台 CPU 资源使用率。

支持用户按比例进行 CPU 超分，调整 CPU 超分比例延时生效，超分获得的可分配核数，以超分结果为准。平台自服务支持超分比例 100% (流畅)、150%、200% (低风险)、250%、300%、350%、400% (高风险)、450%、500%、550%、600%。

具体超分可分配核数以 1 台双路 CPU 的计算节点服务器为例：

- 双路 CPU 即为 2 颗物理 CPU，每颗物理 CPU 为 12 核，开启双线程；

- 每颗 CPU 为 24 核，两颗 CPU 为 48 核，即可分配 48 vCPU；
- 正常情况下，能提供的虚拟机 vCPU 为 48C；

若平台管理员开启 CPU 超分，并设置超分比例为 1:2，即代表可使用的 vCPU 数量是实际 CPU 数量的 2 倍。服务器（48C）在开启 2 倍超分后，可实际创建使用的 vCPU 为 96，即可创建 96C 的虚拟机。但不支持向下修改，即如果已经设置了超分比为 1:2，则不再允许将超分比调为 1:1。

仅支持平台专业的运维人员设置并管理 CPU 超分比，平台管理员可查看平台 CPU 的实际使用量及 vCPU 的使用量。由于开启超分后，可能存在多台虚拟机共用 vCPU 的情况，为不大幅影响虚拟机的性能及可用性，通常建议尽量降低 CPU 超分比例，甚至不建议开启 CPU 超分。

如平台实际共 48 vCPU，经过超分后可创建 96 vCPU 的虚拟机，在虚拟机业务峰值时可能会真正占满 48 vCPU 的性能，通过超分资源运行的虚拟机性能会极速下降，甚至会影响虚拟机的正常运行。

CPU 超分比例需通过长期运行运营的数据进行调整，与平台虚拟机上所运行的业务应用程序有强关联性，需要长期考察平台在峰值业务时需要的 CPU 资源量进行灵活调整。

3.1.2 镜像文件

平台使用 RAW 格式的镜像作为虚拟机的虚拟磁盘文件，即原始镜像。

RAW 镜像会直接当作一个块设备提供给虚拟机使用，由宿主机文件系统管理镜像文件的空洞，如创建一个 100GB 的 RAW 的镜像文件，实际占用空间很小。当虚拟机的 GuestOS 读写磁盘时，会以 CHS 变量方式进行运算并寻址，通过 KVM 驱动将值翻译成 RAW 镜像特有格式进行 IO 操作。

RAW 镜像的优势在于启动虚拟机的效率较高，即启动虚拟机速度较快，相比 QCOW2 格式镜像的虚拟机启动速度快 25%，同时 RAW 镜像支持转换为 QCOW2 格式。

RAW 镜像及运行的虚拟机块设备均会存储于统一分布式存储系统中，方便

虚拟机的迁移和故障恢复。

平台支持导入 QCOW2 和 ISO 格式的虚拟机镜像, 为方便镜像文件的上传、下载及分发, 平台使用占用空间更小的 QCOW2 格式的镜像文件。

当需要使用一个全新的操作系统时, 可以选择使用包含操作系统介质的 ISO 镜像, 直接使用 ISO 介质引导并安装到虚拟机中, 方便用户快速创建自定义操作系统的虚拟机。

当创建虚拟机时, 平台会将第一次使用到的镜像自动转换为 RAW 格式并导入到分布式存储中, 通过分布式存储快照能力实现虚拟机磁盘的快速创建。

转换格式后的镜像及运行的虚拟机块设备均会存储于统一分布式存储系统中, 方便虚拟机的迁移和故障恢复。

支持虚拟机加载 ISO 镜像, 加载后可在系统内挂载/dev/sr0 或设置从该设备引导。并支持设置 ISO 为系统默认引导项, 关机重启后会从 ISO 引导进入装机页面。

3.1.3 GPU 透传

平台支持 GPU 设备透传能力, 为平台用户提供 GPU 虚拟机服务, 让虚拟机拥有高性能计算和图形处理能力。GPU 虚拟机在科学计算表现中比传统架构性能提高数十倍, 可同时搭配 SSD 云硬盘, IO 性能亦在普通磁盘的数十倍以上, 可有效提升图形处理、科学计算等领域的计算处理效率, 降低 IT 成本投入。

GPU 虚拟机与标准虚拟机采用一致管理方式, 包括内外网 IP 分配、弹性网卡、子网及安全组管理, 并可对 GPU 虚拟机进行全生命周期管理, 包括重置密码, 变更配置及监控等, 使用方式与普通的虚拟机一致, 支持多种操作系统, 如 CentOS、Ubuntu、Windows 等, 在不增加额外管理的基础上, 为租户提供快捷的 GPU 计算服务。

为让 GPU 发挥最佳性能, 平台对 GPU、CPU 及内存的组合定义如下:

GPU	CPU	内存
-----	-----	----

1 颗	4 核	8G, 16G
	8 核	16G, 32G
2 颗	8 核	16G, 32G
	16 核	32G, 64G
4 颗	16 核	32G, 64G
	32 核	64G, 128G

平台本身不限制 GPU 品牌及型号，即支持任意 GPU 设备透传，已测试并兼容 GPU 型号为 NVIDIA 的 K80、P40、V100、2080、2080Ti、T4 及华为 Atlas300。

3.1.4 USB 透传

平台支持 USB 设备透传能力，使虚拟机可以使用物理机上的 USB 设备，从而使虚拟机可与宿主机 USB 设备进行数据交互。平台提供 USB Passthrough（设备直通）及 USB Redirection（设备转发）两种模式。

3.1.4.1 USB 直通

- **实现：**直通模式是将主机上的 USB 控制器或具体的 USB 设备直接分配给虚拟机的技术，通过 Qemu hostdev 设备方式挂载到虚拟机实例中，使得虚拟机能够直接访问和控制 USB 设备。
- **优势：**USB 设备直通能够提供更高的性能和更低的延迟，适用于对性能要求较高的 USB 设备，如外部存储设备等。
- **限制：**直通模式下虚拟机只能使用所在宿主机上的 USB 设备，迁移虚拟机时需要卸载 USB 设备。

3.1.4.2 USB 转发

- **实现：**转发模式通过 Qemu redir 设备方式挂载到虚拟机中，底层协议是 TCP 协议，主要原理是将“USB I/O 消息”封装成“TCP/IP 消息”，虚拟化层通过网络访问远端 USB 设备；转发模式下的 USB 设备不可再直通给虚拟机使用。
- **优势：**转发模式相对灵活，虚拟机可以跨节点使用其他宿主机上的 USB 设备，迁移时不需要卸载此 USB 设备。
- **限制：**延迟相对较高，不适用于对性能要求较高的设备。

在实际应用中，选择直通还是转发模式通常取决于具体的需求和应用场景。性能敏感的应用可能更倾向于使用 USB 设备直通，而对性能要求不苛刻的场景可能会选择 USB 设备转发，以保持更大的灵活性。平台提供了相应的功能界面和接口，使用户在创建虚拟机时能够根据需要配置选择对应的模式。

3.1.5 物理机纳管

UCloudStack 平台支持物理机纳管能力，将用户存量的物理机纳管到平台中，支持对物理机进行开机、关机、重启、登陆物理机等操作。

平台通过融合和适配传统的 IPMI 和 Redfish 协议，收集硬件信息和适配远程控制 KVM 镜像，实现物理机信息查询和管理。

支持适配不同的机型，根据用户的个性化需求进行定制化开发。已适配支持的机型包括浪潮 SA5212M4、联想 ThinkSystem SR650、联想 ThinkSystem SR658、新华三 R4900 G2、新华三 UniServer R4900 G3 等。

3.1.6 集群平滑扩容

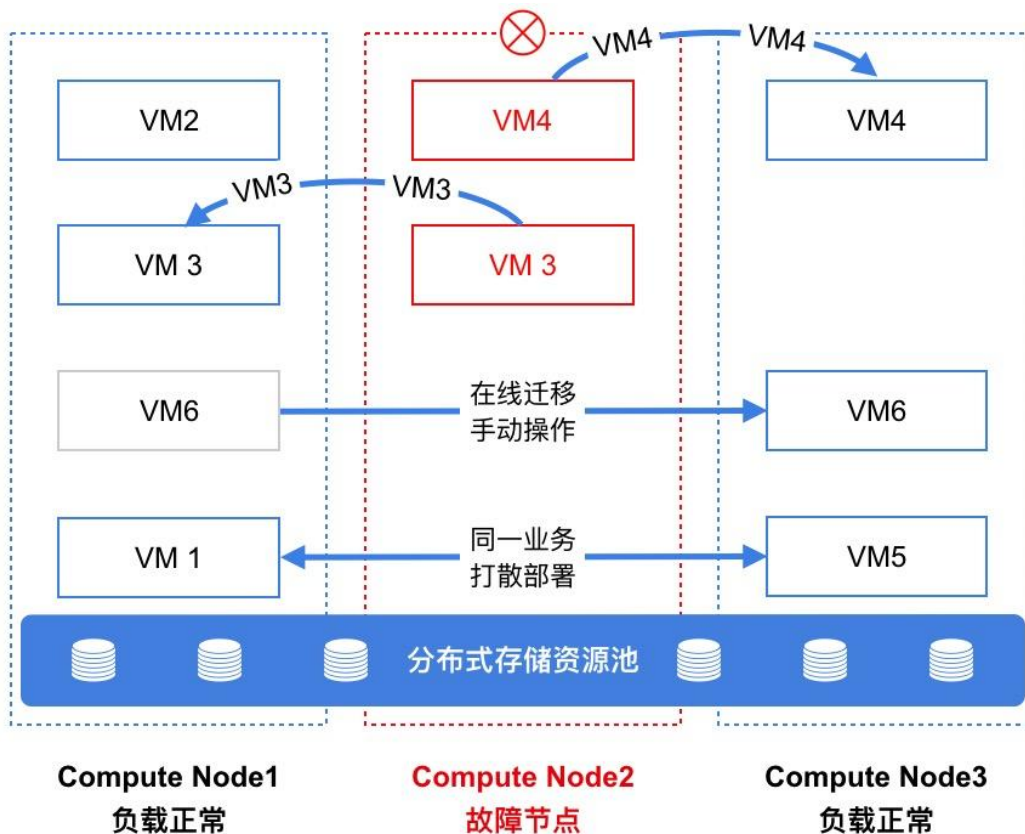
平台支持平滑扩容集群内的计算节点，新增的节点不会影响已有节点及虚拟资源的运行。通过平滑扩容云平台管理员可轻松解决平台因业务增长而带来的资源扩展，包括硬件资源不足、高负载主机维护、新业务上线资源扩容等场景。

UCloudStack 集群扩容可保证节点扩展过程中业务不中断, 虚拟资源均正常运行, 并提供简单快速的部署操作, 支持自动化脚本一键部署上线。在扩容后平台支持在线为节点添加磁盘功能, 使管理员可在不影响平台稳定运行的情况下, 为平台横向及纵向的扩展资源。

平滑扩容成功后, 平台原有的虚拟资源会保持原始状态, 待平台有新的虚拟资源需要运行和部署时, 智能调度平台会将新的虚拟资源 (如虚拟机) 调度至平滑扩容的节点; 若平台有物理机发生故障, 原物理机上的虚拟机会根据调度策略迁移至新扩容的节点。支持平台管理员手动将一台虚拟机迁移至新扩容的节点, 用于平衡平台整体资源使用率。

3.2 智能调度

智能调度是平台虚拟资源调度管理的核心, 由调度模块负责调度任务的控制和管理, 用于决策虚拟机运行在哪一台物理服务器上, 同时管理虚拟机状态及迁移计划, 保证虚拟机可用性和可靠性。



3.2.1 均衡调度

智能调度系统实时监测集群所有节点计算、存储、网络等负载信息，作为虚拟机调度和管理的数据依据。当有新的虚拟资源需要部署时，**调度系统会优先选择低负荷节点进行部署**，确保整个集群节点的负载相对均衡。如上图所示，新创建的虚拟资源将会通过调度检测，自动部署至负载较低的 **Node3** 节点上。

调度系统在优先选择低负荷节点进行虚拟资源部署的同时，分别提供打散部署、在线迁移、宕机迁移等能力，整体保证平台的可靠性。

平台使用分布式存储提供存储服务，如上图所示，虚拟机均运行于分布式存储池之上，且分布式存储池可跨多台物理机构建统一分布式存储资源池。

虚拟机的系统盘、镜像文件及挂载的硬盘均存储于统一分布式存储池中，每台节点均可通过分布式存储池中的虚拟机磁盘文件及配置信息注册一个相同的虚拟机进程，可作用于在线迁移或宕机迁移任务。

3.2.2 亲和策略

亲和反亲和策略，允许用户自定义虚拟机与其它虚拟机或宿主机之间的调度关系，使用户可以根据业务需求、性能要求或其它因素定义虚拟机调度逻辑，以满足特定的架构和运行要求。平台通过亲和反亲和策略提供了灵活的资源调度机制，用户可以根据实际需求动态调整调度策略，以适应不同业务负载和调度需求。

- 亲和策略用于将策略相关的虚拟机实例调度部署在同一物理主机上，可以最大程度地提高它们之间的数据传输和通信效率。
- 反亲和策略，是将虚拟机实例分散在不同物理主机上的调度策略，旨在降低单点故障的风险，提高系统的可靠性和容错性。

3.2.3 在线迁移

在线迁移（虚拟机热迁移）是计划内的迁移操作，即虚拟机不停机的情况下，在不同的物理机之间进行在线跨机迁移。首先是在目标物理机注册一个相同配置的虚拟机进程，然后进行虚拟机内存数据同步，最终快速切换业务到目标新虚拟

机。整个迁移切换过程非常短暂，几乎不影响或中断用户运行在虚拟机中的业务，适用于平台资源动态调整、物理机停机维护、优化服务器能源消耗等场景，进一步增强平台可靠性。

由于采用分布式统一存储，虚拟机在线迁移时只迁移【计算】的运行位置，不涉及【存储】（系统盘、镜像、虚拟硬盘）位置迁移。迁移时仅需通过统一存储内的源虚拟机配置文件在目的主机上注册一个相同配置且状态置为暂停的虚拟机进程，然后反复迁移源虚拟机的内存至目的虚拟机，待虚拟机内存同步一致后，关闭源虚拟机并激活目标虚拟机进程，最后进行网络切换并成功接管源虚拟机业务。

整个迁移任务仅在激活目标虚拟机及网络切换时业务处于短暂中断，由于激活和切换所用时间很短，少于 TCP 超时重传时间，因此源虚拟机业务几乎无感知。同时由于无需迁移虚拟机磁盘及镜像位置，虚拟机挂载的虚拟硬盘迁移后不受影响，可为用户提供无感知且携带存储数据的迁移服务。具体迁移过程如下：

(1) 注册目标虚拟机

- 调度系统使用统一分布式存储内的源虚拟机配置文件在目标主机上注册一个相同配置的虚拟机进程；
- 注册的虚拟机进程为不可提供服务的暂停【**paused**】状态，并通过监听一个 TCP 端口接收迁移数据；
- 注册目标虚拟机的阶段为瞬间完成，通常耗时为几毫秒，此时源虚拟机处于正常提供业务的状态。

(2) 迁移源虚拟机内存

- 在目标虚拟机注册完成的同时，调度系统会立即将源虚拟机的全量内存数据迁移至目标虚拟机；
- 为保证数据迁移的一致性，迁移过程中源虚拟机的内存更新也需要进行同步，因此调度系统通过多次迭代将源虚拟机产生的新内存数据迁移至目标端，耗时与物理机的网络带宽、性能及虚拟机的内存大小有关；

- 内存迁移时源虚拟机正常提供业务，待内存数据反复迭代迁移完成时立即暂停源虚拟机进程，避免产生新的内存数据；
- 源虚拟机进程暂停后，会再进行一次内存数据的同步，保证源端和目标端的数据一致性。

(3) 接管源虚拟机服务

- 完成内存同步的收尾工作，调度系统会关闭源虚拟机并激活目标虚拟机的进程，实现虚拟机平滑运行；
- 虚拟机从源主机迁移至目标主机，系统会将虚拟机的网络切换至目标主机，通过目标主机的虚拟交换机进行通信，成功接管源虚拟机服务。

整个迁移过程中，从源虚拟机暂停至目标虚拟机激活并完成网络切换为停机时间，由于激活虚拟机及网络切换时间非常短暂，通常小于几百毫秒，少于 TCP 超时重传时间，对大多数应用服务来说可忽略不计，因此虚拟机业务几乎不会感知到迁移停机。如智能调度图中的 VM6 默认运行在 Node1 上，管理员通过在线迁移功能手动将 VM6 迁移至 Node3 的流程如下：

- 调度系统收到迁移指令后，会立即使用 VM6 的配置文件在 Node3 节点上注册一个暂停状态的虚拟机进程；
- 立即迁移 VM6 的全量进程数据至 Node3 节点的 VM6'，并反复多次迁移更新内存数据；
- 调度系统暂停 Node1 上的 VM6 虚拟机，再次进行内存数据的迁移并关闭 VM6 虚拟机；
- 激活 Node3 节点上的 VM6 虚拟机进程，完成网络切换并接管 VM6 的业务服务及通信；
- 若 VM6 有挂载的虚拟硬盘，迁移成功后，不影响虚拟硬盘的挂载信息及配置，可正常读写虚拟硬盘。

3.2.4 离线迁移

离线迁移，是虚拟机在关机状态下，在不同的集群之间进行的离线跨集群迁移。离线迁移不涉及存储及数据迁移，更改虚拟机的集群配置，在虚拟机下次启动的时候，重新进入启动流程，重新分配物理机，在目标物理机注册一个相同配置的虚拟机进程，然后进行虚拟机内存数据同步，最终重新启动虚拟机到目标物理机。

离线迁移适用于平台资源动态调整、物理机停机维护、优化服务器能源消耗等场景，进一步增强平台可靠性。

由于采用分布式统一存储，虚拟机的系统盘及写进系统盘的数据均存储在底层分布式存储中，虚拟机离线迁移只迁移【计算】的运行位置，不涉及【存储】（系统盘、镜像、虚拟硬盘）位置迁移，仅需选择可迁移的集群，在下次虚拟机启动时保证网络通信即可。迁移过程如下：

- 更改虚拟机的 xml 配置，修改集群信息，清空物理机信息。
- 虚拟机在关机状态下，不进行任何操作，下一次启动重新走虚拟机启动流程，智能调度到空闲物理主机。
- 调度系统使用统一分布式存储内的源虚拟机配置文件在目标主机上注册一个相同配置的虚拟机进程。

3.2.5 宕机迁移

宕机迁移又称虚拟机高可用 (High Availability)，指平台底层物理机出现异常或故障而导致宕机时，调度系统会自动将其所承载的虚拟资源快速迁移到健康且负载正常的物理机，尽量保证业务的可用性。

整体宕机迁移不涉及存储及数据迁移，新虚拟机可快速在新物理机上运行，平均迁移时间为 90 秒左右，可能会影响或中断运行在虚拟机中的业务。

由于采用分布式统一存储，虚拟机的系统盘及写进系统盘的数据均存储在底层分布式存储中，虚拟机宕机迁移只迁移【计算】的运行位置，不涉及【存储】

(系统盘、镜像、虚拟硬盘) 位置迁移, 仅需在新物理机上重新启动虚拟机并保证网络通信即可。迁移机制说明如下:

- 平台调度管理系统会周期性检测除本物理机之外的所有物理机, 间隔时间为 10 秒;
- 当检测到某物理机出现网络中断, 则会重试 3 次;
- 如果重试 3 次之后都不成功, 就会将此物理机标记为不可达;
- 在所有物理机中, 有超过半数的物理机都标记某台物理机为不可达, 就会判定此物理机为宕机, 此物理机所有的虚拟机会在该集群 (Set) 内进行宕机迁移操作;
- 调度系统使用分布式存储内故障虚拟机的系统盘及数据重新在新物理机上启动虚拟机, 启动过程及状态流转与新建虚拟机一致, 平均启动时间为 30 秒左右;
- 虚拟机在新物理机上启动后, 会将虚拟机网络切换至新物理机;

整个迁移过程, 从检测到故障至迁移成功平均为 90 秒左右。虚拟机启动时间与源虚拟机的组件及配置有关, 如绑定虚拟硬盘; 同时由于虚拟机规格过大、底层物理资源不足、底层硬件故障等原因可能会导致宕机迁移失败, 通常建议尽量保证底层物理资源充足。

如智能调度图中的 Node2 节点故障, 智能调度系统自动将 VM3 和 VM4 分别迁移至 Node1 和 Node3 节点, 具体流程如下:

- 调度系统经过周期性监测及二层检测, 判断 Node2 节点故障, VM3/VM4 两台虚拟机不可用, 需要进行宕机迁移操作;
- 调度系统根据收集的集群节点信息, 使用分布式存储系统中 VM3 的系统盘及数据在 Node1 节点启动 VM3 虚拟机, 并在启动后将 VM3 的网络信息切换至 Node1;
- 使用分布式存储系统中 VM4 的系统盘及数据盘在 Node3 节点启动 VM4 虚拟机, 并在启动后将 VM4 的网络信息切换至 Node3;

宕机迁移的前提是集群中至少有 2 台以上的物理服务器,且在迁移过程中需保证健康节点的资源充足及网络连通性。通过宕机迁移技术,为业务系统提供高可用性,极大缩短由于各种主机物理故障或链路故障引起的中断时间。

3.3 存储虚拟化

云计算平台通过硬件辅助的虚拟化计算技术最大程度上提高资源利用率和业务运维管理的效率,整体降低 IT 基础设施的总拥有成本,并有效提高业务服务的可用性、可靠性及稳定性。在解决计算资源的同时,企业还需考虑适用于虚拟化计算平台的数据存储,包括存储的安全性、可靠性、可扩展性、易用性、性能及成本等。

虚拟化计算 KVM 平台可对接多种类型的存储系统,如本地磁盘、商业化 SAN 存储设备、NFS 及分布式存储系统,分别解决虚拟化计算在不同应用场景下的数据存储需求。

- 本地磁盘: 服务器上的本地磁盘,通常采用 RAID 条带化保证磁盘数据安全。性能高,扩展性差,虚拟化环境下迁移较为困难,适用于高性能且基本不考虑数据安全业务场景。
- 商业化存储: 即磁盘阵列,通常为软硬一体的单一存储,采用 RAID 保证数据安全。性能高,成本高,需配合共享文件系统进行虚拟化迁移,适用于 Oracle 数据库等大型应用数据存储场景。
- NFS 系统: 共享文件系统,性能较低,易用性较好,无法保证数据安全性,适用于多台虚拟机共享读写的场景
- 分布式存储系统: 软件定义存储,采用通用分布式存储系统的标准,将大量通用 x86 廉价服务器的磁盘资源聚合在一起,提供统一存储服务。通过多副本的方式保证数据安全,高可靠、高性能、高安全、易于扩展、易于迁移且成本较低,适用于虚拟化、云计算、大数据、企业办公及非结构化数据存储等存储场景。

每一种类型的存储系统,在不同的存储场景下均有优劣势,虚拟化计算平台

需根据业务特征选择适当的存储系统，用于提供存储虚拟化功能，在某些特定的业务模式下，可能需要同时提供多种存储系统，用于不同的应用服务。

在传统的存储结构中，客户端与单一入口点的集中式存储组件进行通信，可能会限制存储系统的性能和可伸缩性，同时可能带来单点故障。

UCloudStack 平台采用分布式存储系统作为虚拟化存储，用于对接 **KVM** 虚拟化计算及通用数据存储服务，消除集中式网关，使客户端直接与存储系统进行交互，并以多副本/纠删码、多级故障域、数据重均衡、故障数据重建等数据保护机制，确保数据安全性和可用性。

3.3.1 分布式存储

云平台基于 **Ceph** 分布式存储系统适配优化，为虚拟化计算平台提供一套纯软件定义、可部署于通用服务器的高性能、高可靠、高扩展、高安全、易管理且较低成本的虚拟化存储解决方案，同时具有极大可伸缩性。作为云平台的核心组成部分，为用户提供多种存储服务及 **PB** 级数据存储能力，适用于虚拟机、数据库等应用场景，满足关键业务的存储需求，保证业务高效稳定且可靠的运行。

分布式存储服务通过将大量通用服务器的磁盘存储资源融合在一起进行【池化】，构建一个无限可伸缩的统一分布式存储集群，实现对数据中心所有存储资源的统一管理及调度，向虚拟化计算层提供【块】存储接口，供云平台虚拟机或虚拟资源根据自身需求自由分配并使用存储资源池中的存储空间。

同时云平台虚拟化通过 **iSCSI** 及 **FC** 协议对接 **SAN** 商业存储设备，将商业存储作为虚拟化后端存储池，提供存储池管理及逻辑卷分配，可直接作为虚拟机的系统盘及数据盘进行使用，即只要支持 **iSCSI** 或 **FC** 协议的存储设备均可作为平台虚拟化的后端存储，适应多种应用场景；可利旧企业用户的集中存储设备，整体节省信息化转型的总拥有成本。

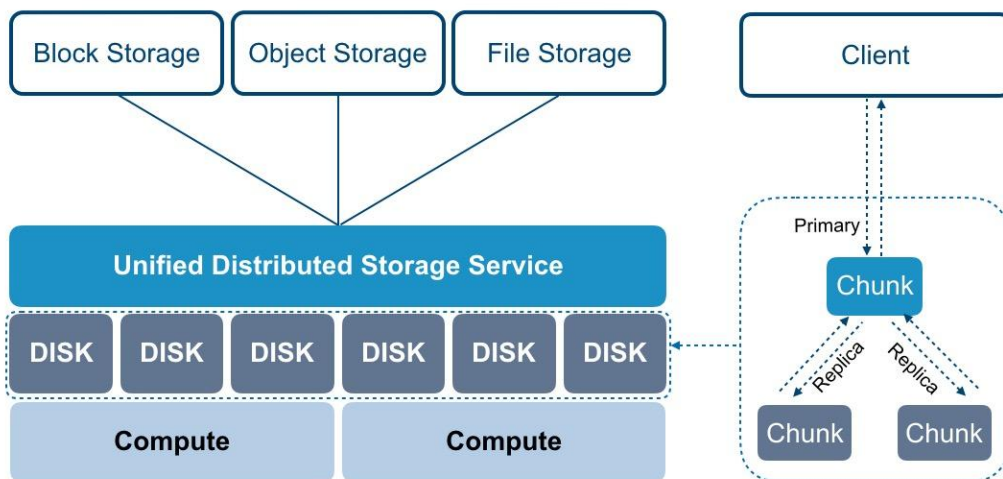
存储功能所见即所得，用户无需关注存储设备的类型和能力，即可在云平台快捷使用虚拟化存储服务，如虚拟磁盘挂载、扩容、增量快照、监控等，云平台用户像使用通用服务器的本地硬盘一样的方式使用虚拟磁盘，如格式化、安装操作系统、读写数据等。云平台管理和维护者可以全局统一配置并管理平台整体虚

虚拟化存储资源，如 QoS 限制、存储池扩容、存储规格及存储策略配置。

分布式存储系统可提供块存储、文件存储及对象存储服务，适用于多种数据存储的应用场景，同时可保证数据的安全性及集群服务的可靠性。在块存储的部署上，通常推荐使用同一类型的磁盘构建存储集群：

- 如超融合计算节点和独立存储节点自带 SSD 磁盘构建为高性能的存储集群；
- 超融计算节点和独立存储节点自带的 SATA/SAS 磁盘构建为普通性能存储集群。
- SATA/SAS 磁盘构建的普通性能存储节点和集群，在平台需通过 SSD 缓存加速的方式提升存储性能，缓存盘可采用 SSD 或 NVME 磁盘介质，平台要求缓存盘容量配比不高于 1:20，数量比不高于 1:5。

分布式存储系统将集群内的磁盘设备通过 OSD 内建不同的存储资源池，分别提供弹性块存储服务、对象存储及文件存储服务，其中块存储服务可供虚拟机直接挂载使用，在数据写入时通过三副本、写入确认机制及副本分布策略等措施，最大限度保障数据安全性和可用性。逻辑架构如下：



分布式存储系统是整个云平台架构不可或缺的核心组件，通过分布式存储集群体系结构提供基础存储资源，并支持在线水平扩容，同时融合智能存储集群、超大规模扩展、多副本与纠删码冗余策略、数据重均衡、故障数据重建、数据清

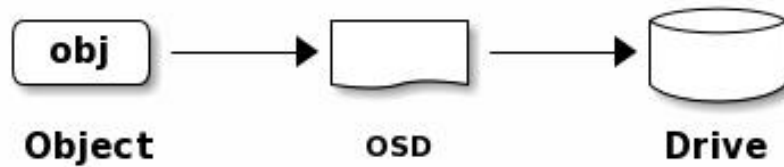
洗、自动精简配置及快照等技术，为虚拟化存储提供高性能、高可靠、高扩展、易管理及数据安全性保障，全方面提升存储虚拟化及云平台的服务质量。

3.3.2 智能存储集群

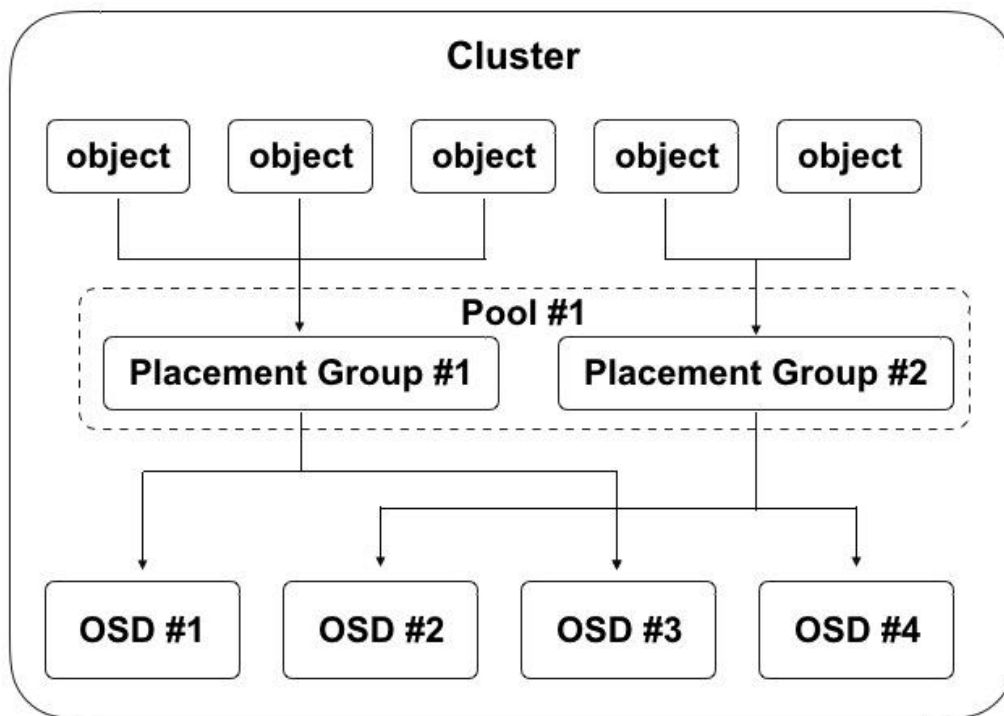
分布式存储集群可包含数千个存储节点，通常至少需要一个监视器和多个 OSD 守护进程才可正常运行及数据复制。分布式智能存储集群消除集中控制网关，使客户端直接和存储单元 OSD 守护进程交互，自动在各存储节点上创建数据副本确保数据安全性和可用性。其中包括的基础概念如下：

- **OSD**：通常一个 OSD 对应物理机一块磁盘、一个 RAID Group 或者一个物理存储设备，主要负责数据存储、处理数据复制、恢复、回填及数据重均衡，并负责向监视器报告检测信息。单集群至少需要两个 OSD，并在物理架构可划分为多个故障域（机房、机架、服务器），通过策略配置使多副本位于不同的故障域中。
- **监视器 Monitor**：实现存储集群的状态监控，负责维护存储集群的 Object、PG 及 OSD 间的映射关系图，为数据存储提供强一致性决策，同时为客户端提供数据存储的映射关系。
- **元数据服务 MDS**：实现文件存储服务时，元数据服务（MDS）管理文件元数据。
- **客户端**：部署在服务器上，实现数据切片，通过 CRUSH 算法定位对象位置，并进行对象数据的读写。通常包括块设备、对象存储及文件系统客户端，读/写操作由 OSD 守护进程处理。
- **CRUSH 算法**：用于保证数据均匀分布的伪随机算法，OSD 和客户端均使用 CRUSH 算法来按需计算对象的位置信息，为存储集群动态伸缩、重均衡和自修复功能提供支撑。

存储数据时，存储集群从客户端（块设备、对象存储、文件系统）接收数据，并将数据分片为存储池内的对象 Object，每个对象直接存储至 OSD 的裸存储设备上，由 OSD 进程处理裸设备上的读写操作。如下图所示：



客户端程序通过与 OSD 或监视器交互获取映射关系数据，在本地通过 CRUSH 算法计算得出对象存储位置后，直接与对应的 OSD 进行通信，完成数据读写操作。为实现分布式存储集群可自主、智能且自我修复的存取数据，智能存储集群通过 CURSH 算法、存储池 Pool、放置组 PG 及 OSD 等多种逻辑概念相互关联承载数据存储流程，逻辑架构图如下：



- 一个集群可逻辑上划分为多个 Pool，Pool 是一个命名空间，客户端存储数据时需指定一个 Pool；
- 一个 Pool 包含若干个逻辑 PG（Placement Group），可定义 Pool 内的 PG 数量和对象副本数量；
- PG 是对象和 OSD 的中间逻辑分层，写对象数据时，会根据 CRUSH

算法计算每个对象要存储的 PG;

- 一个物理文件会被切分为多个 Object, 每个 Object 会被映射到一个 PG, 一个 PG 包含多个 Object;
- 一个 PG 可映射到一组 OSD, 其中第一个 OSD 为主, 其它 OSD 为从, Object 会被均匀分发至一组 OSD 上进行存储;
- 承载相同 PG 的 OSD 间相互监控存活状态, 支持多个 PG 同时映射到一个 OSD。

在存储集群的机制中, 承载相同 PG 的主从 OSD 间需要彼此交换信息, 确保彼此的存活状态。客户端首次访问会首先从监视器获取映射关系的数据, 存储数据时会与 OSD 对比映射关系数据的版本。由上图示意图得知, 一个 OSD 可同时承载多个 PG, 在三副本机制下每个 PG 通常为 3 个 OSD。如上图所示, 数据寻址流程分为三个映射阶段:

1. 将用户要操作的文件映射为存储集群可处理的 Object, 即将文件按照对象大小进行分片处理;
2. 通过 CRUSH 算法将所有文件分片的 Object 映射到 PG;
3. 将 PG 映射到数据实际存储的 OSD 中, 最后客户端直接联系主 OSD 进行对象数据存储操作。

分布式存储客户端从监视器获取集群映射关系图, 并将对象写入到存储池。集群存储数据的逻辑主要取决于存储池的大小、副本数量、CRUSH 算法规则及 PG 数量等。

3.3.3 超大规模扩展

在传统集中式架构中, 中心集群组件作为客户端访问集群的单一入口, 这将严重影响集群的性能和可扩展性, 同时引入单点故障。

在 UCloudStack 分存储存储集群的设计中, 存储单元 OSD 和存储客户端能直接感知集群中的其它 OSD 及监视器信息, 允许存储客户端直接与存储单元

OSD 交互进行数据读写, 同时允许每个 OSD 与监视器及其它节点上的 OSD 直接交互进行数据读写, 这种机制使得 OSD 能够充分利用每个节点的 CPU/RAM, 将中心化的任务分摊到各个节点去完成, 支持超大规模集群扩展能力, 支持 EB 级的存储容量管理。

- OSD 直接服务于客户端, 存储客户端直接与 OSD 进行通信, 消除中心控制器及单点故障, 提升整体集群的性能及可扩展性。
- OSD 之间相互监测彼此的健康状态, 并主动更新状态给监视器, 使监视器可以轻量化部署和运行。
- OSD 使用 CRUSH 算法, 用于计算数据副本的位置, 包括数据重平衡。在多副本机制中, 客户端将对象写入主 OSD 中后, 主 OSD 通过自身的 CRUSH 映射图识别副本 OSD 并将对象复制到副本 OSD 中; 凭借执行数据副本复制的能力, OSD 进程可减轻存储客户端的负担, 同时确保高数据可用性和数据安全性。

为消除中心节点, 分布式存储客户端和 OSD 均使用 CRUSH 算法按需计算对象的位置信息, 避免对监视器上集群映射图的中心依赖, 让大部分数据管理任务可以在集群内的客户端和 OSD 上进行分布式处理, 提高平台的可伸缩性。

在存储集群扩容层面, 支持存储节点水平扩展、增量扩容及数据自动平衡性, 同时集群的整体性能随容量的增长呈正向增长, 继而保证存储系统的性能及高扩展性。

3.3.4 高可用和高可靠

为构建全平台高可用的分布式存储服务, 保证虚拟化计算及应用服务数据存储的可靠性, 分布式存储系统从多方面保证存储服务的稳健运行。

- **基础设施高可用**

存储集群不强行绑定硬件及品牌, 可采用通用服务器及网络设备, 支持存储集群异构。物理网络设备支持 10GE/25GE 底层存储堆叠网络架构, 同时服务器层面均采用双链路, 保证数据读写的 IO 性能及可用性。

- **存储监视器高可用**

集群监视器维护存储集群中 Object、PG 及 OSD 间的主映射图，包括集群成员、状态、变更、以及存储集群的整体健康状况等。OSD 和客户端均会通过监视器获取最新集群映射图，为保证平台服务的可用性，支持监视器高可用，当一个监视器因为延时或错误导致状态不一致时，存储系统会通过算法将集群内监视器状态达成一致。

- **存储接入负载均衡**

对象存储和文件存储接入网关支持负载均衡服务，保证对象存储和文件存储网关高可用，同时为存储网关提供流量负载分发，提升存储的整体性能。在负载均衡的接入机制下，读写 I/O 会均衡到集群中所有网关服务上，当其中一台网关服务器出现异常时，会自动剔除异常网关节点，屏蔽底层硬件故障，提升业务的可用性。

3.3.5 多副本冗余机制

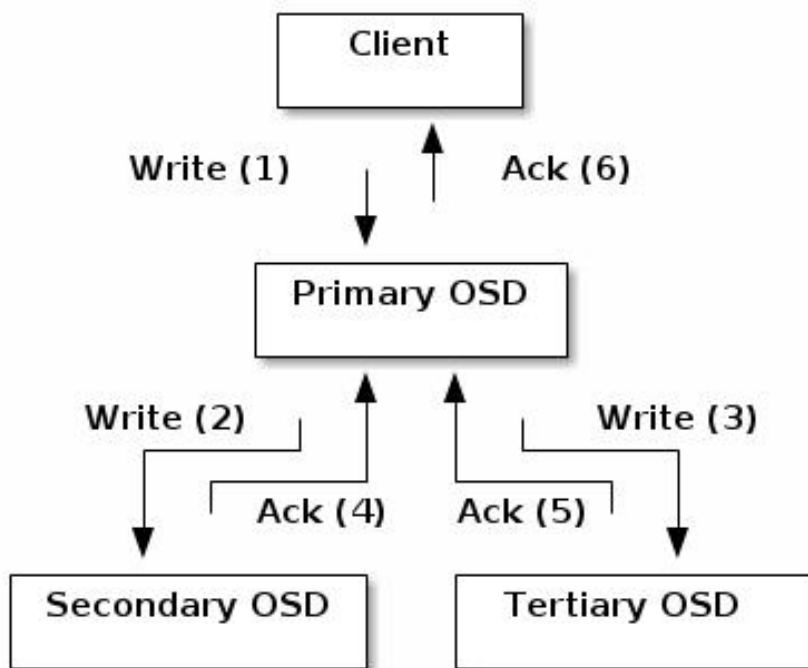
多副本机制是指将写入的数据保存多份的数据冗余技术，并由存储系统保证多副本数据的一致性。UCloudStack 分布式块存储系统默认采用多副本数据备份机制，写入数据时先向主副本写入数据，由主副本负责向其他副本同步数据，并将每一份数据的副本跨节点、跨机柜、跨数据中心分别存储于不同磁盘上，多维度保证数据安全。存储客户端在读取数据会优先读取主副本的数据，仅当主副本数据故障时，由其它副本提供数据的读取操作。

分布式存储系统通过多副本、写入确认机制及副本分布策略等措施，最大限度保障数据安全性和可用性。多副本机制存储数据，将自动屏蔽软硬件故障，当磁盘损坏和软件故障导致副本数据丢失，系统自动检测到并自动进行副本数据备份和同步，不会影响业务数据的存储和读写，保证数据安全性和可用性。本章节以三副本为例，具体描述多副本的工作机制：

(1) 三副本

用户通过客户端写入分布式存储的数据，会根据 Pool 设置的副本数量 3 写

入三份，并按照副本分布策略，分别存储于不同物理主机的磁盘上。分布式存储保证数据安全的副本数量至少为 2 份，以便存储集群可以在降级状态下运行，保证数据安全。



(2) 写入确认机制

如上图所示，三副本在写入过程中，只有三个写入过程全部被确认，才返回写入完成，确保数据写入的强一致性。

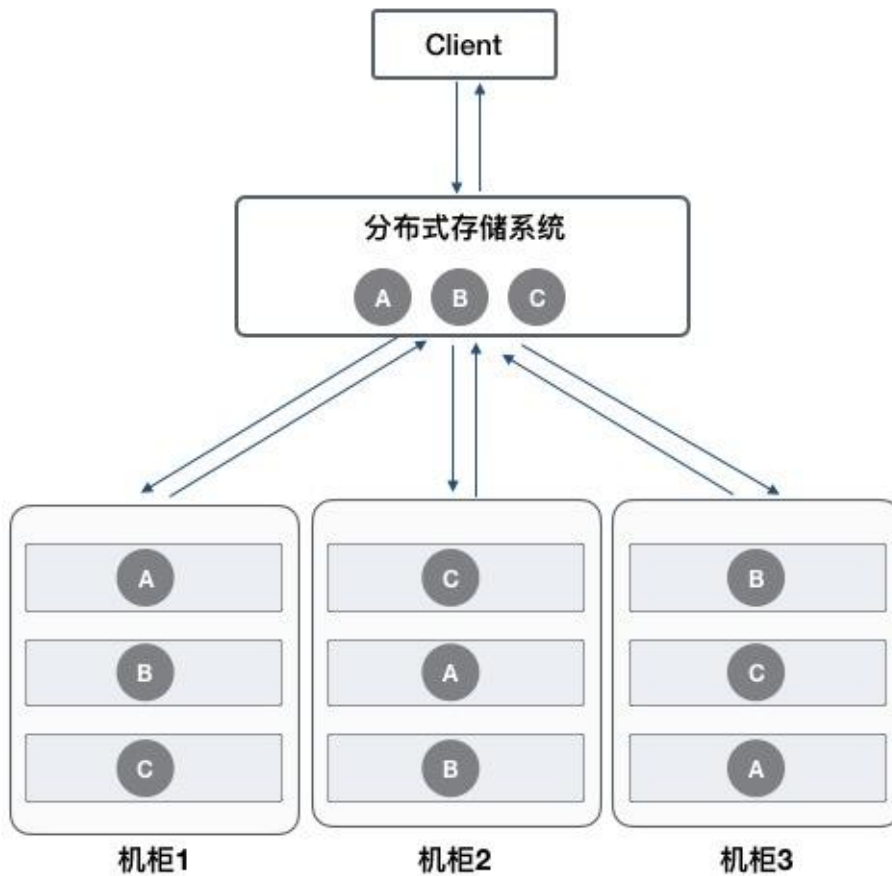
客户端将对象写入到目标 PG 的主 OSD 中，然后主 OSD 通过 GRUSH 映射关系图定位用于存储对象副本的第二个和第三个 OSD，并将对象数据复到 PG 所对应的两个从 OSD，当三个对象副本数据均写入完成，最后响应客户端确认对象写入成功。

(3) 副本分布策略

分布式存储支持副本数据落盘分布策略（多级故障域），使用 CRUSH 算法根据存储设备的权重值分配数据对象，尽量确保对象数据的均匀分布。平台通过定义存储桶类型，支持节点级、机柜级、数据中心级故障域，可将副本数据分布在不同主机、不同机柜及不同数据中心，避免因单主机、单机柜及单数据中心整

体故障造成数据丢失或不可用的故障，保证数据的可用性和安全性。

为保证存储数据的访问时延，通常建议最多将数据副本保存至不同的机柜，若将数据三副本保存至不同的机房，由于网络延时等原因，可能会影响云硬盘的 IO 性能。



如上图所示，客户端通过分布式存储系统写入 ABC 三个对象数据，根据 CRUSH 规则定义的故障域，需要将三个对象的副本分别存储于不同的机柜。以 A 对象为例，存储系统提前设置副本分布策略，尽量保证对象副本分布在不同机柜的服务器 OSD 中，即定义机柜和主机存储桶。当分布式存储系统计算出写入对象的 PG 及对应的 OSD 位置时，会优先将 A 写入到机柜 1 的服务器 OSD 中，同时通过主 OSD 复制副本 A' 至机柜 2 的服务器 OSD 中，复制 A'' 至机柜 3 的服务器 OSD 中，数据全部复制写入成功，即返回客户端对象 A 写入成功。



在存储节点无网络中断或磁盘故障等异常情况时，对象副本数据始终保持为 3 副本。仅当节点发生异常时，副本数量少于 3 时，存储系统会自动进行数据副本重建，以保证数据副本永久为三份，为虚拟化存储数据安全保驾护航。如上图第三个节点发生故障，导致数据 D1-D5 丢失并故障，存储系统会将对象数据的 PG 自动映射一个新的 OSD，并通过其它两个副本自动同步并重建出 D1'-D5'，以保证数据始终为三副本，保证数据安全。

3.3.6 数据重均衡

云平台分布式存储集群在写入数据时，会通过数据分片、CRUSH 映射关系、多副本分布策略尽量保证数据对象在存储池中的均衡。随着存储集群的长期运行及对平台的运维管理，可能会导致存储池内的数据失衡，如存储节点和磁盘扩容、存储部分数据被删除、磁盘和主机故障等。

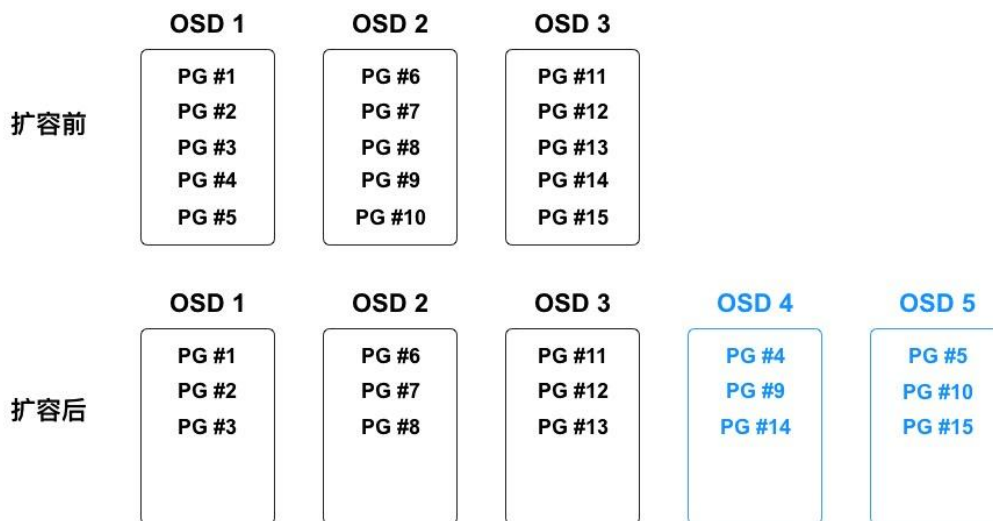
- 存储节点及磁盘扩容后，平台总存储容量增加，新增容量未承载数据存储，导致集群数据失衡；

- 用户删除虚拟机或云硬盘数据，导致集群内出现大量空闲空间；
- 磁盘和主机故障下线后，部分数据对象副本会重建至其它磁盘或主机，故障恢复后处于空闲状态。

为避免扩容及故障导致存储集群数据分布失衡，UCloudStack 分布式存储系统提供数据重均衡能力，在存储集群及磁盘数据发生变更后，通过 CRUSH 规则及时对数据的部分对象进行重新分发和均衡，使存储池中的对象数据尽量均衡，避免产生数据热点及资源浪费，提升存储系统的稳定性及资源利用率。

(1) 集群扩容重均衡

平台支持水平扩展存储节点或在线向存储节点中增加磁盘的方式扩容存储集群的容量，即分布式存储集群支持在运行时增加 OSD 进行存储池扩容。当集群容量达到阈值需要扩容时，可将新磁盘添加为集群的 OSD 并加入到集群的 CRUSH 运行图，平台会按照新 CRUSH 运行图重新均衡集群数据分布，将一些 PG 移入/移出多个 OSD 设备，使集群数据回到均衡状态。如下图所示：



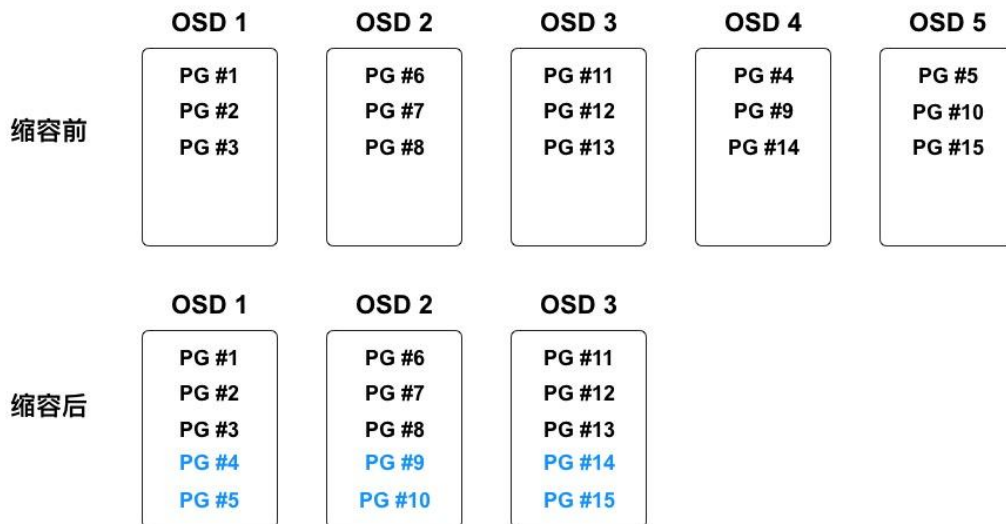
在数据均衡过程中，仅会将现有 OSD 中的部分 PG 迁移到新的 OSD 设备，不会迁移所有 PG，尽量让所有 OSD 均腾出部分容量空间，保证所有 OSD 的对象数据分布相对均衡。如上图中新增 OSD 4 和 OSD 5 后，有三个 PG (PG#4、PG#9、PG#14) 迁移到 OSD 4，三个 PG (PG#5、PG#10、PG#15) 迁移到 OSD 5，使五个 OSD 中映射的 PG 均为 3 个。为避免 PG 迁移导致集群性能整体降低，存储系统会提高用户读写请求的优先级，在系统空闲时间进行 PG 迁移

操作。

! 注意 PG 在迁移过程中，原 OSD 会继续提供服务，直到 PG 迁移完成才将数据对象写入新 OSD 设备。

(2) 集群容量缩减重均衡

存储集群在运行过程中可能需要缩减集群容量或替换硬件，平台支持在线删除 OSD 及节点下线，用于缩减集群容量或进入运维模式。当 OSD 被在集群中被删除时，存储系统会根据 CRUSH 运行图重新均衡集群数据分布，将被删除的 OSD 上的 PG 迁移至其它相对空闲的 OSD 设备上，使集群回到均衡状态。如下图所示：



在数据均衡过程中，仅会将删除 OSD 上的 PG 迁移至相对空闲的 OSD 设备，尽量保证所有 OSD 的对象数据分布相对均衡。如上图即将被删除的 OSD 4 和 OSD 5 上共映射 6 个 PG，删除后分别有 2 个 PG 会被迁移至剩余 3 个 OSD 中，使 3 个 OSD 中映射的 PG 均为 5 个。

(3) 故障数据重均衡

分布式存储在长期运行中会存在磁盘、节点的物理损坏、系统崩溃及网络中断等故障，均会中断节点的存储服务。存储集群提供容错方法来管理软硬件，PG 作为对象与 OSD 的中间逻辑层，可保证数据对象不会直接绑死到一个 OSD 设备，意味着集群可在“降级”模式下继续提供服务。详见数据故障重建。

注意 通过数据重均衡机制，可支持分布式存储集群平滑扩容，包括横向扩容和纵向扩容，即可以在线添加存储节点及存储磁盘。

3.3.7 数据故障重建

根据多副本和 EC 纠删码的保护机制，存储集群在把数据对象通过 CRUSH 写入到指定 OSD 后，OSD 会通过运行图计算副本或数据块的存储位置，并将数据副本或数据块写入到指定 OSD 设备中，通常数据对象会被分配至不同故障域中，保证数据安全性和可用性。

当磁盘损坏或节点故障时，即代表节点部分/全部 OSD 设备下线或无法为 PG 内对象提供存储服务，同时也表示有部分对象数据的副本数量不完整，如 3 副本可能因为磁盘损坏变为 2 副本。故障时对象数据的 PG 被置为“降级”模式继续提供存储服务，并开始进行数据副本重建操作，按照最新 CRUSH 运行图将故障节点或磁盘上的对象数据重映射到其它 OSD 设备上，即重新复制对象数据的副本至其它 OSD 设备，保证副本数量与存储池设置一致。

在 EC 纠删码策略下，节点或磁盘设备故障时会导致部分数据块或校验块丢失，如 4+2 的纠删码数据会丢失一个数据块或校验块，此时对象数据的 PG 被置为“降级”模式继续提供存储服务，并开始进行纠删数据的解码和恢复操作，按照最新 CRUSH 运行图将故障数据块或校验块数据重新恢复至其它健康的 OSD 设备上，保证对象数据的完整性和可用性。

故障数据重建时会遵循存储集群中配置故障域（主机级、机柜级及数据中心级），选择符合故障域定义的 OSD 作为故障数据重建的位置，让同一对象数据的多副本或 EC 数据间位置互斥，避免数据块均位于同一个故障域，保证数据安全性和可靠性。同时为提高故障数据的重建速度，多个故障数据重建任务的 I/O 会并发进行，实现故障数据的快速重建。

故障节点或磁盘恢复后，OSD 被重新加入至集群的 CRUSH 运行图，平台会按照新 CRUSH 运行图重新均衡集群数据分布，将一些 PG 移入/移出多个 OSD 设备，使集群数据回到均衡状态。为保证存储集群的运营性能，副本或纠删码 EC 数据恢复及迁移时，会限制恢复请求数、线程数、对象块尺寸，并提高用户

读写请求的优先级，保证集群可用性和运行性能。

3.3.8 数据清洗

分布式存储集群在长期运行及数据重平衡的过程中，可能会产生一些脏数据、缺陷文件及系统错误数据。如一块 OSD 磁盘损坏，集群在重均衡后重建数据至其它 OSD 设备，当故障 OSD 设备恢复后可能还存储着之前数据的副本，这些副本数据在集群重新平衡时需及时进行清洗。

分布式存储的 OSD 守护进程可进行 PG 内对象的清洗，即 OSD 会比较 PG 内不同 OSD 的各对象副本元数据，如果发现有脏数据、文件系统错误及磁盘坏扇区，会对其进行深度清洗，以确保数据的完整性。

3.3.9 自动精简配置

自动精简配置 (Thin Provisioning)，又称【超额申请】或【运行时空间】，是一种利用虚拟化技术减少物理存储部署的技术。通过自动精简配置，可以用较小的物理容量提供较大容量的虚拟存储空间，且真实的物理容量会随着数据量的增长及时扩展，可最大限度提升存储空间的利用率，并带来更大的投资回报。

UCloudStack 云平台分布式存储系统支持自动精简配置，在创建块存储服务时，分配逻辑虚拟容量呈现给用户，当用户向逻辑存储容量中写入数据时，按照存储容量分配策略从物理空间分配实际容量。如一个用户创建的云硬盘为 1TB 容量，存储系统会为用户分配并呈现 1TB 的逻辑卷，仅当用户在云硬盘中写入数据时，才会真正的分配物理磁盘容量。若用户在云硬盘上存储的数据为 100GB，则云硬盘仅使用存储池的 100GB 容量，剩余的 900GB 容量可以供其它用户使用。

云平台分布式存储系统支持对真实物理容量的监控，可提供真实物理已使用容量和逻辑的已分配容量。通常建议真实已使用容量超过总容量的 70% 时对存储集群进行扩容。自动精简配置类似 CPU 超分的概念，即可供租户创建使用的存储容量可大于物理总容量，自动按需分配物理存储空间给块存储设备，消除已分配但未使用的存储空间浪费。

通过自动精简配置，平台管理员无需对业务存储规模进行细化且准确预判，更不需提前为每个业务做精细的空间资源规划和准备，配合逻辑存储卷的容量分配策略，有效提升运维效率及存储空间的整体利用率。

3.3.10 存储功能简介

平台通过软件定义的分布式存储重新定义数据存储服务，基于通用服务器构建统一存储层，为应用提供块、对象及文件存储服务，同时提供多种数据接口，用户无需关注底层存储设备及架构，即可在云平台构建并使用存储服务，适用于虚拟化、云计算、大数据、物联网及企业应用等使用场景。

3.3.10.1 块存储服务

平台基于分布式存储系统为云平台租户提供块设备，即云硬盘服务，为计算虚拟化的虚拟机提供持久化存储空间的块设备。具有独立的生命周期，支持随意绑定/解绑至多个虚拟机使用，并能够在存储空间不足时对云硬盘进行扩容，基于网络分布式访问，为云主机提供高安全、高可靠、高性能及可扩展的数据磁盘。

根据底层物理存储设备的磁盘介质不同，云平台可为租户提供普通和高性能多种架构类型的云硬盘：

- 普通云硬盘使用 SATA/SAS 磁盘作为存储介质，并采用 SSD/NVME 作为缓存加速盘，保证普通云盘性能。
- 性能型云硬盘使用 SSD/NVME 磁盘作为存储介质。

云硬盘数据均通过 PG 映射及三副本机制进行存储，并在分布式存储系统的基础之上通过块存储系统接口为用户提供云硬盘资源及全生命周期管理。

支持组建多个存储集群，如 SATA 存储集群和 SSD 存储集群，并支持虚拟机跨集群挂载集群上的块存储服务。

- 分布式块存储服务直接通过物理网络进行挂载，无需通过 overlay 网络进行挂载和传输。
- 通过 libvirt 融合分布式存储 rbd 和 qemu，qemu 通过 librbd 操作分布式

存储。

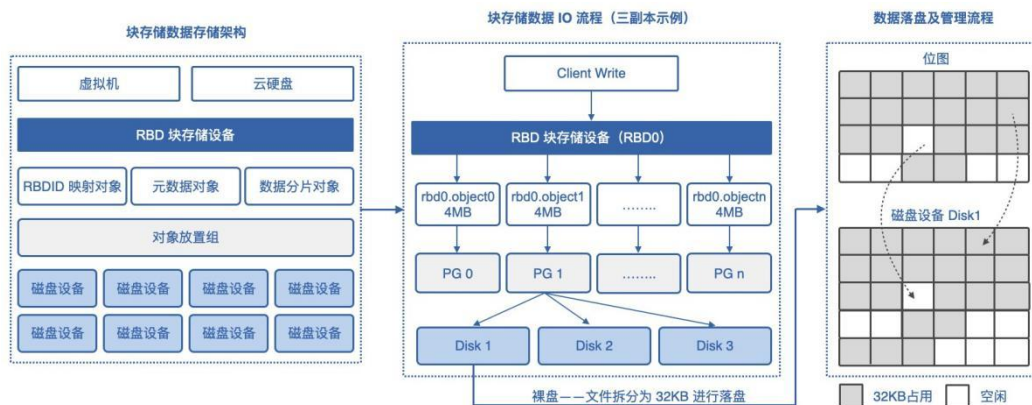
- 虚拟化进程与分布式存储进程通过本机&跨物理机内网进行通信。

不同存储集群间，对象数据的存储完全隔离。一个存储集群中不同块存储设备的存储策略完全隔离，互不干扰。分布式存储系统为虚拟机系统盘、镜像及云硬盘提供统一存储及管理，提高虚拟机与系统盘、云硬盘的数据传输效率，实现虚拟机快速创建及恢复，并支持系统盘和云硬盘的在线快速扩容和迁移。

在业务数据安全方面，云平台分布式存储支持磁盘快照能力，可降低因误操作、版本升级等导致的数据丢失风险，是平台保证数据安全的一个重要措施。支持对虚拟机的系统盘和数据盘进行手动或定时快照，在数据丢失或损坏时，可通过快照快速恢复本地业务的数据，实现业务分钟级恢复，包括数据库数据、应用数据及文件目录数据等。

3.3.10.2 数据存储机制

私有云的块存储服务采用分布式统一存储系统，由统一存储提供 RBD 接口为虚拟机提供系统盘、镜像及云硬盘服务。本节通过块存储数据存储架构、块存储数据 IO 流程及数据落盘管理流程对数据存储及删除机制进行说明。



(1) 块存储数据存储架构

虚拟机和云硬盘创建后，会在分布式存储系统中分别生成一个 RBD 块存储设备，即 KVM 引擎客户端读写数据的载体，同时针对一个块存储设备会生成 RBDID 映射对象、元数据对象及数据分片对象。

- **RBDID 映射对象**：指每个 RBD 块存储设备在存储系统中映射的 ID，作为全局唯一标识符，如 RBD0 对应的标识符为 RBD00001。
- **元数据对象**：指 RBD 块存储设备的元数据描述信息，包括块设备的创建时间、更新时间、属性、容量等。
- **数据分片对象**：RBD 块存储设备的数据分片对象文件，每个分片默认为 4MB，分片数量取决于 RBD 设备的大小，如 400MB 的云硬盘，分片数量即为 100 个对象文件。

所有的对象文件分别会通过算法计算对象的 PG 存储放置组，三副本模式下，一个放置组通常对应三个磁盘设备，即 RBDID 对象、元数据对象及数据分片的所有对象均会对应一个放置组，同时会将对象数据写入放置组对应的三个磁盘设备中。

(2) 块存储数据 IO 流程

虚拟机和云硬盘的虚拟化客户端在写数据至 RBD 块设备时，会自动对数据进行切片操作。如上图中 RBD 块设备为 RBD0，每个分片大小为 4MB，即会将写入的数据切分为 4M 大小的对象文件，同时包括元数据对象及 RBDID 映射对象。每个对象文件都有一个名字，即 rbd 设备+object+序号，如 rbd0.object0。

每一个 rbd.objectn 的对象文件通过放置组进行副本位置的分配，放置组通过 Cursh 算法定位出三个磁盘设备，作为对象文件的存储位置，即数据及元数据会首先进行对象文件的拆分，并根据放置组及磁盘设备的对应关系，分别存储至存储系统中的所有磁盘中。

(3) 数据落盘及管理流程

分布式存储系统使用裸盘进行磁盘管理及数据落盘操作，在进行对象文件 rbd.objectn 文件的存储和落盘时，会通过存储管理系统将每一个对象文件再次进行拆分进行存储，即通过位图的方式计算拆分后文件的在物理磁盘上的存储位置，将每个 4MB 对象文件拆分后存储至磁盘设备中，默认拆分大小为 32KB。

在写数据时根据位图计算出 32KB 文件在磁盘介质上的存储位置，同时在位

置上将占用的位置标示为 1（占用），未被占用的磁盘位置标示为 0（空闲）。

整体存储数据的过程，会将文件拆成 4MB 大小的对象文件，对应至不同的磁盘设备；同时在落盘时再次将 4MB 文件拆分成 32KB 大小的块存储至磁盘设备中。

(4) 删除数据机制

根据上面存储数据和落盘状况，存储在分布式存储系统中的文件被两次拆分成 32KB 的块文件，完全打散写入至整个存储集群的所有磁盘中，包括存储文件的元数据文件；在读取数据或找回数据时，需通过元数据计算数据是由哪些对象文件组成，同时需要结合磁盘位图计算对象数据中由哪些 32KB 的块数据组成，即其中一个 32KB 的数据是无法读取或恢复一个文件，必须将文件打散存储在存储集群中所有磁盘的 32KB 数据组合为一个对象文件，再通过元数据拼接对象文件，才可读取和恢复一个文件。

在平台上删除虚拟机、云硬盘或删除虚拟机中的文件时，存储系统会将文件的元数据进行删除，同时到磁盘管理的位图中将相关的 32KB 块置为 0（仅将块置为空闲，不真正清除数据），即用于恢复数据和读取数据的元数据被清除，同时 32KB 的块被置为空闲，可以被其它数据占用和写入。

- 若 32KB 块空间被其它数据占用后，则之前的数据会被新的数据覆盖；
- 若 32 KB 块空间未被其它数据占用，则可通过恢复软件找回的 32KB 数据，但 32KB 数据由于无元数据及位图，无法找出其它关联的 32KB 数据及相对应的对象文件，保证数据的安全性。

3.4 网络虚拟化

网络是虚拟化计算和分布式存储为云平台提供服务时不可或缺的核心部分，通常可采用硬件定义的 UnderLay 网络或软件定义的 OverLay 网络与虚拟化计算对接，为云平台提供多应用场景的网络及信息传输服务。

- **UnderLay 网络**：传统 IT 架构中硬件方式定义的单层物理网络，由物理设备和物理链路组成，即当前数据中心物理基础转发架构层——物理底

层承载网，包括一切现有的传统网络技术，负责互联互通。常见的物理设备有交换机、路由器、负载均衡、防火墙、IDS/IPS 等。

- **OverLayer 网络**: 虚拟网络，基于底层 UnderLayer 网络架构上叠加隧道技术构建的逻辑网络，实现网络资源虚拟化，以软件的方式在虚拟化平台上完整再现物理网络的功能。

OverLayer 网络的核心是隧道技术，只负责虚拟化计算资源的网络通信，具有独立的控制面和转发面（SDN 的核心理念）。对于连接到 OverLayer 的终端设备（例如服务器）来说，物理网络是透明的，从而可以实现承载网络和业务网络的分离。

作为云计算核心技术之一的虚拟化计算已被数据中心普遍应用，UnderLayer 网络和 OverLayer 网络均可作为虚拟化计算提供网络服务。随着业务规模的增长，虚拟机数量的快速增长和迁移已成为一个常态性业务，如果采用传统 IT 架构中硬件方式定义的 UnderLayer 网络，可能会给云平台带来一些问题：

- **网络隔离能力限制**

UnderLayer 主流的网络隔离技术是 VLAN，由于 IEEE 802.1Q 中定义的 VLAN ID 为 12 比特，仅能实现 4096 个 VLAN，无法满足大二层网络中标识大量租户或租户群的需求。同时由于 Vlan 技术会导致未知目的广播数据在整网泛滥，无节制消耗网络交换能力与带宽，仅适合小规模云计算虚拟化环境。

- **虚拟机迁移范围受网络架构限制**

为保证虚拟机热迁移，需保持虚拟机的 IP 地址和 MAC 地址保持不变，即要求业务网络为二层网络且需具备多路径的冗余备份和可靠性。传统物理网络 STP、设备虚拟化等技术部署反锁且不适合大规模网络，限制虚拟机的网络扩展性，通常仅适用于数据中心内部网络。

为大规模网络扩展的 TRILL/SPB/FabricPath/VPLS 等技术，虽可解决规模问题，但均需网络中的软硬件进行升级而支持此类新技术，增加云计算平台的部署成本。

- **虚拟机规模受网络规格限制**

在传统二层网络环境下，数据报文是通过查询 MAC 地址表进行二层转发，而网络设备 MAC 地址表的容量限制了虚拟机的数量。若选择适配较大容量 MAC 地址表的网络设备，则会提升网络建设成本。

- **部署缓慢且僵化**

虚拟化计算快速部署及灵活扩展特性上，均需网络提供强有力的支撑。传统网络中虚拟机部署业务及上线，均需对系统及网络设备进行繁琐的配置，甚至需要改变物理设备部署位置，降低业务发布效率的同时，难以快速响应新业务灵活部署及发布。

基于上述的问题和场景，可在 UnderLayer 网络基础架构上采用 OverLayer 网络解决方案，构建大二层虚拟网络，实现业务系统间网络隔离，并通过 NFV 实现网络中所需的各类网络功能和资源，按需灵活的调度资源，功能所见即所得，从而实现云计算平台中的网络虚拟化，满足虚拟化计算对网络的能力需求。

- **网络隔离能力**

OverLayer 网络虚拟化提供多种隧道隔离技术，如 VXLAN、GRE、NVGRE、STT 等，均引入类似 Vlan 的用户隔离标识，并对隔离标识进行极大扩展，如 VXLAN 支持 24 比特，可支持千万级以上的网络隔离标识。

- **隧道路由网络**

OverLayer 通过隧道技术，将二层以太报文封装在三层 IP 报文之上，通过路由的方式在网络中分发传输。路由网络本身无特殊网络结构限制，具备大规模扩展能力和高性能转发能力，同时路由三层网络会缩小二层广播域，大幅降低网络广播风暴的风险，具备很强的故障自愈能力和负载均衡能力。通过 OverLayer 技术的路由网络，虚拟机迁移不受网络架构限制，企业部署的现有网络便可用于支撑新的云计算业务。

- **大规模虚拟机规模**

虚拟机发出的数据包封装在 IP 数据包中，对网络只表现为封装后的网络参

数，即隧道端点的地址。因此极大的降低大二层网络（UnderLay）对 MAC 地址表容量的需求，可支撑大规模虚拟机场景。

● 快速灵活部署

基础网络不感知虚拟网络业务变化，OverLay 网络中应用部署的位置将不受限制，网络功能所见即所得，支持即插即用、自动配置下发及自动运行，可快速并灵活的部署业务，并支持业务在虚拟网络中进行迁移和变更。

网络虚拟化通过结合软件定义网络（SDN Software Defined Network）和网络功能虚拟化（NFV Network Function Virtualization）提供服务。SDN 是一种全新的网络架构，核心思想是通过标准化技术（如 openflow）将网络控制面和数据转发面进行分离，由控制器统一计算并下发流表，进而实现对网络流量集中化、灵活化、细粒度的控制。NFV 是指具体网络设备的虚拟化，使用通用服务器和软件实现并运行网络功能，如虚拟网卡、虚拟交换机、虚拟防火墙等，实现网络功能灵活配置、快速部署及定制编程能力。

SDN 和 NFV 是高度互补关系，各有侧重，分别从不同角度提供解决方案满足不同虚拟化场景的网络需求。SDN 通过将控制平台和数据转发面分离实现集中的网络控制，而 NFV 技术是通过软硬件分离，实现网络功能虚拟化。二者的关系如下：

- SDN 技术在流量路由方面所提供的灵活性，结合 NFV 的虚拟化架构，可更好地提升网络的效率，提高网络整体的敏捷性。
- NFV 不依赖 SDN，可在无 SDN 的情况下进行虚拟化部署，但 SDN 中控制和数据转发分离可改善 NFV 网络性能、易用性及可维护性，可实现 NFV 的快速部署及网络构建。

UCloudStack 通过 OVS+VXLAN 的 OverLay 网络及软件定义的 SDN 控制器，构建大二层虚拟网络，实现业务系统间网络隔离；并通过 NFV 实现网络中所需的各类网络功能和资源，用于对接 KVM 虚拟化计算服务，结合分布式网络架构为平台提供高可用、高性能且功能丰富的网络虚拟化能力及管理。

3.4.1 分布式网络

UCloudStack 基于 [OVS](#) (Open vSwitch) 组件, 通过 [VXLAN](#) 隧道封装技术实现隔离的虚拟网络, 并结合软件定义的 SDN 控制器, 为虚拟化计算平台提供一套纯软件定义、可运行于 x86 通用服务器的高性能、高可用、高可靠、易管理及较低成本的分布式网络解决方案。

作为云平台的核心组成部分, 为云平台所有虚拟资源提供全方位的网络转发及通信能力, 提供与物理网络相同的功能特性和性能保证, 且通过虚拟化提供网络资源分配、灵活部署及自动恢复等自动化运维能力, 满足网络功能的虚拟化的同时, 保证网络的可靠性。

分布式网络通过纯软件定义的方式在 x86 通用服务器上提供云计算所有网络功能, 无需网络硬件设备支撑 SDN 或 OverLay 特性, 即所有虚拟网功能特性及业务流量转发均由计算节点中的虚拟网络组件提供, 物理网络交换机设备仅承载平台物理节点间通信的数据转发, 因此物理网络仅需支持 Vlan、Trunk、LACP、IPV6、堆叠等特性即可。

平台虚拟化网络功能所见即所得, 用户无需关第底层设备类型及网络架构, 即可在云平台构建使用虚拟网络服务, 如虚拟私有网络 VPC、子网、弹性网卡、外网 IP、高可用 VIP、NAT 网关、负载均衡、防火墙及 VPN 等。

云平台用户像使用物理网络一样的方式使用虚拟网络, 如将虚拟机加入一个隔离网络、分配 IP 地址、配置外网 IP 访问外网, 或者通过 NAT 网关使多台虚拟机通过一个外网 IP 地址访问外网等。

云平台管理员可以像物理网络的管理员一样, 对全局网络资源进行统一配置、监控及管理, 如 IP 地址规划、外网 IP 地址池管理、网络设备资源管理及 QoS 配置等。

3.4.2 分布式架构

UCloudStack 采用 OVS 作为虚拟交换机, VXLAN 隧道作为 OverLay 网络隔离手段, 通过三层协议封装二层协议, 用于定义虚拟私有网络 VPC 及不同虚

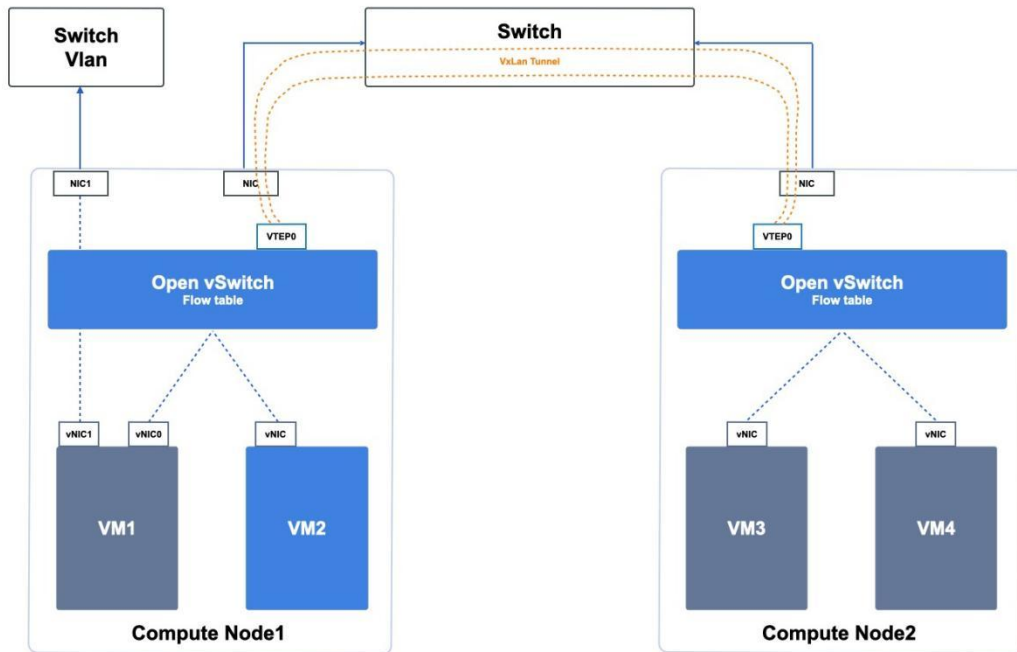
拟机 IP 地址之间数据包的封装和转发。

私有网络 (VPC——Virtual Private Cloud) 是一个属于用户的、逻辑隔离的二层网络广播域环境。在一个私有网络内, 用户可以构建并管理多个三层网络, 即子网 (Subnet), 包括网络拓扑、IP 网段、IP 地址、网关等虚拟资源作为租户虚拟机业务的网络通信载体。

私有网络 VPC 是虚拟化网络的核心, 为云平台虚拟机提供内网服务, 包括网络广播域、IP 网段、IP 地址等, 是所有 NVF 虚拟网络功能的基础。VPC 网络基于 VXLAN 协议, 不同网络之间二层完全隔离。

- 平台通过 VXLAN 定义并封装的 VPC 网络使用 VXLAN 头部 VNI (VXLAN Network Identifier, 3 字节) 字段作为全局唯一网络标识符, 即 VPCID (类似物理网络中的 VlanID)。
- 根据 VXLAN RFC 7348 的描述, VNI 字段包含一个由 3 个 8 位字节组成的数字封装器, 用于验证和标识 VXLAN 数据包的来源。
- VNI 字段长度为 24 位, 每一个 VXLAN 隧道号对应一个 VPC 网络, 即平台可支持 1600 (2 的 24 次方) 万个 VPC 网络。

平台 OverLay 网络数据面组件以分布式的方式部署于每个计算节点服务器, 结合自研的 SDN 虚拟网络控制器下发流表, 提供虚拟网络及 NFV 组件的实现、隔离、流表分发、数据封装及数据传输等功能, 实现可弹性、高安全、高可靠及绝对隔离的虚拟化网络。如下图所示:



OVS 是虚拟网络数据通路的核心路径, 每个计算节点开始提供服务时, SDN 控制器会自动下发属于当前节点的流表到虚拟交换机, 告知每个虚拟资源的网络应该如何通信。VXLAN 则提供虚拟网络跨物理主机访问时的数据封装及网络隧道。OVS 在所有计算节点上为分布式结构, SDN 控制器所属的管理控制模块为集群架构, 结合物理网络及链路的冗余架构, 整体提升虚拟网络的可用性。

如上图所示, 云平台 OverLay 网络组件分布式运行在所有计算节点, 即每个计算节点均部署 OVS+VXLAN 等组件:

- 虚拟网络流表分发服务为高可用架构, 仅做流表分发不透传生产网络传输。
- 分布式架构, 无集中网络转发节点, 所有生产网络仅在计算节点上传输, 无需通过管理服务或流表分发服务进行转发, 避免集中网络转发节点成为性能瓶颈。
- 每个计算节点仅承载运行在本机的虚拟机网络转发和传输, 单节点故障, 不影响其它节点的虚拟网络通信。
- 管理服务和流表分发服务故障, 不影响已部署好的虚拟资源运行及通信。

- 分布式存储直接通过物理网络进行挂载, 无需通过 OverLay 网络进行挂载和传输, 提升存储性能和可用性。
 - 通过 libvirt 融合分布式存储 rbd 和 qemu, qemu 通过 librbd 操作分布式存储;
 - 虚拟化进程与分布式存储进程通过本机&跨物理机内网进行通信。

云平台管理服务仅作为管理角色, 不承担网络组件部署及生产网络传输。分布式网络架构将业务数据传输分散至各个计算节点, 除业务逻辑等北向流量需要管理服务外, 所有虚拟化资源的业务实现等南向流量均分布在计算节点或存储节点上, 即平台业务扩展并不受管理节点数量限制。

3.4.3 通信机制

云平台通过 VXLAN 隧道及分布式网络架构提供完全隔离的虚拟网络, 通过定义虚拟私有网络为虚拟化计算提供与物理网络 VLAN 类似的网络功能, 具体通信原理如下:

- 相同 VPC 网络中, 同一个物理主机上的虚拟资源可直接通过 OVS 进行网络数据通信;
- 相同 VPC 网络中, 跨物理主机虚拟资源间的数据均通过 VXLAN 隧道封装送至物理网络上进行传输;
- 不同 VPC 间使用的隧道 ID 不同, 在网络上处于两个逻辑的路由平面, 使 VPC 间的网络天然隔离, 即不同 VPC 网络间虚拟资源不可通信;
- 不同 VPC 网络间资源内网不通, 可通过 VPC 互通的路由功能将不同 VPC 间的网络打通。

如分布式网络架构图所示, 假设 VM1、VM3、VM4 属于同一个 VPC 网络, VM2 属于独立的 VPC 网络。则虚拟机内网和外网通信机制如下:

(1) 内网络通信及限制

- VM1 和 VM2 属于不同 VPC 网络, 由于 VPC 网络的隔离性, VM1 无法

和 VM2 进行网络通信。


- VM3 和 VM4 属于同一个 VPC 网络且在同一台物理主机，可以直接通过 Open vSwitch 的流表进行通信。
 - VM1 和 VM3/VM4 属于同一个 VPC 网络，默认可进行网络通信，但由于 VM1 与其它两个 VM 不在同一台物理主机，则需要借助 VXLAN 隧道进行数据封装，并通过物理网络进行传输，具体过程如下：
 - VM1 发送数据至 Compute1 的 OVS，OVS 查询流表 VPC 信息，得知目的虚拟机 VM3 位于 Compute2 节点；
 - OVS 将数据包发送至 Compute1-vtep0 设备，对数据包进行 VXLAN 三层封装，并在两个节点 vtep0 间建立 VXLAN 隧道；
 - VXLAN 封装数据报文后，根据此 IP 包的目的地地址及路由信息，将报文通过 Compute1 的网络接口及物理网络 Switch 送达 Compute2 节点；
 - Compute2-vtep0 设备通过物理网络收到 VXLAN 报文后，对数据包进行 VXLAN 解封装；
 - 数据报文解封装后，通过 OVS 流表 VPC 信息，将报文转发至 VM3/VM4；
- VM2 和 VM3/VM4 属于不同 VPC 网络，由于 VPC 网络隔离性无法直接进行网络通信。

(2) 外网络通信及限制

如上图所示，本文假设 VM1 已绑定一个外网 IP 地址，通过外网 IP 访问外网为例进行描述，VM 通过 NAT 网关进行外网访问可参照【NAT 网关】。

- 用户通过云平台申请外网 IP 并绑定至虚拟机 VM1 时，云平台系统直接将外网 IP 及网关相关信息配置至虚拟机默认外网虚拟网卡 eth1 接口；

- 虚拟机需要访问外网时将数据包发送至 OVS, OVS 查询流表外网 IP 路由相关信息, 将数据包直接发送至 Compute1 外网网卡, 通过物理网络配置的路由或 Vlan 与互联网进行通信;
- 外网需要访问 VM1 时, 数据包会通过物理网络 Compute1 的外网网卡发送至 OVS, OVS 查询流表外网 IP 路由相关信息, 将数据包直接发送至 VM1 外网虚拟网卡 eth1 接口。

 **说明** 网络通信会受安全组规则的限制, 即在流量进出虚拟网卡时均会根据安全组规则再次进行筛选。

3.4.4 SDN 控制器

SDN 控制器即 OverLay 虚拟网络的控制面, 负责虚拟网络流表生成及下发管理, 通过控制器可自动对云平台的虚拟网络功能和组件进行参数及流表规则配置, 无需人工干预, 提升平台网络管控能力及运维效率。

每个计算节点开始提供服务时, 控制器会自动生成并下发属于当前节点的流表到节点虚拟交换机, 告知每个虚拟资源的网络应该如何通信。网络控制器和智能调度系统一样, 由【Schedule Manager 核心调度及管理模块】提供虚拟网络控制及管理, 支持集群架构, 结合物理网络及链路的冗余架构, 整体提升虚拟网络的可用性。

- 每一个地域仅需部署一套高可用 (主备模式) 的 Schedule Manager, 可在两台或多台节点上进行部署。
- 当部署网络控制模块所在的主计算节点服务器物理故障时, 部署调度模块的备计算节点将自动接替调度服务, 保证核心调度及流表控制服务的可用性。
- 网络控制器仅承载流表分发和控制服务, 不透传生产网络传输。

网络控制器高可用架构全部故障, 仅影响新建虚拟资源的流表下发及管理, 不影响已部署的虚拟资源运行及通信。由于网络控制器及云管理服务不承载生产网络转发和传输, 均由分布在所有计算节点的 OVS 组件进行所属虚拟机的网络

传输，所以计算节点在水平扩展的同时，承载生产流量的虚拟网络也同步进行了扩展，整体提升平台的可用性和可靠性。

3.4.5 网络功能简介

平台通过软件定义网络（SDN）对传统数据中心物理网络进行虚拟化，虚拟化网络功能所见即所得，用户无需关注底层设备类型及网络架构，即可在云平台构建使用虚拟网络服务，包括私有网络 VPC、网络隔离、弹性网卡、外网 IP、NAT 网关、负载均衡、防火墙（安全组）及 VPN 连接等网络服务，承载云平台上虚拟资源的网络通信及安全。

- **VPC 网络**：软件定义虚拟专有网络，用于租户间数据隔离。提供自定义 VPC 网络、子网规划及网络拓扑，可将虚拟机加入私有网络和子网，为虚拟机提供二层网络服务。
- **网络隔离能力**：由 VPC 提供的逻辑隔离的二层网络广播域环境，为云平台用户或子帐号提供网络隔离能力，不同 VPC 网络间网络完全隔离，不可进行通信。
- **弹性网卡**：一种可随时附加到虚拟机的弹性网络接口，支持绑定和解绑，可在多个虚拟机间灵活迁移，为虚拟机提供高可用集群搭建能力，同时可实现精细化网络管理及廉价故障转移方案。
- **外网 IP**：用于 VM、负载均衡及 NAT 网关等资源的互联网接入。支持多运营商线路接入并可调整外网 IP 的带宽上限。
- **高可用 VIP**：用于 VM 资源的内网虚拟 IP。归属于 VPC 内某个子网内的可漂移内网 IP，可将 HaVIP 与高可用服务结合。
- **NAT 网关**：企业级 VPC 网关，为云平台资源提供 SNAT 和 DNAT 代理，支持外网和物理网两种网络地址转换能力，并支持 VPC 级、子网级及实例级 SNAT 规则。
- **负载均衡**：基于 TCP/UDP/HTTP/HTTPS 协议将网络访问流量在多台虚拟机间自动分配的控制服务，类似于传统物理网络的硬件负载均衡器。

用于多台虚拟机间实现流量负载及高可用，提供内外网 4 层和 7 层监听及健康检查服务。

- **安全组**：虚拟防火墙，提供出入双方向流量访问控制规则，定义哪些网络或协议能访问资源，用于限制虚拟资源的网络访问流量，支持 TCP、UDP、ICMP 及多种应用协议，为云平台提供必要的安全保障。
- **VPN**：VPN 网关服务，提供可容灾的高可用 VPN 服务，配合 VPC、本地网关及公网服务三者共同使用。用户可选用多种加密及认证算法，保证隧道的可靠性。

云平台虚拟网络为用户提供丰富的网络功能，同时也提升平台可运营及可运维性，为云平台管理员提供网络规划配置、监控、QoS 限制及网络资源管理，让管理员可以像管理物理网络一样管理虚拟网络。

- **网络规划配置**：支持平台管理员对私有网络 IP 地址池、外网 IP 地址池、物理网络混合接入等进行对接及配置管理。
- **网络监控**：支持平台管理员监控平台所有 IP 及网络资源的使用率、流量及可用性，保证平台的可用性。
- **网络 QoS**：支持平台管理员配置网络 QoS，用于控制和限制内/外网络的带宽，用于避免用户间争取网络资源性能，保证所有虚拟网络组件的可用性。
- **网络资源管理**：支持平台管理员查看并管理平台所有的网络资源，包括物理网络资源和虚拟网络资源。

3.5 复用公有云

优刻得私有云基于 UCloud 公有云基础架构，复用内核及核心虚拟化组件，将公有云架构私有化部署，历经 10 年的大规模磨炼和验证，保证平台底层的稳定性。具有自主可控、稳定可靠、持续进化及开放兼容等特点，让企业轻松构建和公有云架构一致的云基础设施

同时复用公有云的核心模块和架构，专业的内核及测试团队及时跟进兼容性问题、安全问题、性能问题，确保平台时刻拥有一个健壮的底座；同时使用户拥有公有云一致用户体验的自服务平台。

根据权威机构评测，UCloudStack 私有云平台代码自研率达 96% 以上，为业界领先水平。在国产化替代的大背景下，自主可控显得尤为重要。非 OpenStack 二次开发，产品演进不受开源项目干扰和限制，紧跟客户需求，为用户解决实际问题。

信创版私有云除基础 IaaS 产品外，UCloudStack 信创云平台同时提供高可靠的负载均衡、NAT 网关、IPSecVPN、数据库服务、缓存服务、文件存储、对象存储等 PaaS 类产品服务。

此外，丰富的运营功能，组织架构管理、账号权限、服务目录、资源审批流程、报表统计，帮助 IT 管理中心提高管理效率，减少重复运维工作，实现服务化转型同时有效提升数据中心资源利用率。

3.6 一云多芯架构

UCloud 优刻得推出了自研的 UCloudStack 优钛私有云，并于 2020 年推出信创版，助力企业快速构建自主可控的云底座。UCloudStack 信创版私有云，提供通过信创互认证的 IaaS 和 PaaS 功能，兼容硬件、操作系统到上层应用的全信创生态。

- **异构资源统一管理**：一云多芯的异构资源管理，建设信创资源池的同时利旧和纳管传统架构资产，方便不同业务可按需选用合适的资源类型。
- **软硬件生态的快速适配**：专业的适配团队，快速完成硬件、操作系统、中间件等的适配，充分的测试保证兼容稳定性，满足用户不同场景需求。
- **多数据中心统一管理**：通过多地域统一管理能力实现多数据中心、多集群的统一调度和管理，提供统一运维和运营的一致体验，简化整体管理。

私有云平台从芯片到应用进行了全面的适配，CPU 已完成鲲鹏、飞腾、海光、龙芯、兆芯、申威主流国产芯片的适配，宿主机层面也已实现银河麒麟和统

信等国产操作系统的全面兼容。



私有云支持“一云多芯”，让企业在信创建设过程中得以平滑的过渡。支持纯软件交付、云融合一体机交付、云融合机柜交付、云数据中心交付等多种形式，可与用户既有基础设施结合，打造统一的资源池，最大限度降低用户 IT 成本。



针对特定的 CPU 架构进行调优适配，提升平台整体性能，如鲲鹏 ARM 架构服务器，根据虚拟化使能调优指南，充分发挥多核架构优势，释放极致算力。

3.7 混合云架构

优刻得混合云架构通过整合公有云、私有云、自建数据中心及客户托管数据中心等资源，通过托管云网络高速互联互通，结合同构私有云资源调度体系，统一运营运维管理，统一安全服务保障体系，提供兼具安全合规、经济可控及全面服务覆盖的一站式托管服务，业务自由灵活部署，IT 资源高效利用，达成收益风险平衡。



混合云架构集成同构的全栈私有云能力，涵盖虚拟化、SDN 网络、裸金属、分布式存储、数据库缓存、大数据平台及统一云管平台等服务能力，实现统一运营运维管理，统一安全服务保障，并与公有云内网互通，可随时进行业务和数据的云上云下互通迁移。

用户基于 UCloud 混合云 2.0 解决方案，可将稳态业务部署至 UCloud 自建云计算中心的同构私有云和大数据平台上，满足企业合规业务的云化、实现资源智能调度及有效利用；同时将需要资源弹性扩容及相关 PaaS 组件的敏态业务部署至 UCloud 公有云，满足业务弹性扩展需求的同时，降低 IT 投入成本。

(1) “东数西算”数据中心布局，打造稳定基础设施

UCloud 在全球部署了 32 个数据中心可用区，其中，两大自建云计算中心分别位于“东数西算”国家算力网络枢纽节点的内蒙乌兰察布和上海青浦。这些云计算中心共同构建了云网融合、安全稳定、智能敏捷、绿色低碳的数字信息基

基础设施，为客户提供整模块的深度定制混合云解决方案。



(2) 全栈资源调度服务，一站式整合云上云下资源

UCloud 混合云 2.0 基于与公有云同源同构的全线私有化产品，涵盖金翼物理机、优钛私有云 UCloudStack、统一分布式存储 UCloudStor、智能大数据平台 USDP 及多云管理平台 UCMP，可为用户提供资源调度及运营运维等一站式云基础设施服务，整合云上云下资源，全面提升业务能效。

- 公有云

UCloud 公有云拥有海量弹性资源，可提供计算、存储、网络、数据库、中间件、大数据、人工智能、云分发、视频服务、云安全及云运维等全面云服务。

- 金翼物理机 UXZONE

混合云基础架构平台金翼专区 UXZONE，把繁琐冗长、重资产的企业 IT 基础设施建设，打包成定制、便捷、轻资产的服务交付给用户，并整合私有云、存储及大数据等资源调度解决方案，实现一体化交付与持续运维服务。

- 私有云 UCloudStack

优钛私有云 UCloudStack 基于 UCloud 公有云基础架构，提供虚拟化、SDN、分布式存储、运营运维及云管平台等核心能力，同时可提供数据库、缓存、对象及文件存储等 PaaS 组件服务。3 节点即可构建生产环境且可平滑扩容，不强行绑定硬件及品牌，可同时部署 X86 和 ARM、MIPS 混合架构。借助私有云管理平台提供的资源智能调度和超分特性，可满足资源高可用和资源高复用的需求和场景，为用户降本增效。

- **统一分布式存储 UCloudStor**

UCloudStor 统一分布式存储，可以在统一的底层存储架构上，提供高性能块存储、文件存储和大容量对象存储服务，存储结构化、半结构化和非结构化等各类数据，使得存储在统一底座上的数据可以自由流通，提升数据利用率。此外，UCloudStor 支持替换 VMware 虚拟化下的 vSan 存储，满足客户对数据安全的需求。

- **智能大数据平台 USDP**

USDP 是 UCloud 自主研发的轻量级、智能化的大数据基础服务平台，提供一站式大数据集群管理和运维能力，全面兼容开源生态，支持部署 Hadoop、Hive、HBase、Spark、Flink、Presto、Atlas、Ranger 等众多开源组件，并进行智能运维管理。

USDP 可为混合云托管的大数据分析业务，构建海量数据的流批一体及数据湖仓一体架构，实现对数据质量、可用性、可靠性、安全性等多方面的数据治理；结合 UCloud 公有云及统一存储的能力，可帮助用户实现大数据分析和建模，同时可获取丰富的 PaaS 组件以及数据统一存储的能力。

- **混合云多云管理平台 UCMP**

UCMP 打造了包含裸金属管理、云上资源管理、智能告警治理、自动化运维为一体的全链路统一管理平台，支持对公有云、私有云等各类数据进行整合，为多云企业打通混合的协调管理能力，真正实现“一个界面统一纳管”、精细化运营。

(3) 贴身运营运维服务，打通云上云下运维体系

基于公有云强大技术专家团队，UCloud 还提供专属托管方案咨询设计、应用迁移及混合云交付等服务，并接入强大的公有云运维体系，90 秒快速响应确保问题闭环；同时可提供全链路的安全等保测评和服务交付。

(4) 显著降低成本投入，全面提升业务上云收益

当企业的云上业务增长到一定规模后，对云资源的投入会到达公有云的成本

临界点, 采用 UCloud 混合云 2.0 架构的自建云计算中心和私有云进行业务部署, 可合理分配公有云和私有云资源, 有效降低 IT 总投入成本。

- UCloud 自建云计算中心为低能耗机房, 得益于地理优势, 机房物业电力成本低廉, 硬件成本可节省 20%, 同时可为标准机柜租赁提供 30% 左右的成本优化空间。
- UCloud 提供的金翼托管物理机均为公有云大批量集采, 具备价格优势, 可在服务器设备上减少一次性投入成本。
- 稳态业务部署于混合云服务器或私有云平台上, 服务器利用率和生命周期运营可控, 折旧年限可达 6~8 年, 费用均摊。
- 结合私有云资源超高复用比运营配置, 可进一步缩减总投入的硬件资源, 综合对比公有云可有效降低 50% 左右的支出成本, 全面提升财务收益。

优刻得混合云架构基于低能耗自建云计算中心和集采优势的服务器为基础设施, 整合云上云下资源, 托管混合云业务。采用私有云超高复用比高效部署业务, 并提供全面的自动化、智能化运营运维管理服务, 且资源独享、业务可控, 硬件可达 6 年折旧, 综合对比公有云可有效降低 50% 左右的支出成本, 全面提升财务收益。

UCloud 已为多个客户提供混合云 2.0 解决方案, 助力客户业务在合规安全的要求下, 实现降本增效。某客户在公有云上有 500 台云主机, 以 8C16G 2TB 配置使用 6 年为例, 按照公有云售卖价格计算, 总成本约为 3600 万; 将业务迁移至 UCloud 混合云 2.0, 通过自建云计算中心, 采用同构全栈私有云的计算存储分离结构部署业务后, 在拥有 20% 的资源储备量的情况下, 6 年总成本约为 2500 万, 较公有云节省 30% 以上, 有效降低 IT 资源总投入成本。



UCloud 混合云 2.0 解决方案在安全合规的前提下，为用户提供弹性敏捷且兼具高性价比的混合云服务，实现云上云下统一管理，打造稳固的数字底座。

4 核心产品服务

4.1 基本概念

4.1.1 地域

地域 (Region) 指 UCloudStack 云平台物理数据中心的地理区域, 是云平台中的一个逻辑概念, 指资源部署的物理位置分类, 可对应机柜、机房或数据中心, 如上海、北京、杭州、主数据中心、备数据中心等。

通常一个数据中心对应一套 UCloudStack 云平台, 可支持部署多个计算和存储集群; 数据中心之间资源和网络完全物理隔离, 可通过一套管理平台管理遍布各地数据中心的私有云平台。

地域在平台也称为数据中心, 通常数据中心之间完全隔离以保证最大程度的稳定性和容错性。作为平台最大的资源定义, 一个地域即部署一套 UCloudStack 云平台。

平台默认内置一个地域, 管理服务通过本地数据中心云平台提供的 API 端点管理地域内计算、存储及网络资源。支持对数据中心内资源的生命周期管理, 包括计算集群、存储集群、外置存储、基础镜像及自制镜像等资源的查看和维护。

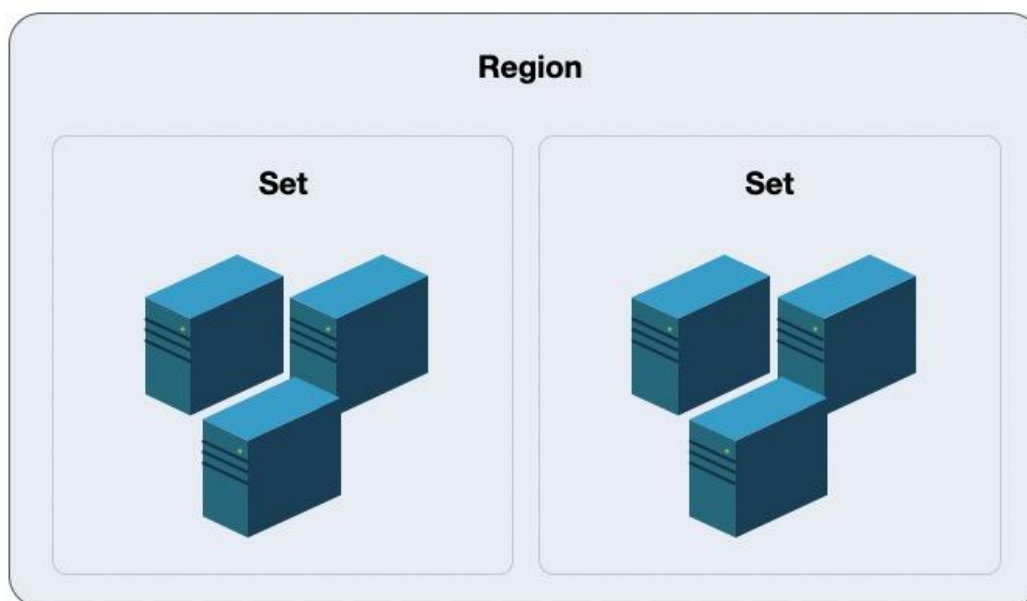
- 不同地域间完全物理隔离, 云平台资源创建后不能更换地域;
- 不同地域间网络完全隔离, 资源内网不能互通, 可通过公网或专线进行网络通信;
- 私有网络 VPC 和负载均衡服务支持相同地域部署。

4.1.2 集群

集群 (Set) 是 UCloudStack 物理资源的逻辑划分, 用于区分不同配置规格及不同存储类型的服务器节点, 如 X86 计算集群、ARM 计算集群、SSD 存储集群或商业存储集群等。

一个数据中心可支持部署多个计算和存储集群, 一个集群通常由一组配置、

用途相同的物理节点组成，且服务节点一般具有相同的 CPU/内存、磁盘类型及操作系统。



- 一个地域可包含多个集群，使用统一云管理平台进行集群管理和运营，云资源仅支持在单集群调度；
- 一个集群至少由 3 台服务器节点组成，集群内服务器须具有相同的 CPU/内存、磁盘类型及操作系统；
 - 服务器为计算&存储融合节点时，不同磁盘类型的节点划分为一个集群，如 SSD 计算节点集群；
 - 服务器为独立存储节点时，不同磁盘类型的节点划分为一个集群，如 SATA 存储节点集群；
- 通常一个集群的服务器建议接入同一组接入交换机，业务数据网络仅在集群内进行传输；
- 若采用独立存储节点，可将其划分为一个独立集群进行磁盘挂载。虚拟机仅支持跨集群挂载分布式块存储设备，用于数据存储。

云平台支持将 X86、ARM、GPU 等异构计算集群统一管理，并可统一管理 SSD、STAT、NVME 多种架构存储集群。

用户可将虚拟资源部署于不同的计算集群,并分别对虚拟资源挂载不同存储集群的块存储设备。同时云平台虚拟化可通过 ISCSI 协议对接 IPSAN 商业存储设备,为云平台虚拟机提供集群中高性能块存储服务,同时可利旧企业用户的集中存储设备,整体节省信息化转型的总拥有成本。

管理员控制台可对数据中心的计算集群、存储集群及外置存储集群进行便捷的管理和维护,同时平台可对集群进行权限控制,用于将部分物理资源独享给一个或部分租户使用,适用于专属私有云场景。

4.1.2.1 计算集群

计算集群是一组配置、用途相同的计算节点(物理机)组成,用于部署并承载平台上运行的虚拟计算资源。一个数据中心可部署多个不同类型的计算集群,如 X86 集群、ARM 集群、GPU 集群等,不同的集群可运行不同类型的虚拟机资源,如 GPU 集群可为租户提供 GPU 虚拟机,ARM 集群可为租户提供基于 ARM 或国产化 OS 的虚拟机。

为保证虚拟机高可用,平台基于集群维度提供虚拟化智能调度策略,包括打散部署、在线迁移、离线迁移及宕机迁移,即虚拟资源可在集群内的所有计算节点中进行调度、部署及迁移,提升业务的可用性。

- **打散部署:** 平台租户创建虚拟机时默认会将创建的虚拟机尽量打散部署于集群内的所有节点上,保障硬件或软件故障等异常情况下租户业务服务的可用性。
- **在线迁移:** 手动将一台虚拟机从集群的一个物理机迁移到另一台物理机,释放源物理机的资源,支持随机分配和指定物理节点两种模式。
- **离线迁移:** 手动将一台关机的虚拟机从一个集群迁移到另一个集群,调整集群的机器数量,支持指定集群迁移。
- **宕机迁移:** 运行虚拟机的物理机出现异常或故障导致宕机时,调度系统会自动将其所承载的虚拟资源快速迁移至集群内健康且负载正常的物理机,尽量保证业务的可用性。

基于在线迁移、离线迁移和宕机迁移的逻辑，通常在部署上推荐将相同 CPU 和内存配置的物理机节点规划为一个计算集群，避免因 CPU 架构或配置不一致，导致虚拟机迁移后异常或无法启动。

默认情况下平台会根据 CPU 平台架构设定集群名称，管理员可根据平台自身使用情况修改集群名称；同时支持管理员管理计算集群内的物理机和计算实例。

集群默认对所有租户开放权限，平台支持对计算集群进行权限控制，用于将部分物理计算资源独享给一个或部分租户使用，适用于专属私有云场景。修改集群权限后，集群仅可对指定的租户开放并使用，无权限的租户无法查看并使用受限的集群创建虚拟资源。

4.1.3 存储集群

存储集群为平台分布式块存储集群，通常由一组配置相同的存储节点（物理机）组成，用于部署并承载分布式存储资源。一个数据中心可部署多个不同类型的存储集群，如 SSD 集群、SATA 集群、容量型集群、性能型集群等，不同的集群可提供不同类型的云盘源，如 SSD 存储集群可为租户提供 SSD 类型的云硬盘。

平台通过分布式存储集群体系结构提供基础存储资源，并支持在线水平扩容，同时融合智能存储集群、多副本机制、数据重均衡、故障数据重建、数据清洗、自动精简配置、QOS 及快照等技术，为虚拟化存储提供高性能、高可靠、高扩展、易管理及数据安全性保障，全方面提升存储虚拟化及云平台的服务质量。

分布式存储集群默认支持 3 副本策略，写入数据时先向主副本写入数据，由主副本负责向其他副本同步数据，并将每一份数据的副本跨磁盘、跨服务器、跨机柜分别存储于不同磁盘上，多维度保证数据安全。在存储集群中存储服务器节点无网络中断或磁盘故障等异常情况时，副本数据始终保持为 3 副本，不区分主副本和备副本；当存储节点发生异常副本数量少于 3 时，存储系统会自动进行数据副本重建，以保证数据副本永久为三份，为虚拟化存储数据安全保驾护航。

默认情况下平台会根据存储架构设定集群名称，管理员可根据平台自身使用

情况修改集群名称；同时支持管理员管理存储集群。

集群默认对所有租户开放权限，平台支持对存储集群进行权限控制，用于将部分物理存储资源独享给一个或部分租户使用，适用于专属私有云场景。修改集群权限后，集群仅可对指定的租户开放并使用，无权限的租户无法查看并使用受限的集群创建云盘资源。

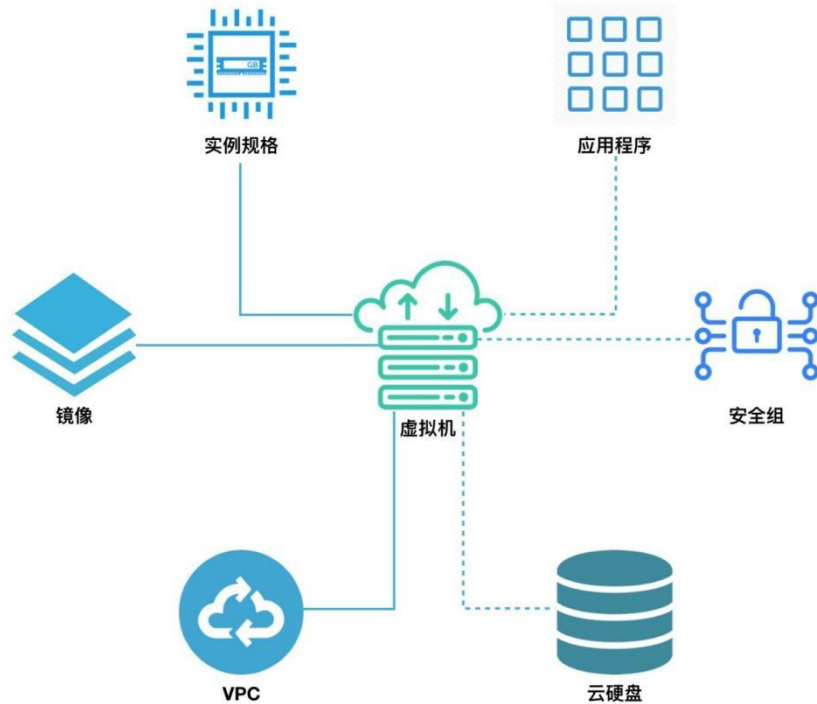
4.2 虚拟机

4.2.1 概述

虚拟机是 UCloudStack 云平台的核心服务，提供可随时扩展的计算能力服务，包括 CPU、内存、操作系统等基础的计算组件，并与网络、磁盘等服务结合提供完整的计算环境。通过与负载均衡等服务结合共同构建 IT 架构。

- 云平台通过 KVM(Kernel-based Virtual Machine)将物理服务器计算资源虚拟化，为虚拟机提供计算资源；
- 一台虚拟机的计算资源只能位于一台物理服务器上，当物理服务器负载较高或故障时，自动迁移至其它健康的物理服务器；
- 虚拟机计算能力通过虚拟 CPU(vCPU)和内存表示，存储能力通过云存储容量和性能体现；
- 虚拟机管理程序通过控制 vCPU、内存及磁盘的 QoS，用于支持虚拟机资源隔离，保证多台虚拟机在同一台物理服务器上互不影响。

虚拟机是云平台用户部署并运行应用服务的基础环境，与物理计算机的使用方式相同，提供创建、关机、断电、开机、重置密码、重装系统、升降级等完全生命周期功能；支持 Linux、Windows 等不同的操作系统，并可通过 VNC、Spice 等方式进行访问和管理，拥有虚拟机的完全控制权限。虚拟机运行涉及资源及关联关系如下：



如图所示，实例规格、镜像、VPC 网络是运行虚拟机必须指定的基础资源，即指定虚拟机的 CPU 内存、操作系统、虚拟网卡及 IP 信息。在虚拟机基础之上，可绑定云硬盘、弹性 IP、弹性网卡及安全组，为虚拟机提供数据盘、公网 IP、弹性网络及网络防火墙，保证虚拟机应用程序的数据存储和网络安全。

在虚拟化计算能力方面，平台提供 GPU 设备透传能力，支持用户在平台上创建并运行 GPU 虚拟机，让虚拟机拥有高性能计算和图形处理能力。支持透传的设备包括 NVIDIA 的 K80、P40、V100、2080、2080Ti、T4 及华为 Atlas300 等。

4.2.2 实例规格

实例规格是对虚拟机 CPU 内存的配置定义，为虚拟机提供计算能力。CPU 和内存是虚拟机的基础属性，需配合镜像、VPC 网络、云硬盘、安全组及密钥，提供一台完整能力的虚拟机。

- 默认提供 1C2G、2C4G、4C8G、8C16G、16C32G、32C64G 等实例规格；
- 支持自定义实例规格，提供多种 CPU 内存组合，以满足不同应用规模

和场景的负载要求；

- 支持升降级虚拟机 CPU 和内存配置，可通过更改实例规格进行调整；
- 实例规格通过关机后变更，需重新启动虚拟机生效；
- 实例规格与虚拟机生命周期一致，虚拟机被销毁时，实例规格即被释放。

平台支持自定义规格，创建虚拟机规格支持根据不同的集群创建不同的规格，即可为不同的机型创建不同的规格，租户创建虚拟机选择不同机型时，即可创建不同规格的虚拟机，适应不同集群硬件配置不一致的应用场景。可分别定义 CPU 和内存：

- **CPU 规格 (C)**：支持除 1 以外，以 2 为步长进行增加，如 1C、2C、4C、6C，最大可支持 240C。
- **内存规格 (G)**：支持除 1 以外，以 2 为步长进行增加，如 1G、2G、4G、6G，最大可支持 1024G。

创建出的规格即可被所有租户看到并使用，可根据业务需求在不同的集群中创建不同的规格。

4.2.3 生命周期管理

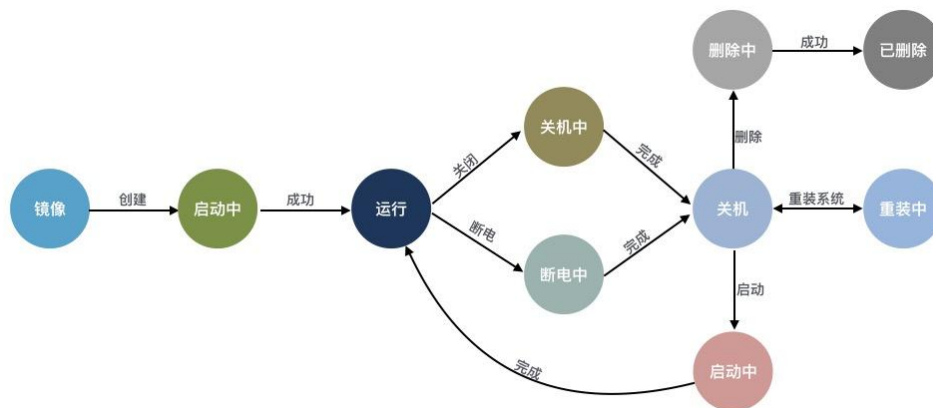
平台为虚拟机提供完整生命周期管理，用户可自助创建虚拟机，并对虚拟机进行关机、断电、开机、重置密码、重装系统、升降级配置、热升级、制作镜像、修改业务组、修改名称/备注、修改告警模板及删除等基本操作；同时支持与虚拟机相关联资源的绑定和解绑管理，包括弹性网卡、云硬盘、外网 IP 及安全组等。

- 关机是对虚拟机操作系统的正常关机，断电是将虚拟机强制关机；
- 重装系统即更换虚拟机镜像，Linux 仅支持更换 Linux 类型镜像，Windows 仅支持更换 Windows 类型镜像；
- 升降级配置是对虚拟机的规格配置进行升级或降级的变更操作；
- 热升级指在虚拟机开机 (running) 状态下，支持升级虚拟机的 CPU、

内存，不支持在线降级操作。

- 销毁虚拟机会自动删除实例规格、系统盘及默认虚拟网卡，同时会自动解绑相关联的虚拟资源；
- 一个虚拟机支持绑定多个云硬盘、弹性网卡、外网 IP 及安全组。

虚拟机完整生命周期包括启动中、运行、关机中、断电中、关机、启动中、重装中、删除中及已删除等资源状态，各状态流转如下图所示：



4.2.4 镜像服务

镜像（Image）是虚拟机实例运行环境的模板，通常包括操作系统、预装应用程序及相关配置等。虚拟机管理程序通过指定的镜像模板作为启动实例的系统盘，生命周期与虚拟机一致，虚拟机被销毁时，系统盘即被销毁。平台虚拟机镜像分为基础镜像和自制镜像。

4.2.4.1 基础镜像

基础镜像是由 UCloudStack 官方提供，包括多发行版 Centos、Ubuntu 及 Windows 等原生操作系统。


基础镜像默认所有租户均可使用，默认提供的镜像包括 Centos 6.5 64、Centos 7.4 64、Windows 2008r2 64、Windows 2012r2 64、Ubuntu 14.04 64、Ubuntu 16.04 64。

- 基础镜像均经过系统化测试，并定期更新维护，确保镜像安全稳定的运行和使用；
- 基础镜像为系统默认提供的镜像，仅支持查看及通过镜像运行虚拟机，不支持修改；

Linux 镜像默认系统盘为 40GB，Windows 镜像默认系统盘为 40GB，支持创建时进行系统盘容量扩容，也可以在虚拟机创建后做系统盘扩容操作（需要用户手动进入虚拟机内部进行文件系统扩容操作）。

支持管理将租户自制或导入的镜像复制为基础镜像，作为默认基础镜像共享给平台所有租户使用；同时支持管理员修改基础镜像的名称备注及删除基础镜像。

支持重装系统，即更换虚拟机镜像，Linux 虚拟机仅支持更换 Centos 和 Ubuntu 操作系统，Windows 虚拟机仅支持更换 Windows 其它版本的操作系统。

 **注意** Windows 操作系统镜像为微软官方提供，需自行购买 License 激活。

4.2.4.2 自制镜像

自制镜像由租户或管理员通过虚拟机自行制作或自定义导入已有的自有镜像，可用于创建虚拟机，除平台管理员外，平台的租户自身也有权限查看和管理。

- 支持管理员和租户制作、导入和导出自定义镜像；同时管理员可导出镜像仓库中的所有自制镜像。
- 支持管理员和租户通过自制镜像创建虚拟机、删除自制镜像、修改自制镜像名称。
- 支持平台管理员和租户自有 ISO 镜像导入，自制镜像以及 ISO 镜像归属于云平台租户。

ISO 镜像是一种将光盘或 DVD 中的数据以文件的形式保存在计算机硬盘上的方法。当需要使用一个全新的操作系统时，可以选择使用包含操作系统介质的

ISO 镜像，直接使用 ISO 介质引导并安装到虚拟机中。通过这种方式，可以快速部署一个全新的云主机。在部署完成后通常会进行以下操作，用于进一步提升将来的部署效率。

- 将安装完操作系统的虚拟机，修改符合企业标准的系统参数，并转化为一个标准的云主机镜像，方便下一次更加高效的创建一个符合企业信息化标准的全新云主机实例。
- 进一步部署相关的业务系统，转化为一个包含业务系统的云主机镜像，以提升下一次部署相同业务时的效率，同时也保障业务系统的符合企业信息化标准规范。

自制镜像和 ISO 镜像可用于创建虚拟机，并支持用户下载虚拟机镜像到本地，同时镜像管理支持查看镜像、修改名称和备注、从镜像创建虚拟机、导入镜像、下载镜像及删除镜像等生命周期管理。

为方便平台镜像模板文件的共享，平台支持管理员将一个自制镜像复制为一个基础镜像，使一个租户的自制镜像共享给所有租户使用，适用于运维部门制作模板镜像的场景，如自制镜像操作系统的漏洞修复或升级后，制作一个自制镜像并复制为基础镜像，使所有租户可使用新的镜像文件升级虚拟机系统。

4.2.4.3 镜像存储

基础镜像和用户自制镜像默认均存储于分布式存储系统，保证性能的同时通过三副本保证数据安全。

- 镜像支持 QCOW2 格式，可将 RAW、VMDK 等格式镜像转换为 QCOW2 格式文件，用于 V2V 迁移场景。
- 所有镜像均存储于分布式存储系统，即镜像文件会分布在底层计算存储超融合节点磁盘上。
- 若为独立存储节点，则分布存储于独立存储节点的所有磁盘上。
- 地域的镜像只能创建本地域的虚拟机，不支持跨地域镜像创建虚拟机。

4.2.5 虚拟机存储

虚拟机的系统盘和数据盘存储支持块存储和商业存储作为后端存储,并统一池化为云硬盘,用户可以像使用物理机硬盘一样的格式化并建立文件系统来使用云硬盘,

针对虚拟机的云硬盘和数据安全,平台支持云硬盘加密特性,使用 LUKS 加密规范来对磁盘全盘加密,保护用户的数据不被未经授权的访问者获取,甚至在磁盘丢失或被盗的情况下也可以保证数据的机密性。

(1) 云硬盘

一种基于分布式存储系统为虚拟机和数据库服务提供持久化存储空间的块设备。云硬盘基于网络分布式访问,为虚拟机提供高安全、高可靠、高性能及可扩展的数据磁盘。可用于虚拟机的系统盘和数据盘。

- 支持对云硬盘类型的系统盘进行扩容、快照及加密。
- 支持对云硬盘类型的数据盘进行绑定、解绑、扩容、快照、续费及加密。
- 支持通过云硬盘创建并启动虚拟机。
- 支持将云硬盘设置为共享盘,多个虚拟机同时进行挂载使用。
- 单个虚拟硬盘可支持的 32TB 存储容量,同时单个虚拟机磁盘可支持 200 个快照。
- X86 架构虚拟机可支持挂载 25 块云硬盘,ARM 和龙芯架构虚拟机可支持挂载 3 块云硬盘。

支持对云硬盘本身进行全生命周期管理,包括云盘创建、查看、绑定、解绑、扩容、克隆、删除、续费、快照、设为共享盘等,详见【云硬盘】章节描述。

(2) 商业存储 SAN

云平台虚拟化支持对接商业存储设备,如 IPSAN、FCSAN 等存储阵列,为云平台虚拟机提供集群中高性能块存储服务,同时可利用旧企业用户的集中存储设备,整体节省信息化转型的总拥有成本。

- 支持将 SAN 存储类型的 LUN 磁盘作为虚拟机的系统盘和数据盘。
- 支持 SAN 存储类型的磁盘创建并启动虚拟机，并支持对磁盘进行加密。
- 支持将 SAN 存储类型的磁盘设置为共享盘，多个虚拟机同时进行挂载使用。
- 支持对 SAN 存储类型的磁盘进行绑定和解绑操作，同时支持管理员将 LUN 分配给不同的租户进行使用。

平台仅将商业存储的 LUN 作为存储卷进行使用，不对存储卷本身进行管理，如 LUN 的创建、映射、扩容、快照、备份、回滚、克隆等。

(3) 共享云硬盘

共享云硬盘是一种数据块级存储设备，能够同时支持多个云服务器并发读写访问。这种存储设备具有多挂载点、高可靠性等特点，适用于需要支持集群和高可用性（HA）能力的关键企业应用场景，多个云服务器可以同时访问一个共享云硬盘。

支持将云硬盘、SAN 存储 LUN 设备设备为共享盘，并作为虚拟机的数据盘，使多个虚拟机同时对共享盘进行数据读写操作。同时支持对共享盘进行创建、绑定、解绑、扩容、克隆、续费及删除等操作。

4.2.6 存储热迁移

存储热迁移是平台提供的在不中断虚拟机运行的情况下可以动态更换虚拟机存储的能力，支持对 Intel/AMD x86 架构虚拟机的任意磁盘进行动态更换。该能力允许在虚拟机运行的同时进行存储迁移变更，可以在不影响业务连续性的情况下执行存储维护、变更或优化操作，以提高系统的可用性。

4.2.6.1 应用场景

存储热迁移的能力不仅限于系统盘，同时支持对虚拟机中任意数据盘的动态更换，通常应用于如下场景：

- **存储空间均衡**

通过创建新的存储集群对平台进行存储扩容后,通常面临着新旧存储集群空间使用不均衡的问题。存储热迁移的能力使得在不中断虚拟机业务的前提下,可以灵活地调整存储资源,实现存储空间的均衡分配,从而有效缓解原有存储集群的存储压力,提升整体系统的性能和稳定性。

- **更换高性能存储**


随着业务对虚拟机磁盘性能需求的提升,可能需要将虚拟机磁盘更换为性能更高的存储集群。通过存储热迁移的能力,可以在不中断虚拟机服务的情况下,将虚拟机磁盘迁移到高性能存储上,满足业务对性能的更高要求,提升应用的响应速度和整体处理能力。

- **存储设备下线**

在存储设备需要进行维护或下线的情况下,存储热迁移能够实现虚拟机磁盘的平稳迁移,将数据从即将下线的存储设备迁移到长期稳定可用的存储集群中。这种无缝迁移的过程保证了虚拟机业务的连续性,同时让管理员可以灵活地进行存储设备的维护工作,确保整个环境的可维护性和稳定性。

4.2.6.2 迁移模式

平台支持通过内置分布式存储为虚拟机提供虚拟硬盘,存储热迁移能力支持在不同存储集群之间进行热迁移。迁移时选择目标存储集群即可,平台会自动在目标集群创建新存储并在迁移完毕后自动删除旧存储。

 **注意** 共享盘允许多个虚拟机同时挂载使用,不支持针对共享盘进行存储热迁移。

4.2.6.3 迁移过程

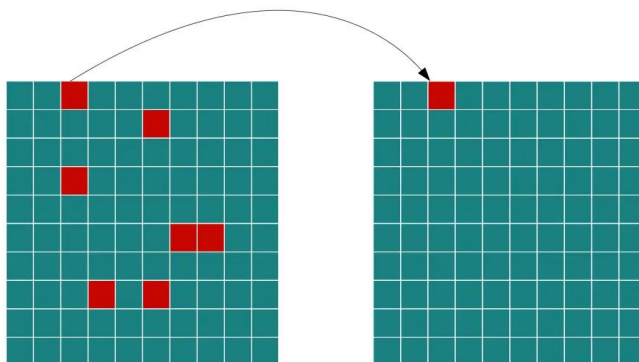
运行中的虚拟机在不断变更存储中的数据,为确保在迁移完成时目标存储包含迁移过程中的所有变更数据,平台通过迁移准备、数据迁移、迁移收尾三个阶段实现完整的存储热迁移。

(1) 迁移准备

将磁盘按 **block** 为单位组织成一个数组，块级别的组织结构使得系统能够更加细致地管理和操作数据，便于系统实现对数据的分批迁移。同时开启脏块记录机制用于数据变更追踪，系统在进行数据迁移时可以有选择性地迁移仅包含脏块的部分，降低了迁移的成本和复杂性。

(2) 数据迁移

首先进行磁盘的全量数据迁移，依次将每个 **block** 迁移到目标存储。通过将整个磁盘的数据迁移到目标存储，建立一个基准状态，用于确保后续增量迁移的准确性和完整性。



然后通过多次迭代，将迁移过程中虚拟机产生的新数据迁移到目标存储。迭代过程逐渐收敛的特性使得迁移操作更具渐进性和可控性。**Qemu** 将边迁移边记录剩下的脏数据大小，随着迭代的进行，剩余的脏数据不断减小，最终收敛到一个较小的值。迭代收敛的过程确保了最终一致性，当脏数据逐渐趋近零时，系统达到了一种稳定状态，**Qemu** 进程就会暂停，从而避免产生新的脏数据，以便进行迁移收尾工作。

(3) 迁移收尾

在虚拟机暂停之后，整个迁移过程进入第三阶段的收尾工作，这一阶段的主要任务是确保迁移数据的完整性和一致性。**Qemu** 进程会将剩余的磁盘脏数据一次性同步到目标端，完成时虚拟机新旧存储的数据将会一致。

4.2.7 虚拟机网络

4.2.7.1 虚拟网卡

虚拟网卡 (Virtual NIC) 是虚拟机与外部通信的虚拟网络设备，创建虚拟机时随 VPC 网络默认创建的虚拟网卡。虚拟网卡与虚拟机的生命周期一致，无法进行分离，虚拟机被销毁时，虚拟网卡即被销毁。有关 VPC 网络详见 VPC 网络章节。

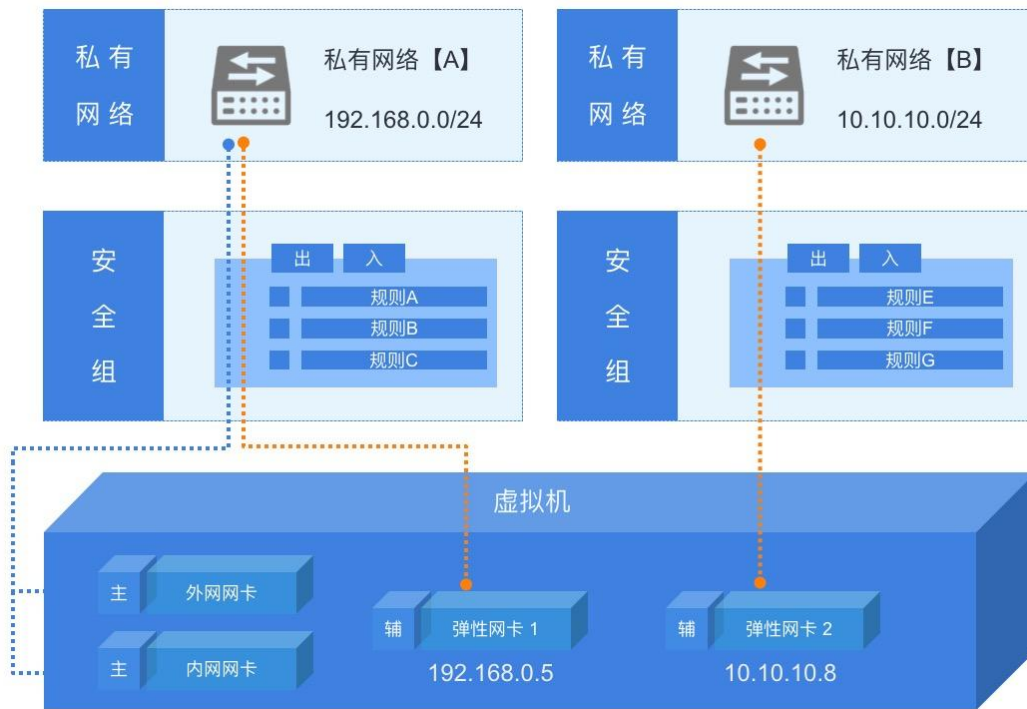
虚拟网卡基于 Virtio 实现，QEMU 通过 API 对外提供一组 Tun/Tap 模拟设备，将虚拟机的网络桥接至宿主机网卡，通过 OVS 与其它虚拟网络进行通信。

- 每个虚拟机默认会生成 2 块虚拟网卡，分别承载虚拟机内外网通信。
- 在虚拟机启动时，根据选择的 VPC 子网自动发起 DHCP 请求以获取内网 IP 地址，并将网络信息配置在一块虚拟网卡上，为虚拟机提供内网访问。
- 虚拟机启动后，可申请公网 IP（外网 IP）绑定至虚拟机，提供互联网访问服务。
 - 绑定的外网 IP 会自动将公网 IP 信息配置在另一块虚拟网卡上，为虚拟机提供外网访问；
 - 一个虚拟机支持绑定 50 个外网 IPv4 和 10 个 IPv6 地址。
- 不支持修改虚拟网卡的 IP 地址，手动修改的 IP 地址将无法生效。
- 每块虚拟网卡支持绑定一个安全组，提供网卡级别安全控制。
- 支持虚拟网卡 QoS 控制，提供自定义设置虚拟网卡的出/入口带宽。

平台默认提供 2 块虚拟网卡，若业务有 2 块以上网卡需求可通过绑定弹性网卡，为虚拟机提供多网络服务。

4.2.7.2 弹性网卡

弹性网卡 (Elastic Network Interface, ENI) 是一种可随时附加到虚拟机的弹性网络接口, 支持绑定和解绑, 可在多个虚拟机间灵活迁移, 为虚拟机提供高可用集群搭建能力, 同时可实现精细化网络管理及廉价故障转移方案。



弹性网卡与虚拟机自带的默认网卡均为为虚拟机提供网络传输的虚拟网络设备, 分为内网网卡和外网网卡两种类型, 同时均会从所属网络中分配 IP 地址、网关、子网掩码及路由相关网络信息。

- 内网类型的弹性网卡所属网络为 VPC 和子网, 同时从 VPC 中自动或手动分配 IP 地址。
- 外网类型的弹性网卡所属网络为外网网段, 同时会从外网网段中自动或手动分配 IP 地址, 且分配的 IP 地址与弹性网卡生命周期一致, 仅支持随弹性网卡销毁而释放。
- 当网卡类型为外网时, 网卡会根据所选外网 IP 的带宽规格进行计费, 用户可根据业务需要, 选择适合的付费方式和购买时长。

弹性网卡具有独立的生命周期, 支持绑定和解绑管理, 可在多个虚拟机间自

由迁移；虚拟机被销毁时，弹性网卡将自动解绑，可绑定至另一台虚拟机使用。

弹性网卡具有地域（数据中心）属性，仅支持绑定相同数据中心的虚拟机。一块弹性网卡仅支持绑定至一个虚拟机，x86 架构虚拟机可支持绑定 6 块弹性网卡，ARM 架构虚拟机可支持绑定 3 块网卡。

外网弹性网卡被绑定至虚拟机后，不影响虚拟机默认网络出口策略，包含虚拟机上弹性网卡绑定的外网 IP 在内，以第一个有默认路由的 IP 作为虚拟机的默认网络出口，用户可设置某一个有默认路由的外网 IP 为虚拟机默认网络出口。

每块弹性网卡仅支持分配一个 IP 地址，并可根据需要绑定一个安全组，用于控制进出弹性网卡的流量，实现精细化网络安全管控；如无需对弹性网卡的流量进行管控，可将弹性网卡的安全组置空。

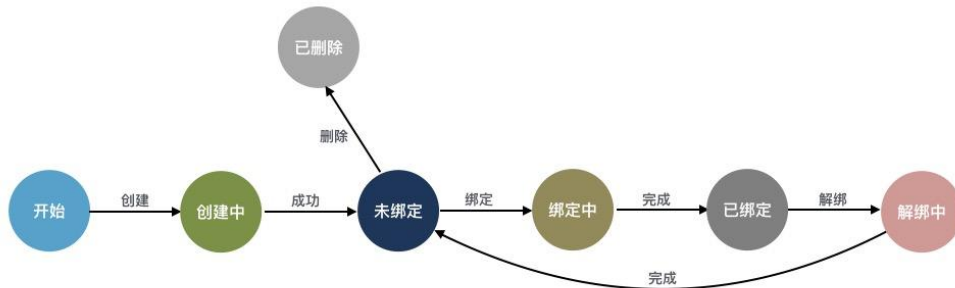
用户可通过平台自定义创建网卡，并对网卡进行绑定、解绑及修改安全组等相关操作，对于外网弹性网卡还可进行【调整带宽】操作，用于调整外网弹性网卡上的外网 IP 地址的带宽上限。

弹性网卡具有地域、网卡类型、VPC、子网、外网网段、外网 IP 带宽、IP 及安全组等属性，支持创建、绑定、解绑、绑定安全组、解绑安全组及删除弹性网卡等生命周期管理。

- 地域：弹性网卡仅支持绑定至相同地域的虚拟机。
- 网卡类型：弹性网卡的网络接入类型，支持 VPC 内网和 EIP 外网两种类型。
- VPC/子网：一块内网弹性网卡仅支持加入至一个 VPC 和子网，创建后无法修改 VPC 和子网。
- 外网网段：一块外网弹性网卡仅支持从一个外网网段中分配 IP 地址，创建后无法修改。
- 外网 IP 带宽：外网网卡分配 IP 地址的带宽。
- IP 地址：支持手动指定和自动获取弹性网卡在子网或外网网段内的 IP 地址，一块弹性网卡仅支持 1 个 IP 地址，创建后无法修改 IP 地址；

- 安全组：每块弹性网卡支持绑定一个安全组，提供网卡级别安全控制，详见安全组；
- MAC 地址：每块弹性网卡拥有全局唯一 MAC 地址；

弹性网卡整个生命周期包括创建中、未绑定、绑定中、已绑定、解绑中、已删除等状态，状态流转如下图所示：



4.2.7.3 VPC 网络

私有网络（VPC —— Virtual Private Cloud）是一个属于用户的、逻辑隔离的二层网络广播域环境。在一个私有网络内，用户可以构建并管理多个三层网络，即子网（Subnet），包括网络拓扑、IP 网段、IP 地址、网关等虚拟资源作为租户虚拟机业务的网络通信载体。

私有网络 VPC 是虚拟化网络的核心，为云平台虚拟机提供内网服务，包括网络广播域、子网（IP 网段）、IP 地址等，是所有 NVF 虚拟网络功能的基础。私有网络是子网的容器，不同私有网络之间是绝对隔离的，保证网络的隔离性和安全性。

VPC 网络具有数据中心属性，每个 VPC 私有网络仅属于一个数据中心，数据中心间资源和网络完全隔离，资源默认内网不通。租户内和租户间 VPC 网络默认不通，从不同维度保证租户网络和资源的隔离性。

虚拟机在创建时必须选择一个 VPC 网络和子网，不可进行变更。同时虚拟机从 VPC 的子网中分配 IP 地址、网关、DNS 等。有关 VPC 和子网详见【VPC 网络】章节描述。

4.2.7.4 外网 IP

平台支持 IPv4/IPv6 双栈网络，每个虚拟机最多支持绑定 50 个 IPv4 和 10 个 IPv6 外网 IP 地址，默认以第一个有默认路由的 IP 地址（包括外网弹性网卡的 IP 地址）作为虚拟机的默认网络出口。

支持在虚拟机中查看已绑定的外网 IP 地址及网络路由，虚拟机访问外网的流量直接通过虚拟网卡透传至物理网卡与外部网络通信，提升网络传输的性能。

外网 IP 信息包括虚拟机及绑定的外网弹性网卡 IP，仅支持将有默认路由的外网 IP 设为虚拟机默认网络出口。可通过列表信息查看已绑定外网 IP 的详细信息及相关管理操作，已绑定外网 IP 信息包括 IP 地址、IP 版本、IP 类型、出口、带宽、路由类型、绑定时间及状态。

- IP 指当前已绑定外网 IP 的 IP 地址及网段名称（网段是由平台管理员自定义的外网 IP 地址池）；
- IP 版本是指当前已绑定外网 IP 的 IP 版本，包括 IPv4 和 IPv6；
- IP 类型是指当前已绑定外网 IP 的 IP 类型，包括直通和 NAT；
- 出口指当前 IP 是否为虚拟机的默认出口，一台虚拟机最多支持两个默认出口（IPv4 和 IPv6 各一个）；
- 带宽指当前 IP 地址的带宽上限，带宽上限由申请外网 IP 地址时指定；
- 路由类型指当前 IP 地址所属网段下发路由的类型（网段路由策略由平台管理员自定义），包括默认路由和非默认路由，仅支持将有默认路由的外网 IP 设为虚拟机默认网络出口。
 - 默认路由类型指虚拟机绑定该 IP 地址时，会自动下发目标地址为 0.0.0.0/0 的路由到虚拟机中；
 - 非默认路由指虚拟机绑定该 IP 地址时，会下发管理员为网段配置的指定目标地址路由，如为虚拟机下发目标地址为 10.0.0.0/24 的路由；

- 若绑定至虚拟机的多个外网 IP 地址均为默认路由类型，默认以第一个有默认路由的 IP 地址作为虚拟机的默认出口。

用户可通过外网 IP 管理控制台的操作项，进行外网 IP 地址的绑定、解绑及设为默认出口操作，并支持批量解绑。有关外网 IP 的详细描述详见【外网 IP】。

绑定至虚拟机的外网弹性网卡的 IP 地址同时会展示至外网 IP 列表，支持设为出口操作但不支持解绑，可通过解绑弹性网卡进行弹性网卡外网 IP 的解绑和释放。

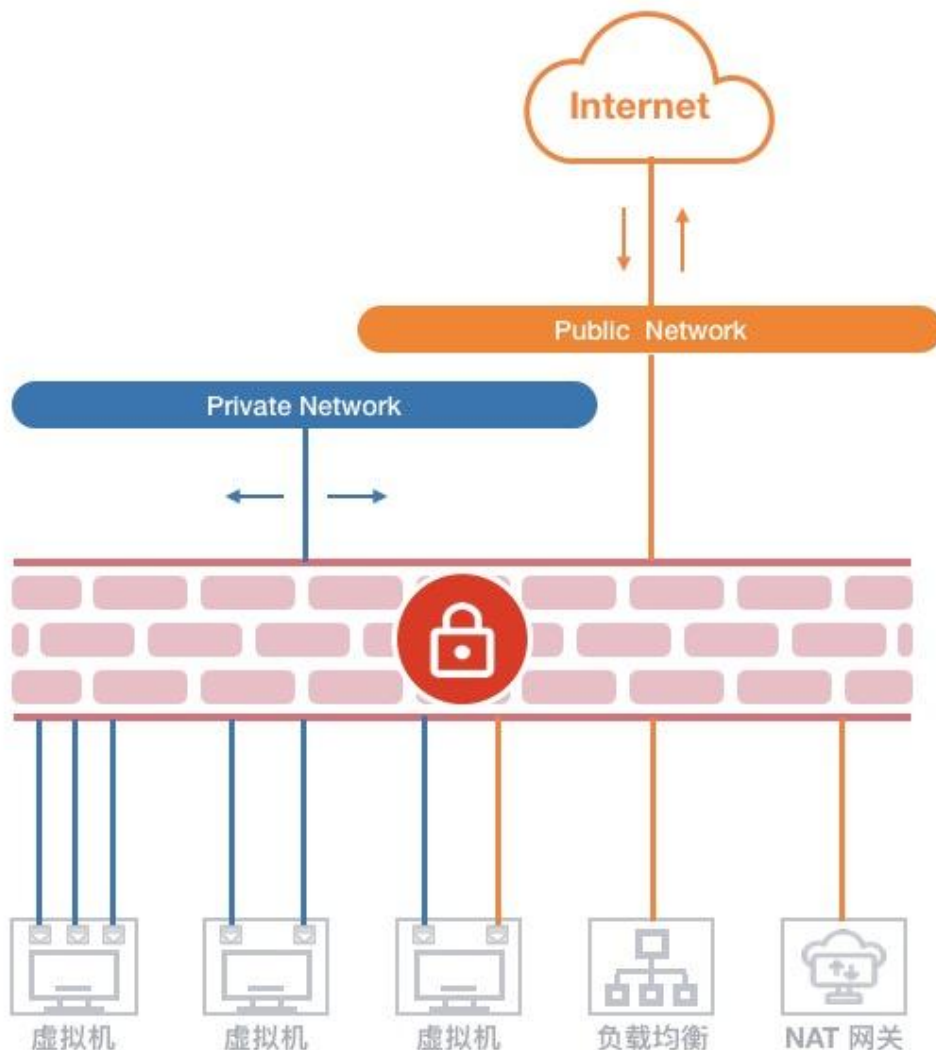
4.2.8 安全组

安全组 (Security Group) 是一种类似 [IPTABLES](#) 的虚拟防火墙，提供出入双方向流量访问控制规则，定义哪些网络或协议能访问资源，用于限制虚拟资源的网络访问流量，支持 IPv4 和 IPv6 双栈限制，为平台提供必要的安全保障。

4.2.8.1 实现机制

平台安全组基于 Linux Netfilter 子系统，通过在 [OVS](#) 流表中添加流表规则实现，需开启宿主机 IPv4 和 IPv6 包转发功能。每增加一条访问控制规则会根据网卡作为匹配条件，生成一条流表规则，用于控制进入 OVS 的流量，保证虚拟资源的网络安全。

安全组仅可作用于同一个数据中心内具有相同安全需求的虚拟机、弹性网卡、负载均衡、NAT 网关、MySQL、Redis、文件存储及对象存储等服务实例，工作原理如下图所示：



安全组具有独立的生命周期，可以将安全组与虚拟机、弹性网卡、负载均衡、NAT 网关绑定在一起，提供安全访问控制，与之绑定的虚拟资源销毁后，安全组将自动解绑。

- 安全组对虚拟机的安全防护针对的是一块网卡，即安全组是与虚拟机的默认虚拟网卡或弹性网卡绑定在一起，分别设置访问控制规则，限制每块网卡的出入网络流量；
- 安全组与提供外网 IP 服务的虚拟外网网卡绑定，通过添加出入站规则，对南北向（虚拟机外网）的访问流量进行过滤；
- 安全组与提供私有网络服务的虚拟网卡或弹性网卡绑定，通过添加出入站规则，控制东西向（虚拟机间及弹性网卡间）网络访问；

- 安全组与外网类型的负载均衡关联，通过添加出入站规则，可对进出外网负载均衡的外网 IP 流量进行限制和过滤，保证外网负载均衡器的流量安全；
- 安全组与 NAT 网关绑定，通过添加出入站规则，可对进入 NAT 网关的流量进行限制，保证 NAT 网关的可靠性和安全性；
- 一个安全组支持同时绑定至多个虚拟机、弹性网卡、NAT 网关及外网负载均衡实例；
- 虚拟机支持绑定一个内网安全组和一个外网安全组，分别对应虚拟机默认的内网网卡和外网网卡上，其中外网安全组对绑定至虚拟机的所有外网 IP 地址生效；
- 弹性网卡仅支持绑定一个安全组，与虚拟机默认网卡绑定的安全组相互独立，分别限制对应网卡的流量；
- 外网负载均衡和 NAT 网关实例仅支持绑定一个安全组，可更换安全组应用不同的网络访问规则。

支持创建虚拟机时不指定安全组，支持虚拟机启动后再进行调整，支持随时修改安全组的出入站规则，新规则生成时立即生效，可根据需求调整安全组出入方向的规则。支持安全组全生命周期管理，包括安全组创建、修改、删除及安全组规则的创建、修改、删除等生命周期管理。

4.2.8.2 安全组规则

安全组规则可控制允许到达安全组关联资源的入站流量及出站流量，提供双栈控制能力，支持对 IPv4/IPv6 地址的 TCP、UDP、ICMP 等协议数据包进行有效过滤和控制。

每个安全组支持配置多条规则，根据优先级对资源访问依次生效。**规则为空时，安全组将默认拒绝所有流量；规则不为空时，除已生成的规则外，默认拒绝其它访问流量。**


支持有状态的安全组规则，可以分别设置出入站规则，对被绑定资源的出入流量进行管控和限制。每条安全组规则由协议、端口、地址、动作、优先级及方向六个元素组成：

- 协议：支持 TCP、UDP、ICMPv4、ICMPv6 四种协议数据包过滤。
 - ALL 代表所有协议和端口，ALL TCP 代表所有 TCP 端口，ALL UDP 代表所有 UDP 端口；
 - 支持快捷协议指定，如 FTP、HTTP、HTTPS、PING、OpenVPN、PPTP、RDP、SSH 等；
 - ICMPv4 指 IPv4 版本网络的通信流量，ICMPv6 指 IPv6 版本网络的通信流量。
- 端口：源地址访问的本地虚拟资源或本地虚拟资源访问目标地址的 TCP/IP 端口。
 - TCP 和 UDP 协议的端口范围为 1~65535 ；
 - ICMPv4 和 ICMPv6 不支持配置端口。
- 地址：访问安全组绑定资源的网络数据包来源地址或被安全组绑定虚拟资源访问的目标地址。
 - 当规则的方向为入站规则时，地址代表访问被绑定虚拟资源的源 IP 地址段，支持 IPv4 和 IPv6 地址段；
 - 当规则的方向为出站规则时，地址代表被绑定虚拟资源访问目标 IP 地址段，支持 IPv4 和 IPv6 地址段；
 - 支持 CIDR 表示法的 IP 地址及网段，如 120.132.69.216 、 0.0.0.0/0 或 ::/0 。
- 动作：安全组生效时，对数据包的处理策略，包括“接受”和“拒绝”两种动作。
- 优先级：安全组内规则的生效顺序，包括高、中、低三档规则。

- 安全组按照优先级高低依次生效，优先生效优先级高的规则；
- 同优先级的规则，优先生效精确规则。
- 方向：安全组规则所对应的流量方向，包括出站流量和入站流量。
- 描述：每一条安全组规则的描述，用于标识规则的作用。

安全组支持数据流表状态，规则允许某个请求通信的同时，返回数据流会被自动允许，不受任何规则影响。即安全组规则仅对新建连接生效，对已经建立的链接默认允许双向通信。

如一条入方向规则允许任意地址通过互联网访问虚拟机外网 IP 的 80 端口，则访问虚拟机 80 端口的返回数据流（出站流量）会被自动允许，无需为该请求添加出方向允许规则。

 **注意** 通常建议设置简洁的安全组规则，可有效减少网络故障。

4.2.8.3 端口组和 IP 组

创建安全组规则时，除了通过自定义端口方式指定端口信息外，平台提供的端口组功能允许用户更灵活地定义安全组规则。用户可以将一组端口和协议组织成一个逻辑单元，并在安全组规则中引用该端口组，从而简化规则的管理和维护。

平台提供的 IP 组功能提供了类似的灵活性，允许用户指定包含多个 IP 地址、网段或连续地址段的规则。这使得在规则定义中，可以更加便捷地表示一组相关的网络地址，从而提高规则的可读性和管理效率。

端口组和 IP 组功能允许用户在单个规则中定义多个协议、端口及 IP 信息，从而简化规则的配置。用户无需为每个协议或端口创建单独的规则，而是可以通过选择端口组及 IP 组一次性定义相关的协议、端口和 IP 地址，这样的简化有助于降低配置的复杂性。通过集中管理相关信息和减少规则数量有助于降低配置错误的风险。

(1) 端口组

创建端口支持以下格式:

- 单个端口, 如: 80
- 多个单端口, 如: 80,443
- 连续端口段, 如: 3306-20000

单个端口组可包含多条“协议:端口”信息, 如:

```
TCP:80,443,3306-10000
UDP:53,1000
```

(2) IP 组

创建 IP 支持以下格式:

- 单个 IP, 如: 10.0.0.1 或 FF05::B5
- 网段, 如: 10.0.1.0/24 或 FF05:B5::/60
- 连续地址段, 如: 10.0.0.1-10.0.0.100

单个 IP 组可包含多个 IP 信息, 如:

```
10.0.1.10
172.16.1.0/24
10.0.1.100-10.0.1.200
```

4.2.9 隔离组

隔离组, 又称亲和反亲和策略, 允许用户自定义虚拟机与其它虚拟机或宿主机之间的调度关系, 意味着用户可以根据业务需求、性能要求或其它因素定义虚拟机调度逻辑, 以满足特定的架构和运行要求。

隔离组提供了灵活的资源调度机制, 用户可以根据实际需求动态地调整隔离组的配置, 以适应不同业务负载和调度需求。隔离组根据策略对象类型分为虚拟机组和节点组。

4.2.9.1 虚拟机组

策略对象类型为虚拟机组的隔离组，用于控制一组虚拟机之间或多组虚拟机之间的调度关系，支持亲和、反亲和两种策略。

(1) 亲和策略

亲和策略用于将策略相关的虚拟机实例调度部署在同一物理主机上，可以最大程度地提高它们之间的数据传输和通信效率。物理主机内部的虚拟机通信会经过高速内部网络，从而实现更高的性能和更低的通信延时。在同一物理主机上运行的虚拟机之间的通信不需要经过物理网络，减少了对网络带宽的占用，有助于减轻网络拥塞和提高整体系统的可扩展性。

亲和策略通常适用于有业务关联性的虚拟机实例，例如构成同一业务服务的多个虚拟机，虚拟机之间通常需要频繁地进行通信和数据共享，因此将它们调度到同一物理主机上有助于提高整体服务的性能。

- 非强制亲和：当物理节点资源小于虚拟机资源需求时，系统将会忽略亲和策略，选择其它合适的物理主机创建虚拟机。
- 强制性亲和：当物理节点资源无法满足虚拟机资源需求时，则虚拟机会因不满足强制性策略而阻塞调度。

(2) 反亲和策略

反亲和策略，也被称为非亲和性策略，是将虚拟机实例分散在不同物理主机上的调度策略，旨在降低单点故障的风险，提高系统的可靠性和容错性。通过分散调度降低了某一物理主机发生故障对整个系统造成影响的风险，当某个节点出现问题时，其它节点上的虚拟机可以继续运行，确保整个系统的持续性，以提高系统的可靠性、容错性和弹性。

- 非强制反亲和：当集群节点资源不能满足相关虚拟机完全分散调度时，系统将会忽略反亲和策略，从而出现多个虚拟机调度到相同物理主机的情况。
- 强制性反亲和：当集群节点资源不能满足相关虚拟机完全分散调度时，

则虚拟机会因不满足强制性策略而阻塞调度。

4.2.9.2 节点组

策略对象类型为节点组的隔离组，用于控制虚拟机和物理主机之间的调度关系。创建后支持加入目标节点及虚拟机实例，支持亲和、反亲和两种策略。

(1) 亲和策略

将组内的虚拟机实例调度部署在已加入组内的物理主机上，以实现虚拟机启动时的定向调度。

- 非强制亲和：当物理节点资源小于虚拟机资源需求时，系统将会忽略亲和策略，选择其它合适的物理主机创建虚拟机。
- 强制性亲和：当物理节点资源无法满足虚拟机资源需求时，则虚拟机会因不满足强制性策略而阻塞调度。

(2) 反亲和策略

将组内的虚拟机实例避免调度在已加入组内的物理主机上。

- 非强制反亲和：当集群节点资源不能满足虚拟机资源需求时，系统将会忽略反亲和策略，从而出现组内虚拟机调度到组内物理主机的情况。
- 强制性反亲和：当集群节点资源不能满足虚拟机资源需求时，则虚拟机会因不满足强制性策略而阻塞调度。

4.2.10 USB 透传

平台支持 USB 透传功能，物理机 USB 设备可直接透传至该物理机上所运行的虚拟机，USB 设备包含以下两种模式：

- 直通

将 USB 设备加载到此物理机上的虚拟机，迁移虚拟机时需要卸载此 USB 设备，通常用于对 USB 设备性能有要求的场景。

- 转发

将 USB 设备加载到此物理机所在计算集群，通过网络转发 USB 设备内的数据，迁移虚拟机时不需要卸载此 USB 设备。

4.2.11 VNC 登录

VNC (Virtual Network Console) 是平台为用户提供的一种通过 WEB 浏览器连接虚拟机的登录方式，适用于无法通过远程登录客户端 (如 SecureCRT、PuTTY 等) 连接虚拟机的场景。通过 VNC 登录连到虚拟机，可以查看虚拟机完整启动流程，并可以像 SSH 及远程桌面一样管理虚拟机操作系统及界面，支持发送各种操作系统管理指令，如 CTRL+ALT+DELETE。

支持用户获取虚拟机的 VNC 登录信息，包括 VNC 登录地址及登录密码，适用于使用 VNC 客户端连接虚拟机的场景，如桌面云场景。为确保 VNC 连接的安全性，每一次调用 API 或通过界面所获取的 VNC 登录信息有效期为 300 秒，如果 300 秒内用户未使用 IP 和端口进行连接，则信息直接失效，需要重新获取新的登录信息；同时用户使用 VNC 客户端登录虚拟机后，300 秒内无任何操作将会自动断开连接。

支持用户获取虚拟机的 Spice 登录信息，包括 Spice 登录地址及登录密码，同样适用于使用 Spice 客户端连接虚拟机的场景，如桌面云场景，与 VNC 连接一致，限制有效期 300 秒，保证连接的安全性。

4.2.12 自定义启动源

平台支持自定义虚拟机启动源，不仅可以采用常规方式选择镜像进行虚拟机创建，还可以充分发挥灵活性，选择已有的虚拟硬盘作为系统盘进行虚拟机的创建和启动。

在选择已有盘作为启动源时，要求所选盘中包含完整的操作系统，以确保虚拟机能够在启动时正常运行。该功能赋予用户极大的灵活性，使虚拟机的创建方式更为快捷、个性化，为用户提供了定制化的虚拟机创建方式。

4.2.13 自定义主机名称

支持自定义虚拟机主机名称，用于自动设置虚拟机操作系统内部的计算机名。批量创建时会在当前填写主机名添加有序后缀。

- Windows 系统，长度为 2~15 个字符。允许使用大小写字母、数字或连接符 (-)。不能以连字符 (-) 开头或结尾，不能连续使用连字符 (-)，也不能仅使用数字。
- Linux 系统，长度为 2~63 个字符。允许使用大小写字母、数字、点号 (.) 或连接符 (-)。不能以点号 (.) 或连字符 (-) 开头或结尾，不能连续使用点号 (.) 或连字符 (-)，也不能仅使用数字。

4.2.14 自定义 DNS

虚拟机默认 DNS 指定为 114.114.114.114，用于提供通用域名解析服务。同时系统支持用户自定义配置最多两个自定义 DNS 地址。

企业内部网络通常会使用自己的 DNS 服务器来处理内部域名解析，用户可以选择自定义虚拟机的 DNS 配置，以满足特定网络环境或业务需求。通过自定义 DNS，虚拟机能够适应特定的企业网络环境，确保内部域名的正确解析。

4.2.15 自定义 MAC

虚拟机的 MAC 地址为平台随机分配，以避免冲突和确保网络的唯一性。为满足特定场景下如授权与指定 MAC 地址绑定的需求，用户可以选择自定义虚拟机的 MAC 地址，在控制台上即可对关机状态下的虚拟机进行 MAC 地址设定。

在自定义 MAC 地址时，需要确保所选择的地址在整个网络中是唯一的，平台会对输入的 MAC 地址进行重复性校验以避免冲突。

4.2.16 自定义引导方式

虚拟机默认通过 BIOS 固件引导，对于启动盘磁盘格式是 GPT 的镜像需要通过 UEFI 方式引导。创建虚拟机时，引导方式默认和所选镜像保持一致，平台

支持自定义选择引导方式，支持 BIOS 和 UEFI。

- BIOS 使用 Master Boot Record (MBR) 进行引导，使用文本模式的启动界面，限制在 2TB 以下的硬盘容量，启动速度相对较慢，标准成熟，对于一些老旧的硬件和操作系统具有较好的兼容性。
- UEFI 使用 GUID Partition Table (GPT) 进行引导，支持图形化的用户界面，提供更直观的操作和信息显示，可以支持大于 2TB 的硬盘容量，启动速度更快，支持并行加载驱动和应用程序。

4.2.17 自定义 CPU 启动模式

支持虚拟机自定义选择 CPU 启动模式，分为默认 (Custom) 和直通 (Host-passthrough)，方便用户根据使用场景灵活选择。

- Custom 模式通过提炼 CPU 通用指令集，最大限度保障了虚拟机在不同宿主机之间热迁移时的兼容性。
- Host-passthrough 模式将宿主机的 CPU 指令集全部透传给虚拟机，可以最大限度的使用宿主机 CPU 指令集，但是在热迁移时要求目的节点的 CPU 和源节点的完全一致。

4.2.18 自定义高可用模式

支持设定虚拟机的高可用模式，提供【永不停止】和【无】两种选项，方便用户根据业务类型选择合适的高可用性模式。

- 选择【永不停止】模式的虚拟机将会开启高可用模式，确保虚拟机关闭后能够自动重启，无论是因为宿主机故障、虚拟机异常关闭或则其它原因，虚拟机将会尽快重新启动，使其处于长期运行状态，以维持业务的连续性和可用性。
- 选择【无】模式的虚拟机不启用高可用模式，虚拟机关闭后将不会自动重启，对于不需要长期运行的虚拟机，便于用户更灵活的控制运行状态。

4.3 GPU 虚拟机

4.3.1 概述

平台提供 GPU 设备透传能力，支持用户在平台上创建并运行 GPU 虚拟机，让虚拟机拥有高性能计算和图形处理能力。

GPU 虚拟机可以提供更好的成本效益。通过共享和灵活分配 GPU 资源，可以更有效地利用硬件资源，降低硬件投资和运营成本。同时虚拟机的动态调整和弹性扩展功能，可以根据实际需求进行资源分配，避免资源浪费。

- 支持用户选择 GPU 颗数，选择 GPU 规格创建 GPU 虚拟机，GPU 虚拟机与虚拟机的管理功能和生命周期一致。
- 支持用户关机状态下修改虚拟机配置、解绑 GPU。

支持透传的设备包括 NVIDIA 的 K80、P40、V100、2080、2080Ti、T4 及华为 Atlas300 等。

针对 GPU 虚拟机，平台支持最高配置 4 颗 GPU 芯片，为使 GPU 虚拟机发挥最佳性能，平台限制最小 CPU 内存规格为 GPU 颗数的 4 倍以上：

- 1 颗 GPU 芯片最小需要 4 核 8G 规格
- 2 颗 GPU 芯片最小需要 8 核 16G 规格
- 4 颗 GPU 芯片最小需要 16 核 32G 规格

4.3.2 应用场景

- GPU 资源共享

GPU 虚拟机允许多个用户共享同一台物理服务器上的 GPU 资源。每个虚拟机实例可以分配物理服务器上的一个或多个 GPU 资源，以满足不同用户的需求。

- 高性能图形处理

GPU 虚拟机提供了强大的图形处理能力，可以加速图形密集型任务，如游戏渲染、图像处理和视频编码等。通过虚拟化技术，多个用户可以同时享受到高性能的图形处理能力。

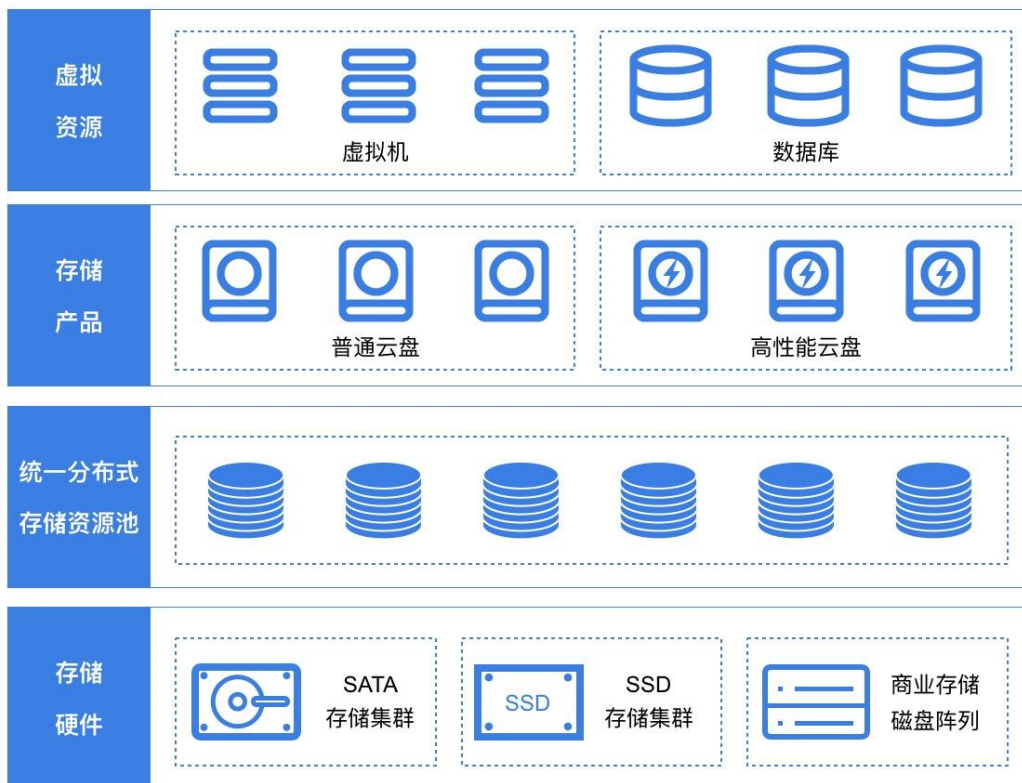
- GPU 加速计算

GPU 虚拟机不仅可以用于图形处理，还可以用于加速通用计算任务。虚拟机实例可以利用 GPU 的并行计算能力，加速科学计算、机器学习、数据分析、AI 训练、AI 推理等工作负载。

4.4 云硬盘

4.4.1 概述

云硬盘是一种基于分布式存储系统为虚拟机提供持久化存储空间的块设备。具有独立的生命周期，支持随意绑定/解绑至多个虚拟机使用，并能够在存储空间不足时对云硬盘进行扩容，基于网络分布式访问，为云主机提供高安全、高可靠、高性能及可扩展的数据磁盘。



存储系统兼容并支持多种底层存储硬件，如通用服务器（计算存储超融合或独立通用存储服务器）和商业存储，并将底层存储硬件分别抽象不同类型集群的存储资源池，由分布式存储系统统一调度和管理。在实际应用场景中，可以将普通 SATA 接口的机械盘统一抽象为【SATA 存储集群】，将 SSD 全闪磁盘统一抽象为【SSD 存储集群】，分别由统一存储封装后提供平台用户使用。

如示意图所示，将 SATA 存储集群的资源封装为普通云盘，将 SSD 全闪存存储集群的资源封装为高性能云盘。平台的虚拟机和数据库服务可根据需求挂载不同存储集群类型的磁盘，支持同时挂载多种集群类型的云硬盘。云平台管理员可通过管理员控制台自定义存储集群类型的别名，用于标识不同磁盘介质、不同品牌、不同性能或不同底层硬件的存储集群，如 EMC 存储集群、SSD 存储集群等。

通常 SSD 磁盘介质的云硬盘的性能与容量的大小成线性关系，容量越大提供的 IO 性能越高，如对 IO 性能有强烈需求，可考虑扩容 SSD 磁盘介质的云硬盘。

分布式存储底层数据通过 PG 映射的方式进行数据存储，同时以多副本存储的方式保证数据安全，即写入至云平台存储集群的数据块会同时保存多份至不同服务器节点的磁盘。

多副本存储的数据提供一致性保证，可能导致写入的多份数据因误操作或原始数据异常导致数据不准确；为保证数据的准确性，云平台提供硬盘快照能力，将云盘数据在某一时间点的数据文件及状态进行备份，在数据丢失或损坏时，可通过快照快速恢复数据，包括数据库数据、应用数据及文件目录数据等，可实现分钟级恢复。

4.4.2 功能与特性

云硬盘由统一存储从存储集群容量中分配，为平台虚拟资源提供块存储设备并共享整个分布式存储集群的容量及性能；同时通过块存储系统为用户提供云硬盘资源及全生命周期管理，包括云硬盘的创建、绑定、解绑、扩容、克隆、快照及删除等管理。

云硬盘容量是由统一存储的从存储集群容量中分配的，所有云硬盘共享整个

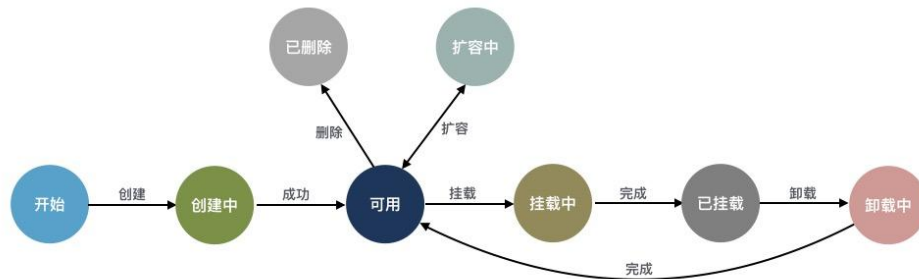
分布式存储池的容量及性能。

- 支持云硬盘创建、挂载、卸载、磁盘扩容、删除等生命周期管理，单块云硬盘同时仅能挂载一台虚拟机。
- 支持在线和离线的方式扩容磁盘容量，磁盘扩容后需要在虚拟机的操作系统进行磁盘容量的扩容操作。
- 为保证数据安全性及准确性，云硬盘仅支持磁盘扩容，不支持磁盘缩容。
- 云硬盘最小支持 10G 的容量，步长为 1GB，可自定义控制单块云硬盘的最大容量。
- 云硬盘具有独立的生命周期，可自由绑定至任意虚拟机或数据库服务，解绑后可重新挂载至其它虚拟机；
- X86 架构的虚拟机最多支持绑定 25 块云硬盘，ARM 架构虚拟机最多支持绑定 3 块云硬盘；
- 支持云硬盘克隆，即将云硬盘内的数据复制成为一个新的云硬盘；
- 支持对云硬盘进行快照备份，包括虚拟机的系统盘快照及弹性云盘快照，并可从快照回滚数据至云硬盘，用于数据恢复和还原场景；
- 支持对全局及每一块云硬盘的 QoS 进行配置，可根据不同业务模式调整磁盘的性能，以平衡平台整体性能；
- 支持设置存储集群类型权限，即可以将部分存储资源设置为租户独享，满足需要独享底层存储资源的场景。
- 支持从云硬盘创建虚拟机，云硬盘需要有能正常启动的镜像系统。

支持自动精简配置，在创建云硬盘时，仅呈现分配的逻辑虚拟容量。当用户向逻辑存储容量中写入数据时，按照存储容量分配策略从物理空间分配实际容量。如一个用户创建的云硬盘为 1TB 容量，存储系统会为用户分配并呈现 1TB 的逻辑卷，仅当用户在云硬盘中写入数据时，才会真正的分配物理磁盘容量。

高性能型云硬盘的性能与容量的大小成线性关系，容量越大，提供的 IO 性

能越高，如果对 IO 性能有强烈需求，可考虑扩容性能型云硬盘。UCloudStack 云硬盘完整生命周期包括创建中、可用、挂载中、已挂载、卸载中、扩容中、已删除等资源状态，各状态流转如下图所示：



4.4.3 应用场景

(1) 普通虚拟硬盘 (SATA+SSD 缓存)

- 适用于对容量要求较高且数据不被经常访问或 I/O 负载低的应用场景；
- 需要低成本并且有随机读写 I/O 的应用环境，如大型视频、音乐、离线文档存储等；

(2) 高性能虚拟硬盘 (SSD/NVME)

- 适用于 I/O 负载高且数据经常被读写的应用场景；
- 中大型关系数据库；
- 中大型开发测试环境；
- 中大型实时响应服务类环境；

4.5 共享云盘

共享云硬盘是一种支持多个云服务器并发读写访问的数据块级存储设备，具备多挂载点、高并发性、高性能、高可靠性等特点。主要应用于需要支持集群、HA (High Available, 指高可用集群) 能力的关键企业应用场景，多个云服务器可同时访问一个共享云硬盘。

用户可通过指定共享硬盘的类型、容量及名称即可快速创建一块云硬盘，作

为虚拟机的共享数据盘。

支持将云硬盘、SAN 存储 LUN 设备设备为共享盘，并作为虚拟机的数据盘，使多个虚拟机同时对共享盘进行数据读写操作。同时支持对共享盘进行创建、绑定、解绑、扩容、克隆、续费及删除等操作。

4.6 快照服务

平台分布式存储支持磁盘快照能力，可降低因误操作、版本升级等导致的数据丢失风险，是平台保证数据安全的一个重要措施。

快照是某一时间点一块虚拟硬盘的数据状态文件，可以理解虚拟硬盘某个时刻的数据备份，虚拟硬盘的数据写入和修改不会对已创建的快照造成影响。

支持定时快照策略，即一个可周期性执行的自动创建快照的策略，快照策略与快照分离，拥有独立的生命周期。在实际应用中，磁盘快照可降低因误操作、版本升级等导致的数据丢失风险，可大致应用于以下业务场景：

- **容灾备份**：定时为虚拟硬盘制作快照，当系统出现问题时，可快速回退，避免数据丢失。
- **版本回退**：在业务做重大升级时，建议预先做好快照，当升级版本出现系统问题无法修复时，可通过快照恢复到历史版本。

用户可为某块虚拟硬盘创建快照，同时支持对虚拟机系统盘进行快照备份。为保证数据及磁盘的安全：

- 仅支持对未绑定及已绑定的硬盘进行快照操作，若硬盘在扩容或快照中，无法进行快照备份；
- 创建快照时，不可进行磁盘挂载/卸载及修改虚拟机状态（如开机或关机），否则可能会导致快照创建异常；
- 快照仅捕获已写入硬盘的数据，不包含应用程序或操作系统缓存在内存中的数据，建议在暂停对硬盘的 I/O 操作后进行快照制作，如关机或卸载硬盘。

平台支持对已绑定虚拟机的系统盘及数据盘进行快照操作, 并支持快照回滚操作, 即将虚拟硬盘回滚到快照时刻的数据状态, 以满足数据恢复的应用场景。同时支持通过快照创建虚拟硬盘。

(1) 回滚快照

将虚拟硬盘回滚到快照时刻的数据状态, 应对快照数据恢复的应用场景。回滚时虚拟硬盘必须处于未绑定或绑定的虚拟机必须处于关机状态, 仅支持正常状态的快照进行回滚操作。

(2) 从快照创建虚拟硬盘

创建的硬盘大小与快照的原始硬盘大小相等, 继承加密属性; 从快照创建虚拟硬盘, 该虚拟硬盘只能与快照所对应的原始虚拟硬盘归属同一存储集群, 可以用系统盘快照创建的虚拟硬盘创建虚拟机。

(3) 快照删除

平台采用 Copy-on-Write (COW, 写时复制) 快照技术。系统通过元数据记录每个快照引用的数据块, 元数据包含了关于每个数据块在哪个快照中被引用的信息。多个快照可以引用相同的数据块, 系统不需要在存储上复制数据块的多个副本, 从而有效减少了存储空间占用。当用户删除某个快照时, 系统会检查该快照引用的数据块, 并且只会删除仅有该快照引用的数据, 多个快照引用的数据块则不会被删除, 以保障其它任意快照的完整性。

4.7 商业存储服务

4.7.1 概述

云平台默认提供分布式存储作为虚拟化的后端存储, 为云平台用户提供高可用、高性能、高可靠及高安全的存储服务。同时云平台虚拟化支持对接商业存储设备, 如 IP SAN 等存储阵列, 为云平台虚拟机提供集群中高性能块存储服务, 同时可利用旧企业用户的集中存储设备, 整体节省信息化转型的总拥有成本。

外置存储服务是云平台为企业用户提供的商业存储服务, 目前支持 iSCSI

协议、FC 协议对接商业存储，将商业存储作为虚拟化后端存储池，提供存储池管理及逻辑卷分配，可直接作为虚拟机的系统盘及数据盘进行使用，即只要支持 ISCSI 协议、FC 协议的存储设备均可作为平台虚拟化的后端存储，适应多种应用场景。

4.7.2 功能与特性

平台支持存储设备的对接和管理，并支持将存储设备中的 LUN 分配给租户，由租户将 LUN 分配或挂载至虚拟机的系统盘或数据盘，进行数据的读写，具体功能特性如下：

- 支持存储设备资源池的录入管理，并支持一键扫描 ISCSI 设备、FC 设备中已创建的 LUN 存储卷信息。
- 支持将已扫描的 LUN 存储卷分配给平台租户，使租户有权限使用磁盘作为虚拟机的系统盘或数据盘。
- 支持租户将有权限的 LUN 存储卷信息作为虚拟机的系统盘，使虚拟机直接运行直商业存储中，提升性能。
- 支持租户将有权限的 LUN 存储卷信息作为虚拟机的数据盘。
- 支持将存储卷重新分配给平台其它租户。

基于以上功能特性，平台可支持直接使用商业存储设备作为虚拟化的后端存储，为虚拟机提供传统商业存储设备的存储空间，同时不影响商业存储中的其它 LUN 为其它业务提供存储服务。

平台基于 ISCSI 协议、FC 协议对接商业存储，在对接中需要将存储设备的 LUN 映射到平台计算节点，使平台计算节点上运行的虚拟机可直接使用映射的 LUN；同时为保证虚拟机的高可用，需要将 LUN 同时映射到一个集群内的所有计算节点，即所有计算节点均可挂载并使用映射的存储卷，以保证宕机迁移时可在每个计算节点挂载该存储卷信息。

- 当虚拟机所在的计算节点故障时，平台会自动触发虚拟机宕机迁移，即将虚拟机迁移至计算集群内正常的计算节点上，使虚拟机可正常提供服务。

务。

- 虚拟机使用的 LUN 存储卷已被映射到集群内所有计算节点，当虚拟机在集群内迁移至新节点后，可直接使用已映射的 LUN 存储启动虚拟机的系统盘或数据盘，并正常挂载至虚拟机，保证虚拟机迁移后业务正常。

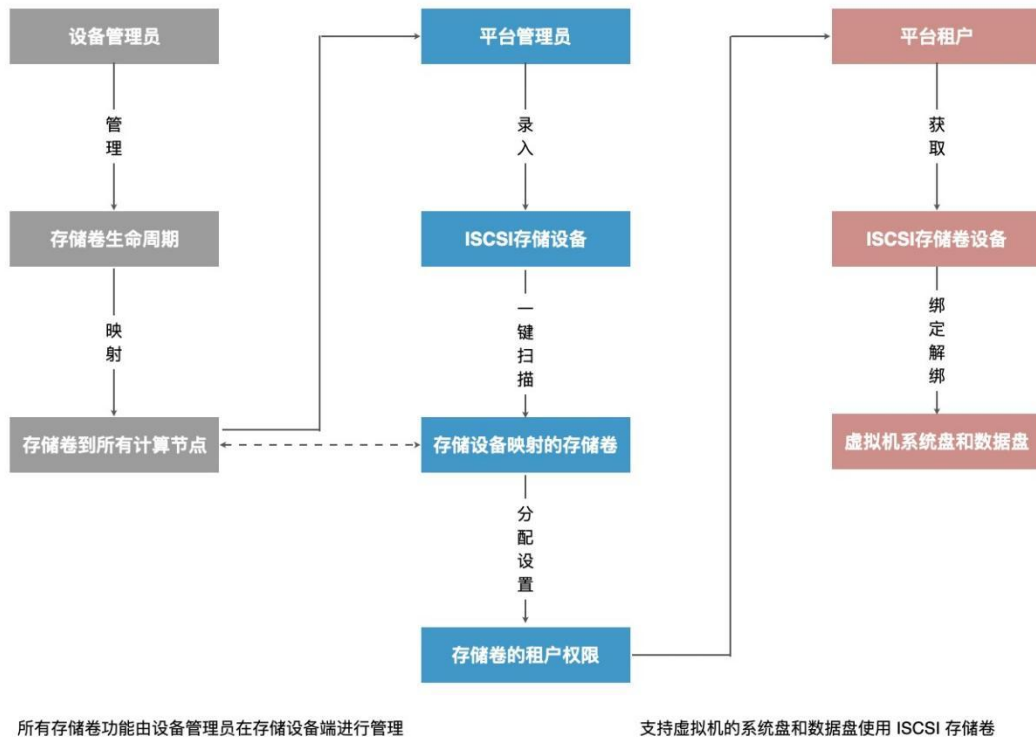
ISCSI 协议、FC 协议各有侧重，ISCSI 基于 TCP/IP 协议，设备对协议的支持一致性好；FC 协议速度快，需要购置专门的交换机，支持用户根据需求随意搭配。

平台仅将商业存储的 LUN 作为存储卷进行使用，不对存储卷本身进行管理，如 LUN 的创建、映射、扩容、快照、备份、回滚、克隆等。

4.7.3 使用流程

在使用外置存储前，需要平台管理者或存储设备管理者，将外置存储与平台的计算节点网络打通，使计算节点可与存储设备间直接内网可互相通信。

物理存储设备及网络准备好后，即可与平台进行对接并使用平台提供的外置存储服务，整个对接过程需要存储设备管理员、平台管理员及平台租户三个角色进行操作，其中与平台相关的为平台管理员和平台租户的操作，如下图流程所示：



1. 存储设备管理员管理存储卷

所有存储卷的管理均由存储设备管理员自行在商业存储的管理系统上进行操作，包括存储卷（Lun）的创建和映射，同时包括存储卷的扩容、快照、备份及删除等相关生命周期管理。

2. 存储设备管理员映射存储卷至集群计算节点

创建好的 Lun，由存储设备管理员在存储设备上映射到所有计算节点（如果新增计算节点，需再次进行映射），同时也可进行多路径映射。

3. 平台管理员录入并管理存储设备

- 针对 ISCSI 存储

存储卷 LUN 映射成功后，由【平台管理员】在管理控制台“外置存储-ISCSI”中进行 ISCSI 存储池或存储设备的录入，录入时需要指定存储设备的 ISCSI 地址，如 172.18.12.8:8080。

- 针对 FC 存储：无需录入设备地址，可通过平台直接扫描添加。

4. 平台管理员扫描已映射的 LUN 信息

- **ISCSI 存储:**

录入的存储设备后，由【平台管理员】在存储设备中一键扫描 ISCSI 存储设备中已被映射至集群节点上的存储卷设备及信息。

- **FC 存储:**

存储卷 LUN 映射成功后，由【平台管理员】在管理控制台“外置存储-FC SAN”中进行扫描，即可将系统中存在的存储设备添加到平台。

5. 平台管理员为租户分配 LUN 设备

由【平台管理员】将扫描成功的 LUN 存储卷设备指定给租户，一个存储卷同一时间仅支持分配给一个租户，分配后租户在外置存储设备中即可查询已分配的存储卷设备，并可进行创建虚拟机或挂载虚拟机。

6. 平台租户使用 LUN 存储卷设备

平台租户通过控制台外置存储可直接查询已分配的存储卷，并在创建虚拟机时指定系统盘类型为外置存储，或者也可直接将 LUN 存储卷直接挂载给已有虚拟机，作为虚拟机的数据盘进行使用。

平台租户使用外置存储服务的前提是存储卷已映射并分配给租户，租户只需要简单的绑定即可便捷的使用平台提供的外置存储设备，并可进行弹性绑定、解绑及设为共享硬盘。

4.8 私有网络

4.8.1 VPC 概述

平台通过软件定义网络（SDN）对传统数据中心物理网络进行虚拟化，采用 OVS 作为虚拟交换机，VXLAN 隧道作为 OverLay 网络隔离手段，通过三层协议封装二层协议，用于定义虚拟私有网络 VPC 及不同虚拟机 IP 地址之间数据包的封装和转发。

私有网络（VPC——Virtual Private Cloud）是一个属于用户的、逻辑隔离

的二层网络广播域环境。在一个私有网络内,用户可以构建并管理多个三层网络,即子网 (Subnet),包括网络拓扑、IP 网段、IP 地址、网关等虚拟资源作为租户虚拟机业务的网络通信载体。

私有网络 VPC 是虚拟化网络的核心,为云平台虚拟机提供内网服务,包括网络广播域、子网 (IP 网段)、IP 地址等,是所有 NVF 虚拟网络功能的基础。私有网络是子网的容器,不同私有网络之间是绝对隔离的,保证网络的隔离性和安全性。

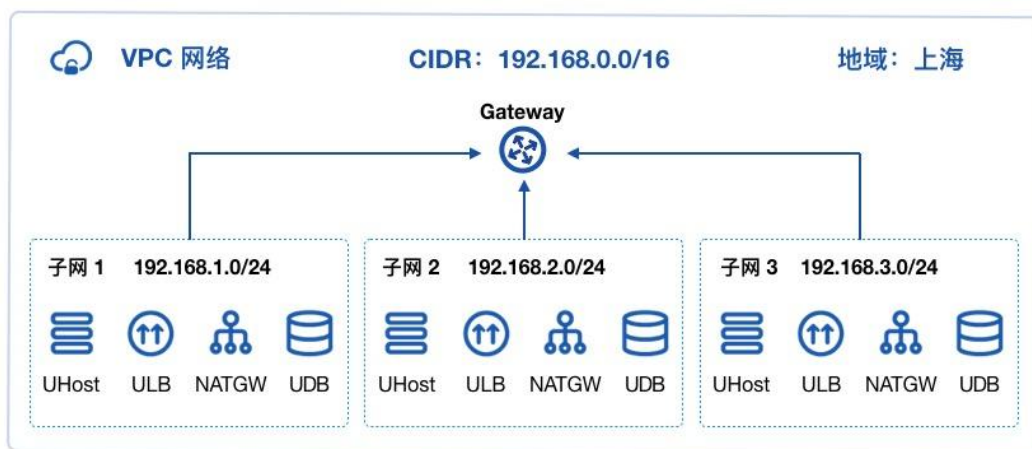
可将虚拟机、负载均衡、弹性网卡、NAT 网关等虚拟资源加入至私有网络的子网中,提供类似传统数据中心交换机的功能,支持自定义规划网络,并通过安全组对虚拟资源 VPC 间的流量进行安全防护。

说明 可通过 IPSecVPN、专线及外网 IP 接入等方式将云平台私有网络及虚拟资源与其它云平台或 IDC 数据中心组成一个按需定制的混合云网络环境。

VPC 网络具有数据中心属性,每个 VPC 私有网络仅属于一个数据中心,数据中心间资源和网络完全隔离,资源默认内网不通。租户内和租户间 VPC 网络默认不通,从不同维度保证租户网络和资源的隔离性。

4.8.2 VPC 逻辑结构

一个 VPC 网络主要由私有网络网段和子网两部分组成,如下图所示:



(1) 私有网络网段

VPC 网络所属的 CIDR 网段，作为 VPC 隔离网络的私网网段。关于 CIDR 的相关信息，详见 CIDR。创建 VPC 网络需指定私有网段，平台管理员可通过管理控制台自定义 VPC 私有网络的网段，使租户的虚拟资源仅使用管理员定义网段的 IP 地址进行通信。平台 VPC 私有网络 CIDR 默认支持的网段范围如下表所示：

网段	掩码范围	IP 地址范围	默认/可配置
10.0.0.0/8	8 ~ 29	10.0.0.0 - 10.255.255.255	可配置项
10.0.0.0/16	16 ~ 29	10.0.0.0 - 10.10.255.255	默认配置
172.16.0.0/16 ~ 172.29.0.0/16	16 ~ 29	172.16.0.0 - 172.29.255.255	可配置项
192.168.0.0/16	16 ~ 29	192.168.0.0 - 192.168.255.255	默认配置

由于 DHCP 及相关服务需占用 IP 地址，私有网络 CIDR 网段不支持 30 位掩码的私有网段。

(2) 子网

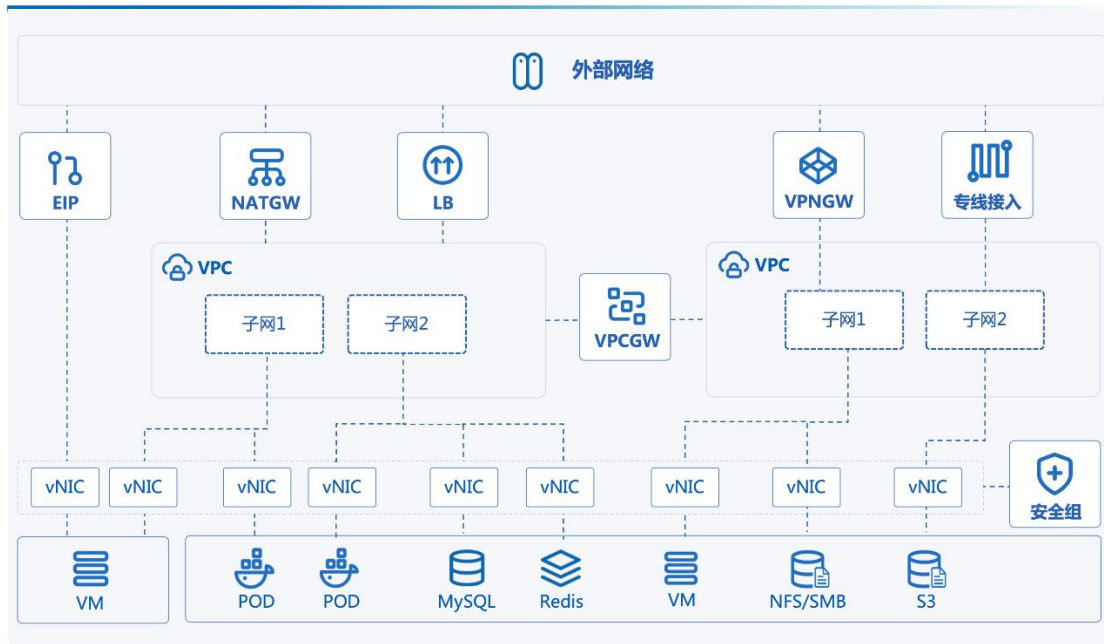
子网 (Subnet) 是 VPC 私有网络的基础网络地址空间，用于虚拟资源间内网连接。

- 一个私有网络至少由一个子网组成，子网的 CIDR 必须在 VPC 的 CIDR 网段内；
- 同一私有网络内子网间通过公共网关连接，资源默认内网互通，可部署虚拟机、负载均衡、NAT 网关及 IPsecVPN 网关等；
- 同一个 VPC 子网间默认通过公共网关进行互通；
- 子网 CIDR 网段掩码最小为 29 位，不支持 30、32 位掩码的子网网段；
- 每个子网中，使用第一个可用 IP 地址作为网关，如 192.168.1.0/24 的网关地址是 192.168.1.1。

当子网中存在虚拟资源时，不允许删除并销毁私有网络和子网资源。

4.8.3 VPC 连接

平台对常用网络设备均进行软件定义及组件抽象，通过将 VPC 网络与虚拟机、弹性网卡、外网 IP、安全组、NAT 网关、负载均衡、VPN 网关等组件连接，可快速构建和配置繁杂的网络环境及混合云场景，如下图所示：



- 虚拟机默认内网网卡（创建时自带的虚拟网卡）加入同一个 VPC 网络实现虚拟机间网络通信，并可通过安全组保证虚拟机东西向流量安全。
- 虚拟机默认外网网卡（创建时自带的虚拟网卡）可直接绑定多个外网 IP 地址实现 Internet 访问，同时可绑定与 IDC 物理网络相连的外网 IP 地址实现物理网络打通，结合安全组管控虚拟机南北向流量的同时，构建安全可靠的混合接入环境。
- 虚拟机的弹性网卡加入相同 VPC 网络的子网，实现精细化网络管理及廉价故障转移方案，同时将安全组与弹性网卡绑定，通过安全组规则多维度保障私有网络及虚拟资源的安全。
- 相同 VPC 网络的虚拟机可通过 NAT 网关及外网 IP 连接，共享外网 IP 访问 Internet 或 IDC 数据中心网络，并可通过 DNAT 端口映射对外提供业务服务。

- 相同 VPC 网络的虚拟机加入至内网 LB 后端服务节点，提供 VPC 网络内负载均衡服务。
- 相同 VPC 网络的虚拟机加入到外网 LB 后端服务节点，结合 LB 关联的外网 IP，提供外网负载均衡服务。
- 相同 VPC 网络的虚拟机通过 VPC 网关（VPC 互通）可与不同 VPC 网络的虚拟机进行内网互联，实现 VPC 间互通。
- 通过 IPSecVPN 网关打通不同 VPC 间的网络，使两个 VPC 间的虚拟机可直接进行内网通信。
- 采用 IPSecVPN 网关或专线接入网关将平台与本地 IDC 数据中心及第三方云平台连通，构建安全可靠的混合云环境。

外网 IP 可用于打通 IDC 数据中心的物理网络，应用与虚拟机直接与物理机进行内网通信的场景；IPSecVPN 网关用于打通第三方云平台或 IDC 数据中心的虚拟网络，应用于不同云平台间通过 VPN 安全连接场景。

4.8.4 功能与特性

平台 VPC 网络基于租户控制台和 API 提供隔离网络环境、自定义子网、子网通信及安全防护等功能，并可结合硬件及 DPDK 等技术特性提供高性能的虚拟网络。

- **隔离的网络环境**

私有网络基于 OVS（Open vSwitch）组件，通过 VXLAN 隧道封装技术实现隔离的虚拟网络。每一个 VPC 网络对应一个 VXLAN 隧道号（VNI），作为全局唯一网络标识符，为租户提供一张独立且完全隔离的二层网络，可通过在私有网络中划分多个子网作为虚拟资源的通信载体，用于连通多个虚拟资源。不同的 VPC 网络间完全隔离，无法直接通信。

- **自定义子网**

支持在一个 VPC 网络内进行三层网络规划，即划分一个或多个子网。提供

自定义 IP 网段范围、可用 IP 网段及默认网关，可在子网中通过虚拟机部署应用程序和服务。支持在子网中增加多个弹性网卡，分别指定子网中的 IP 地址，并绑定至部署应用程序的虚拟机，用于精细化管理应用服务的网络访问。

● 子网通信

每一个子网都属于一个广播域，VPC 网络默认提供网关服务，同一个 VPC 内不同子网通过网关进行通信。

● 安全防护

云平台提供内网安全组和外网防火墙，通过协议、端口为虚拟资源提供多维度安全访问控制，同时基于虚拟网卡及虚拟实例的网络流量进行上下行的 QoS 控制，全方位提高 VPC 网络的安全性。安全组为有状态安全层，可分别设置出入方向的安全规则，用于控制并过滤进出子网 IP 的数据流量。

● 高性能虚拟网络

SDN 网络分布式部署于所有计算节点，节点间通过 20GE 冗余链路进行通信，并通过所有计算节点负载内网流量，为云平台提供高可靠及高性能的虚拟网络。

云平台在保证网络隔离、网络规模、网络通信及安全的同时，为租户和子账号提供 VPC 子网的创建、修改、删除及操作审计日志等全生命周期管理。用户创建虚拟机、NAT 网关、负载均衡、VPN 网关等虚拟资源时可指定需加入的 VPC 网络和子网，并可查询每个子网的可用 IP 数量。

VPC 网络具有数据中心属性，不同数据中心之间的虚拟资源默认内网不互通，同数据中心内不同 VPC 间默认内网不互通，同一个 VPC 的所有子网和资源默认内网互通。仅支持指定相同数据中心的虚拟资源到 VPC 网络中，且每个 VPC 网络的子网网段必须在 VPC 网络的 CIDR 网段中。

平台会通过管理员配置的 VPC 网络，为每个租户和子账号提供默认的 VPC 网络和子网资源，方便用户登录云平台快速部署业务。

4.8.5 自定义路由

提供自定义网络规划能力，支持子网和虚拟机级别的自定义路由能力，用于控制子网出流量的走向。用户可通过自定义路由配置子网内路由策略，包括目的地址、下一跳类型、下一跳，下一跳类型指虚拟机、VIP 和自定义类型等。

(1) 目的端口

目的端即为您要转发到的目标网段，目的网段描述仅支持网段格式，如果您希望目的端为单个 IP，可设置掩码为 32（例如 172.16.1.1/32）。

(2) 下一跳类型

下一跳类型	默认/可配置
Local	不可编辑，提供 VPC 互通能力
NAT 网关	NAT 网关，不可编辑，NATGW 下发的路由
IPSecVPN	IPSecVPN，不可编辑，IPSecVPN 下发的路由
公共服务	公共服务，不可编辑
VIP	VIP，可编辑，VIP
虚拟机	虚拟机，可编辑，虚拟机资源
自定义	自定义，可编辑，自定义地址

(3) 下一跳

指定具体跳转到的下一跳实例，如网关或云服务器 IP 等。

4.8.6 网络拓扑

VPC 网络子网提供网络拓扑，用于查看子网的使用情况和拓扑状态。

支持用户查看子网详细使用情况，各资源占用子网 IP、绑定资源 ID、MAC 等详细信息。

4.8.7 VPC 互通

网络互通功能用于实现同租户两个 VPC 之间的网络互通，租户可以通过网络互通功能将两个 VPC 之间建立连接，让用户可以使用私有 IP 地址在两个 VPC 之间进行通信，体验类似两个 VPC 在同一个网络中。

- 配置网络互通时，两端 VPC 的网段（CIDR）不能重叠，否则可能会造成路由冲突，导致配置不生效。
- 两个 VPC 之间不能同时建立多个 VPC 连接。
- VPC 中存在连接时，VPC 网关不能关闭。

支持 VPC 网络互通连接和断开，并可支持用户查看当前 VPC 网络已联通的 VPC 网络。

4.9 组播

4.9.1 概述

作为一种与单播（Unicast）和广播（Broadcast）并列的通信方式，组播（Multicast）技术能够有效地解决单点发送、多点接收的问题，从而实现了网络中点到多点的高效数据传送，能够节约大量网络带宽、降低网络负载。

利用网络的组播特性方便地提供一些新的增值业务，包括在线直播、网络电视、远程教育、远程医疗、网络电台、实时视频会议等互联网的信息服务领域。

组播是主机间一对多的通讯模式，组播是一种允许一个或多个组播源发送同一报文到多个接收者的技术。组播源将一份报文发送到特定的组播地址，组播地址不同于单播地址，它并不属于特定某个主机，而是属于一组主机。一个组播地址表示一个群组，需要接收组播报文的接收者都加入这个群组。

- **组播组**

用 IP 组播地址进行标识的一个集合。任何用户主机（或其他接收设备），加入一个组播组，就成为该组成员，可以识别并接收发往该组播组的组播数

据。

- 组播源

信息发送者称为“组播源”，一个组播源可以同时向多个组播组发送数据，多个组播源也可同时向一个组播组发送报文。组播源通常不需要加入组播组，由源端 DR 负责管理组播源的注册和 SPT (Shortest Path Tree) 的建立。

4.9.2 组播组成员

所有加入某组播组的主机便成为该组播组的成员，组播组中的成员是动态的，主机可以在任何时刻加入或离开组播组。组播组成员可以广泛地分布在网络中的任何地方。

4.9.3 组播路由器

支持三层组播功能的路由器或交换机。组播路由器不仅能够提供组播路由功能，也能够在与用户连接的末梢网段上提供组播组成员的管理功能。

4.9.4 组播地址

IANA (Internet Assigned Numbers Authority, 互联网编号分配委员会) 将 D 类地址空间分配给 IPv4 组播使用。IPv4 地址一共 32 位，D 类地址最高 4 位为 1110，因此地址范围从 224.0.0.0 到 239.255.255.255，具体分类及含义详见下表描述：

地址范围	含义
224.0.0.0 ~ 224.0.0.255	永久组地址。IANA 为路由协议预留的 IP 地址（也称为保留组地址），用于标识一组特定的网络设备，供路由协议、拓扑查找等使用，不用于组播转发。
224.0.1.0 ~ 231.255.255.255 233.0.0.0 ~ 238.255.255.255	ASM 组播地址，全网范围内有效。说明：其中，224.0.1.39 和 224.0.1.40 是保留地址，不建议使用。

232.0.0.0 ~ 232.255.255.255	缺省情况下的 SSM 组播地址，全网范围内有效。
239.0.0.0 ~ 239.255.255.255	本地管理组地址，仅在本地管理域内有效。在不同的管理域内重复使用相同的本地管理组地址不会导致冲突。

4.9.5 组播转发机制

在组播模型中，IP 报文的地址字段为组播组地址，组播源向以此目的地址所标识的主机群组传送信息。因此，转发路径上的组播路由器为将组播报文传送到各个方位的接收站点，往往需要将从一个入接口收到的组播报文转发到多个出接口。

为保证组播报文在网络中的传输，必须依靠单播路由表或者单独提供给组播使用的路由表（如 MBGP 路由表）来指导转发；

为处理同一设备在不同接口上收到来自不同对端的相同组播信息，需要对组播报文的入接口进行 RPF（Reverse Path Forwarding，逆向路径转发）检查，以决定转发还是丢弃该报文。RPF 检查机制是大部分组播路由协议进行组播转发的基础。

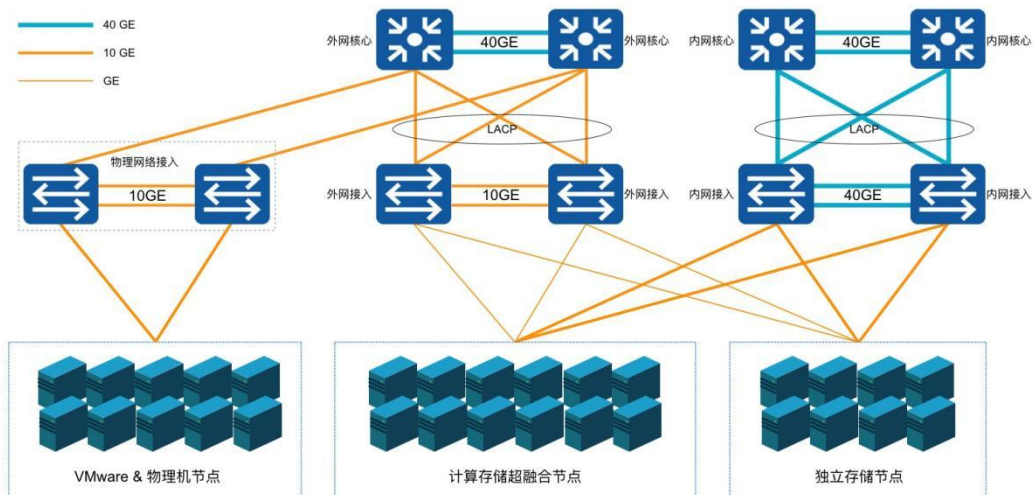
4.10 外网 IP

外网弹性 IP（Elastic IP Address，简称 EIP），是平台为用户的虚拟机、NAT 网关、VPN 网关及负载均衡等虚拟资源提供的外网 IP 地址，为虚拟资源提供平台 VPC 网络外的网络访问能力，如互联网或 IDC 数据中心物理网络，同时外部网络也可通过 EIP 地址直接访问平台 VPC 网络内的虚拟资源。

EIP 资源支持独立申请和拥有，用户可通过控制台或 API 申请 IP 网段资源池中的 IP 地址，并将 EIP 绑定至虚拟机、NAT 网关、负载均衡、VPN 网关上，为业务提供外网服务通道。

4.10.1 物理架构

在私有云平台中，允许平台管理员自定义平台外网 IP 资源池，即由平台管理员自定义平台访问外网的方式，外网 IP 网段资源池在添加至云平台前，需要通过物理网络设备下发至计算节点连接的交换机端口。



如上图物理架构示意图所示，所有计算节点需要连接网线至物理网络的外网接入交换机，并在物理网络的交互机上配置所连接端口允许透传 Vlan 的网络访问方式，使运行在计算节点上虚拟机可通过外网物理网卡直接与外部网络进行通信：

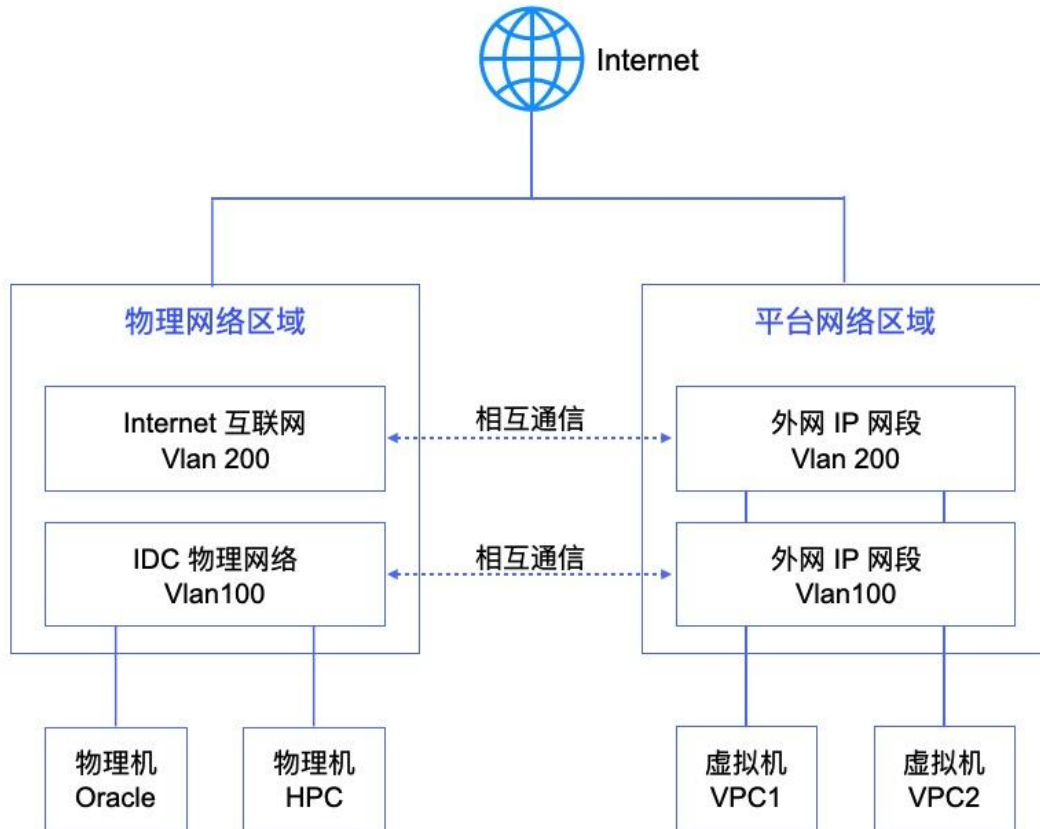
- 若通过外网 IP 访问互联网，需要物理网络设备上将自定义的外网 IP 网段配置为可直通或 NAT 到互联网；
- 若通过外网 IP 访问 IDC 数据中心的物理网络，需要在物理网络设备上将自定义的外网 IP 网段配置为可与 IDC 数据中心网络通信，如相同的 Vlan 或 Vlan 间打通等。

物理网络架构为高可用示意图，实际生产环境架构可进行调整，如内外网接入交换机可合并为一组高可用接入交换机，通过不同的 Vlan 区分内外网等。

4.10.2 逻辑架构

物理网络架构及配置确认后，在平台层面需要分别添加互联网 IP 网段和 IDC

物理网段至云平台 IP 网段资源池中，租户可申请不同网段的 EIP 地址，并将通往不同网络的 EIP 地址绑定至虚拟机默认外网网卡，使虚拟机可通过外网 IP 地址同时访问互联网和 IDC 数据中心物理网络。



如逻辑架构图所示，用户在平台中分别添加通往 Internet(Vlan200)和通往 IDC 物理网络 (Vlan100) 的网段至云平台。网段举例如下：

- Vlan200 的网段为 106.75.236.0/25，配置下发默认路由，即虚拟机绑定网段的 EIP 将会自动下发目标地址为 0.0.0.0/0 的默认路由；
- Vlan100 的网段为 192.168.1.0/24，仅下发当前网段路由，即虚拟机绑定网段的 EIP 仅下发目标地址为 192.168.1.0/24 的指定路由。

租户可分别申请 Vlan200 和 Vlan100 的 EIP 地址，并可将两个 EIP 同时绑定至虚拟机。平台会将 EIP 地址及下发路由直接配置至虚拟机外网网卡，并通过 SDN 控制器下发流表至虚拟机所在的物理机 OVS，物理机 OVS 通过与物理机外网网卡接口及交换机进行互联，通过交换机设备与互联网或 IDC 物理网络进行通信。

当虚拟机需要访问互联网或物理网络时，数据会通过虚拟机外网网卡直接透传至物理机的 OVS 虚拟交换机，并通过 OVS 流表将请求转发至物理机外网网卡及物理交换机，经由物理交换机的 Vlan 或路由配置将数据包转发至互联网或 IDC 物理网络区域，完成通信。

如上图 VPC1 网络的虚拟机同时绑定了 Vlan100 和 Vlan200 网段的 EIP 地址，Vlan100 EIP 为 192.168.1.2，Vlan200 EIP 为 106.75.236.2。平台会直接将两个 IP 地址直接配置至虚拟机的外网网卡，通过虚拟机操作系统可直接查看配置到外网网卡的 EIP 地址；同时自动下发两个 IP 地址所属网段需要下发的路由到虚拟机操作系统中，虚拟机的默认路由指定的下一跳为 Vlan200 互联网网段的网关，使虚拟机可通过 106.75.236.2 IP 地址与互联网进行通信，通过 192.168.1.2 与物理网络区域的 Oracle 及 HPC 高性能服务器进行内网通信。

整个通信过程直接通过虚拟机所在物理机的物理网卡进行通信，在物理网卡和物理交换机性能保障的前提下，可发挥物理网络硬件的最佳转发性能，提升虚拟机对外通信的转发能力。同时所有外网 IP 流量均可通过平台安全组在平台内进行流量管控，保证虚拟机访问平台外部网络的安全性。

4.10.3 EIP 通信模式

平台的外网 IP 与物理网络通信支持直通和 NAT-EIP 两种模式，适应多种访问外网的场景。

- **直通模式**

直接将 EIP 地址配置至虚拟机，在虚拟机中下发 EIP 网段的默认路由，通过虚拟网卡直接透传至物理机网络，与平台外网进行通信，可有效减少网络性能损耗，适于对 EIP 性能要求较高的应用场景。

- **NAT-EIP 模式**

虚拟机通过公共 VPC 网关 (NAT 网关) 与平台外网进行通信，该模式不会侵入虚拟机网络配置，即在虚拟机中不会进行外网 IP 地址及路由配置，适用于虚拟机仅提供单个 IP 地址向外提供服务的场景。

当使用 NAT-EIP 绑定模式绑定云主机时，云主机内部仅可查看虚拟机所属 VPC 网络的 IP 地址及相关信息。虚拟机通信时通过 VPC 网关上的 EIP 进行对外通信，VPC 网关可服务于该 VPC 下所有子网的虚拟机。

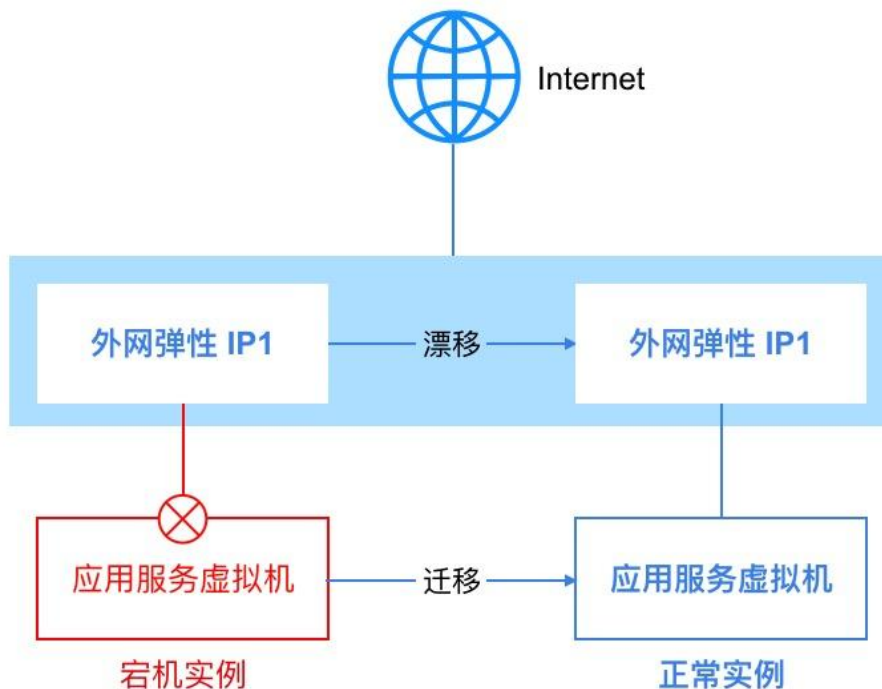
对比直通 EIP 直接绑定在云主机的模式，NAT-EIP 代理模式可避免部分应用无法处理多 IP 的路由信息，导致应用运行异常，提高虚拟机外网通信的兼容性。

当无需使用 NAT-EIP 时，可通过禁用 VPC 网关关闭 NAT-EIP 能力。对于已经创建的 VPC 网关，也可通过启用 VPC 网关使用 NAT-EIP，虚拟机可同时绑定直通模式外网 IP 和 NAT 模式外网 IP。

注意 虚拟机镜像特性不支持 qemu-ga 时，创建虚拟机不可申请直通 EIP，可在虚拟机创建完成后绑定 NATEIP，绑定 NATEIP 需要开启 VPC 网关。开启 VPC 网关会消耗平台 2C2G 计算资源。

4.10.4 功能特性

EIP 为浮动 IP，可随故障虚拟机恢复漂移至健康节点，继续为虚拟机或其它虚拟资源提供外网访问服务。



当一台虚拟机所在的物理主机发生故障时，智能调度系统会自动对故障主机上的虚拟机进行宕机迁移操作，即故障虚拟机会在其它健康的主机上重新拉起并提供正常业务服务。若虚拟机已绑定外网 IP，智能调度系统会同时将外网 IP 地址及相关流表信息一起漂移至虚拟迁移后所在的物理主机，并保证网络通信可达。

- 支持平台管理员自定义外网 IP 资源池，即自定义外网 IP 网段，并支持配置网段的路由策略。租户申请网段的外网 IP 绑定至虚拟资源后，下发目的路由地址的流量自动以绑定的外网 IP 为网络出口。
- 外网 IP 网段支持下发默认路由和指定路由，下发默认路由代表默认所有流量均以绑定的外网 IP 为出口，指定路由为管理员指定目的地址的流量以绑定的外网 IP 为出口。
- 提供 IPv4/IPv6 双栈能力，管理员可自定义管理 IPv4 和 IPv6 网段资源池，并支持同时绑定 IPv4/IPv6 地址到虚拟机，为虚拟机提供双栈网络通信服务。
- 支持外网 IP 网段的权限管控，可指定所有租户或部分租户使用，未被指定的租户无权限申请并使用网段 EIP。
- EIP 具有弹性绑定的特性，支持随时绑定至虚拟机、NAT 网关、负载均衡、VPN 网关等虚拟机资源，并可随时解绑绑定至其它资源。
- 虚拟机支持绑定 50 个外网 IPv4 和 10 个外网 IPv6 地址，以第一个有默认路由的外网 IP 作为虚拟机的默认网络出口。
- 提供外网 IP 网段获取服务，支持租户手动指定 IP 地址申请 EIP，并提供 IP 地址冲突检测，方便用户业务网络地址规划。
- 平台管理员可自定义外网 IP 网段的带宽规格，租户可在带宽规格范围内配置外网 IP 的带宽上限。
- 仅支持 QEMU-Agent 机器绑定直通模式弹性外网 IPv6。

外网 IP 具有数据中心属性，仅支持绑定相同数据中心的虚拟资源。用户可

通过平台自定义申请 EIP，并对 EIP 进行绑定、解绑、调整带宽等相关操作。

4.11 高可用 VIP

4.11.1 概述

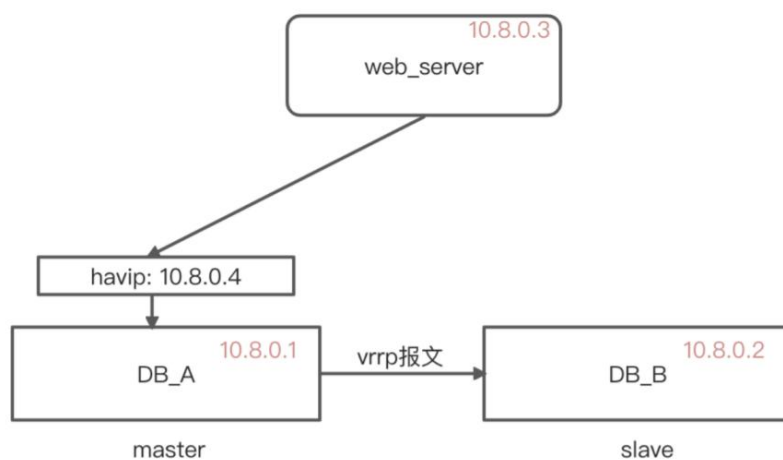
高可用 VIP (High available Virtual IP Address, 简称 HAVIP)，高可用虚拟 IP 地址，是归属于 VPC 内某个子网内的可漂移内网 IP，用户可将 HAVIP 与高可用服务结合，以便在服务出现故障时进行服务入口的漂移，以实现服务的高可用。

用户可通过 API 接口或控制台申请高可用 VIP，用于服务的高可用，创建高可用 VIP 前需保证账户至少拥有一个 VPC 网络和子网。

支持高可用 VIP 的申请、更新、删除等管理操作，支持更新 VIP 关联虚拟机，并支持替换、删除、新增虚拟机。

4.11.2 工作机制

HAVIP 作为一个不绑定特定设备的浮动 IP，通常和高可用软件 (Keepalived、Heartbeat、Failover Cluster) 配合使用，用于搭建高可用主备集群，如 HA 负载均衡、主备版数据库等。本文以 Keepalived 为例介绍 HAVIP 的工作原理，示例图如下：



- Master 和 Slave 均安装 Keepalived，配置从控制台申请出来的 HAVIP

为 VRRP VIP，分别设置优先级（priority 值）。

- Keepalived 中的 VRRP 协议通过对比两台虚拟机的初始优先级大小，选举出 Master 服务器。
- Master 服务器向外发送 ARP 报文，宣告 VIP，实现 VIP 和 MAC 的地址映射更新（arp 缓存）。
- 宣告 VIP 生效后，真正对外提供服务的服务器为 Master 服务器，通信的内网 IP 为 HAVIP。
- Master 服务器周期性发送 VRRP 报文给 Slave 服务器。若 Master 服务器异常，Slave 服务器在一定时间内没有收到 VRRP 报文，则会将自己设置为 Master，并对外发送 ARP 更新（GARP），报文携带自己的 MAC 地址。
- Slave 服务器将作为 Master 服务器对外提供通信服务，外部访问的报文将转发至 Slave 处理，直至实现 Realserver 切换。

4.12 NAT 网关

4.12.1 产品概述

NAT 网关（NAT Gateway）是一种类似 NAT 网络地址转换协议的 VPC 网关，为云平台资源提供 SNAT 和 DNAT 代理，支持互联网或物理网地址转换能力。平台 NAT 网关服务通过的 SNAT 和 DNAT 规则分别实现 VPC 内虚拟资源的 SNAT 转发和 DNAT 端口映射功能。

- **SNAT 规则**

通过 SNAT 规则实现 VPC 级、子网级及虚拟资源实例级的 SNAT 能力，使不同维度的资源通过 NAT 网关访问外网。

- **DNAT 规则**

通过 DNAT 规则，可配置基于 TCP 和 UDP 两种协议的端口转发，将 VPC

内的云资源内网端口映射到 NAT 网关所绑定的外网 IP，对互联网或 IDC 数据中心网络提供服务。

作为一个虚拟网关设备，需要绑定外网 IP 作为 NAT 网关的 SNAT 规则出口及 DNAT 规则的入口。NAT 网关具有地域（数据中心）属性，仅支持相同数据中心下同 VPC 虚拟资源的 SNAT 和 DNAT 转发服务，

虚拟机通过 NAT 网关可访问的网络取决于绑定的外网 IP 所属网段在物理网络上的配置，若所绑定的外网 IP 可通向互联网，则虚拟机可通过 NAT 网关访问互联网；若所绑定的外网 IP 可通向 IDC 数据中心的物理网络，则虚拟机通过 NAT 网关访问 IDC 数据中心的物理网络。

4.12.2 应用场景

用户在平台使用虚拟机部署应用服务时，有访问外网或通过外网访问虚拟机的应用场景，通常我们会在每一台虚拟机上绑定一个外网 IP 用于和互联网或 IDC 数据中心网络进行通信。真实环境和案例中，可能无法分配足够的公网 IP，即使公网 IP 足够也无需在每一台需要访问外网的虚拟机上绑定外网 IP 地址。

- **共享 EIP**

通过 SNAT 代理，使多台 VPC 内网虚拟机共享 1 个或多个外网 IP 地址访问互联网或 IDC 数据中心的物理网络。

- **屏蔽真实 IP**

通过 SNAT 代理，多台 VPC 内网虚拟机使用代理 IP 地址通信，自动屏蔽真实 IP 内网地址。

- **VPC 内网虚拟机提供外网服务**

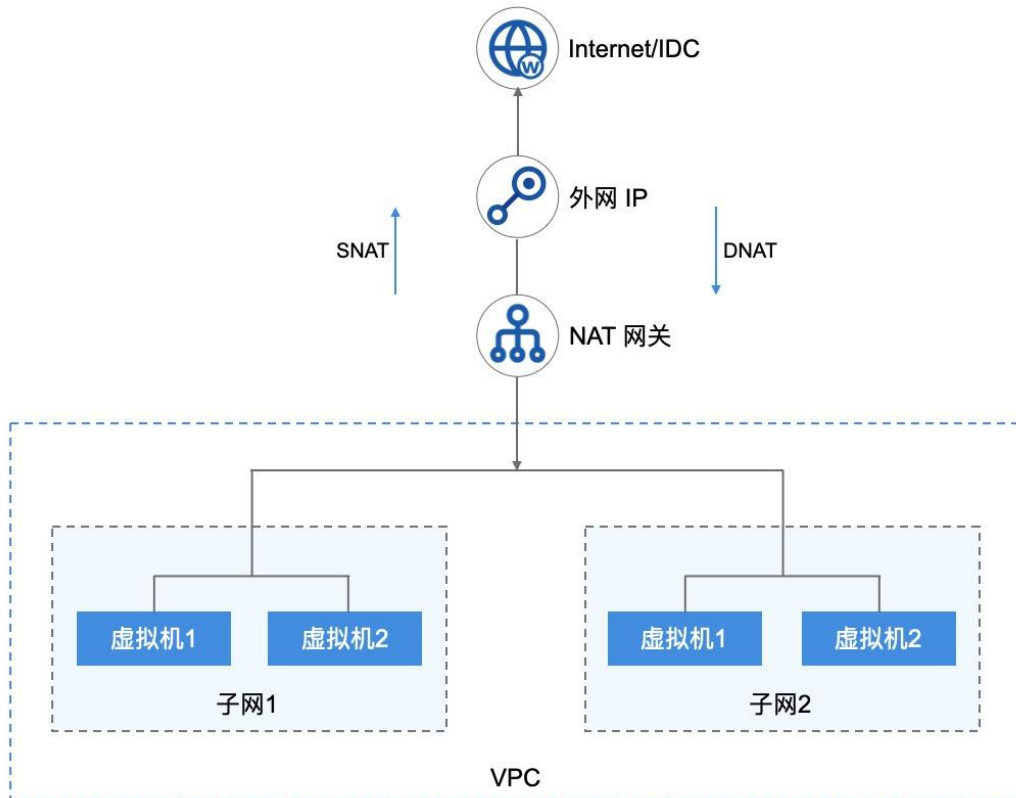
通过 DNAT 代理，配置 IP 及端口转发，对互联网或 IDC 数据中心的网络提供业务服务。

4.12.3 架构原理

平台产品服务底层资源统一，NAT 网关实例为主备高可用集群架构，可实

现 NAT 网关故障自动切换，提高 SNAT 和 DNAT 服务的可用性。同时结合外网 IP 地址，根据 NAT 配置为租户虚拟资源提供 SNAT 和 DNAT 代理。

在产品层面，租户通过申请一个 NAT 网关，指定 NAT 网关可允许通信的子网，通过绑定【外网 IP】使多子网下虚拟机与互联网或 IDC 数据中心物理网进行通信，具体逻辑架构图如下：



- 平台支持同 VPC 多子网虚拟机使用 NAT 网关访问互联网或 IDC 数据中心网络。
- 当多个子网中未绑定外网 IP 的虚拟机关联 NAT 网关时，平台将自动在虚拟机中下发访问外网的路由。
- 虚拟机通过下发的路由，将访问外网的数据通过 NAT 网关透传至已绑定的【外网 IP】。
- 透传至外网 IP 的数据通过平台 OVS 及物理网卡将数据包发送至物理交换机，完成数据 SNAT 的通信。
- 当外网需要访问 VPC 中的虚拟机服务时，可通过 NAT 网关端口转发，

使互联网或 IDC 物理网通过 NAT 网关已绑定的 IP+端口访问 VPC 内网服务。

注意 NAT 网关实例底层由虚拟机进行构建，虚拟机配置为 2C4G。

4.12.4 功能特性

云平台提供高可用 NAT 网关服务，并支持网关的全生命周期管理，包括外网 IP、SNAT 规则及 DNAT 端口转发及监报告警，同时为 NAT 网关提供网络及资源隔离的安全保障。

一个 VPC 允许创建 20 个 NAT 网关，相同 VPC 下所有 NAT 网关中 SNAT 规则不可重复，即 20 个 NAT 网关中的 SNAT 规则不允许重复。场景举例：

- 当 NATGW(VPC: 192.168.0.0/16) 中创建了子网 (192.168.0.1/24) 的 SNAT 规则，则相同 VPC 下 NATGW 不可在创建子网 (192.168.0.1/24) 为源地址的 SNAT 规则，当 NATGW01 中该子网规则删除后，才可进行创建。
- 当 NATGW(VPC: 192.168.0.0/16) 中创建了 VPC 级别的规则，则相同 VPC 下不可在创建 VPC 级别的规则。
- 当 NATGW(VPC: 192.168.0.0/16) 中创建了虚拟机 (192.168.1.2) 的 SNAT 规则，则相同 VPC 下 NATGW 不可在创建虚拟机 (192.168.1.2) 为源地址的 SNAT 规则。

4.12.4.1 SNAT 规则

NAT 网关通过 SNAT 规则支持 SNAT (Source Network Address Translation 源地址转换) 能力，每条规则由源地址和目标地址组成，即将源地址转换为目标地址进行网络访问。平台 SNAT 规则支持多种场景的出外网场景，即源地址包括 VPC、子网、虚拟机三种类型：

- **VPC 级别**

指 NAT 网关所属 VPC 下的所有虚拟机可通过 NAT 网关访问外网。

- **子网级别**

指 NAT 网关所属 VPC 下被指定子网中的所有虚拟机可通过 NAT 网关访问外网。

- **虚拟机级别**

指 NAT 网关所属 VPC 下被指定的虚拟机才可通过 NAT 网关访问外网。

通常只需要指定一条 VPC 类型的 SNAT 规则, 即可实现 NAT 网关所属 VPC 网络下所有虚拟机访问外网的能力, 规则内的源地址资源必须与 NAT 网关处于相同的 VPC 网络。

规则的目标地址为 NAT 网关绑定的外网 IP 地址, 通过规则策略即可将源地址在 VPC、子网、虚拟机的 IP 地址转换为网关绑定的外网 IP 进行网络通信, 即通过 SNAT 规则虚拟机可在不绑定外网 IP 的情况下与平台外网进行通信, 如访问 IDC 数据中心网络或互联网。

SNAT 规则中不同源地址类型的规则优先级不同, 以优先级高的规则为准:

(1) 源地址为 VPC

- NAT 网关所属 VPC 下所有虚拟机均可通过 NAT 网关访问外网。

(2) 源地址为子网 CIDR

- 子网下虚拟机可通过 NAT 网关访问外网。
- 每个子网仅可创建一条 SNAT 规则, 不允许重复。
- 支持为子网下虚拟机单独配置 SNAT 规则, 优先级高于源地址为子网的 SNAT 规则。

(3) 源地址为虚拟机 IP

- 虚拟机可通过 NAT 网关访问外网。
- 每个虚拟机 IP 仅可创建一条 SNAT 规则, 不允许重复。

- 源地址为虚拟机 IP 的 SNAT 规则优先级高于源地址为子网的 SNAT 规则。
- SNAT 规则的目标地址可以为 NAT 网关已绑定的外网 IP。

注意 一个 NAT 网关默认可创建 100 条 SNAT 规则。

用户配置 SNAT 规则后，NAT 网关会自动下发默认路由至源地址匹配的虚拟机，使虚拟机通过 SNAT 规则的外网 IP 访问外网。具体通信逻辑如下：

- 虚拟机未绑定 IPv4 外网 IP，则默认通过 NAT 网关访问外网。
- 虚拟机已绑定 IPv4 外网 IP 且存在默认网络出口，则通过虚拟机默认网络出口访问外网。
- 虚拟机已绑定 IPv4 外网 IP 且无默认网络出口，则通过 NAT 网关访问外网。

虚拟机通过 NAT 网关访问外网时，使用的外网 IP 取决于 SNAT 规则的配置，会将 IP 地址作为虚拟机的出口。

4.12.4.2 DNAT 规则

NAT 网关支持 DNAT（Destination Network Address Translation 目的地址转换），也称为端口转发或端口映射，即将外网 IP 地址转换为 VPC 子网的 IP 地址提供网络服务。

支持 TCP 和 UDP 两种协议的端口转发，支持对端口转发规则进行生命周期管理。用户可通过端口转发为 NAT 网关配置端口映射，将 VPC 子网内虚拟机内网端口映射到 NAT 网关的外网 IP，使虚拟机可对外网提供服务。

每条规则由协议、源 IP（外网 IP）、端口、目的 IP（虚拟机 IP）、目的端口五元组组成，即将源 IP 的端口请求转发至目的 IP 的端口，使用户直接通过源 IP 地址访问 VPC 内网虚拟机提供的服务。

- 协议：指 DNAT 端口转发规则的转发协议，支持 TCP 和 UDP，创建时

必须指定，默认为 TCP。

- 源 IP: DNAT 端口转发规则的源 IP 地址, 即 NAT 网关的所绑定的外网 IP, 一条规则仅支持一个外网 IP。
- 源端口: DNAT 端口转发规则的源端口, 即 NAT 网关所绑定的外网 IP 暴露出来的端口。仅支持指定未创建的源端口, 相同协议下不支持重复的源端口规则。
- 目的 IP: DNAT 端口转发规则的目的 IP, 即 NAT 网关所属 VPC 网络下虚拟机的内网 IP 地址。目的 IP 地址不受 SNAT 规则限制, 即一台虚拟机可同时添加 SNAT 规则和 DNAT 规则。
- 目的端口: DNAT 端口转发规则的目的端口, 即目的 IP 虚拟机对外提供服务的端口。
 - 端口范围为 1~65535。
 - 目的端口可与源端口相同或不同, 如源端口为 TCP:80, 目的端口为 TCP:8080, 即代表将源 IP 地址的 TCP 80 端口流量转发至目的 IP 地址 TCP 8080 端口。
 - 支持创建同一目的 IP 重复的目的端口, 如将两个源地址为的 80 端口均转发至同一个目的地址的相同端口进行业务数据处理。

相同协议情况, 不支持重复的源端口规则。同时 DNAT 规则支持多端口映射规则, 即支持指定源端口为连续范围, 如 1024~1030。指定端口范围时, 目的端口范围的数量必须与源端口一致。

NAT 网关绑定外网 IP 时, 端口转发规则为 VPC 子网内的虚拟机提供互联网外网服务, 可通过外网访问子网内的虚拟机服务。

4.12.4.3 网关外网 IP

NAT 网关支持绑定默认路由类型的 IPv4 外网 IP 地址, 为 NAT 网关指定子网的虚拟资源提供共享的外网 IP 资源池, 以提供更加灵活便捷的 SNAT 及 DNAT

能力。

支持为 NAT 网关外网 IP 地址的绑定和解绑，绑定后可用于创建 SNAT 和 DNAT 规则，使资源通过指定的外网 IP 访问外网，并通过外网 IP 端口访问 VPC 内的虚拟资源。

支持将外网 IP 从 NAT 网关解绑，解绑后相关 SNAT 和 DNAT 的目标/源 IP 将被置空，可重新进行绑定及修改。

用户可通过外网 IP 管理查看 NAT 网关已绑定的外网 IP 地址及信息，同时支持对 NAT 网关的外网 IP 进行绑定和解绑操作。

NAT 网关不支持绑定 IPv6 及非默认路由类型的 IPv4 外网 IP 地址。

4.12.4.4 监报告警

平台支持对 NAT 网关进行监控数据的收集和展示，通过监控数据展示每一个 NAT 网关的指标数据，同时支持为每一个监控指标设置阈值告警及通知策略。支持的监控指标包括网络出/带宽、网络出/包量及连接数。

支持查看一个 NAT 网关多时间维度的监控数据，包括 1 小时、6 小时、12 小时、1 天、7 天、15 天及自定义时间的监控数据。默认查询数提成为 1 小时的数据，最多可查看 1 个月的监控数据。

4.12.4.5 NAT 网关高可用

NAT 网关实例支持高可用架构，即至少由 2 个虚拟机实例构建，支持双机热备。支持在线将单机版的 NAT 网关实例升级至主备版，启用 NAT 网关实例的高可用。

当一个 NAT 网关的实例发生故障时，支持自动在线切换到另一个虚拟机实例，保证 NAT 代理业务正常。同时基于外网 IP 地址的漂移特性，支持在物理机宕机时，保证 SNAT 网关出口及 DNAT 入口的可用性。

4.12.4.6 NAT 网关安全

NAT 网关的网络访问控制可以关联安全组给予安全保障，通过安全组的规则可控制到达 NAT 网关所绑定外网 IP 的入站流量及出站流量，支持 TCP、UDP、ICMP 等协议数据包的过滤和控制。

安全组及安全组的规则支持对已关联安全组的 NAT 网关的流量进行限制，仅允许安全组规则内的流量透传安全组到达目的地。为保证 NAT 网关的资源和网络安全，平台为 NAT 网关提供资源隔离及网络隔离机制：

(1) 网关高可用

支持用户将单机版 NAT 网关升级为主备版，满足 NAT 网关高可用场景。当其中一个网关的虚拟机实例因故障宕机时，备实例会自动转换为主实例持续提供 NAT 网关的 SNAT 和 DNAT 服务。

(2) 资源隔离

- NAT 网关具有数据中心属性，不同数据中心间 NAT 网关资源物理隔离；
- NAT 网关资源在租户间相互隔离，租户可查看并管理账号及子账号下所有 NAT 网关资源；
- 一个租户内的 NAT 网关资源，仅支持绑定租户内同数据中心的 VPC 子网资源；
- 一个租户内的 NAT 网关资源，仅支持绑定租户内同数据中心的外网 IP 资源；
- 一个租户内的 NAT 网关资源，仅支持绑定租户内同数据中心的安全组资源。

(3) 网络隔离

- 不同数据中心间 NAT 网关资源网络相互物理隔离；
- 同数据中心 NAT 网关网络采用 VPC 进行隔离，不同 VPC 的 NAT 网关资源无法相互通信；

- NAT 网关绑定的外网 IP 网络隔离取决于用户物理网络的配置，如不同的 Vlan 等。

4.13 负载均衡

4.13.1 产品概述

负载均衡 (Load Balance) 是由多台服务器以对称的方式组成一个服务器集合，每台服务器都具有等价的地位，均可单独对外提供服务而无须其它服务器的辅助。平台负载均衡服务 (简称 LB—Load Balance) 是基于 TCP/UDP/HTTP/HTTPS 协议将网络访问流量在多台虚拟机间自动分配的控制服务，类似于传统物理网络的硬件负载均衡器。

通过平台负载均衡服务提供的虚拟服务地址，将相同数据中心、相同 VPC 网络的虚拟机添加至负载均衡转发后端，并将加入的虚拟机构建为一个高性能、高可用、高可靠的应用服务器池，根据负载均衡的转发规则，将来自客户端的请求均衡分发给服务器池中最优的虚拟机进行处理。

支持内外网两种访问入口类型，分别提供 VPC 内网和 EIP 外网的负载访问分发，适应多种网络架构及高并发的负载应用场景。提供四层和七层协议的转发能力及多种负载均衡算法，支持会话保及健康检查等特性，可自动隔离异常状态虚拟机，同时提供 SSL Offloading 及 SSL 证书管理能力，有效提高整体业务的可用性及服务能力。

LB 支持收集并展示负载流量各种网络指标的监控数据，并可根据告警模板进行监控报警及通知，保证业务的正常运行。当前负载均衡为接入的虚拟机服务池提供基于 NAT 代理的请求分发方式，在 NAT 代理模式下，所有业务的请求和返回数据都必须经过负载均衡，类似 LVS 的 NAT 工作模式。

4.13.2 应用场景

平台提供外网和内网两种类型的负载均衡服务，分别对应外网服务和内网服务两种场景。用户可根据业务需求，选择创建对外公开或对内私有的负载均衡实

例，平台会根据负载均衡类型分别分配外网 IP 地址或 VPC 私有网络的 IP 地址，即负载均衡的服务访问地址。

(1) 外网类型的负载均衡使用场景

部署在平台的业务服务需要构建高可用虚拟机集群，且需对互联网提供统一访问入口。

部署在平台的业务服务需要构建高可用虚拟机集群，且需对 IDC 数据中心提供统一访问入口。

(2) 内网负载均衡使用场景

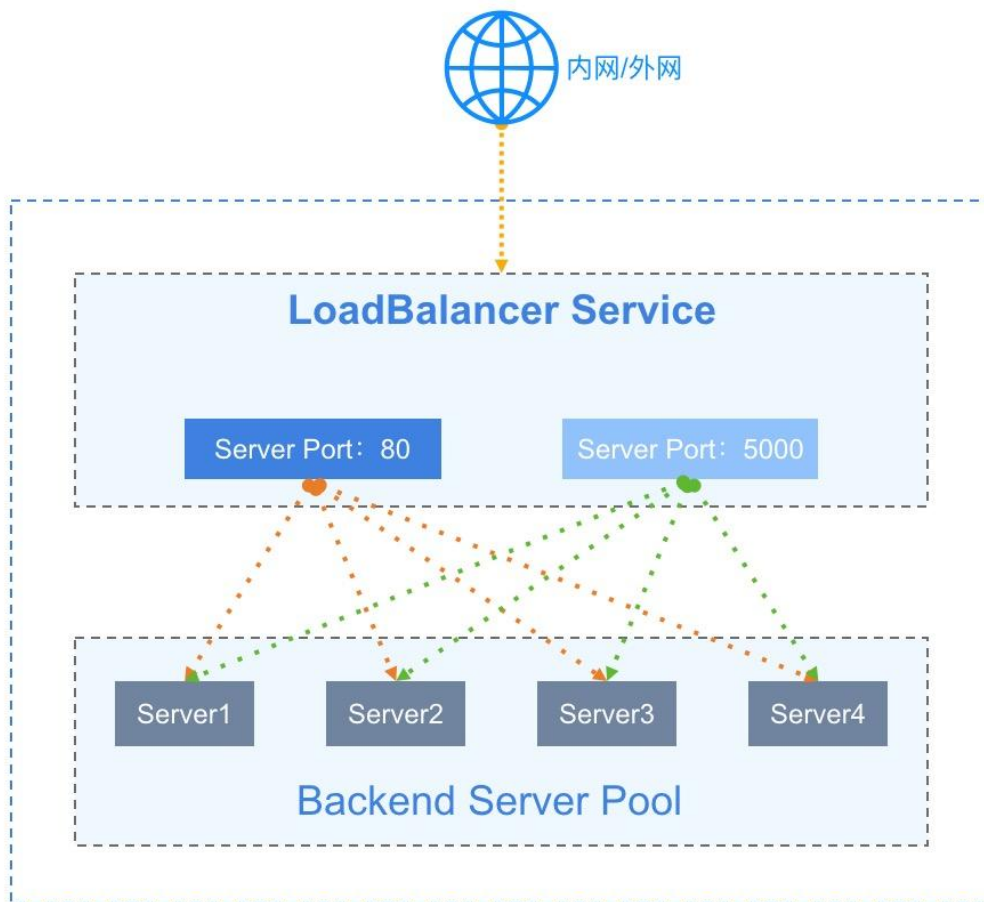
部署在平台的业务服务需要构建高可用虚拟机集群，且仅需对 VPC 内网提供统一访问入口。

部署在 VPC 私有网络的虚拟机集群需要对其它用户或服务屏蔽真实 IP 地址，对客户端提供透明化服务。

用户也可将负载均衡服务分配的 IP 地址与自有域名绑定在一起，通过域名访问后端应用服务。

4.13.3 架构原理

一个提供服务的负载均衡，主要由 LB 实例 (LoadBalancer)、虚拟服务器 (VServer)、后端服务器 (Backend Real Server) 三部分组成。如架构图所示：



- **LoadBalancer (LB)**：负载均衡实例为主备高可用集群架构，可实现负载均衡器故障自动切换，提高接入负载均衡服务的可用性。同时结合内外网 IP 地址，根据 VServer 配置的监听器，将虚拟机加入到 Backend 成为 Real Server，以实现业务的流量均衡与服务容错。
- **Virtual Server (VServer)**：监听器，每个监听器是一组负载均衡的监听端口配置，包含协议、端口、负载算法、会话保持、连接空闲超时及健康检查等配置项，用于分发和处理访问 LB 的请求。
- **Backend Server Pool**：后端一组虚拟机服务器池，实际处理请求的真实服务器 (RealServer)，即真实部署业务的虚拟机实例。
- **外网 IP (EIP)**：外网弹性 IP 地址，绑定至外网类型的 LB 实例上，对互联网或 IDC 数据中心提供业务负载均衡访问入口。
- **内网 IP (Private IP)**：内网 IP 地址，内网类型 LB 实例提供服务的访

问地址，通常是由创建内网负载均衡器时指定的 VPC 自动分配。

负载均衡器用于承载 VServer 及访问入口，VServer 负责访问入口地址的端口监听及请求分发。当负载均衡器接受到来自客户端的请求后，会通过一系列负载均衡算法，将访问请求路由分发到后端虚拟机服务器池进行请求处理，同时由 VServer 将处理结果返回给客户端。

- 通过加权轮询、最小连接数及基于源地址的负载均衡调度策略，进行业务请求流量转发，满足多场景业务负载需求，如加权轮询是按照后端服务器的权重进行请求转发，权重越大转发的请求越多。
- 通过会话保持机制，在请求会话的生命周期内，会将来自同一个客户端的会话转发至同一个虚拟机进行处理，适用于 TCP 长连接等应用场景。
- 通过健康检查机制，监控 RealServer 的运行状况及业务可用性，确保只将流量分发至业务健康的虚拟机。当后端虚拟机业务不可访问时，调度器会停止向虚拟机分发负载流量；待虚拟机业务恢复正常后，会将虚拟机重新加入至 VServer 后端并分发流量至虚拟机。

负载均衡器的工作模式为 NAT 请求代理，请求和返回均由负载均衡器进行转发和处理，即后端 RealServer 虚拟机处理请求后，会将请求返回给负载均衡，由负载均衡将结果返回给客户端。

负载均衡服务的负载均衡实例底层由虚拟机进行构建，默认为 2C2G 配置，支持用户对负载均衡的 CPU 配置进行变更，可变更的范围为 1C、2C、4C、8C。

4.13.4 功能特性

4.13.4.1 基础功能

平台负载均衡服务提供四层和七层转发能力，支持内网和外网两种网络入口，在多种负载调度算法基础之上支持健康检查、会话保持、连接空闲超时、内容转发及 SSL Offloading 和 SSL 证书管理等功能，保证后端应用服务的可用性和可靠性。

负载均衡实例所在节点的集群类型，由平台管理员自定义，如 x86 机型和 ARM 机型，通过 ARM 机型创建的实例为 ARM 版负载均衡实例，已适配国产芯片、服务器及操作系统。

- 支持内网和外网两种类型负载均衡器，满足 VPC 内网、IDC 数据中心及互联网服务负载均衡应用场景。
- 提供四层和七层业务负载分发能力，支持基于 TCP、UDP、HTTP 及 HTTPS 协议的监听及请求转发。
- 支持加权轮询、最小连接数和基于源地址的的负载调度算法，满足不同场景的流量负载业务。
 - 加权轮询：基于权重的轮询调度，负载均衡器接收到新的访问请求后，根据用户指定的权重，按照权重概率分发流量至各后端虚拟机，进行业务处理；
 - 最小连接数：基于后端服务器最小连接数进行调度，负载均衡器接收到新的访问请求后，会实时统计后端服务器池的连接数，选择连接数最低的虚拟机建立新的连接并进行业务处理；
 - 源地址：基于客户端源 IP 地址的调度策略，采用哈希算法将来源于相同 IP 地址的访问请求均转发至一台后端虚拟机进行处理。
- **监控数据**：负载均衡级别提供每秒连接数、每秒出/入流量、每秒出/入包数量的监控及告警；VServer 监听器级别提供连接数、HTTP 2XX、HTTP 3XX、HTTP 4XX、HTTP 5XX 等监控数据及告警。
- **安全控制**：通过安全组对外网负载均衡的访问进行安全管控，仅允许安全组规则内的流量透传负载均衡到达后端真实服务器，保证业务负载的安全性。

4.13.4.2 会话保持

提供会话保持功能，在会话生命周期内，保证同一个客户端的请求转发至同

一台后端服务节点上。四层和七层分别采用不同的方式进行会话保持。

针对 UDP 协议，基于 IP 地址保证会话保持，将来自同一 IP 地址的访问请求转发到同一台后端虚拟机进行处理，支持关闭会话 UDP 协议的会话保持；

针对 HTTP 和 HTTPS 协议，提供 Cookie 植入的方式进行会话保持，支持自动生成 KEY 和自定义 KEY。自动生成 KEY 是由平台自动生成 Key 进行植入，自定义 Key 是由用户自定义 Key 进行植入。

4.13.4.3 连接空闲超时

支持 TCP、HTTP 及 HTTPS 协议的连接空闲超时配置，自动中断在超时时间内一直无访问请求的连接。

客户端向 LB 地址发送的请求，在平台会维护两个连接，一个由客户端到 LB，一个由 LB 到后端虚拟机；

连接空闲超时是指由客户端到 LB 的连接空闲超时，若在超时周期内没有发送或接收任何数据，将自动中断从客户端到 LB 的连接；

默认连接空闲超时周期为 60 秒，即在建立连接后的 60 秒内一直没有新的数据请求，将自动中断连接。

4.13.4.4 健康检查

支持端口检查和 HTTP 检查，根据规则对后端业务服务器进行业务健康检查，可自动检测并隔离服务不可用的虚拟机，待虚拟机业务恢复正常后，会将虚拟机重新加入至后端并分发流量至虚拟机。

- 端口检查：针对四层和七层负载均衡，支持按 IP 地址+端口的的方式探测后端服务节点的健康状况，及时剔除不健康的节点；
- HTTP 检查：针对七层负载均衡，支持按 URL 路径和请求 HOST 头中携带的域名进行健康检查，筛选健康节点。

4.13.4.5 七层高级特性

- **内容转发**: 针对七层 HTTP 和 HTTPS 协议的负载均衡, 支持基于域名和 URL 路径的流量分发及健康检查能力, 可将请求按照域名及路径转发至不同的后端服务节点, 提供更加精准的业务负载均衡功能。
- **SSL 证书**: 针对 HTTPS 协议, 提供统一的证书管理服务和 SSL Offloading 能力, 并支持 HTTPS 证书的单向和双向认证。SSL 证书部署至负载均衡, 仅在负载均衡上进行解密认证处理, 无需上传证书到后端业务服务器, 降低后端服务器的性能开销。
- **获取监听器协议**: HTTP 监听器支持附加 HTTP header 字段, 通过 X-Forwarded-Proto 获取监听器的协议。
- **附加 HTTP HOST**: HTTP 监听器支持附加 HTTP header 字段, 通过 Host 附加 HOST 域名至 HTTP 请求中, 用于适配需要检测 HTTP 头 HOST 字段的业务。
- **HTTP 重定向 HTTPS**: 支持将 HTTP 访问重定向至 HTTPS, HTTPS 是加密数据传输协议, 安全性较高。
- **访问日志**: 支持 7 层负载均衡开启日志管理功能, 转储周期 5 分钟, 日志默认保留 6 个月, 过期自动删除。

4.13.4.6 获取客户端真实 IP

HTTP 获取客户端真实 IP: HTTP 监听器支持附加 HTTP header 字段, 通过 X-Forwarded-For 和 X-Real-IP 获取客户端真实 IP 地址。TCP 支持通过 Proxy Protocol 获取客户端真实地址, 确保您的后端服务节点支持 Proxy Protocol 即可。

TCP 获取客户端真实 IP: TCP 监听器采用 Nginx 官方的 Proxy-Protocol 方案。使 LB TCP 监听收到客户端的请求后, 在转发请求至后端服务节点时, 将客户端的源 IP 地址封装在 TCP 请求数据包中, 发送给后端服务节点, 使服务端通过解析 TCP 数据包后即可获取客户端 IP 地址。

Proxy Protocol 是一种 Internet 协议，用于将连接信息从请求连接的源传送到请求连接的目的地，通过为 TCP 报文添加 Proxy Protocol 报头来获取客户端源 IP，因此需要后端服务节点做相应的适配工作，解析 Proxy Protocol 报头以获取客户端源 IP 地址。详见：[Proxy-Protocol 官网文档](#)。

4.13.5 负载均衡高可用

负载均衡为用户提供业务级别的高可用方案，可以将业务应用同时部署至多个虚拟机中，通过负载均衡和 DNS 域名的方案设置流量均衡转发，实现多业务级别的流量负载均衡。

当大并发流量通过负载均衡访问虚拟机业务时，可通过最小连接数、加权轮询等算法，将请求转发给后端最健壮的虚拟机进行处理，请通过负载均衡将请求结果返回给客户端，保证业务可用性和可靠性。

用户可通过智能 DNS 服务，将两个数据中心的负载均衡实例同时绑定至一个域名，使用 DNS 实现跨数据中心的业务容灾方案。

4.13.6 SSL 证书

负载均衡支持 HTTPS 负载转发及 SSL 证书装载能力，确保用户业务受到加密保护并得到权威机构的身份认证。针对 HTTPS 协议的服务器证书和客户端证书，平台提供统一的证书管理服务，包括证书的上传、绑定、删除操作。

证书无需上传到服务节点，解密处理在负载均衡上进行，降低后端服务器的 CPU 开销，即 HTTPS 协议的监听器仅实现客户端至负载均衡器的 HTTPS 请求和 SSL 加解密，负载均衡至后端服务节点依然采用 HTTP 协议转发请求。

在上传和创建证书前需确认需要上传的证书类型，包括服务器证书和客户端证书，并按照证书格式要求上传或输入证书内容至平台。

- 服务器证书

用户证明服务器的身份，HTTPS 检查服务器发送的证书是否是由自己信赖的中心签发。部署并配置于负载均衡服务器中，为负载均衡后端服务节点的

网站提供 SSL 服务器证书及验证。单向认证和双向认证均需要上传服务器证书和私钥内容。

● 客户端证书

客户端 CA 公钥证书用于验证客户端证书的签发者，HTTPS 双向认证中需验证客户端提供的证书，才可成功建立连接。网站服务器用 CA 证书验证客户端证书的签名，如果没有通过验证，则拒绝连接。仅在双向认证时需要上传客户端证书并绑定到 VServer 监听器。

证书具有地址 (数据中心) 属性，仅支持关联相同数据中心的负载均衡资源，若一个证书需要在多个数据中心同时使用，需要在多个数据中心同时创建并上传证书。

负载均衡 SSL 证书支持用户上传 .crt 和 .pem 格式的证书文件，当 SSL 证书被 VServer 监听器关联时，平台会自动读取文件中的证书内容并装载至负载均衡 VServer 监听器中，使用户 HTTPS 应用通过 SSL 证书进行加解密。

证书文件格式支持 Linux 环境下 PEM 或 CRT，不支持其他格式的证书，需进行证书格式转换才可上传。用户也可通过直接输入证书内容创建证书，在上传证书或输入证书内容前，需确保证书、证书链及私钥内容符合证书的格式要求。

- 若证书是 Root CA 机构颁发的唯一证书，则无需额外的证书，配置的站点即可被浏览器等访问设备认为可信
- 若证书是通过中级 CA 机构颁发的证书，则拿到的证书文件包含多份证书，需要人为将服务器证书与中间证书合并在一起填写或上传，俗称证书链。
- 在上传服务器证书时，需要用户同时上传证书的私钥内容。

4.13.7 LB 安全性

(1) 网关高可用

支持用户将单机版负载均衡升级为主备版，满足负载均衡网关高可用场景。

当其中一个负载均衡的虚拟机实例因故障宕机时，备实例会自动转换为主实例持续提供负载均衡服务。

(2) 资源隔离

- 负载均衡具有数据中心属性，不同数据中心间负载均衡资源物理隔离；
- 负载均衡资源在租户间相互隔离，租户可查看并管理账号及子账号下所有负载均衡资源；
- 一个租户内的负载均衡资源，仅支持绑定租户内同数据中心的 VPC 子网资源；
- 一个租户内的负载均衡资源，仅支持绑定租户内同数据中心的外网 IP 资源；
- 一个租户内的负载均衡资源，仅支持绑定租户内同数据中心的安全组资源。

(3) 网络隔离

- 不同数据中心间负载均衡资源网络相互物理隔离；
- 同数据中心负载均衡网络采用 VPC 进行隔离，不同 VPC 的负载均衡资源无法相互通信；
- 负载均衡绑定的外网 IP 网络隔离取决于用户物理网络的配置，如不同的 Vlan 等。

4.14 IPsecVPN 服务

4.14.1 背景

用户在使用云平台部署并管理应用服务时，会有部分业务部署于 IDC 数据中心环境的内网或第三方公/私有云平台上，如 Web 服务部署于公有云平台，应用和数据库等应用部署于私有云，构建公有云和私有云混合部署环境。

在混合云的应用场景中，可以通过专线的方式将两端网络的内网直接打

通，且较好的保证网络可靠性和性能。但由于专线成本较高，仅适用于部分对网络时延要求较高的业务，为节省成本并与第三方平台建立点对点的网络通信，云平台提供 VPN 网关-IPsecVPN 连接的服务能力，允许平台侧 VPC 子网的资源直接与第三方平台内网的主机进行通信，同时也可为平台不同 VPC 网络间提供连接服务。

4.14.2 概述

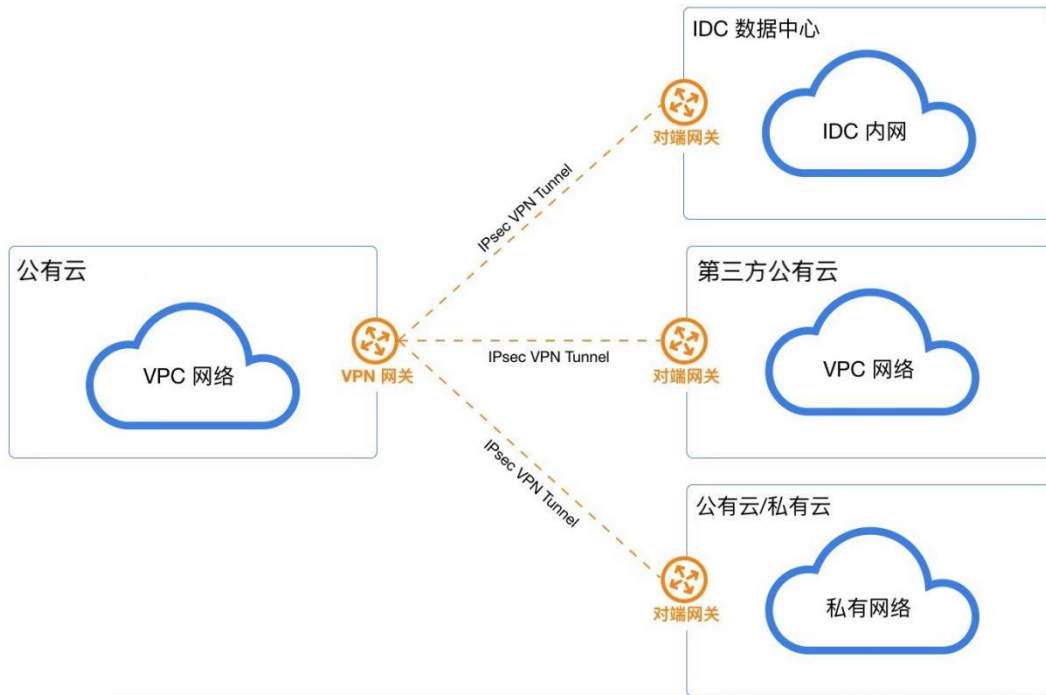
IPsec VPN 是一种采用 IPsec 协议加密的隧道技术，由 Internet Engineering Task Force ([IETF](#)) 定义的安全标准框架，在互联网上为两个私有网络提供安全通道，通过加密保证连接的安全。有关 IPsec 可参考 [RFC2409](#) (IKE—Internet Key Exchange 因特网密钥交换协议) 和 [RFC4301](#) (IPsec 架构)。

云平台 IPsecVPN 服务是基于 Internet 的网络连接服务，采用 IPsec (Internet Protocol Security) 安全加密通道实现企业数据中心、办公网络与平台 VPC 私有网络的安全可靠连接，同时也可使用 VPN 网关在 VPC 之间建立加密内网连接。网关服务为可容灾的高可用架构，同时支持用户选择多种加密及认证算法，并提供 VPN 连接健康检测及连接日志，保证隧道连接的可靠性、安全性及管理便捷性。

通过 IPsecVPN 服务，用户可将本地数据中心、企业分支机构与私有云平台的 VPC 私有网络通过加密通道进行连接，也可将用于不同 VPC 之间的加密连接。对端设备或系统仅需支持 IPsec 的 IKEv1 或 IKEv2，即可通过配置与平台的 VPN 网关进行互连，如通用网络设备或配置 IPsecVPN 的服务器。

4.14.3 逻辑架构

VPN 网关 IPsecVPN 服务由 VPN 网关、对端网关及 VPN 隧道连接三部分组成。



(1) VPN 网关

平台侧 VPC 网络建立 IPsecVPN 连接的出口网关，通过关联 VPC 和外网 IP 与对端网关的 IPsecVPN 进行连接，用于平台私有网络和外部网络（如 IDC、公有云、私有云）之间建立安全可靠的加密网络通信。

支持用户将单机版 VPN 网关升级为主备版，满足 VPN 网关高可用场景。当其中一个网关的虚拟机实例因故障宕机时，备实例会自动转换为主实例持续提供 VPN 网关的通信服务。

注意 VPN 网关实例底层由虚拟机进行构建，虚拟机配置为 2C2G。

(2) 对端网关

运行于外部网络端 IPsecVPN 网关的公网 IP 地址，即与私有云平台 VPN 网关进行隧道连接的网关 IP 地址，支持 NAT 转发的网关地址。

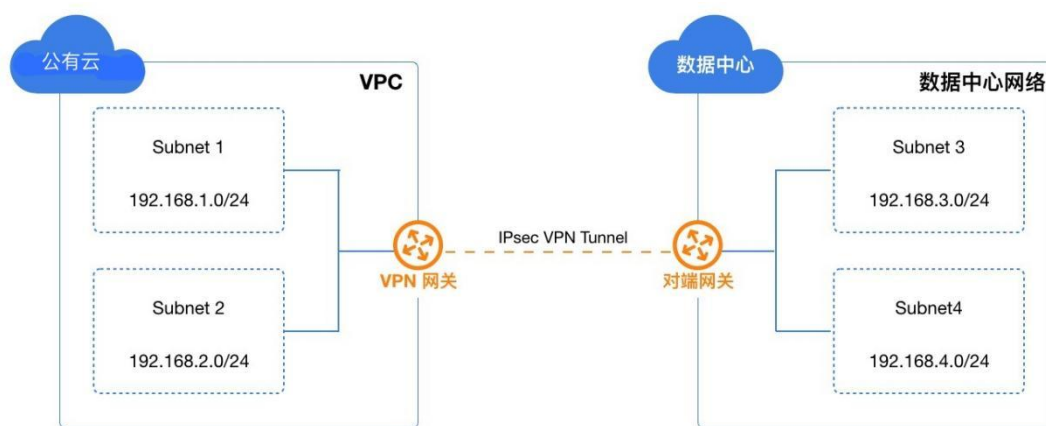
对端网关可以认为是与当前平台建立 VPN 连接的第三方私有云平台、IDC 数据中心及公有云平台的 VPN 网关 IP 地址。若远端网络 VPN 网关使用的是内网地址，需提供内网地址被 SNAT 后的固定公网 IP 地址。

(3) VPN 隧道

连接 VPN 网关和对端网关的加密隧道，结合相应的加密认证算法及策略，为平台 VPC 私有网络和外部私有网络建立加密通信的隧道连接。

一个 VPN 网关有且必须关联 1 个 VPC 网络和 1 个外网 IP 地址，与对端网关相对应，通过 VPN 隧道进行连接。IPsecVPN 支持点到多点的连接特性，使得 VPN 网关与对端网关可以为一对一或一对多的连接关系，即一个 VPN 网关可以同时与多个对端网关建立隧道。

VPN 隧道支持平台多个 VPC 子网与对端网络的多个网段通过隧道进行加密通信，平台 VPC 子网的网段与对端网络的网段不可重叠（本端与对端子网重叠会影响网络的正常通信）。



如上图案例所示，在云平台中的 VPC 网络已拥有 2 个子网，分别为 subnet1 (192.168.1.0/24) 和 subnet2 (192.168.2.0/24)。在远端 IDC 数据中心下有 2 个内网网段，分别为 subnet3 (192.168.3.0/24) 和 subnet4 (192.168.4.0/24)。

- 私有云平台 VPN 网关绑定 VPC 子网，并使用外网 IP 地址作为网络出口及远端数据中心的对端网关。
- 远端数据中心的平台的网关绑定数据中心子网，并使用另一个公网 IP 地址作为网络出口及私有云平台的的对端网关。
- 两端 VPN 网关分别建立 IPsecVPN 隧道，使用相同的预共享密钥及加密认证策略，经过第一阶段的 IKE 认证及第二阶段的 IPsec 认证，建立 VPN 连接通道。

- 两端网络的子网分别通过 VPN 隧道与对端网络的子网进行通信，打通跨数据中心、跨云平台的内网，构建混合云环境。

IPsecVPN 通道在 Internet 网络中构建并运行，公网的带宽、网络阻塞、网络抖动会直接影响 VPN 网络通信的质量。

4.14.4 VPN 隧道建立

在建立 IPsecVPN 安全通道时，需要先在两个网关间建立 SA (Security Association 安全联盟)。SA 是 IPsec 的基础，是通信网关间对连接条件的约定，如网络认证协议 (AH、ESP)、协议封装模式、加密算法 (DES、3DES 和 AES)、认证算法、协商模式 (主模式和野蛮模式)、共享密钥及密钥生存周期等。SA 安全联盟的建立需要在两端网关上均约定并配置相同的条件，以确保 SA 可以对两端网关进行双向数据流通信保护。

标准 IPsecVPN 建立 SA 的方式有手工配置和 IKE 自动协商两种，私有云平台 VPN 网关服务使用 IKE 协议来建立 SA。IKE 协议建立在由 ISAKMP (Internet Security Association and Key Management Protocol, 互联网安全联盟和密钥管理协议) 定义的框架上，具有一套自保护机制，可在不安全的网络上安全地认证身份、交换及密钥分发，为 IPsec 提供自动协商交换密钥并建立 SA 服务。

- **身份认证**：支持预共享密钥 (pre-shared-key) 认证，确认通信两端的身份，并在密钥产生之后对身份数据进行加密传送，实现对身份数据的安全保护。
- **交换及密钥分发**：DH (Diffie-Hellman, 交换及密钥分发) 算法是一种公共密钥算法，通信两端在不传输密钥的情况下通过交换一些数据，计算出共享的密钥。

IKE 通过两个阶段为 IPsec 进行密钥协商并建立 SA:

1. 第一阶段: 通信两端彼此间建立一个已通过身份认证和安全保护的通道，即建立一个 IKE SA，作用是为两端之间彼此验证身份，并协商出 IKE SA，保护第二阶段中 IPsec SA 协商过程。支持 IKE V1 和 V2 版本，其

中 V1 版本支持主模式 (Main Mode) 和野蛮模式 (Aggressive Mode) 两种 IKE 交换方法。

2. 第二阶段: 用第一阶段建立的 IKE SA 为 IPsec 协商安全服务, 即为 IPsec 协商具体的 SA, 建立用于最终的 IP 数据安全传输的 IPsec SA。

IKE 为 IPsec 协商建立 SA, 并将建立的参数及生成的密钥交给 IPsec, IPsec 使用 IKE 协议建立的 SA 对最终 IP 报文加密或认证处理。通过 IKE 协议可为 IPsecVPN 提供端与端之间的动态认证及密钥分发, 通过自动建立 IPsec 参数, 降低手工配置参数的复杂度; 同时由于 IKE 协议中每次 SA 的建立均需运行 DH 交换过程, 可有效保证每个 SA 所使用密钥的互不相关, 增加 VPN 通道的安全性。

VPN 隧道成功建立连接后, 将自动为所属 VPC 关联的本端子网下发到对端子网的路由, 使本端子网访问远端私有网络的需求通过 VPN 网关及隧道进行转发, 完成整个链路的打通。

4.14.5 VPN 隧道参数

IPsecVPN 隧道 SA 协商建立需要配置相应的参数信息, 包括隧道的基本信息、预共享密钥、IKE 策略及 IPsec 策略配置信息。两端的 VPN 在建立的过程中, 需保证预共享密钥、IKE 策略及 IPsec 策略配置一致, IKE 策略指定 IPsec 隧道在协商阶段的加密和认证算法, IPsec 策略指定 IPsec 在数据传输阶段所使用的协议及加密认证算法。具体参数信息如下表所示:

(1) 基本信息

- 名称/备注: VPN 隧道连接的名称和备注。
- VPN 网关: VPN 隧道挂载的 VPN 网关, 即隧道运行在云平台端的所属 VPN 网关。
- 对端网关: VPN 隧道挂载的对端网关, 即对端网关的互联网出口 IP 地址, 如 IDC 数据中心的 VPN 网关。
- 本端网段: VPN 网关所在 VPC 网络内需要和对端网络 (如 IDC 数据中

心) 互通的子网, 如 192.168.1.0/24。本端网段用于第二阶段协商, 不可与对端网段重叠。

- 对端网段: IDC 数据中心或第三方云平台中需要与本端网段 VPN 通信的子网, 如 192.168.2.0/24。对端网段用于第二阶段协商, 不可与本端网段重叠。

(2) 预共享密钥

Pre Shared Key: IPsecVPN 连接的密钥, 用于 VPN 连接的协商, 在 VPN 连接协商过程中, 需保证本端与对端的密钥一致。

(3) IKE 策略

- 版本: IKE 密钥交换协议的版本, 支持 V1 和 V2。V2 版对 SA 的协商过程进行简化且更加适应多网段场景, 推荐选择 V2 版本。
- 认证算法: 为 IKE 协商过程中的报文提供认证, 支持 md5、sha1 和 sha2-256 三种认证算法。
- 加密算法: 为 IKE 协商过程中的报文提供加密保护, 支持 3des、aes128、aes192、aes256 四种加密算法。
- 协商模式: IKE v1 的协商模式, 支持主模式 (main) 和野蛮模式 (aggressive)。
 - 主模式在 IKE 协商时需经过 SA 交换、密钥交换、身份验证三个双向交换阶段 (6 个消息), 而野蛮模式仅需要经过 SA 生成/密钥交换和身份验证两次交换阶段 (3 个消息)。
 - 由于野蛮模式密钥交换与身份认证一起进行无法提供身份保护, 因此主模式的协商过程安全性更高, 协商成功后信息传输安全性一致。
 - 主模式适用于两端设备的公网 IP 固定的场景, 野蛮模式适用于需要 NAT 穿越及 IP 地址不固定的场景。
- DH 组: 指定 IKE 交换密钥时使用的 Diffie-Hellman 算法, 密钥交换的

安全性及交换时间随 DH 组的扩大而增加，支持 1、2、5、14、24。

- 1: 采用 768-bit 模指数 (Modular Exponential, MODP) 算法的 DH 组。
 - 2: 采用 1024-bit MODP 算法的 DH 组。
 - 5: 采用 1536-bit MODP 算法的 DH 组。
 - 14: 采用 2048-bit MODP 算法的 DH 组。
 - 24: 带 256 位的素数阶子群的 2048-bit MODP 算法 DH 组。
- 本端标识: VPN 网关的标识, 用于 IKE 第一阶段协商。支持 IP 地址和 FQDN (全称域名)。
 - 对端标识: 对端网关的标识, 用于 IKE 第一阶段协商。支持 IP 地址和 FQDN (全称域名)
 - 生存周期: 第一阶段 SA 的生存时间, 在超过生存周期后, SA 将被重新协商, 如 86400 秒。

(4) IPSec 策略

- 安全传输协议: IPSec 支持 AH 和 ESP 两种安全协议, AH 只支持数据的认证保护, ESP 支持认证和加密, 推荐使用 ESP 协议。
- IPSec 认证算法: 为第二阶段用户数据提供的认证保护功能, 支持 md5 和 sha1 两种认证算法。
- IPSec 加密算法: 为第二阶段用户数据提供的加密保护功能, 支持 3des、aes128、aes192 和 aes256 四种加密算法, 使用 AH 安全协议时不可用。
- PFS DH 组: PFS (Perfect Forward Secrecy, 完善的前向安全性) 特性是一种安全特性, 指一个密钥被破解, 并不影响其他密钥的安全性。PFS 特性为第二阶段协商的 Diffie-Hellman 密钥交换算法, 支持的 DH 组为支持 1、2、5、14、24 与关闭 (Disable), Disable 适用于不支持

PFS 的客户端。

- 生存周期：第二阶段 SA 的生存时间，在超过生存周期后，SA 将被重新协商，如 86400 秒。

4.14.6 应用场景

VPN 网关 IPsecVPN 服务是基于 Internet 的网络连接服务，通过 IPsec 安全加密通道实现企业数据中心、办公网络与平台 VPC 私有网络的安全可靠连接，同时用户也可使用 VPN 网关在 VPC 之间建立加密内网连接。网关服务为可容灾的高可用架构，同时支持用户选择多种加密及认证算法，并提供 VPN 连接健康检测及连接日志，可满足不同的应用场景。

- VPC 到本地数据中心的连接：通过 IPsecVPN 服务将本地数据中心的内网主机和 VPC 网络的虚拟资源进行连接，构建混合云服务模式。
- VPC 到公有云 VPC 的连接：通过 IPsecVPN 服务将第三方公有云 VPC 私有网络和私有云 VPC 网络的虚拟资源进行连接，构建多云混合服务模式。
- VPC 到第三方私有云内网的连接：通过 IPsecVPN 服务将第三方私有云的 VPC 私有网络和 UCloudStack VPC 网络的虚拟资源进行连接，构建多云混合服务模式。
- VPC 到 VPC 的连接：通过 IPsecVPN 服务将 VPC 与的另一个 VPC 网络进行连接，实现 VPC 打通的场景。

4.14.7 使用流程

使用 VPN 网关 IPsec 服务前，需要明确场景并根据不同场景部署 VPN 及连接：

- 租户根据需要创建本端 VPC 网络及子网，并在子网中部署虚拟机。
- 租户根据需求指定 VPN 网关所在的 VPC 网络外网 IP 地址、安全组等参数创建高可用 VPN 网关。

- 租户根据对端网关的 IP 地址创建对端网关。
- 租户根据需求指定 VPN 隧道基本参数、预共享密钥、IKE 策略及 IPsec 策略部署 IPsec VPN 隧道。
- 用户使用一致的 VPN 隧道参数对远端网关设备的 VPN 进行配置。（远端网关设备指 IDC 数据中心的 VPN 路由设备、不同于本端 VPC 的 VPN 网关或第三方云平台的 VPN 网关等）
- 根据需求配置 VPC 私有网络中需要通信主机的路由，若可以自动下发路由，则无需配置路由。
- 测试网络连通性，如本端 VPC 子网中虚拟机 ping 远端私有网络中的 IP 地址，验证通信是否正常。

通常情况下，IKE 协议采用 UDP 的 500 和 4500 端口进行通信，IPsec 的 AH 和 ESP 协议分别使用 51 或 50 号协议来工作，因此为保障 IKE 和 IPsec 的正常运行，需要确保应用 IKE 和 IPsec 配置的网关设备或防火墙已开放以上端口和协议的流量。

注意 IPsecVPN 服务是基于互联网的加密通信服务，在使用 IPsecVPN 前需确认两端网关均有固定或 NAT 后的互联网 IP 地址。

4.15 裸金属

4.15.1 概述

裸金属为用户提供了统一纳管存量裸金属的能力，用户在控制台即可对已纳管的裸金属进行电源管理、访问远程控制台和查看硬件监控，装机等基础运维操作。

平台管理员可对裸金属进行全生命周期管理以及分配已经添加的裸金属给平台租户使用。

支持适配不同的机型，根据用户的个性化需求进行定制化开发。已适配支持的机型包括浪潮 SA5212M4、联想 ThinkSystem SR650、联想 ThinkSystem

SR658、新华三 R4900 G2、新华三 UniServer R4900 G3

支持裸金属服务器的 IPMI 网络与平台网络打通，将裸金属添加到云平台进行统一管理，并支持批量导入裸金属设备。

支持将添加至云平台的裸金属设备分配给租户，使租户及子账号可在控制台直接申请裸金属服务器进行使用。

支持租户对已申请的裸金属进行电源管理，访问远程控制台、查看硬件监控、重装系统等常用运维操作。如开机、关机、重启、准备控制台、控制台登录、释放控制台、装机、重装、强制关机及关机并重启等。

4.15.2 使用流程

在使用裸金属服务前，必须提前准备好裸金属设备，并根据需求将裸金属服务器的 IPMI 网络及业务网络与平台网络进行打通，在通过平台录入设备信息，将设备添加给租户管理。裸金属服务的使用流程分为【平台管理员流程】和【租户流程】两大部分，具体如下：

1. 硬件环境装备

准备好硬件环境，配置物理网络交换机及服务器 IPMI 网络，使平台物理网络与 IPMI 网络可互相通信。

2. 为租户添加裸金属

由【平台管理员】为租户添加裸金属信息，包括服务器的名称、IPMI IP、IPMI User、IPMI PassWord、机架位置、标签及租户邮箱，支持批量导入裸金属信息。

3. 裸金属管理

由【平台租户】对已申请的裸金属进行生命周期管理，支持开启、关机、重启、准备控制台、控制台登录、释放控制台、装机、重装、强制关机及关机并重新开机。

平台租户在获得平台管理员所授权使用的裸金属时，租户在控制台即可对已纳管的裸金属进行电源管理、访问远程控制台、查看硬件监控、重装系统等常用

的运维操作。

平台管理员可对裸金属进行全生命周期管理以及分配已经添加的裸金属给平台租户使用。生命周期的管理内容包含：添加、查看、删除等，通过一系列配置后租户通过控制台即可对裸金属进行电源管理、访问远程控制台和查看硬件监控等基础运维操作。

4.16 弹性伸缩

4.16.1 概述

弹性伸缩 (Auto Scaling) 是指在业务需求增长时自动增加计算资源 (虚拟机) 以保证计算能力, 在业务需求下降时自动减少计算资源以节省成本; 同时可结合负载均衡及健康检查机制, 满足请求量波动和业务量稳定的场景。

用户可通过弹性伸缩服务, 定制弹性伸缩组及伸缩策略, 在伸缩组内资源量达到策略定义的阈值后, 根据定制的虚拟机模板自动增减虚拟机数量或配置, 提升业务部署及运维的效率。

弹性伸缩服务支持水平伸缩和垂直伸缩两种类型, 帮助用户保持应用系统的可用性, 并允许用户定义扩缩容策略, 自动添加、删除和云主机实例 (水平伸缩), 或者为云主机增加 CPU 和内存配置 (垂直伸缩), 为既定或实时的需求, 提供有保障的云服务模式。

4.16.2 水平伸缩

水平伸缩负责将组内的实例数量维持在“期望”的水位, 添加/缩减虚拟机的动作均由伸缩组进行操作, 支持“自动伸缩”和“固定数量”两种模式维护伸缩组内的实例数量, 适应多种自动伸缩场景。伸缩组是通过伸缩策略中对于伸缩规则及伸缩实例数来维护组内实例的期望水平, 并通过虚拟机模板创建新的实例。

- 自动伸缩模式依据伸缩器的伸缩策略维护伸缩组中的实例数量;
- 固定数量模式依据用户指定的实例数量维护伸缩组中的实例, 即固定数量

模式无需指定伸缩策略。

支持伸缩组预热时间，使虚拟机创建成功后，在预热时间内拉起应用程序以承接业务流量；同时支持虚拟机类型伸缩组关联负载均衡，为伸缩组中的虚拟机业务提供负载均衡服务，同时通过监听器的健康检查机制，判断伸缩组中所有实例的业务健康状况，自动剔除业务不健康的实例并新增健康实例到业务集群。

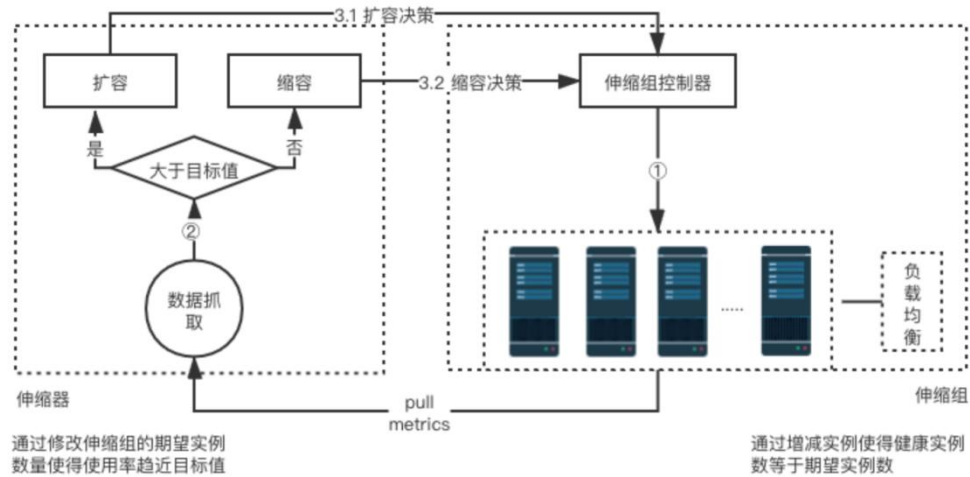
作为自动伸缩服务的最核心模块，支持水平伸缩的全生命周期管理，包括创建伸缩组、查看伸缩组、修改伸缩组、启动/禁用伸缩组、绑定/解绑负载均衡、查看伸缩日志及删除伸缩组等。

水平弹性伸缩的策略管理功能有助于维护伸缩组的运行状况和可用性。考虑到用户业务的多样性，伸缩服务提供多维度的指标，提升应对复杂业务场景的云资源弹性需求。

- 基于 CPU 阈值作为伸缩条件。
- 基于内存阈值作为伸缩条件。
- 基于 TCP 连接数作为伸缩条件。
- 基于整体并发数作为伸缩条件。
- 基于后端的延迟响应作为伸缩条件。

4.16.2.1 逻辑架构

水平伸缩从逻辑架构上可分为三部分，分别为伸缩组、伸缩器及虚拟机模板。



- 伸缩组：负责将组内的实例数量维持在“期望”的水位，添加/缩减虚拟机的动作均由伸缩组进行操作，支持“自动伸缩”和“固定数量”两种模式维护伸缩组内的实例数量。
- 伸缩器：即伸缩策略，用于定义伸缩组内虚拟机伸缩的规则，支持定义伸缩组最小及最大实例数量，并可配置是否允许缩容。
 - 虚拟机类型根据 CPU、内存使用率的阈值触发伸缩动作，
 - 监听器类型根据负载均衡中 Vserver 的七层连接数、七层 QPS、七层响应时间、四层连接数、四层 CPS 的阈值触发伸缩动作
- 虚拟机模板：用户根据需求自定义虚拟机模板，用于弹性伸缩时自动创建虚拟机的模板，同时支持通过虚拟机模板手动创建虚拟机。

伸缩组定义好伸缩模式后，伸缩组的实例“期望”值由伸缩策略接管并动态修改，最终由伸缩组负责虚拟机的动态扩容和缩容，新增虚拟机实例时会根据虚拟机模板创建新的虚拟机实例。

4.16.2.2 伸缩组工作流程

伸缩组内的虚拟机实例可定义预热时间，指为虚拟机创建成功后需要一定的时间拉起应用程序以承接业务流量。因此在伸缩组发起创建虚拟机的请求后，在虚拟机创建成功并处于运行中状态时，伸缩组中虚拟机的状态为“启动中”，代

表虚拟机在预热中，待超过预热时间后，会自动转换为“运行”，代表虚拟机为健康状态。

伸缩组每 15 秒获取一次被其控制的所有虚拟机状态，判断是否需要添加或删除实例。若伸缩组关了负载均衡，则由负载均衡判断伸缩组内的实例是否健康，若不健康具体流程如下：

- **健康实例等于期望值**

伸缩组会自动将不健康 (基于三个周期健康检测的判断) 的实例移出伸缩组，并执行删除虚拟机操作。

- **健康实例大于期望值**

选择将最晚创建的健康虚拟机实例移出伸缩组，并执行删除虚拟机操作，同时将不健康的实例移出伸缩组并执行删除操作。

- **健康实例小于期望值**

伸缩组会自动以虚拟机模板发起创建实例操作，并将实例数量维持在期望值，同时会将不健康的实例移出伸缩组并执行删除操作。

4.16.2.3 伸缩器工作流程

伸缩器会根据伸缩策略中设置的最小和最大实例值，每 15 秒采集一次伸缩组中健康实例的伸缩阈值数据，用于判断是否需要扩容或缩容伸缩组中的实例。

- **扩容**：若伸缩组中健康实例的伸缩阈值大于伸缩策略定义的阈值，则会触发伸缩组进行扩容实例操作。
- **缩容**：通常伸缩组中健康实例的伸缩阈值小于伸缩策略定义的阈值，则会触发伸缩组进行缩容实例操作。

为避免频繁的缩容导致伸缩组内集群服务震荡，缩容时会获取伸缩组过去 10 分钟内所有健康实例的伸缩指标监控数据平均值，用于判断是否需要缩容伸缩组中的实例。

4.16.2.4 功能特性

水平伸缩通过伸缩组、伸缩策略及虚拟机模板共同维护集群内虚拟机的实例数量，虚拟机类型可结合负载均衡对伸缩组内虚拟机实例的业务健康进行检测并及时剔除处于不健康状态的虚拟机实例，保证整体集群业务的可用性和可靠性。监听器类型的伸缩组创建时需要绑定负载均衡。

虚拟机伸缩类型：支持设置 CPU、内存使用率作为伸缩指标，可后续绑定负载均衡为伸缩组中资源提供负载均衡服务。

监控器伸缩类型：支持 TCP、HTTP、HTTPS 三种协议类型，选择 TCP 协议时伸缩指标可选四层连接数、四层 CPS 作为伸缩指标，选择 HTTP 以及 HTTPS 协议时可选择七层连接数、七层 QPS、七层响应时间作为伸缩指标。

- **虚拟机模板：**伸缩组为组内新增虚拟机实例时所使用的虚拟机模板，即会根据所选虚拟机模板为组内新增虚拟机实例。
- **预热时间：**伸缩组内虚拟机实例创建成功后的预热时间，在预热时间内虚拟机可拉起应用程序以承接业务流量，预热中的虚拟机实例在伸缩组中处于【启动中】状态。创建时必须指定，默认为 300s。
- **成员数量设置：**可设置最小成员数量和最大成员数量，调度后的资源数量不会超过设置的成员数量上下限。如果想要达到固定数量的效果，可以将最大成员数量、最小成员数量设置成一致。
- **是否允许缩容：**允许缩容时伸缩组会在未满足指标时减少资源数量，缩容时会给虚拟机内的业务进程发送 SIGTERM 信号，最多等待 90 秒后关闭并删除虚拟机，业务进程可利用该机制优雅关闭。
- **伸缩指标：**
 - 虚拟机指标 CPU、内存（单位%）
 - 监听器指标七层连接数（个）、七层 QPS（个/s）、七层响应时间（ms）、四层连接数（个）、四层 CPS（个/s）

- 指标阈值：统计标准 10min 内所有虚拟机监控数据的平均值

支持用户查看伸缩组的伸缩事件和已添加至伸缩组的实例信息，用于查看自动伸缩组所有执行动作及原因，方便用户对伸缩组集群业务进行维护。

支持用户启用或禁用一个伸缩组，伸缩组禁用后即为不可用状态，将不会在触发伸缩策略执行实例伸缩和健康检查，禁用伸缩组不影响伸缩组中已存在实例的正常运行。

水平伸缩服务提供伸缩组中所有实例的平均监控指标数据，并可通过告警模板对监控数据进行告警配置，在使伸缩指标触发扩缩容时，为用户发送告警邮件。

4.16.3 垂直伸缩

垂直伸缩是指根据伸缩指标对虚拟机的 CPU、内存配置及 EIP 带宽维持在“期望”水位，添加/缩减的动作均由伸缩组进行操作。

垂直伸缩使用户可通过虚拟机的 CPU/内存利用率为伸缩指标对虚拟机进行规格扩容，需要虚拟机支持热升级特性；同时支持对 EIP 带宽的伸缩，通过带宽使用率按需伸缩。

垂直伸缩支持虚拟机和 EIP 两种资源类型的伸缩组设置，其中虚拟机不支持缩容，需要支持热升级。

- 虚拟机伸缩指标包括 CPU、内存。
- EIP 伸缩指标为带宽使用率，并支持缩容操作。
- 扩容不能超过最大规格限制，缩容不能低于初始规格。
- 每次扩容后会有两分钟的冷却时间，刚扩容结束监控不一定准确，冷却时间可以保证下次获取监控准确。

4.17 备份服务

4.17.1 概述

备份服务是一种提供数据备份和恢复功能的服务。它允许用户将关键数据和文件复制到另一个存储介质，以便在数据丢失、损坏或灾难恢复时进行恢复。备份服务针对不同的云服务支持不同的备份类型：

支持对 MySQL 服务、Redis 服务、对象存储及文件存储等实例和数据进行定时自动备份和手动备份。

备份服务通过存储池和备份任务计划进行相关服务的备份操作，存储池为备份数据提供数据存储能力；备份任务计划提供备份任务的具体策略，并根据策略进行手动或自动的备份操作。

备份任务策略支持根据备份源类型、备份资源、备份类型、存储池及保留时间进行多维度定义，满足 PaaS 产品的备份场景需求。

平台会保存定时器执行的任务列表及执行结果记录，支持用户在定时器中查看每个任务的执行记录。

4.17.2 存储池

备份服务为用户提供备份数据的存储池管理，当前备份服务的存储池类型支持对象存储。

支持用户获取存储池上已关联的备份任务，并支持用户在平台上绑定存储池、更新存储池及解绑存储池。

4.17.3 备份任务

平台支持各服务的备份任务管理，用于云服务数据和实例的备份场景。支持用户查看备份计划的信息，如备份计划的状态、存储池、源数据类型、源数据地域、源数据、备份策略、定时器、备份类型、备份保留时间(天)、创建时间及更新时间等；同时支持用户从备份计划的可用备份数据创建实例。

备份任务计划支持指定备份策略、备份源类型、备份资源、存储池、保留时间等。

(1) 备份策略

备份策略支持选择定时器和手动备份两种类型, 可根据场景选择适合的备份策略。

- 定时器策略: 提供自动化任务功能, 可用于定期执行一系列任务, 可在指定的周期重复执行, 且每个任务支持多个资源批量操作。
 - 重复执行支持每天、每周、每月的指定时间执行任务。
 - 每天支持单小时或每个小时进行定时任务的执行操作。
 - 每周支持星期一至星期日单小时或每个小时执行操作。
 - 每月支持每一天单小时或每个小时进行定时任务的执行操作。
- 手动备份: 提供手动执行任务的能力, 制定好的备份计划, 可进行手动操作, 每操作一次则执行一次。

(2) 备份源类型

备份源类型支持对 MySQL 服务、Redis 服务、对象存储实例及文件存储实例进行自动备份和手动备份。各服务的备份机制如下:

- MySQL 服务: 支持逻辑备份、物理备份及快照备份。
 - 逻辑备份支持选择数据库的全部表或部分表进行备份操作。
 - 物理备份是指备份数据库的数据文件。
 - 快照备份是指将 MySQL 服务的实例进行快照。
- Redis 服务: 支持逻辑备份。
- 对象存储: 支持快照备份, 对对象存储实例进行快照操作。
- 文件存储: 支持快照备份, 对文件存储实例进行快照操作。

(3) 备份资源

备份资源是指在当前备份源类型下，用户所拥有的具体资源，支持对多个资源进行批量备份操作。

(4) 存储池

存储池是备份计划中备份资源的备份数据存储资源，当前存储池资源仅支持对象存储，在创建备份前必须创建一个对象存储实例作为备份数据的存储池。

(5) 保留时间

支持设置备份的保留时间，超期后会自动删除。保留时间默认为 30 天，可在制订备份任务策略时进行自定义，时间范围可配置 1~90 天。

5 PaaS 产品服务

5.1 MySQL 服务

5.1.1 产品概述

MySQL 服务是平台基于关系型数据库 MySQL 提供的数据库 PaaS 服务，支持单机版和主备版两种机型，并提供数据库实例、从库管理、升级机型、备份、监控、事件日志及参数配置等特性，满足应用服务数据库构建和运维场景需求。

MySQL 服务支持单机版和主备份两种类型，分别满足单机场景和高可用场景的数据库服务。通过平台 MySQL 服务，用户可一键部署 MySQL 数据库实例，并提供主备高可用实例机制。

支持用户在部署时选择 MySQL 实例所在的集群，以适配不同的硬件环境；同时支持用户对 MySQL 实例的内存进行配置，如 2GB、32GB 等，支持自定义 MySQL 实例的内存规格。

MySQL 数据库实例的存储默认提供 40GB 系统盘，支持用户自定义挂载数据盘，进行 MySQL 数据的存储。用户在部署时可选择 MySQL 数据盘所在的存储集群（如 HDD 三副本存储集群、SSD 多副本存储集群等），支持用户自定义数据存储的容量。

MySQL 服务可与虚拟机部署至相同的 VPC 网络，使虚拟机可直接通过内网对 MySQL 数据库进行读写操作。用户也可对 MySQL 服务绑定外网 IP 地址，对云平台外的数据中心网络提供 MySQL 数据库服务，并通过外网安全组的安全组策略，保证网络访问的安全性。

平台支持 MySQL 实例及数据库各种监控数据，并可根据监控模板进行监控报警及通知，保证数据库的正常运行。同时支持获取并查看 MySQL 数据库的操作日志及事件日志。

当前平台仅支持 MySQL 5.7 版本，支持为 MySQL 实例创建从库，并对从库进行全生命周期管理，同时支持 MySQL 的重置密码、配置升级、升级为主备

版，并支持 MySQL 服务的参数配置，通过参数模板快速应用并配置实例。

5.1.2 实例管理

MySQL 服务支持对 MySQL 实例进行全生命周期管理，通过平台一键部署单机版或主备高可用版的 MySQL 服务，用户可通过指定集群、机型、内存容量、数据盘类型及容量、版本、参数模板、VPC、子网、外网 IP、外网安全组、名称及密码等相关基础信息自动创建一个 MySQL 实例。

- 集群：MySQL 实例所在的计算集群，即 MySQL 服务的虚拟机实例所运行的物理位置。
- 机型：支持单机版和高可用版，单机版可在实例创建后升级为主备版。
- 内存容量：MySQL 实例的可使用的内存容量，默认支持 1G、2G、4G、8G、16G 及 32G，支持平台管理自定义内存规格。
- 数据盘类型及容量：支持为 MySQL 实例挂载不同存储集群的云硬盘，并支持自定义配置数据盘容量，容量范围为 10-32000GB，调整步长以 1GB 为单位。
- 版本：当前平台仅支持 MySQL 5.7 版本。
- 参数模板：MySQL 实例启动时所使用的配置参数模板，支持用户自定义参数模板，并将模板应用到实例。
- VPC/子网：MySQL 实例运行时所属的 VPC 网络和子网网段，同时为 MySQL 实例提供的 IP 地址。
- 外网 IP：支持为 MySQL 实例挂载外网 IP 地址，为平台外网提供 MySQL 数据库服务。
- 外网安全组：支持为绑定外网 IP 的 MySQL 实例绑定安全组，通过安全组规则的出入规则配置，保证 MySQL 对外提供服务的网络安全性。
- 通用信息配置：支持为 MySQL 配置名称、备注、登录密码及标签等设置，通过登录密码，用户可使用客户端在网络可通的情况下，管理

MySQL 数据库或数据表。

MySQL 实例 CPU 规格根据选择的内存变化，4G 内存以下为 2C；若内存大于等于 4GB 小于 32GB，则 CPU 为 4C；若内存大于等于 32GB，则 CPU 为 8C。

注意 以上所描述的规格及资源均为单机版容量，或为主备高可用版本，则 CPU、内存、存储等容量规格为 2 倍。

支持用户查看 MySQL 实例信息，如包括状态、机型、集群、存储类型、版本、IP、内存容量、数据盘容量、VPC、子网、安全组、计费方式、项目组、创建时间及过期时间等。

同时支持对 MySQL 实例进行从库管理、重置密码、应用参数模板、配置升级、升级主备版、绑定外网 IP、解绑外网 IP、修改安全组、续费、修改告警模板及修改 IP 等管理。

- 从库管理：平台支持用户对 MySQL 主库进行从库创建及管理。
- 重置密码：平台支持用户对 MySQL 登录密码进行重置。
- 应用参数模板：将一个参数模板应用于当前实例。
- 配置升级：支持用户对 MySQL 进行配置升级操作，包括内存容量和数据盘容量更改。
- 升级主备版：平台支持用户对单机版 MySQL 进行升级至主备版操作，计算集群、内存容量、存储集群、数据盘容量不可修改。
- 绑定外网 IP：支持用户对 MySQL 实例绑定外网 IP，对外网提供 MySQL 服务，每个 MySQL 支持最多绑定一个外网 IP。
- 解绑外网 IP：支持用户对已绑定外网 IP 的 MySQL 进行解绑。
- 修改安全组：支持用户对 MySQL 实例的外网安全组进行修改。
- 续费：支持用户对 MySQL 进行手动续费操作。
- 修改告警模板：支持用户修改 MySQL 实例的告警模板。

- 修改 IP 地址：支持用户修改 MySQL 的内网 IP(VIP)地址。

5.1.3 从库管理

平台支持用户对 MySQL 主库进行从库创建及管理，每个 MySQL 主库支持最多创建 5 个从库。从库的管理与主库的生命周期基本一致。

从库所属地域必须与主库地域一致，用户进行从库创建时支持选择从库实例的计算集群、内存容量、数据盘集群、数据盘容量、外网 IP、安全组及机型等。

5.1.4 参数配置

平台支持用户对 MySQL 进行参数配置相关操作，包括修改实例参数、应用参数模板、导入参数文件、导出为模板、导出参数文件等。

- 修改实例参数：平台支持用户对 MySQL 实例的通用参数进行修改，修改参数即时生效。
- 应用参数模板：支持将参数模板应用到 MySQL 实例。
- 导入参数文件：平台支持用户对 MySQL 进行导入参数文件。
- 导出为模板：将当前 MySQL 实例的参数配置导出为一个参数模板，可应用于其它 MySQL 实例。
- 导出参数文件：将当前 MySQL 实例的参数配置导出为一个文件。

5.1.5 备份管理

平台 MySQL 支持逻辑备份、物理备份及快照备份三种方式对 MySQL 数据库进行备份操作，并可结合备份服务进行自动备份任务的执行，全面保证 MySQL 数据库数据的安全性。

- 逻辑备份支持选择数据库的全部表或部分表进行备份操作。
- 物理备份是指备份数据库的数据文件。
- 快照备份是指将 MySQL 服务的实例进行快照。

平台支持对三种 MySQL 备份进行全面管理，如查看备份信息、删除备份及从备份创建。

- 查看备份信息：平台支持用户查看备份管理信息，如资源 ID、状态、所属存储池、来源备份计划、保留时间(天)、保存路径、创建时间、更新时间及到期时间等。
- 删除备份：支持用户删除备份数据，并支持批量删除。
- 从备份创建：支持用户从备份创建 MySQL 实例。

5.1.6 监控告警

平台支持 MySQL 服务的监控告警服务，包括监控数据和监控告警。

支持监控 MySQL 实例和从库实例的 CPU 使用率、内存使用率、磁盘使用率、QPS、Buffer Pool 容量、行锁平均锁定时间、活跃连接数、每秒连接数、表锁数、每秒网卡入流量及每秒网卡出流量等监控数据。

同时支持用户自定义 MySQL 告警模板，统一进行告警规则及告警通知的配置和管理，提高运维和监控效率。

5.1.7 日志事件

MySQL 服务支持用户获取并查看实例的操作日志及事件日志。

操作日志包括操作 (API) 名称、所属模块、地域、关联资源 ID、操作者、操作结果、操作时间，并可通过操作结果及操作时间进行筛选。

事件日志包括事件类型、事件等级、事件内容、事件发生次数、开始时间、更新时间，可通过事件周期进行筛选。

5.1.8 参数模板

平台为 MySQL 服务提供参数模板，用户可将 MySQL 的参数配置预定义至参数模板中，便于后续快速构建 MySQL 实例，并可应用至已有的 MySQL 实例。

支持参数模板的创建、应用到实例、下载参数模板、删除参数模板等管理。

- 创建参数模版：支持用户创建 MySQL 参数模板，创建方式可通过复制现有模板或导入模板文件。
- 应用到实例：支持将参数模板应用到 MySQL 实例，使用参数模板中的参数配置 MySQL 实例。
- 下载参数模板：支持用户下载 MySQL 参数模板为一个文件。
- 删除参数模版：支持用户删除 MySQL 自定义的参数模板，删除后已应用该模板的 MySQL 不受影响。

5.2 Redis 服务

5.2.1 产品概述

Redis 服务是平台基于 Redis 提供的缓存 PaaS 服务，支持单机版和主备版两种机型，并提供缓存实例、从库管理、升级内存、升级主备版、清理数据、备份、监控、事件日志及参数配置等特性，满足应用服务缓存构建和运维场景需求。

Redis 服务支持单机版和主备份两种类型，分别满足单机场景和高可用场景的缓存服务。通过平台 Redis 服务，用户可一键部署 Redis 缓存实例，并提供主备高可用实例机制。

支持用户在部署时选择 Redis 实例所在的集群，以适配不同的硬件环境；同时支持用户对 Redis 实例的内存进行配置，如 2GB、32GB 等，支持自定义 Redis 实例的内存规格。

Redis 缓存实例的 CPU 规格默认为 2C，不可进行变更；存储默认提供 40GB 系统盘，ARM 版 Redis 实例的系统盘为 100GB。

Redis 服务可与虚拟机部署至相同的 VPC 网络，使虚拟机可直接通过内网对 Redis 进行读写操作。用户也可对 Redis 服务绑定外网 IP 地址，对云平台外的数据中心网络提供 Redis 数据服务，并通过外网安全组的安全组策略，保证网络访问的安全性。

平台支持 Redis 实例的各种监控数据，并可根据监控模板进行监控报警及通

知，保证缓存服务的正常运行。同时支持获取并查看 Redis 缓存服务的操作日志及事件日志。

当前平台仅支持 Redis 4.0 版本，支持为 Redis 实例创建从库，并对从库进行全生命周期管理。

5.2.2 实例管理

Redis 服务支持对 Redis 实例进行全生命周期管理，通过平台一键部署单机版或主备高可用版的 Redis 服务，用户可通过指定集群、机型、内存容量、参数模板、VPC、子网、外网 IP、外网安全组、名称及密码等相关基础信息自动创建一个 Redis 实例。

- 集群：Redis 实例所在的计算集群，即 Redis 服务的虚拟机实例所运行的物理位置。
- 机型：支持单机版和高可用版，单机版可在实例创建后升级为主备版。
- 内存容量：Redis 实例的可使用的内存容量，默认支持 1G、2G、4G、8G、16G 及 32G，支持平台管理自定义内存规格。
- 参数模板：Redis 实例启动时所使用的配置参数模板，支持用户自定义参数模板，并将模板应用到实例。
- VPC/子网：Redis 实例运行时所属的 VPC 网络和子网网段，同时为 Redis 实例提供的 IP 地址。
- 外网 IP：支持为 Redis 实例挂载外网 IP 地址，为平台外网提供 Redis 缓存服务。
- 通用信息配置：支持为 Redis 配置名称、备注、登录密码及标签等设置，通过登录密码，用户可使用客户端在网络可通的情况下，管理 Redis 缓存服务。

注意 以上所描述的规格及资源均为单机版容量，或为主备高可用版本，则 CPU、内存、存储等容量规格为 2 倍。

支持用户查看 Redis 实例信息，如包括状态、机型、IP 和端口、实例容量、VPC、子网、安全组、计费方式、项目组、创建时间、过期时间等。

同时支持对 Redis 缓存实例进行从库管理、重置密码、应用参数模板、升级内存、升级主备版、清理数据、绑定外网 IP、解绑外网 IP、修改安全组、续费、修改告警模板及修改 IP 等管理。

- 从库管理：平台支持用户对 Redis 主库进行从库创建及管理。
- 重置密码：平台支持用户对 Redis 登录密码进行重置。
- 应用参数模板：将一个参数模板应用于当前实例。
- 升级内存：支持用户对 Redis 进行内存配置升级。
- 升级主备版：平台支持用户对单机版 Redis 进行升级至主备版操作，计算集群、内存容量不可修改。
- 清理数据：支持用户对 Redis 实例进行数据清理，清除数据即时生效且影响线处业务，建议在业务低峰期操作。
- 绑定外网 IP：支持用户对 Redis 实例绑定外网 IP，对外网提供 Redis 服务，每个 Redis 支持最多绑定一个外网 IP。
- 解绑外网 IP：支持用户对已绑定外网 IP 的 Redis 进行解绑。
- 修改安全组：支持用户对 Redis 实例的外网安全组进行修改。
- 续费：支持用户对 Redis 进行手动续费操作。
- 修改告警模板：支持用户修改 Redis 实例的告警模板。
- 修改 IP 地址：支持用户修改 Redis 的内网 IP(VIP)地址。

5.2.3 从库管理

平台支持用户对 Redis 主库进行从库创建及管理，每个 Redis 主库支持最多创建 5 个从库。从库的管理与主库的生命周期基本一致。

从库所属地域必须与主库地域一致，用户进行从库创建时支持选择从库实例

的计算集群、内存容量、外网 IP、安全组及机型等，其中内存容量不可低于主库的内存容量。

5.2.4 参数配置

平台支持用户对 Redis 进行参数配置相关操作，包括修改实例参数、应用参数模板、导入参数文件、导出为模板、导出参数文件等。

- 修改实例参数：平台支持用户对 Redis 实例的通用参数进行修改，修改参数即时生效。
- 应用参数模板：支持将参数模板应用到 Redis 实例。
- 导入参数文件：平台支持用户对 Redis 进行导入参数文件。
- 导出为模板：将当前 Redis 实例的参数配置导出为一个参数模板，可应用于其它 Redis 实例。
- 导出参数文件：将当前 Redis 实例的参数配置导出为一个文件。

5.2.5 备份管理

平台 Redis 支持逻辑备份的方式对 Redis 进行备份操作，并可结合备份服务进行自动备份任务的执行，全面保证 Redis 数据的安全性。

平台支持对 Redis 备份进行全面管理，如查看备份信息、删除备份及从备份创建。

- 查看备份信息：平台支持用户查看备份管理信息，如资源 ID、状态、所属存储池、来源备份计划、保留时间(天)、保存路径、创建时间、更新时间及到期时间等。
- 删除备份：支持用户删除备份数据，并支持批量删除。
- 从备份创建：支持用户从备份创建 Redis 实例。

5.2.6 监控告警

平台支持 Redis 服务的监控告警服务，包括监控数据和监控告警。

支持监控 Redis 实例和从库实例的 CPU 使用率、内存使用率、内存使用量、磁盘使用率、连接数量、QPS、Key 总个数、Key 过期数、Key 过期数、未命中次数及命中率等监控数据。

同时支持用户自定义 Redis 告警模板，统一进行告警规则及告警通知的配置和管理，提高运维和监控效率。

5.2.7 日志事件

Redis 服务支持用户获取并查看实例的操作日志及事件日志。

操作日志包括操作（API）名称、所属模块、地域、关联资源 ID、操作者、操作结果、操作时间，并可通过操作结果及操作时间进行筛选。

事件日志包括事件类型、事件等级、事件内容、事件发生次数、开始时间、更新时间，可通过事件周期进行筛选。

5.2.8 参数模板

平台为 Redis 服务提供参数模板，用户可将 Redis 的参数配置预定义至参数模板中，便于后续快速构建 Redis 实例，并可应用至已有的 Redis 实例。

支持参数模板的创建、应用到实例、下载参数模板、删除参数模板等管理。

- 创建参数模版：支持用户创建 Redis 参数模板，创建方式可通过复制现有模板或导入模板文件。
- 应用到实例：支持将参数模板应用到 Redis 实例，使用参数模板中的参数配置 Redis 实例。
- 下载参数模板：支持用户下载 Redis 参数模板为一个文件。
- 删除参数模版：支持用户删除 Redis 自定义的参数模板，删除后已应用该模板的 Redis 不受影响。

5.3 文件存储

5.3.1 概述

文件存储是云平台提供的 NFS 文件服务器，可以与虚拟机实例或本地服务器搭配使用。文件存储提供标准 NFS 文件访问协议，协议版本为 NFSv4，支持 POSIX 文件接口。

文件存储服务提供实例管理、文件管理、QOS 配置、备份管理、日志事件、监报告警等特性，满足文件共享存储系统的构建和运维场景需求。

通过文件存储服务，用户可一键构建文件存储实例，用户只需在虚拟机实例中安装文件存储客户端，使用标准挂载命令挂载创建的文件系统，就可轻松地在多个实例间共享文件。

支持用户在部署时选择文件存储实例所在的计算集群和存储集群，以适配不同的硬件环境。

- 文件存储计算实例默认配置为 2C2G，不可进行变更。
- 文件存储实例的存储默认 x86 系统盘为 40GB，ARM 系统盘为 100GB。
- 支持用户自定义文件存储实例的存储容量，用户在部署时可选择实例数据存放的存储集群(如 HDD 三副本存储集群、SSD 多副本存储集群等)，支持用户自定义数据存储的容量。

文件存储服务可与虚拟机部署至相同的 VPC 网络，使虚拟机可直接通过内网挂载或访问文件存储 NFS 协议。用户也可对文件存储服务绑定外网 IP 地址，对云平台外的数据中心网络提供 NFS 存储服务，并通过外网安全组的安全组策略，保证网络访问的安全性。

平台支持文件存储实例的各种监控数据，并可根据监控模板进行监控报警及通知，保证存储服务的正常运行。同时支持获取并查看文件存储服务的操作日志及事件日志。

同时平台支持在线扩容文件存储容量，并支持配置文件存储实例的硬盘

QoS, 用于控制文件存储服务的 IOPS 和读写带宽。

5.3.2 实例管理

文件存储服务支持对文件存储服务实例进行全生命周期管理, 通过平台一键部署文件存储服务, 用户可通过指定计算集群、存储集群、容量、VPC、子网、外网 IP、外网安全组、项目组、文件存储名称等相关基础信息创建文件存储等基础信创自动创建一个文件存储实例。

- 计算集群: 文件存储实例所在的计算集群, 即文件存储服务的虚拟机实例所运行的物理位置。
- 存储集群: 文件存储实例数据存放的存储集群, 如 SSD 存储集群或 HDD 存储集群。
- 容量: 文件存储服务的存储容量, 支持的容量范围为 100 ~ 32000G, 并支持管理员自定义文件存储容量规格。
- VPC/子网: 文件存储实例运行时所属的 VPC 网络和子网网段, 同时为文件存储实例提供服务 IP 地址。
- 外网 IP: 支持为文件存储实例挂载外网 IP 地址, 为平台外网提供文件存储访问服务。
- 外网安全组: 支持为绑定外网 IP 的文件存储实例绑定安全组, 通过安全组规则的出入规则配置, 保证文件存储对外提供服务的网络安全性。
- 通用信息配置: 支持为文件存储配置名称、备注及标签等设置。

支持用户查看文件存储实例信息, 如包括状态、存储容量、挂载地址、VPC、子网、计费方式、项目组、创建时间及过期时间等。

同时支持对文件存储实例进行扩容、QoS 配置、文件管理、绑定外网 IP、解绑外网 IP、修改安全组、修改告警模板及修改 IP 等管理。

- 实例扩容: 平台支持用户扩容文件存储的容量, 适应于业务发生变化需扩容文件存储容量的场景。仅支持扩容文件存储容量, 不支持文件存储

容量的扩容。

- **QoS 配置：**支持配置文件存储实例的硬盘 QoS，用于控制文件存储服务的 IOPS 和读写带宽。硬件介质和容量会影响硬盘的读写 IOPS 和带宽速率，若配置的 QOS 超过硬件本身性能，以硬件性能为准。系统会默认分配 QOS 值，如需取消一块硬盘的限速功能，可将 IOPS 和带宽均配置为 0。
- **绑定外网 IP：**支持用户对文件存储实例绑定外网 IP，对外网提供文件存储服务。
- **解绑外网 IP：**支持用户对已绑定外网 IP 的文件存储进行解绑。
- **修改安全组：**支持用户对文件存储实例的外网安全组进行修改。
- **续费：**支持用户对文件存储实例进行手动续费操作。
- **修改告警模板：**支持用户修改文件存储实例的告警模板。
- **修改 IP 地址：**支持用户修改文件存储的内网 IP(VIP)地址。

5.3.3 实例扩容

平台支持用户扩容文件存储的容量，适应于业务发生变化需扩容文件存储容量的场景。平台仅支持扩容文件存储容量，不支持文件存储容量的扩容。

文件存储容量扩容范围即当前硬盘类型的规格，支持的容量范围为 100GB~32000GB。

扩容文件存储容量会对费用产生影响，按小时付费的硬盘，扩容容量下个付费周期按新配置扣费；按年按月付费的硬盘，扩容容量即时生效，并按比例自动补差价。

更改容量，即文件存储需要扩容的容量。平台已展示当前文件存储的容量大小，由于不支持扩容，扩容时更改容量必须大于当前容量大小。

5.3.4 监控告警

平台支持文件存储服务的监控告警服务，包括监控数据和监控告警。

支持监控文件存储实例和从库实例的 CPU 使用率、内存使用率、存储容量、当前存储量、存储容量使用率、存储总写吞吐、存储总读吞吐等。

同时支持用户自定义文件存储告警模板，统一进行告警规则及告警通知的配置和管理，提高运维和监控效率。

5.3.5 日志事件

文件存储服务支持用户获取并查看实例的操作日志及事件日志。

操作日志包括操作（API）名称、所属模块、地域、关联资源 ID、操作者、操作结果、操作时间，并可通过操作结果及操作时间进行筛选。

事件日志包括事件类型、事件等级、事件内容、事件发生次数、开始时间、更新时间，可通过事件周期进行筛选。

5.3.6 备份管理

平台文件存储支持快照备份的方式对文件存储实例进行备份操作，并可结合备份服务进行自动备份任务的执行，全面保证文件存储数据的安全性。

平台支持对文件存储备份进行全面管理，如查看备份信息、删除备份及从备份创建。

- 查看备份信息：平台支持用户查看备份管理信息，如资源 ID、状态、所属存储池、来源备份计划、保留时间(天)、保存路径、创建时间、更新时间及到期时间等。
- 删除备份：支持用户删除备份数据，并支持批量删除。
- 从备份创建：支持用户从备份创建文件存储实例。

5.3.7 文件管理

文件存储服务支持在产品控制台直接进行文件管理，包括文件/目录查看、文件上传、创建目录、删除目录、删除文件、下载文件、显示隐藏文件等。

- 文件目录查看：支持以目录路径结构分别查看上传的目录或文件，支持查看文件的文件名称、文件大小、更新时间等。
- 创建目录：支持用户在控制台创建目录。
- 删除目录：支持用户删除目录，删除后目录中的文件会同步删除。
- 删除文件：支持用户删除文件存储实例中的文件，支持批量操作。
- 下载文件：支持直接下载文件存储服务中的文件。
- 显示隐藏文件：支持显示在目录中的隐藏文件。

5.4 对象存储

5.4.1 概述

对象存储服务 OSS (Object Storage Service) 是云平台提供的 S3 对象存储服务，基于标准 S3 接口为上层应用提供非结构化数据的对象存储，具有高扩展、高可靠、高安全、易接入等特性。

平台对象存储服务兼容亚马逊云标准 S3 协议和 API 接口，整体为云原生设计，即使在高负载的情况下也可以高效利用 CPU 和内存资源，适合私有云场景。

用户可通过对象存储服务提供的 RESTful API 接口、SDK 及对象存储 WEB 桶管理客户端通过网络快速存储或访问对象存储中的海量数据，同时支持通过 S3 接口对存储数据进行备份和归档操作。

对象存储服务提供实例管理、桶管理、QoS 配置、备份管理、日志事件、监报告警等特性，满足对象存储系统的构建和运维场景需求。

通过对象存储服务，用户可一键构建对象存储实例，用户只需在平台上创建

对象存储实例，便可以在任何应用、任何时间、任何地点通过存储和访问任意类型的数据。

支持用户在部署时选择对象存储实例所在的计算集群和存储集群，以适配不同的硬件环境。

- 对象存储计算实例默认配置为 2C2G，CPU 可进行调整，默认支持 2C、4C、8C 等，支持管理员自定义对象存储实例规格。
- 对象存储实例的存储默认 x86 系统盘为 40GB，ARM 系统盘为 100GB。
- 支持用户自定义对象存储实例的存储容量，用户在部署时可选择实例数据存放的存储集群(如 HDD 三副本存储集群、SSD 多副本存储集群等)，支持用户自定义数据存储的容量。

对象存储实例网络必须归属一个 VPC 网络，使虚拟机可直接通过内网挂载或访问对象存储 S3 接口进行数据读写。用户也可对对象存储服务绑定外网 IP 地址，对云平台外的数据中心网络提供对象存储服务，并通过外网安全组的安全组策略，保证网络访问的安全性。

平台支持对象存储实例的各种监控数据，并可根据监控模板进行监控报警及通知，保证存储服务的正常运行。同时支持获取并查看对象存储服务的操作日志及事件日志。

支持在控制台直接进行对象存储服务的存储桶管理，并可在存储桶中直接进行文件上传及管理。同时支持存储桶 WROM 锁定和多版本能力，满足对象存储安全场景的需求。

同时平台支持在线扩容对象存储容量，并支持配置对象存储实例的硬盘 QoS，用于控制对象存储服务的 IOPS 和读写带宽。

5.4.2 实例管理

对象存储服务支持对对象存储服务实例进行全生命周期管理，通过平台一键部署对象存储服务，用户可通过指定计算集群、存储集群、CPU、容量、VPC、子网、外网 IP、外网安全组、项目组、对象存储名称等相关基础信息创建对象

存储等基础信创自动创建一个对象存储实例。

- **计算集群**: 对象存储实例所在的计算集群, 即对象存储服务的虚拟机实例所运行的物理位置。
- **存储集群**: 对象存储实例数据存放的存储集群, 如 SSD 存储集群或 HDD 存储集群。
- **容量**: 对象存储服务的存储容量, 支持的容量范围为 100 ~ 32000G, 并支持管理员自定义对象存储容量规格。
- **VPC/子网**: 对象存储实例运行时所属的 VPC 网络和子网网段, 同时为对象存储实例提供 S3 服务的 IP 访问地址。
- **外网 IP**: 支持为对象存储实例挂载外网 IP 地址, 为平台外网提供对象存储访问服务。
- **外网安全组**: 支持为绑定外网 IP 的对象存储实例绑定安全组, 通过安全组规则的出入规则配置, 保证对象存储对外提供服务的网络安全性。
- **通用信息配置**: 支持为对象存储配置名称、管理员密码备注及标签等设置。

支持用户查看对象存储实例信息, 如状态、存储容量、域名、VPC、子网、计费方式、项目组、创建时间及过期时间等。

同时支持对对象存储实例进行扩容、QoS 配置、桶管理、绑定外网 IP、解绑外网 IP、修改安全组、重置密码、修改告警模板及修改 IP 等管理。

- **实例扩容**: 平台支持用户扩容对象存储的容量, 适应于业务发生变化需扩容对象存储容量的场景。仅支持扩容对象存储容量, 不支持对象存储容量的缩容。
- **QoS 配置**: 支持配置对象存储实例的硬盘 QoS, 用于控制对象存储服务的 IOPS 和读写带宽。硬件介质和容量会影响硬盘的读写 IOPS 和带宽速率, 若配置的 QOS 超过硬件本身性能, 以硬件性能为准。系统会默认分配 QOS 值, 如需取消一块硬盘的限速功能, 可将 IOPS 和宽带

均配置为 0。

- 绑定外网 IP：支持用户对对象存储实例绑定外网 IP，对外网提供对象存储服务。
- 解绑外网 IP：支持用户对已绑定外网 IP 的对象存储进行解绑。
- 修改安全组：支持用户对对象存储实例的外网安全组进行修改。
- 重置密码：支持用户重置对象存储的管理员登录密码。
- 续费：支持用户对对象存储实例进行手动续费操作。
- 修改告警模板：支持用户修改对象存储实例的告警模板。
- 修改 IP 地址：支持用户修改对象存储的内网 IP(VIP)地址。

5.4.3 实例扩容

平台支持用户扩容对象存储的容量，适应于业务发生变化需扩容对象存储容量的场景。平台仅支持扩容对象存储容量，不支持对象存储容量的缩容。

对象存储容量扩容范围即当前硬盘类型的规格，默认为 100GB~1024 GB。

扩容对象存储容量会对费用产生影响，按小时付费的硬盘，扩容容量下个付费周期按新配置扣费；按年按月付费的硬盘，扩容容量即时生效，并按比例自动补差价。

更改容量，即对象存储需要扩容的容量。平台已展示当前对象存储的容量大小，由于不支持缩容，扩容时更改容量必须大于当前容量大小。

5.4.4 监控告警

平台支持对象存储服务的监控告警服务，包括监控数据和监控告警。

支持监控对象存储实例和从库实例的 CPU 使用率、内存使用率、对象个数、存储容量、当前存储量、存储容量使用率、存储总写吞吐、存储总读吞吐等。

同时支持用户自定义对象存储告警模板，统一进行告警规则及告警通知的配

置和管理，提高运维和监控效率。

5.4.5 日志事件

对象存储服务支持用户获取并查看实例的操作日志及事件日志。

操作日志包括操作 (API) 名称、所属模块、地域、关联资源 ID、操作者、操作结果、操作时间，并可通过操作结果及操作时间进行筛选。

事件日志包括事件类型、事件等级、事件内容、事件发生次数、开始时间、更新时间，可通过事件周期进行筛选。

5.4.6 备份管理

平台对象存储支持快照备份的方式对对象存储实例进行备份操作，并可结合备份服务进行自动备份任务的执行，全面保证对象存储数据的安全性。

平台支持对对象存储备份进行全面管理，如查看备份信息、删除备份及从备份创建。

- 查看备份信息：平台支持用户查看备份管理信息，如资源 ID、状态、所属存储池、来源备份计划、保留时间(天)、保存路径、创建时间、更新时间及到期时间等。
- 删除备份：支持用户删除备份数据，并支持批量删除。
- 从备份创建：支持用户从备份创建对象存储实例。

5.4.7 桶管理

存储桶管理是平台为用户提供的控制台对象存储桶及文件管理服务，支持私有、公共读、公共读写三种访问类型，并支持多版本及桶锁定等特性。

- 私有类型：私有类型的存储桶中所有文件必须获取拥有者的授权才可进行访问。
- 公共读：公共读类型的存储桶中的所有文件均可直接通过 URL 进行文件的读访问。

- 公共读写：公共读类型的存储桶中的所有文件均可直接通过 URL 进行文件的读写操作。

支持存储桶多版本保护管理，包括开启和关闭版本保护，开启后存储桶内上传的同名对象文件会保留历史版本。通过多版本保护，可对特定时间点的对象进行恢复和下载。

桶锁定是指将桶进行锁定，支持一次写入多次读取(WORM)模型存储对象，保证存储桶内对象文件无法被修改或删除，对关键数量实行写保护，杜绝病毒破坏或非法篡改。桶锁定天数可进行自定义配置，支持范围为 1~36500 天。

注意 对象锁定依赖多版本功能，开启后无法关闭。

5.4.8 桶文件管理

平台支持对桶进行创建和删除管理，并可对桶内文件进行管理，文件管理包括文件/目录查看、文件上传、创建目录、删除目录、删除文件、下载文件、清除当前桶及上传列表查看等。

- 文件目录查看：支持以目录路径结构分别查看上传的目录或文件，支持查看文件的文件名称、文件大小、更新时间等。
- 创建目录：支持用户在控制台创建目录。
- 删除目录：支持用户删除目录，删除后目录中的文件会同步删除。
- 删除文件：支持用户删除文件存储实例中的文件，支持批量操作。
- 下载文件：支持直接下载对象存储桶中存放的文件。
- 清除当前桶：清空存储桶将删除您存储桶中的所有文件、历史版本、文件碎片、删除后数据不可恢复和访问。
- 上传列表查看：支持用户实时查看上传列表的进度。

除桶文件基本管理外，支持历史版本的查看，包括历史版本的文件名、文件大小、更新时间及还原等。

6 运维运营管理

6.1 统一管理服务

云平台为底层异构计算、存储、网络等基础设施提供统一管理服务，支持将 x86、ARM、龙芯、申威、GPU、SSD 存储、HDD 存储、商业存储等多种架构物理资源软件定义为统一抽象为计算和存储集群资源，为上层客户提供云资源服务能力。

平台以不同的集群划分不同配置、不同架构、不同用途的服务节点，可在一个数据中心下部署多个不同类型的计算集群或存储集群，并通过一套云平台统一进行资源分发配置和管理，打平不同架构资源的管理面。

用户在平台上进行不同架构资源的部署时，只需选择适合的集群即可创建出相应架构的资源，如选择 GPU 集群，则可以创建 GPU 虚拟机，并可对所有资源进行统一的生命周期管理。

同时针对不同架构的资源，云平台通过云资源化的统一抽象，统一提供资源接入、账号认证、资源管理、计量计费、配额管理、监控告警、日志事件、统计分析及权限控制等，从下而上形成一个统一的整体，满足从数据中心到云服务的管理与运营场景需求，

云平台统一管理服务为企业用户提供租户视角和管理视角两套控制台，租户控制台用于平台租户虚拟资源管理，管理控制台用于平台管理者对云平台整体的运营运维。

- 租户控制台是为租户（主/子帐号）提供的 Web 云资源管理平台，支持租户主/子帐号管理及权限控制，可对云平台所有产品服务的虚拟资源进行全生命周期管理，同时可对虚拟资源进行监控告警、计量计费及日志审计等管理。
- 管理控制台是为平台管理者及运维人员提供的 Web 运营运维管理平台，使用管理员账号统一管理整个云平台全局，拥有平台所有管理权限，包括全平台账号认证、多租户管理、资源管理、流程审批、计量计费、监

控告警、审计日志、平台管理及运维迁移等管理功能模块，同时提供大屏监控服务，支持自由布局灵活投屏，提升平台资源的数据可视化效果。

为方便整个云平台资源的统一运维和运营，云平台在账号认证体系上提供平台管理账号、多租户管理、多地域管理全局资源视图、物理资源管理、虚拟资源管理、QoS 配置、资源模板、标签管理、监控告警、通知组、操作日志、资源事件、回收站、计量计费、审批流程、报表统计、大屏监控及 API 控制台等服务，全面覆盖云平台运营运维的使用场景。

6.2 平台管理账号

6.2.1 管理员概述

平台组织和账号管理分为租户侧和管理员侧两个维度。管理员侧认证通过系统管理员账号及权限提供。管理员用于全局管理和运营整个云平台，可通过管理员账号管理云平台的地域、集群、租户、管理员账号、资源、计费、审批、安全及平台全局配置等。

在部署私有云的客户视角，平台账号分为系统管理员和租户。系统管理员 Admin 账号是管理员控制台账号，可创建系统管理员和地域管理员，系统管理员和地域管理员都属于管理员，管理员角色体系内容如下：

角色等级	默认	数据范围	功能范围
系统级	系统管理员	系统+地域	地域管理、集群管理、物理资源管理、云资源管理、网络服务管理、容器服务管理、账号和组织管理、运营与管理、运维与管理、监控大屏、自定义 UI 管理、全局配置管理、服务目录管理、充值管理、订单管理、交易管理
	系统只读管理员	系统+地域	地域查看、集群查看、物理资源查看、云资源查看、网络服务查看、容器服务查看、账号和组织查看、运营与查看、运维与查看、监控大屏、自定义 UI 查

			看、全局配置查看、服务目录查看、充值查看、订单查看、交易查看
地域级	地域管理员	地域	地域管理、集群管理、物理资源管理、云资源管理、网络服务管理、容器服务管理、账号和组织管理-我的账号、运营与管理、运维与管理、监控大屏、全局配置-产品策略管理、服务目录管理、订单管理、交易管理
	地域只读管理员	地域	地域查看、集群查看、物理资源查看、云资源查看、网络服务查看、容器服务查看、账号和组织管理-我的账号查看、运营查看、运维查看、监控大屏、全局配置-产品策略查看、服务目录查看、订单查看、交易查看

系统管理员拥有平台所有管理权限，用于全局管理和运营整个云平台。可通过管理员账号管理云平台的地域、集群、租户、管理员、资源、计费、审批、安全及平台全局配置。

地域管理员拥有平台特定地域下的管理权限，用于运营整个云平台。可通过管理员账号管理云平台特定地域下的集群、资源、计费、审批及平台全局配置。

平台账号体系使用邮箱地址作为平台的登录账号，在使用前需确保提供的账号邮箱地址可用，方便接收告警邮件或找回密码等。

平台默认使用 `admin@ucloud.cn` 邮箱地址作为平台的管理员账号，使用管理员账号登录控制台即会自动登入管理员控制台，使用租户账号登录控制台会自动登入租户控制台。

6.2.2 管理员账号安全

作为平台管理员账号，拥有平台最高权限，针对管理员账号自身的安全平台提供修改登录密码、找回密码、双因子验证登录、登录访问限制、修改登录邮箱等安全防护功能。

- 修改登录密码：在支持在使用管理员账号时，更改管理员密码。

- 找回密码：平台支持管理员账号在忘记密码时通过控制台自主找回密码，找回密码时需通过邮箱进行验证，需确保管理员添加的账号为真实可用的邮箱
- 双因子验证：平台为管理员账号提供免费的基于 TOTP（Time-Based One-Time Password Algorithm）登录二次认证服务，开通本服务后，管理员每次登录控制台均需通过授权认证。支持国密硬件版和普通软件版，用户可根据需要通过部署进行配置，
- 登录访问限制：为保证账号登录的安全及对特定安全场景的需求，平台提供账号登录访问限制能力，为管理员账号设置登录控制台和访问 API 的客户端 IP 地址。
 - 配置后管理员账号只能从指定的 IP 登录或发起 API 访问，保证管理员登录及资源管理的安全性。
 - 支持配置多个 IP 地址或 IP 地址段，多个 IP 地址/段间使用英文逗号进行分隔。
 - 配置的 IP 地址或 IP 地址段为白名单模式，即配置的 IP 地址/段客户端才可正常登录控制台或访问 API。
 - 默认不指定任何 IP，代表不限制登录控制台和访问 API 的客户端 IP 地址，即默认全网可访问登录控制台。
- 修改登录邮箱：平台支持管理员可通过管理员控制台右上角管理员账号头像中的【修改登录邮箱】来进行修改管理员的账号邮箱地址，用于将管理员账号修改为真实可用且实际需要接收告警邮件的邮箱地址。

同时平台支持获取系统管理员和地域管理员的 API 密钥信息，用于通过 API 接口查询平台全局资源及运行信息。

6.2.3 管理员账号管理

(1) 自定义创建管理员

平台管理员账号可创建系统管理员和地域管理员，地域管理员不可创建管理员，创建管理员时需要添加一个账号邮箱作为管理员的账户。

可同步设置账号首次登陆强制修改密码，同时支持对账户进行管理设置管理类别、是否只读及开通地域。

(2) 冻结/解冻管理员

冻结管理员是指将一个管理员进行锁定，被成功冻结的管理员将无法登录云平台行。仅支持状态为【使用中】的管理员进行冻结操作。

当管理员被冻结后，管理员的状态为冻结中，支持具有管理权限的系统管理员解冻管理员，管理员解冻后，可正常登录控制台。

(3) 地域授权管理

地域授权管理可管理一个地域管理员在地域下的授权情况。只有在授权地域下，地域管理员才可在该地域正常使用服务。企业可以根据云平台实际运营情况管理系统管理员在对应地域的开通情况。

(4) 删除管理员

支持具有管理权限的系统管理员删除管理员账号，管理员账号被删除后，无法再登录平台。

6.2.4 管理员权限管理

平台支持通过角色对账号进行权限管理，以实现平台细粒度的权限控制。

角色是一组权限的集合，通过对资源权限进行集合配置管理，可支持的资源包括虚拟机、镜像、虚拟硬盘、快照、安全组、资源模板、监控告警、扁平网络等资源，并支持对资源的增、删、改、查及相关管理 API 的操作权限进行开启和关闭，以满足多场景用户权限管理及控制场景。

平台提供内置角色，根据地域性或全局性以及只读或管理维度，默认提供如下四种角色：


- 系统管理员：包含了平台所有资源操作和管理授权，属于平台最高权限

角色。

- 系统只读用户：对平台所有资源有查看权限，通常为平台运维管理人员赋予该角色。
- 地域管理员：相比系统管理员角色，无账号管理等全局性权限。为账号赋予角色的同时可选定部分地域，以实现账号对指定地域的所有资源有操作和管理授权。
- 地域只读用户：对指定地域的所有资源有查看授权，通常为地域运维管理人员赋予该角色。

内置角色是平台根据应用场景提供的默认角色，可进行快速权限配置，若内置角色无法满足场景需求，支持用户自定义创建角色，并对角色进行修改、删除管理，同时支持查看一个角色已绑定的账号。

创建管理员账号时，支持设定授权范围，【系统管理员】表示授权所有地域，【地域管理员】只对选定的地域进行授权。同时通过所选角色对账号进行功能授权，即账号只拥有角色中定义的权限集合。平台通过授权范围及角色控制账号的权限以及资源范围，以实现平台细粒度的权限控制。

 **说明** 角色授权时，支持多个角色一起授权，最终账号获得的权限为多个角色的权限的并集，即所有角色权限的集合。

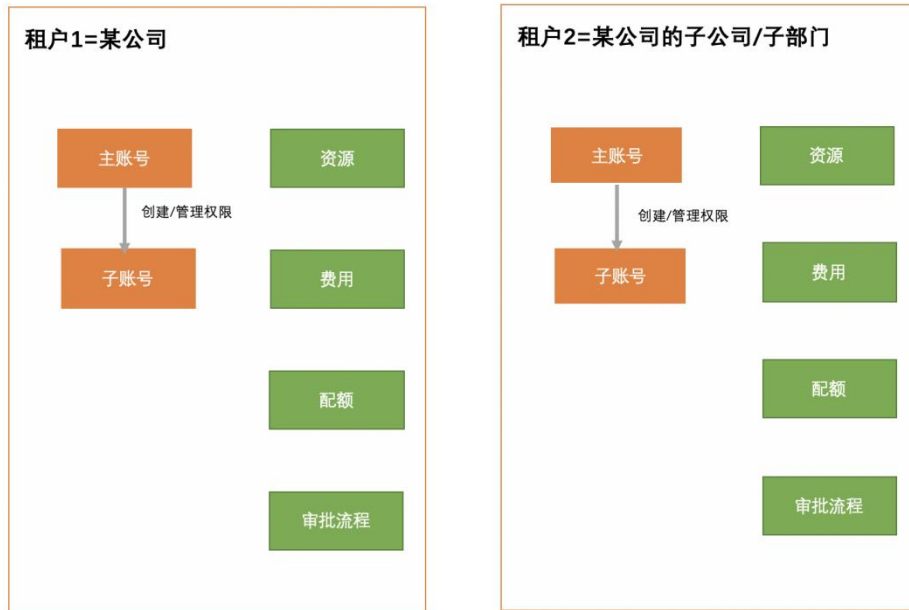
6.3 多租户管理

6.3.1 概述

平台支持多租户模式，用于有多级组织架构的企业，可将租户作为一个单独的公司进行运营，有效实现权限管理，降低总公司、子公司及不同部门资源混用造成的风险，并可实现对资源的审计。

租户主要为企业用户提供组织架构管理，是平台中一组资源的集合，提供资源隔离、子账号管理、权限控制、配额及价格配置等能力。不同租户间资源通过VPC网络及权限实现强隔离，所有主账号和子账号的资源、使用的费用、资源

的配额及流程审批均以租户为单位。



租户可由平台系统管理员创建，创建租户时默认会生成一个用户账号，此账号即为租户的主账号，也是租户下的管理人员；注册成功后需要管理员进行充值才可进行使用。平台多租户管理的基本概念如下：

- **主账号**

一个租户必须有一个主账号，主账号默认有租户下所有资源及管理的全部权限。可通过主账号创建和管理子账号，并管理子账号的权限。

- **子账号**

子账号是主账号创建的用户，子账号在租户下的权限由主账号控制。一个租户可拥有多个子账号，支持对子账号进行资源管理的权限控制。

- **人员**

企业中的人员，人员需要使用账号登录云平台使用资源。

- **角色**

权限的集合，为用户和成员组赋予权限可获得调用相关 API 进行资源操作的能力。

- **项目组**

以项目组为维度进行资源规划,可为一个具体项目或者业务建立独立的资源池,实现资源更细粒度地管理。同时针对子账号的授权也是基于项目组维度进行授权。项目组只是逻辑上面的分组,不具有资源隔离的作用,租户所有资源均需要属于某个项目组。

- **流程审批**

为满足企业对核心云资源,如虚拟机、云硬盘、外网 IP 等资源使用的管控需求所引入的云资源工单审批流程。

- **审批管理**

审批管理仅平台管理员 `admin` 可以进行操作。

平台支持系统管理员对全局租户进行运营和管理,同时为租户的主账号提供自服务管理能力。

6.3.2 租户管理

支持管理员创建租户,创建租户时需要添加一个主账号作为租户的初始管理人员,支持同步设置账号登陆安全策略,首次登陆强制修改密码。

创建时可对租户进行管理设置,包括是否开启资源审批,自动审批,以及设置租户开通使用的地域。

账户创建后余额默认为 0,需要进入管理员的充值管理页面,为租户充值后,租户下的账号才可正常创建和使用资源。

平台系统管理员可对平台的所有租户进行生命周期管理,如激活账号、修改名称、冻结、解冻、登录访问限制、开关审批流程、地域授权、修改邮箱、删除租户。同时可对租户进行主/子账号管理、配额管理、订单管理、交易管理、充值管理、价格管理及资源概览查看。

- **激活租户:** 平台用户自行注册的账号默认为未激活状态,需用户通过激活链接激活账号后才可进行使用,同时平台支持管理员手动激活账号,

使账号可正常登录并管理平台。

- 冻结租户：将一个租户进行锁定，被成功冻结的租户，主/子账号将无法登录云平台，不影响租户内已创建的资源及业务的正常运行。
- 解冻租户：租户解冻后，默认租户中的所有账号均会被解冻，并可正常登录控制台。
- 设置登录限制：登录访问策略决定可以登录控制台和访问 API 的客户端 IP 地址，配置后账号只能从指定的 IP 登录或发起 API 访问，默认不指定任何 IP，代表不限制登录控制台和访问 API 的 IP 地址。
- 开关辑审批流程：开启或关闭一个租户下所有账号变更资源的审批流程方式，包括开启/关闭资源审批及开启/关闭自动审批。
- 地域授权管理：通过地域授权管理可管理一个租户在地域下的授权情况。只有在授权地域下，租户下的账号才可在该地域正常使用服务。
- 修改邮箱：支持管理员修改一个租户主账号的登录邮箱地址。
- 修改密码：支持管理员修改一个租户主账号的登录密码。
- 删除租户：支持管理员删除租户，删除前先进行账号冻结。
 - 若租户下存在资源，则不可进行删除，需清空所有资源，才可进行租户删除。
 - 删除租户后，租户下所有用户账号被清除，无法使用账号登录平台。
- 子账号管理：支持查看并管理租户下已有主账号和子账号，包括查看账号信息、修改密码、冻结账号及解冻账号。
- 订单管理：支持管理员查看单租户内的资源订单信息。
- 交易管理：支持管理员查看租户内的交易记录信息。
- 资金管理：管理员可为租户进行充值操作，并可查看单租户的充值记录信息；同时可为租户进行提现操作，并可查看租户的提现记录信息。

- 配额管理：管理员可查看每个租户的当前配额信息，并可通过配置每个租户的在每个地域的资源配额。
- 价格管理：管理员可查看租户当前每个产品的价格，并支持管理员对租户每种产品设置不同的折扣，如 9 折。
- 资源概览：支持管理员查看每个租户当前的资源使用概览。

在第三方账号体系方面，支持 OAuth 2.0 登录认证，用户可通过将企业内 OAuth 统一认证登录系统与云平台进行对接，使用企业统一登录用户即可登录并使用云平台的资源。

在平台成功对接 OAuth 认证后，在登录页面会提供第三方登录入口。企业用户可通过自有的 OAuth 统一认证平台登录跳转至云平台，同时也可通过云平台第三方登录入口通过统一用户密码认证登录云平台，

6.3.3 租户自服务

平台提供主账号自服务模式，支持自助注册流程，用户可通过注册链接，自动化的进行注册并使用云平台。账户创建后余额默认为 0，需通过平台管理员进行账号充值，为租户充值后，租户下的账号才可正常创建和使用资源。

主账号自服务支持找回密码、修改登录密码、开启登录保护及登录访问限制等管理，并支持自行查看账号的基本信息、配额信息及面向开发者的 API 密钥。

- 找回密码：支持主账号在忘记密码时通过控制台自主找回密码，找回密码时需通过邮箱进行验证。
- 修改登录密码：平台支持用户修改账号登录密码，修改密码需要验证旧密码，若忘记旧密码，可联系管理员在后台帮助修改密码。
- 开启登录保护：平台提供基于 TOTP (Time-Based One-Time Password Algorithm) 的免费登录二次认证服务，开通本服务后，账号登录控制台均需通过授权认证，支持国密硬件版和普通软件版，用户可根据需要通过部署进行配置。

- 登录访问限制：支持主账号自行设置可登录控制台和访问平台 API 的客户端 API 地址。配置后租户下的所有账号只能从指定的 IP 登录或发起 API 访问，有效保证账户登录及资源的安全性。
- 账号信息：支持主账号自行获取账号的基本信息、API 密钥及配额信息。
 - API 密钥：提供当前账号的 API 密钥，用于管理并使用 API 接口。
 - 配额信息：提供当前账号的配额信息，包括虚拟机数量、虚拟机 CPU、虚拟机内存、VPC 数量、子网数量、云硬盘数量、去硬盘容量、弹性网卡数量、外网 IP 数量、负载均衡数量、SSL 证书数量、虚拟机模板数量、NAT 网关数量、安全组数量、对象存储数量、文件存储数量、VPN 网关数量、对端网关数量、VPN 隧道数量、弹性伸缩组数量、伸缩策略数量、VIP 数量等。

6.3.4 账号权限管理

6.3.4.1 概述

平台支持对租户下的子账号进行权限管理，以项目组、角色权限、人员（子账号）及资源的授权绑定，实现平台资源级权限控制，如将租户下部分虚拟机授权给一个人员进行管理。具体权限工作机制和流程如下：

(1) 以项目组进行资源分组管理

平台以项目组为维度进行资源规划，可为一个具体项目或者业务建立独立的资源池，实现资源更细粒度地管理。同时针对子账号的授权基于项目组维度进行授权。

项目组帮助解决资源的精细化管理以及授权管理等复杂性问题。租户所有资源均需要属于某个项目组，分组的资源根据用户的角色授权决定用户是否对组中资源具有权限。

项目组只是逻辑上面的分组，不具有资源隔离的作用。主账号默认拥有租户下所有资源的管理权限，在主账号下创建资源时可不进这行项目组的配置。子账

号创建资源时若不配置项目组则代表资源为默认项目组，若配置项目组则仅授权了项目组权限的子账号可进行查看或管理。

(2) 以角色权限集群资源权限

角色是一组权限的集合，包括系统默认内置角色和自定义角色。内置角色是平台根据应用场景提供的默认角色，可进行快速权限配置，若内置角色无法满足场景需求，可对角色进行自定义配置。

角色通过对资源权限进行集合配置管理，可支持的资源包括虚拟机、VPC、自制镜像、外网 IP、组播、NAT 网关、负载均衡、裸金属、硬盘、快照、备份服务、弹性伸缩、VIP、VPN 网关、商业存储、安全组、资源模板、MySQL、Redis、对象存储、文件存储、监控告警、API 控制台及弹性伸缩等所有资源，并支持对所有资源的增、删、改、查及相关管理 API 的操作权限进行开启和关闭，以满足多场景用户权限管理及控制场景。

用户可将平台所有资源类型的 API 操作项分类整合，形成符合特定场景的角色，并将角色和资源授权给子账号，使子账号具备所授权资源的角色权限，实现资源级精细化权限管控。

(3) 以人员管理管理子账号并进行角色授权管理

人员即子账号，是云平台内的一个账号实体，代表组织内需要访问云资源的人员，租户下支持创建一到多个子账号。

可通过角色授权进行权限管理，将角色、项目组资源与人员进行关联，实现将全部或所有资源以一种角色权限的集合授权给子账号，使子账号具备授权资源的角色及权限。

支持通过人员管理对租户下所有账号进行管理，建议谨慎地创建和管理组织下的账号，并为账号添加上适当的角色授权，防止权限扩大化，导致资源管理的混乱，影响企业 IT 资源使用的安全。

通过账号权限管理的配置，可分别分配系统管理员、安全管理员及安全审计员三种角色，并分别赋予系统资源管理权限、安全配置和权限配置、日志事件审

计权限，亦可实现三权分立的账号管理场景。

6.3.4.2 项目组管理

用户在创建资源时可选择资源所属的项目组，将资源加入到一个项目组，并可在后续将项目组的资源及角色权限与子账号进行关联，使子账号对项目组内的资源有相应的管理权限。

平台支持对项目组进行创建、删除、查看等基本管理，同时支持对项目组进行资源的转入转出管理。当项目组中存在资源时，无法删除项目组，需要将资源转移到其他项目组。

项目组管理通过转入转出资源管理项目组下的资源，具体说明如下：

- 转入资源到项目组时，可选择不在当前项目组下的所有资源，可以在资源列表处查看到资源所属项目组。
- 转出资源到其他项目组时，仅支持指定当前项目组下的资源，且转出到的项目组只能指向其他项目组。
- 资源被转出到其他项目组后，有该项目组授权的用户将无法再查看和管理此资源。
- 资源转入/转出时以产品类型维度，一次可以转入/转出同一产品类型下的多个资源。

注意 VPN 网关、负载均衡、伸缩组等相关产品服务均有多个维度资源，在转入转出时需将关联资源分配到同一组内。

6.3.4.3 角色管理

角色是一组权限的集合，可将平台所有资源类型的 API 操作项分类整合，形成符合特定场景的角色，并将角色和资源授权给子账号，使子账号具备所授权资源的角色权限，实现资源级精细化权限管控。

角色管理支持系统内置角色，默认包括管理员和只读用户两种角色：

- 管理员角色包含了租户下所有云资源操作和组织管理操作的授权。若对子账号无需进行精细地权限管理，可直接使用系统管理员为子账号进行角色授权，更快捷地进行操作。
- 只读用户：包含所有资源和组织管理功能的查看权限。

注意 平台提供自定义角色的能力，支持用户自定义权限集群创建角色，并对角色进行修改、删除管理，同时支持用户查看一个角色已绑定的子账号。

6.3.4.4 人员管理

人员即子账号，是云平台内的一个账号实体，代表组织内需要访问云资源的人员，租户下支持创建一到多个子账号。

可通过角色授权进行权限管理，将角色、项目组资源与人员进行关联，实现将全部或所有资源以一种角色权限的集合授权给子账号，使子账号具备授权资源的角色及权限。

支持通过人员管理对租户下所有账号进行管理，可查看当前租户下所有子账号的相关信息，同时支持子账号创建、冻结子账号、解冻子账号、删除子账号及角色授权管理。

- 创建子账号：具有子账号创建权限的账号可进行子账号的创建管理，并支持设计子账号的名称、邮箱、密码、授权范围、角色，并可开启首次登录修改密码。
- 冻结子账号：冻结账号是指将一个账号进行锁定，冻结后将不允许登录控制台。
- 解冻子账号：解冻账户是指解冻一个已冻结的账号，被成功解冻的的用户，可登录管理控制台。
- 删除子账号：可对账户进行删除，删除账号后用户无法再使用此账号密码登录平台。删除后，可再使用账号邮箱在平台注册或创建新的账号。
- 角色授权管理：支持对子账号进行角色授权管理，将全部资源或指定项

目组的资源，以某个角色授权给子账号，使子账号对指定资源拥有指定角色的权限。

注意 角色授权时，支持多个角色一起授权，最终子账号获得的权限为多个角色的权限的并集，即所有角色权限的集合。

6.4 多地域管理

6.4.1 概述

地域 (Region) 是云平台中的一个逻辑概念，指资源部署的物理位置分类，可对应机柜、机房或数据中心。通常一个数据中心对应一套 UCloudStack 云平台，可支持部署多个计算和存储集群；数据中心之间资源和网络完全物理隔离，可通过一套管理平台管理遍布各地数据中心的私有云平台。

- 地域在平台也称为数据中心，通常数据中心之间完全隔离以保证最大程度的稳定性和容错性。
- 作为平台最大的资源定义，一个地域即部署一套云平台。平台默认内置一个地域，管理服务通过本地数据中心云平台提供的 API 端点管理地域内计算、存储及网络资源。
- 支持对数据中心内资源的生命周期管理，包括计算集群、存储集群、外置存储、基础镜像及自制镜像等资源的查看和维护。

多地域管理是指在一个组织或企业内部建立和管理多个数据中心，并将多个数据中心部署在不同的地理位置，提供更高的可用性、容错性和灵活性，使组织能够更好地满足地理分布和业务需求。

多地域管理支持在多个数据中心部署 UCloudStack 私有云平台，多套云平台通过统一管理界面进行运维运营管理，租户可在统一控制台使用多个地域下的私有云资源。

在私有云的部署结构上可以将一个机柜、模块或机房当作一个地域或数据中心，部署一套私有云平台，可应用于边缘节点构建、双活数据中心、两地三中心

或政务云构建等场景。

多套 UCloudStack 云平台均通过上层统一云管平台进行统一管理，多套云平台共享一套账号认证、计量计费、监控告警、审计日志、API 网关、操作界面及相关通用套件，每个数据中心的云平台资源均部署于自己的地域中，通过地域标识保证资源的隔离和安全性。

6.4.2 多地域特性

多地域管理通过一套控制台对多个地域的多套 UCloudStack 平台进行统一管理和调度，具有多地域部署、统一管理、安全隔离及平滑扩展等特性。

● 多地域部署

多地域管理允许将私有云数据中心分布在不同的地理位置上，可以是不同的城市、国家或洲际，用于减少地理位置带来的单点故障风险，提高系统的可靠性和可用性。

● 统一管理

为方便管理和监控多个地域的私有云数据中心，平台提供统一管理服务，支持对各个地域的集中管理，包括资源配置、监控和故障排除等能力。租户可在平台一键切换地域，并进行资源创建和管理；平台管理员可通过管理控制台统一管理整个私有云多地域架构，并为租户分配地域使用权限。

● 安全与隔离

私有云多地域管理，提供安全和隔离机制，以确保各个地域之间的数据和资源不被未授权的访问。具备网络隔离、访问控制手段来实现，确保地域之前的独立性和安全性。

● 平滑扩展

多地域管理具备平滑的扩展能力，以应对业务的变化和需求的增长。当需要增加新的地域或调整已有地域的规模时，多地域管理机制不影响现有业务，进行扩展和调整，而不会对整个架构产生太大的影响。

6.4.3 多地域管理能力

面向租户，平台支持资源多地域统一管理，通过地域切换按钮可进行不同地域的管理，在不同的产品服务页面，切换地域按钮将会仅展示和操作当前地域的资源和管理。

面向平台管理员，平台支持管理当前云平台已部署的所有地域，包括地域信息的查看、编辑地域信息、查看地域资源用量统计及地域资源管理等。

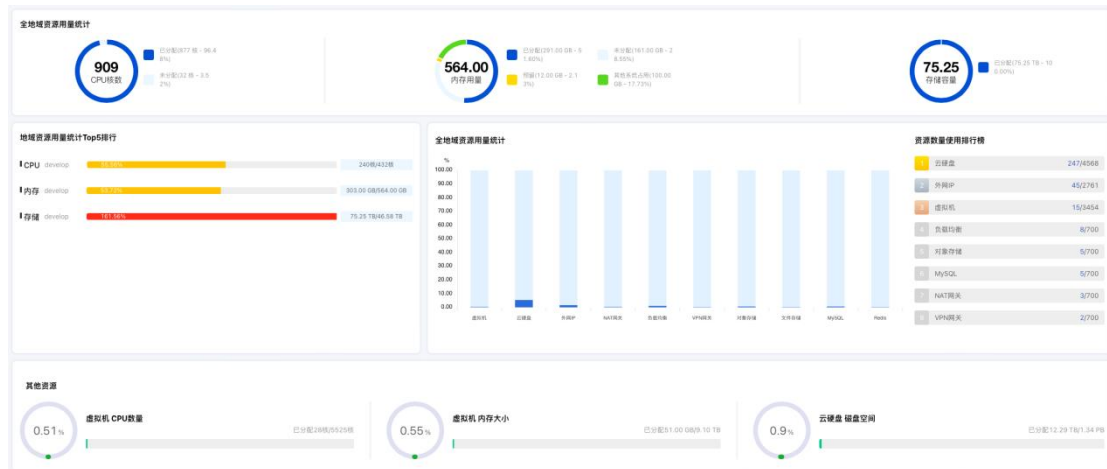
同时支持对地域进行运营管理，将地域授权给一个或部分租户，满足地域资源独享的专属云场景。

- 地域信息获取：支持系统管理员查看平台管理的所有地域列表，以及查看地域下 CPU 核数用量、内存用量、存储用量物理 GPU 用量及授权状态。
- 地域资源用量统计：支持查看当前地域以及全部地域的资源使用情况，按照已分配个数和总配额生成扇形图，已分配数量为用户创建的资源总数，总配额为当前地域所有租户配额总和。支持查看多地域中资源使用率最高的前五个地域排行展示。
- 地域资源管理：提供全平台所有每个地域物理资源的生命周期管理和运维能力，使平台管理员可通过控制台统一管控地域中整体物理资源，包括物理机资源（节点）、物理机纳管、镜像资源及外网网段资源。
- 地域授权管理：通过地域授权管理可管理一个地域管理员在地域下的授权情况。只有在授权地域下，租户才可在该地域正常使用服务。企业可以根据云平台实际运营情况管理系统管理员在对应地域的开通情况。
- 私有云平台支持全局资源管理，包括物理资源管理和虚拟资源管理，全面满足云平台运营管理者对云平台资源的自主可控性需求。

6.5 全局资源视图

平台提供全局资源管理视图，针对多地域资源使用情况提供总体视图展示，

支持查看单地域和全部地域资源的使用情况，方便平台管理者统计所有物理资源及云资源的使用状况，如使用率、使用量及资源分配数量等。



支持平台所有地域或单个地域的资源用量统计，如 vCPU、内存及存储资源的资源总量、已分配量及未分配量。

支持平台所有地域或单个地域的配额用量统计，如虚拟机 CPU 数量、虚拟机内存大小、磁盘空间、虚拟机资源数量、云硬盘资源数量、EIP 资源数量、NAT 网关资源数量及相关产品的配额的总数及已分配数量。

支持平台所有地域或单个地域的 vCPU、vCPU、内存、存储资源的使用率排行前五统计。

6.6 物理资源管理

物理资源管理是对平台的物理及事实存在的资源进行管理和调度，包括物理节点管理、集群管理、镜像管理、外网网段管理及专线接入管理等，集群管理详见基本概念【**集群**】章节描述。

6.6.1 物理节点管理

节点管理是指对地域内的所有计算节点的管理，包括查看节点信息、锁定、解锁、进入维护模式、退出维护模式及修改告警模板等，同时可支持对每个计算节点中已存在的计算实例、USB 设备、网络设备、磁盘设备及集群信息进行查看和管理。

(1) 节点信息

支持查看节点的资源 ID、IP 地址、CPU 信息（如 CPU 型号、CPU 核心、CPU 槽数、CPU 线程）、CPU 总量、内存总量、序列号、状态、架构、节点类型、地域及 NTP 信息。

同时支持查看节点所属的集群信息，如集群 ID、集群类型、CPU 用量、内存用量、GPU 用量、物理 GPU 用量、vGPU 用量、地域等。

(2) 计算实例

支持管理员通过节点详情页面，查看节点中的计算实例列表及信息，包括名称、计算实例 ID、资源 ID、所属租户、节点 IP、镜像 ID、GPU、CPU、内存、状态、创建时间及更新时间。

(3) 监控告警

支持查看节点机器的监控信息，包括：网卡入带宽、网卡出带宽、硬盘读吞吐、硬盘写吞吐、平均负载、内存使用率、空间使用率、硬盘读此书、网卡入包量、硬盘写次数、网卡出包量、CPU 使用率、TCP 连接数、阻塞进程数。

支持管理员根据节点监控指标修改相应的告警模板，用于监控节点的健康状态，并支持多种方式的告警通知。

(4) 锁定节点

节点被锁定后，新建计算实例不会被调度至计算节点，不影响节点内已有计算实例，可配合节点进入维护模式功能，以实现节点维护、升级等操作。

(5) 解锁节点

管理员将锁定的节点进行解锁，可对外提供计算服务，计算实例可被调度并部署至节点。

(6) 进入维护模式

当需要维护节点时，比如扩展内存、升级、修复硬件等维护场景下，平台支持将节点进入维护模式，使节点上的虚拟资源自动迁移至同计算集群中其他物理

节点上，使节点处于空闲状态，确保对物理节点维护时不影响平台的虚拟资源运行，保证业务的可用性。节点进入维护模式前必须保证节点状态为已锁定。

(7) 退出维护模式

退出维护模式是指将节点重新加入至调度系统，为平台提供计算能力。仅支持状态为【维护模式】的节点退出维护模式，退出维护模式，将会自动恢复并进入至锁定状态，需进行解锁才可加入智能调度系统以提供计算能力。

(8) USB 设备管理

支持管理员通过节点页面，查看节点中 USB 设备列表及信息，包括设备名、设备 ID、状态、厂商、类型、序列号、虚拟机、USB 版本、所属租户，并支持对 USB 进行租户分配。

6.6.2 镜像管理

镜像是虚拟机所使用的镜像模板文件，如 CentOS、Windows、Ubuntu、Debian 等操作系统模板文件，平台的镜像文件均为 QCOW2 格式。镜像管理是平台为虚拟机提供的镜像仓库，支持基础镜像和自制镜像两种类型：

基础镜像是由平台官方默认提供，包括多发行版 Centos、Debian、Ubuntu 及 Windows 等原生操作系统；

自制镜像由租户或管理员通过虚拟机自行导出或自定义导入的自有镜像，可用于创建虚拟机，除平台管理员外仅账号自身有权限查看和管理。

(1) 基础镜像

平台默认会提供多发行版 Centos、Debian、Ubuntu 及 Windows 等原生操作系统的基础镜像，基础镜像默认所有租户均可使用。

- 支持管理将租户自制或导入的镜像复制为基础镜像，作为默认基础镜像共享给平台所有租户使用。
- 支持修改镜像的使用权限，赋予不同租户对应的镜像版本。
- 支持管理员修改基础镜像的名称备注及删除基础镜像。

(2) 自制镜像

自制镜像由租户或管理员通过虚拟机自行导出或自定义导入的自有镜像，可用于创建虚拟机，除平台管理员外仅账号自身有权限查看和管理。

- 支持管理将为租户导入自定义镜像，租户或平台管理员将第三方业务虚拟机以镜像的方式迁移到平台镜像仓库，作为业务迁移的重要通道。
- 支持用户导入 Linux 和 Windows 发行版及自定义镜像，并支持 X86 架构和 aarch64 两种系统架构镜像的导入。
- 云平台的镜像格式默认为 RAW，用户上传 VHD、VMDK、QCOW2、OVA、ISO 等格式的镜像时，需先将镜像转换为 QCOW2 格式的镜像才可导入。
- 并支持管理员将租户的虚拟机导出为自制镜像，并支持管理员下载镜像仓库中的所有自制镜像。
- 支持通过自制镜像创建虚拟机、删除自制镜像、修改自制镜像名称。

为方便平台镜像模板文件的共享，平台支持管理员将一个自制镜像复制为一个基础镜像，使一个租户的自制镜像共享给所有租户使用，适用于运维部门制作模板镜像的场景，如自制镜像操作系统的漏洞修复或升级后，制作一个自制镜像并复制为基础镜像，使所有租户可使用新的镜像文件升级虚拟机系统。

6.6.3 外网网段管理

外网网段是平台对外通信的网络，一般由管理员或运维人员通过物理网络分配并配置至云平台。外网网段是平台为租户分配外网 IP 的 IP 资源池，支持 IPv4 和 IPv6 两种 IP 类型，并支持配置网段路由并自动下发路由至平台虚拟机。

平台在部署时默认为配置一段外网网段，如果平台业务需求，也可由管理员在管理控制台上自助添加 IP 网段，在添加 IP 网段前需要保证物理交换机上已为节点外网网络配置 Vlan 及相关网段信息。

网段管理仅作为平台管理员将物理网络上的网段信息录入至云平台，使云平

台的租户可申请外网网段是的 IP 地址作为虚拟资源的外网 IP，与平台外网进行通信。

支持管理员对外网 IP 的网段进行维护及管理，包括 IP 网段、网关、外网网卡、VLAN、路由及网段权限等配置，方便云平台管理员对外网 IP 地址池的管理，同时支持 IPv4 和 IPv6 双栈 IP 资源池管理。

- 支持通过私有 IP 地址段模拟外网网段，在交换机或上层路由将私有 IP 地址段 NAT 到互联网。
- 支持为每个网段配置路由策略，即租户申请网段的外网 IP 绑定至虚拟资源后，下发目的路由地址的流量自动以绑定的外网 IP 为网络出口。路由策略提供默认路由、指定路由及暂不指定三种模式：
 - 默认路由：即下发路由的目的地址为 0.0.0.0/0，代表默认所有流量均以绑定的外网 IP 为出口。
 - 指定路由：即管理员指定目的地址（如 10.0.2.0/24）的流量以绑定的外网 IP 为出口。
 - 暂不指定：即该网段不自动下发路由，仅可通过此外网 IP 地址与本网段进行通信。
- 支持管理员为云平台添加 IPv4 或 IPv6 版本的网段，使平台租户可同时申请 IPv4 和 IPv6 版本的外网 IP，并绑定至虚拟机提供网络服务。
- 支持管理对每个网段的开放范围进行控制及修改，默认为所有租户（所有租户可申请并使用网段 IP），支持配置为部分租户（指定的租户才可申请并使用网段 IP，未指定租户无法查看并申请网段 IP）。
- 支持管理员对每个网段添加标签，使 EIP 申请和创建虚拟机时可根据标签过滤出需要的线路。
- 支持管理员自定义外网网段 IP 地址的带宽规格，定义每个网段租户可申请的带宽范围。

为方便管理员和运维人员，平台提供外网网段的查看、修改等生命周期管理。

外网网段管理与平台部署的物理网络及架构拓扑紧密相关, 在维护外网网段前需确保物理网络配置完善后, 至平台录入 IP 网段后才可使用。

6.6.4 专线接入管理

专线接入用于搭建用户本地数据中心与私有云 VPC 之间高速、低时延、稳定安全的专属连接通道。通过专线接入, 可帮助用户实现混合云架构的网络数据面打通。

提供自助专线接入能力, 支持设置本端与远端 IPv4 地址、远端子网网段、网卡、Vlan 及带宽。

支持管理对每个专线的开放范围进行控制及修改, 默认为所有租户 (所有租户可使用专线接入与指定 VPC 互通), 支持配置为部分租户 (指定的租户才可使用专线接入与指定 VPC 互通)。

支持有专线接入权限的租户在 VPC 互通中指定专线接入资源与 VPC 建立内网互信通道。

支持用户修改远端子网网段, 适用于专接入远端网段变更的场景。

支持用户调整专线接入的带宽和专线接入的资源标签, 并支持管理员对专线接入资源进行删除。

6.7 虚拟资源管理

平台为管理员提供全平台所有租户的虚拟资源全生命周期运营和管理能力, 使平台管理员可通过控制台统一管控平台的整体虚拟资源。

虚拟资源管理包括租户端的所有产品服务, 如虚拟机、虚拟机模板、隔离组、弹性网卡、VPC 网络、外网 IP、高可用 VIP、组播、安全组、负载均衡、NAT 网关、VPN 网关、云硬盘、快照、商业存储磁盘、文件存储、对象存储、MySQL、Redis、备份、监报告警、资源事件等资源。

支持平台管理员为指定租户创建资源, 创建的资源归属租户所有, 同时支持管理员对平台所有租户的虚拟资源进行全生命周期管理, 如创建、删除、修改、

绑定、解绑等。

平台管理员为租户的虚拟资源绑定关联资源时，仅支持指定归属于租户的资源进行绑定，如仅支持管理员将归属于租户的外网 IP 绑定至相同租户的虚拟机。

6.8 QoS 配置管理

平台全局默认提供全局云硬盘 QoS 配置，即新创建的云盘会根据平台公式赋予 QoS 值，限制平台用户对磁盘性能强行占用。

支持管理员对平台所有租户的云硬盘自定义设置 QoS 值，仅当全局 QoS 配置开启时，管理员为每个云硬盘自定义的 QoS 才可生效。同时支持对虚拟机系统盘进行 QoS 配置。

(1) 读/写 IOPS

当磁盘的 Arch 架构为 HDD 时，可设置的读/写 IOPS 范围为 0~50000，默认值为 1000，配置为 0 不限速。

当磁盘的 Arch 架构为 SDD 时，可设置的读/写 IOPS 范围为 0~50000，默认值为计算公式根据当前硬盘容量计算的值，配置为 0 不限速。

(2) 读写带宽 (MBps)

当磁盘的 Arch 架构为 HDD 时，可设置的读/写带宽范围为 0~1000Mbps，默认为 100，配置为 0 则不限速。

当磁盘的 Arch 架构为 SSD 时，可设置的读/写带宽范围为 0~1000Mbps，默认为计算公式根据前当前硬盘容量计算的值，配置为 0 则不限速。

硬盘扩容容量后，会根据计算公式重新计算新容量的 QoS 值，根据计算的 QoS 值重新设置硬盘的 QoS。

- 若硬盘扩容前设置的 QoS 值 < 新容量 QoS 值，则以新容量 QoS 值为准。
- 若硬盘扩容前设置的 QoS 值 > 新容量 QoS 值，以扩容前设置的值为准。

硬件介质和容量会影响硬盘的读写 IOPS 和宽带速率，若配置的 QOS 超过

硬件本身性能，以硬件性能为准。系统会默认分配 QoS 值，如需取消一块硬盘的限速功能，可将 IOPS 和宽带均配置为 0。

QoS 配置同时针对 MySQL、对象存储、文件存储类的 PaaS 层产品提供 QoS 配置管理，用于限制 PaaS 产品使用磁盘的 IOPS 和带宽。

- 支持管理员对平台所有租户的文件存储自定义设置 QoS 值。
- 支持管理员对平台所有租户的对象存储自定义设置 QoS 值。
- 支持管理员对平台所有租户的 MySQL 自定义设置 QoS 值。

6.9 资源模板

资源模版支持租户预定义创建资源的参数配置，保存到模版中，便于后续快速创建，以及结合水平弹性伸缩完成业务节点的快速伸缩。

云平台用户可以通过指定机型、规格、镜像、云硬盘、VPC 网络、公网 IP、安全组及虚拟机相关基础信息一键创建虚拟机模板，用于从模板创建虚拟机实例。

- 虚拟机模板仅作为一资源创建的模板配置，不占用实际资源。
- 支持通过资源模板一键创建资源，创建时可进行配置变更。
- 支持用户更新资源模板的配置项和标签，并支持对资源模板进行克隆。
- 支持用户删除资源模板，删除后对通过资源模板创建的资源无影响。

6.10 标签管理

6.10.1 概述

标签用于标记各项云资源，从不同维度对具有相同特征的云资源进行分类、搜索和聚合，让资源管理变得更加方便。标签由一对键值对 (key:value) 构成，用户可根据需求自定义键值对内容，绑定不同资源。

- 支持标签批量创建，单次创建，删除标签功能。

- 支持查看资源，展示该条标签下所有绑定的资源。
- 支持绑定资源，可选择不同地域下不同资源类型进行绑定。
- 支持解绑资源，可批量解绑。

同时，标签支持资源创建时选择需要的标签进行添加，支持在资源界面对标签进行添加与删除操作。资源界面将会展示当前资源所绑定的标签键值对。

支持统一的搜索入口，可根据 **key/value**，资源 ID，资源类型，三个维度进行绑定资源到查询，灵活操作资源，可在云资源界面以及标签管理界面进行搜索，方便查询管理较大数量的标签，以及快速的匹配资源。

6.10.2 资源类型

租户侧支持的资源类型包括虚拟机、虚拟机模版、镜像、VPC、子网、云硬盘、弹性网卡、快照、弹性 IP、负载均衡、SSL 证书、安全组、IP 组、端口组、NAT 网关、Redis、MySQL、对象存储、文件存储、VPN 网关、对端网关、隧道、伸缩组、VIP、组播、隔离组。

管理侧支持的资源类型包括：外网网络、专线接入。

6.10.3 使用限制

- **标签命名限制**

标签键以及 value 值支持最大 127 位字符，不能为空，区分大小写-标签 key 以及 value 内容支持 utf-8 格式表示的大小写数字、汉字、数字、空格以及特殊字符

- **数量规范**

1 个资源最多可以绑定 50 个标签-1 个标签包含 1 个标签键和 1 个标签值 (tagKey:tagValue) -1 个资源上的同一个标签键只能对应 1 个标签值-单次批量创建标签数量最多不超过 5 个

- **资源状态限制**

虚拟机除过删除，删除中和失败的资源不能更新标签，其他状态下可修改资源绑定的标签内容。

6.11 监控告警

6.11.1 概述

监控告警是平台全线产品的运维监控及告警服务，提供全线资源实时监控数据及图表信息，可根据监控数据批量为资源设置告警策略，并在资源故障或监控指标超过告警阈值时，以邮件的方式给予通知及预警；同时监控告警服务实时为用户提供资源告警状态，让用户精准掌控业务和各云产品的健康状况，全方位保障业务的可靠性和安全性。

监控告警服务提供监控图表、告警模板、告警记录及通知组四大架构功能，整体架构功能均以监控数据为基准：

- 云平台通过智能化数据采集系统，租户对虚拟机、云硬盘、EIP、负载均衡、NAT 网关、弹性伸缩、VPN 网关，对象存储，文件存储等资源指定的监控指标数据进行完整挖掘；管理员可对节点、计算集群、存储集群指定的监控指标数据进行完整挖掘。
- 云平台将采集来的监控数据存储至数据库中，并根据指定规则对数据进行检索及统计，通过指定的时间维度及数据粒度以图形化的方式显示监控图表。
- 基于已有的监控数据，用户可通过配置告警模板，为指定的监控指标指定告警阈值、持续时间、重要程度、通知组及选择对比方式，可通过设置告警持续时间，判定区分不同等级的告警及通知。
- 可为告警模板配置通知组，指定在发生告警时通知事件的通知人及通知方式。
- 在告警期间，可通过告警记录查询实时告警信息，以判断故障的发生时间和重要程度。

6.11.2 监控图表

监控图表指平台将智能化采集的资源运行数据，根据指定的资源及指标等筛选规则进行检索并统计，通过指定的数据粒度及时间维度以图形化的方式显示监控图表。通过监控图表，用户可以直观的查看并了解平台上已运行虚拟资源的性能、容量及网络状态等状态，及时了解资源的健康状况及故障节点。

平台为租户构建的虚拟机、弹性 EIP、负载均衡、NAT 网关、弹性伸缩、VPN 网关、对象存储、文件存储分别提供多种监控指标的实时和历史监控图表，并可根据监控指标项配置相关告警模板，用于阈值超标时给予告警及通知。

- 虚拟机监控图表：通过虚拟机详情页面的监控信息栏可查看单台虚拟机的监控信息，包括网卡出/入带宽、网卡出/入包量、磁盘读/写吞吐、磁盘读/写次数、平均负载、空间使用率、内存使用率、CPU 使用率、GPU 使用率、GPU 总显存、GPU 显存使用量、GPU 总消耗、GPU 平均功耗、GPU 温度；
- 弹性 EIP 监控图表：通过 EIP 详情页面的监控信息可查看单个 EIP 资源的监控信息，包括网卡出带宽使用率、入带宽、出带宽、入包量、出包量；
- 负载均衡监控图表：通过负载均衡详情页面的监控信息可分别查看负载均衡实例和 VServer 监听器的监控信息，监控图表包括 LB 每秒连接数、LB 每秒网卡出/入流量、LB 每秒网卡出包数量、VServer 连接数、HTTP 2XX、HTTP 3XX、HTTP 4XX、HTTP 5XX；
- NAT 网关监控图表：通过 NAT 网关详情页面的监控信息可查看单个 NAT 网关的监控信息，包括网卡入带宽、网卡出带宽、连接数、网卡入包量、网卡出包量。
- VPN 网关监控图表：通过 VPN 网关服务详情页面的监控信息可查看单个网关的监控信息，包括网关出/入带宽、网关出带宽使用率、网关出/入包量。

- VPN 隧道监控图表：通过 VPN 隧道服务详情页面的监控信息可查看单个隧道的监控信息，包括隧道出/入带宽、隧道出/入包量及隧道健康状况。
- 对象存储监控图表：通过对象存储服务详情页面的监控信息可查看，包括 CPU 使用率、内存使用率、对象数量、存储容量、当前存储量、存储容量使用率、存储总写吞吐、存储总读吞吐。
- 文件存储监控图表：通过文件存储详情页面的监控信息可查看，包括 CPU 使用率、内存使用率、存储容量、当前存储量、存储容量使用率、存储总写吞吐、存储总读吞吐。

同时整个平台为全局节点、计算集群、存储集群提供多种监控指标的实时监控图表，并可按照用户的需求对告警模版进行配置，用于阈值超标的时给予告警和通知。

- 节点监控图表：通过节点详情页面的监控信息可查看，包括 CPU 使用率、内存使用率、GPU 使用率。
- 计算集群监控图表：通过计算集群详情页面的监控信息可查看，包括 CPU 分配率、内存分配率、GPU 分配率。
- 存储集群监控图表：通过存储集群详情页面的监控信息可查看，包括存储分配率。

监控图表可根据时间维度展示实时监控数据，同时支持查看 1 小时及自定义时间的监控数据及图表信息。

6.11.3 告警模板

告警模板是平台监控告警服务为用户提供的一种批量设置资源告警的功能，通过预先定义模板中的告警规则及通知规则，将模板中定义的规则应用到虚拟资源；若虚拟资源的监控指标数据达到或超过告警规则中设定的阈值及条件，则根据通知规则中定义的通知方式发送告警通知到指定的联系人。

根据不同的资源类型，可定制不同监控指标及阈值的告警规则，并可选择将

监控指标应用至关联资源的单个网卡或磁盘设备，满足多种应用场景下的监控报警需求。

- 告警模板是由多条告警规则及关联资源构成的；
- 一个告警模板仅支持绑定一种类型资源，包括虚拟机、弹性 EIP、负载均衡、NAT 网关、弹性伸缩、VPN 网关、对象存储、文件存储、节点、计算集群及存储集群等。
- 每个告警模板可包含多条告警规则，每条告警规则包含监控指标、对比方式、告警阈值、持续时间、重要程度及通知组；
- 每个告警模板仅支持绑定一个通知组，每个通知组可包含多个通知人，支持邮件的通知方式。

云平台支持用户和平台管理员为资源创建告警模板，并提供告警模板的全生命周期管理，用户可通过告警模板自定义所对应资源类型的告警规则，通过告警规则对资源进行监控指标的告警触发和处理。

告警规则是告警模板的核心，每个告警模板均由 1 条或多条告警规则组成。被绑定至告警模板的资源监控指标数据会根据告警规则中定义的阈值触发相关告警策略，并通过告警规则中的通知方式进行告警信息的通知，以便快速入处理告警或故障。

- 监控指标：仅可选择告警模板资源类型所包含的监控指标，一条告警规则仅支持一个监控指标；
- 对比方式：指监控指标的实际数据与告警阈值的比较方式，代表当前告警规则的告警逻辑，包括 \geq 、 \leq 、 $>$ 、 $<$ ：
 - 当选择 \geq 时，即代表监控数据大于或等于阈值时触发一次告警周期；
 - 当选择 \leq 时，即代表监控数据小于或等于阈值时触发一次告警周期；
 - 当选择 $>$ 时，即代表监控数据大于阈值时触发一次告警周期；

- 当选择<时，即代表监控数据小于阈值时触发一次告警周期；
- 告警阈值：指监控指标数据的临界值，与监控指标数据进行对比，符合对比方式即触发一次告警周期，如 CPU 使用率的告警阈值为 80，对比方式为大于等于，即 CPU 使用率大于等于 80%即触发一次告警周期；
- 持续时间：监控指标数据触发阈值持续的时间，持续时间内均达到告警阈值才会触发告警；
- 重要程度：用户可根据业务需要在创建告警规则时选择合适的等级，分为一般、重要、危险三种，在告警记录中可根据重要程度进行记录的筛选；
- 通知组：即触发告警周期且需要发送通知时，发送告警通知的方式及联系人。

6.11.4 告警记录

告警记录告警模板定义的所有告警记录及信息，通过告警记录可查阅实时及历史告警信息，包括告警的指标说明、模版类型、标签、当前值、状态、重要程度及告警时间。

- 指标说明：触发当前告警记录的资源监控指标项，即数据来源；
- 模版类型：触发当前告警记录的资源类型及资源；
- 标签：显示磁盘空间使用率指标的数据盘信息；
- 当前值：即触发告警或恢复告警时当前告警记录监控指标的数据值；
- 状态：告警记录的当前状态，分为触发中、待触发、未触发，可根据需求进行状态的筛选；
- 重要程度：根据监控规则显示当前告警记录的重要程度，包含危险、重要、一般，可根据需求进行告警记录的筛选；
- 告警时间：触发告警规则的具体时间。

6.12 通知组

通知组是监控报警发送告警通知的方式及联系人信息，通过对用户邮箱、Webhook 地址的记录，将不同资源告警通过邮件或 Webhook 方式通知给通知人，以便划分全责，精细化处理告警通知。

在使用监控告警模板时，需要先创建一个通知组，添加相关联系人信息，并设置通知组的通知方式，以便关联告警模板。

通知组是一组通知人的组合，可以包含一个或多个联系人，在资源发生告警时会通过所设置的通知方式至所有通知人。

通知人是指告警规则发送通知的具体联系人，同一个联系人，可以加入多个通知组，支持邮件通知和 Webhook 通知两种方式。

- 邮件通知：支持配置联系人姓名及邮箱地址信息，用于发送告警通知至配置的联系人邮箱。
- Webhook 通知：支持配置 Webhook 地址及发送警告信息的请求方式，请求方法支持 GET 和 POST 两种方式进行信息传输。

支持用户对通知组及组内的通知人进行自定义配置，并支持用户对已有的通知方式进行修订及删除。

6.13 操作日志

6.13.1 操作日志

操作日志是指用户在控制台或 API 对资源进行的操作行为及登录登出平台的审计信息。操作日志会记录用户在平台中的所有资源操作，提供操作记录查询及筛选，通过操作日志可实现安全分析、资源变更追踪以及合规性审计。

通过操作日志用户可查看整个平台单个地域或全部地域所有的资源操作及平台审计日志，同时可通过 API 查询租户内所有资源的操作日志及审计信息。

操作日志支持的资源模包括面向租户的裸金属、虚拟机、USB、资源模板、

镜像、VPC、云硬盘、外置存储、弹性网卡、外网 IP、VIP、组播、安全组、负载均衡、NAT 网关、VPN 网关、对象存储、文件存储、MySQL、Redis、自动伸缩、监控告警模板、定时器、账户、API 控制台、备份服务等。同时面向平台管理员可提供运维方面的操作日志。

平台支持获取日志的操作名称、所属模块、地域、关联资源、操作者、操作结果、备注及操作时间等信息；并支持导出用户的操作审计日志为本地 Excel 表格，方便账户管理和运营。

- 操作名称：指操作日志的操作名称，包括调用 API 的接口名称及操作的界面展示名称，如调整带宽。
- 所属模块：指操作日志操作的资源类型，如虚拟机类型。
- 地域：操作资源所属的地域。
- 关联资源：操作日志对应的资源标识符，并可查看一个操作中所有关联的资源标识，如绑定弹性 IP 对应的虚拟机 ID 和外网 IP 的 ID。
- 操作者：操作日志对应的操作者，可追溯到具体的主账号和子账号。
- 操作结果：操作日志的结果，如操作成功、操作失败、参数异常、存储集群物理资源不足等。
- 备注：操作日志的备注信息。
- 操作时间：操作日志的操作时间。

为方便用户便捷的查看操作审计日志，支持日志的筛选和搜索检索，包括所属模块、操作状态及查询时间范围等纬度。

所属模块支持所有产品模块的筛选，同时支持查看全部资源的日志及审计信息，即不对所属模块进行筛选。操作状态支持状态为成功、失败的日志筛选，同时支持查看全部状态的日志和审计信息。查询时间范围支持 1 小时及自定义时间的日志筛选，最长可查询半年的操作日志。

6.13.2 通知规则

操作日志通知规则对资源操作日志进行监控并通过通知组进行事件信息的告知。支持用户为操作日志配置通知规则，当资源操作日志符合通知规则要求时，即发送监控邮件到通知组内的成员。

支持用户配置操作日志通知规则的监控地域、通知组、监控模块及监控级别等信息，并支持用户和平台管理员对操作日志的通知规则进行修改和删除。

- 监控地域：通知规则的地域信息。
- 通知组：邮件通知的通知组信息，仅支持选择一个通知组。
- 监控模块：监控的资源模块内容，如虚拟机、云硬盘等，支持批量配置多个模块。
- 监控级别：操作日志的操作结果，包括操作成功、操作失败。

6.14 资源事件

6.14.1 资源事件

资源事件是用于对云平台核心资源的部分操作及状态进行记录及通知，如资源生命周期状态的变化、操作运维执行情况等。

资源事件记录用户在资源类型的部分核心操作事件，提供事件详细记录查询及筛选，并可配合通知规则及时通知用户、定位问题。

资源事件支持的资源类型包括虚拟机、NAT 网关、VPN 网关、隧道、负载均衡、对象存储、文件存储、MySQL 及 Redis。支持获取资源事件的资源 ID、资源类型、事件类型、事件等级、事件内容、事件发生次数、开始时间及更新时间。

- 资源 ID：指资源事件监控的资源 ID。
- 资源类型：当前资源事件记录所指定的资源类型。

- 事件类型：分为生命周期变化事件和操作运维事件，如虚拟机调度，虚拟机开关机，挂载磁盘等。
- 事件等级：事件等级的类型，包括正常、警告、错误。
- 事件内容：详细记录触发事件的具体信息。
- 发生次数：记录该事件累计触发次数。
- 开始时间：第一次资源事件发现的时间。
- 更新时间：第二次及以后触发资源事件的时间。

资源事件对整个平台及所有用户资源事件进行记录，为方便用户便捷的查看资源事件日志，支持事件日志的筛选和搜索检索，包括所属地域、资源类型、资源及事件周期等维度；同时平台资源事件日志支持用户导出资源事件为本地 Excel 表格，方便用户查看和定位。

6.14.2 通知规则

资源事件通知规则对事件日志进行监控并通过通知组进行事件信息的告知。支持用户为资源事件配置通知规则，当事件日志符合通知规则要求时，即发送监控邮件到通知组内的成员。

支持用户配置资源事件通知规则的监控地域、通知组、监控模块及监控级别等信息，并支持用户和平台管理员对资源事件的通知规则进行修改和删除。

- 监控地域：通知规则的地域信息。
- 通知组：邮件通知的通知组信息，仅支持选择一个通知组。
- 监控模块：监控的资源模块内容，如虚拟机、负载均衡等。
- 监控级别：对实例正常运行的影响程度进行划分，含正常、警告、错误。

6.15 回收站

6.15.1 概述

平台提供资源回收站，是平台或租户资源删除或欠费自动释放的暂时保留区，用户删除的资源包括虚拟机、磁盘、EIP、自制镜像等资源，会在删除后自动进入回收站中。

云平台资源被用户手动删除及费用过期时，会自动进入回收站暂时留存。进入回收站中的资源平台默认保留时间为 360000 秒，可通过云平台管理员进行自定义保留时间的设置。

保留期间用户可在回收站中查看资源的信息，如资源 ID、资源名称、资源类型、过期时间、删除时间、是否自动销毁及预定销毁时间等。

- 资源状态：当前资源的状态，包括已删除、销毁中。
- 资源类型：当前留存资源的资源类型，包括虚拟机、硬盘、外网 IP、自制镜像等；
- 过期时间：指当前资源的费用过期时间，仅当资源类型为需计费的资源时有效，如虚拟机、磁盘、外网 IP；
- 删除时间：指当前留存资源被手动删除或费用过期进入回收站的时间；
- 是否自动销毁：指当前留存资源是否会在留存期间自动销毁，可通过云平台管理控制台设置保留期后是否自动销毁资源：
 - 若云平台全局配置为回收站资源自动销毁，则到达保留期后，将自动销毁资源；
 - 若云平台全局配置为回收站资源不自动销毁，则资源将永久留存在回收站，可通过手动恢复或销毁资源；
- 销毁时间：指当前留存资源将被自动销毁的时间，仅当云平台全局配置为回收站资源自动销毁时有效。

在回收站中的资源支持恢复、续费及销毁操作，保留时间到期后，资源会被彻底销毁，不可恢复。

为方便租户对回收站资源的维护，支持对进入回收站的资源进行批量操作，包括批量恢复资源、批量销毁资源及批量续费资源。

6.15.2 恢复资源

恢复资源是指手动恢复被误删而进入回收站的资源。

若资源被用户手动删除且无欠费的情况下，可直接通过恢复资源操作进行恢复；

若资源因账户欠费而自动进入回收站，则恢复资源时，需联系云平台管理员对账号进行充值后，通过“续费”操作对资源进行续费后，在进行资源恢复；

若全局未开启资源自动续费且账户余额充足，资源过期后会自动进入回收站，恢复资源时，需要先通过“续费”操作对资源进行续费后，在进行资源恢复。

6.15.3 续费资源

续费资源是指对资源的费用周期进行续费，仅支持需计费的资源进行“续费”操作。因欠费或费用到期自动进入回收站的资源被成功续费后，才可进行恢复操作。资源续费的周期根据计费方式会有所区别：

- 资源按小时计费时：一次续费操作可续费 1 个小时，N 次续费操作即续费 N 个小时。
- 资源按月计费时：一次续费操作可续费 1 个月，N 次续费操作即续费 N 个月。
- 资源按年计费时：一次续费操作可续费 1 年，N 次续费操作即续费 N 年的费用周期。

支持对回收站中的资源进行批量续费，批量续费时资源会按照资源的计费方式自动续费一个周期，如按时计费的资源，则自动续费一个小时；按月购买的资源，则自动续费一个月。

6.15.4 销毁资源

销毁资源是指用户手动销毁留存在回收站的资源，资源被销毁后无法恢复。支持用户批量对回收站中的资源进行批量销毁，以方便对账户的清理和维护。

6.16 计量计费

6.16.1 概述

平台为用户资源分配和使用提供计量计费服务，需计费的资源均支持按时、按年、按月三种计费方式，支持资源的计费、扣费、续费及过期回收等订单管理操作，同时基于基于账户提供充值、扣费等交易管理。

子账号共享主账号的账户余额，通过子账号创建的资源可直接通过共享余额进行扣费，并可通过主账号或子账号查看账户的交易流水及订单明细。

平台资源计费均为预付费模式，即无论按时、按年、按月付费，在资源创建时都需保证账户余额可满足一个计费周期的扣费，下一个计费周期开始前即进行扣费。

- 按时计费：一小时为一个计费周期，资源按照每小时的单价进行预扣费；
- 按月计费：一个月（非自然月）为一个计费周期，资源按照每个月的单价进行预扣费；
- 按年计费：一年（顺延年）为一个计费周期，资源按照每年的单进行预扣费；按年按月按时购买的资源支持随时升降级配置并在升级配置后自动补齐差价。

账户余额不足下一个计费周期时，资源即会自动进入回收站，需要对资源账号及资源进行续费操作后，才可恢复使用；对于 文件存储、对象存储、外网网卡、NAT 网关、VPN 网关、负载均衡资源，账户余额不足下一个计费周期时，资源会自动进行删除。

云平台管理员在全局开启“资源自动续费”且账户余额充足时，则资源在下

一个计费周期会进行自动续费操作；若云平台管理员在全局关闭“资源自动续费”且账户余额充足时，则资源在下一个计费周期会自动进入回收站，需在回收站对资源进行续费操作，并恢复资源。

资源在创建时，所有计费资源的计费计价均会通过资源计价器按照计费方式进行展示，用于确认订单的费用。每个计费周期内的资源均支持释放和删除，当账户余额不足时，可通过云平台管理员进行充值。

在财务运营管理方面，平台提供完整的财务管理能力，包括订单管理、交易管理、充值管理、价格配置四大模块。

6.16.2 资源计价格

资源计价器为用户提供资源付费方式的选择，并展示付费模式下所有资源的信息及资源的“购买”确认按钮。

计价器中付费方式支持用户选择时、月、年，分别代表按时计费、按月计费、按年计费，其中选择月和年时，可以选择购买的月份数量和年份数量。

- 月份可选择 1~11，分别代表 1 个月或 11 个月；
- 年份可选择 1~5，分别代表 1 年或 5 年；

合计费用指当前订单中所有计费资源一个计费周期的费用合计，如一个虚拟机订单中，包括指定的 CPU 内存、云盘(若有)、EIP(若有)等资源按照付费方式的费用合计。

用户创建资源后，平台即从账号余额扣除合计费用金额，并产生一个新购订单及一笔扣费的交易流水。

6.16.3 资金管理

资金管理为平台运营管理提供租户账号充值及提现管理能力，使租户可使用租户的账户余额进行资源使用，同时支持将当前账号余额进行提现扣除。

6.16.3.1 充值管理

云平台充值管理为管理运营者提供租户资金充值通道, 支持平台内部充值和外部渠道充值两种, 管理员可按照使用需求为租户充值金额。

外部渠道充值来源分为银行转账、支付宝支付、微信支付、新浪支付四种。平台内部充值是平台为租户赠送的余额。单次充值最小金额为 100, 最大为 500000 元。

管理员可通过充值管理查看租户及整个云平台的充值记录信息, 支持获取充值单号、充值租户 ID、主账号名称、主账号邮箱、充值渠道、充值金额及充值时间等信息。

- 充值单号: 充值记录在云平台的唯一标识。
- 充值租户 ID: 充值的租户 ID。
- 主账号名称: 充值的租户下的主账号名称。
- 主账号邮箱: 充值的主账号邮箱。
- 充值渠道: 充值的渠道。
- 充值金额: 充值的金额数。
- 创建时间: 充值产生的时间。

支持管理员下载平台所有充值记录信息为本地 Excel 文件, 方便平台运营管理和报表统计。

6.16.3.2 提现管理

云平台充值管理为管理运营者提供租户资金提现通道, 支持平台账户和外部账户两种, 管理员可按照使用需求, 为租户进行余额的提现扣除, 提现金额不可超过账户当前余额。

管理员可通过提现管理查看租户及整个云平台的提现记录信息, 包括提现单

号、租户 ID、主账号名称、主账号邮箱、源账号类型、提现金额、创建时间等。

- 提现单号：提现记录在云平台的唯一标识。
- 提现租户 ID：提现的租户 ID。
- 主账号名称：充值租户下的主账号名称。
- 主账号邮箱：充值的主账号邮箱。
- 源账号类型：充值的金额类型，分为平台账户和外部账户。
- 提现金额：提现的金额数。
- 创建时间：提现产生的时间。

支持管理员下载平台所有提现记录信息为本地 Excel 文件，方便平台运营管理和报表统计。

6.16.4 价格配置

6.16.4.1 基准价格

价格配置即平台全局的产品定价，平台支持对 5 个维度的资源项进行定价，包括 CPU、内存、磁盘、外网 IP、GPU。单个云服务的实际出售价格，根据云服务所涉及到的计费资源项进行累加。

租户价格默认继承价格配置中的价格，可由管理员自定义租户价格的折扣，以适应平台运营的需求。

资源项	计费类型	计费规则
CPU	小时、月、年	每个集群每核 CPU 价格
内存	小时、月、年	每个集群每 GB 价格
硬盘	小时、月、年	每个集群每 GB 价格
外网 IP	小时、月、年	每个网段每 MB 价格，可以定义带宽不同梯度的价格
GPU	小时、月、年	每个集群每颗 GPU 价格

平台在初始化时，会对所有计费项进行初始定价，如果需要修改可在管理平

台/运营与管理/价格配置中进行调整。如果平台不需要计费，可以将所有计费项的价格设置为 0。

支持管理员通过价格配置控制台查看当前平台上每个地域下所有产品（计费资源）的价格信息，包括计费因子、属性、计费类型、计费规则、价格、单位、创建时间更新时间。

管理员可在全局价格配置处对每一个计费资源项的全局基准价格进行配置，价格修改后资源项的价格在全局进行变更；管理员对租户自定义的折扣不变，但最终折扣价会随着租户的折扣率进行变更。

- 全局资源项的小时基准定价被更新后，按小时付费的服务将在下一个计费周期按照新的基准价进行扣费。
- 全局资源项的月、年基准定价被更新后，对于已支付的按月和按年服务无影响，在下一个计费周期将按照新的基准价进行扣费。

管理员可在全局价格配置列表上对计费资源项进行价格更新，支持设置单个资源项针对不同集群的基准价。以虚拟机 CPU 为例，可设置单核 vCPU 的每小时价格为 0.2431 元，表示单核 vCPU 的小时单价为 0.2431 元。

同时平台针对外网 IP 的网段带宽，支持按梯度区间定价。如设置 0M~5M 的全局基准价格为 1 元，5M~99999999M 的全局基准价格为 10 元，提升平台计费的维护性。

6.16.4.2 租户价格配置

管理员可针对租户的每个产品修改在地域及不同集群的价格，计费因子包括 CPU、内存、磁盘、外网 IP、GPU。租户创建虚拟资源时，通过计费因子的费用合计按照付费方式进行扣费。

- 针对 CPU、内存、GPU 等计算计费因子在不同的计算集群可定义并展示不同的价格及折扣。
- 针对磁盘在不同的存储集群可定义并展示不同的价格和折扣。

- 针对外网 IP 可展示并定义全局的价格和折扣。
- 针对不同的外网网段可展示并定义不同的价格和折扣。

租户在平台的资源价格默认继承平台全局的价格配置, 管理员可通过租户价格列表查看每种资源在不同集群的基准价格, 同时支持自定义每个租户的产品价格折扣, 设置单个资源针对租户在不同集群的价格,

其中折扣为百分数, 如 90 为基准价格的 9 折, 即将租户 CPU 按月付费的价格打 9 折。修改后即会生效, 租户已创建的资源不受影响, 新创建的 CPU 资源将按照新的价格进行扣费。

修改租户的产品折扣只对本租户生效, 不影响其它租户的价格, 满足针对不同客户类型提供不同云资源价格的场景, 提升运营效率。

6.16.5 订单管理

订单管理是平台为用户提供的订单查询及统计服务, 通过订单管理可以查看平台所有租户账号及子账号订单记录, 支持查看某个地域、1 天、3 天、7 天、14 天、30 天及自定义时间的历史订单记录。

平台管理员可查看平台所有租户的订单记录数据, 租户主账号与所有子账号的订单管理及数据相同, 可通过一个账号查看租户内所有订单记录。

用户对资源进行创建、续费、变更配置或删除时, 会分别产生新购、续费、升级、降级及退单等类型订单。订单的信息包括订单号、订单类型、资源 ID、地域、订单金额、平台账户、外部账户、创建时间等。

- 订单号: 指当前订单的全局唯一标识符;
- 订单类型: 当前订单的类型, 包括新购、续费、升级、降级及退单五种类型。
 - 新购是指用户新创建的计费资源, 包括虚拟机、云硬盘、弹性 IP、外网网卡、文件存储、对象存储、NAT 网关、VPN 网关、负载均衡、文件存储、对象存储、MySQL、Redis 等;

- 续费是指预付费资源每一个计费周期续费时产生的订单，包括手动续费和系统自动续费；
 - 升级是指按时按月按年计费的资源变更配置时产生的续费订单，如升级带宽、升级虚拟机配置等；
 - 降级是指按时按月按年计费的资源变更配置时产生的续费订单，如降级带宽、降级虚拟机配置等；
 - 退单是指按时按月按年计费的资源被删除时产生的退费订单。
- 资源 ID：产生当前订单的资源标识符；
 - 地域：当前订单资源所在的区域；
 - 订单金额：当前订单金额，即订单在新购、续费、升级所付的费用及退单、降级所退的费用（退费展示为负值）；
 - 平台账户：当前订单平台账户支付的金额；
 - 外部账户：当前订单外部账户支付的金额；
 - 创建时间：当前订单记录的生成时间，一个计费周期产生一个订单记录。

支持平台租户、子账号及管理员下载订单管理记录信息为本地 Excel 文件，方便平台运营管理和报表统计。

6.16.6 交易管理

交易管理是平台为用户提供的账号金额相关的收支明细，包括扣费、充值、退费及统计服务，支持查看某个地域、1 天、3 天、7 天、14 天、30 天及自定义时间的历史交易记录

平台管理员可查看平台账号及子账号所有交易流水记录，租户主账号和所有子账号的交易管理及数据相同，可通地一个账号查看租户内所有交易记录信息。包括交易单号、交易类型、支出、收入、外部充值金额、平台充值金额及交易时间等。

- 交易单号: 当前交易记录在全局唯一的 ID 标识符, 以 **trade** 作为开头;
- 交易类型: 当前交易记录的类型, 根据平台对资源的不同操作, 分别包括充值、扣费和退费:
 - 充值指平台管理员通过后台为租户进行的充值操作;
 - 扣费指系统针对每个资源生命周期的计费操作, 如创建资源时, 进行扣费操作;
 - 退费指系统针对每个资源生命周期的计费操作, 如删除资源时, 进行退费操作;
- 支出: 当前交易记录所扣费的金额, 仅当交易类型为扣费时有效, 充值显示为 0.00 ;
- 收入: 当前交易记录进账的金融, 当交易类型为充值和退费时有效, 扣费显示为 0.00 ;
- 外部充值余额: 当前账户在当前交易记录发生后的外部充值余额;
- 平台充值余额: 当前账户在当前交易记录发生后的内部充值余额;
- 交易时间: 当前交易记录发生时间。

管理员可查看所属租户及平台全局所有交易信息, 同时可通过自定义查询时间查看一定时间周期内产生的交易记录, 并支持下载交易管理信息为本地 Excel 文件, 方便平台运营管理。

6.16.7 账单管理

账单管理包括账单总览、资源账单、账单明细。其中账单总览可以查看费用趋势以及本月账单汇总, 资源账单与账单明细支持筛选导出功能。

6.16.7.1 账单总览

账单总览支持查看单个租户及整个平台的费用趋势及本月账单汇总信息。

(1) 费用趋势

管理员通过费用趋势可自定义费用类型查看云平台或单个租户在近六个月内产生的总交易费用信息。

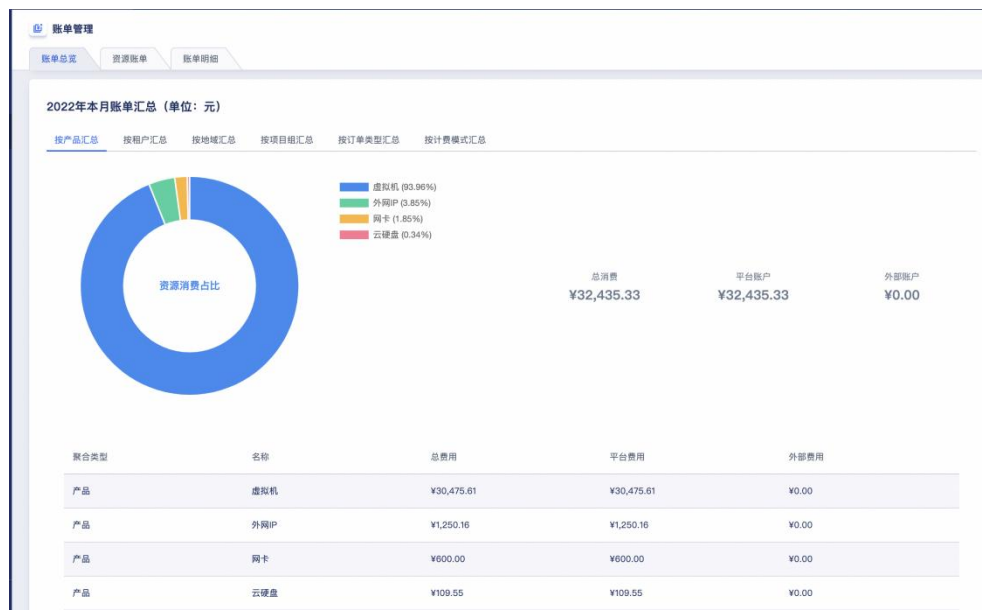
支持按租户进行批量查询费用趋势，同时可按照总费用、平台账户及外部账户分别查看费用趋势走向信息。

租户也可通过费用趋势查询租户内半六个月内产生的总交易费用信息。

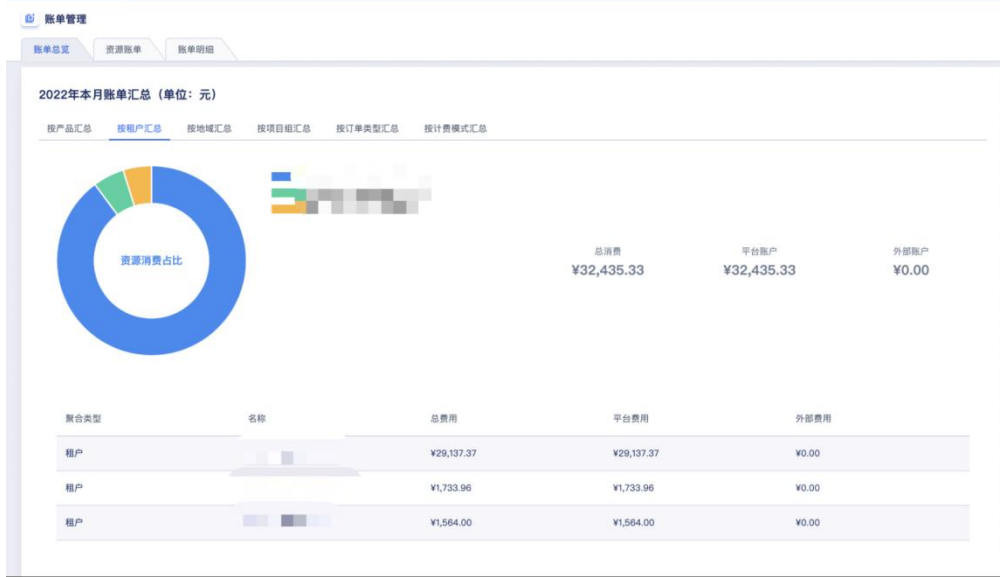
(2) 本月账单汇总

本月账单汇总从按产品汇总、按租户汇总、按地域汇总、按项目组汇总、按订单类型汇总及按计费模式汇总六个方面用饼图展示，列表包括聚合类型、名称、总费用、平台费用及外部费用。

● 按产品汇总



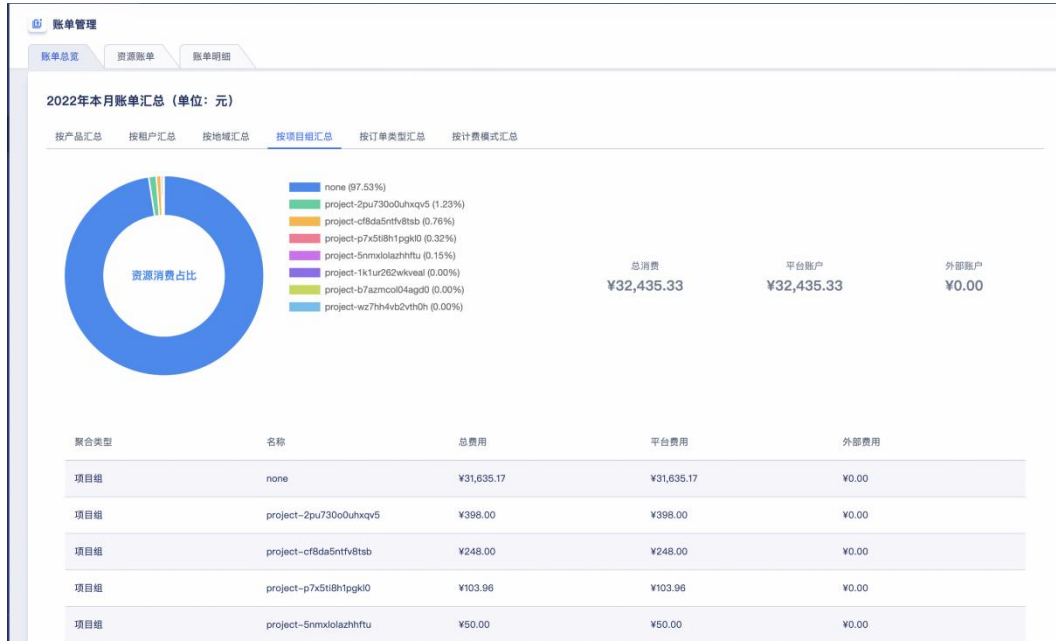
● 按租户汇总



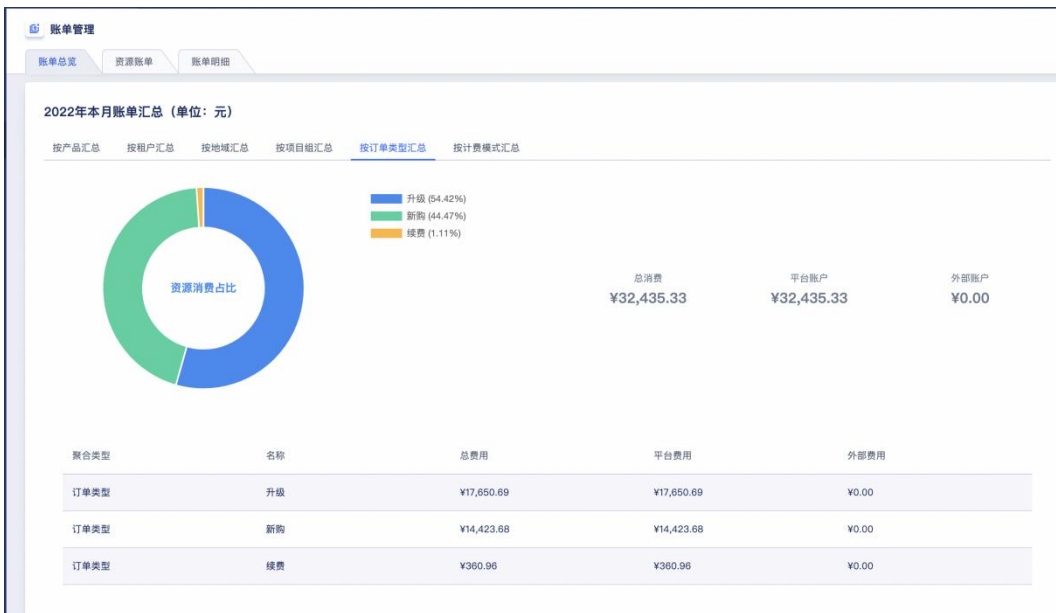
● 按地域汇总



● 按项目组汇总



● 按订单类型



● 按计费模式汇总



6.16.7.2 资源账单

管理员和租户均可从账单周期/所属产品/计费模式/所属租户/所属地域/所属项目六个维度查看云平台及租户内的资源账单信息。

资源账单信息包括资源 ID、地域、租户 ID、主账号名称、主账号邮箱、所属产品、所属项目、计费模式、总费用、平台账户、外部账户及交易时间。

- 资源 ID: 账单的全局唯一标识符
- 地域: 资源所在的地域信息
- 租户 ID: 产生订单的租户信息
- 主账号名称: 充值的主账号下的主账号名称
- 主账号邮箱: 充值的主账号邮箱
- 所属产品: 云平台的产品, 包括虚拟机、云硬盘、外网 IP、VPN 网关、负载均衡、NAT 网关、网卡、Redis、MySQL、文件存储、对象存储等。
- 所属项目: 本次交易资源所绑定的项目
- 计费模式: 按小时、月、年的计费模式
- 总费用: 本次交易的总费用
- 平台账户: 本次交易消费平台账户的金额

- 外部账户：本次交易消费外部账户的金额
- 交易时间：本次交易产生的时间

平台支持管理员和租户从账单周期、所属产品、计费模式、所属租户、所属地域、所属项目六个维度筛选资源账单，并导出到本地 Excel 文件，方便平台运营管理和报表统计。

6.16.7.3 账单明细

管理员和租户均可从账单周期/所属产品/订单类型/计费模式/所属租户/所属地域/所属项目七个维度查看云平台的账单明细。

账单明细信息包括资源 ID、交易单号、交易类型、订单号、订单类型、地域、租户 ID、主账号名称、主账号邮箱、所属产品、所属项目、计费模式、总费用、平台账户、外部账户及交易时间。

- 资源 ID：账单的全局唯一标识符
- 交易单号：交易记录在云平台的唯一标识
- 交易类型：账户充值和扣费均会生成一次交易记录，因此交易类型包括账户余额充值、免费账户充值及扣费
- 订单号：订单在云平台的唯一标识符
- 订单类型：包括升级和新购两种
- 地域：资源所在的地域信息
- 租户 ID：产生订单的租户信息
- 主账号名称：充值的主账号下的主账号名称
- 主账号邮箱：充值的主账号邮箱
- 所属产品：云平台的产品
- 所属项目：本次交易资源所绑定的项目

- 计费模式：按小时、月、年的计费模式
- 总费用：本次交易的总费用
- 平台账户：本次交易消费平台账户的金额
- 外部账户：本次交易消费外部账户的金额
- 交易时间：本次交易产生的时间

平台支持管理员和租户从账单周期、所属产品、订单类型、计费模式、所属租户、所属地域、所属项目七个维度筛选账单明细，并导出到本地 Excel 文件，方便平台运营管理和报表统计。

6.17 审批流程

6.17.1 概述

随着信息化数字转型在政企、教育、金融、制造等行业的实践和应用，企业对资源管理的标准化、流程化管理需求日益旺盛，对于云化资源同样需要设置标准的审批流程，满足平台资源的申请、审批业务的使用流程需求。

针对企业云化资源的管理，云平台为企业管理者提供的自助模式的资源审批服务，用于制定信息系统云化资源的标准使用流程，在租户或子账号需要使用或管理资源时，按照流程中定义的审批人和审批层级完成审批后，由平台自动化交付用户所需业务资源。

平台审批流程由平台管理员进行定义和发布，并由平台管理者设置是否为一个租户设置开通审批流程，支持手动审批和自动审批。

- **手动审批**

租户下主账号和子账号进行虚拟资源操作时需要走申请、审批流程，待审批通过后，平台会自动为用户创建或操作所需资源，并生成一条审批记录用于追溯。

- **自动审批**

租户下主账和子账号进行虚拟资源操作无需人工介入，系统将自动审批通过，并自动生成一条审批记录用于保留相关申请记录和审批记录。

开通资源审批的前提是设置审批流程，用于定义租户申请资源时，需要多少层级的审批，每一层级由谁进行审批，所有层级均通过后才可进行资源的创建和变更操作。为满足企业多种场景的审批业务，平台内置默认审批流程。

默认审批流程提供简单的审批逻辑，仅支持 1 级审批，当平台管理者为租户开启资源审批流程后，租户及子账号下资源的创建及变更申请统一由【平台管理员】进行审批，即平台管理员审批通过后，平台将自动执行资源的变更操作。

审批流程支持多种资源的变更操作，如虚拟机、云硬盘、VPC、外网 IP 及负载均衡、弹性网卡等资源，同时支持的变更资源生命周期管理操作，如下：

- 虚拟机：创建虚拟机、修改配置、扩容系统盘、扩容数据盘；
- 云硬盘：创建云硬盘、扩容磁盘；
- VPC 网络：创建 VPC；
- 外网 IP：创建外网 IP、调整带宽；
- 负载均衡：创建负载均衡。
- 弹性网卡：调整 IP 带宽

6.17.2 使用流程

(1) 管理员为租户开启审批流程

管理员通过租户列表为租户开启审批流程，同时可开自动审批。开启后租户在创建虚拟机时即需要发起审批流程，待管理员进行审批后，平台会自动发起虚拟机的创建操作。

(2) 租户申请变更资源

由租户在租户控制台发起资源变更操作（本文以申请虚拟机为示例）：

- 进入虚拟机控制台，通过【创建虚拟机】进入虚拟机创建引导页面；

- 填写申请名称和申请备注，按照虚拟机的创建要求输入其它必填信息，点击【立即购买】提交申请。
- 提交申请后，页面会自动跳转至【申请管理】页面，并在申请列表中自动新增一条待审批的申请记录。

(3) 管理员审批

由平台管理员 **admin** 账号通过【审批管理】中的待办对租户的申请进行通过或拒绝的审批，完成审批流程。

若平台管理员通过申请，则租户的申请状态变更为【处理中】，并会自动执行虚拟机的创建操作，可通过虚拟机列表查看正在创建的虚拟机资源，待资源创建成功后，申请状态变更为【成功】。

若平台管理员拒绝申请，则租户的申请状态变更为【已拒绝】，申请的资源变更将不被执行，可联系平台管理员或查看审批备注了解拒绝原因。

审批结束后，管理员可通过审批管理中的已办审批列表查看平台上所有的审批记录及资源详细信息，方便后续针对平台和业务进行审计。

6.17.3 开启审批

平台支持管理员创建租户时为租户开启审批流程，同时支持为已创建的租户开启或关闭审批流程，方便平台的管理和运营。

平台默认审批流程是由平台管理员对资源申请进行审批，管理员新建租户时，支持为租户开启/关闭审批流程，并支持设置开启/关闭自动审批。

开启自动审批后，租户的主/子账号提供资源申请后，将自动进行审批，无需人工干预即可完成资源的审批和创建。

平台支持管理员为已创建的租户开启或关闭审批流程，同时支持开启或关闭自动审批

6.17.4 审批管理

审批管理为平台管理员提供整个云平台所有的审批流程记录, 包括待办和已办两个部分。

- 待办: 指租户发起资源变更申请后, 需要管理员进行审批的记录。
- 已办: 指管理员已经处理过的审批记录, 包括通过和拒绝的所有记录。

(1) 查看审批记录

平台管理可在审批管理列表查看平台所有待办和已办审批记录, 已办列表主要展示已由平台管理员审批过的申请记录; 待办列表展示尚未被审批的申请记录。

通过待办和已办列表信息, 管理员可分别查看需要处理及已处理的审批记录的列表信息, 包括申请名称、资源类型、操作、账号邮箱、账号 ID、创建时间、审批结果等。

待办列表的操作项中支持管理员对审批记录进行通过和拒绝操作, 同时为方便平台运营支持管理员下载审批管理的待办及已办信息为本地 Excel 文件。

支持管理员查看每条审批记录对应的申请详情, 包括申请的基本信息、资源信息、关联资源以及处理记录等。

租户在对资源进行变更并提交申请后, 可查看申请记录及申请状态。租户可查看租户下所有已提交的申请记录, 包括手工审批和自动审批的所有记录, 并可查看申请信息、申请涉及到的资源的信息、申请的处理记录、申请的关联资源。

(2) 通过申请

租户发起资源变更申请后, 管理员审批管理的待办列表中会生成一条审批记录, 支持管理员对租户的申请进行通过操作, 即同意用户的资源变更申请。审批通过后将自动为申请租户执行资源操作, 下发资源或对资源执行变更。

(3) 拒绝申请

租户发起资源变更申请后, 管理员审批管理的待办列表中会生成一条审批记

录，支持管理员对租户的申请进行拒绝操作，即拒绝用户的资源变更申请。审批拒绝后申请将直接结束流程，不再执行资源操作。

6.18 报表统计

报表统计是平台用于汇总和分析平台内各种资源数据的机制，包括资源用量统计及资源统计表，将各种数据整理成易于理解和分析的形式，提高平台整体运营和管理的效率。

6.18.1 资源用量统计

资源用量是云平台聚合全平台资源监控，根据多维度查询和指标分析，展示资源使用情况，支持管理员创建监控报告并导出 Excel 表格。

支持创建资源用量报告，从资源用量周期、租户信息、项目组信息、资源类型四个维度对控制台资源进行统计。

支持虚拟机、计算集群和存储集群三种资源类型的自定义时间周期资源用量报告，支持范围为 1 小时 ~ 6 个月的用量，可选择 1 天/3 天/7 天/14 天/30 天/自定义，自定义可将开始时间和结束时间精确到小时。

- **计算集群资源用量报表信息**

地域、资源类型、资源 ID、资源名称、架构、CPU 超分比例、CPU 总量、CPU 分配量、CPU 分配率、内存总量、内存分配量、内存分配率及统计时间。

- **存储集群资源用量报表信息：**

地域、资源类型、资源 ID、资源名称、集群架构、总量、已分配量、已分配率、已使用量、已使用率及统计时间。

- **虚拟机资源用量报表信息：**

地域、资源 ID、资源名称、资源类型、租户 ID、项目组 ID、状态、CPU 规格、内存规格、GPU 规格、CPU 平均使用率、CPU 最大使用率、内存

平均使用率、内存最大使用率、磁盘分区使用情况及统计时间。

当资源类型为虚拟机时，可查看虚拟机 CPU 使用率分布和虚拟机内存使用率分布，以使用率为横轴，虚拟机数量为纵轴进行统计。

支持资源用量统计的自动创建策略，通过制定保留数量、重复周期、执行时间、资源用量周期。

- 保留数量：需要保留的数量，超出此数量的最旧的将被删除。
- 重复周期支持单次、每天、每周、每月、间隔多种模式，单次执行默认当天执行，若执行时间已过则为次日执行；间隔支持按分钟或和小时进行间隔执行。
- 资源用量周期：生成时间范围内的资源用量报告。

支持查看平台所有创建的资源用量报表，并支持管理员删除资源用量报告。方便运营数据的统计，平台支持导出资源用量报告为本地 Excel 表格文件。

6.18.2 资源统计报表

资源统计表是呈现平台各类云资源清单的离线 Excel 表格，用户可以统一收集各云资源的基本信息，用于报表分析和自定义的数据处理。

平台支持虚拟机、云盘、快照、弹性网卡、VPC 资源、安全组、外网 IP、VIP、NAT 网关、VPN 网关、负载均衡、Redis、MySQL 等信息报表统计。

- 虚拟机资源：名称、资源描述、资源 ID、状态、所属租户、宿主机 IP、VPC ID、VPC 名称、子网 ID、子网名称、集群名称、镜像、GPU、CPU、内存、系统盘总量、数据盘总量、CPU 使用率、内存使用率、磁盘使用率、IP、计费方式、项目组、项目组名称、标签、高可用标签、创建时间、过期时间。
- 云盘资源：名称、资源描述、资源 ID、状态、是否加密、所属租户、租户 ID、集群架构、集群、硬盘类型、硬盘容量、绑定资源类型、绑定资源 ID、计费方式、项目组、项目组名称、标签、创建时间、过期时间。

- 快照资源：名称、资源描述、资源 ID、状态、是否加密、所属租户、租户 ID、硬盘 ID、硬盘类型、项目组 ID、项目组名称、标签、快照来源、创建时间。
- 弹性网卡资源：名称、资源描述、资源 ID、状态、所属租户、租户 ID、网卡类型、IP 地址、所属网络、绑定资源类型、绑定资源 ID、绑定资源名称、安全组、项目组、项目组名称、标签、创建时间、过期时间、付费方式。
- VPC 资源：名称、资源描述、资源 ID、状态、所属租户、租户 ID、网段、子网数量、VPC 网关状态、项目组、项目组名称、标签、创建时间。
- 安全组资源：名称、资源描述、资源 ID、状态、所属租户、租户 ID、规则数量、绑定资源数量、项目组 ID、项目组名称、标签、创建时间。
- 外网 IP 资源：名称、备注、资源 ID、状态、所属租户、租户 ID、IP 地址/网段、IP 版本、带宽、绑定资源 ID、路由类型、计费方式、项目组、项目组名称、标签、创建时间、过期时间。
- VIP 资源：名称、资源描述、资源 ID、状态、租户邮箱、租户 ID、关联资源、关联类型、VIP 类型、IP 地址、所属网络、项目组、项目组名称、标签、计费方式、创建时间、过期时间。
- NAT 网关资源：名称、资源描述、资源 ID、状态、机型、集群、租户邮箱、租户 ID、IP、IP 状态、VPC ID、VPC 名称、子网、子网名称、安全组、安全组名称、项目组、项目组名称、标签、创建时间、过期时间、计费方式。
- VPN 网关资源：名称、资源描述、资源 ID、状态、机型、集群、租户邮箱、所属租户、IP、VPC、子网 ID、隧道数量、计费方式、项目组、项目组名称、标签、创建时间、过期时间。
- 负责均衡资源：名称、资源描述、资源 ID、状态、机型、集群、租户邮箱、租户 ID、外网 IP、内网 IP、VPC ID、VPC 名称、子网 ID、子网名称、安全组 ID、安全组名称、vServer 数量、计费方式、项目组、项

目组名称、标签、创建时间、过期时间

- **Redis 资源：**角色、名称、资源描述、资源 ID、状态、所属租户、租户 ID、机型、IP 和端口、VPC ID、VPC 名称、子网 ID、子网名称、安全组、安全组名称、实例容量、计费方式、项目组 ID、项目组名称、标签、创建时间、过期时间。
- **MySQL 资源：**角色、名称、资源描述、资源 ID、状态、机型、集群、存储类型、版本、所属租户、租户 ID、ID、内存容量、数据盘容量、VPC ID、VPC 名称、子网、安全组、安全组名称、计费方式、项目组 ID、项目组名称、标签、创建时间、过期时间。

6.19 大屏监控

监控大屏是平台为企业提供的云平台资源可视化大屏，主要展示平台宏观维度的监控数据，帮助企业云平台运营者快速了解平台的整体运行情况，支持自定义拖拽模块及全屏展示，方便管理员进行分屏管理。



- **通知和告警：**展示最近 5 条平台告警信息。
- **物理机 TOP5：**展示 CPU 使用率、硬盘读吞吐、硬盘写吞吐、内存使用率在前 5 名的物理节点 IP 。
- **虚拟机 TOP5：**展示 CPU 使用率、硬盘读吞吐、硬盘写吞吐、内存使

用率在前 5 名的虚拟机 ID 。

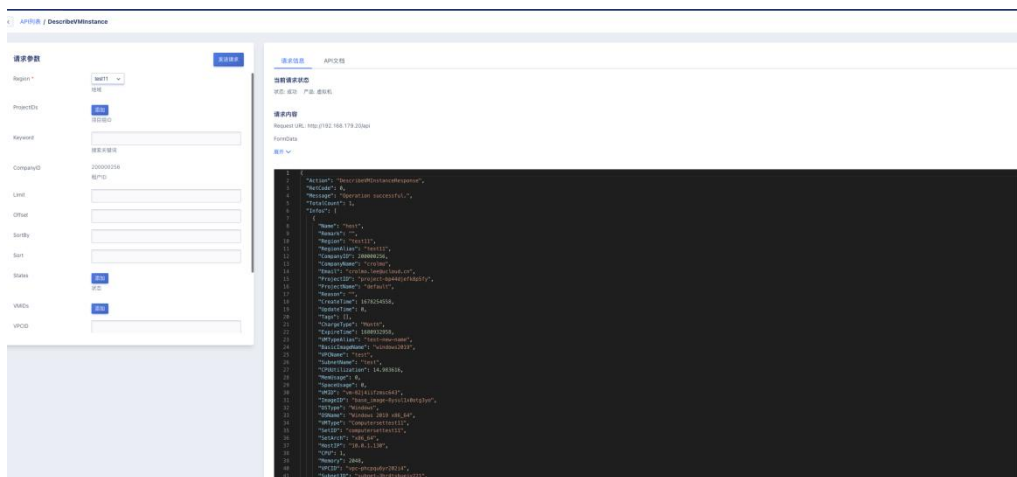
- 裸金属 TOP5: 展示 CPU 使用率、内存使用率在前 5 名的虚拟机 ID 。
- 资源分配: 展示云平台 CPU、内存、存储的总容量以及已分配容量的百分比。
- 资源概览: 展示物理机总量以及状态分布 (可用、锁定)、计算集群和存储集群的数量分布、虚拟机总量以及状态分布 (运营、关机、其他) 。

6.20 API 控制台

云平台 API 控制台, 通过界面提供对云平台 API 的调用和参数的解释说明。按照租户权限展示可以调用的 API 列表。

租户可以在页面添加请求参数, 发送请求对当前账号的线上资源操作, 请求发送成功后会在请求信息中展示当前请求状态, 请求内容, 以及响应结果, 在 API 文档中展示响应文档, 展示响应值的参数、类型及对应描述。

租户可按照产品模块查看对应产品子模块的 API 列表及 API 接口相关信息, 支持按照 API 名称和描述进行模糊搜索。



支持在 API 控制台直接发送请求, 添加请求参数后点击发送请求, 返回请求信息, 包含当前请求状态、请求内容、响应信息。

7 云平台管理

7.1 客制化能力

云平台为客户提供平台客制化能力，支持管理员自定义云平台 UI 展示样式，包括网站基本设置、监控大屏及登录页设置，如修改平台 logo 图片、登录页图片及标题等。

7.1.1 自定义网站展示 UI

网站设置是平台为企业和管理员提供的客制化能力，包括网站 Favicon 图片、网站 Title、云平台 Logo 图片，即自定义云平台的 Logo 及浏览器标志等。同时支持设置是否展示帮助文档、收藏夹、默认语言及默认币种等配置。

- 网站 Favicon 图片：浏览器标签页上展示的 Favicon 图片，必须为 ico 格式，最大不超过 100KB，推荐尺寸 48px*48px。
- 网站 Title：浏览器标签页上展示的网站说明，如中立安全可信赖的云计算服务商，支持中英文及特殊字符。
- 云平台 Logo 图片：租户和管理员控制台导航栏上方的 logo，允许管理员自定义，图片支持 png、jpeg、jpg 格式，最大不超过 200KB，推荐尺寸 352px*72px。
- 帮助文档：允许管理员开启或关闭控制台上的帮助文档的展示。
- 收藏夹：允许管理员开启或关闭控制上的收藏夹功能。
- 默认语言：允许管理员开启或半闭英文控制台。
- 默认币种：允许管理员修改默认币种，如 CNY 或 USD。

7.1.2 自定义监控大屏 UI

云平台提供监控大屏默认为出厂设置，允许平台管理员自定义监控大屏的标题，同时支持管理员配置是否在监控大屏展示裸金属的监控信息。

7.1.3 自定义登录页 UI

登录页设置是平台为企业和管理员提供的登录客制化能力，包括登录页标题、登录页标题颜色、背景图、登录页面 Logo、登录页联系电话、登录页输入框位置、登录页面版权、登录页描述信息、登录面输入框背景透明、OAuth 登录开关及 OAuth 认证配置信息等。

- 登录页标题：代表登录框上的标题描述，如私有云，可支持中英文及特殊字符。
- 登录页标题颜色：代表登录框上标题的颜色，支持白色、黑色、红色、黄色、绿色、青色、蓝色、棕色、紫色、橙色、灰色及金色，以适应不同背景图片上文字的可读性。
- 背景图：代表登录框后的背景图片，图片支持 png、jpeg、jpg 格式，最大不超过 500KB。
- 登录页面 Logo：图片支持 png、jpeg、jpg 格式，最大不超过 200KB，推荐尺寸 352px*72px
- 登录页联系电话：登录页的联系电话。
- 登录页输入框位置：支持居中、居左和居右。
- 登录页面版权：支持自定义登录页面的版本说明信息。
- 登录页描述信息：登录面的描述信息，登录页账户输入区域位置设置为居中时不展示该信息。
- 登录面输入框背景透明：可设置输入框为透明色，如不透明，背景色为白色。
- OAuth 认证：支持 OAuth 第三方统一登录，支持开启和关闭。开启时可进行 OAuth 服务端认证配置。
 - 认证服务器地址：OAuth 服务提供商的认证服务器地址，用于向用户发出认证请求并授权给客户端

- 客户端 ID: 由 OAuth 服务提供商分配给客户端的唯一标识符, 用于标识该客户端
- 请求响应类型: OAuth 认证流程中请求的响应类型, 通常为 code, 表示请求授权码
- 验证信息: 用于保护授权过程中的跨站请求伪造攻击 (CSRF) 攻击。该值将随着授权请求一起发送到 OAuth 服务提供商, 然后在授权过程结束时返回客户端, 以便客户端验证该值与请求时是否匹配

7.2 平台系统配置

7.2.1 邮箱配置

邮箱设置是指平台邮件服务的配置, 主要功能是找回密码、监报告警邮件的接收和发送。平台支持管理员定义邮箱的是否支持 SSL、发件人邮箱地址、发件人邮箱密码、邮件服务器 IP、邮箱服务器 Port 及邮件主题前缀。

- 邮箱支持 SSL: 配置邮箱是否支持 SSL。
- 发件人邮箱地址:配置发件人的邮箱地址。
- 发件人邮箱密码: 配置发件人邮箱密码
- 邮箱服务器地址:设置发件邮箱的 IP 地址
- 邮箱服务器端口: 设置发件邮箱的端口, 默认值为 994, 范围支持 0-65535。
- 邮箱主题前缀: 配置平台发送的提醒邮件的主题前缀, 如私有云、政务云等。

平台部署时默认必须提供邮箱设置, 避免无法接收找回密码及监报告警邮件。邮箱配置完成后, 支持用户对邮箱是否配置正确进行测试。

7.2.2 磁盘设置

7.2.2.1 全局磁盘 QoS

硬盘管理支持管理员对平台全局云硬盘开启或关闭 QoS 控制，以保证平台所有租户云盘资源的性能可靠性。

(1) 平台默认全局开启磁盘 QoS，即代表平台全局硬盘 QoS 生效，包括新建硬盘和已有硬盘的 QoS。

- 硬盘默认创建出来会根据平台计算公式赋予 QoS 值。
- 已有硬盘的 QoS 根据已赋予的默认值或管理员修改的值生效。
- 配置为开启时，管理员为每个硬盘自定义的 QoS 才可生效。

(2) 配置为关闭时，平台全局硬盘 QoS 失效，包括新建硬盘和已有硬盘的 QoS。

- 新创建的硬盘 QoS 不受限制。
- 已有硬盘的 QoS 不受限制。
- 配置为关闭时，管理员为每个硬盘自定义的 QoS 不会生效。

(3) 硬盘扩容容量后，会根据计算公式重新计算新容量的 QoS 值，根据计算的 QoS 值重新设置硬盘的 QoS。

- 若硬盘扩容前设置的 QoS 值 < 新容量 QoS 值，则以新容量 QoS 值为准。
- 若硬盘扩容前设置的 QoS 值 > 新容量 QoS 值，则以扩容前设置值为准。

7.2.2.2 全局磁盘设置

平台针对磁盘支持配置共享盘绑定虚拟机数量，并支持自定义每个云硬盘可创建的快照数量。

- 共享盘绑定虚拟机数量：支持限制共享云硬盘和共享外置存储盘可同时绑定的虚拟机数量，范围为 2~30 个。

- 单个硬盘快照数量：设置单个硬盘创建快照的数量上限，默认为 10 个，支持的范围为 0~200 个，0 代表单个硬盘创建快照的数量不受限制，管理员可根据平台实际使用情况调整上限数值。

7.2.3 网络设置

网络设置支持管理员对平台的 VPC 网段和内网带宽进行配置。

- VPC 网段设置：支持管理员为平台租户设置 VPC 可使用的 CIDR 网络，支持配置多段 VPC 网段。如 10.0.0.0/16,172.16.0.0/16,192.168.0.0/16。
- 内网带宽是指平台 VPC 网络内网带宽的 QoS，支持带宽 QoS 范围为 0~8192Mbps。

7.2.4 计费配置

(1) 平台计费功能

支持开启和关闭平台的计费功能，若关闭平台计费功能，平台所有资源的购买费用均为零，但系统依然会有交易和扣费记录产生。如无特殊情况，请勿关闭自动续费功能，防止资源过期。

(2) 自动续费开关

管理员为平台设置【资源是否自动续费】，默认值为是，即租户在账户余额充足时，将自动对租户下已有的计费资源进行续费。

若管理员关闭资源自动续费时，将不再对租户下资源进行自动续费，资源到期后会即变更为已过期状态，7 天内不进行续费将会自动进行删除或进入回收站。

7.2.5 回收策略

云平台支持管理员对全局资源开启或关闭删除能力，若开启则允许平台所有租户删除资源。若关闭资源删除能力，平台所有租户无法删除资源。

为保证资源的保留期和安全性，平台在资源删除时提供回收站能力，允许平

台管理员自定义回收站策略，如资源是否自动销毁、自动销毁周期及过期资源是否自动删除进入回收站等。

(1) 回收站资源是否自动销毁

设置资源进入回收站后，是否在固定周期后销毁，默认为开启状态；如果实际使用中，无需自动销毁进入回收站的资源，可关闭此配置。

(2) 回收站资源自动销毁周期

资源删除进入回收站后自动删除的周期，仅当开启回收站资源自动销毁时有效，默认值为 360000 秒，支持的范围为 1~360000 秒。

(3) 过期资源是否自动删除进回收站

开启后，虚拟机、云硬盘、外网 IP 等资源过期 7 天后会自动进入回收站。进入回收站时，虚拟机关联的资源将会被解绑。若平台的过期资源无需删除，则会自动变更为已过期状态，待 7 天后自动进行删除。

7.3 平台数据备份

平台数据备份服务是针对平台自身的数据库及配置文件进行备份，保证平台本身的数据安全性。支持平台数据库和配置文件备份。

- 平台数据库备份：平台部署后会自动生成一条自动备份策略。备份文件存放地址：/data/backups/taishancms-xxx.gz，默认每天备份一次。
- 平台配置文件备份：平台部署后会自动生成一条自动备份策略，备份文件存放地址：/data/backups/configs-xxx.tar.gz，默认每天备份一次。

默认定时策略为每小时备份一次，过期时间为 10 天。支持管理员修改定时策略，并支持平台管理员查看自动策略的每次执行记录。

自定义策略支持过期时间和重复周期配置，其中过期时间支持 1~99999 天，重复周期支持单次、每天、每周、每月及间隔等策略。

7.4 服务目录

服务目录是为云平台管理员提供一个统一的云服务管理入口，可快速查看和管理平台整体云服务的开通和授权情况，包括基础服务和高级服务两类。



基础服务：平台基础服务目录，包括虚拟机、镜像、VPC、云硬盘、网卡、快照、外网 IP、负载均衡、资源模版、安全组、NAT 网关、VPN 网关、弹性伸缩、VIP、监控告警、组播、隔离组、运维。

高级服务：平台高级组件服务目录，包括应用商店、Redis、MySQL、对象存储、文件存储、备份服务、API 控制台、外置存储、裸金属及 USB 透传。

已授权的服务为平台已经通过 License 开通授权的服务，未授权服务为平台暂未授权的服务。针对已授权的服务，平台支持管理员关闭和开启服务，并支持管理员为服务添加或移除租户，满足平台租户服务运营需求。

(1) 关闭服务

平台支持将已授权的服务整体关闭，服务关闭后，平台所有用户无法再使用此服务。

若有租户在服务下已存在资源，则无法关闭服务，关闭前需确保所有租户下此服务资源已被清空。

关闭服务功能使用场景较少，仅当云平台需要整体下线一项云服务时进行操作。

(2) 开启服务

平台支持将已关闭的服务进行整体开启，服务开启后，平台已开添加至服务的用户可正常使用该服务。

(3) 为服务移除租户

平台支持为一个服务移除租户，满足平台运营中需要指定某项或者某几项云服务开放给部分租户的使用场景。

当租户被移除后，租户下的所有账号将无法使用此服务，登录后将直接提示服务处于未授权状态。

(4) 为服务添加租户

平台支持为一个服务添加租户，已被移除服务授权的租户，可以通过添加租户的方式重新进行授权。为服务添加租户后，租户下的所有账号均可正常使用此云服务。

7.5 自定义规格

规格配置是平台为企业和管理员提供的自定义规格能力，管理人员可通过自定义规格配置调整云平台上架云产品服务的规格类型，包括虚拟机、硬盘、外网 IP 等。

- 虚拟机支持定义 CPU 和内存规格；
- 硬盘支持定义租户可创建云盘的容量范围；
- 外网 IP 支持定义租户可申请外网 IP 的带宽范围。

平台针对虚拟机、硬盘、外网 IP 会默认提供建议型的规格，管理员可根据

企业需求对规格进行变更，包括查看、删除、修改等。

(1) 创建规格

平台支持创建虚拟机 CPU/内存的规格，云硬盘和外网 IP 的规格由平台默认生成，仅支持修改。

创建虚拟机规格支持根据不同的集群创建不同的规格，即可为不同的集群创建不同的规格，租户创建虚拟机选择不同集群时，即可创建不同规格的虚拟机，适应不同集群硬件配置不一致的应用场景。可分别定义 CPU 和内存：

- CPU 规格支持 (C)：除 1 以外，以 2 的倍数进行增加，如 1C、2C、4C、6C，最大值为 240C。
- 内存规格支持 (G)：除 1 以外，以 2 的倍数进行增加，如 1G、2G、4G、6G，最大值为 1024G。

(2) 修改规格

创建出的规格即可被所有租户看到并使用，可根据业务需求在不同的集群中创建不同的规格，同时支持管理员对已创建和默认生成的规格值进行修改。

- 虚拟机产品规格支持修改 CPU 和内存值，其中 CPU 可指定 1、2、4、8、16、24、32、64；内存可过年费定 1、2、4、8、16、24、32、64、128，可根据业务需要自定义组合 CPU 和内存规格。
- 硬盘产品规格支持修改容量范围的最小容量和最大容量，其中最小容量和最大容量可设置的范围为 10GB 到 8000G，即平台允许租户创建的云硬盘最小为 10GB，最大支持到 32000GB，可根据业务需要调整最小值和最大值。
- 外网 IP 规格支持修改带宽范围的最小带宽和最大带宽，其中最小带宽和最大带宽可设置的范围为 1Mb 到 10000Mb，即平台允许租户申请的外网 IP 最小为 1Mb 最大为 10000Mb，可根据业务需要调整最小值和最大值。

(3) 删除规格

支持管理员删除指定自定义虚拟机规格，不支持删除硬盘和外网 IP 的规格。规格删除后平台租户即不可在所属集群中创建当前规格的虚拟机，但不影响通过该规格创建资源的正常运行。

虚拟机规格在每个集群内会生成一条无法删除的默认规格，以避免平台上所有规格均被删除，导致无法创建虚拟机。

7.6 配额管理

配额 (quota) 是一个租户 (包含子账号) 针对每种虚拟资源在一个地域下可创建的数量或容量限制。通过限制每个租户拥有的资源配额，可有效共享并合理分配云平台物理资源，提升资源利用率的同时，满足云平台上每一个账户的资源需求。

配额管理生效的前提是全局配置中【是否开启配额管理】的配置项为是，否则配额管理的配置值则不生效。

云平台全局提供每种资源在每个数据中心的默认配额，即每个租户创建时默认提供的资源配额模板。平台管理员可通过租户管理中的配额管理自定义每个租户的资源配额。租户主账号及所拥有的子账号不可自定义修改资源配额数量，仅提供查看配额。

子账号和主账号共享租户的资源配额，即每种资源配额为主账号和所包含的所有子账号可创建的资源数量或容量之和。如租户对于云硬盘的配额为 10，则租户的主账号及所有子账号可创建的云硬盘数量上限不可超过 10 个。

对于虚拟机、云硬盘、外网 IP 删除或未续费进入回收站的资源，不占租户资源配置，恢复资源时会检查资源配额。

管理员可根据平台实际使用情况调整默认配额，调整后及时生效，平台所有已有租户的配额将按照新的配额标准执行，新创建的租户也会按照新的配额标准设置配额限制。

目前配额限制主要包括虚拟机、虚拟机模版、镜像、VPC 数量、VPC 子网、云硬盘、网卡、外网 IP、安全组、负载均衡、负载均衡 SSL 证书、VPN 网关、

NAT 网关数量等。支持管理员查看平台全局配额配置，同时支持管理员对配额项进行修改操作。

平台支持为单个租户调整配额，修改全局配额设置不影响为租户自定义的配额设置。全局配额后修改后，若租户已有的产品数量超过设置值，不影响租户已有的资源运行；若租户将资源删除后，则无法再创建超过配额值的资源。

7.7 巡检服务

一键巡检，是平台提供的用来检查云平台健康情况的特性能力。通过对平台管理节点、计算节点的巡检项扫描，检查平台节点 CPU、内存、磁盘等资源的使用情况，使管理员更方便地对问题进行评估。

巡检主要是对平台进行全面扫描，包括管理节点的时间源同步检查、CPU 使用率、内存使用率检查、磁盘使用率检查；计算节点的物理机 CPU 平均使用率检查、物理机内存使用率检查、物理机系统盘已用容量检查等，一键巡检内容如下：

巡检类型	巡检项	巡检项含义	结果展示	巡检建议
管理节点	时间源一致性检查	检查是否设置时间源同步	提供节点当前时间源，和推荐时间源	若检测到时间源与集群内其他节点时间源不一致或物理机系统时钟未与时间源同步，请 SSH 登录对应系统，检查时间源配置
	CPU 使用率检查	检查云平台管理节点 CPU 的使用占比	提供当前占比，若超过 80%，提供最高使用率的五个进程	若检测到云平台 CPU 使用率在 10 分钟内持续的超过 80% 的使用率，请尽快联系平台相关人员进行热升级或问题评估，以继续正常使用本平台功能
	内存使用率检查	检查云平台管理节点内存的使用占比	提供当前占比，若超过 80%，提供最高使用率的五个进程	若检测到云平台内存使用率在 10 分钟内持续的超过 80% 的使用率，请尽快联系平台相关人员进行热升级或问题评估，以继续正常使用本平台功能

	磁盘容量检查	检查云平台管理节点磁盘的使用占比	提供当前占比, 若超过 70%, 提供占比最高的十个文件路径和文件大小	若检测到云平台磁盘数据容量已占用管理节点所在磁盘超过 70% 的容量, 请尽快联系平台相关人员进行磁盘检查和扩容, 以继续正常使用本平台功能
	管理服务检查	检查云平台管理服务的运行情况	提供当前服务名称及状态, 若服务异常, 提供异常的节点 IP	若检测到服务状态异常, 请根据提供的节点信息, SSH 登录对应系统, 检查服务的状态
计算节点	物理机 CPU 平均使用率检查	检查云平台上物理机 CPU 平均使用率	提供当前占比, 若超过 80%, 提供最高使用率的五个进程	若检测到物理机 CPU 平均使用率超过 70%, 请登录物理机系统, 确认物理机上是否存在异常进程。若未存在异常进程, 建议考虑对集群进行扩容
	物理机内存使用率检查	检查云平台上物理机内存平均使用率	提供当前占比, 若超过 80%, 提供最高使用率的五个进程	若检测到物理机内存使用率超过 80% 甚至 90%, 请立即登录物理机系统, 检查物理机上是否存在业务异常, 并按需优化运行业务。必要时, 建议对集群进行扩容
	物理机系统盘已用容量检查	检查云平台上物理机系统盘使用率和使用量	提供当前占比, 若超过 70%, 提供占比最高的十个文件路径和文件大小	若检测到物理机系统盘容量使用率超过 70% 甚至 90%, 请立即登录至物理机系统, 检查并清理对业务无影响的数据

支持用户下载巡检报告, 通过浏览器将巡检报告下载到本地。通过巡检报告可查看详细报告内容, 如地域、节点、名称、巡检项、巡检结果、现状、分数、建议以及最高使用率。

当巡检结果异常时, 展示当前设备参数现状及针对性的建议, 并展示导致结果异常的主要文件名称及大小, 使管理人员及时了解物理机状态并介入处理。同时针对已完成巡检任务的巡检报告, 平台支持管理员删除巡检报告。

7.8 统一授权

统一授权支持客户按需对基础服务模块和增值服务模块分开授权, 支持用户按照 x86/arm 架构区分授权节点, 根据业务需求选择各模块授权的生效时间和失效时间, 平台通过授权证书激活保证了密钥不可克隆验证的唯一性。

平台为用户提供完整的授权管理能力, 包括授权管理和节点管理两大模块。

7.8.1 授权管理

管理员可通过【信息采集】下载当前平台硬件及服务授权需求信息，并将信息发送给平台运营平台，运营平台确定用户需求并生成授权证书，并由平台管理员将证书上传至支平台。

通过授权管理用户可查看平台基础许可、拓展许可和服务许可，并可详细了解产品授权的状态、生效时间、失效时间和数量限制等信息。

- 基础许可：支持云计算基础设施管理系统套件-标准版、云计算基础设施管理系统套件-信创版和分布式块存储套件的授权；
- 拓展许可：支持云套件功能拓展、高级网络扩展、裸金属服务、USB 透传服务、GPU 服务、弹性伸缩服务、商用存储服务、文件存储服务、对象存储服务、MySQL 5.7 服务、Redis 4.0 服务和异构平台迁移软件服务的授权。
- 服务许可：支持云计算基础实施管理系统+增值服务 5*8 维保服务、云计算基础实施管理系统+增值服务 7*24 维保服务和金牌 VIP 维保服务。

7.8.2 节点管理

用户可通过节点管理查看平台所有节点授权状态及节点信息情况等。如节点 IP 地址、序列号、角色、授权状态、CPU 型号、CPU 总量、内存总量和架构等。

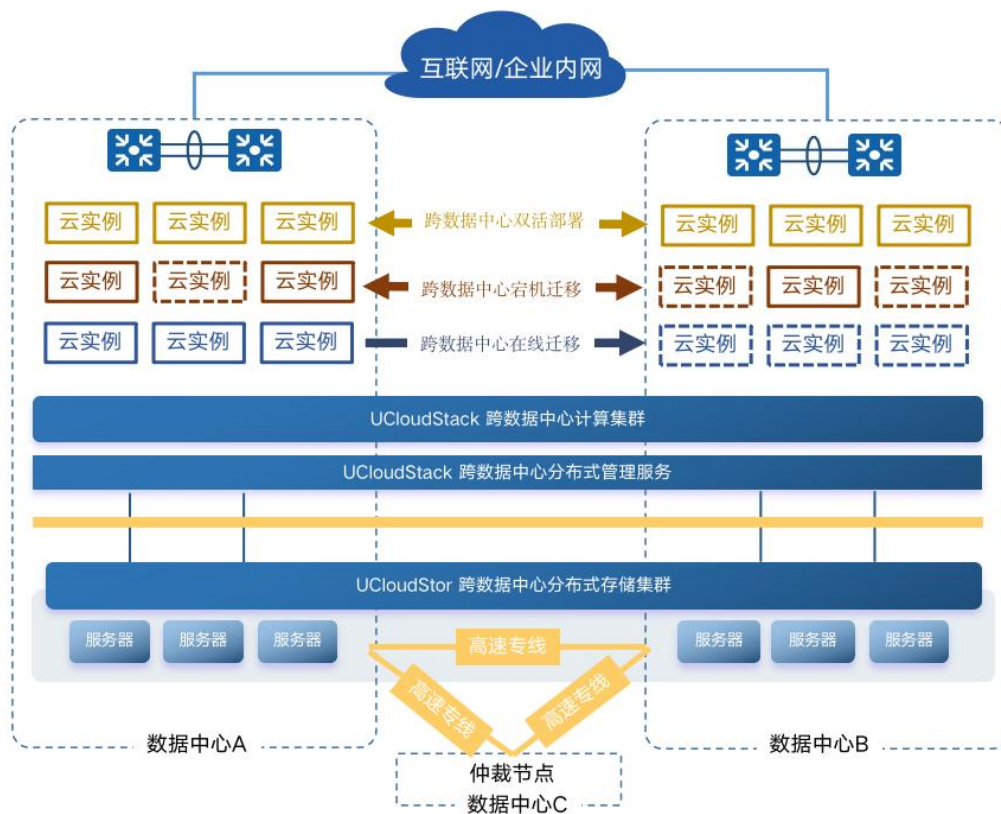
同时用户可通过节点管理查看授权节点的基本信息、CPU 信息和 Memory 信息。

8 双活数据中心

8.1 概述

出于数据隐私和安全性考量，私有云解决方案成为构建数字化转型的基础底座，通过“同城双活”及“两地三中心”的高可用架构保障生产环境稳定性和业务过程连续性。同时私有云在企业数字化转型中可提供更加快速灵活的 IT 资源交付和管控，支撑业务创新和变革。

私有云平台具备多数据中心部署和统一管理能力，帮助用户降低双活数据中心的建设门槛，快速实现业务跨数据中心的故障转移保障机制，从而进一步提升并保障业务连续性。



双活数据中心方案采用两个相互独立、互为备份的数据中心，可理解为在两个数据中心采用一套集群统一建设一套私有云平台，并共享一个互联网/企业内网出口。

当一个数据中心出现故障或宕机，可在另一个数据中心之间进行实时数据同

步和故障切换，避免业务系统因单点故障而中断，确保业务任何情况都能保持稳定运行。同时双活数据中心具备高度的可扩展性，可根据客户的需求进行自定义配置和扩容，满足不同业务场景下的需求。

通过平台提供的双活数据中心能力，对数据安全和业务连续性保障进行全面梳理和支撑，提升系统可靠性和连续性，助力企业在数字化转型中创造优质价值。

8.2 部署结构

在部署结构上，需将计算集群和存储集群的服务节点均一分为二，分别部署于两个数据中心，如 10 台计算节点，6 台存储节点：

- 则在 A 数据中心部署 5 台计算节点和 3 台存储节点，B 数据中心部署 3 台存储节点。
- 将 A 数据中心的 5 台计算节点和 B 数据中心的 5 台计算节点构建为一个计算集群
- 将 A 数据中心的 3 台存储节点和 B 数据中心的 3 台存储节点构建为一个存储集群。

两个数据中心的节点间网络可为二层或三层网络通信，需将两个数据中心间通过专线打通内网，保证两个数据中心间网络性能，避免影响分布式存储或虚拟机存储服务的性能。

由于双活数据中心仅在物理上部署在两个数据中心，逻辑上为一套云平台的一个集群，故平台 VPC 内网可直接通过物理网络进行通信，创建在两个数据中心计算节点上的虚拟机实例均可在一个 VPC 网络内。

两个数据中心共享统一的互联网/企业网出口，即外网 EIP 为统一出口，创建在两个数据中心计算节点上的虚拟机实例均可绑定统一出口配置的外网 EIP，并可通过外网 IP 与互联网或企业内网进行通信，并支持随实例进行浮动迁移。

双活数据中心的存储节点共同构建为一套存储集群，共享一套分布式存储系统，将存储管理服务分别部署于两个数据中心和一个仲裁数据中心，避免数据中

心脑裂，保证数据中心存储服务的可用性。

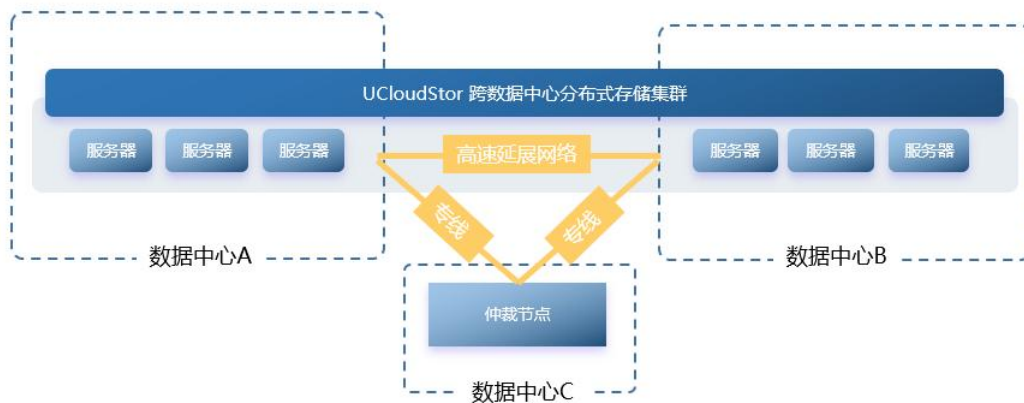
分布式存储数据冗余机制通过 4 副本的方式将分别将 2 副本存放于 2 个数据中心，采用数据同步、数据复制、多级故障域及故障自恢复等技术，实现数据在不同数据中心间的强一致同步，提高数据的可用性和可靠性。

云平台管理服务分别部署于两个数据中心和一个仲裁数据中心，采用分布式系统架构，保障云平台调度和管理服务的健壮性和可用性，使私有云在多个数据中心健康运行。

8.3 双活机制

(1) 跨数据中心强一致性数据保障机制，防止单点故障和数据丢失

平台内置分布式存储，作为私有云解决方案的核心存储系统。依托跨数据中心的存储双活能力，可实现跨数据中心强一致性数据保障，是双活数据中心方案的核心技术之一。

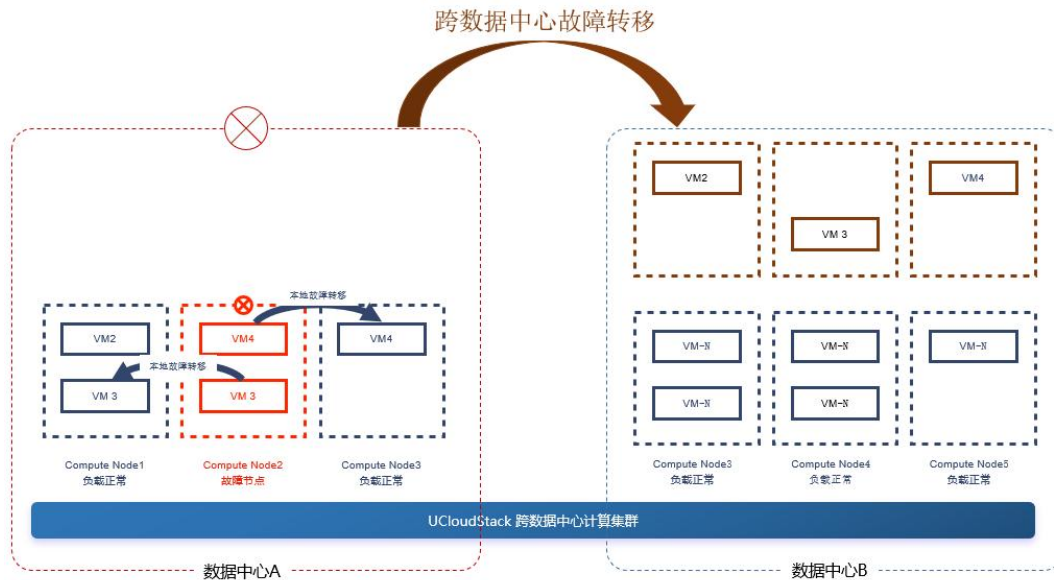


通过将数据存储在多个数据中心的，并采用数据同步和数据复制等技术，实现数据在不同数据中心之间的双向同步和备份，提高数据的可用性和可靠性，防止数据中心单点故障和数据丢失。在数据中心之间的网络质量符合方案要求的前提下，可以实现 RPO=0，RTO≈0，保证数据零丢失。

(2) 跨数据中心故障转移机制，有效降低故障恢复时间

平台默认提供本地故障转移调度策略和机制，当物理服务器发生宕机或故障

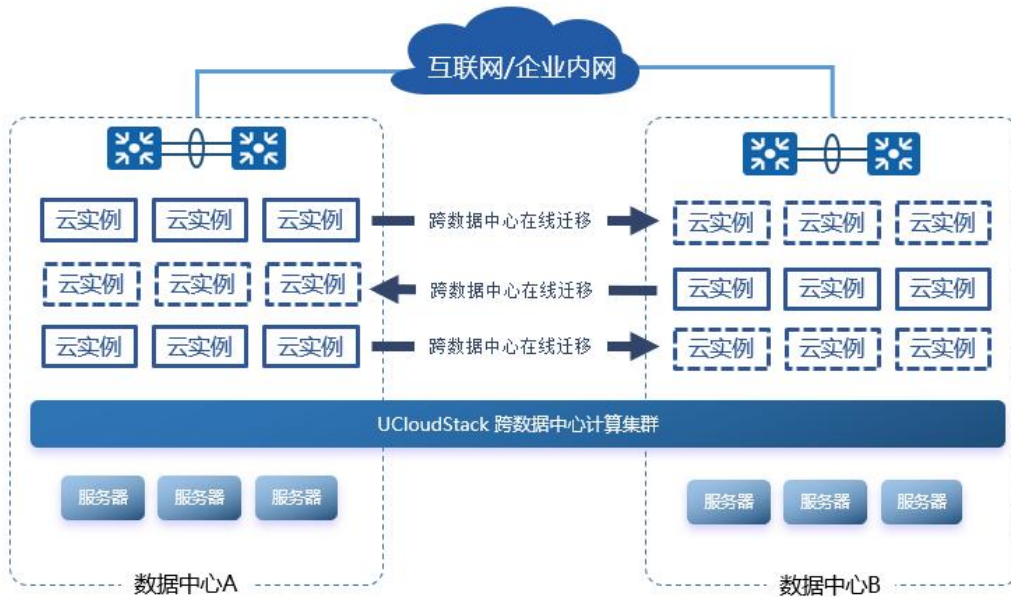
时，实现在同数据中心内进行故障转移，即：将故障物理机的云实例，向另一台有空闲资源的物理机上迁移并启动，从而大幅降低系统的故障恢复时间， $RPO=0$ ， $RTO<5min$ 。



跨数据中心的故障转移机制，在优先采用本地调度策略的基础上，增加多数数据中心调度属性，当数据中心出现极端的故障时，在对应用不做任何改造的条件下，将云实例迁移至健康数据中心的物理机，实现业务系统跨数据中心的容灾恢复， $RPO=0$ ， $RTO<5min$ 。

(3) 跨数据中心在线迁移机制，多数据中心资源平衡

在线迁移是计划内的云主机热迁移操作，即：云主机内的业务应用保持着持续对外服务的同时，云主机在不同的物理机之间进行在线跨物理机迁移，业务应用近似无感知。



跨数据中心在线迁移机制，即提供多数据中心迁移能力，使在线迁移不受限于同一数据中心。跨数据中心在线迁移机制，可以有效的进行多数据中心之间的资源平衡，以及计划内的跨数据中心热迁移。

(4) 跨数据中心分布式管理服务机制，保障系统健康运行

跨数据中心分布式管理是一种基于分布式系统架构的管理服务，用于支撑私有云平台本身健壮性的一组管理服务，同时可保障私有云在多个数据中心健康运行。该服务机制支持跨数据中心运行，通过将管理功能和资源分布到不同的数据中心，以实现跨数据中心的分布式管理和协作。

- **管理服务自愈能力**

基于分布式系统的建设原理，通过智能化和自动化的管理策略，可以自动监控和维护多个数据中心内的健康状况，减少人工干预和管理成本。在面临数据中心级别故障或异常情况时，管理服务可自动检测、定位、诊断和修复，从而保证云平台的稳定性和可靠性。

- **可视化监控**

提供全面的可视化监控和报告功能，帮助管理员了解云平台的状态和性能，及时发现和解决问题。

- **统一管理接口**

提供统一的管理接口和管理策略，方便管理员对整个系统进行集中化的管理和协作。

8.4 双活收益

(1) 降低双活数据中心建设门槛

传统的双活数据中心建设是一项较为复杂的集成类项目，项目周期长、涉及的软硬件产品多、运维成本高、建设效果参差不齐。

平台将双活数据中心建设所需的基础能力标准化和产品化，客户在建设过程中无需集成第三方产品，采用私有云平台标准的建设步骤即可快速完成双活数据中心的建设。同时配合平台轻量化特性，有效降低双活数据中心的建设成本。

(2) 进一步提升应用的业务连续性

通过双活数据中心建设，可在不改造业务系统的情况下，实现业务系统的跨数据中心故障转移机制，即跨数据中心宕机迁移和跨数据中心在线迁移，从而再次提升系统的业务连续性。

(3) 为跨数据中心双活应用的建设夯实基础

通过双活数据中心建设，即完成必要的双活数据中心基础建设，并内置一定的故障转移机制，为双活应用的改造提供充分准备。

用户仅需为应用增加跨数据中心的访问流量调度机制和业务系统本身的跨数据中心高可用，即可完成跨数据中心的的双活应用改造。

8.5 方案场景

私有云双活数据中心解决方案，可帮助客户进一步提升业务系统的可靠性，同时保障数据安全和隐私性，通过降低建设建设门槛和复杂性，提升建设效率，赋能客户加快数字化转型的进程。适用场景如下：

1. 高可用性要求较高的应用场景，如金融、电商、物流等行业，在业务高

峰期和数据中心故障时，能够保证系统的稳定运行。

2. 考虑建设灾备数据中心的客户，可在数据中心出现故障时，快速将应用切换到另一个数据中心，保证业务的连续性和数据安全性。
3. 建设双活数据中心需要付出高额成本的客户，平台双活数据中心可实现异地备份和故障转移，降低运营成本；同时支持在多个数据中心之间实现负载均衡和资源共享，有效提高资源利用率和运营效率。
4. 对于数据安全和合规性的有较高要求的客户，平台双活数据中心，提供数据备份和异地容灾能力，保证客户业务的数据的安全性和完整性，同时基于私有云的管理机制及安全等保，全面满足监管和合规要求。

私有云全面的双活数据中心能力，已在多个行业客户案例中得到验证。如某专注于智能营销云的大型集团，需针对数据爬虫、网站及大数据分析业务资源进行多数据中心部署和有效运营。

为保障客户业务的连续性和安全性，私有云平台提供双活数据中心能力，通过跨数据中心的数据强一致性保障、故障转移及跨数据中心统一管理等机制，为客户提供跨城双活灾备方案，使客户通过一套云平台统一管理并调度多数据中心资源，在提升业务连续性和数据可靠性的同时，大幅降低运营成本、提高资源利用效率。

9 平台安全性

私有云平台提供多维度且全面的安全保障体系，包括控制台安全、账号认证授权、网络安全控制、数据存储安全及日志审计等体系，并结合信息安全等级保护三级保证云平台和业务的安全性。

9.1 控制台安全性

私有云平台的管理控制台通过注册功能开启关闭、多次登录失败冻结、密码登录有效期、禁止多点登录、无操作自动退出、密码过期强制修改、自定义密码复杂度及密码不符合规则时强制修改等多方面保证平台的安全性和排它性。

- 注册功能开启/关闭：支持管理员为全局开启或关闭注册功能，开启后用户可通过平台自服务注册租户账号，关闭后租户账号仅支持管理员在平台进行创建。
- 多次登录失败冻结：支持管理员为平台开启或关闭全局账号登录多次失败冻结功能，开启后账号登录输入密码错误超过次数将被冻结。同时支持设置多次登录失败锁定时长，支持设置 15~1440 分钟。
- 密码登录有效期：支持管理员为平台全局账号默认开启密码到期强制修改策略，保证账号安全。默认设置为 90 天，可修改天数，每 90 天账号登录控制台时会强制要求修改账户密码。
- 密码过期强制修改：支持管理员为平台全局设置密码过期强制修改，开启后平台所有账号密码过期后必须强制修改；若关闭则密码过期后不强制要求修改。
- 禁止多点登录：支持管理员为平台全局账号开启多点登录，以适应对平台账号不同的登录需求。关闭时平台账号支持多个客户端同时登录，即一个账号在不同的客户端均可同时登录并管理平台资源。开启时代表平台账号仅支持单点登录，即一个账号同一时间仅支持在一个客户端进行登录，在其它客户端进行登录时将会自动退出已登录的客户端连接。

- 无操作自动退出时长：支持管理员为平台全局设置空闲时长，保证控制台资源和数据的安全。默认值为 30 分钟，即代表控制台在 30 分钟内无任何操作即会自动退出，支持设置 15~1440 分钟。
- 自定义密码复杂度：支持管理员对平台账号和虚拟机密码的长度进行配置，支持自定义配置为 6-30 位；同时支持自定义密码复杂度，如密码须包含有大写字母、小写字母、数字、特殊符号(除空格)中的两种或以上,不能包含[A-Z],[a-z],[0-9]和[() `~!@#\$%^&*~+=_[]{}:; '<> ,.?/]之外的非法字符。
- 密码不符合规则强制修改：支持管理员配置用户账号密码不符合长度及复杂度规则时强制修改。

9.2 账号认证授权

平台账号提供平台管理员账号和多租户账号，管理员账号用于整体平台资源管理及配置，多租户账号用于租户及子账号虚拟资源管理和使用。基于账号认证安全，平台提供多租户隔离、物理资源隔离、账号认证安全、角色权限授权、账号冻结、API 签名及审批流程等保障体系。

- 多租户隔离：支持多租户隔离，提供资源隔离、子账号管理、权限控制、配额及价格配置等能力。不同租户之间资源完全隔离，互不影响，从账号认证层面进行资源隔离；同时不同租户间资源可通过 VPC 网络及权限实现强隔离。
- 物理资源隔离：提供专属私有云方案，支持对计算存储集群物理资源进行权限控制，用于将部分物理资源独享给一个或部分用户使用，从物理层面保证资源隔离和安全性。
- 账号认证安全：管理员账号、租户主账号及子账号支持登录密码修改、登录邮箱修改、找回密码、双因子验证、数字证书、国密认证、API 密钥鉴权及登录访问限制等多种方式安全防护保障，保证支平台 API 访问入口及控制台的认证安全。

- 角色权限授权：支持平台资源级权限控制，以项目组为维度进行细粒度资源规划和管理，整合资源 API 操作权限为角色，通过角色授权将角色、项目组、子账号进行关联，使部分或全部资源以一种角色权限的集合授权给子账号，使子账号具备授权资源的角色权限，实现资源级精细化权限管控，达到达到不同人员以不同权限访问资源的效果。
- 账号冻结：支持平台对管理员账号、租户账号及子账号进行冻结，当账号被冻结时，无法进行登录和管理操作，保证平台账号的安全性。
- API 签名：云平台向用户开放产品服务资源操作 API 接口，并提供 API 接口访问密钥对（公钥和私钥），在用户调用 API 接口时，支持将公钥作为参数包含在每一个请求中发送，私钥负责生成请求串的签名，保证云平台 API 访问入口的安全性。
- 审批流程：针对企业云化资源的管理，云平台为企业管理者提供的自助模式的资源审批服务，用于制定信息系统云化资源的标准使用流程，在租户或子账号需要使用或管理资源时，按照流程中定义的审批人和审批层级完成审批后，由平台自动化交付用户所需业务资源。

9.3 网络安全控制

平台网络分为内网和外网两部分，分别通过 VPC 网络隔离、内网安全组、外网安全组、带宽 QoS、EIP 规格及 IPsecVPN 连接保证资源网络隔性及南北性流量安全和控制。

- VPC 网络：提供 VPC 隔离网络，VPC 内默认内网不通，租户内和租户间不同 VPC 网络默认不通，保证资源的网络隔离性。
- 内网安全组：内网虚拟防火墙，提供东西向出入双方向流量的访问控制规则，支持 TCP、UDP、ICMP、GRE 等协议数据包的过滤和控制，用于限制虚拟资源的東西向网络访问流量。
- 外网安全组：外网虚拟防火墙，基于外网 IP 提供南北向出入双方向流量的访问控制规则，支持 IPv4 和 IPv6 的 TCP、UDP、ICMP、GRE

等协议数据包的过滤和控制，用于限制南北向网络访问流量。

- **带宽 QoS:** 支持带宽 QoS 配置，即对 VPC 网络内网带宽的 QoS 进行配置，避免多租户虚拟资源对内网带宽的争抢。
- **EIP 规格:** 提供外网 IP 带宽规格上限配置能力，保证平台访问外网时的网络可靠性。
- **IPSecVPN:** 提供可容灾的高可用 IPSecVPN 服务，采用 IPSec 安全加密通道实现企业数据中心、办公网络与平台 VPC 私有网络的安全可靠连接，支持多种加密及认证算法并提供 VPN 连接健康检测及连接日志，保证隧道连接的可靠性、安全性及管理便捷性。

9.4 数据存储安全

平台数据存储通过数据保护机制、云存储安全、云存储加密、存储快照、云存储加密、对象存储鉴权、磁盘 QoS、平台数据备份及数据备份服务等安全体系，全面保证底层存储和数据存储的安全性。

- **数据保护机制:** 提供多副本数据冗余机制，通过多副本、写入确认机制及副本分布策略等措施，自动屏蔽软硬件故障并自动进行副本数据备份和同步，保证数据安全性和可用性。
- **云存储安全:** 虚拟机存储在分布式存储系统中的文件被两次拆分成 32KB 的块文件，完全打散写入至整个存储集群的所有磁盘中，读取数据必须经过元数据和其它盘上块文件进行数据整合，保证数据安全。
- **云存储加密:** 针对虚拟机的云硬盘和数据安全，平台提供云硬盘加密特性，使用 LUKS 加密规范来对磁盘全盘加密，保护用户的数据不被未经授权的访问者获取，甚至在磁盘丢失或被盗的情况下也可以保证数据的机密性。
- **存储快照:** 云平台分布式存储提供磁盘快照能力，支持对虚拟机系统盘、云硬盘、共享盘、文件存储实例、对象存储实例等进行手动和自动快照，降低因误操作、版本升级等导致的数据丢失风险，是平台保证数据安全

的一个重要措施。

- 对象存储鉴权：平台对象存储提供全面鉴权认证体系，支持对象存储用户访问流量的密钥认证管理，保证存储数据的安全性。
- 磁盘 QoS：平台全局默认提供全局云硬盘 QoS 配置，即新创建的云盘会根据平台公式赋予 QoS 值，限制平台用户对磁盘性能强行占用，保证平台所有租户云盘资源的性能可靠性。
- 平台数据备份：针对平台自身的数据库及配置文件支持备份特性，保证平台本身的数据安全性。支持平台数据库和配置文件备份。
- 数据备份服务：提供数据备份和恢复功能的服务，允许用户将关键数据和文件复制到另一个存储介质，以便在数据丢失、损坏或灾难恢复时进行恢复。支持对 MySQL 服务、Redis 服务、对象存储及文件存储等实例和数据进行定时自动备份和手动备份。

9.5 日志审计体系

平台提供全面的日志审计能力，包括操作日志及事件管理，实现安全分析、资源变更追踪以及合规性审计。

操作日志审计：平台提供全面操作日志，包括控制台或 API 资源的操作行为及登录登出审计信息。操作日志会记录用户在平台中的所有资源操作，提供操作记录查询及筛选，通过操作日志可实现安全分析、资源变更追踪以及合规性审计。

事件日志审计：平台提供资源事件审计能力，对云平台核心资源的部分操作及状态进行记录及通知，如资源生命周期状态的变化、操作运维执行情况等。资源事件记录用户在资源类型的部分核心操作事件，提供事件详细记录查询及筛选，并可配合通知规则及时通知用户定位问题。

10 平台可靠性

平台可靠性通过数据中心、硬件设施、云平台软件、云平台服务及云平台升级等多维度高可用设计保证平台整体可靠性，进一步保证用户业务连续性。

10.1 数据中心

(1) 双活数据中心

平台具备多数据中心部署和统一管理能力，采用两个相互独立、互为备份的数据中心，可理解为在两个数据中心采用一套集群统一建设一套私有云平台，并共享一个互联网/企业内网出口。

当一个数据中心出现故障或宕机，可在另一个数据中心之间进行实时数据同步和故障切换，避免业务系统因单点故障而中断，确保业务任何情况都能保持稳定运行。同时双活数据中心具备高度的可扩展性，可根据客户的需求进行自定义配置和扩容，满足不同业务场景下的需求。

通过平台提供的双活数据中心能力，快速实现业务跨数据中心的故障转移保障机制，提升系统可靠性和连续性，助力企业在数字化转型中创造优质价值。

(2) 异地容灾

在数据中心维度支持容灾方案，根据不同容灾等级的需求，分别通过专线、SD-WAN、VPN、互联网等互联多个异地数据中心。将业务和数据按需求部署或异地复制到异地数据中心，并通过智能 DNS 进行异地数据中心的业务切换。

(3) 机柜级冗余

网络设备和服务器硬件设备均对称部署于机柜，如 3 台计算节点服务器分别对称部署于 3 个机柜，一个机柜一台服务器，单机柜掉电或故障不影响业务正常运行和使用。

10.2 硬件设施

(1) 内外网物理隔离

内网业务和外网业务在物理网络设备上完全隔离，避免内外网业务相互影响。内网和外网分别使用独立的网络交换机设备，每台服务器提供内网网卡和外网网卡，分别接入至内网交换机和外网交换机。

(2) 网络设备高可用

- 网络设备扩展：网络设备扩展性设计，所有网络设备分为核心和接入两层架构，一套核心可水平扩展几十套接入设备。
- 网络设备冗余：网络设备冗余性设计，所有网络设备均为一组两台堆叠，避免交换机单点故障，实现交换机级别高可用。
- 网络接入冗余：交换机下联接入冗余性设计，所有服务器双上联交换机的接口均做 LACP 端口聚合，避免单点故障，实现交换机互联高可用。

(3) 服务器高可用

- 接入层冗余：服务器网络接入冗余性设计，所有服务器节点均做双网卡绑定，分别接入内网交换机和外网交换机，避免单点故障，实现服务器网络接入高可用。
- 管理节点冗余：管理节点冗余性和扩展性设计，多台管理节点分布式部署，并支持横向扩展，避免管理节点单点故障，实现管理服务高可用。
- 计算节点高可用：云平台通过智能调度系统将虚拟机均衡部署于计算节点，可水平扩展计算节点数量。支持虚拟机在线迁移、故障转移、智能均衡部署，实现计算节点高可用及虚拟机高可用。
- 分布式存储节点：分布式存储冗余性设计，将数据均衡存储于所有磁盘，并通过三副本保证数据安全。数据及副本可放置于不同机柜、不同节点及不同磁盘上，尽可能保证数据安全性；同时分布式存储服务器本身提供冗余性负载设计，节点损坏不影响分布式存储使用并不会丢失数据。

10.3 云平台软件

- 分布式调度：基于分布式服务调用和远程服务调用为租户提供智能调度

模块。智能调度模块实时监测集群和所有服务节点的状态和负载，当某集群扩容、服务器故障、网络故障及配置发生变更时，智能调度模块可根据隔离组自动迁移被虚拟资源到健康的服务器节点，保证云平台的高可靠性和高可用性。

- 分布式网络：1) 云平台采用分布式 **Overlay** 网络部署模式，所有的虚拟网络均部署于所有计算节点，管理节点仅作为管理角色，不承担网络组件部署及生产网络传输，所有生产网络仅在计算节点上传输，无需通过管理节点进行转发。分布式存储直接通过物理网络进行挂载，无需通过虚拟网络进行挂载和传输，虚拟化与存储通过本机或跨物理机内网进行通信。2) 分布式网络架构将业务数据传输分散至各个计算节点，除业务逻辑等北向流量需要管理节点服务外，所有虚拟化资源的业务实现等南向流量均分布在计算节点或存储节点上，即平台业务扩展并不受管理节点数量限制。
- 资源管理：通过分布式资源管理模块，负责集群计算、存储、网络等资源的分配及管理，为云平台租户提供资源配额、资源申请、资源调度、资源占用及访问控制，提升整个集群的资源利用率。
- 安全管理：为租户提供身份认证、授权机制、访问控制等功能。通过 **API** 密钥对和用户名密码等多种方式进行服务间调用及用户身份认证；通过角色权限机制进行用户对资源访问的控制；通过 **VPC** 隔离机制和安全组对资源网络进行访问控制，保证平台的安全性。
- 集群部署：提供自动化部署集群节点的模块，为运维人员提供集群部署、配置管理、集群管理、集群扩容、在线迁移及服务节点下线等功能，为平台管理者提供自动化部署通道。
- 集群监控：提供平台集群物理资源和虚拟资源信息收集、监控及告警。通过自动化获取资源的运行状态信息，并将信息指标化展示给用户；同时提供监控告警规则，通过配置告警规则，对集群的状态事件进行监控及报警，并有效存储监控报警历史记录。

- 业务实现分离：云平台架构从业务逻辑上分为北向接口和南向接口，将云平台的业务逻辑和业务实现进行分离，业务管理逻辑不可用时，不影响虚拟资源的正常运行，整体提升云平台业务可用性和可靠性。业务实现分离后，当云平台业务端（如 WEB 控制台）发生故障时，并不影响已运行在云平台上的虚拟机及运行在虚拟机中的业务，一定程度上保证业务高可用。
- 云平台升级：云平台支持在线扩展计算节点，并支持平滑升级所有服务，保证平台扩容时的业务可用性。

10.4 云平台服务

(1) 弹性计算

- 智能调度：智能调度优先选择低负荷节点进行虚拟资源部署，并提供打散部署、在线迁移、离线迁移及宕机迁移等能力，整体保证云平台的可靠性。
- 部署策略：将多个虚拟机加入隔离组，用于控制云主机的分布以保证业务高可用性。可自定义虚拟机与其他虚拟机或宿主机之间的亲和关系。支持亲和性和反亲和性两种策略类型，可有效提高业务服务的可用性。
- 自制镜像：自制镜像由云平台用户通过虚拟机自行导出的自有镜像，可用于创建虚拟机，提高平台虚拟机资源的可用性。

(2) 网络服务

- 二层隔离：私有网络是一个属于用户的、逻辑隔离的二层网络广播域环境。私有网络是子网的容器，不同私有网络之间是绝对隔离的，保证网强的安全性。
- NAT 网关：云平台提供 NAT 网关，是一种网络地址转换协议的 VPC 公网网关，为云平台资源提供 NAT 代理（SNAT、DNAT），通过 NAT 网关可以让 VPC 子网中未绑定弹性 IP 的虚拟机访问外网，同时可配置端口转发规则使虚拟机对外提供服务，从网络层面为云平台提供高可用

机制。

- 安全组：提供出入双方向流量访问控制规则，定义哪些网络或协议能访问资源，用于限制虚拟资源的网络访问流量，为云平台提供必要的安全保障。
- 负载均衡：通过平台负载均衡服务提供的虚拟服务地址，将同一种业务应用虚拟构建为一个高性能、高可用、高可靠的应用服务器池，支持轮询、加权轮询、最小连接数和源基于源地址的的负载调度算法，并根据负载均衡的转发规则，将来自客户端的请求均衡分发给服务器池中最优的虚拟机进行处理，并支持会话保持和健康检查，自动检测并隔离服务不可用的实例，迅速将故障虚拟机进行切换，确保业务服务的可用性。

(3) 弹性存储

- 分布式统一存储：云硬盘采用大规模分布式存储系统，将整个计算&存储集群中的存储资源虚拟化后，整合在一起对外提供统一的存储服务。分布式存储系统通过三副本、写入确认机制及副本分布策略等措施，最大限度保障数据安全性和可用性。
- 三副本冗余：用户通过虚拟机应用程序写入云硬盘的数据，会根据分布式存储系统三副本机制存储三份，并按照副本分布算法，分别存储于不同物理主机的磁盘上。三副本机制存储数据，将自动屏蔽软硬件故障，磁盘损坏和软件故障，导致副本数据丢失，系统自动检测到并自动进行副本数据备份和同步，不会影响业务数据的存储和读写，保证数据安全性和可用性。
- 写入确认机制：三副本在写入过程中，只有三个写入过程全部被确认，才返回写入完成，确保数据写入的强一致性。
- 数据分布策略：支持副本数据落盘分布策略，可将三副本数据分布在不同磁盘、不同主机、不同机柜甚至不同机房，避免因单主机及单机柜整体故障造成数据丢失或不可用的故障，保证数据的可用性和安全性。为保证云硬盘数据访问时延，通常建议最多将数据副本保存至不同的机

柜，若将数据三副本保存至不同的机房，由于网络延时等原因，可能会影响云硬盘的 IO 性能。

(4) 日志和监控

- 操作日志：云平台提供资源级别操作日志及平台组件化系统操作日志，记录云平台所有操作信息，方便定位故障及相关平台运营信息。
- 审计日志：提供云平台所有登陆登出操作审计信息，包括租户及管理云平台登录登出的信息。
- 资源事件：平台提供核心资源的部分操作及状态进行记录及通知，如资源生命周期状态的变化、操作运维执行情况等。
- 监控告警：平台全线产品的运维监控及告警服务，提供全线资源实时监控数据及图表信息，可根据监控数据批量为资源设置告警策略，并在资源故障或监控指标超过告警阈值时，以邮件的方式给予通知及预警，全方位保障业务的可靠性和安全性。

(5) PaaS 服务高可用

- NAT 网关：NAT 网关实例支持高可用架构，即至少由 2 个虚拟机实例构建，支持双机热备。当一个 NAT 网关的实例发生故障时，支持自动在线切换到另一个虚拟机实例，保证 NAT 代理业务正常。同时基于外网 IP 地址的漂移特性，支持在物理机宕机时，保证 SNAT 网关出口及 DNAT 入口的可用性。
- 负载均衡：支持负载均衡主备版，满足负载均衡网关高可用场景。当其中一个负载均衡的虚拟机实例因故障宕机时，备实例会自动转换为主实例持续提供负载均衡服务。
- VPN 网关：支持 VPN 网关主备版，满足 VPN 网关高可用场景。当其中一个网关的虚拟机实例因故障宕机时，备实例会自动转换为主实例持续提供 VPN 网关的通信服务。
- MySQL 服务：支持 MySQL 服务主备版，满足 MySQL 服务高可用场景。

当其中一个 MySQL 实例因故障宕机时，备实例会自动转换为主实例持续提供 MySQL 数据库服务。

- **Redis 服务：**支持 Redis 服务主备版，满足 Redis 服务高可用场景。当其中一个 Redis 实例因故障宕机时，备实例会自动转换为主实例持续提供 Redis 缓存服务。
- **对象存储：**支持对象存储服务主备版，满足对象存储服务高可用场景。当其中一个对象存储实例因故障宕机时，备实例会自动转换为主实例持续提供对象存储服务。
- **文件存储：**支持文件存储服务主备版，满足文件存储服务高可用场景。当其中一个文件存储实例因故障宕机时，备实例会自动转换为主实例持续提供文件存储服务。

(6) 资源配额

- 支持云平台全局配置每种资源的配额，防止资源利用溢出，导致云平台业务不可用，优先保证重要业务的资源分配和使用。
- 支持用户及子帐号级别的资源配额定义，优先保证线上帐号资源分配和使用。

11 灾备服务

云平台通过分布式存储系统保证本地数据的安全性，同时通过远程数据备份服务，为用户提供远程数据备份和容灾备服务，可以将本地云端数据统一归档、备份至远程云平台，保证本地发生重大灾难时，可通过远端数据中心快速恢复业务。容灾方案需考虑两个核心的指标：

- **RTO (RecoveryTime Object)**

恢复时间目标，指数据中心发生灾难后，应用系统从宕机到业务恢复所需的时间，即业务恢复的及时性体现，代表可以容忍业务最长恢复时间。RTO 值越小，代表需越快恢复业务，相对成本也较高；

- **RPO (Recovery Point Object)**

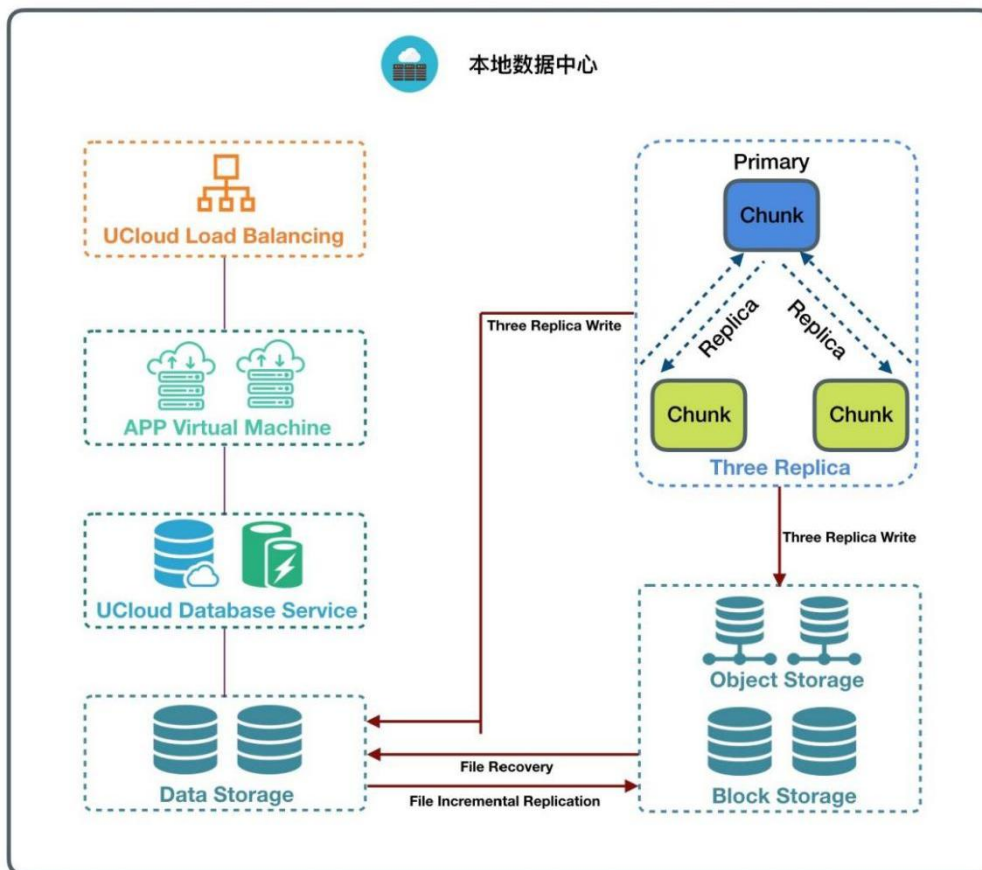
恢复点目标，指数据中心发生灾难后，灾备系统恢复的数据对应的时间点，即应用发生故障时，可以容忍的最大数据丢失量。RPO 值越小，代表数据越重要，需提高对数据备份的频率，相对成本也较高；

RTO 和 RPO 的标准与容灾方案的成本为线性关系，对于 RTO 和 RPO 的需求，需考虑业务系统本身特征及成本等方面因素，详见信息安全技术信息系统灾难恢复规范。

灾备服务支持本地灾备、异地灾备、公有云灾备、两地三中心等多种服务方式，可根据业务特点和需求，灵活选择灾备方式，保证业务的 RTO 和 RPO。

11.1 本地灾备

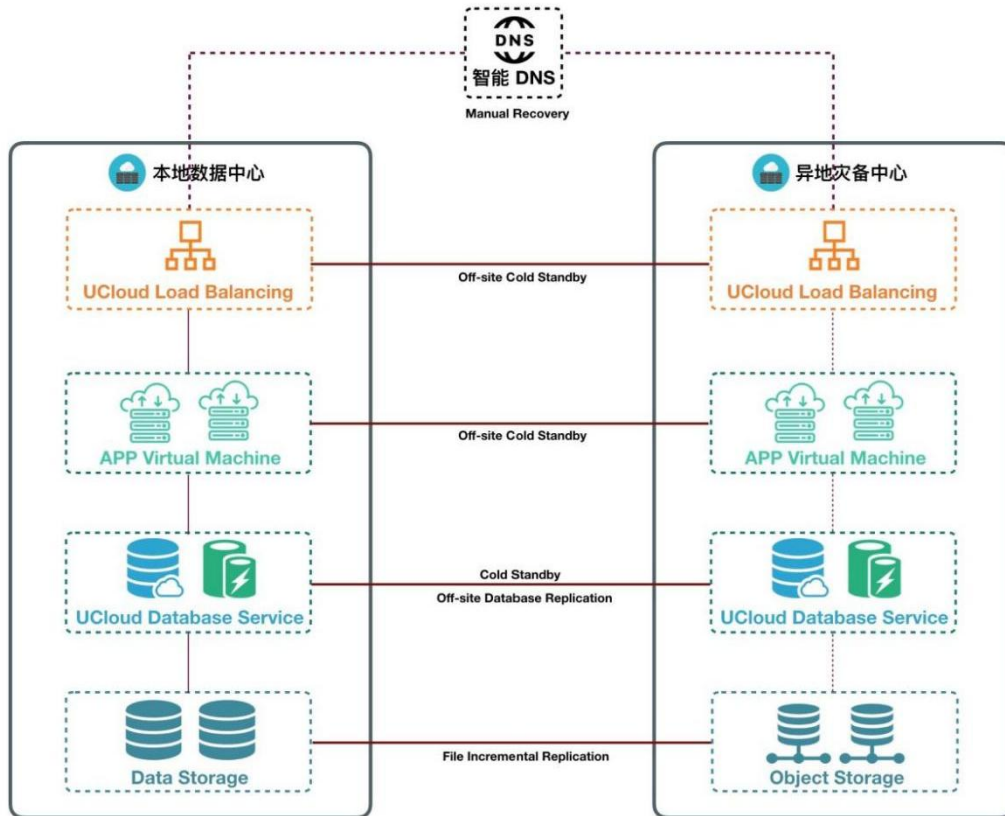
平台通过分布式存储系统、RAID5 及多副本机制，自动屏蔽软硬件故障，磁盘损坏和软件故障，系统自动检测到并自动进行副本数据备份和迁移，保证本地数据安全性。详见分布式存储。同时平台支持将本地虚拟机、镜像、云硬盘、数据库等数据定时增量备份至对象存储服务。本地灾备架构如下图所示：



- 平台支持灵活的备份和恢复策略，可通过不同时间维度，全量或增量的方式备份数据。
- 当本地数据损坏或误删时，可将本地备份数据还原至平台，恢复业务数据及业务运行。
- 当本地数据中心发生灾难时，可通过异地灾备、公有云灾备等方式重建数据中心并恢复业务。

11.2 异地灾备

云平台在保证本地数据中心的业务数据安全的同时提供异地灾备服务，将云业务镜像及数据通过专线、SD-WAN、VPN 或互联网连接以增量的方式复制到异地对象存储服务，确保业务数据 RPO 指标。当本地数据中心发生灾难时，可快速通过异地数据恢复业务。



异地灾备服务支持多种业务部署方式，为云平台业务提供不同 RTO 指标，控制云平台业务灾备成本。

(1) RTO 指标高，业务恢复时间长，成本低

业务部署：在异地灾备中心仅部署对象存储服务，将本地数据中心云业务镜像、业务数据及数据库以全/增量的方式复制到对象存储服务中；

业务恢复：通过将异地灾备中心对象存储服务的备份数据还原至本地，在本地恢复云业务及数据，重建本地数据中心；

(2) RTO 指标低，业务恢复时间短，成本高

业务部署：所有业务应用、数据库、负载均衡分别部署在本地数据中心和异地灾备中心；

- 本地数据中心为 Active 模式，异地灾备中心为 Cold Standby 模式；
- 负载均衡：每个业务的负载均衡实例均在灾备中心各部署一套；

- 虚拟机：针对业务对 RTO 不同需求，灾备中心可部署相同配置或降级配置的虚拟机；
 - 对于 RTO 要求较高的应用，在灾备中心需部署与生产中心相同配置的虚拟机，用于满足业务切换时，可快速恢复业务，并保证业务运行环境的性能；
 - 对于 RTO 要求较低的应用，在灾备中心可部署降级配置的虚拟机，以节省资源和成本；
- 数据库：针对每个业务在灾备中心部署一套相同的数据库服务，灾备中心数据库均为只读模式，两地数据库采用异步方式进行数据复制；
- 存储：灾备中心部署对象存储服务，灾备中心数据库直接连接对象存储进行数据读写；
- 智能 DNS：通过智能 DNS 服务，将业务域名 A 记录配置为本地数据中心 LB 的 IP 地址；

数据复制：针对业务对 RPO 的不同需求，两个数据中心采用多种网络互联；

- 数据库服务将数据通过异步方式复制到灾备中心对象存储服务；
- 虚拟机镜像、业务数据、文件数据以全/增量的方式，从本地数据中心复制到灾备中心对象存储；

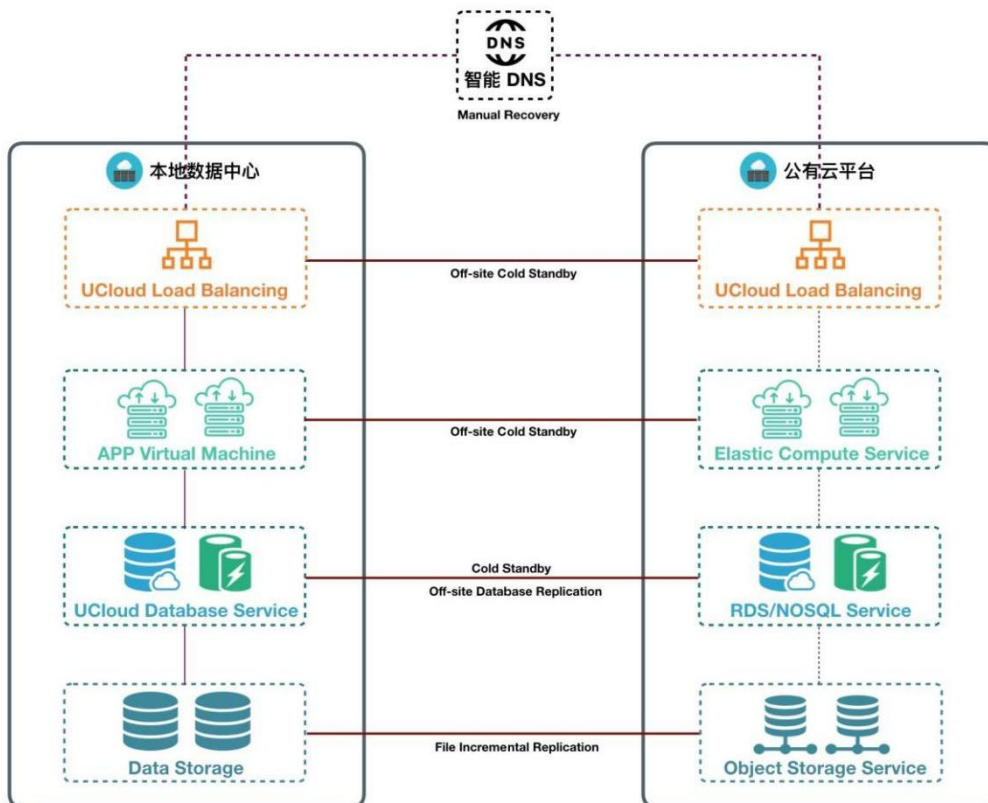
业务恢复：当本地数据中心发生灾难或需要业务切换时，修改业务域名 A 记录和数据库状态；

- 通过智能 DNS 将业务域名 A 记录手动修改为灾备中心业务 LB 的 IP 地址，实现故障切换和业务恢复；
- 将灾备中心业务应用数据库服务修改为读写状态，业务应用数据库直接读写对象存储中的数据；

异地灾备中心与本地数据中心网络互联方式，会影响业务数据备份的频率和完整性；由于异地网络延时的影响，不建议两地数据中心均为 Active 模式。

11.3 公有云灾备服务

云平台提供本地数据中心到公有云平台的灾备服务，将云业务镜像及数据通过专线、SD-WAN、VPN 或互联网连接以增量的方式复制到第三方公有云平台对象存储服务，确保业务数据 RPO 指标。当本地数据中心发生灾难时，可快速在公有云上恢复业务，同时也可将公有云上的业务数据备份还原至本地，重新本地数据中心。



公有云灾备服务支持多种业务部署方式，为云平台业务提供不同 RTO 指标，控制云平台业务灾备成本。

(1) RTO 指标高，业务恢复时间长，成本低

业务部署：在公有云平台仅申请对象存储服务，将本地数据中心云业务镜像、业务数据及数据库以全/增量的方式复制到对象存储服务中。

业务恢复：通过将公有云对象存储服务的备份数据还原至本地，在本地恢复业务，重建本地数据中心；使用对象存储备份数据，在公有云平台直接部署业务

云主机、负载均衡、数据库等服务，恢复业务；

(2) RTO 指标低，业务恢复时间长，成本高

业务部署：所有业务应用、数据库、负载均衡分别部署在本地数据中心和公有云平台；

- 本地数据中心为 **Active** 模式，公有云平台为 **Cold Standby** 模式；
- 负载均衡：每个需要负载均衡的业务均在公有云申请一个负载均衡实例，并将业务云主机加入后端；
- 云主机：针对业务对 **RTO** 不同需求，公有云可部署与本地数据中心相同或降级配置的云主机；
 - 对于 **RTO** 要求较高的应用，公有云需部署与生产中心相同配置的云主机，用于满足业务切换时，可快速恢复业务，并保证业务运行环境的性能；
 - 对于 **RTO** 要求较低的应用，公有云可部署降级配置的云主机，以节省资源和成本；
- 数据库服务：针对每个业务在公有云平台部署一套相同的数据库服务，云平台数据库服务均为只读模式，两地数据库采用异步方式进行数据复制；
- 对象存储服务：公有云平台部署对象存储服务，公有云平台业务应用数据库直连对象存储进行数据读写；
- 智能 DNS：通过智能 DNS 服务，将业务域名 A 记录配置为本地数据中心 LB 的 IP 地址；

数据复制：针对业务对 **RPO** 的不同需求，本地数据中心和公有云平台间采用多种网络互联；

- 数据库服务将数据通过异步方式复制到公有云平台对象存储服务；
- 虚拟机镜像、业务数据、文件数据以全/增量的方式，从本地数据中心复

制到公有云平台对象存储服务；

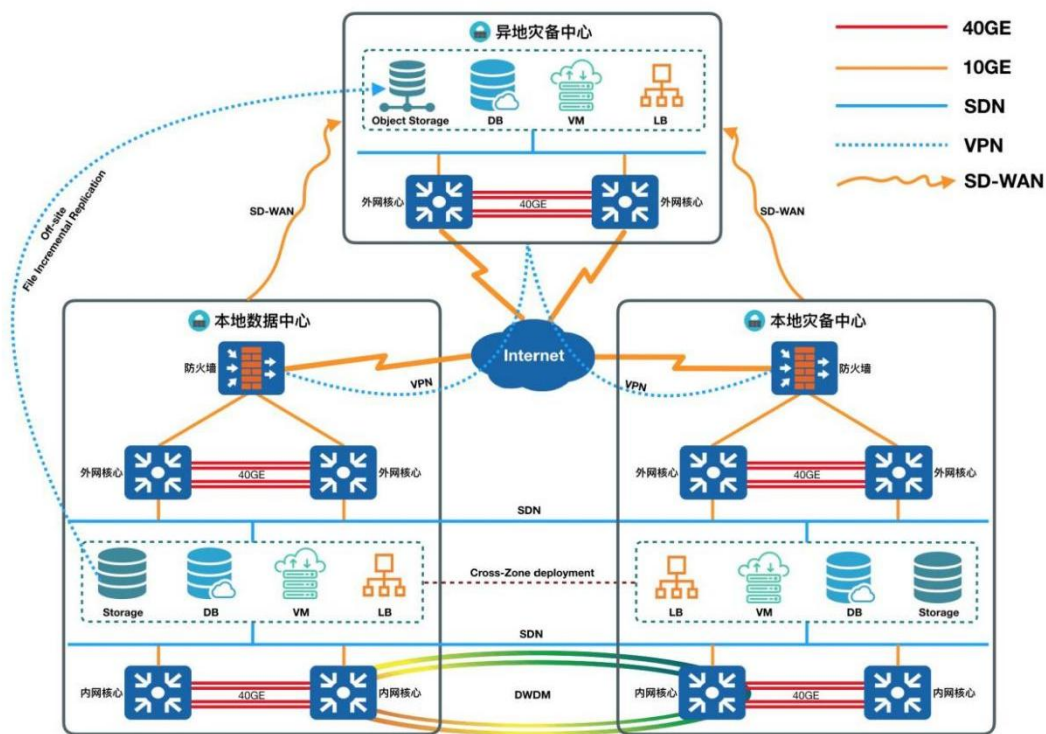
业务恢复：当本地数据中心发生灾难或需要业务切换时，修改业务域名 A 记录和数据库状态；

- 通过智能 DNS 将业务域名 A 记录手动修改为公有云平台业务 LB 的 IP 地址，实现故障切换和业务恢复；
- 将公有云平台业务应用数据库服务修改为读写状态，业务应用数据库直接读写对象存储中的数据；

异地灾备中心与公有云平台网络互联方式，会影响业务数据备份的频率和完整性；由于网络延时的影响，不建议公有云平台业务为 Active 模式。

11.4 灾备网络架构

灾备服务网络分别同城双中心网络和异地灾备网络，不同灾备服务构建方式，通过不同的网络链路进行互联互通。生产中心与异地灾备中心可通过 SD-WAN、专线、VPN、互联网等方式进行网络联通和数据复制，可根据业务对 RPO 的需求，选择不同的网络连接方案。



(1) 同城双中心

- 本地数据中心和同城灾备中心通过 DWDM 链路，将同城双中心内网核心进行物理互联，并通过三层将双中心二层网络打通，保证网络负载均衡条件，网络时延小于 2ms；
- 同城双中心分别通过 WAN 链接与互联网连通，承载同城双中心的外网接入；
- 同城双中心的负载均衡、VPC、虚拟机、数据库及存储跨可用区部署，并保证跨可用区高可用；

(2) 异地灾备中心

异地灾备中心与同城双中心通过多种方式进行网络联通，用于数据复制和数据库复制；

异地灾备网络互联方式包括 SD-WAN、专线、VPN、互联网等，可根据业务对 RPO 的需求及对于成本的考虑进行网络互联方案选择；

- SD-WAN/专线
 - 通过 SD-WAN、专线的方式将同城双中心的外网核心与异地灾备外网进行互联；
 - 线路质量好，数据复制和同步速度较快，异地灾备业务 RPO 可以得到保证；
 - RPO 指标低，成本高。
- Internet/VPN
 - 直接通过 Internet 或 VPN 的方式将同城双中心的外网和异地灾备外网互联；
 - 网络质量无法保证，数据复制和同步速度较慢，异地灾备业务 RPO 不能得到保证；
 - RPO 指标高，成本低。

11.5 灾备切换

灾备服务根据业务场景分为计划内和计划外切换；根据灾备服务方式分为同城和异地切换。

- 计划内指业务灾备演练和云平台运维，生产中心并未发生灾难或故障，多用于验证灾备服务能力；
- 计划外指生产中心发生大规模灾难，如地震、电子故障、病毒攻击等，生产中心已彻底损坏；
- 同城切换指同城双中心中的某个数据中心发生灾难的业务切换，合理部署业务可实现同城双中心自动容灾，无需用户介入切换，业务自动恢复；
- 异地切换指同城双中心均发生灾难，无法提供服务，需手工进行业务恢复和切换；
- 通常情况下，同城双中心自动进行业务灾备切换，无需人工介入，下文仅对异地灾备切换进行描述。

11.5.1 计划内切换

- 数据比对：人工介入比对同城双中心和异地灾备中心业务数据和资源数据的一致性及完整性；
- 停止生产中心业务、网络、负载均衡或关闭硬件设施电源等；
- 检查异地灾备中心业务虚拟机中业务运行状态，并检查虚拟机是否在业务负载均衡实例的后端；
- 修改灾备中心业务数据库状态为读写状态，测试通过负载均衡服务地址，访问业务服务地址的状态；
- 通过智能 DNS 将业务域名 A 记录手动修改为异地灾备中心业务 LB 的 IP 地址，测试业务服务状态；
- 若异地灾备中心仅部署对象存储服务，即仅有生产中心业务数据的备

份, 需要在异地灾备中心或生产中心准备运行业务的基础 IaaS 及 PaaS 环境, 通过备份数据逐个还原业务虚拟机、负载均衡、数据库、存储等。


11.5.2 计划外切换


- 检查并确认同城双中心均已故障或不可用;
- 检查异地灾备中心业务虚拟机中业务运行状态, 并检查虚拟机是否在业务负载均衡实例的后端;
- 修改灾备中心业务数据库状态为读写状态, 测试通过负载均衡服务地址, 访问业务服务地址的状态;
- 通过智能 DNS 将业务域名 A 记录手动修改为异地灾备中心业务 LB 的 IP 地址, 测试业务服务状态;
- 修复生产中心或在异地重新同城灾备中心, 并部署相关服务, 对数据和业务进行同步及复制;


若异地灾备中心仅部署对象存储服务, 即仅有生产中心业务数据的备份, 需要在异地灾备中心准备运行业务的基础 IaaS 及 PaaS 环境, 通过备份数据逐个还原业务虚拟机、负载均衡、数据库、存储等。


11.5.3 灾备回切

生产中心或同城双中心故障恢复或重新后, 将异地新生产中心业务切回至原生产中心, 即本地数据中心。灾备回切属于计划内, 切换流程与计划内切换一致。

 **说明** 为保证私有云所有文档的统一样，需要统一使用 WPS 软件进行文档的撰写和编制，请一定使用 WPS。

 **注意** 为保证私有云所有文档的统一样，需要统一使用 WPS 软件进行文档的撰写和编制，请一定使用 WPS。

 **警告** 为保证私有云所有文档的统一样，需要统一使用 WPS 软件进行文档的撰写和编制，请一定使用 WPS。

 **危险** 为保证私有云所有文档的统一样，需要统一使用 WPS 软件进行文档的撰写和编制，请一定使用 WPS。

以上说明、注意、警告、危险均用表格进行实现，表格框的【文字环绕】均设置为【无】，并且表格本身在文档中【居中】。

11.6 表格样式

表格整体在文档中居中（非表格单元格内的文本），表格单元格内容根据实际需要进行居中或居左。

参数	类型	说明	必填
Text	类型 A	这是说明文字	Yes
Text	类型 B	这是说明文字	Yes
Text	类型 C	这是说明文字，这是说明文字，这是说明文字，这是说明文字，这是说明文字这是说明文字这是说明文字。	Yes

- **表格文字**: 整体文字均采用 5 号宋体, 英文采用 5 号 Arial。
- **表格标题行**: 第一行标题文字加粗, 同时底纹为黑色。
- **文字行距**: 表格内文字段落采用 1 倍行距。
- **段落间距**: 段前 0.5 行的间距, 段后 0.5 行的间距。

注意 表格自身居中放置在文档中, 居中前需保证表格前无缩进 2 字符。

11.7 代码样式

在撰写技术文档过程中, 通常会进行命令行或代码的编写, 本文档中采用灰色表格作为代码块承载, 如下示例:

```
import hashlib
import urlparse
import urllib

def _verify_ac(private_key, params):
    # 请求参数串
    items=params.items()
    # 将参数串排序
    items.sort()

    # 拼接
    params_data = "";
    for key, value in items:
        params_data = params_data + str(key) + str(value)
    params_data = params_data + private_key

    # 生成的 Signature 值
    sign = hashlib.shal()
    sign.update(params_data)
    signature = sign.hexdigest()

    return signature
```

代码字体整体采用【Courier New】小五号字体, 中英文一致。字体颜色为灰色。文字环绕为【无】, 以方便代码块后续正文的格式准确性。