

阿里云

ZStack for Alibaba Cloud

运维手册

产品版本 : V2. 2. 0

文档版本 : 20171016









法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

表 1: 格式约定

格式	说明	样例
	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险：重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告：重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于警示信息、补充说明等，是用户必须了解的内容。	 注意：导出的数据中包含敏感信息，请妥善保管。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 注：您也可以通过按 Ctrl + A 选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
粗体	表示按键、菜单、页面名称等UI元素。	单击 确定。
<i>courier</i>字体	命令。	执行 <i>cd /d C:/windows</i> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<i>bae log list --instanceid Instance_ID</i>
[]或者[a b]	表示可选项，至多选择一个。	<i>ipconfig [-all -t]</i>
{ }或者{a b}	表示必选项，至多选择一个。	<i>switch</i> { <i>stand slave</i> }

目录

法律声明.....	I
通用约定.....	I
1 引言.....	1
2 概述.....	3
2.1 产品概述.....	3
2.2 产品功能.....	4
2.2.1 专有云功能.....	5
2.2.2 混合云功能.....	12
3 系统架构.....	15
3.1 专有云架构.....	15
3.1.1 ZStack功能架构.....	15
3.1.2 ZStack资源结构.....	17
3.2 混合云架构.....	20
3.2.1 身份认证.....	20
3.2.2 互连网络.....	21
3.2.3 资源管理.....	22
3.2.4 业务实现.....	23
4 网络配置.....	24
5 安装部署.....	27
5.1 环境准备.....	27
5.2 安装 ZStack for Alibaba Cloud.....	45
5.2.1 ZStack for Alibaba Cloud管理节点模式.....	47
5.2.2 ZStack for Alibaba Cloud计算节点模式.....	82
5.2.3 ZStack for Alibaba Cloud OCFS2存储节点模式.....	84
5.2.4 ZStack for Alibaba Cloud专家模式.....	86
5.3 管理ZStack for Alibaba Cloud.....	87
5.4 升级ZStack for Alibaba Cloud.....	87
5.5 裸机管理.....	90
6 登录.....	92
7 云资源池.....	93
7.1 云主机.....	93
7.2 镜像.....	100
7.3 云盘.....	103
7.4 计算规格.....	105
7.5 云盘规格.....	105
8 硬件设施.....	106
8.1 区域.....	106
8.2 集群.....	106
8.3 物理机.....	108
8.3.1 GPU设备透传.....	109
8.4 主存储.....	110
8.4.1 本地存储.....	111
8.4.2 NFS.....	112

8.4.3 Ceph.....	113
8.4.4 Fusionstor.....	113
8.4.5 Shared Mount Point.....	114
8.5 镜像服务器.....	115
9 网络.....	117
9.1 二层网络.....	117
9.2 三层网络.....	118
9.3 扁平网络.....	119
9.4 云路由网络.....	120
9.5 VXLAN.....	121
9.6 网络服务.....	122
9.6.1 安全组.....	122
9.6.2 虚拟IP.....	123
9.6.3 弹性IP.....	124
9.6.4 端口转发.....	125
9.6.5 负载均衡.....	126
9.6.6 IPsec隧道.....	126
10 vCenter.....	128
10.1 基础资源.....	128
10.2 云主机.....	129
10.3 网络.....	129
10.4 镜像.....	129
11 平台管理.....	131
11.1 性能统计.....	131
11.2 报警.....	131
11.3 计费.....	131
11.4 定时器.....	132
11.5 控制台服务.....	132
11.6 用户管理.....	133
11.7 设置.....	135
11.7.1 全局设置.....	135
11.7.1.1 基础设置.....	135
11.7.1.2 高级设置.....	138
12 管理节点高可用.....	151
12.1 管理节点高可用方案.....	151
13 基本运维.....	152
13.1 安装部署.....	152
13.2 启动服务.....	153
13.3 zstack-ctl命令.....	154
13.4 zstack-cli命令.....	155
13.5 重连物理机.....	156
13.6 重连主存储.....	157
13.7 重连镜像服务器.....	157
13.8 云主机异常.....	157
13.9 日志分析.....	164
13.10 网络脚本.....	165
14 其他配置运维.....	169
14.1 安装多个管理节点.....	169
14.2 配置独立的MySQL服务器.....	170

14.3 配置独立的RabbitMQ服务器.....	171
14.4 zstack.properties核心配置文件.....	171
14.5 ZStack计算节点密码变更.....	171
14.6 ZStack管理节点更改物理环境.....	172
14.7 ZStack数据库备份与恢复.....	172
14.8 本地存储空间扩容.....	172
15 常见问题.....	174
16 授权.....	205
术语表.....	207

1 引言

产品版本

目前与本文档相对应的产品版本为：ZStack for Alibaba Cloud V2.2.0

读者对象

本文档详述了ZStack for Alibaba Cloud V2.2.0在安装部署使用过程中需要注意的事项，以及管理运维过程中的常见问题解决方案。本文档主要适用于以下读者：

- 部署运维工程师
- 技术支持工程师
- 对ZStack有兴趣研究的相关人员

版本更新

2.2.0

2017/10/16主要更新：

1. 公有网络创建云主机
2. 自定义DHCP模式
3. 新增系统网络
4. 云主机根云盘扩容
5. 浏览器添加镜像（目前支持ImageStore类型镜像服务器）
6. 管理节点高可用：多网络配置
7. 跨Ceph存储数据迁移
8. 增强Ceph存储功能
9. 增强混合云功能
10. 增强VDI功能
11. LDAP自定义过滤规则
12. 增强裸机管理
13. 单集群支持多类型主存储（目前支持本地存储+NFS/SMP类型）
14. 更换License支持本地上传
15. 共享存储指定存储网络，增强云主机高可用
16. 多个场景新增进度条、操作助手和帮助文档，优化UI交互

17.优化部分业务逻辑

2.1.0

2017/08/14第一次正式发布。

2 概述

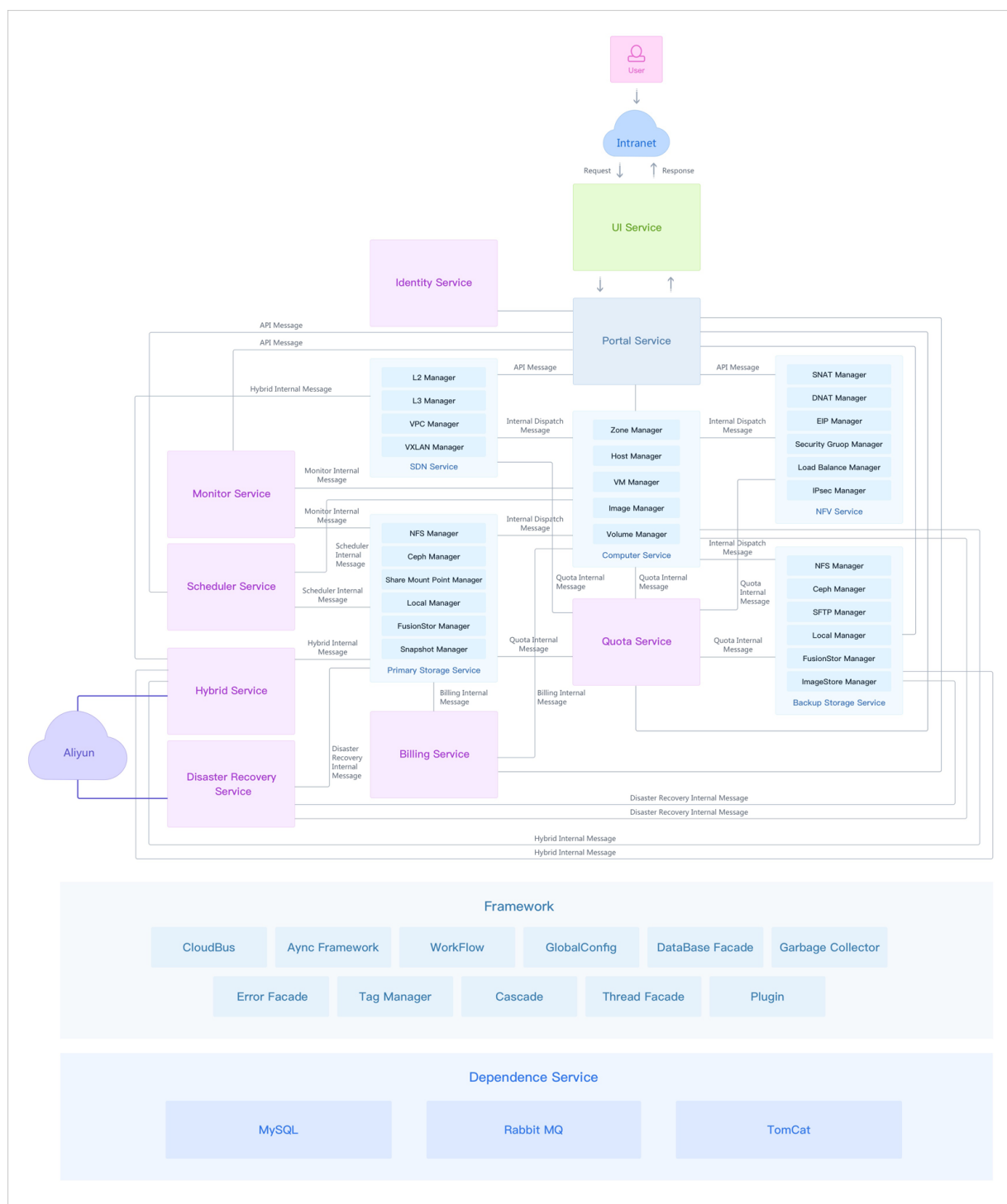
2.1 产品概述

ZStack是下一代开源的云计算IaaS（基础架构即服务）软件。它主要面向未来的智能数据中心，通过提供灵活完善的APIs来管理包括计算、存储和网络在内的数据中心资源。用户可以利用ZStack快速构建自己的智能云数据中心，也可以在稳定的ZStack之上搭建灵活的云应用场景，例如VDI（虚拟桌面基础架构）、PaaS（平台即服务）、SaaS（软件及服务）等。

通过对ZStack云引擎的深度定制，阿里云和ZStack联合推出了具有混合云功能的ZStack for Alibaba Cloud，其结合了ZStack专有云的简单、强壮、弹性、智能以及阿里云公共云的领先、安全、稳定等特点，并以"云+端"的形式提供了一套无缝集成的混合云管理方案。

系统架构如[图 1: 系统架构示意图](#)所示：

图 1: 系统架构示意图



2.2 产品功能

ZStack for Alibaba Cloud不仅拥有轻量级专有云ZStack的完整功能，还提供了一套无缝集成的混合云管理方案。

2.2.1 专有云功能

ZStack作为产品级专有云平台，提供了对用户数据中心的计算、存储、网络等资源的管理和调度。用户使用ZStack可以快速配置专有云环境，并快速创建云主机、分配云盘和自动配置云主机网络。

ZStack企业版功能列表：

类别	特性	ZStack企业版
区域	接管vCenter	支持对现有数据中心中的VMware虚拟化环境进行管理；能够查看VMware vCenter Server所管理的vSphere服务器资源和虚拟机资源；能够在虚拟数据中心中使用VMware vSphere资源；在VMware vCenter集群中完成对云主机的常用操作
	管理多个区域	用户可以根据实际情况创建并管理多个区域，一般情况下可将一个物理数据中心归为一个Zone来管理；用户根据不同的业务需求，每个Zone内建立自己独立的集群、主存储、网络等资源
集群	存储架构	集群内使用同构存储服务，存储服务挂载到集群，提供云主机高可用
	物理主机	集群内管理物理主机
	云主机	集群内管理云主机
	集群功能	提供高可用特性
	网络服务	支持VLAN、VXLAN网络加载到集群并统一管理、提供网络自助服务（IP池管理和弹性网络）
物理主机	虚拟化	支持KVM虚拟化技术，支持VMware虚拟化
	资源设定超分	支持CPU、内存和存储空间设定超分比例，适应云环境资源使用
	嵌套虚拟化	支持KVM/ESXi嵌套虚拟化，云主机内部开启CPU硬件虚拟化功能
	实时监控	采集物理主机的CPU、内存、存储和网络运行数据，提供图形可视化
	停用与启用	对物理主机设定可用属性，以便停止在该物理主机上创建云主机
	维护模式	对物理主机设定维护状态，设定维护模式后，物理主机上的云主机将会迁移（共享存储）

类别	特性	ZStack企业版
	裸机管理	<ul style="list-style-type: none"> 通过PXE技术，使管理员自动化完成对新上线物理裸机的批量部署 支持对裸机进行远程电源管理 支持VNC无人值守模式
	GPU透传	支持物理机PCI设备透传，让云主机拥有高性能计算和图形处理能力。
	USB透传	支持USB透传，满足多种USB应用场景
	操作日志	展示物理主机执行任务的事件审计
云主机	批量操作	批量创建/删除多个云主机
	云主机生命周期	支持创建、停止、启动、重启、关闭电源、删除、暂停、恢复等基本生命周期控制
	根云盘扩容	支持根云盘扩大容量，方便修改云主机配置
	云主机控制台	用户可通过终端方式访问云主机，而不依赖云主机远程工具，支持控制台设置密码
	云主机快照	<ul style="list-style-type: none"> 在云主机运行过程中进行快照 在线快照（支持ImageStore/Ceph/FusionStor类型的镜像服务器） 关机快照（支持ImageStore/Sftp/Ceph/FusionStor类型的镜像服务器）
	云盘在线快照	在使用云盘的过程中进行快照
	云主机在线修改密码	支持Windows/Linux的云主机在线修改密码
	云主机QGA开关	灵活控制qemu guest agent的状态
	云主机RDP模式开关	针对VDI用户界面，启用后默认以RDP模式打开控制台
	云主机显卡切换	支持选择云主机显卡类型：qxl、cirrus、vga
	云主机显卡透传	支持英伟达和AMD GPU设备透传给云主机
	User Data导入	支持创建云主机时导入User Data
	云主机克隆	<ul style="list-style-type: none"> 基于云主机快速克隆若干个云主机 在线克隆（支持ImageStore/Ceph类型的镜像服务器） 关机克隆（支持ImageStore/Ceph类型的镜像服务器）
	重置云主机	支持恢复云主机状态为模板初始状态

类别	特性	ZStack企业版
	根云盘扩容	支持关机状态下的云主机根云盘扩容
	基于ISO部署	基于ISO系统光盘部署云主机，引导安装系统
	基于模板部署	基于系统模板创建云主机
	制作镜像模板	基于当前某个云主机制作模板
	创建镜像	<ul style="list-style-type: none"> 云主机运行中在线创建镜像 在线创建镜像（支持ImageStore/Ceph类型的镜像服务器） 关机创建镜像（支持ImageStore/Sftp/Ceph/FusionStor类型的镜像服务器）
	云主机启动顺序	调整云主机的启动顺序，用于切换ISO引导
	动态加载、卸载云盘	云主机可动态加载和卸载云盘，支持优化驱动模型，支持SCSI WWN号唯一识别
	动态加载卸载网卡	云主机可动态加载和卸载网卡，支持设置默认网卡
	共享云盘	支持Ceph存储下多云主机共享使用同一数据云盘
	实时性能监控	采集云主机的CPU、内存、存储和网络运行数据，提供图形可视化
	高可用特性	物理主机故障，云主机自动重启
	在线修改云主机CPU/内存	支持在线修改云主机配置，不用重启VM
	实时更新云盘和网络QoS	提供云盘和网络的限速能力，避免单个云主机占用过量资源
	SSH密钥注入	支持Linux和BSD操作系统SSH密钥注入，支持创建和删除密钥
	自定义计算规格	支持自定义计算规格，满足各种应用资源消耗特性
	自定义标签	支持自定义标签，满足查询和编写定时任务
	资源删除保护	云资源删除后，将移入回收站，提供恢复和确认销毁
	冷迁移	<ul style="list-style-type: none"> 支持本地存储类型主存储上的云主机进行关机状态迁移 支持Ceph类型主存储上的云主机跨存储设备冷迁移
	在线迁移	支持所有主存储类型上的云主机进行在线迁移
	操作日志	展示云主机操作过程的事件审计

类别	特性	ZStack企业版
	Windows系统性能优化	提供Windows云主机性能优化加速
	USB重定向	支持将VDI客户端USB设备重定向至云主机
镜像管理	系统模板	支持系统模板，支持QCOW2和RAW格式，自动匹配镜像类型
	ISO镜像	支持ISO镜像，支持从ISO镜像引导云主机
	镜像上传	支持URL上传和本地浏览器上传
	镜像迁移	支持Ceph类型主存储上的云主机跨存储设备迁移
镜像仓库	镜像存放	存放镜像数据，包括ISO和系统模板
	镜像导出	支持镜像导出下载链接
	标准系统镜像	支持标准的系统，支持Windows、红帽、Ubuntu和其他开源Linux系统
	预设运行镜像	支持众多的软件运行环境，支持Windows IIS和Dot Net Framework运行环境，支持Linux Tomcat、JAVA、Apache Web、Jboss、PHP、Node JS、Golang、Python等语言和运行环境，支持数据库Oracle、MySQL、Postgres、Mongodb、Influxdb、Cassandra和Redis等数据库服务；支持广泛的应用中间件
	预设应用镜像	支持众多的应用系统，论坛BBS、社交SNS、博客Blog、微博的常用应用系统；支持phpmyadmin等运维管理应用；支持厂商提供的应用镜像
	自定义镜像	支持管理员根据标准系统镜像和预设运行镜像，定义满足自身业务系统运行环境的镜像，以增量方式保存镜像内容，并实现智能去重功能
存储管理	本地存储	支持云盘存放到物理机本地
	NFS存储	<ul style="list-style-type: none"> 支持云盘存放到NFS协议存储，物理主机共享访问 共享文件系统管理节点高可用方案 支持指定存储网络，支持存储网络和管理网络分离，增强云主机高可用
	共享挂载存储	<ul style="list-style-type: none"> 支持云盘存放到POSIX兼容的共享存储，支持iSCSI/FC存储 共享文件系统管理节点高可用方案

类别	特性	ZStack企业版
		<ul style="list-style-type: none"> 支持指定存储网络，支持存储网络和管理网络分离，增强云主机高可用
	Ceph存储	<ul style="list-style-type: none"> 支持共享云盘 超融合管理节点高可用方案 支持指定不同性能的磁盘卷创建云盘 支持云盘存放到Ceph分布式存储 支持数据冷迁移 支持指定存储网络，支持存储网络和管理网络分离，增强云主机高可用 支持创建Ceph pool，并设置显示名
	FusionStor存储	<ul style="list-style-type: none"> 支持云盘存放到FusionStor分布式存储 超融合管理节点高可用方案 支持指定存储网络，支持存储网络和管理网络分离，增强云主机高可用
	多主存储支持	支持同一集群挂载多个主存储，包括：多个本地存储、多个NFS存储、一个本地存储和一个NFS/SMP
网络管理	VLAN二层隔离	支持VLAN 802.1q作为网络隔离手段
	VXLAN网络	支持VXLAN网络，有效解决云数据中心逻辑网段不足、上层交换机MAC地址溢出等问题、支持云主机的跨地域迁移
	分布式扁平网络	支持云主机直接使用真实网络IP资源
	分布式弹性网络	支持云主机使用虚拟网络地址，与真实网络映射
	分布式DHCP服务	支持云主机自动获取分配的IP地址
	网络地址空间预留	支持预留网络地址空间，以便与物理网络混合使用
	动态和静态分配IP	支持动态分配IP地址，支持指定使用某个IP地址
	多级网络管理	支持云主机接入多个网络，构建复杂场景的业务
	虚拟IP的QoS设置	支持对虚拟IP做QoS限制，对网络服务的高效分配管理
	MTU	自定义限制网络传输数据包的大小
	VPC	支持创建VPC的全生命周期管理，包括：创建，删除，修改，网络的加载卸载，云路由网络的所有网络服务
	公有网络	<ul style="list-style-type: none"> 支持创建云主机 支持为网络服务提供虚拟IP

类别	特性	ZStack企业版
	系统网络	可作为管理网络、存储网络等使用
	云路由网络	支持基于云路由的弹性IP
		支持基于云路由的端口转发
		支持基于云路由的负载均衡
		支持基于多个EIP绑定同一个云主机网卡
		支持基于云路由的IPsec隧道服务
		支持一个云路由设备接多个公有网络
		支持配置静态路由表
		支持分布式DHCP提升服务性能
定时任务	定时对象	支持云主机、云盘的定时操作
	定时操作	可对云主机关闭/重启，云盘快照等设置定时操作
安全管理	三层安全策略	支持基于TCP/UDP端口的安全策略
	安全组统一管理	支持安全组统一管理云主机安全策略，实现组内互通，组间策略
性能分析	云主机性能统计	支持云主机CPU、内存、磁盘和网络的性能分析，排序筛选
	物理主机性能统计	支持物理主机CPU、内存、磁盘和网络的性能分析，排序筛选
实时监控	物理主机监控	支持物理主机运行实时监控，显示CPU、内存、磁盘和网络时序监控图
	云主机监控	支持云主机运行实时监控，显示CPU、内存、磁盘和网络时序监控图
	报警	<ul style="list-style-type: none"> 支持物理机、主存储、镜像服务器等失联引起的故障报警 支持云主机或物理机负载异常的邮件报警
审计	资源审计	支持ZStack所有资源的审计查询。用户能对该资源的所有操作行为审计，有效保障用户在云环境下核心数据的安全
	导出csv文件	支持云主机列表和物理机列表导出为csv表格，方便统计分析处理
账户管理	账户和用户管理	账户管理功能，分为账户和用户，其中账户是资源计量团体，用户可定义操作权限
	LDAP账户	<ul style="list-style-type: none"> 支持添加LDAP账户，并绑定ZStack普通账户

类别	特性	ZStack企业版
		<ul style="list-style-type: none"> 支持自定义清除规则
	账户云资源配额	支持自定义分配账户最大可用资源，包括云主机运行数量、CPU、内存、云盘数量、云盘总容量、镜像数量、镜像总容量和弹性IP数量
	用户组权限分配	支持用户组权限分配，统一编排用户权限
	用户操作权限分配	支持对用户进行权限分配
	云主机更改所有者	支持变更云主机所有者，指定云主机所属账户
	云盘更改所有者	支持变更云盘所有者，指定云盘所属账户
	计算规格指定分配	支持计算规格共享特性，可指定账户是否可使用
	镜像资源指定分配	支持镜像资源共享特性，可指定账户是否可使用
	云盘规格指定分配	支持云盘规格共享特性，可指定账户是否可使用
	网络资源指定分配	支持网络资源共享特性，可指定账户是否可使用
	全局配置	管理员可以直接在UI上对很多特性进行全局配置 <ul style="list-style-type: none"> 所有的全局配置都有一个默认值 更新全局配置并不需要重启管理节点
	修改admin账户密码	忘记admin账户的登录密码，可以使用zstack-ctl reset_password还原默认值
计费	自定义计费单价	支持自定义CPU、内存、系统云盘和数据云盘的计费单价，其计费单价支持秒、分、小时和天；支持删除某时段的计费设置
	基于账户计费	基于账户进行计费，统计账户各项目消费情况
	灵活计费单价	动态可调的计费单价，满足周期性促销需求
访问	TUI	支持常用运维操作，定制化OS界面
	图形界面	支持Web图形界面的访问云管理平台，账户（用户名密码方式或LDAP方式）和用户支持图形界面登录访问
	命令行	支持通过命令行方式访问云管理平台，命令行支持全功能访问，账户和用户支持命令行登录访问
	API接口	支持全功能的API交付，API支持消息总线访问和HTTP接口访问
操作助手	智能提示	对ZStack的核心操作给出智能的环境检查和操作指导

类别	特性	ZStack企业版
VDI	解决方案	<ul style="list-style-type: none"> 通过定制客户端，支持SPICE，RDP，VNC等协议，并进行了优化 支持指定VDI网络 支持USB重定向，兼容多种USB设备 支持设置独立VDI网络 支持多屏显示 支持麦克风
UI导航	快速入口	增加快速进入产品与服务的入口，并支持高亮标注
UI信息导出	列表信息csv导出	导出云主机和物理机主列表的信息，离线管理便于图表编辑
license	更换license	支持本地浏览器上传license
管理节点	管理节点高可用	<ul style="list-style-type: none"> 支持基于Ceph、FusionStor的超融合场景 支持基于NFS、SMP的共享文件系统场景 支持多网络灵活配置
安装	一键安装	一条命令，30分钟完成从裸机到云平台的安装部署
升级	无缝升级	ZStack支持低版本至高版本的无缝升级
	增量升级	支持增量升级，大幅提高升级速度
	环境升级	可以指定只升级部署环境，通过专家模式自定义安装升级

2.2.2 混合云功能

ZStack for Alibaba Cloud支持管纳阿里云的ECS和VPC服务，统一的管理平台让用户操作阿里云的资源如同操作本地资源一样稳定快捷。

目前对于阿里云的管控界面包含如下功能：

类别	特性	ZStack for Alibaba Cloud
数据中心	阿里云地域管理	<ul style="list-style-type: none"> 查看阿里云地域列表 支持地域的添加和删除；以及地域下资源同步 阿里云地域特性： <p>一般情况下，建议选择与目标用户所在地域最为接近的数据中心，以进一步提升用户访问速度</p>

类别	特性	ZStack for Alibaba Cloud
		在基础设施、BGP网络品质、服务质量、云服务器操作使用与配置等方面，阿里云国内地域数据中心无明显差异。国内BGP网络可以保证全国地域的快速访问
	可用区管理	<ul style="list-style-type: none"> 查看阿里云可用区列表 支持可用区的添加和删除，以及可用区资源的一键同步 阿里云可用区特性： <p>同一可用区内的ECS实例网络延时更小；</p> <p>同一地域内的可用区之间内网互通，且可用区之间故障隔离；</p> <p>是否将ECS实例放在同一可用区内，主要取决于对容灾能力和网络延时的要求</p>
ECS	ECS生命周期管理	包括创建（支持批量创建）、启动、停止、重启、同步、删除ECS云主机，以及支持修改ECS云主机名称和简介、显示付费方式、修改系统用户密码
	ECS云主机控制台	通过ZStack管理界面即可打开ECS云主机控制台，以及设置控制台密码
	安全组、EIP管理	包括安全组和安全组规则的创建、远程同步、查看、阿里云端删除、本地删除；以及EIP的创建、同步、查看、加载到ECS、从ECS卸载及删除
	ECS镜像管理	支持镜像的删除、同步；支持本地镜像上传为ECS自定义镜像，以及同步阿里云系统镜像，支持查看上传进度
	ECS数据云盘管理	支持数据云盘的创建、删除、同步；支持云主机加载/卸载数据云盘；以及修改云盘名称和简介、显示付费方式
网络	VPC管理	<ul style="list-style-type: none"> 支持VPC的创建、同步、查看、阿里云端删除以及本地删除 支持虚拟交换机的创建、同步、查看、阿里云端删除以及本地删除 支持VPC内虚拟路由器的同步、查看以及路由条目的创建、同步、查看、阿里云端删除、本地删除
	高速通道	<ul style="list-style-type: none"> 支持快速建立高速通道，配置双边路由 支持边界路由器的同步、查看 支持路由器接口的同步、查看

类别	特性	ZStack for Alibaba Cloud
	VPN	<ul style="list-style-type: none"> 支持VPN网关的同步、查看、本地删除 支持VPN用户网关的创建、同步、查看、阿里云端删除、本地删除 支持VPN连接管理： <ul style="list-style-type: none"> VPN连接的创建、同步、查看、修改、阿里云删除、本地删除 IPsec配置的创建、查看、删除 Ike配置的创建、查看、删除 快速建立VPN连接
其它	密钥管理	支持AccessKey (包括AccessKey ID以及AccessKey Secret) 在本地的添加、删除、查看以及默认设置；支持多个AccessKey的添加
	支持对象存储 OSS	包括OSS bucket的添加、同步、查看、阿里云端删除、本地删除
	时区配置	支持配置时区以便部署到海外不同站点

3 系统架构

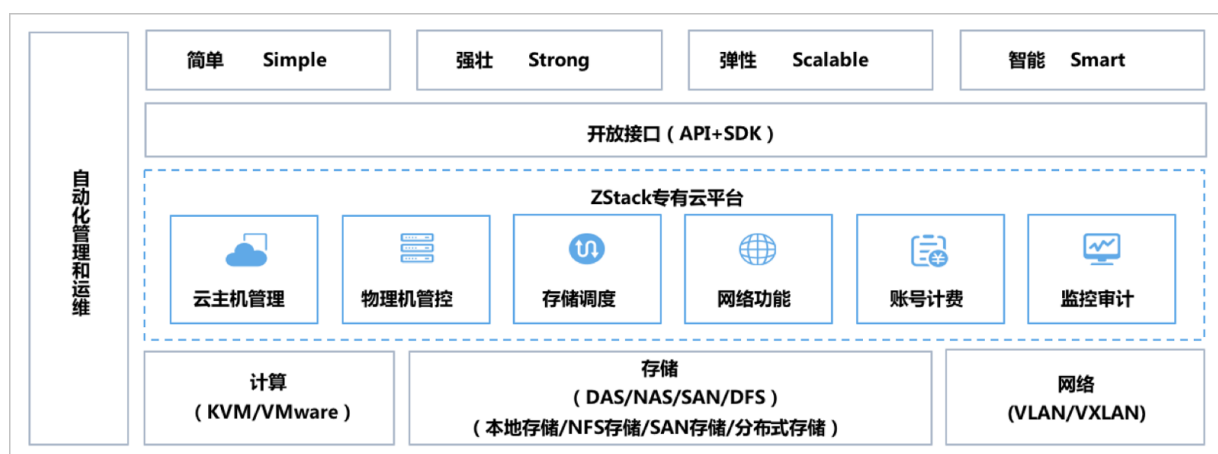
3.1 专有云架构

ZStack作为新一代轻量级专有云管理平台，通过提供灵活完善的APIs来管理包括计算，存储和网络在内的数据中心各种资源。

3.1.1 ZStack功能架构

ZStack功能架构如图 2: ZStack功能架构所示：

图 2: ZStack功能架构



ZStack提供了对企业数据中心基础设施的计算、存储、网络等资源的管理，底层支持KVM和VMware虚拟化技术，支持DAS/NAS/SAN/DFS等存储类型，支持本地存储、NFS存储、SAN存储、分布式块存储，支持VLAN/VXLAN等网络模型。

ZStack的核心云引擎，使用消息总线RabbitMQ同数据库MariaDB及各服务模块进行通信，提供了云主机管理、物理机管控、存储调度、网络功能、账号计费、实时监控等功能。ZStack还提供了Java和Python的SDK，且支持Restful APIs进行资源调度管理。基于ZStack打造的专有云管理平台充分体现专有云的4S优势，即：简单Simple、强壮Strong、弹性Scalable、智能Smart。

ZStack核心架构设计特点：

1. 全异步架构：异步消息、异步方法、异步HTTP调用。

- ZStack使用消息总线RabbitMQ进行各服务的通信连接，在调用服务时，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，就会触发回调函数回复任务结果。异步消息可以并行处理。

- ZStack服务之间采用异步消息进行通信，对于服务内部，一系列相关组件或插件，也是通过异步方法来调用，调用方法与异步消息一致。
- ZStack采用的插件机制，给每个插件设置相应的代理程序。ZStack为每个请求设置了回调URL在HTTP的包头，任务结束后，代理程序会发送应答给调用者的URL。
- 基于异步消息、异步方法、异步HTTP调用这三种方式，ZStack构建了一个分层架构，保证了所有组件均能实现异步操作。
- 基于全异步架构机制，单管理节点的ZStack每秒可并发处理上万条API请求，还可同时管理上万台服务器和数十万台云主机。

2. 无状态服务：单次请求不依赖其他请求。

- ZStack的计算节点代理、存储代理、网络服务、控制台代理服务、配置服务等，均不依赖其他请求，一次请求可包含所有信息，相关节点无须维护存储任何信息。
- ZStack使用一致性哈希环对管理节点、计算节点或者其他资源以UUID为唯一ID进行认证的哈希环处理，消息发送者无需知道待处理消息的服务实例，服务也无须维护、交换相关的资源信息，服务只需单纯的处理消息即可。
- ZStack管理节点间共享的信息非常少，两个管理节点即可满足高可用性和可扩展性需求。
- 无状态服务机制让系统更为健壮，重启服务器不会丢失任何状态信息，数据中心的弹性扩展和伸缩性维护更为简单。

3. 无锁架构：一致性哈希算法。

- 一致性哈希算法保证了同一资源的所有消息均被同一个服务实例来处理。这种聚合消息到特定节点的方法，降低了同步与并行的复杂度。
- ZStack使用工作队列来避免竞争锁的问题，串行任务以工作队列的方式保存在内存中，工作队列可对任意资源的任意操作进行并行处理来提高系统并行度。
- ZStack基于队列的无锁架构，使得任务可以简单地控制并行度，从而提升系统性能。

4. 进程内微服务：微服务解耦。

- ZStack使用消息总线对各服务进行隔离控制，例如，云主机服务、身份认证服务、快照服务、云盘服务、网络服务、存储服务。所有的微服务都集合在管理节点的进程内，各服务之间利用消息总线进行交互，所有消息发送到消息总线后，再通过一致性哈希环选择目的服务进行转发处理。
- 进程内微服务，以星状架构实现各服务独立运行，将高度集中的控制业务进行解耦，实现了系统的高度自治和高度隔离，任何服务出现故障并不影响其他组件。可靠性与稳定性得到有效保障。

5. 全插件结构：插件支持横向扩展。

- ZStack使用中任何新加入的插件对目前其他的插件没有任何影响，均是独立自主提供服务。
- ZStack支持策略模式和观察者模式进行插件设计。策略插件会继承父类的接口然后执行具体实现；观察者插件，会注册listener进行监控内部的业务逻辑的事件变化，当应用内部发现事件时，插件会对此事件做出自响应，在插件自身的代码里执行相应的业务流。
- ZStack支持插件的横向扩展，云平台可以快速更迭，而整体系统架构依然健壮。

6. 工作流引擎：顺序管理，出错回滚。

- ZStack工作流基于XML对每个工作流程进行清晰定义，在任何步骤出现错误均可按照原本执行路径进行回滚，清理掉执行过程的垃圾资源。
- 每个工作流还可以包含子工作流用于扩展业务逻辑。

7. 标签系统：支持业务逻辑变更，增加资源属性。

- ZStack支持利用系统标签和插件机制对原本的业务逻辑进行扩展变更。
- 使用标签机制，可对资源进行分组划分，支持对指定标签进行资源搜索。

8. 瀑布流架构：支持资源的级联操作。

- ZStack使用Cascade Framework对资源管理进行瀑布状的级联操作，对资源进行卸载或者删除时，会对相关的资源进行级联操作。
- 资源也可以通过插件形式加入到瀑布框架中，加入或者退出瀑布框架，并不影响其他资源。
- 级联机制使得ZStack的配置灵活轻便，快速满足客户资源配置的变更。

9. 全自动化Ansible部署：Ansible无代理自动部署。

- ZStack使用Ansible进行无代理的全自动化安装依赖、配置物理资源，部署代理程序，全过程对用户透明，无须额外干预，可透过重连代理程序对代理进行升级。

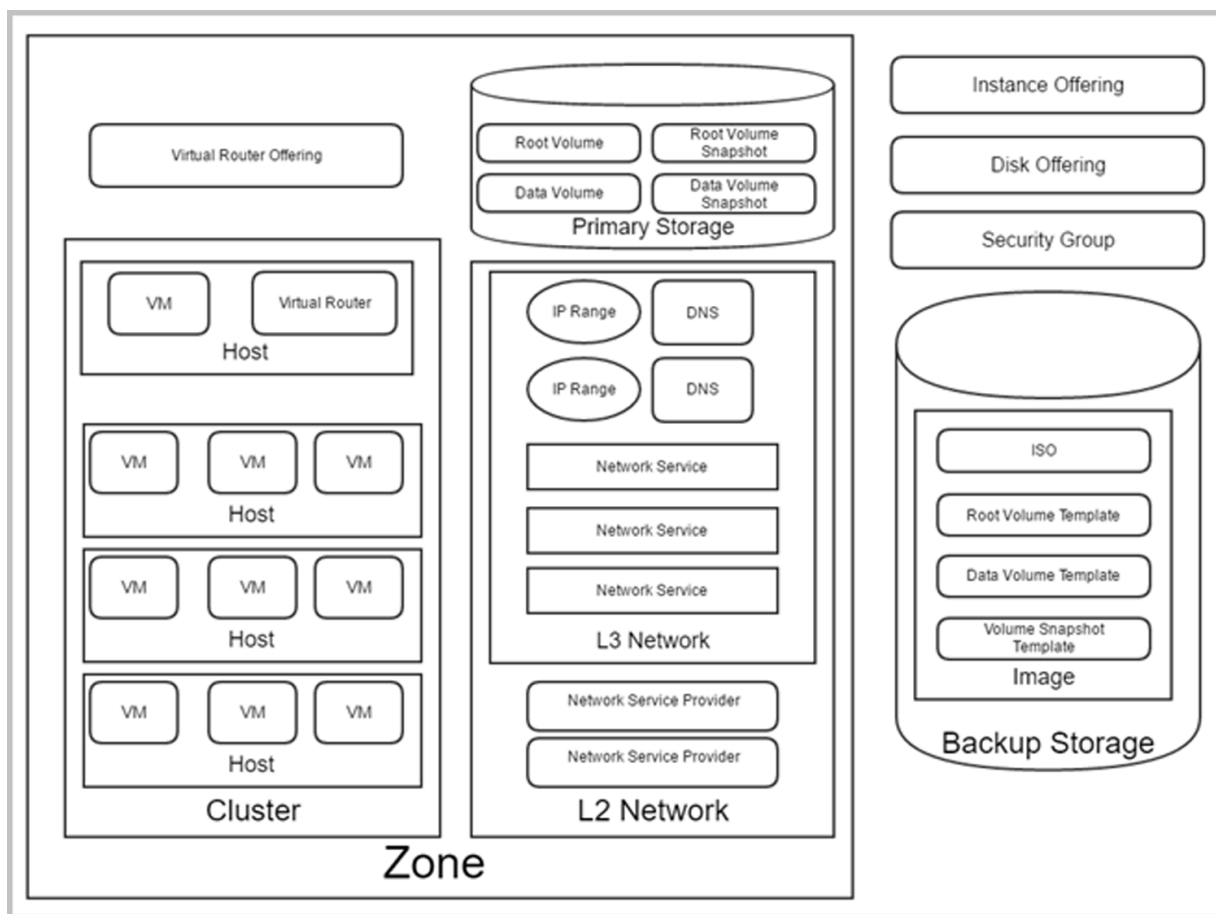
10. 全API查询：任意资源的任意属性均可查询。

- ZStack支持数百万个条件的资源查询，支持全API查询，支持任意组合。

3.1.2 ZStack资源结构

ZStack在本质上是云资源的配置管理系统。ZStack管理的相关资源在结构上如[图 3: ZStack资源结构](#)所示：

图 3: ZStack资源结构



ZStack主要包括以下资源：

- **区域：**ZStack中最大的一个资源定义，包括集群、二层网络、主存储等资源。
- **集群：**集群是一群计算节点（物理机）的逻辑组合。
- **物理机：**为运行云主机实例提供计算、网络、存储资源的物理主机。
- **主存储：**用于存储云主机磁盘文件的存储服务器。支持本地存储、NFS、Ceph、Fusionstor、Shared Mount Point等类型。
- **镜像服务器：**备份存储服务器，主要用于保存云主机的镜像模板文件。
- **二层网络：**二层广播域，进行二层相关的隔离，支持 L2NoVlanNetwork、L2VlanNetwork、VxlanNetworkPool、VxlanNetwork类型。
- **三层网络：**云主机使用的网络配置，包含了IP地址范围、网关、DNS、网络服务等。
- **计算规格：**云主机的CPU、内存、磁盘带宽、网络带宽的数量或大小规格定义。
- **云盘规格：**云主机使用的云盘的大小规格定义。
- **云主机：**直接交付给用户使用的虚拟主机系统，是ZStack的核心组成部分。
- **镜像：**云主机所使用的镜像模板文件，包括ISO和Image类型。
- **根云盘：**云主机的系统云盘，用于支撑云主机的系统运行。

- 数据云盘：云主机使用的数据云盘，一般用于扩展的存储使用。
- 网络服务模块：用于提供网络服务的模块。在UI界面已隐藏。
- 网络服务：给云主机提供的各种网络服务，主要包括安全组、虚拟IP、弹性IP、端口转发、负载均衡、IPsec隧道等。
- 快照：采用增量机制对云盘在特定时间点上的数据进行备份。
- 安全组：给云主机提供三层网络防火墙控制。
- 云路由规格：指定云路由设备使用的CPU、内存、云路由镜像、管理网络、公有网络等资源定义。
- 云路由设备：为云主机提供DHCP、DNS、SNAT、弹性IP、端口转发、负载均衡、IPsec、安全组等各种网络服务的定制云主机。

ZStack资源间存在以下关系：

- 父子关系：例如集群和物理机，物理机和云主机。
- 兄弟关系：例如集群和二层网络，集群和主存储。
- 祖先和后裔关系：例如集群和云主机，区域和物理机。
- 朋友关系：例如主存储和镜像服务器，区域和镜像服务器。



注：主存储和镜像服务器的关系为：

- 创建VM时，主存储会从镜像服务器下载复制云主机的镜像模板文件作为缓存。
- 创建镜像时，主存储会将根云盘拷贝到备份存储进行保存为模板。

ZStack资源均含有以下基本属性：

- UUID：通用唯一识别码UUIDv4 (Universally Unique Identifier) 来唯一标识一个资源。
- 名称：用于标记资源的可读字符串，名称可以重复，一般为必选项。
- 描述：也称之为简介，用于概述资源，可选项。
- 创建日期：资源创建的日期。
- 上次操作日期：资源上次被更新的时间。

ZStack资源一般都支持CRUD操作：

- 创建：创建或者添加新的资源。
- 查询：读取查询资源信息。
- 更新：更新资源信息。

- 删除：删除资源，ZStack使用的瀑布框架级联机制，使得父资源被删除后，相关子资源和后裔资源均会被删除。

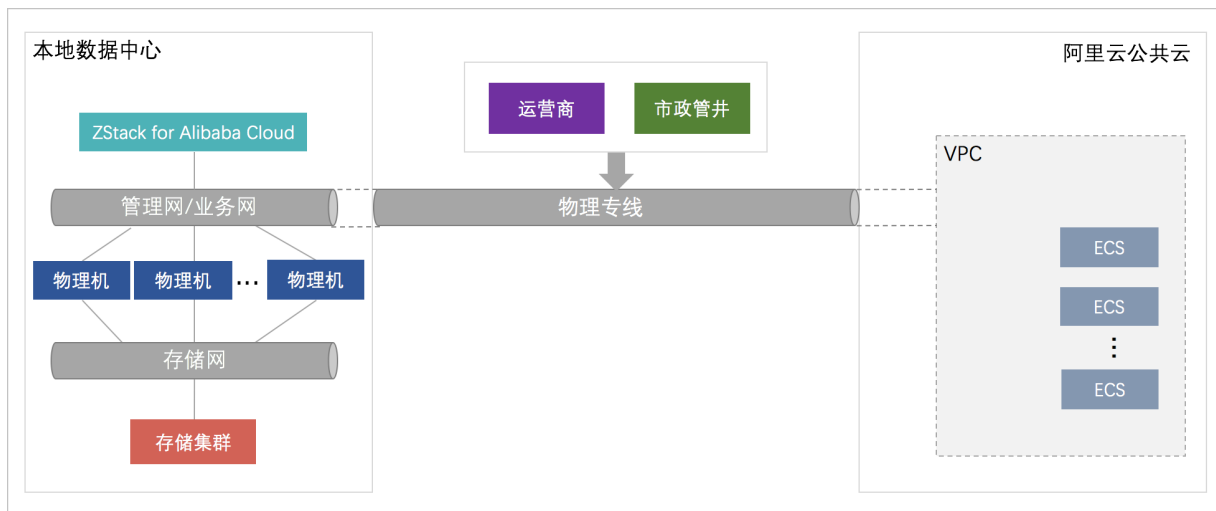
3.2 混合云架构

基于ZStack云引擎深度定制的ZStack for Alibaba Cloud，提供了一套无缝集成的混合云管理方案。

由于ZStack采用进程内微服务架构，因此ZStack for Alibaba Cloud的部署与ZStack完全一样，并不引入新的模块。但ZStack管理节点要求能够访问公网，以便调用阿里云公共云的OpenAPI。

如图 4: 混合云物理部署架构所示，通过物理专线构建"本地—远程"互连网络，从而连通本地数据中心和阿里云公共云。

图 4: 混合云物理部署架构



ZStack for Alibaba Cloud混合云功能模块主要有：身份认证、互连网络、资源管理和业务实现。

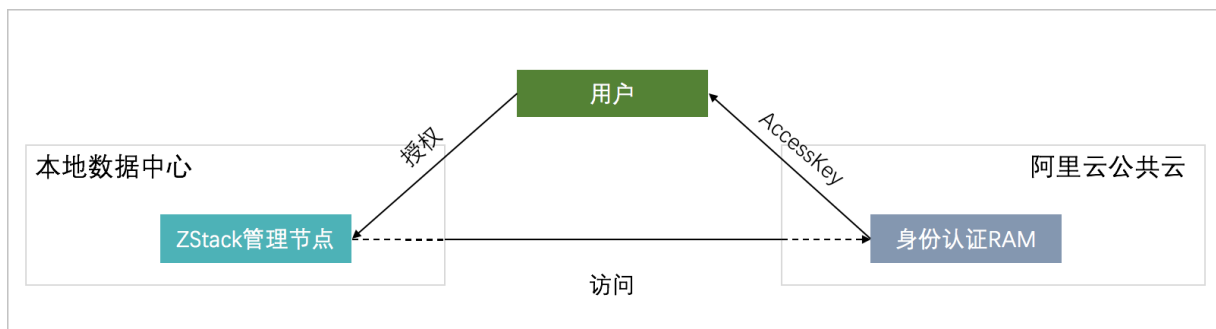
3.2.1 身份认证

ZStack for Alibaba Cloud实现了阿里云公共云的账户身份认证RAM对接，采用授权子账户AK（AccessKey以及KeySecret）信息远程访问。

- 企业资产管理（拥有全局的权限管理），可创建面向混合云平台的子账户，例如hybrid_cloud，并授予一定的资源访问权限，包括ECS、VPC、虚拟交换机和OSS等。
- 企业资产管理将该子账户的AK信息提供给信息技术部门，则可导入ZStack for Alibaba Cloud混合云平台，授予管理阿里云公共云的资源权限。

- 若企业资产管理员需终止或回收该子账户，可登陆到阿里云RAM身份认证系统，执行禁用或删除操作而无需协调信息技术部门。

图 5: 混合云身份认证功能模块图



3.2.2 互连网络

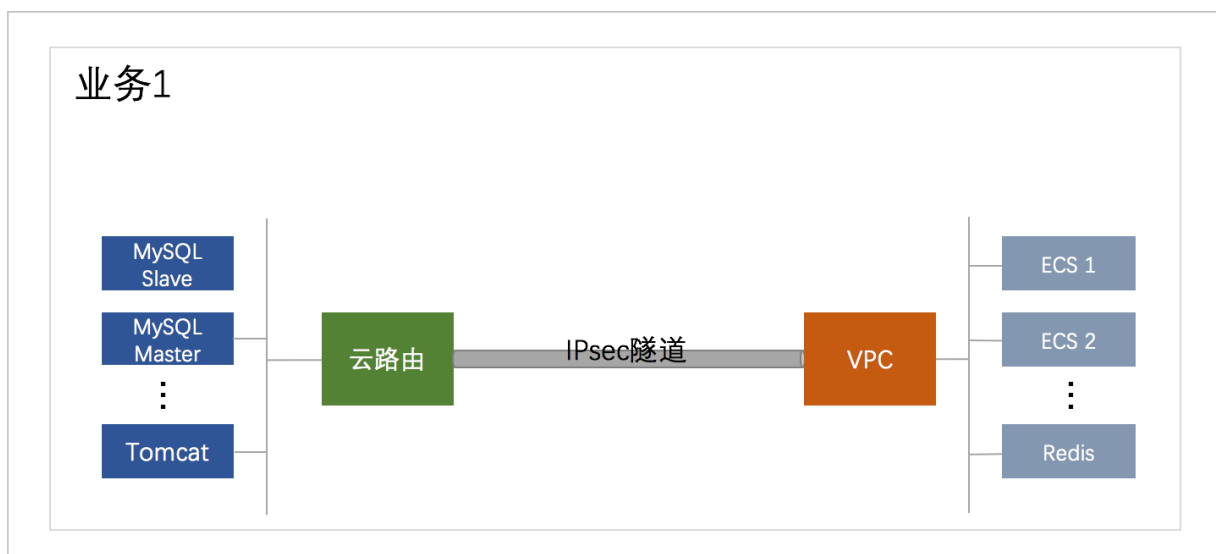
实现IPsec隧道和高速通道（Express Connect）连接本地专有云和阿里云公共云，使得“本地—远程”在三层网络可达下互访。“本地—远程”的互连网络，是混合云核心基础设施。

ZStack for Alibaba Cloud混合云平台支持IPsec隧道和高速通道构建互连网络。

IPsec隧道：

IPsec隧道实现了本地端的云路由与远程端的阿里云公共云的VPC模块互联，并以数据加密的方式传送数据。由于互联网网络延迟较大、带宽成本递增，一般情况下常用于演示测试与数据备份场景。如图 6: IPsec隧道所示：

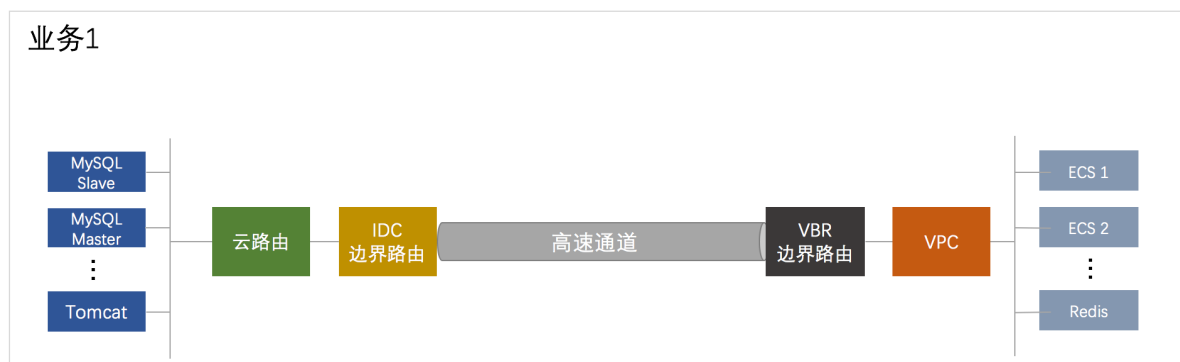
图 6: IPsec隧道



高速通道：

高速通道则是一款便捷高效的网络服务，能提供“本地—远程”网络环境间的高速、稳定、安全的私网通信，有效提高网络拓扑的灵活性和跨网络通信的质量和安全性。混合云平台能基于高速通道实现快速的数据传输，保障所承载的业务架构稳定运行。如图 7: 高速通道所示：

图 7: 高速通道

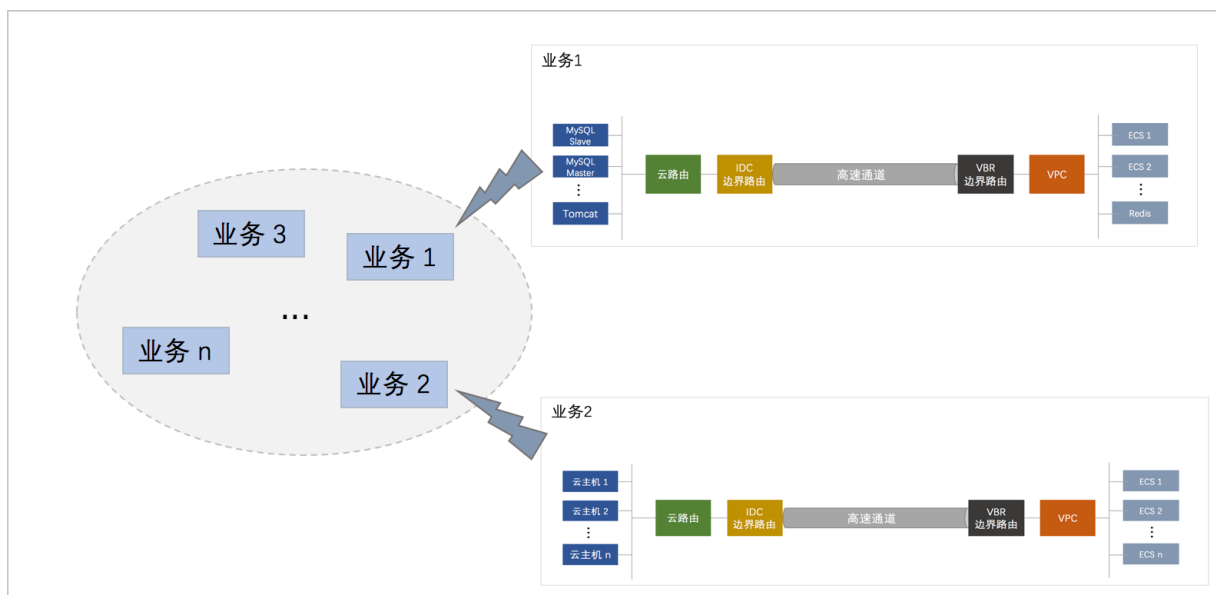


3.2.3 资源管理

通过授权子账户，访问阿里云公共云账户里的资源，包括管理ECS、OSS、VBR、VPC和虚拟交换机等服务。

- 基于上述的身份认证模块，管理员无需再登陆访问阿里云公共云门户控制台，直接通过ZStack for Alibaba Cloud混合云平台本地登陆后，就可以访问并管理公共云端的资源。目前，混合云平台支持管理的阿里云公共云资源包括：ECS、OSS、VBR、VPC和虚拟交换机等服务。
- ZStack for Alibaba Cloud混合云平台提供直观简捷的图形界面（UI）和全功能开放接口（API），以便信息系统管理员可见即可得操作，以及基于统一API实现上层的业务调度逻辑。

图 8: 混合云资源管理功能模块



3.2.4 业务实现

基于上述的网络基础设施和管理控制方式，实现灵活弹性的业务系统架构。

混合云平台建成后，可在其上部署灵活多维的业务模式，场景列举如下：

- 应用与数据分离：

数据库和分析服务保留在专有云，缓存服务和Web应用部署在公共云（在混合云管理面板直接部署），从而既满足数据库关键数据自有，也满足业务灵活扩展和互联网访问的安全性。

- 数据归档备份：

企业可以选择对长久归档的数据（切片和加密）传送到阿里云对象存储（OSS），实现低成本的数据灾备方案，满足成本和随时访问恢复的需求。

- 弹性业务架构：

企业可以将持续稳定的业务负载量部署在专有云环境，将业务瞬时或周期高峰负载分摊到公共云，利用公共云的海量计算资源（相对专有云）满足资源申请与及时释放。

4 网络配置

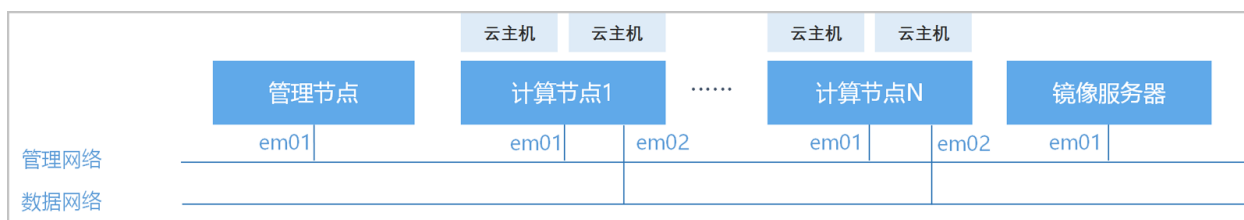
ZStack支持复杂的网络场景，根据业务需求和实际环境，下面列出了几种常见的网络架构：

1. 本地存储+扁平网络架构

其架构如图 9: 本地存储 + 扁平网络所示，此架构支持以下特性：

- 物理机使用本地存储来存储云主机的系统盘和数据盘。
- 云主机数据保护需要依赖物理机的RAID阵列技术。
- 物理机和云主机在同一个二层网络，云主机IP可与物理机IP互通。
- 支持分布式DHCP、分布式弹性IP、Userdata服务。
- 云主机在线迁移受限。
- 成本低廉，并发性能高，适用于软件开发测试环境、应用层可提供高可用的场景。

图 9: 本地存储 + 扁平网络



管理网络：

- 负责管理各计算节点及存储节点资源的网络。
- 图示的em01即为管理节点使用的网卡，是管理节点用于和计算节点及镜像服务器通信的网络。

数据网络：

- 给云主机使用的网络。
- 图示的em02即为创建云主机使用的网络，云主机的网络如果设置与物理机同一网段，就可以与物理机互通，IP互通时，注意IP段不可重叠。
- 可以设置Vlan或VXLAN进行二层隔离，网络隔离后，云主机只能在私网内使用，无法直接访问外部网络，需要使用弹性IP机制来实现网络互通。

镜像服务器：

- 镜像服务器可以使用单独的服务器提供服务。
- 在小型的数据中心，也可与管理节点使用同一台机器。

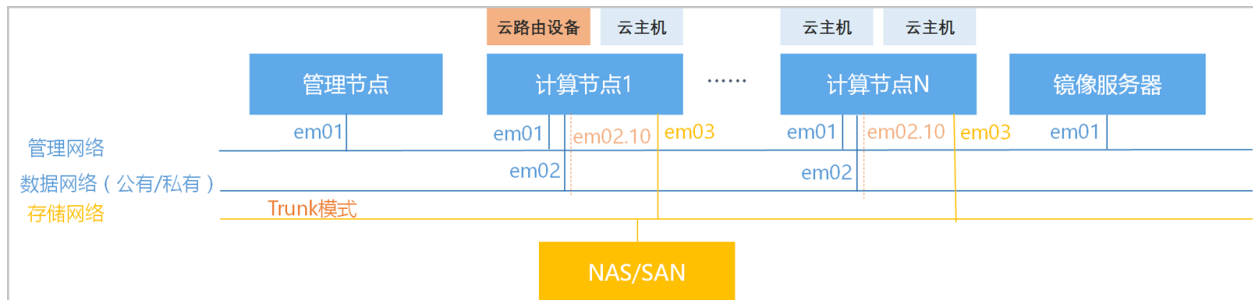
2. NAS/SAN + 云路由网络架构

其架构如图 10: NAS/SAN + 云路由网络所示，此架构支持以下特性：

- 云主机数据存放在集中存储，云盘数据通过存储双控和RAID技术提供保护。
- 需要准备存储网络，准备集中存储，建议万兆以上网络，云主机受高可用保护。
- 添加到ZStack时，可添加为NFS或Shared Mount Point主存储。
- 构建了存储网，云主机支持高可用，数据支持安全保护，IO性能有瓶颈，建议8~10计算节点共用一个存储。
- 使用云路由设备提供DNS、SNAT、弹性IP、端口转发、负载均衡、IPsec、安全组等网络服务。
- 支持分布式DHCP网络。

- 需要一定的投资，适用于传统应用、应用数据紧密、SLA高可用保证、数据安全场景。

图 10: NAS/SAN +云路由网络



管理网络：

- 负责管理各计算节点及存储节点资源的网络。
- 图示的em01即为管理节点使用的网卡，是管理节点用于和计算节点及镜像服务器通信的网络。

数据网络：

- 同一个物理接线通过Vlan设置为公有网络和私有网络。
- 其中公有网络为不带Vlan的em02，可以访问互联网，私有网络为Vlan编号为10的em02，交换机端需设置trunk模式，云主机可以通过云路由设备访问互联网，云路由网络使用云路由设备来提供各种网络服务，云路由设备的网络默认IP为公有网络的IP地址。



注：

如果物理机网卡较少，管理网络和数据网络可以合并为同一个网络。

存储网络：

- 用于提供NAS/SAN存储的网络。
- 建议使用万兆以太网或FC，加速存储访问。图示使用的是em03来专门提供存储网络流量的访问。

镜像服务器：

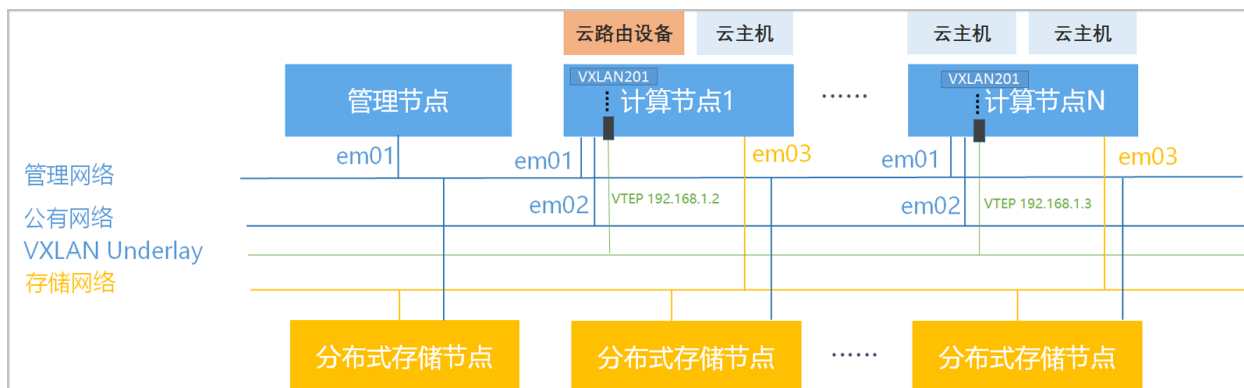
- 镜像服务器可以使用单独的服务器提供服务。
- 在小型的数据中心，也可与管理节点使用同一台机器。

3. Ceph+VXLAN云路由网络架构

其架构如图 11: Ceph + VXLAN云路由网络所示，此架构支持以下特性：

- 云主机数据存放在分布式存储，数据通过多副本和纠删码机制提供保护。
- 需要准备存储网络，云主机受高可用保护。
- Ceph主存储和Ceph镜像服务器使用同一套分布式存储。
- 构建了存储网，云主机支持高可用，数据支持多副本，计算和存储都可以横向扩展。
- 使用VXLAN的VTEP提供二层逻辑子网，VXLAN Underlay网络供云主机访问。
- 使用云路由设备提供DNS、SNAT、弹性IP、端口转发、负载均衡、IPsec、安全组等网络服务。
- 支持分布式DHCP网络。
- 适用于租用平台、通用型应用、应用数据紧密、SLA高可用保证、大规模环境。

图 11: Ceph + VXLAN云路由网络



管理网络：

- 负责管理各计算节点及存储节点资源的网络。
- 图示的em01即为管理节点使用的网卡，是与计算节点和分布式存储通信的网络。

公有网络：

- 可以访问外部互联网的网络。
- 云主机可以通过云路由设备访问互联网，云路由网络使用云路由设备来提供各种网络服务，云路由设备的网络默认IP为公有网络的IP地址。
- 图示使用的em02来提供公有网络服务。如果网卡不足，也可将管理网络和公有网络合并为一个网络使用。

VXLAN Underlay网络：

- 给云主机使用的VXLAN Underlay私有网络。
- 基于三层IP地址的二层逻辑网络，例如图示采用的是192.168.1.x/24的VTEP来提供VNI为201的私有网络。

存储网络：

- 用于提供分布式块存储的网络。
- 用于提供分布式块存储的网络，建议使用万兆以太网或FC，加速存储访问。
- 例如图示的em03的存储网络。Ceph分布式存储至少需要有一个Mon IP可与管理网络互通。

5 安装部署

本章主要介绍ZStack for Alibaba Cloud V2.2.0的安装/升级过程。

5.1 环境准备

前提条件

安装/升级ZStack for Alibaba Cloud都必须使用ZStack for Alibaba Cloud定制版ISO，特性如下：

1. 基于CentOS 7.2深度定制，友好的TUI管理界面，支持多种系统配置；
2. 安装ZStack for Alibaba Cloud无需连接外网，也无须配置yum源，就可以实现完全离线安装；
3. 提供四种安装模式：管理节点模式、计算节点模式、存储节点模式、专家模式，用户按需选择即可；
4. 采用系统默认的网卡命名规则；
5. 默认选项：**DATE&TIME**为亚洲东八区，**LANGUAGE**为English(United States)，**KEYBOARD**为English(US)。

操作步骤

1. 安装ZStack for Alibaba Cloud之前，请管理员准备好以下必要的软件包，以便安装部署过程顺利执行。
 - ZStack for Alibaba Cloud定制版ISO
 - 文件名称：ZStack_Alibaba_Cloud-x86_64-DVD-2.2.x.iso
 - 下载地址：点击[这里](#)
 - ZStack for Alibaba Cloud安装包
 - 文件名称：ZStack_Alibaba_Cloud-installer-2.2.x.bin
 - 下载地址：点击[这里](#)



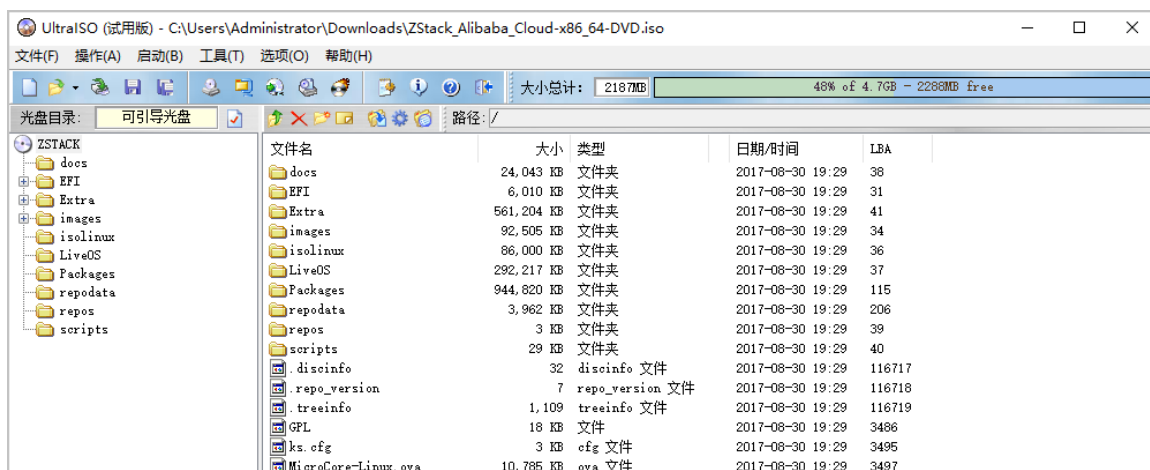
注：软件下载后，需通过MD5校验工具核对校验码，以确保软件完整无损。

2. 使用UltraISO，将此ISO镜像刻录到U盘。

a) 在UltraISO打开ISO镜像。

打开UltraISO，点击文件按钮，选择打开已下载好的ISO镜像文件，如[图 12: 在UltraISO打开ISO镜像](#)所示：

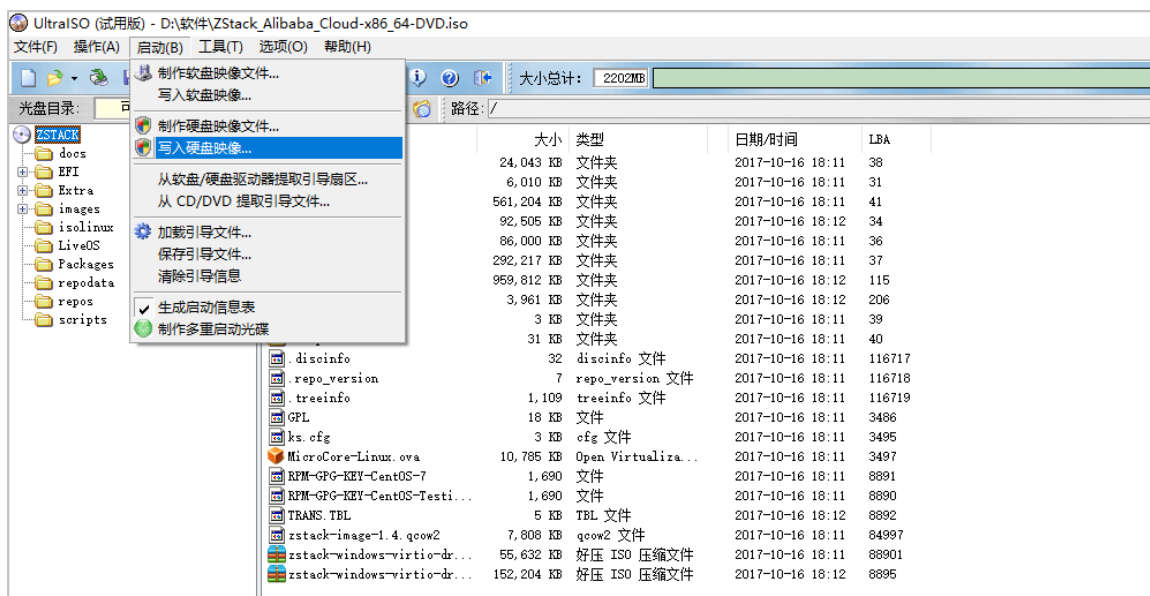
图 12: 在UltraISO打开ISO镜像



b) 写入硬盘镜像。

在UltraISO，点击启动 > 写入硬盘映像，如图 13: 在UltraISO写入硬盘映像所示：

图 13: 在UltraISO写入硬盘映像



c) 在硬盘驱动器列表选择相应的U盘进行刻录。



注:

- 如果系统只插了一个U盘，则默认以此U盘进行刻录和写入，在刻录前，注意备份U盘之前的内容。
- 其他选项，按照默认设置，无须额外配置，点击写入。

如图 14: 在UltraISO确认写入ISO镜像所示：

图 14: 在UltraISO确认写入ISO镜像



d) 在新界面中点击是进行确认，UltraISO将会把ISO镜像刻录到U盘。

e) 此时U盘可用来作为启动盘，支持Legacy模式和UEFI模式引导。

3. 安装操作系统。

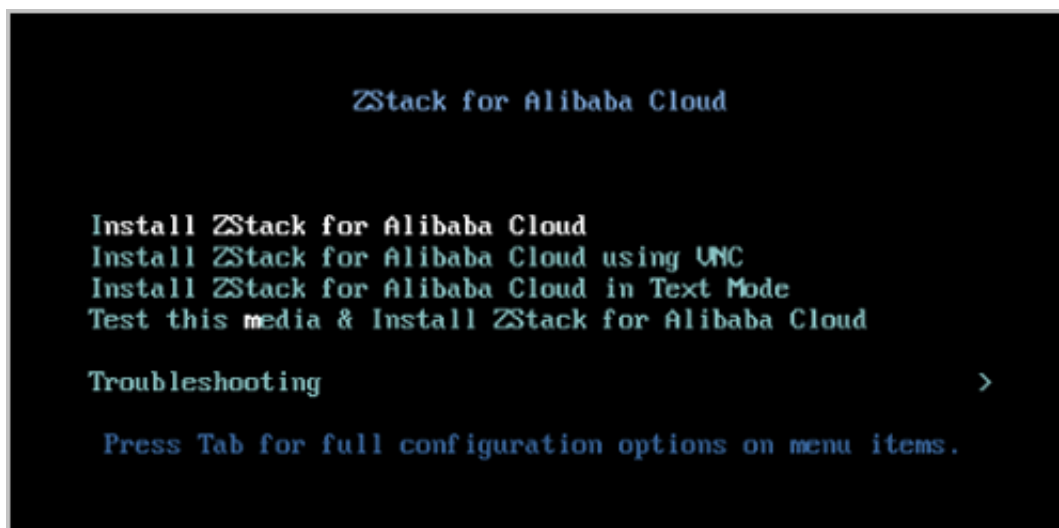
a) 管理员需要预先在服务器进行以下配置：

- 确认服务器内硬盘的数据已作备份，安装过程会覆盖写入；
- 进入BIOS，开启CPU VT选项；开启超线程HT选项；
- 进入阵列卡配置合适的RAID级别，以提供一定的数据冗余特性；
- 设置U盘为第一启动顺序。

b) 以上设置完毕后，服务器重启或上电后，进入安装导航。

如图 15: U盘引导界面所示，进入ISO引导安装界面，默认选择Install ZStack for Alibaba Cloud开始安装操作系统。

图 15: U盘引导界面



注:

- ZStack for Alibaba Cloud提供了三种安装方式：图形界面安装、通过VNC安装和字符界面安装，用户可根据实际情况选择。
- 建议图形界面安装是最好的选择。
- 考虑到某些服务器是不带VGA接口的，只能通过串口连接，这时用户可以选择VNC或者Text Mode。

c) 进入系统安装界面后，已经预先配置如下默认选项，一般情况下管理员无需更改配置。

d) 选择安装模式。

在系统安装界面，点击**SOFTWARE SELECTION**进入服务器安装模式候选，如图 16: 选择安装模式所示：



注:

- 有四种安装模式可供选择：
 1. ZStack for Alibaba Cloud Management Node : ZStack for Alibaba Cloud管理节点模式
 2. ZStack for Alibaba Cloud Compute Node : ZStack for Alibaba Cloud计算节点模式

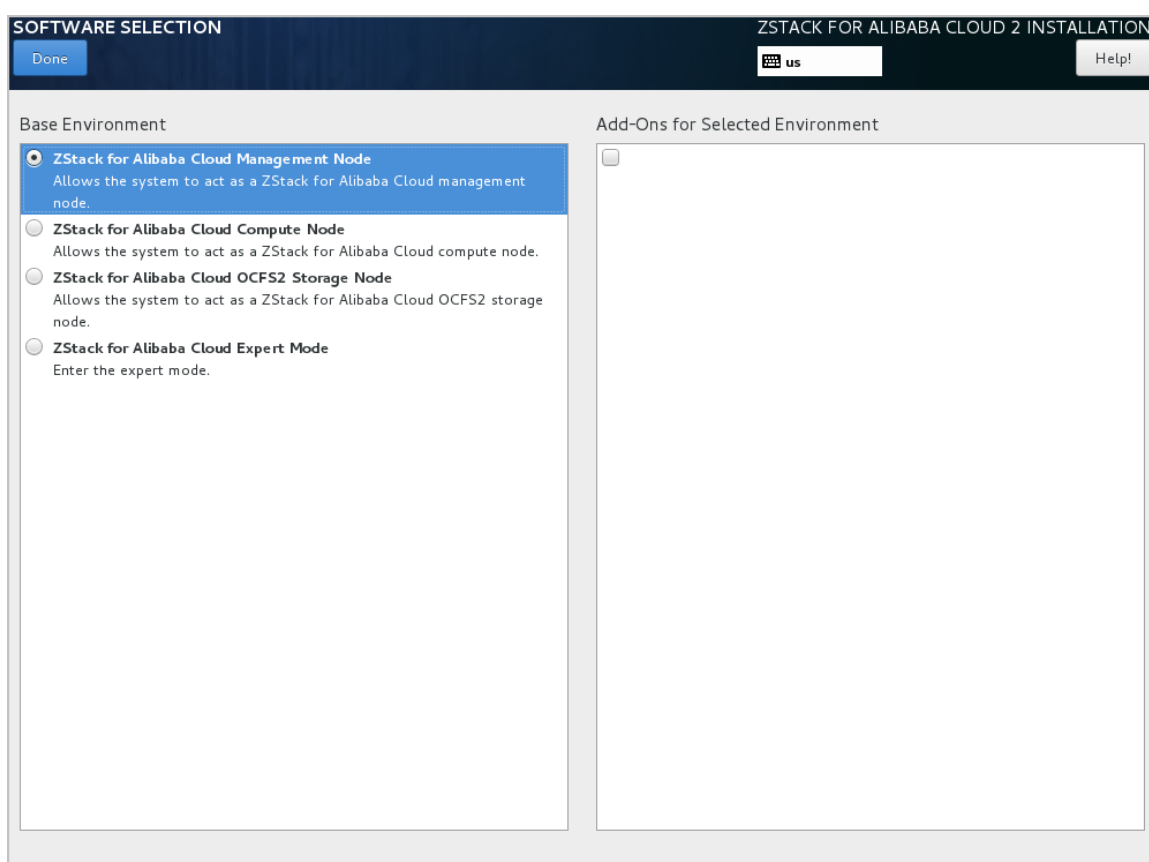
3. ZStack for Alibaba Cloud OCFS2 Storage Node : ZStack for Alibaba Cloud OCFS2存储节点模式

4. ZStack for Alibaba Cloud Expert Node : ZStack for Alibaba Cloud专家模式

关于这四种模式的详细介绍将在下文展开。

- 在四种模式中，首次安装建议选择**ZStack for Alibaba Cloud Management Node**。

图 16: 选择安装模式



e) 安装模式选择完后，配置硬盘分区。

在系统安装界面，点击**INSTALLATION DESTINATION**进入硬盘分区配置界面，如[图 17: 系统预先默认设置#自动硬盘分区](#)所示：

图 17: 系统预先默认设置：自动硬盘分区

INSTALLATION DESTINATION ZSTACK FOR ALIBABA CLOUD 2 INSTALLATION

Done us Help!

Device Selection
Select the device(s) you'd like to install to. They will be left untouched until you click on the main menu's "Begin Installation" button.

Local Standard Disks

500 GiB

Virtio Block Device
vda / 992.5 KiB free

Disks left unselected here will not be touched.

Specialized & Network Disks

Add a disk...

Disks left unselected here will not be touched.

Other Storage Options

Partitioning
☒ Automatically configure partitioning. ☐ I will configure partitioning.
☐ I would like to make additional space available.

Encryption
☐ Encrypt my data. You'll set a passphrase next.

[Full disk summary and boot loader...](#) 1 disk selected; 500 GiB capacity; 992.5 KiB free

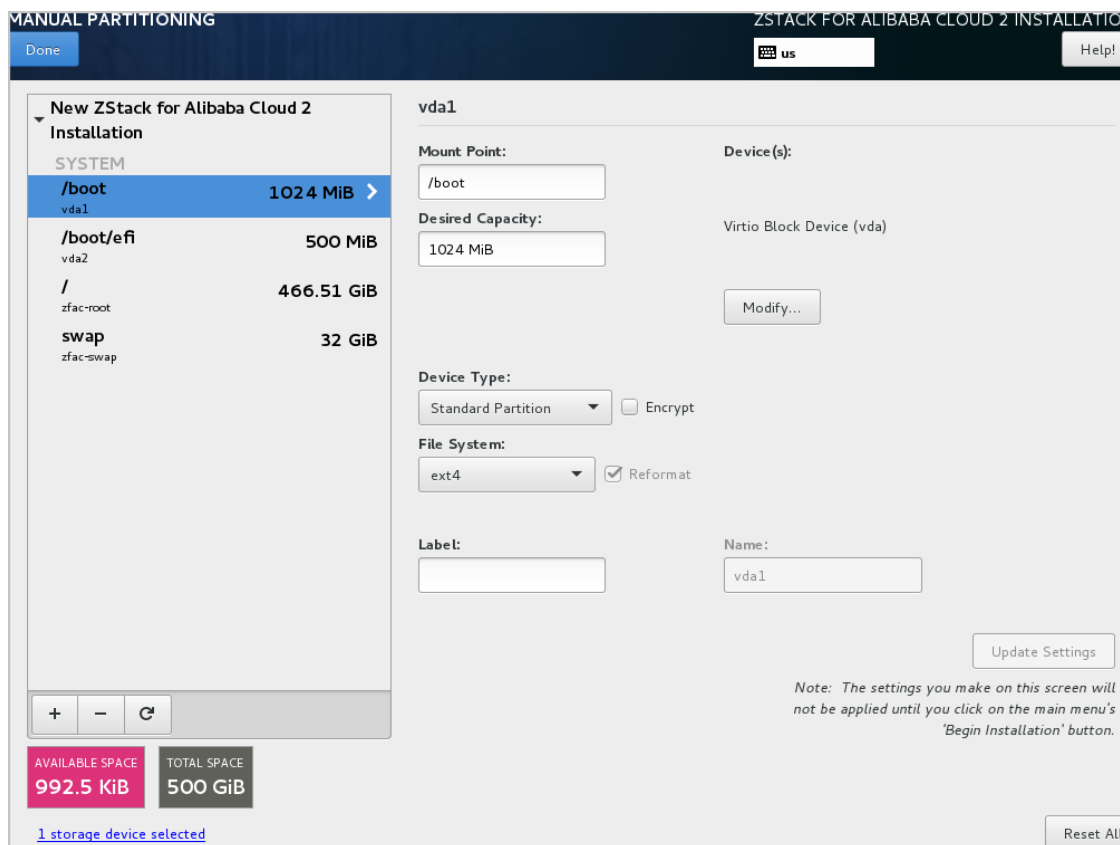
系统预先默认设置：**Automatically configure partitioning**，建议遵循默认设置，执行自动硬盘分区。

如果选择自定义手动分区，建议如下：

- 分区模式有UEFI 模式和Legacy模式两种，应与BIOS设置的引导模式一致。
 - UEFI 模式
 - `/boot`：创建分区 1GB
 - `/boot/efi`：创建分区 500MB
 - `swap`（交换分区）：创建分区 32GB
 - `/`（根分区）：配置剩下容量
 - Legacy模式
 - `/boot`：创建分区 1GB
 - `swap`（交换分区）：创建分区 32GB
 - `/`（根分区）：配置剩下容量
- 以上数值为建议分区容量（硬盘总容量在300G以上）

- Legacy模式不支持单盘容量大于2T，而UEFI模式没有此限制，且还支持GPT分区，因此推荐采用UEFI模式来分区，如图 18: 推荐UEFI模式分区所示：

图 18: 推荐UEFI模式分区



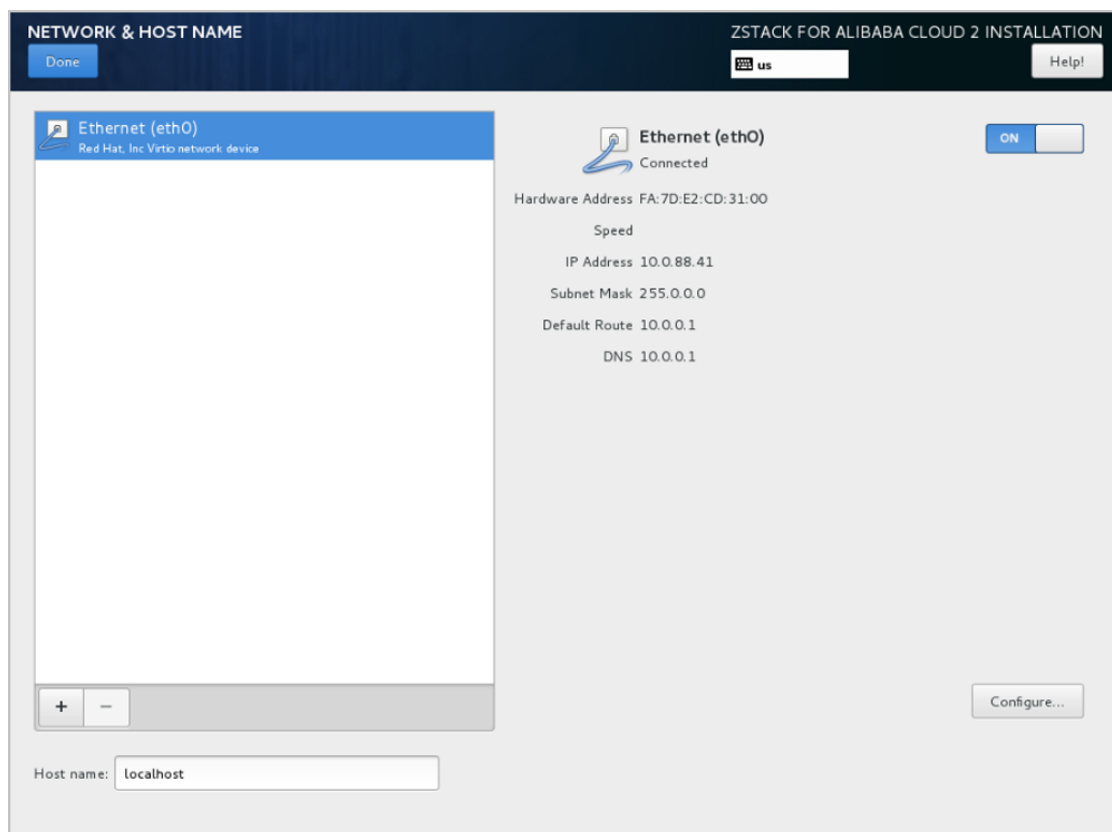
f) 硬盘分区配置完后，进入网络配置。

1. 配置网卡。

在系统安装界面，点击**NETWORK & HOST NAME**进入网卡配置主界面，如图 19: 网卡配置主界面所示。

1. 选中待配置网卡：如eth0
2. 开启网卡：选择On
3. 查看获取的DHCP地址

图 19: 网卡配置主界面

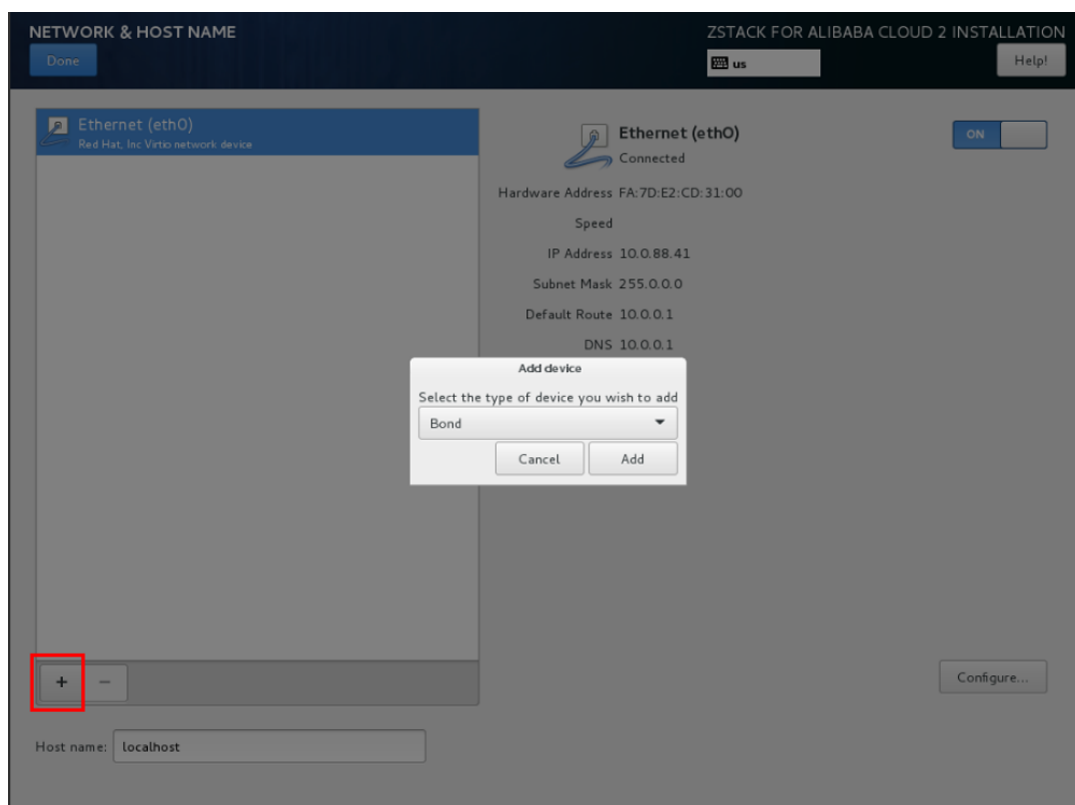


2. 网卡归一化。

a. 添加一个Bond设备。

点击左下角的"+"号，弹出**Add device**界面，在下拉菜单中选择**Bond**，点击**Add**，如图 20: 添加一个Bond设备所示：

图 20: 添加一个Bond设备



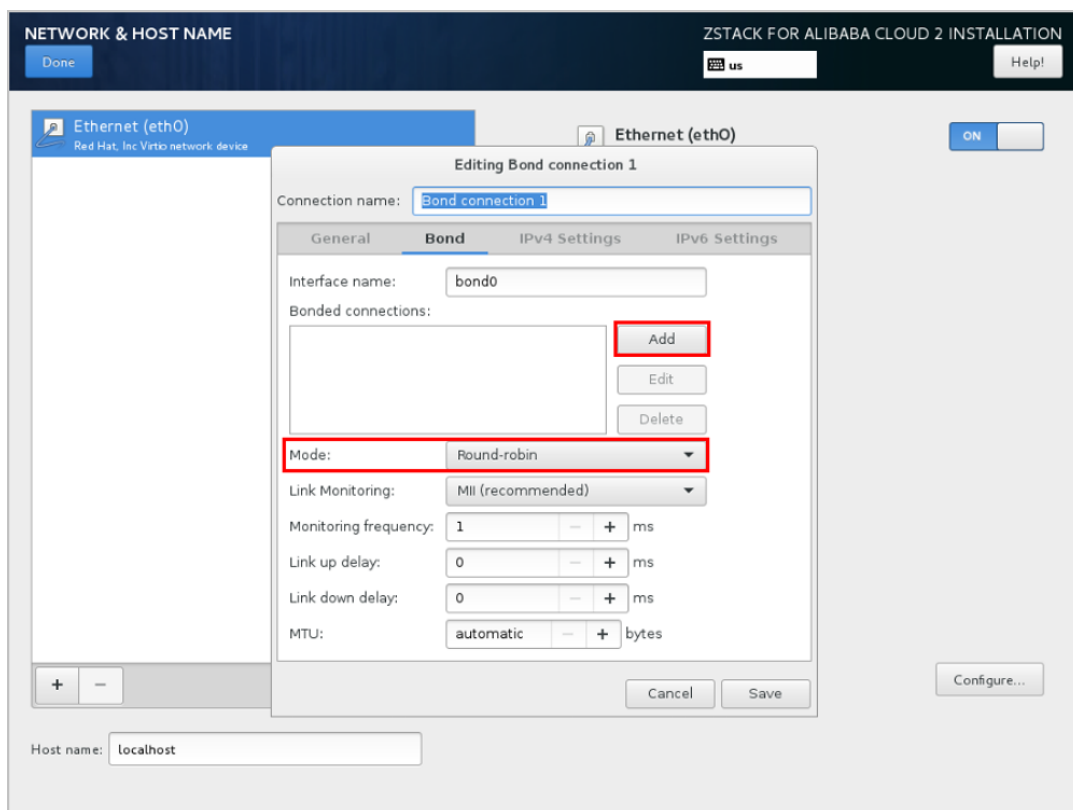
b. 配置Bond。

弹出**Editing Bond connection 1**界面的Bond子页面，如图 21: 配置Bond所示，用户需手动配置的主要有两项：

- **Add**：添加Bond Slave，详见[步骤3](#)
- **Mode**：选择Bond模式，详见[步骤4](#)

其它可选择默认或按需自定义设置。

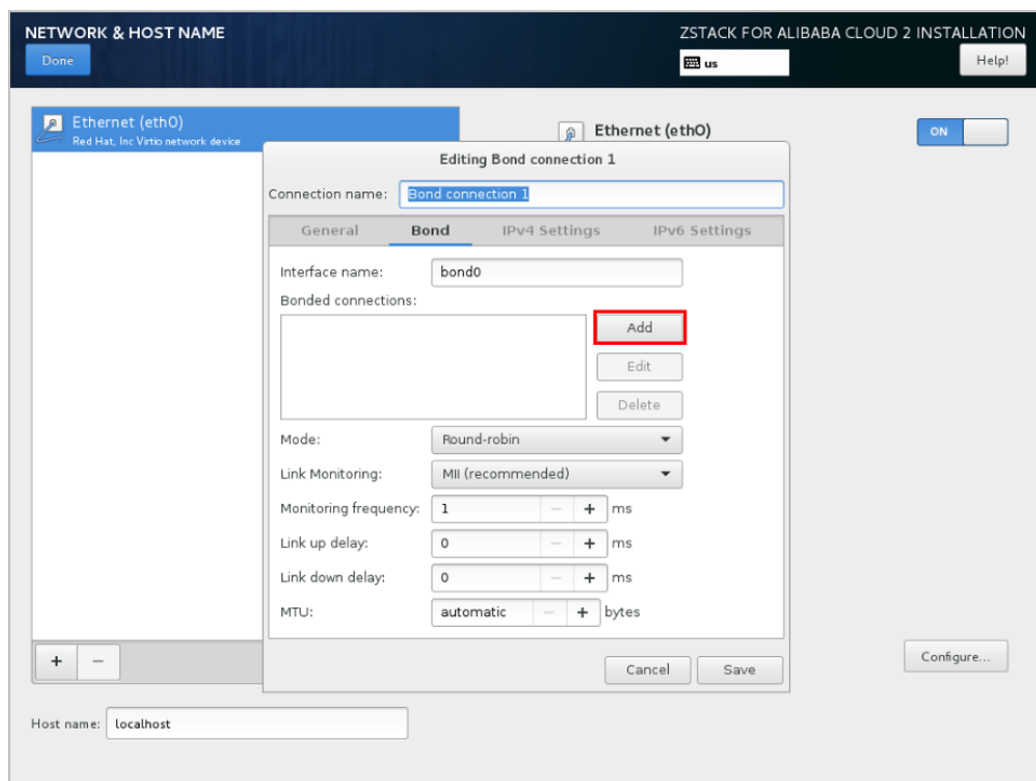
图 21: 配置Bond



c. 添加Bond Slave。

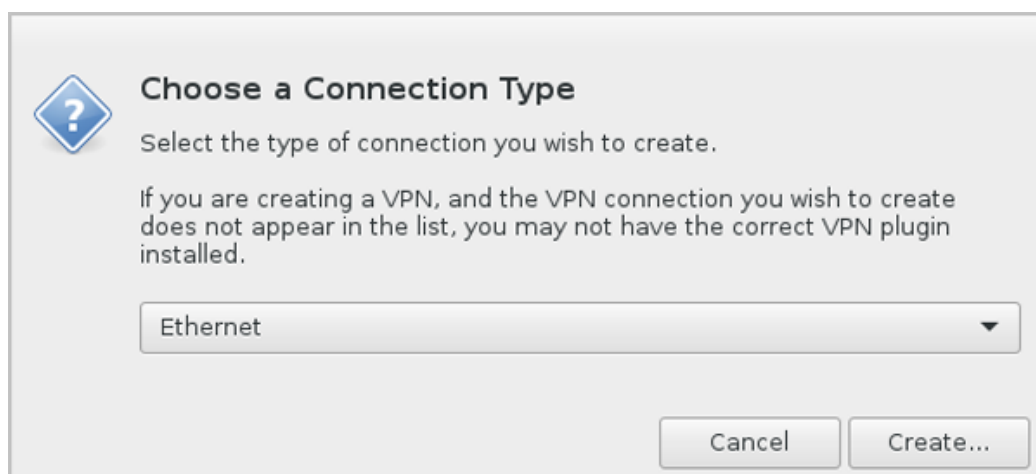
- 在Bond配置界面，点击**Add**，添加Bond Slave，如图 22: 添加Bond Slave所示：

图 22: 添加Bond Slave



- 弹出Choose a Connection Type界面，在下拉菜单中选择Bond Slave连接类型，如Ethernet，点击Create...，如图 23: 选择Bond Slave连接类型所示：

图 23: 选择Bond Slave连接类型



- 弹出Editing bond0 slave1界面的Ethernet子页面，在Device下拉菜单中选择需要Bond的Slave设备，如eth0（相应MAC地址），其它选择默认或按需自定义设置，点击Save，如图 24: 选择Bond Slave设备所示：

图 24: 选择Bond Slave设备

Editing bond0 slave 1

Connection name:

General **Ethernet** 802.1x Security DCB

Device:

Cloned MAC address:

MTU: bytes

- 至此，Bond Slave已成功添加。

d. 选择Bond模式。

在Bond配置界面，**Mode**下拉菜单中，按需选择Bond模式，如**Active backup**（主备模式），其它选择默认或按需自定义设置，点击**Save**，如[图 25: 选择Bond模式](#)所示：

图 25: 选择Bond模式

Editing Bond connection 1

Connection name:

General Bond IPv4 Settings IPv6 Settings

Interface name:

Bonded connections:

bond0 slave 1	<input type="button" value="Add"/>
	<input type="button" value="Edit"/>
	<input type="button" value="Delete"/>

Mode: ▼

Primary:

Link Monitoring: ▼

Monitoring frequency: ms

Link up delay: ms

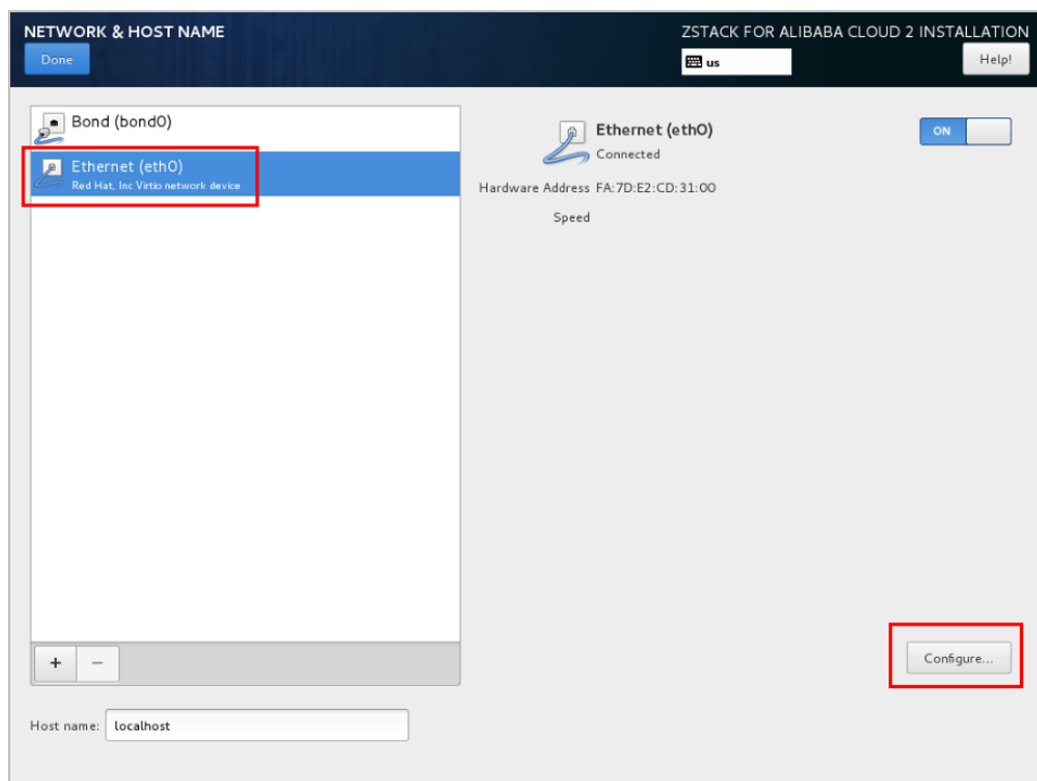
Link down delay: ms

MTU: bytes

e. Bond Slave的IPv4设置为禁用。

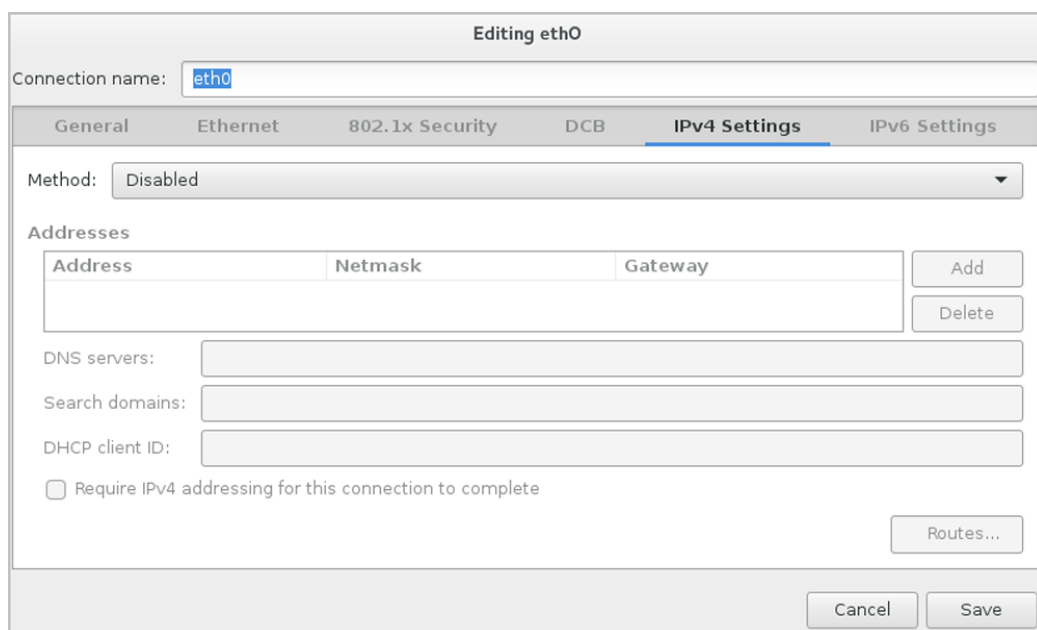
- 回到网卡配置主界面，选中Bond Slave（如eth0），点击**Configure...**，如图 26: 打开Bond Slave配置界面所示：

图 26: 打开Bond Slave配置界面



- 进入Editing eth0界面的IPv4 Settings子页面，Method下拉菜单中，选择Disabled，点击Save，如图 27: Bond Slave的IPv4设置为禁用所示：

图 27: Bond Slave的IPv4设置为禁用



f. 网卡归一化完成。

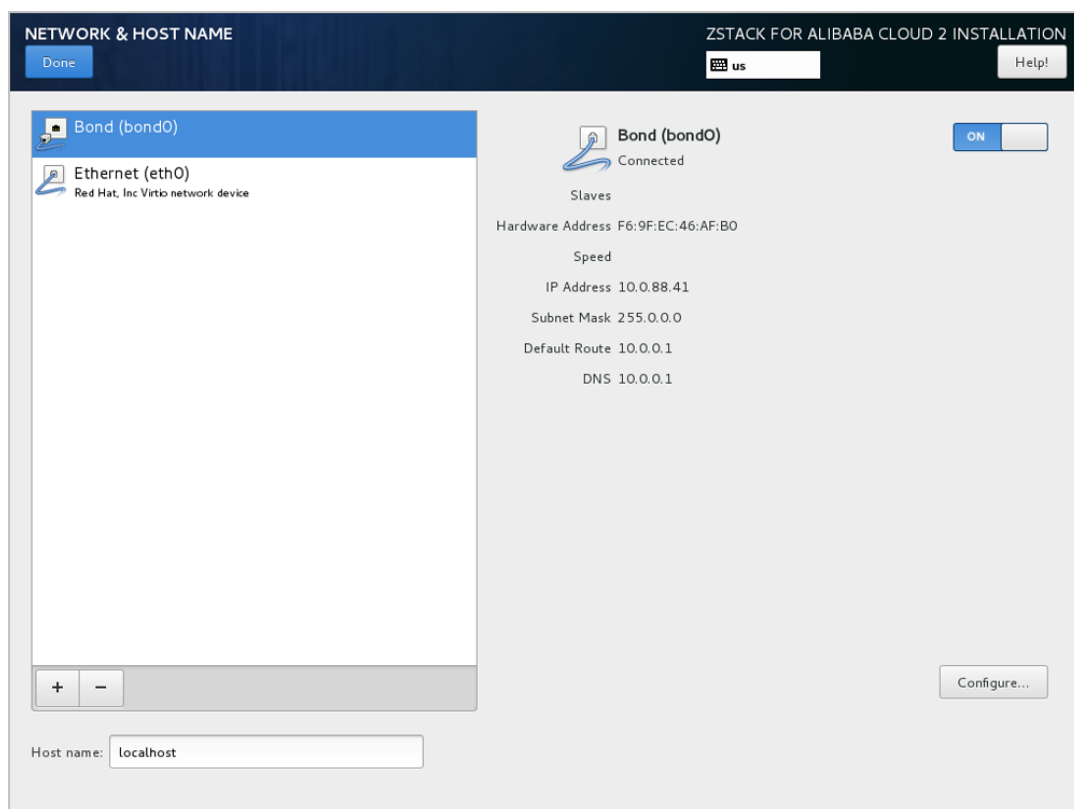
如图 28: 网卡归一化完成所示：



注:

请检查Bond配置项，必须保证**On**开启且配置了地址，同时Bond Slave（如eth0）也**On**开启，否则ZStack for Alibaba Cloud无法正常安装。

图 28: 网卡归一化完成



3. 如果Bond无法获取DHCP地址，需手动配置Bond的静态地址，按如图 29: 配置Bond静态IP所示操作。

- 在图 19: 网卡配置主界面，选中Bond，点击**Configure...**，打开Bond配置界面。
- 进入Bond的**IPv4 Settings**选项页。
- 在**Method**列表选择**Manual**以进行手动配置。
- 点击**Add**增加新的配置条目。
- 根据实际情况配置网卡地址信息。
- 点击**Save**保存。

图 29: 配置Bond静态IP

Editing Bond connection 1

Connection name:

General Bond IPv4 Settings IPv6 Settings

Method:

Addresses

Address	Netmask	Gateway
192.168.200.10	24	192.168.200.1

DNS servers:

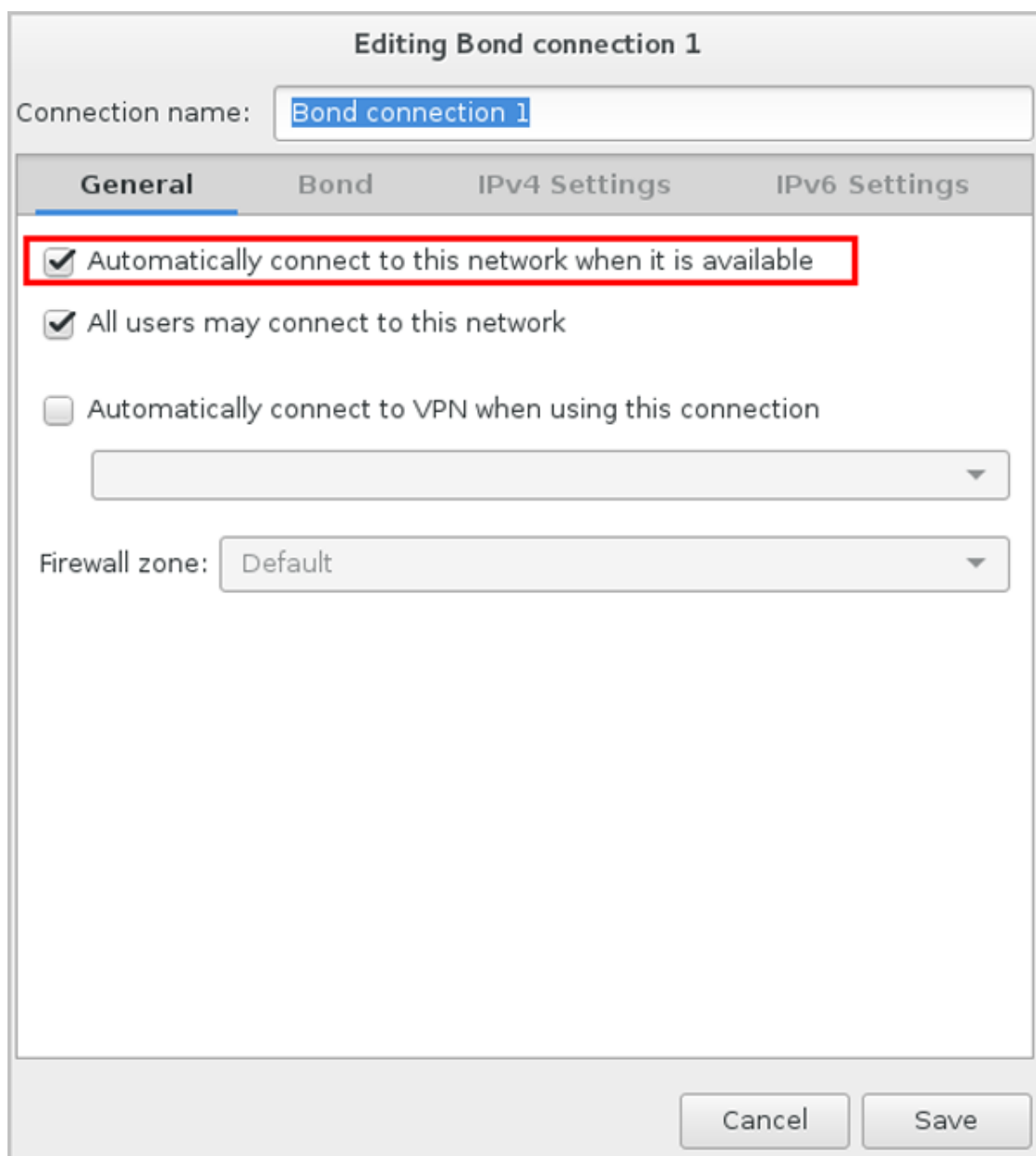
Search domains:

DHCP client ID:

☐ Require IPv4 addressing for this connection to complete

4. 设置Bond自动连接，按[图 30: 设置Bond自动连接](#)所示操作。
 - a. 在[图 19: 网卡配置主界面](#)，选中Bond，点击**Configure...**，打开Bond配置界面。
 - b. 进入**General**选项页。
 - c. 确认已勾选**Automatically connect to this network when it is available**
 - d. 点击**Save**保存。

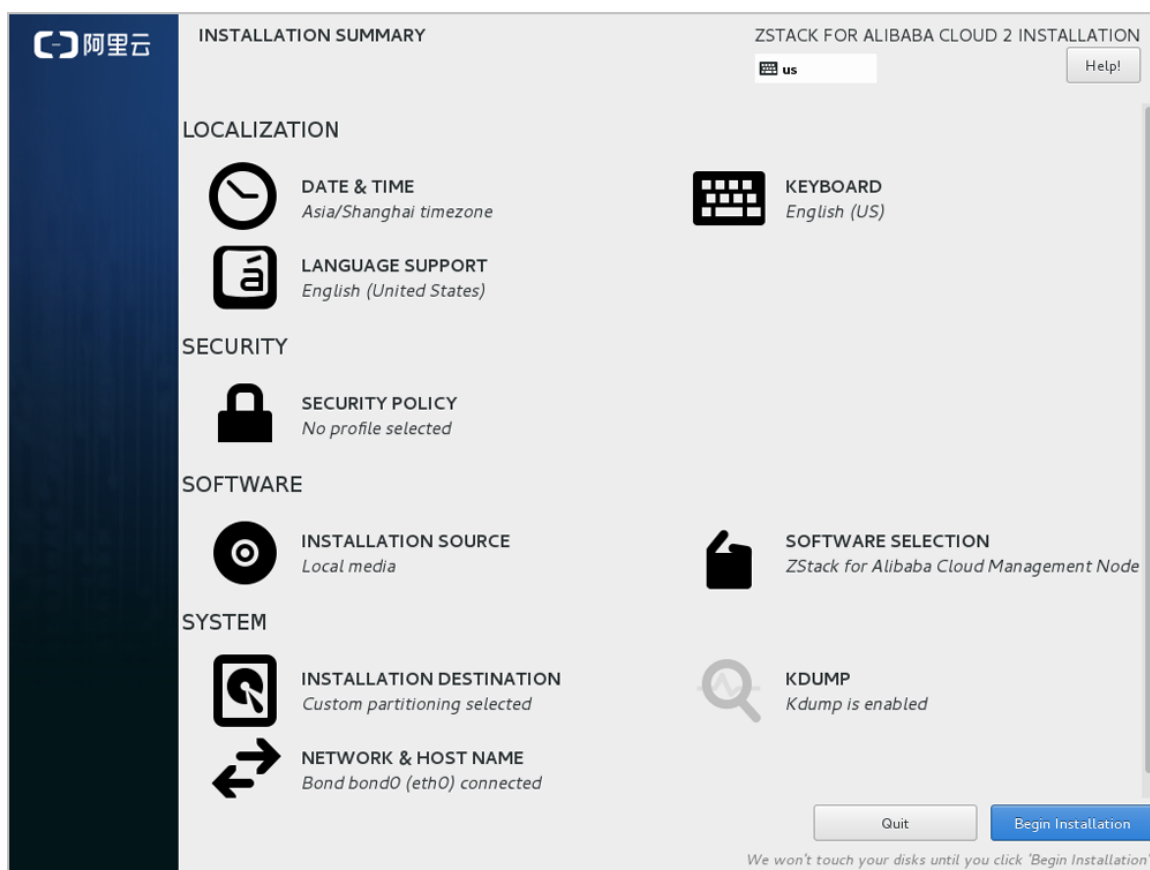
图 30: 设置Bond自动连接



g) 网络配置完后，回到系统安装主界面，点击**Begin Installation**开始安装。

如图 31: 点击`Begin Installation`所示：

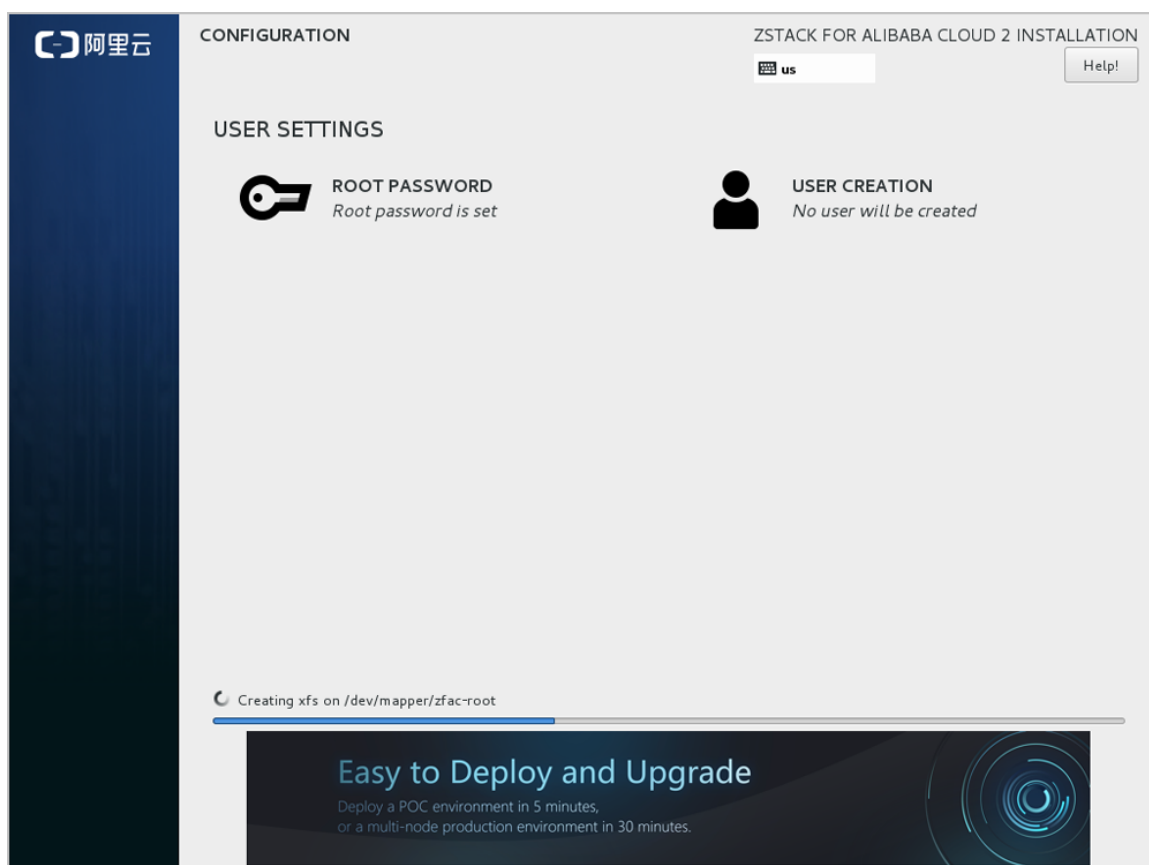
图 31: 点击**Begin Installation**



h) 安装过程自动进行，安装过程中请设置**ROOT PASSWORD**。

如图 32: 系统安装过程界面所示：

图 32: 系统安装过程界面



i) 安装完毕后，点击重启，即可启动进入ZStack for Alibaba Cloud定制版CentOS7.2系统。

后续操作

- 选择管理节点模式/计算节点模式/存储节点模式，系统重启后会自动安装对应的ZStack for Alibaba Cloud安装包。
- 选择专家模式，系统重启后进入shell界面，由高级用户自定义安装。

5.2 安装 ZStack for Alibaba Cloud

本节主要介绍ZStack for Alibaba Cloud V2.2.0的四种安装模式以及TUI的功能介绍。

ZStack for Alibaba CloudTUI简介

ZStack for Alibaba Cloud TUI是专为 ZStack for Alibaba Cloud集群中物理服务器准备的一套用户界面，其意义包含两方面：

- 分流UI的部分功能

将针对服务器的配置密码、配置网络、重启机器等操作从UI中剥离出来，集中显示在TUI中。

- 降低管理员登录服务器的频率

这是为了在降低物理机维护难度的同时，保护物理机内部的配置不被损坏。

用户可以用**Ctrl + Alt + F1**，再使用**Ctrl + Alt + F2**进入命令行模式。按下**Ctrl + Alt + F11**可以进入保留终端，用户可以在里面执行常规命令，但是请谨慎使用，以免对系统造成破坏，影响 ZStack for Alibaba Cloud 服务运行。

ZStack for Alibaba Cloud 定制版 ISO 四种安装模式

ZStack for Alibaba Cloud 定制版 ISO 提供了以下四种安装模式：

- 管理节点模式
- 计算节点模式
- 存储节点模式
- 专家模式

四种安装模式的步骤介绍：

1. 管理节点模式

- 安装基础系统
- 安装 MariaDB、RabbitMQ 等 ZStack 依赖包
- 安装企业版管理节点 TUI
- 自动安装并启动 ZStack 及其 UI

2. 计算节点模式

- 安装基础系统
- 安装 Libvirt、Qemu 等虚拟化组件
- 安装计算节点 TUI

3. OCFS2 存储节点模式

- 安装基础系统
- 安装 Libvirt、Qemu 等虚拟化组件
- 安装 UEK 内核及 OCFS2 组件
- 安装存储节点 TUI

4. 专家模式

- 安装基础系统

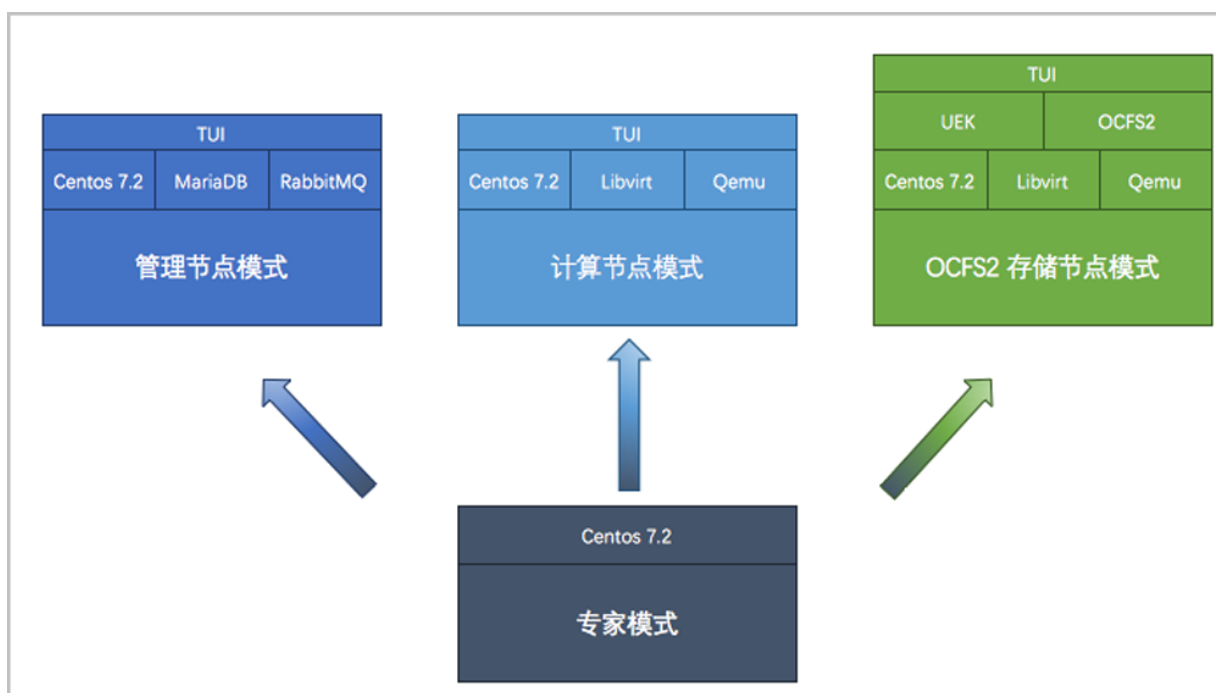


注:

- 专家模式基本等同于CentOS 7.2最小安装模式。
- 当管理员需要做更深度的定制时，可以选择进入专家模式，手动转化为其它三种模式中的任何一种。

四种安装模式如[图 33: 四种安装模式示意图](#)所示：

图 33: 四种安装模式示意图

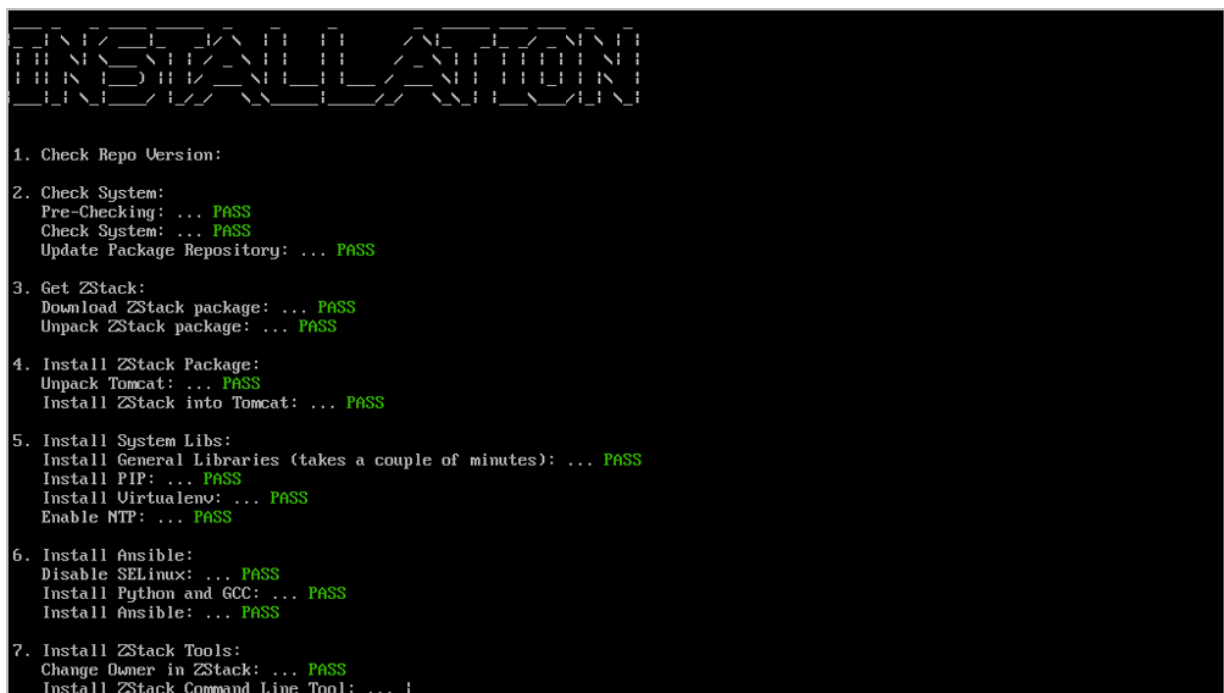


5.2.1 ZStack for Alibaba Cloud管理节点模式

自动安装ZStack for Alibaba Cloud管理节点

如果选择企业版/社区版管理节点模式，重启后会自动安装ZStack for Alibaba Cloud管理节点，安装完成后将自动进入TUI，如[图 34: 自动安装ZStack for Alibaba Cloud管理节点](#)所示：

图 34: 自动安装ZStack for Alibaba Cloud管理节点



管理节点TUI主界面

ZStack for Alibaba Cloud管理节点TUI主界面实时显示物理机的主要信息，如图 35: 管理节点TUI主界面所示:

图 35: 管理节点TUI主界面

ZStack for Alibaba Cloud Management Node	
NODE INFO MM Status: Running UI Status: Running Version : ZStack-aliyun 2.1.0.8 UI : http://10.0.88.41:5000 License Type: Trial	TIME INFO Boot Time: 2017-09-01 14:17 Local Time: 2017-09-01 14:51 Running Time: 34 mins 5 secs Load Average: 0.10, 0.07, 0.09
DISK INFO /: 486.488GB /var/lib/zstack: 486.488GB /var/log/zstack: 486.488GB /usr/local/zstack: 486.488GB	NIC INFO bond0 10.0.88.41 MM eth0 UP
HOST INFO Total : 0 Enabled : 0 Connected: 0	VM INFO Total : 0 Running: 0 Stopped: 0
CPU USAGE 0.0%	MEM USAGE 48.4% #####

- **NODE INFO :**

显示管理节点当前状态，包括服务是否运行、UI是否运行、ZStack版本、UI地址以及License类型等

- **TIME INFO :**

显示服务器启动时间、当前时间、服务器运行时长、服务器平均负载等

- **DISK INFO :**

显示ZStack默认安装目录以及日志目录等的剩余磁盘空间

- **NIC INFO :**

显示管理节点所有网卡的信息，包括物理网卡和逻辑网卡，其中DOWN表示网卡关闭、UP表示网卡启动、MN表示该网卡为管理网卡

- **HOST INFO :**

显示当前ZStack集群共有多少台计算节点，其中Enabled和Connected的各有多少：

- **VM INFO :**

显示当前ZStack集群共有多少台云主机，其中Running和Stopped的各有多少；

- CPU USAGE和MEM USAGE :

分别显示该管理节点的实时资源利用率。



注:

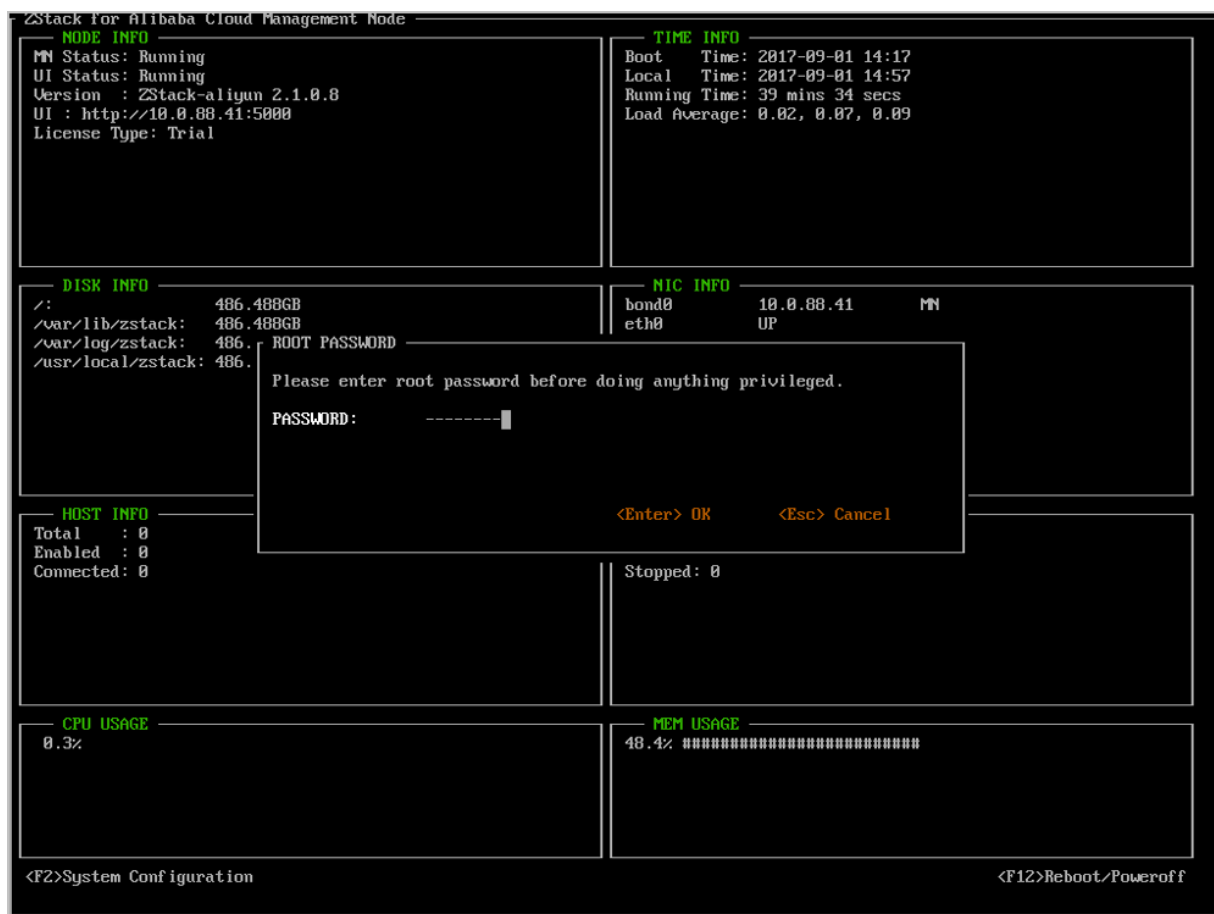
主界面底部还提供了进入系统配置和重启/关机两个入口，根据提示按下**F2**或者**F12**即可进入相应窗口。

系统密码

在主界面按下**F2**或者**F12**时，需要首先输入系统密码，因为无论系统配置还是重启/关机都属于特权操作。

下图中显示了输入系统密码的窗口，根据提示输入**ROOT**密码，回车即可。

图 36: 输入密码

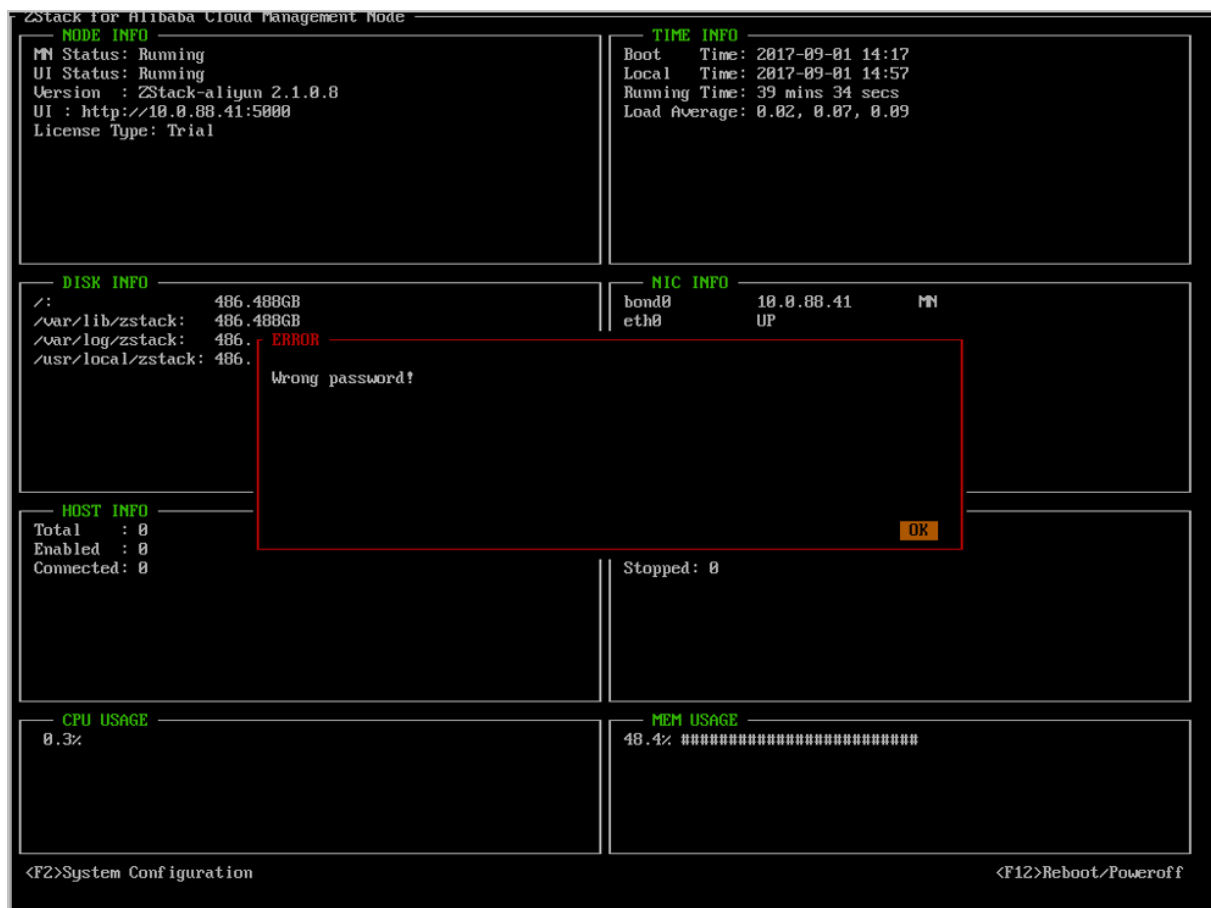


若密码不正确，操作会被阻止，TUI返回主界面。



注：任何时候，只要按下ESC键，就可以退出当前窗口，返回系统配置界面。

图 37: 输入密码错误



系统配置

系统配置的功能包括：

- 修改密码
- 重命名网卡
- 测试网络
- 配置网络
- 配置管理网络
- 配置控制台代理
- 收集日志
- 备份数据库
- 启动/关闭/重启ZStack服务

- 重装ZStack服务
- 重装ZStack服务并删除数据库
- 终端信息

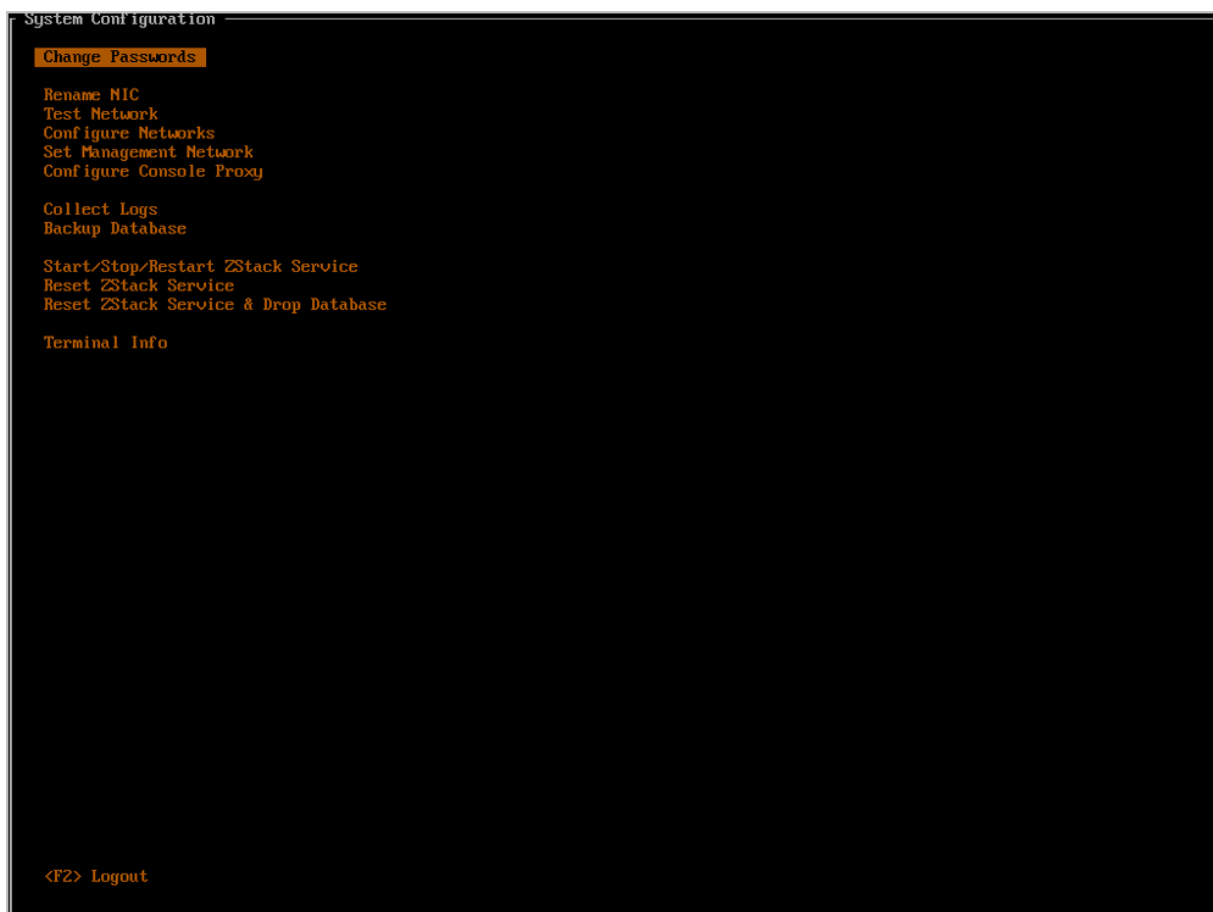


注:

- 系统配置界面提供若干系统配置入口。使用上下键移动光标，选择需要的配置按钮，按下回车即可进入相应配置界面。对于熟悉Vim的用户，也可以使用j和k实现光标上下移动。
- 根据提示，按下F2可以退出至主界面。再次回到系统配置界面时需要重新输入ROOT密码。

如图 38: 系统配置所示：

图 38: 系统配置



以下将逐一介绍系统配置各功能条目。

1. 修改密码

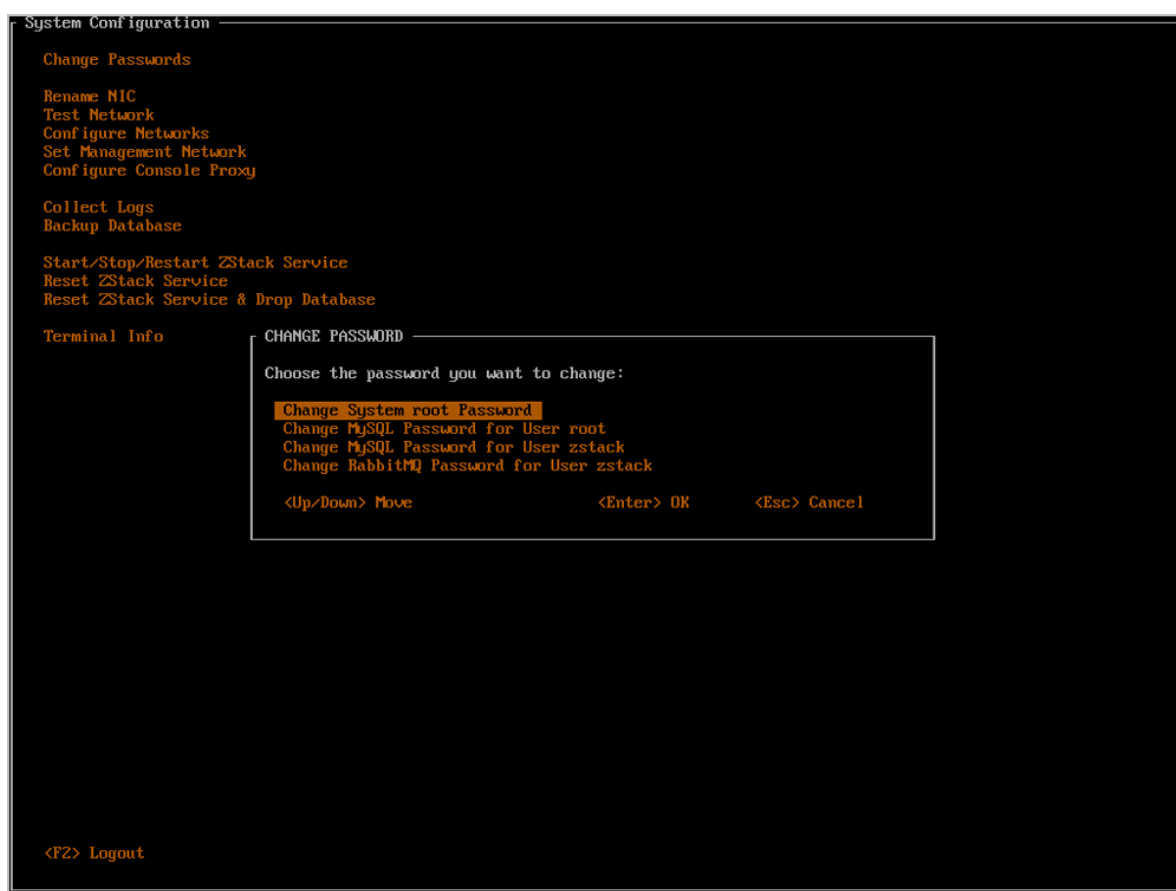
将光标移动至**Change Passwords**处，按下回车，即可进入密码配置窗口。

该配置项集中了管理节点需要的四种密码修改动作：

1. 修改系统root密码
2. 修改root账号的MySQL密码
3. 修改ZStack账号的MySQL密码
4. 修改ZStack账号的RabbitMQ密码

如图 39: 四种密码修改动作所示：

图 39: 四种密码修改动作

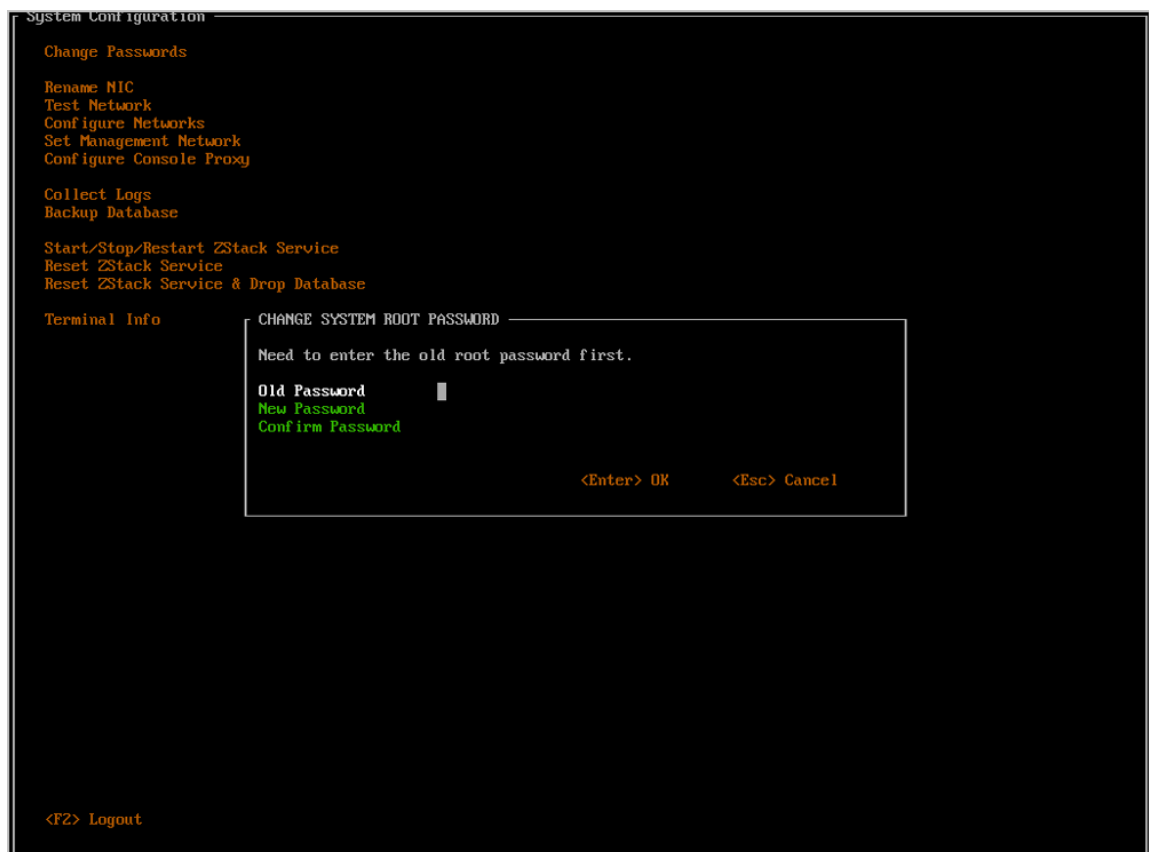


四种密码修改动作具体介绍：

1. 修改系统root密码

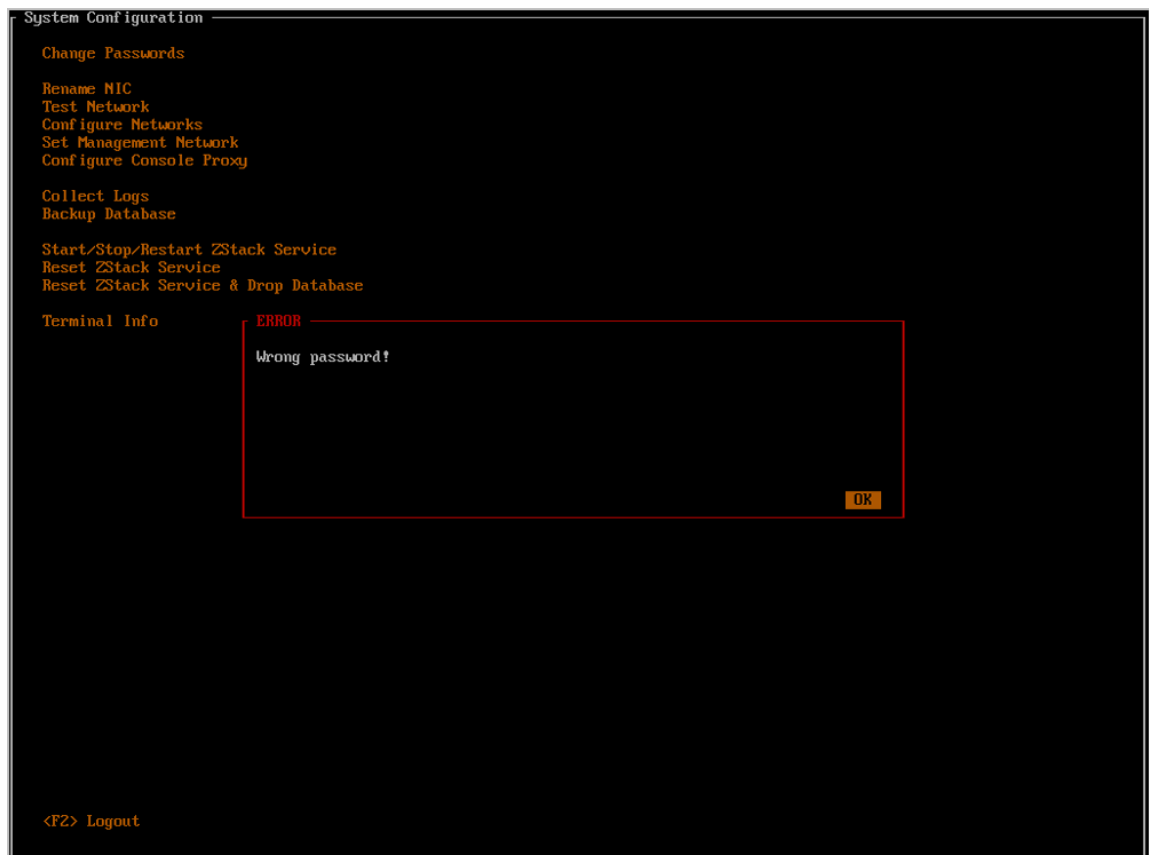
根据提示输入旧系统密码、新系统密码以及新密码确认，回车即可，如图 40: 修改系统root密码所示：

图 40: 修改系统root密码



如果旧系统密码不正确、或者密码确认不一致，都会弹出错误窗口，如图 41: 错误窗口所示：

图 41: 错误窗口



需要根据提示，按Y键以确认，如[图 42: 修改系统root密码的确认界面](#)所示；修改系统root密码成功，如[图 43: 修改系统root密码成功](#)所示。

图 42: 修改系统root密码的确认界面

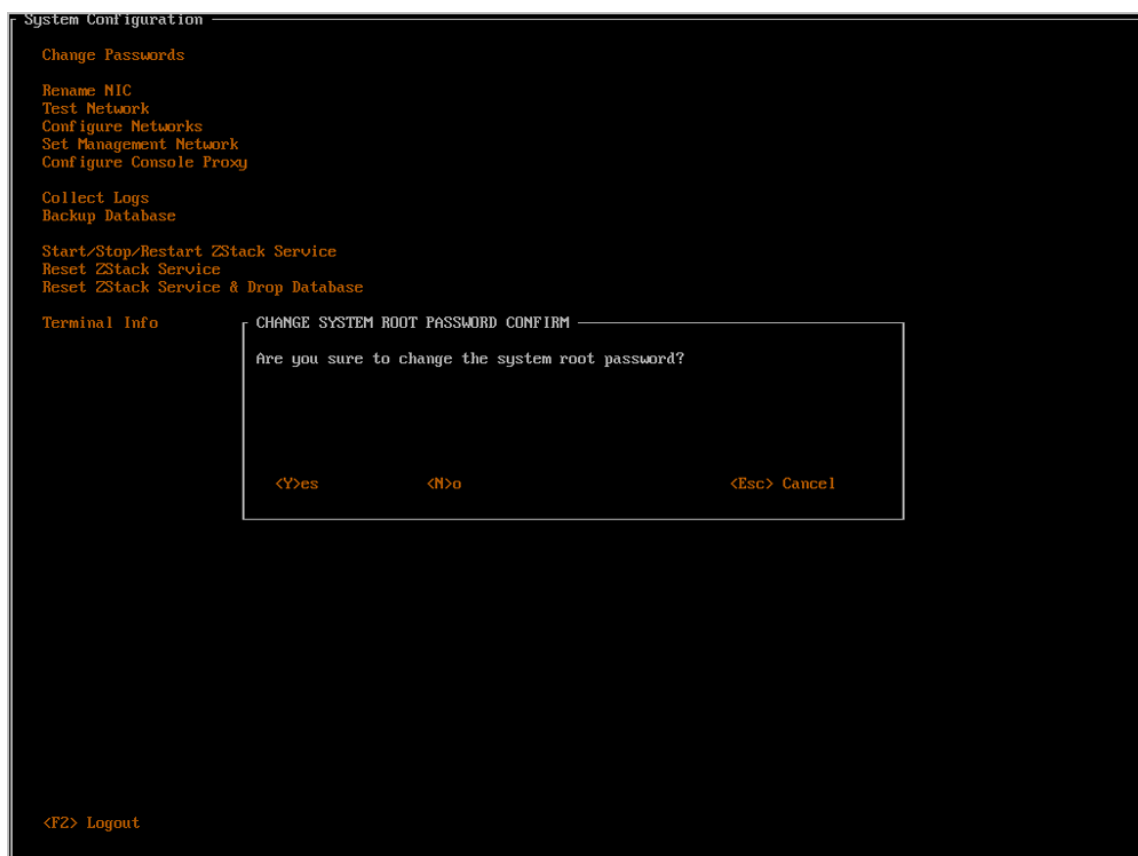
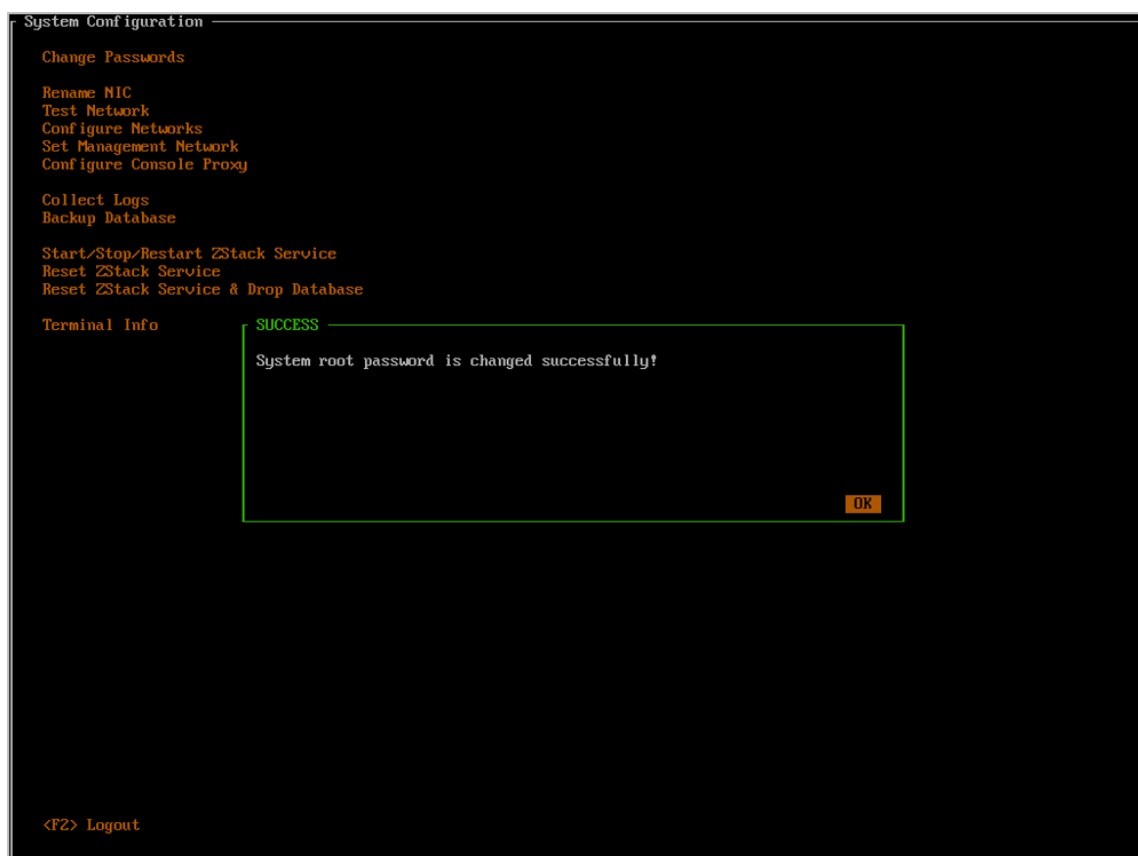


图 43: 修改系统root密码成功

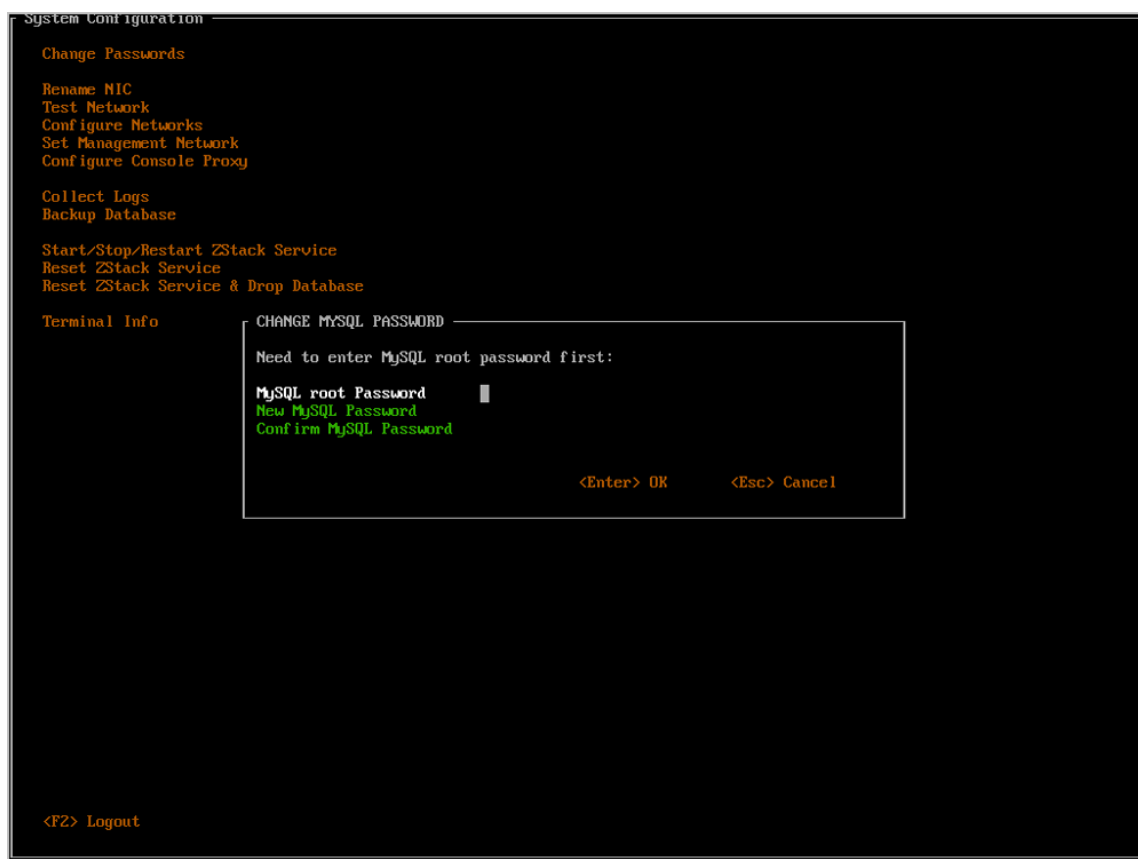


2. 修改root账号/ZStack账号的MySQL密码

- 修改root账号/ZStack账号的MySQL密码，需要MySQL的root权限，因此首先需要输入MySQL root密码，然后根据提示输入新密码和密码确认，回车即可。
- 如果MySQL root密码不正确、或者密码确认不一致，都会弹出错误窗口。
- 此操作需要重启ZStack服务，耗时较长，需要用户按Y键确认。

如图 44: 修改root账号/ZStack账号的MySQL密码所示：

图 44: 修改root账号/ZStack账号的MySQL密码

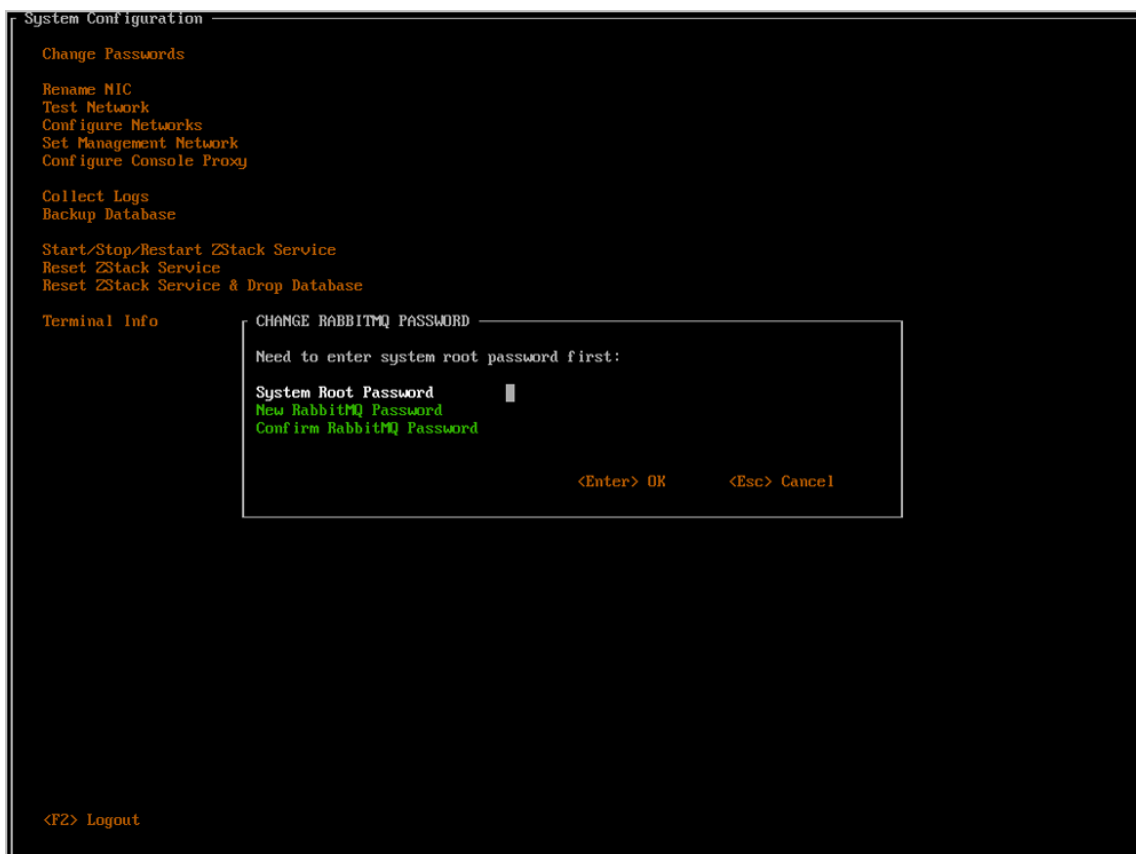


3. 修改ZStack账号的RabbitMQ密码

- 修改ZStack账号的RabbitMQ密码需要管理员权限，因此首先需要输入系统root密码，然后根据提示再输入新密码和密码确认，回车即可。
- 如果系统密码不正确、或者密码确认不一致，都会弹出错误窗口。
- 此操作需要重启ZStack服务，耗时较长，需要用户按Y键确认。

如图 45: 修改ZStack账号的RabbitMQ密码所示：

图 45: 修改ZStack账号的RabbitMQ密码



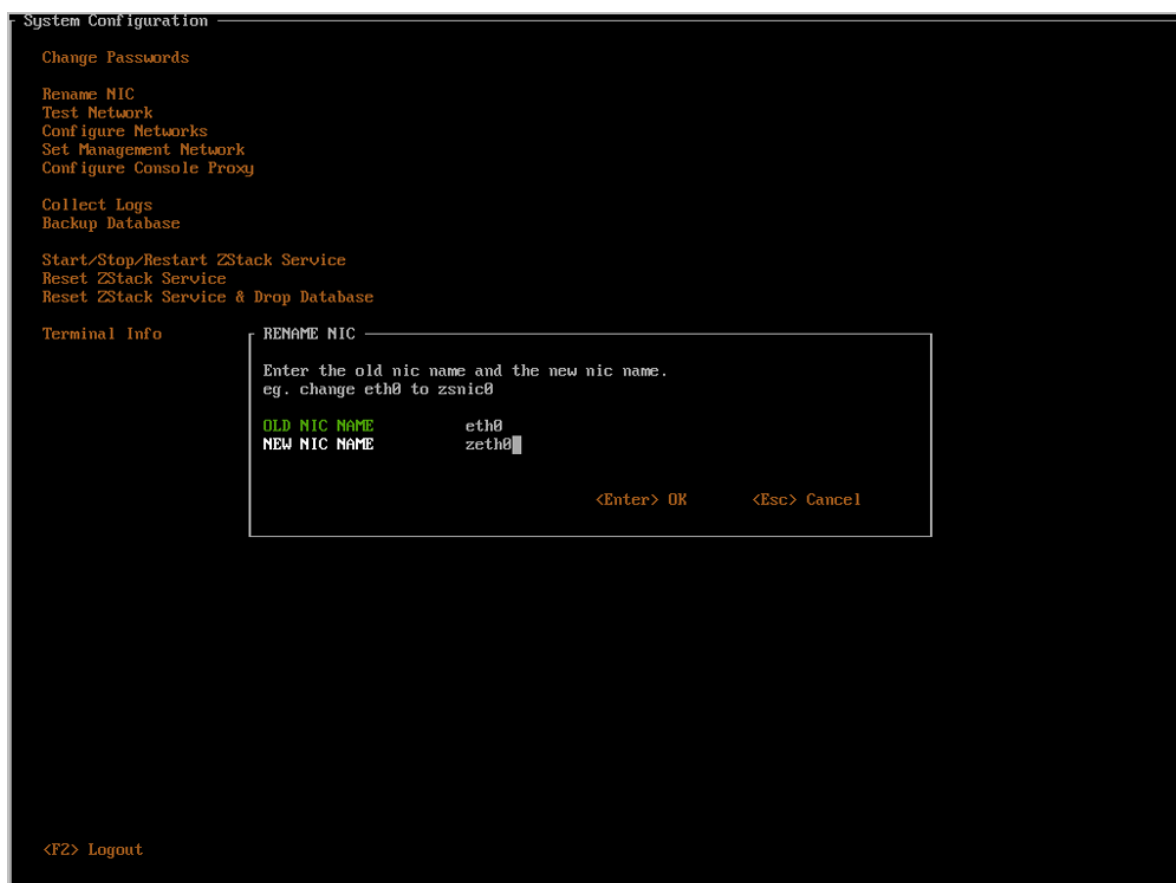
2. 重命名网卡

将光标移动至**Rename NIC**处，按下回车，即可进入重命名网卡窗口。

- 用户可以在此修改网卡名。
- 修改完成后，按下回车，此操作需要按Y键确认。

如图 46: 重命名网卡所示：

图 46: 重命名网卡



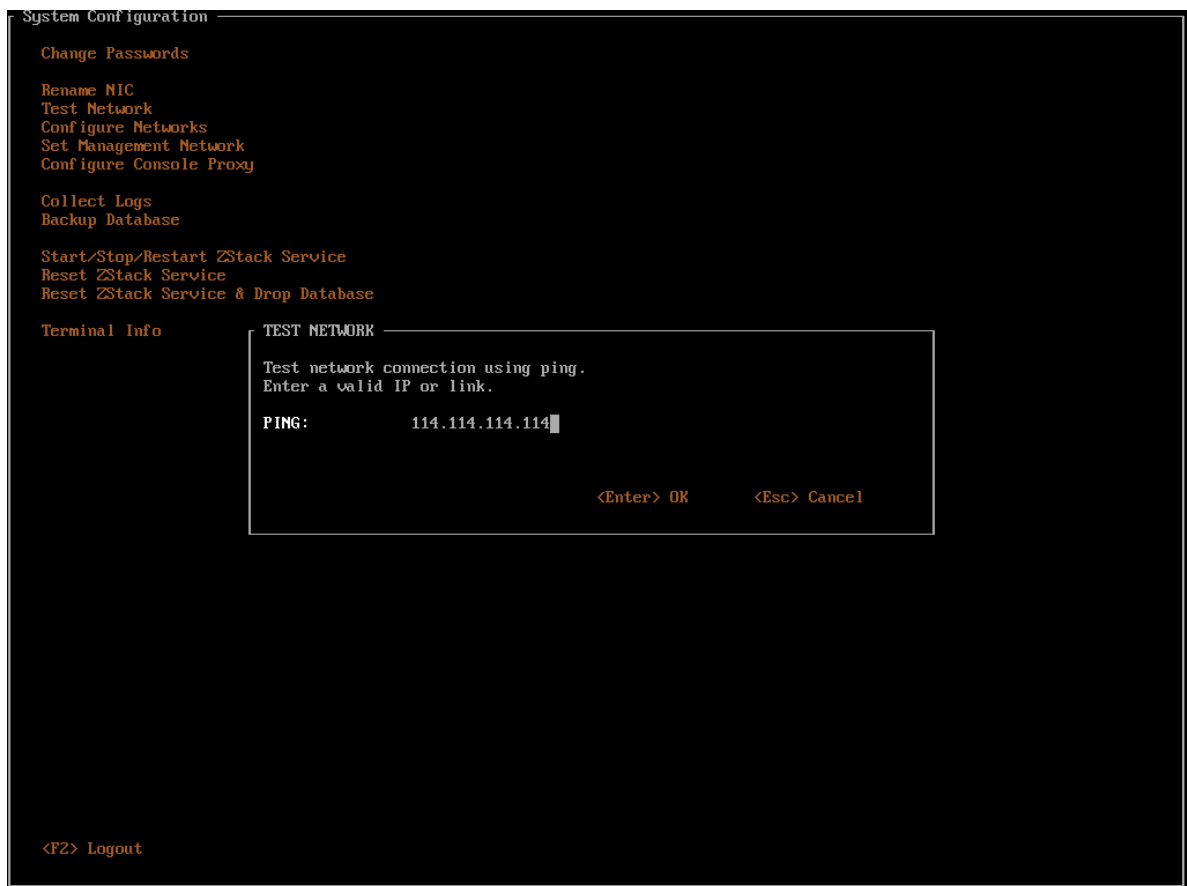
3. 测试网络

将光标移动至**Test Network**处，按下回车，即可进入网络测试窗口。

这里默认填写了114.114.114.114，用户测试外网连接；若要测试内网连接，请自行填写内网IP。

如图 47: 测试网络所示：

图 47: 测试网络

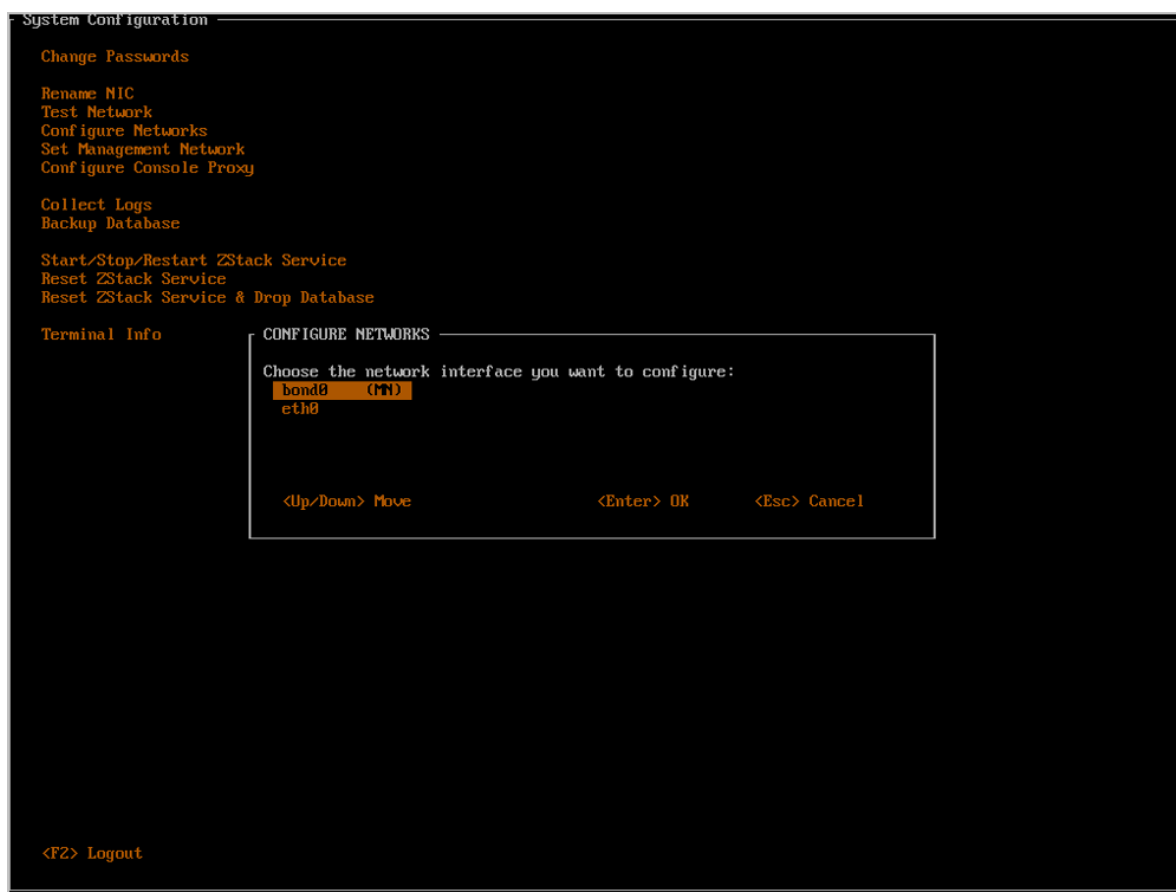


4. 配置网络

将光标移动至**Configure Networks**处，按下回车，即可进入网络配置窗口。

如图 48: 配置网络所示：

图 48: 配置网络



这里罗列了管理节点中所有处于启动状态的网卡，包括物理网卡和逻辑网卡。由于前文已做网卡归一化，这里仅需关心Bond的配置即可。

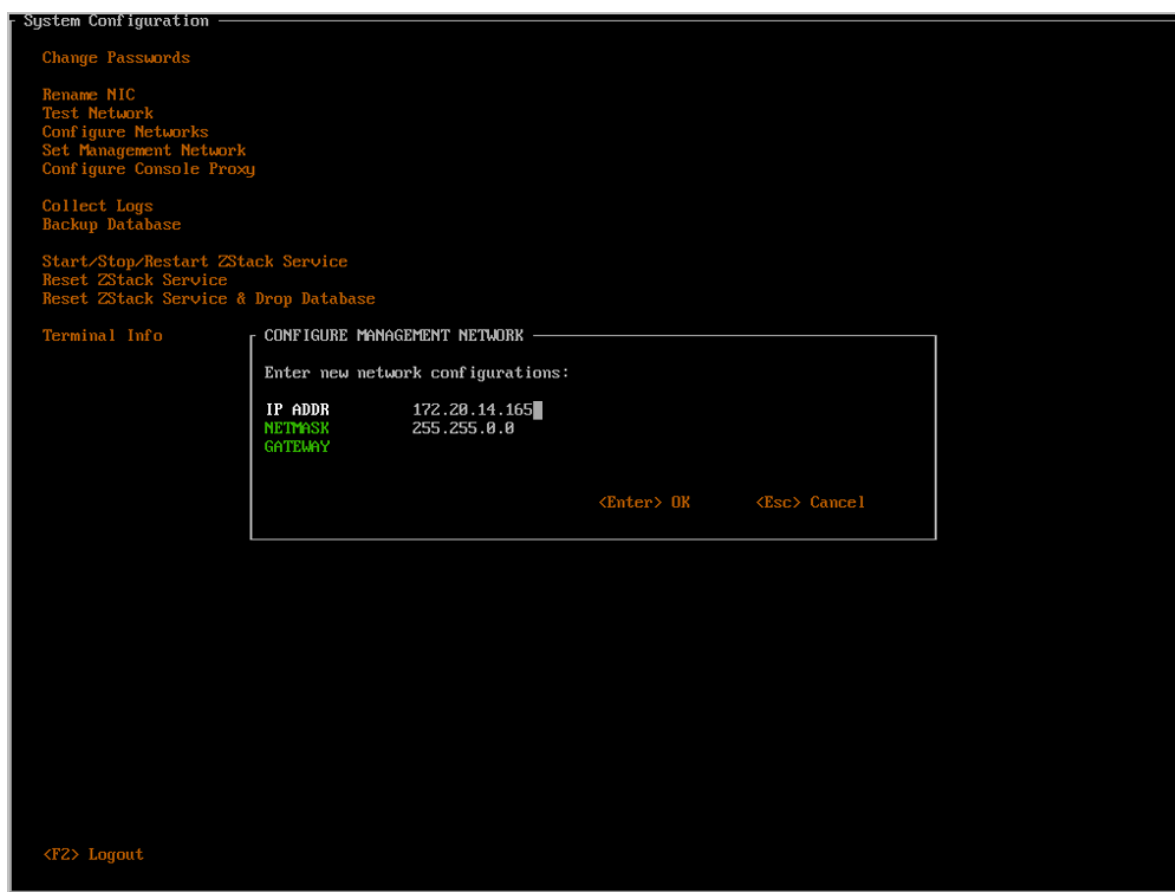
将光标移动至目标Bond，如bond0，再次回车，即可进入该设备的配置界面。



注：如果目标Bond后带有（MN）的标记，则意味着修改管理网络配置，需要较长的等待时间，请谨慎操作。

如图 49: 填写信息所示：

图 49: 填写信息



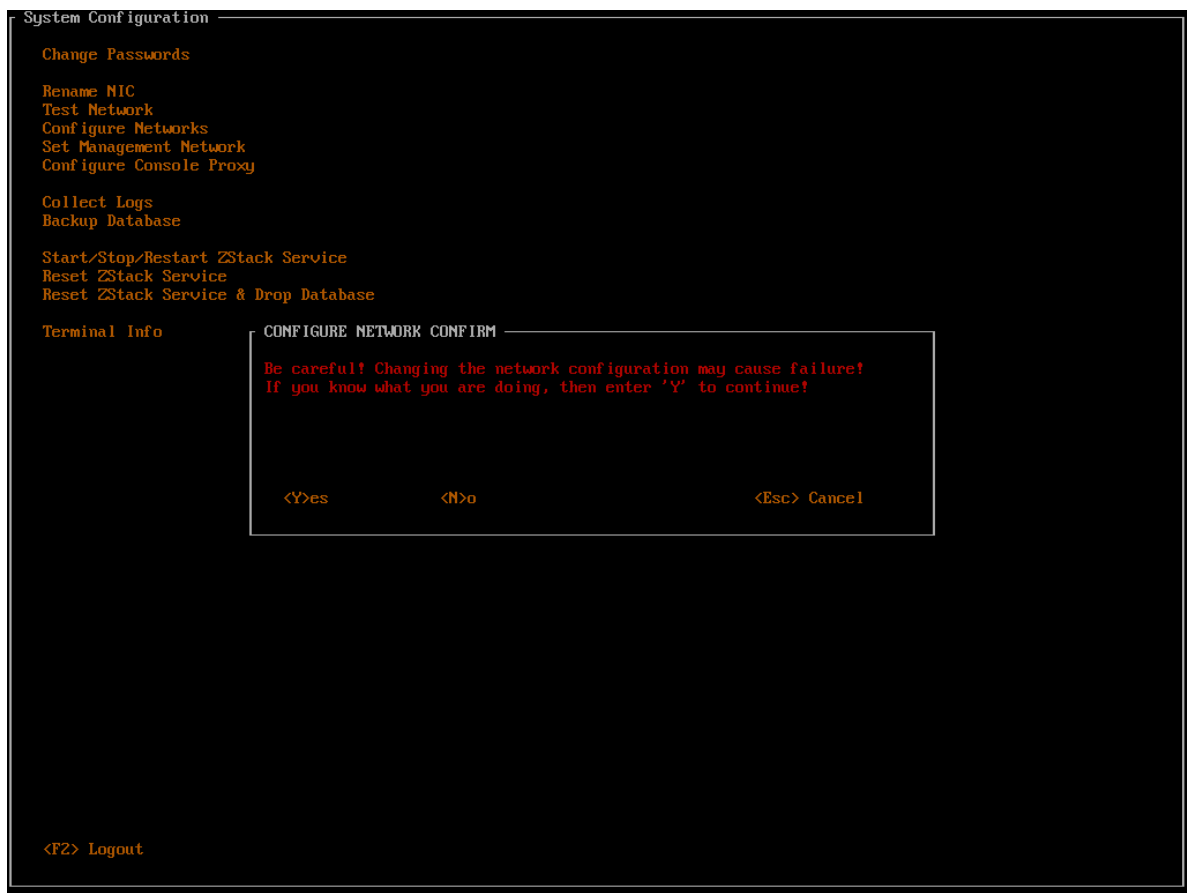
所选Bond设备的已有配置会默认显示在输入框中（网关除外），用户可根据实际情况修改。

如果出现以下任何一种情况，均认为输入有误：

- 有某个或某些输入项为空
- 有某个或某些输入项不是合法的IP地址
- IP地址和网关不在掩码所确定的同一个子网内

确认界面如图 50: 确认界面所示：

图 50: 确认界面



5. 配置管理网络

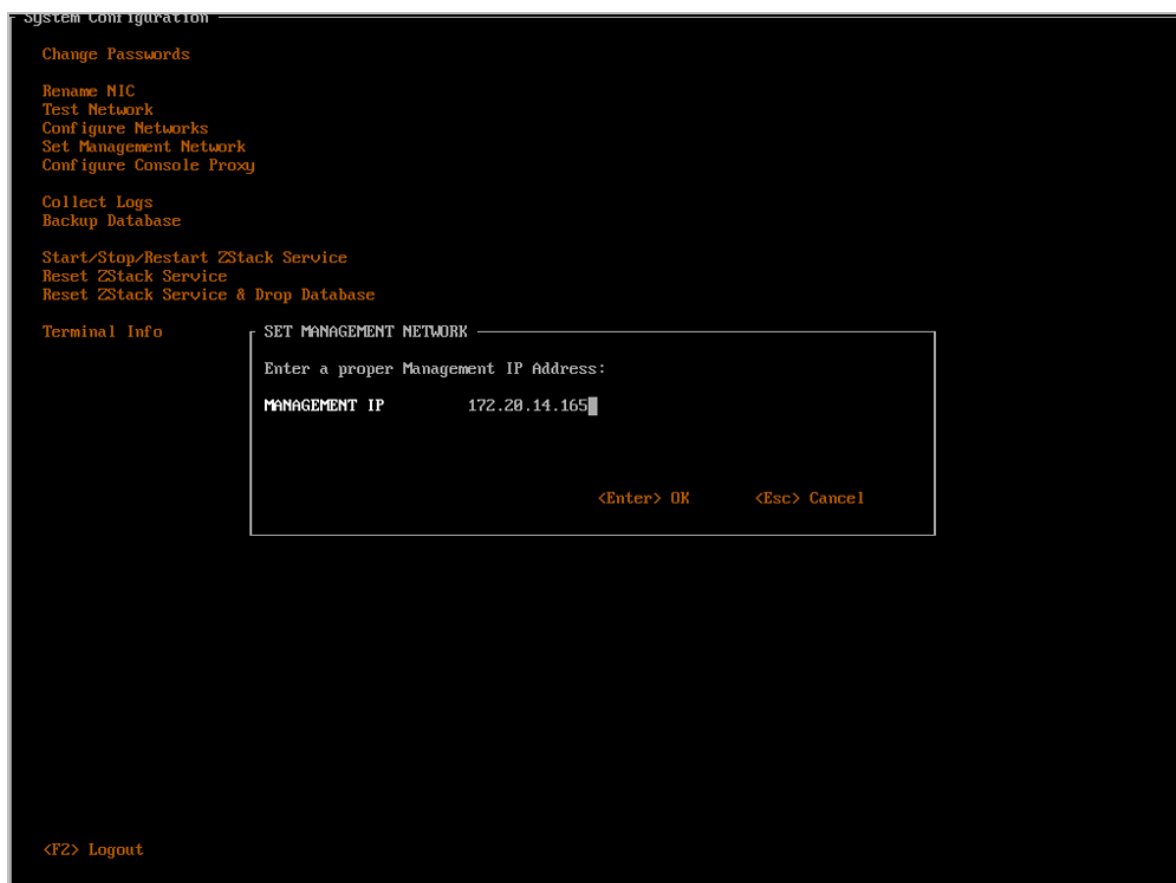
将光标移动至**Set Management Network**处，按下回车，即可进入管理网络配置窗口。



注：配置管理网络耗时较长，请谨慎操作。

如图 51: 配置管理网络所示：

图 51: 配置管理网络



确认界面如图 52: 确认界面所示：

图 52: 确认界面



6. 配置控制台代理

将光标移动至**Configure Console Proxy**处，按下回车，即可进入终端代理配置窗口。

输入代理IP和代理端口，回车即可。

控制台代理的原值将被读出并填充在输入项内，用户可按实际需要进行修改。

如果出现以下任何一种情况，均认为输入有误：

- 代理地址和端口均为空
- 代理地址不是合法的IP地址
- 端口号不是1024至65535之间的数字

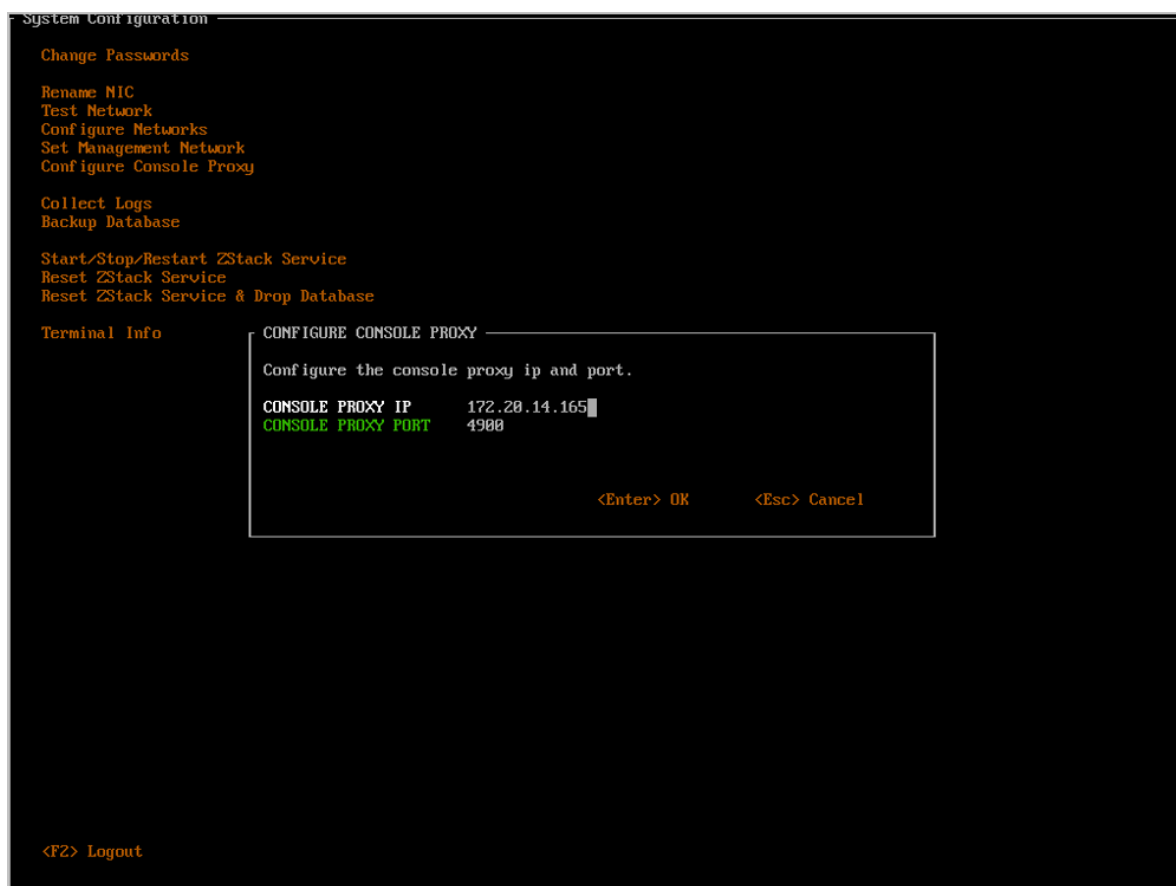


注：

- 此操作需要重置RabbitMQ服务，耗时较长。
- 需要用户按Y键确认。

如图 53: 配置控制台代理所示：

图 53: 配置控制台代理

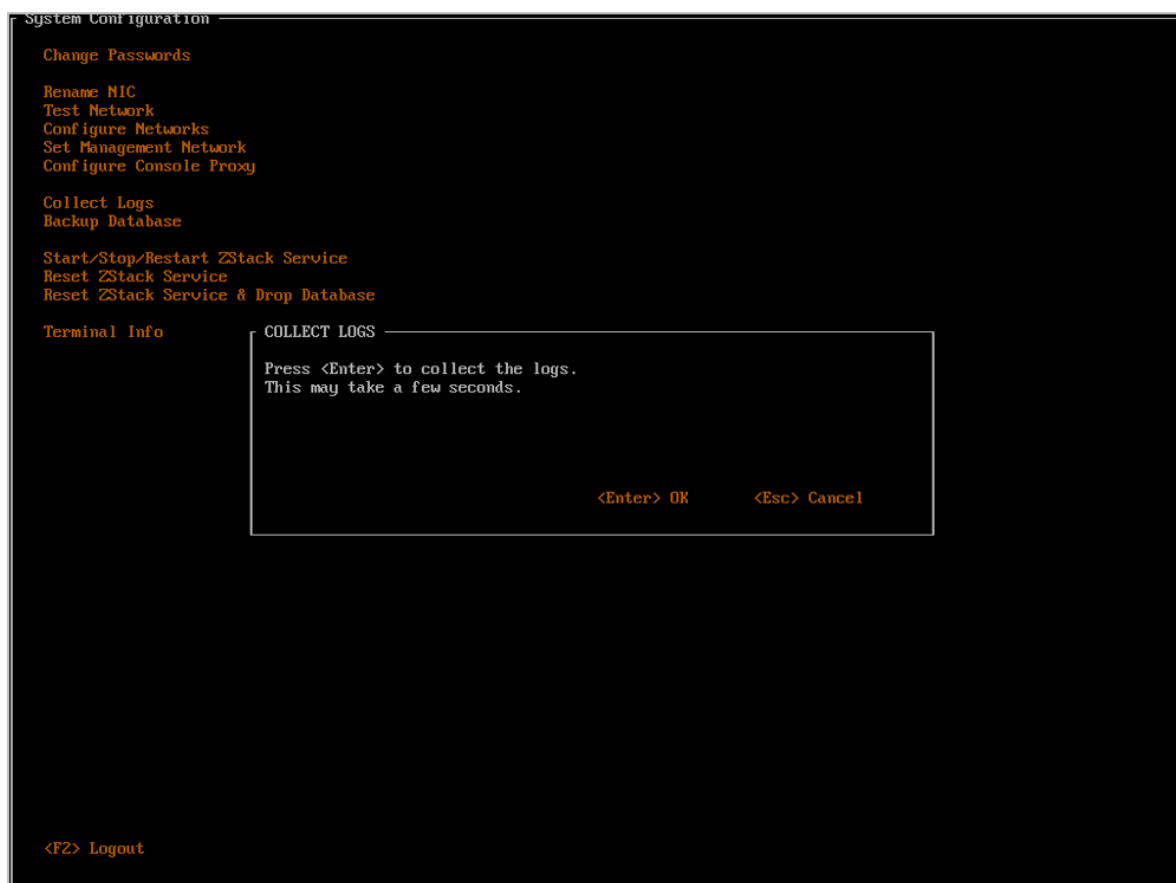


7. 收集日志

将光标移动至**Collect Logs**处，按下回车，即可进入日志收集窗口。

如图 54: 收集日志所示：

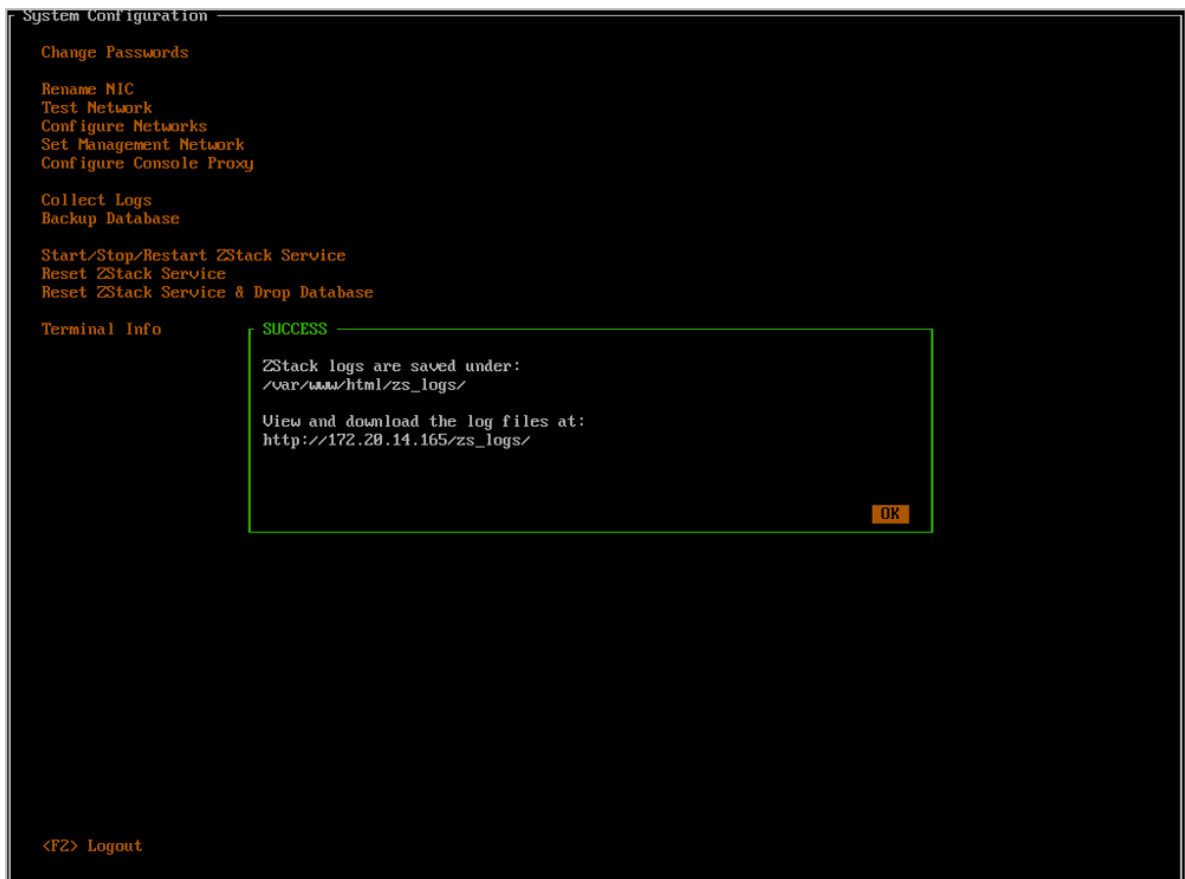
图 54: 收集日志



点击回车，即可将整个集群的日志收集，并导出至HTTP服务器中，以供下载/在线浏览。

如图 55: 日志导出成功所示：

图 55: 日志导出成功



根据提示，用户可以通过浏览器直接访问所有日志内容，支持下载/在线浏览日志。

如图 56: 导出日志至HTTP服务器中所示，

图 56: 导出日志至HTTP服务器中

Index of /zs_logs

Name	Last modified	Size	Description
 Parent Directory		-	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-15.tar.gz	2017-08-01 21:15	1.8M	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-15/	2017-08-01 21:15	-	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-34.tar.gz	2017-08-01 21:34	1.9M	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-34/	2017-08-01 21:34	-	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-40.tar.gz	2017-08-01 21:40	1.9M	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-40/	2017-08-01 21:40	-	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-55.tar.gz	2017-08-01 21:55	1.9M	
 collect-log-ZStack 2.1.0.44-2017-08-01 21-55/	2017-08-01 21:55	-	

图 57: 在线浏览日志

```

2017-08-01 13:55:20,268 DEBUG [Platform] {} linked global property[org.zstack.storage.backup.imagestore.ImageStoreBackupStorageGlobalProperty.QEMU_PACKAGE_NAME], value: zsta
2017-08-01 13:55:20,268 DEBUG [Platform] {} linked global property[org.zstack.storage.backup.imagestore.ImageStoreBackupStorageGlobalProperty.AGENT_PORT], value: 8001
2017-08-01 13:55:20,268 DEBUG [Platform] {} linked global property[org.zstack.storage.backup.imagestore.ImageStoreBackupStorageGlobalProperty.AGENT_URL_SCHEME], value: http
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.storage.backup.imagestore.ImageStoreBackupStorageGlobalProperty.AGENT_URL_ROOT_PATH], value:
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.EXIT_VM_ON_BOOT_FAILURE], value: false
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.BEAN_REF_CONTEXT_CONF], value: beanRefContext.xml
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.VM_TRACER_ON], value: true
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.PROFILER_WORKFLOW], value: false
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.PROFILER_HTTP_CALL], value: false
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.EXIT_VM_ON_STOP], value: true
2017-08-01 13:55:20,269 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.CHECK_BOX_TYPE_IN_INVENTORY], value: false
2017-08-01 13:55:20,270 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.PID_FILE_PATH], value: /usr/local/zstack/management-server.pid
2017-08-01 13:55:20,270 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.CONSOLE_PROXY_OVERRIDDEN_IP], value: 10.0.150.10
2017-08-01 13:55:20,270 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.EXPOSE_SIMULATOR_TYPE], value: false
2017-08-01 13:55:20,270 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.SHADOW_ENTITY_ON], value: true
2017-08-01 13:55:20,270 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.LOCALE], value: zh_CN
2017-08-01 13:55:20,270 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.USER_HOME], value: /usr/local/zstack
2017-08-01 13:55:20,270 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.REST_FACADE_READ_TIMEOUT], value: 300000
2017-08-01 13:55:20,271 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.REST_FACADE_CONNECT_TIMEOUT], value: 15000
2017-08-01 13:55:20,271 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.IS_UPGRADE_START], value: false
2017-08-01 13:55:20,271 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.SHADOW_ENTITY_ON], value: false
2017-08-01 13:55:20,271 DEBUG [Platform] {} linked global property[org.zstack.core.CoreGlobalProperty.CONSOLE_PROXY_PORT], value: 4900
2017-08-01 13:55:20,271 DEBUG [Platform] {} linked global property[org.zstack.storage.volume.VolumeGlobalProperty.ROOT_VOLUME_FIND_MISSING_IMAGE_UUID], value: false
2017-08-01 13:55:20,271 DEBUG [Platform] {} linked global property[org.zstack.storage.volume.VolumeGlobalProperty.SYNC_VOLUME_SIZE], value: false
2017-08-01 13:55:20,271 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.SERVER_IPS], value: [10.0.150.10]
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.CLOSE_TRACKER], value: false
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.TRACKER_GARBAGE_COLLECTOR_INTERVAL], value: 600
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.MESSAGE_LOG_FILTER_ALL], value: true
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.MESSAGE_LOG], value: null
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_USERNAME], value: zstack
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_PASSWORD], value: zstack.password
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_VIRTUAL_HOST], value: null
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_HEART_BEAT_TIMEOUT], value: 60
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_CONNECTION_TIMEOUT], value: 10
2017-08-01 13:55:20,272 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_RETRY_DELAY_ON_RETURN], value: 5
2017-08-01 13:55:20,273 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_RECOVERABLE_SEND_TIMES], value: 5
2017-08-01 13:55:20,273 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.RABBITMQ_NETWORK_RECOVER_INTERVAL], value: 1
2017-08-01 13:55:20,273 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.COMPRESS_NON_API_MESSAGE], value: false
2017-08-01 13:55:20,273 DEBUG [Platform] {} linked global property[org.zstack.core.cloudbus.CloudBusGlobalProperty.CHANNEL_POOL_SIZE], value: 100

```

8. 备份数据库

将光标移动至**Backup Database**处，按下回车，即可进入数据库备份窗口。

回车即可导出数据库至HTTP服务器中，以供下载。

如图 58: 备份数据库和图 59: 备份数据库成功所示：

图 58: 备份数据库

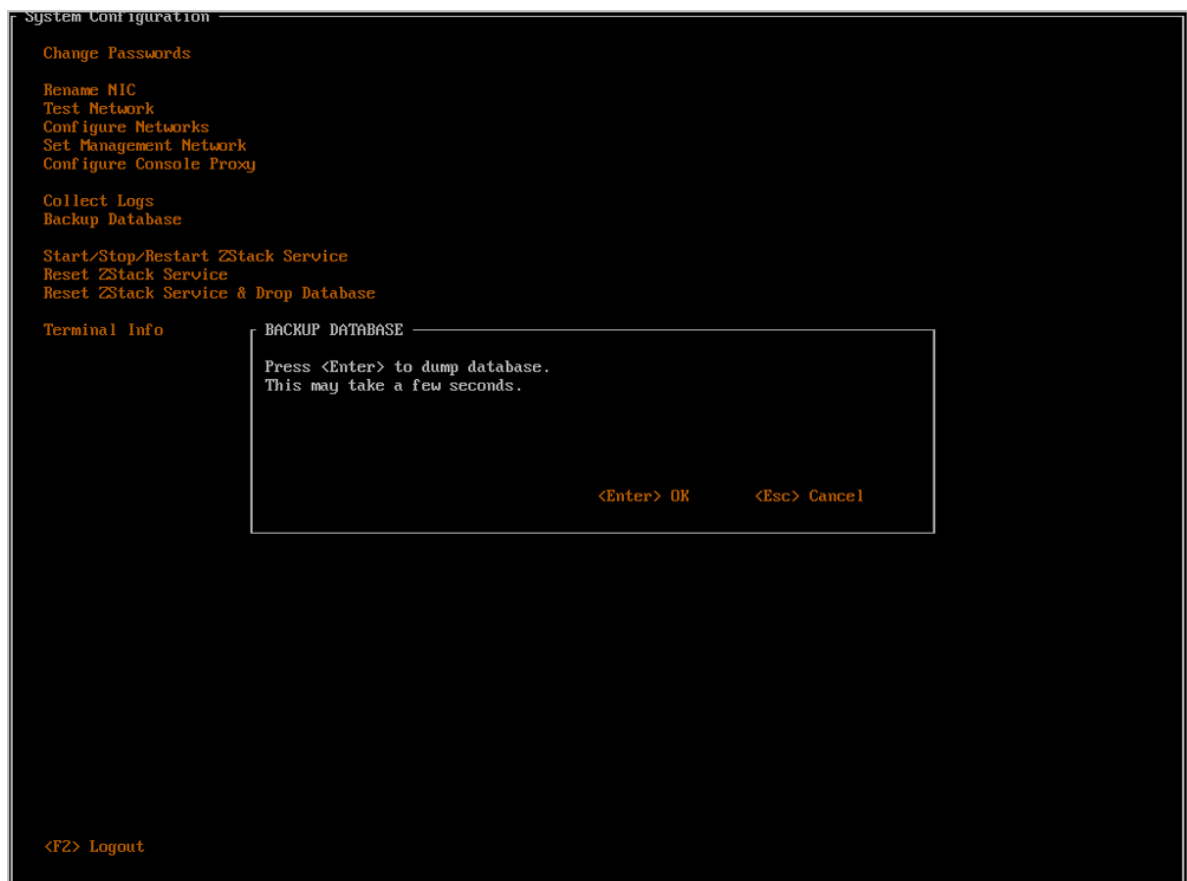
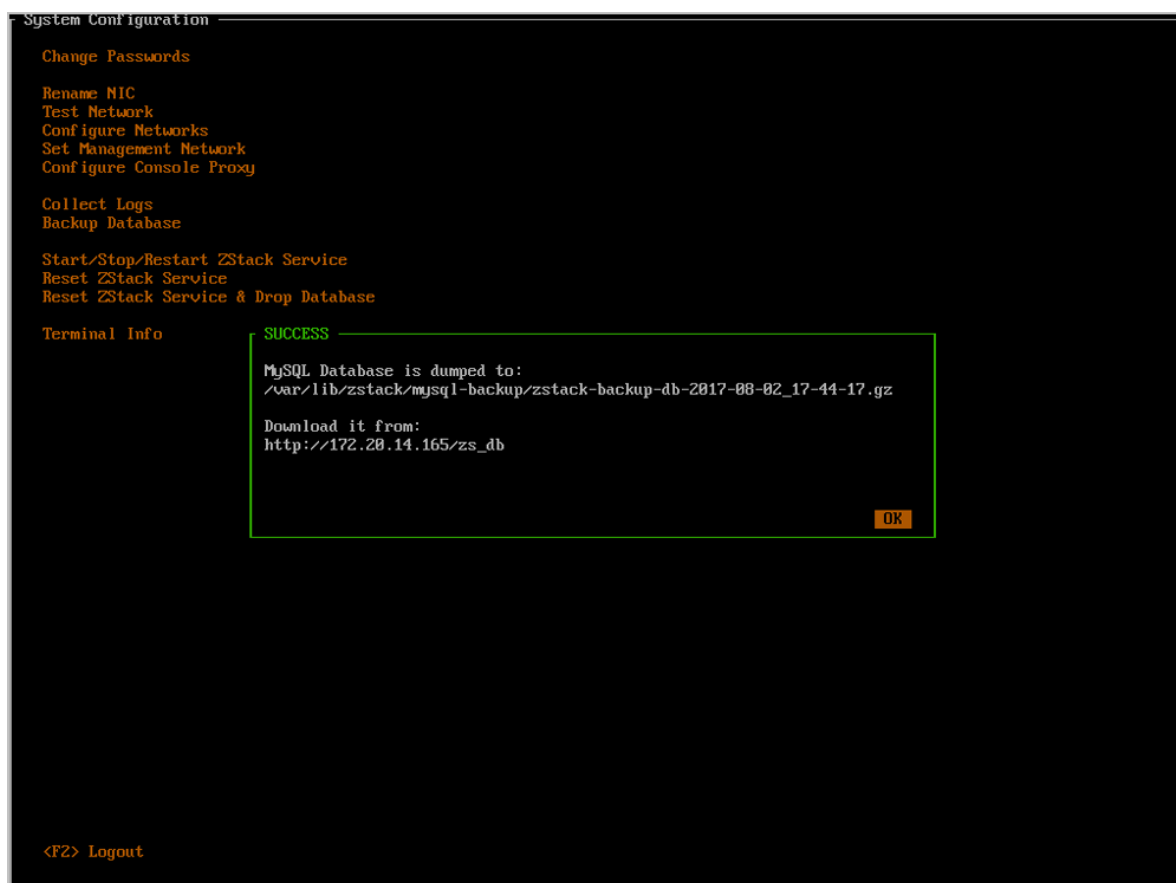


图 59: 备份数据库成功



9. 启动/关闭/重启ZStack服务

将光标移动至**Start/Stop/Restart ZStack Service**处，按下回车，即可进入启动/关闭/重启ZStack服务窗口。

回车后，再次移动光标，选择具体操作并回车。

如图 60: 启动/关闭/重启ZStack服务所示：

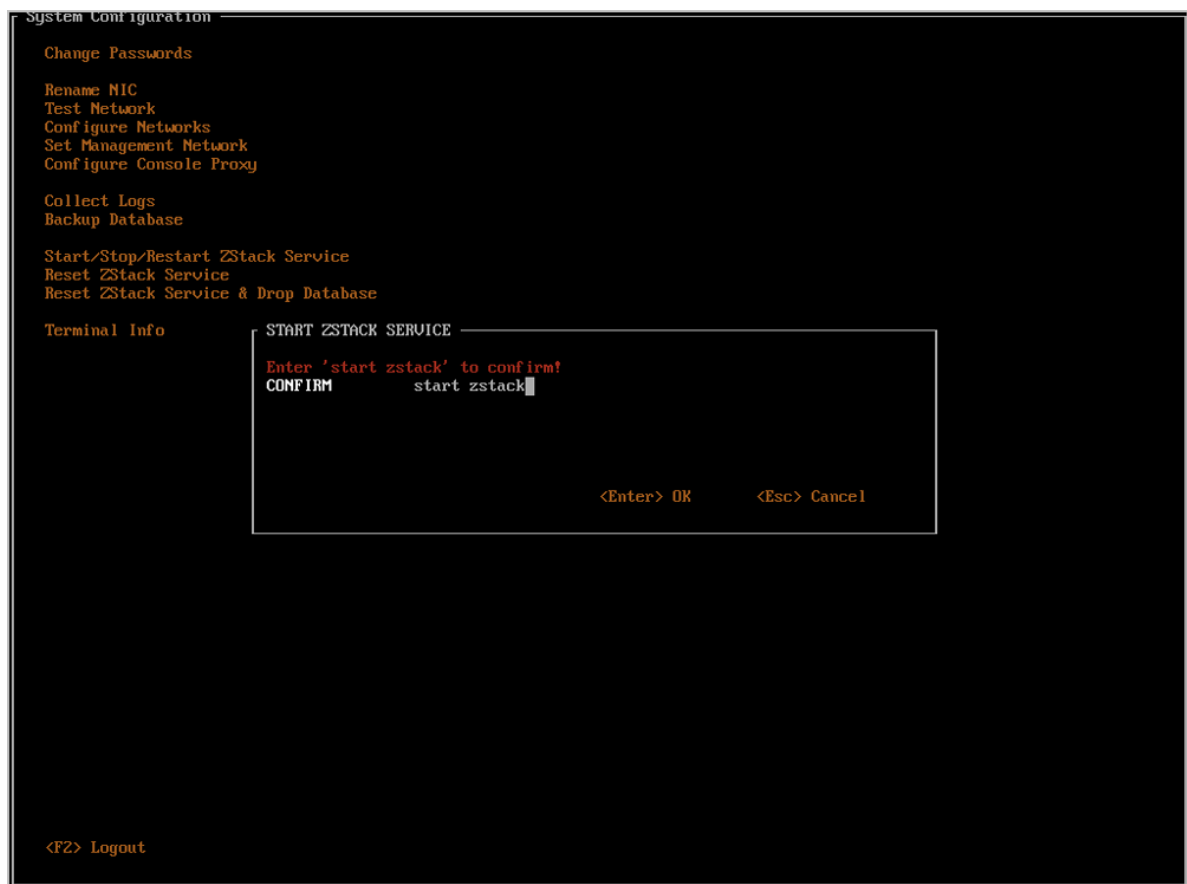
图 60: 启动/关闭/重启ZStack服务



根据提示，输入**start zstack**以确认启动ZStack、输入**stop zstack**以确认停止ZStack、输入**restart zstack**以重启ZStack：

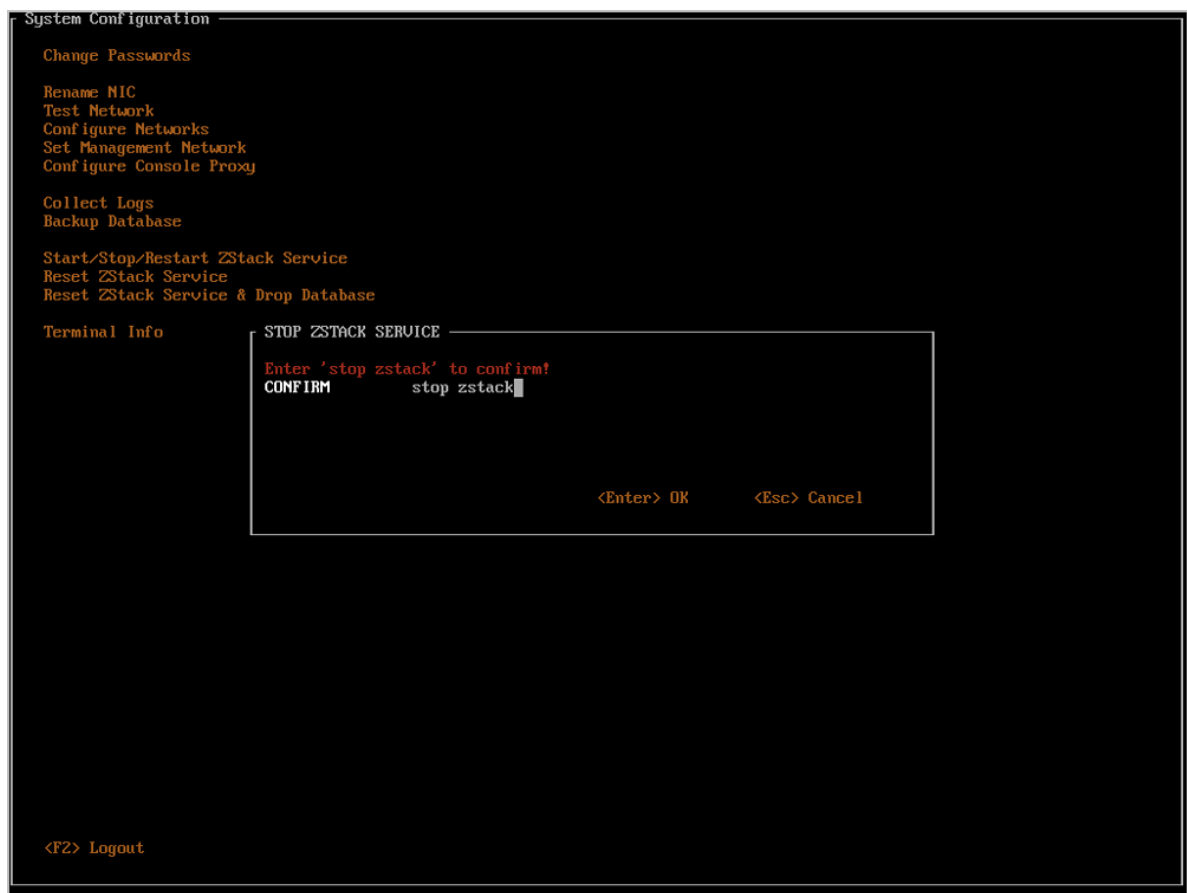
如图 61: 启动ZStack服务所示：

图 61: 启动ZStack服务



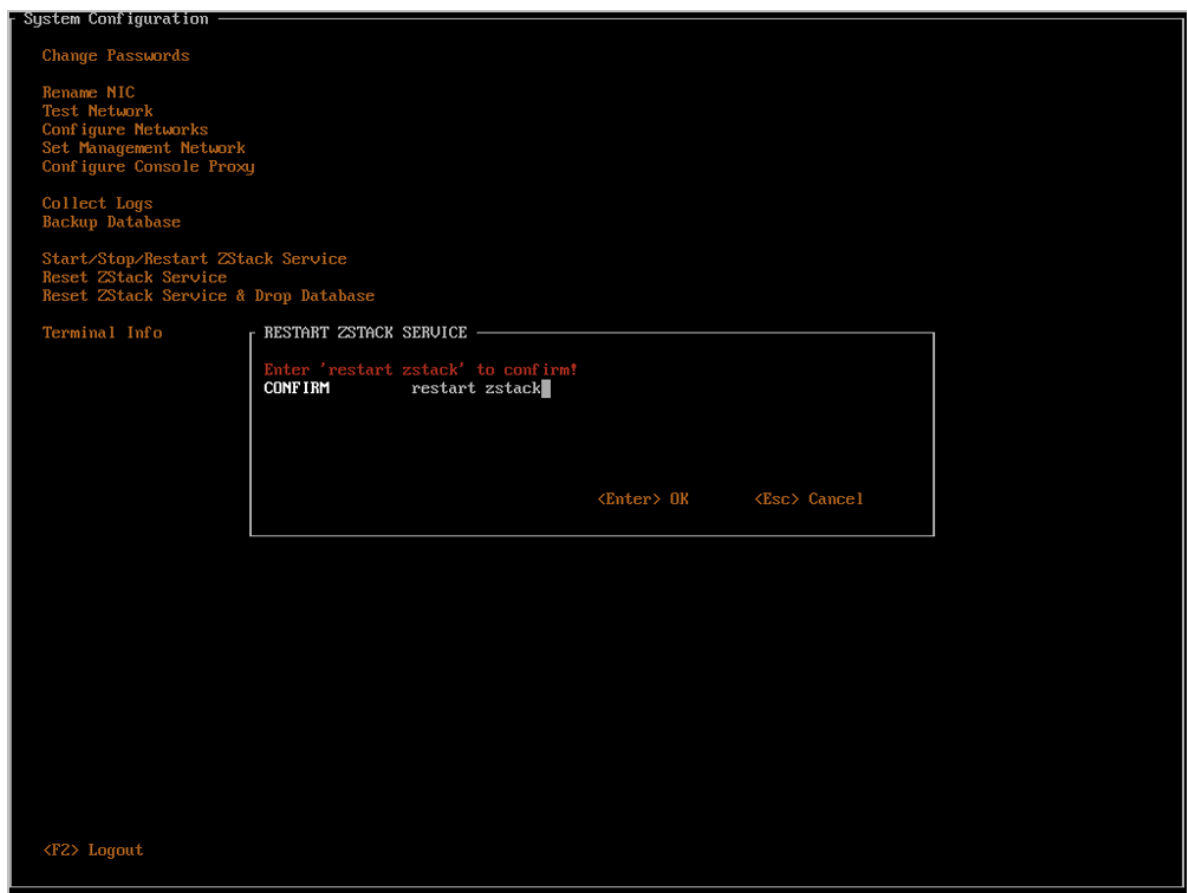
如图 62: 关闭ZStack服务所示：

图 62: 关闭ZStack服务



如图 63: 重启ZStack服务所示：

图 63: 重启ZStack服务



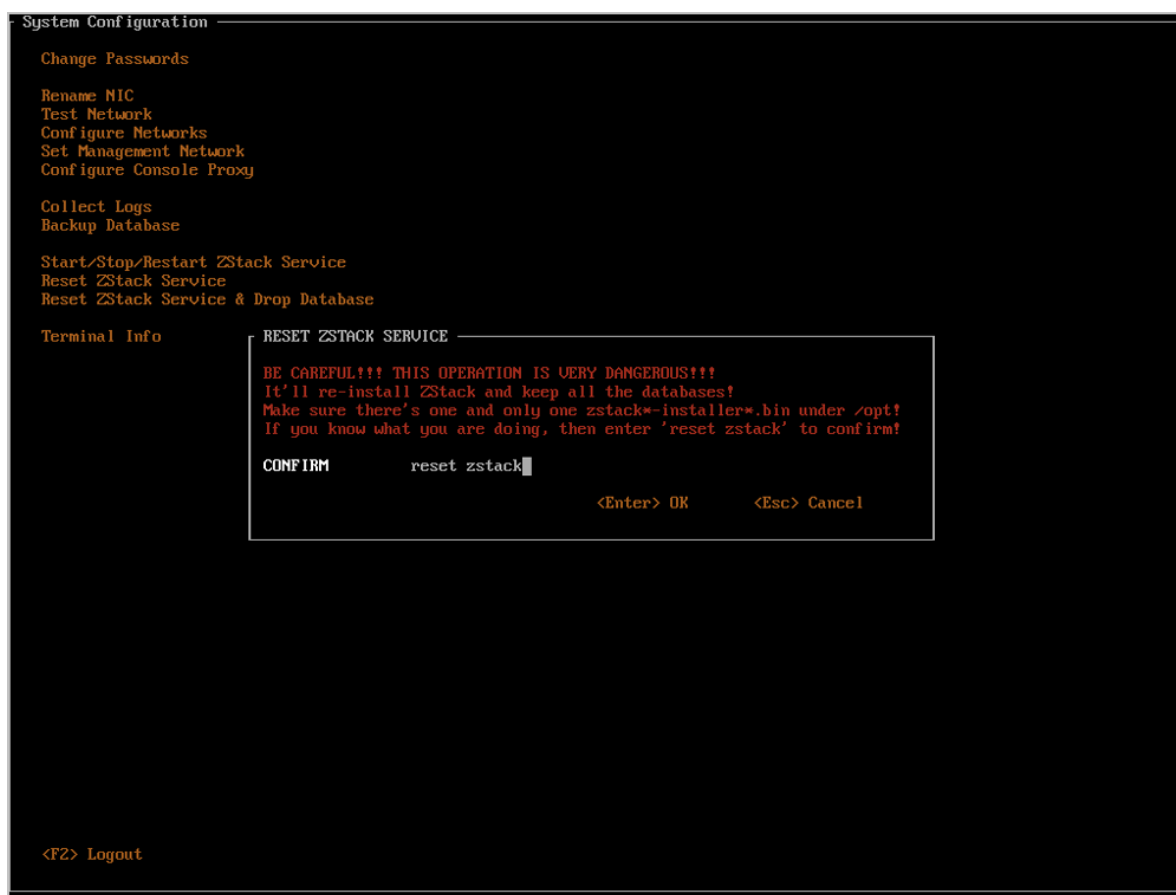
10. 重装ZStack服务

将光标移动至**Reset ZStack Service**，按下回车，即可进入重装ZStack服务窗口。

该操作属于极其危险的操作，需要用户输入**reset zstack**，才能回车确认。

如图 64: 重装ZStack服务所示：

图 64: 重装ZStack服务



11. 重装ZStack服务并删除数据库

将光标移动至**Reset ZStack Service&Drop Database**处，按下回车，即可进入重置ZStack和数据库窗口。

该操作与上一操作一样属于极其危险的操作，而且在重装ZStack的同时还会清空已有的数据库，需要用户输入**reset zstack drop database**，才能回车确认。



注：请谨慎使用此功能！

如图 65: 重装ZStack服务并删除数据库所示

图 65: 重装ZStack服务并删除数据库



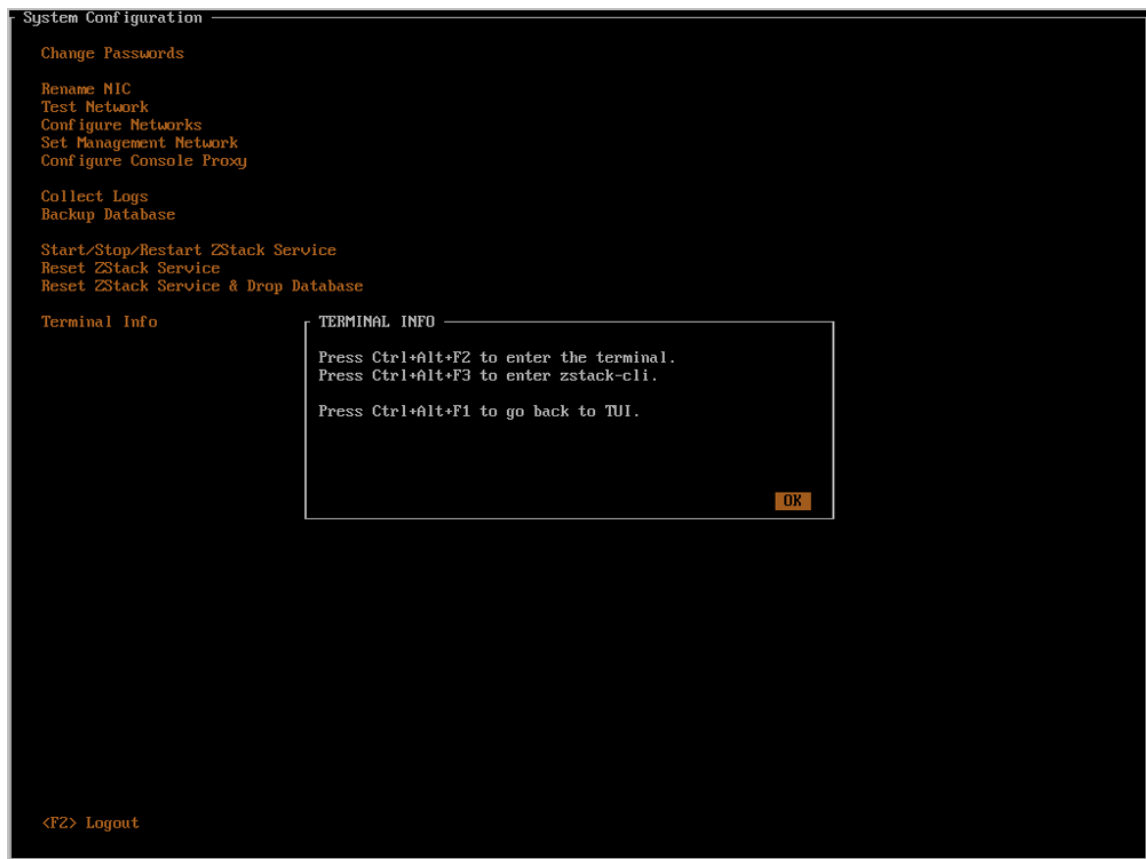
12. 进入终端

ZStack TUI为用户保留了进入终端的入口。将光标移动至**Terminal Info**处，回车即可看到入口信息。

- 按下**Ctrl + Alt + F2**可以进入保留终端，用户可以在里面执行常规命令，但是请谨慎使用，以免对系统造成破坏，影响ZStack服务运行。
- 按下**Ctrl + Alt + F3**可以进入**zstack-cli**命令行界面。
- 任何时候都可以通过按下**Ctrl + Alt + F1**返回ZStack TUI。

如图 66: 进入终端提示所示

图 66: 进入终端提示



重启/关机

在主界面按下**F12**即可进入重启/关机界面。

用户选择将光标移动至**Reboot**或**Poweroff**按钮，回车即可进入确认界面。

如图 67: 重启/关机所示：

图 67: 重启/关机



两种操作都需要用户根据提示输入**REBOOT**或**POWEROFF**才可以回车确认，以免误操作。

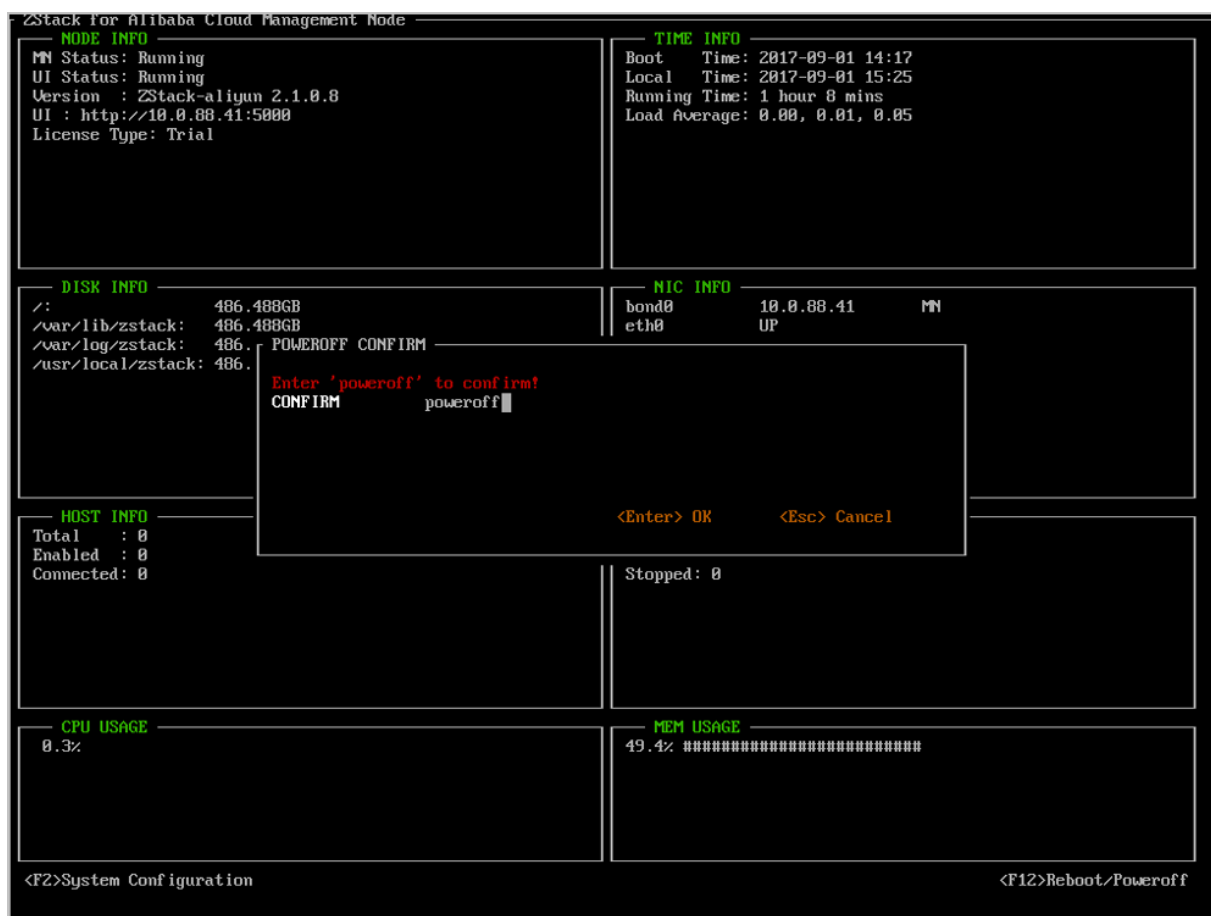
如图 68: 重启确认所示：

图 68: 重启确认



如图 69: 关机确认所示：

图 69: 关机确认



5.2.2 ZStack for Alibaba Cloud计算节点模式

如果用户选择计算节点模式，重启后会自动安装ZStack for Alibaba Cloud计算节点，安装完成后将自动进入TUI。



注：部分场景下，需要all in one的模式来搭建ZStack for Alibaba Cloud，这时应选用ZStack for Alibaba Cloud管理节点模式安装。

计算节点TUI主界面

计算节点TUI拥有管理节点TUI的部分功能，可以视为精简版的管理节点TUI，使用方法与[ZStack for Alibaba Cloud管理节点模式](#)相同。

- **VM_INFO :**

列出了当前计算节点所运行的云主机信息，包括UUID、CPU核心数、内存容量和允许状态等。

- 其他信息模块与管理节点意义相同，不再赘述。

如图 70: 计算节点TUI主界面所示：

图 70: 计算节点TUI主界面

```
ZStack for Alibaba Cloud Compute Node
NODE INFO
Running Time: 1 min 15 secs
Load Average: 0.24, 0.10, 0.04
CPU (4Core): 0.00%
MEM (7.66GB): 2.20%

DISK INFO
/: 487.884GB
/var/log/zstack: 487.884GB

NIC INFO
bond0 10.0.182.81
eth0 UP

UM INFO
-----
UUID CPU MEM STATE
-----
-----

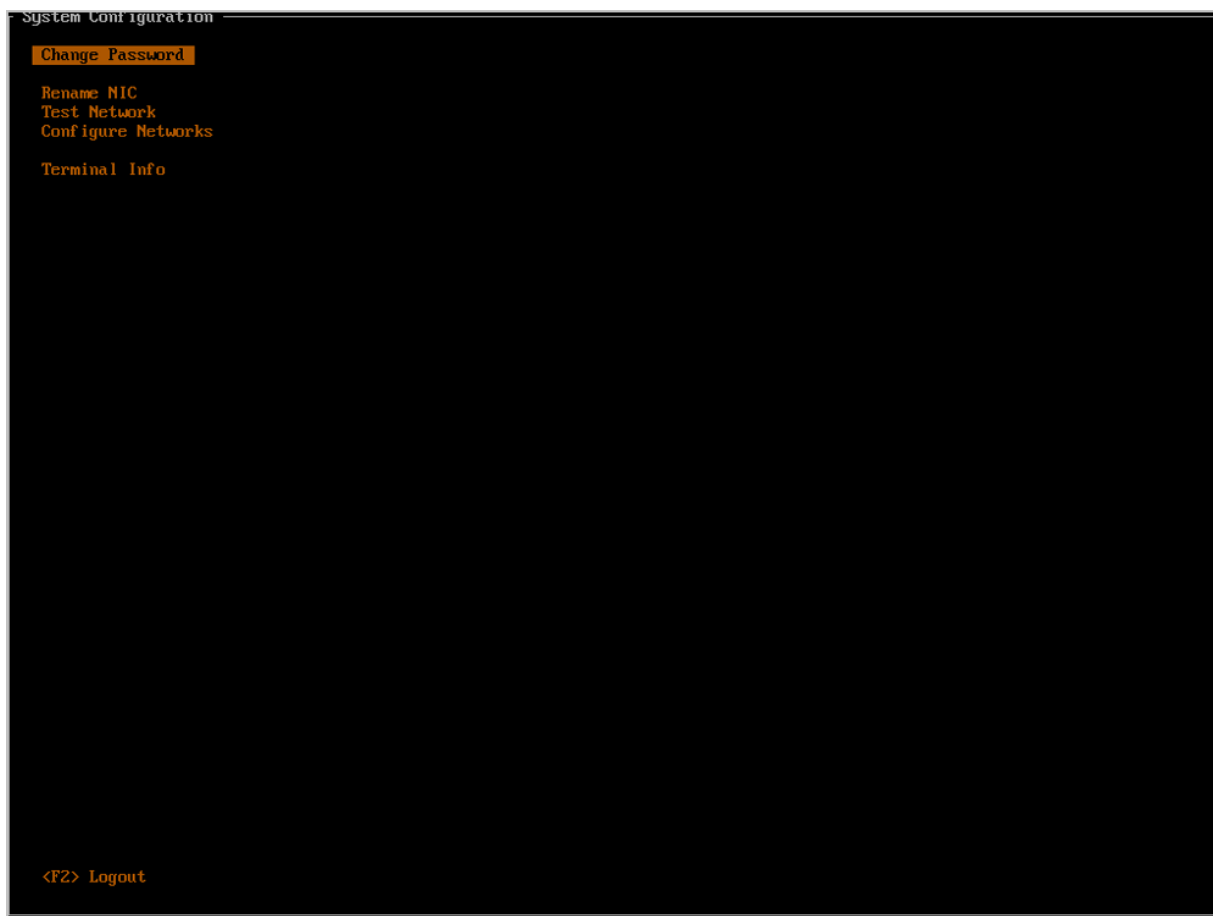
<F2>System Configuration <F12>Reboot/Poweroff
```

系统配置

计算节点系统配置与管理节点系统配置相比，仅拥有其中部分功能条目的配置。

如图 71: 系统配置所示：

图 71: 系统配置



补充说明：管理节点可以添加计算节点的个数在技术上无限制，跟license有关。程序模拟过100万个计算节点。

5.2.3 ZStack for Alibaba Cloud OCFS2存储节点模式

如果用户选择OCFS2存储节点模式，重启后会自动安装ZStack for Alibaba Cloud OCFS2存储节点，安装完成后将自动进入TUI。



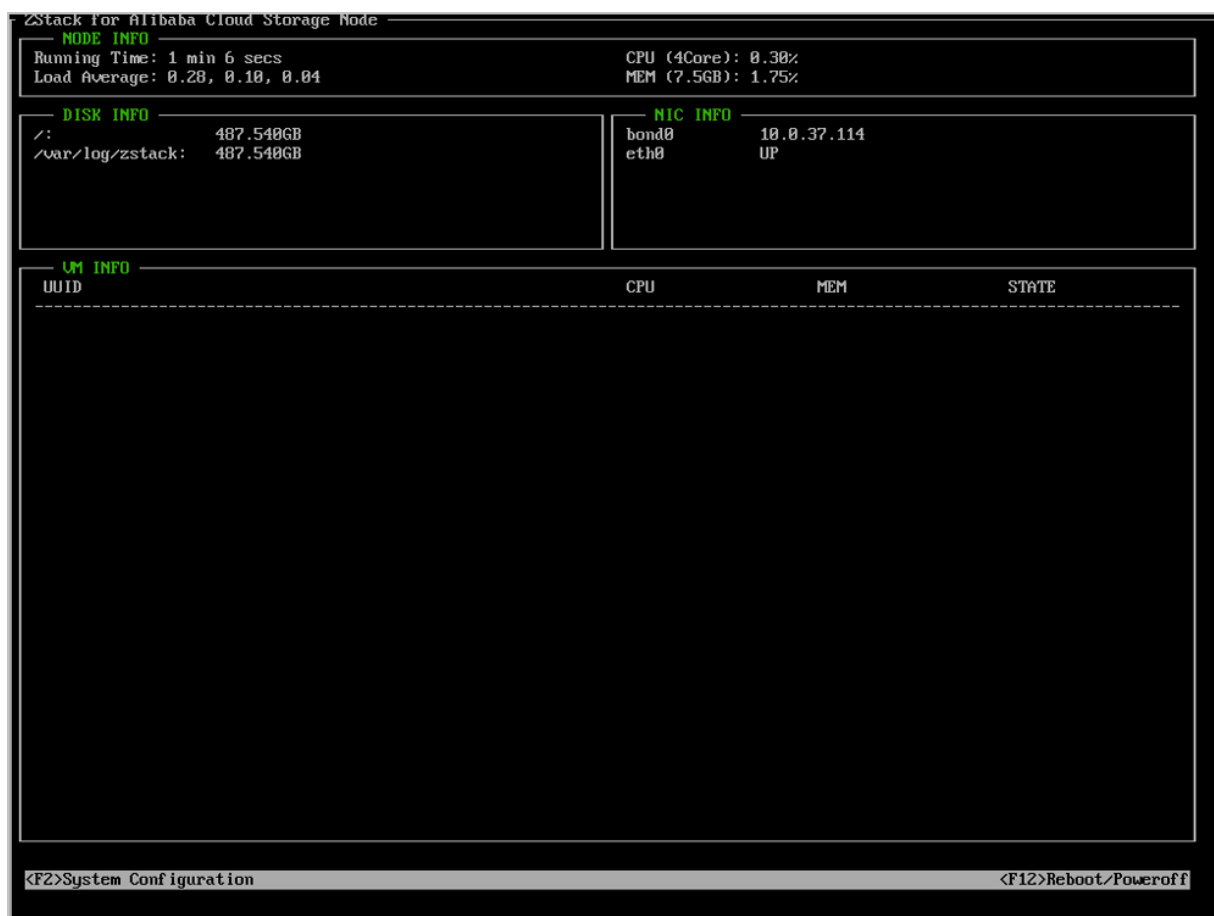
注：部分场景下，需要all in one模式来搭建，这时应选用ZStack for Alibaba Cloud管理节点模式安装。

OCFS2存储节点TUI主界面

存储节点TUI与计算节点TUI极为相似，使用方法与[ZStack for Alibaba Cloud计算节点模式](#)相同。

如图 72: [OCFS2存储节点TUI主界面](#)所示：

图 72: OCFS2存储节点TUI主界面

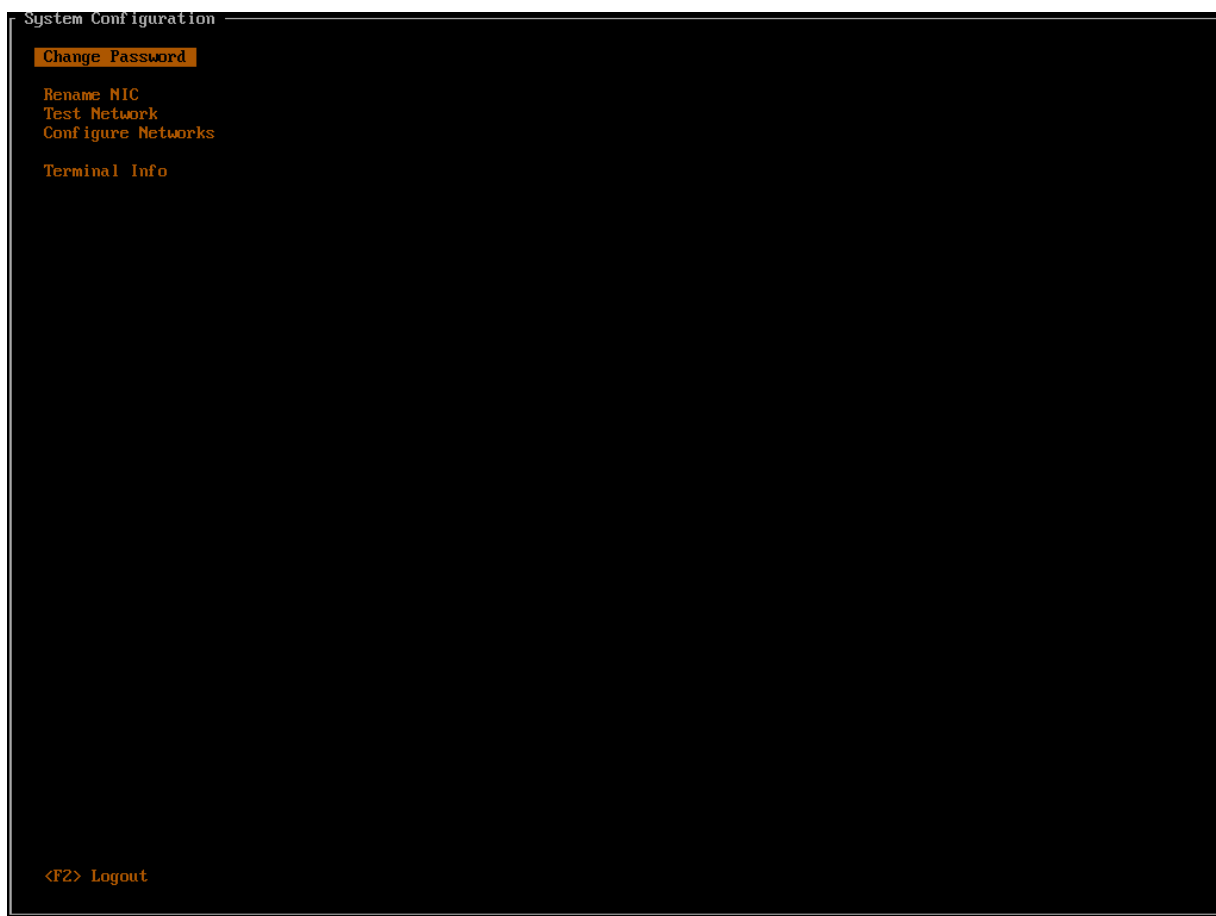


系统配置

存储节点系统配置与管理节点系统配置相比，仅拥有其中部分功能条目的配置。

如图 73: 系统配置所示：

图 73: 系统配置

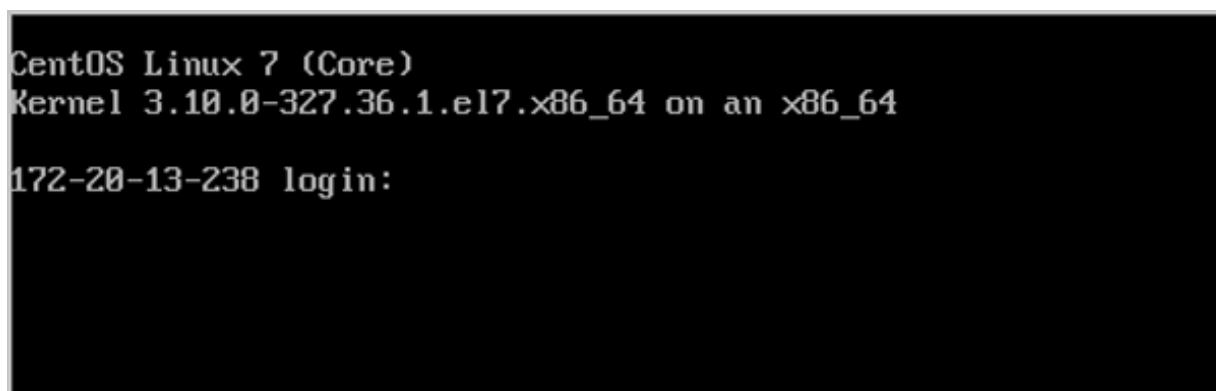


5.2.4 ZStack for Alibaba Cloud专家模式

背景信息

如果用户选择专家模式，重启后会进入终端界面，如[图 74: 终端界面](#)所示：

图 74: 终端界面



安装完专家模式后，用户可根据实际场景需求，自定义安装所需环境。

5.3 管理ZStack for Alibaba Cloud

操作步骤

1. 在首次安装后，系统将自动启动ZStack for Alibaba Cloud服务。
2. 管理节点重启后，ZStack for Alibaba Cloud服务将自动开机自启。
3. 在管理节点因维护或其他异常原因停止服务后，需手动启动服务。

启动ZStack for Alibaba Cloud服务的方法为：

```
[root@localhost ~]#zstack-ctl start
#此命令将同时启动管理节点和WEB UI服务
```

4. 用户可以使用zstack-ctl status命令查看ZStack for Alibaba Cloud管理节点相关服务的运行状态。

```
[root@172-20-14-154 ~]# zstack-ctl status
ZSTACK_HOME: /usr/local/zstack/apache-tomcat/webapps/zstack
zstack.properties: /usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/
zstack.properties
log4j2.xml: /usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/log4j2.xml
PID file: /usr/local/zstack/management-server.pid
log file: /usr/local/zstack/apache-tomcat/logs/management-server.log
MN status: Running [PID:18418]
UI status: Running [PID:19215] http://172.20.14.154:5000
version: 2.2.0 (ZStack 2.2.0.67)
```

5. 用户也可以使用zstack-ctl ui_status命令单独查看Web UI服务状态。

```
[root@172-20-12-20 ~]# zstack-ctl ui_status
UI status: Running [PID:19215] http://172.20.14.154:5000
```

6. 在使用过程中如需重启管理节点服务，则需执行：

```
zstack-ctl restart_node
```

7. 在使用过程中不建议全部停止及重启所有服务。如果确需重启所有服务，可执行以下命令进行重启：

```
zstack-ctl stop && zstack-ctl start
```

5.4 升级ZStack for Alibaba Cloud

ZStack for Alibaba Cloud支持增量升级和离线升级两种方案。

增量升级

相比离线升级方案（即：用户需下载相应版本的ISO并升级本地源，然后升级ZStack for Alibaba Cloud），增量升级方案，用户只需要下载最新的ZStack for Alibaba Cloud安装包，执行升级安装，该安装包会自动检测ISO版本。

1. 在线升级ZStack for Alibaba Cloud之前，请管理员准备好以下必要的软件包，且均存放在管理服务目录/opt/下。

- ZStack for Alibaba Cloud安装包
 - 文件名称：ZStack_Alibaba_Cloud-installer-2.2.x.bin
 - 下载地址：点击[这里](#)

2. 在线升级ZStack for Alibaba Cloud之前，管理员需对数据库进行备份。

3. 管理员执行以下命令升级ZStack for Alibaba Cloud管理服务。

```
[root@zstack-1 opt]# bash ZStack_Alibaba_Cloud-installer-2.2.x.bin -u
```

4. 执行升级安装，该安装包会自动检测ISO版本：

- 如果检测到ISO版本过低，会自动将本地源同步到最新，然后开始ZStack for Alibaba Cloud的安装，如图 75: 自动将本地源同步到最新所示。

图 75: 自动将本地源同步到最新

```

1. Check Repo Version:
  Check local repo version: ... NOT MATCH
  Prepare repo files for syncing: ... PASS
  Install necessary packages: ... PASS
  Test network connection: ... PASS

zstack-local          | 3.6 kB  00:00:00
(1/2): zstack-local/group_gz | 2.1 kB  00:00:00
(2/2): zstack-local/primary_db | 1.1 MB  00:00:00
zstack-online-base    | 3.6 kB  00:00:00
zstack-online-base/group_gz | 2.1 kB  00:00:00
zstack-online-base/primary_db | 1.1 MB  00:00:01
zstack-online-base/group | 19 kB  00:00:00
zstack-online-ceph    | 2.9 kB  00:00:00
zstack-online-ceph/primary_db | 23 kB  00:00:00
zstack-online-uek4    | 2.9 kB  00:00:00
zstack-online-uek4/primary_db | 770 kB  00:00:00
zstack-online-galera   | 2.9 kB  00:00:00
zstack-online-galera/primary_db | 19 kB  00:00:00
zstack-online-gluster  | 2.9 kB  00:00:00
zstack-online-gluster/primary_db | 29 kB  00:00:00
  
```

- 如果在同步本地源过程中报错，将会提示用户采用离线升级方案，如图 76: 同步本地源过程中报错所示。

图 76: 同步本地源过程中报错


```
INSTALLATION

1. Check Repo Version:
  Check local repo version: ... NOT MATCH
  Prepare repo files for syncing: ... PASS
  Install necessary packages: ... PASS
  Test network connection:
  FAIL

Reason: The current local repo is not suitable for ZStack-aliyun installation.
Syncing local repo with repo.zstack.io has been failed too.
Please download proper ISO and upgrade the local repo first.
```



注:

- 整个增量升级过程需在联网状态下进行。

离线升级

1. 离线升级ZStack for Alibaba Cloud之前，请管理员准备好以下必要的软件包，且均存放在管理服务目录/opt/下。

- ZStack for Alibaba Cloud定制版ISO
 - 文件名称：ZStack_Alibaba_Cloud-x86_64-DVD-2.2.x.iso
 - 下载地址：点击[这里](#)
- ZStack for Alibaba Cloud安装包
 - 文件名称：ZStack_Alibaba_Cloud-installer-2.2.x.bin
 - 下载地址：点击[这里](#)
- 升级脚本
 - 文件名称：zstack-upgrade
 - 下载地址：点击[这里](#)



注：软件下载后，需通过MD5校验工具核对校验码，确认与发行信息一致。

2. 离线升级ZStack for Alibaba Cloud之前，管理员需对数据库进行备份。

3. 管理员执行以下命令升级ZStack for Alibaba Cloud管理服务。

```
#离线升级有两种方式:
# 1. 升级本地仓库和管理服务
[root@zstack-1 opt]# bash zstack-upgrade ZStack_Alibaba_Cloud-x86_64-DVD-2.2.x.iso
# 2. 如果先升级本地仓库再升级管理服务:
[root@zstack-1 opt]# bash zstack-upgrade -r ZStack_Alibaba_Cloud-x86_64-DVD-2.2.x.iso
```

```
[root@zstack-1 opt]# bash ZStack_Alibaba_Cloud-installer-2.2.x.bin -u
```



注:

升级ZStack for Alibaba Cloud只需升级管理节点，计算节点会自动完成升级。

5.5 裸机管理

裸机管理表示使用PXE服务实现物理裸机的远程自动系统部署，并对裸机进行远程电源管理。

裸机管理相关操作定义如下：

- 安装服务：在管理节点指定网卡上部署PXE服务，配置DHCP和TFTP服务，可远程引导安装裸机系统。
- 裸机安装：配置裸机的IPMI信息，并使用IPMI进行远程电源管理，并配置裸机安装系统。
- DHCP侦听网卡：管理节点上连通裸机安装网络的网卡设备编号，例如eth0，要求此网卡连接到裸机安装网络，且已配置IP地址。
- DHCP起始IP/结束IP：用于定义遍历DHCP服务的IP地址范围，可留空不填，系统将根据DHCP侦听网卡IP侦测并过滤已用的IP地址作为IP范围。
- 停用PXE服务：表示停止使用此PXE服务，停用后，会导致裸机安装服务不可用，需启用才可继续使用PXE服务。
- 添加裸机：添加物理裸机的IPMI信息，使用IPMI进行远程管理裸机，需提前配置IPMI，输入IPMI的用户名、密码，裸机的安装网络与管理节点的DHCP监听网卡应在同一个二层，可获取DHCP服务。
- 安装系统：给裸机安装部署操作系统，需输入以下字段：
 1. root密码：提供裸机打算预置的root系统密码；
 2. 启用VNC：启用VNC界面可用来查看控制安装过程，建议启用。
 3. 启用无人值守：启用无人值守安装，将所有硬盘均划归为单个LVM分区，把大部分容量分给根分区。如果需要定制分区，可选否，并在VNC界面自定义分区。
 4. PXE启动网卡：表示选择哪个网卡来安装部署系统，多网络的情况下，需确保此网卡已接入交换机且与管理节点提供PXE服务的网卡互通。
 5. 配置网卡：可对系统当前网卡设置静态IP地址。对应的设置应与物理环境一致，否则网卡会工作异常。
 6. 配置网卡绑定：可对系统内多个网卡实现绑定，实现模式1或模式4的绑定。其中模式1表示平衡轮询策略，模式4表示动态链路聚合模式。模式4需要在安装完系统后，在交换机端对对应的链路设置链路聚合。

- 重装系统：重装系统表示对当前裸机重新安装系统，当前系统内的所有数据会重新覆盖，需谨慎操作。

PXE服务使用时需注意：

- 会在管理节点部署PXE服务，配置TFTP和DHCP服务，远程引导安装裸机系统；
- 管理节点用于裸机安装的二层网络，不能存在其他DHCP服务器，以免冲突；
- 管理节点上提供PXE服务的网卡，需配置IP地址；
- 管理节点需存在IPMI网络，可通过IPMI远程管理物理机；
- 裸机的启动项是硬盘优先，系统会自动配置PXE启动，安装系统后会从硬盘引导；
- 裸机的安装网络与管理节点的DHCP监听网卡应在同一个二层，可获取DHCP服务。

6 登录

ZStack在安装部署完毕后，提示的URL一般为管理节点IP附带5000端口，假定管理节点IP地址为172.20.1.20，那么使用UI进行控制的URL需要输入http://172.20.1.20:5000

建议使用chrome或firefox打开浏览器。

系统首次登录时，默认账户名：admin，默认初始密码：password。

默认登录时效为2小时，超时需重新登录。超时时间可以在全局设置界面进行配置。

ZStack登录方式支持：

- 账户登录：需输入账户名和账户密码。
- 用户登录：需输入账户名和用户名和用户密码。
- LDAP登录：需输入LDAP UID和LDAP密码。LDAP需提前设置LADP服务器和LDAP账户绑定。



注：登录失败需检查管理节点服务，确保管理服务和UI服务均正常运行。

7 云资源池

7.1 云主机

创建云主机需要的最重要的三个输入：

1. 计算规格：

用于确定云主机的CPU、内存大小，在分配资源前，ZStack会筛选那些拥有这些资源的物理机作为候选来创建云主机。

2. 镜像：

用于确定使用哪个镜像来创建云主机。选择镜像时：

- 假如此镜像对应的镜像服务器，为Imagestore或SFTP镜像服务器，则ZStack会自动选择那些挂载了本地存储、NFS或者SharedMountPoint主存储的集群作为备选集群。
- 假如此镜像对应的镜像服务器，为Ceph镜像服务器，则ZStack会自动选择挂载了此Ceph镜像服务器对应的Ceph主存储的集群。
- 假如此镜像对应的镜像服务器，为Fusionstor镜像服务器，则ZStack会自动选择挂载了此Fusionstor镜像服务器对应的Fusionstor主存储的集群。

3. 网络：

用于指定云主机的网络。在选择镜像后，待创建的云主机的可用集群已经确定下来，ZStack会根据此集群挂载的网络来提供可选网络。



注：

- 在确定完毕以上输入后，创建此云主机需要的各种资源已经确定，此时可选的资源被筛选为使用此镜像对应的镜像服务器，接着根据此镜像服务器可选择的集群，根据此集群确定的相关网络。
- 如果有满足条件的多个集群挂载已选择的网络，那么高级选项还可以继续筛选相应的集群、主存储、物理机等资源。

如果存在多主存储时，创建云主机的资源分配策略如下：

- 多个本地主存储，创建云主机时，可指定任意主存储；若未指定，系统自动选择可用容量最充足的本地主存储。
- 多个共享主存储，创建云主机时，可指定任意主存储，若未指定，系统随机分配可用主存储。

- 本地主存储+NFS、本地主存储+Shared Mount Point时，创建云主机时，可指定任意主存储，如果带数据云盘，则需指定数据云盘所使用的主存储；如果不指定主存储，则使用本地主存储来创建云主机。

云主机相关操作定义：

- 启动：

启动云主机，将停止状态的云主机启动起来，使其正常运行。云主机启动默认会使用上次所在的物理机启动，如果此云主机所在集群只挂载了共享存储（一个或多个），则此云主机可在使用此主存储的可用物理机上启动；如果此云主机所在集群只挂载了本地存储（一个或多个），则此云主机只能在上次所在的物理机上启动；如果此云主机所在集群挂载了不同类型的主存储（本地+NFS、本地+Shared Mount Point），如果此云主机使用了本地存储的根云盘或数据盘，则此云主机只能在上次所在的物理机启动，如果此云主机只使用了共享存储（也可带数据云盘），则此云主机可在使用此主存储的可用物理机上启动。

启动云主机失败，需检查以下资源：

- 其所在集群内是否有物理机连接并启用，如果没有可用连接的物理机，会提示找不到可用物理机；
- 其所在集群内物理机的资源（CPU、内存）是否充足，如果所有的物理机各自可用的CPU、内存资源不足以支持此云主机的CPU 内存大小，则会提示找不到可用物理机；
- 其所在集群是否挂载了二层网络，如果集群未挂载二层网络，则此云主机上也不会存在三层网络，则会提示此云主机无网络，无法启动；
- 其所在集群对应的二层网络是否存在三层网络，如果此集群挂载的二层网络均无三层网络，则此云主机上也不会存在三层网络，则会提示此云主机无网络，无法启动；
- 其所在集群是否挂载了主存储，如果集群未挂载主存储（一般发生在卸载主存储后），则会提示此云主机的根云盘文件找不到，需挂载回原本的主存储才可正常启动；
- 其所在集群的主存储可用容量是否充足（创建云主机时判断），如果主存储无法提供云主机需要的容量，则会提示容量不足。

- 停止：

停止云主机，将运行状态的云主机停止起来，使其正常关机。

- 重启：

重启云主机，将运行状态的云主机重新启动，重启过程中会更新其全部配置（包括启动顺序、控制台密码、控制台模式、USB设备重定向、屏幕数量控制）。

- 关闭电源：

关闭电源，将运行状态的云主机直接断电，强制关闭，此操作存在丢失数据的风险，需谨慎操作。

- 暂停：

暂停云主机，云主机继续消耗内存或资源，但是暂时不调度hypervisor，可以恢复到运行状态，也可删除。

- 恢复（Resume）：

从暂停状态恢复正常运行，云主机内可继续运行。

- 删除（destroy）：

如果云主机运行状态，则将其关闭电源并置为已删除状态；如果云主机处于其他状态，则将其直接置为已删除状态。

- 恢复（Recover）：

将VM从已删除状态恢复为停止状态。

- 彻底删除：

彻底删除，会将云主机从ZStack彻底移除，并且移除根云盘真实占用的存储空间。

- 打开控制台：

通过VNC方式访问云主机，要求浏览器所在电脑可以访问控制台代理IP。如果打开控制台返回成功，但是无内容，需确保控制台代理指定的IP地址可以被用户的客户机（笔记本电脑）正常访问，如有必要，需修改控制台代理IP，并重启MN服务。



注：修改步骤如下：

1. 修改/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties里面中consoleProxyOverriddenIp为外部可访问的IP地址。
2. 重启管理节点服务，zstack-ctl restart_node

- 本地存储迁移：

主要是指冷迁移，主要迁移根云盘或数据云盘，冷迁移前需卸载相关的数据云盘，本地存储也支持热迁移，需全局打开支持，且Windows云主机不支持，本地存储在热迁移前也需要先卸载数据云盘。

- 共享存储迁移：

为在线实时迁移，主要拷贝内存及CPU相关寄存器状态，支持跨集群迁移，只需目的集群与原本集群均挂载云主机使用的网络和主存储，多个不同主存储的情况下，加载了不同类型存储后，需要卸载云盘后才可迁移。

- 创建镜像

将云主机封装成镜像，可以使用此镜像再创建更多云主机。

- 在线创建：支持云主机运行时，镜像仓库和Ceph存储在线创建镜像。
- 关机创建：支持云主机关机后，各种镜像服务器均支持创建镜像。

- 创建快照：

针对根云盘或数据云盘在指定的时间点创建的数据映像，可以在数据丢失时，有效恢复数据。

- 在线快照：支持云主机运行时，镜像仓库和Ceph镜像存储支持增量快照备份。
- 关机快照：支持云主机关机后，各种镜像服务器均支持创建快照。

- 恢复快照：

将云主机恢复至当前快照点，云盘内相关的数据会恢复至快照点。恢复快照前，需提前备份云盘数据。

- 克隆：

- 只涉及根云盘的克隆，用户在克隆前对云主机根云盘安装的程序或进行的配置都会复制到新克隆出的云主机。
- 目前镜像仓库和Ceph存储支持在线克隆和关机克隆。
- 克隆会复制此云主机相关的系统标签，例如NeverStop、QGA等系统标签。

- 创建云盘：

从云主机所属集群已挂载的主存储创建数据云盘，在多主存储的情况下，可指定主存储来创建云盘。

- 针对多个本地主存储，默认会选择可用容量大的存储来创建；
- 针对多个NFS主存储，默认会随机选择主存储来创建；
- 针对本地主存储+NFS/Shared Mount Point类型主存储，默认会选择与当前根云盘不在同一个主存储的存储来创建。

- 加载云盘/卸载云盘：

- 加载云盘：可选择云主机所属集群已挂载的主存储上可用云盘来加载。
- 卸载云盘：将云盘从云主机卸载，本地存储冷迁移和存储迁移，需要先卸载相关数据云盘。

- 加载ISO/卸载ISO：

加载或卸载ISO，需要提前添加ISO镜像，加载ISO镜像后，可在云主机内识别并挂载。如果打算进行系统修复或系统重新安装，可以加载ISO，调整启动顺序以ISO优先，并重启云主机，用ISO引导系统。

- 加载SSH Key/删除SSH Key：

SSH Key注入需注意：

- SSH Key注入需要云主机提前安装cloud-init,CentOS可执行yum install cloud-init 来安装Cloud-init，输入指定主机的对应用户的公钥，公钥由ssh-keygen生成，存放在~/.ssh/id_rsa.pub。
 - 安装cloud-init后，ssh密码认证默认关闭，需将/etc/cloud/cloud.cfg的ssh_pwauth选项设置为1才能开启。
 - 添加SSH Key时，如果云主机之前已注入过ssh-key，则需在云主机中手动执行：`rm -rf /var/lib/cloud/instances`，以清理先前配置并重启云主机。
 - 删除SSH Key只删除系统中记录的SSH Key信息，并不从云主机中删除相关的Key信息，并不移除已经生效的Key，如果需要再次注入，需手动清理/var/lib/cloud/instances里面的信息，如果需要删除已经生效的Key，则需手动清理云主机里~/.ssh/authorized_keys的Key。
- 设置高可用：

设置为Neverstop的云主机在资源允许的条件下永不停机。如果云主机的云盘使用了共享存储，则在资源可用的情况下，云主机会始终保持开机，即使停止也会再次启动，假定其所在物理机异常断电关机，此云主机也会在其他可用物理机再次启动。如果云主机的云盘使用了本地存储，在其物理机异常关机后，此云主机不能再次启动，直到物理机恢复可用连接状态。

- 修改计算规格/修改CPU、内存：

修改云主机的CPU、内存、磁盘QoS、网络QoS等，支持：

- 在线修改：需全局开启NUMA模式，再重启或者新建的云主机才能生效。不支持在线修改Windows云主机的计算规格。在线修改计算规格，只能增加CPU、内存的配置，不能降低配置。
 - 关机修改：云主机关机后，可修改云主机的配置信息。
- 设置启动顺序：

在加载ISO后，修改云主机的启动顺序，可用于系统修复或系统重新安装，设置启动顺序

以cdrom优先，需重启云主机生效，只生效一次，下次再重启后会默认再从硬盘引导。用户在第一次创建云主机选择ISO进行安装时，无需选择启动顺序，系统会自行从ISO引导。

- 启动（指定物理机）：

从可用的物理机列表选择指定的物理机来启动云主机，适用于共享存储。使用本地存储的云主机，只能使用根云盘所在的物理机启动。

- 设置控制台密码/取消控制台密码：

设置控制台密码后，打开VNC控制台则需要正确输入控制台密码后才可访问控制台。取消控制台密码后，访问此云主机的控制台则无须输入密码。设置和取消控制台密码需要云主机重启（靠ZStack重启）才会生效。

- QGA选项

开启前需确保云主机已安装并运行qemu-guest-agent；开启后云主机默认支持在线修改密码。

- 在线修改密码：

需要打开QGA选项，且云主机需安装并运行qemu guest agent，支持Linux和Windows云主机，需确保输入的用户名已存在。

- 重置云主机：

将停止状态的云主机恢复到镜像模板的初始状态，云主机内目前已修改的数据会被擦除。重置云主机前，需提前备份云盘数据。

- QoS设置：

限制云主机云盘、网络上下行带宽的设置，无须重启立刻生效，过低的QoS可能导致云主机工作异常。

- 更改所有者：

变更云主机的所有者给其他普通账户。

- 设置RDP模式：

设置VDI场景下的RDP模式，默认以RDP模式打开控制台，设置后立刻生效。

- USB重定向

USB重定向主要用于VDI，将客户端插入的USB设备重定向至云主机，修改后，需重启生效。

- 系统扩容：

将云主机的根云盘容量扩容增加。使用LVM分区的Linux云主机可使用lvextend 直接扩容系统分区。Windows云主机也可以右击磁盘管理进行系统扩容。

- 存储迁移：

将停止状态的云主机或云盘跨Ceph存储迁移至其他Ceph存储。需先卸载掉相关的数据云盘。目前支持Ceph存储。

- 切换控制台模式：

控制台模式支持vnc和spice，vnc可以直接在浏览器上打开控制台，spice协议需spice客户端打开，支持两种模式切换，重启后生效。

- Userdata注入：

Userdata可实现脚本化注入，可参考以下示例：

创建了用户test, 使用了ssh-key，开机写入文件/etc/hosts，创建了/tmp/temp目录，并创建文件写入内容。注意严格缩进，缩进不正确会导致注入失败。

```
#cloud-config
users:
  - name: test
    shell: /bin/bash
    groups: root
    ssh-authorized-keys:
      - ssh-rsa AAAAB3NzaC1yc2E root@10-0-0-18

bootcmd:
  - echo 192.168.1.130 us.archive.ub.com >> /etc/hosts
  - mkdir /tmp/temp

write_files:
  - path: /tmp/ZStack_config
    content: |
      Hello,ZStack!
```

- 修改平台：

修改云主机的平台，可选择Linux、Windows、Windowsvirtio、Other、Paravirtualization模式，其中Linux、Windowsvirtio、Paravirtualization支持Virtio类型的设备驱动，Windows、Other使用Qemu模拟的设备。需重启后生效。

- 修改屏幕数量：

VDI场景下云主机可接入的屏幕数量，默认为1，修改后，需重启生效。

- 加载网卡：

从可用的三层网络列表，选择可用网络加载至云主机。

- 卸载网卡：

将云主机的网卡卸载，如果云主机没有任何网卡，停止后，将无法启动，需至少加载一个网卡。

- 设置静态IP：

设置云主机的静态IP，从可用的IP地址段选择可用IP指定给云主机使用，需避免IP冲突。

- 加载GPU设备：

将物理机的GPU设备加载给云主机使用，使得云主机独占使用GPU设备，拥有GPU设备的云主机不支持在线迁移。

- 卸载GPU设备：

将云主机上已独占的GPU设备卸载。

- 定时任务：

基于定时器，创建定时任务，支持定时关机、定时开机、定时重启、定时快照等任务。

- 报警器：

基于邮箱服务器，对指定资源（CPU、内存、磁盘IO、网络IO）在资源使用率在指定时间段内达到阈值后，会触发报警邮件。

- 恢复根云盘快照、设置云主机静态IP、重置云主机、系统扩容、存储迁移需停止云主机为前提。

- 云主机的快照链如果过长，可能会导致云主机IO性能下降。

- 如果SSH Key注入和Userdata同时注入，Userdata生效，而注入的SSH Key不生效。

- 数据云盘默认采用Virtio设备，需安装相应驱动才可识别使用，Linux各发行版默认已集成Virtio驱动，无须额外安装，Windows需额外安装驱动。

- 共享云盘需要使用virtio scsi类型云盘，可使用ocfs2文件系统，各挂载的云主机可同时进行读写。Ceph存储支持共享云盘。

- GPU透传功能需要物理机支持IOMMU及BIOS打开VT-d选项。

- 云主机多网卡的情况下，修改默认网卡，需要在云主机内部重启网络服务后生效。

7.2 镜像

镜像表示云主机使用的镜像模板，包括ISO和Image，其中Image支持qcow2和raw。镜像保存在镜像服务器上，首次创建云主机时，会下载到主存储上作为镜像缓存。

平台类型决定了创建虚拟机时是否使用KVM virtio驱动（包括磁盘驱动和网卡驱动），支持以下类型：

- Linux: 使用virtio；

- WindowsVirtio: 使用virtio驱动，请确保镜像已安装virtio磁盘和网卡驱动；
- Paravirtualization: 使用virtio驱动。镜像可以是已安装virtio驱动的任何操作系统；
- Windows：不使用virtio，使用Qemu模拟设备。镜像操作系统是未安装virtio的Windows；
- Other：不使用virtio，使用Qemu模拟设备。镜像操作系统可以是任何操作系统。

镜像路径支持两种方法：1. URL，采用指定的URL路径来添加镜像。支持以下两种格式：a.

HTTP/HTTPS 格式，例如：`http://cdn.zstack.io/product_downloads/images/zstack-image.qcow2`

b. 镜像服务器上的绝对路径，支持Sftp镜像服务器和镜像仓库。例如：`file:///opt/zstack-dvd/zstack-image-1.4.qcow2`



注：输入URL时，需确保可被镜像服务器访问，且存在此镜像文件。

2. 本地文件，表示选择当前浏览器可访问的镜像直接上传，支持镜像仓库和Ceph镜像服务器，添加本地文件作为镜像，采用了本地浏览器作为中转上传镜像，需注意请勿刷新或关闭当前浏览器，也不可停止管理节点服务，否则会添加失败。

镜像相关的操作定义：

- 启用：

将停用的镜像启用，可作为候选来创建云主机。

- 停用：

将镜像停用，此镜像不可作为候选使用来创建云主机。

- 导出：

将镜像仓库的镜像导出，可直接下载，导出的镜像默认存放在镜像仓库目录下的export子目录下。针对sftp镜像服务器，可以直接拷贝镜像所在的地址，进行拷贝导出；针对Ceph镜像服务器，可参考镜像所在的Pool信息使用rbd export进行手动导出。

- 全局共享：

将镜像共享给所有账户，使得所有普通账户均可使用。

- 共享：

将镜像共享给指定的一个或多个账户。

- 全局召回：

将共享的镜像全部召回，不再共享。全局召回不影响单独共享，单独指定共享的镜像，普通账户依然可以使用。

- 召回：

将镜像从指定的用户召回，不再共享。

- 更改所有者：

变更镜像的所有者给其他普通账户。

- 存储迁移：

将Ceph存储里面的镜像跨存储迁移至其他Ceph存储。迁移后，新的Ceph存储可使用此镜像，原本的Ceph存储则不再保留原本镜像。进行存储迁移前，需确保目的Ceph存储容量充足。

- 删除：

将镜像删除，删除后可进行彻底删除或恢复。

- 恢复：

将已删除的镜像恢复为就绪状态。

- 彻底删除：

将已删除的镜像彻底删除，并删除镜像服务器上的真实数据。

- 下载已导出镜像：

将已导出的镜像下载到浏览器对应本地硬盘中。

- 复制URL：

将已导出的镜像的URL拷贝出来，可以添加到其他云平台，也可直接使用wget下载此URL。

- 删除已导出镜像：

将已导出的镜像从镜像仓库删除，不再占用存储容量。

- 修改平台：

修改镜像的平台类型，修改后新建的云主机会以新的类型来决定云主机的设备驱动，Linux、Windows virtio、Paravirtualization采用了virtio的驱动，Windows和Other采用了Qemu模拟设备。

- 修改QGA选项：

打开QGA选项，表示镜像里面已经预装并开机运行qemu-guest-agent，使用此镜像创建的云主机，支持在线修改密码。

- 刷新容量：

刷新镜像在镜像服务器的真实容量。

使用镜像时，需注意：

- 使用file格式添加镜像，暂不支持ceph或fusionstor镜像服务器。
- 输入的URL需可被镜像服务器访问，且存在此镜像文件，如果镜像服务器无法访问输入的URL，则会提示下载镜像失败。
- Qemu guest agent 勾选表示镜像已安装了qemu-guest-agent，创建出的云主机默认支持在线修改密码，添加ISO镜像时，此选项禁用。
- admin 可共享镜像给普通账户，分享后，普通账户可使用此镜像创建云主机。
- 普通账户可添加镜像创建云主机，也可使用admin共享的镜像资源，普通账户可以删除自身添加的镜像，不可删除其他镜像。
- 镜像停用后，此镜像不可作为候选来创建云主机。
- 镜像容量为虚拟容量，表示此镜像的虚拟大小，镜像真实容量为镜像占用的镜像服务器的空间，类似du -sm读取的文件大小。
- 镜像仓库类型的镜像支持导出功能，导出后，可使用提供的URL下载镜像。

7.3 云盘

云盘提供了云主机的扩展盘，用于云主机存储扩展。

云盘相关操作定义及注意事项：

- 启用：

将停用的云盘启用，启用后，可以加载到云主机。针对已经加载的云盘，停用不影响其正常工作，在其卸载后就不可再加载。

- 停用：

将启用的云盘停用，停用后，此云盘不可作为候选加载到其他云主机。

- 加载：

将云主机加载到云主机，云主机运行或关机状态均可加载。

- 卸载：

从云主机将云盘卸载，卸载后可加载到其他云主机。

- 迁移：

本地存储的云盘可以在卸载后，迁移至其他物理机，迁移后，可以加载到目的物理机的其他云主机上。

- 创建快照：

针对云盘在指定的时间点创建数据映像，可以在数据丢失时，有效恢复数据。与云主机根云盘创建快照类似。卸载后，均可创建快照。针对已加载到运行状态的云主机的云盘。

- 更改所有者：

变更云盘的所有者给其他普通账户。

- 存储迁移：

将已卸载的云盘跨Ceph存储迁移至其他Ceph存储。需先卸载掉相关的数据云盘。目前支持Ceph存储。

- 删除：

将已挂载、未挂载或未实例化的云盘删除掉，未实例化云盘删除操作为立即删除，删除后系统不再记录此云盘。

- 恢复：

将已删除的云盘恢复到就绪状态。

- 彻底删除：

将已删除的云盘彻底删除，系统将不再记录此云盘，并删除真实的数据。

- 云盘存储在主存储中，需确保主存储可用可连接状态，云盘才可正常工作。
- 创建云盘时，需要指定云盘规格。
- 创建云盘时，如果未指定主存储，则创建出的云盘为未实例化云盘。
- 未实例化云盘不占用存储空间，只有挂载到云主机后，实例化后会占用真实存储空间。
- 指定主存储后，如果选定的云盘规格超过目前主存储的可用虚拟容量，创建云盘会失败，提示主存储可用容量不足。
- 指定的主存储如果是本地存储，则同时需选择挂载此主存储的集群下可用的物理机
- Virtioscsi类型的云盘表示采用VirtioSCSI总线，云盘属于SCSI类型，支持IO多队列，可以通过唯一识别ID (WWN) 识别。
- Ceph类型的主存储支持共享云盘，共享云盘可被多个云主机识别并同时访问。
- 本地存储的云盘在卸载后，支持迁移到其他物理机上，迁移后，可加载到新物理机上的云主机。
- 可对云盘创建快照，表示在特定时间点上的数据进行备份。可在数据丢失时，有效恢复数据。
- 云盘快照均增量快照，恢复后，快照点当时的数据会恢复并覆盖云盘内所有数据，需谨慎操作。
- 多个主存储环境下，可以指定主存储创建云盘，如果未指定主存储，针对本地主存储，云盘默认会从容量大的主存储创建。针对NFS主存储，云盘默认会随机选择一个主存储创建。针对本地主

存储+NFS/Shared Mount Point类型主存储，默认会选择与当前根云盘不在同一个主存储的存储来创建。

- 针对云盘，可设置QoS进行磁盘带宽限速，需注意限速不可过低，过低的QoS 可能导致IO性能过低。

7.4 计算规格

计算规格定义了云主机的CPU、内存、磁盘网络带宽的数量或大小规格

使用计算规格需注意：

- 计算规格的内存需大于16M，过低的规格会无法启动云主机。
- 磁盘网络带宽均为可选项，限制了过低的带宽，可能导致云主机工作异常。
- 首次创建，可按照默认的1CPU-1G内存来配置。
- 计算规格支持在线修改，需要全局配置打开NUMA支持。不支持在线修改Windows云主机。

7.5 云盘规格

云盘规格定义了云盘的大小规格。

云盘规格支持启用、停用、共享、召回等操作。使用ISO安装云主机时，需指定云盘规格来指定根云盘大小。

8 硬件设施

8.1 区域

区域：ZStack中最大的一个资源定义，包括集群、二层网络、主存储等资源。

区域一般对应了数据中心的一个机房。区域进行了设置了相关资源的相互隔离，不同区域内的子资源是不可见的，不能互相发生关系。

在规划数据中心时，建议：

- 同一个物理二层广播域中的主机应该在同一个区域中，可以被划分为一个或多个集群。
- 物理二层广播域不应该跨域多个区域，而应该映射为单个区域中的L2网络。
- 为虚拟机存储云盘提供磁盘空间的物理存储，也就是主存储，不应该跨越多个区域，而应该被映射为单个区域中的主存储。
- 一个数据中心可以一个或多个区域。

一个区域可以挂载一个或多个备份存储，但是UI界面为便于管理和维护，限制了一个区域与备份存储的关系。UI界面上，添加镜像服务器时，默认会挂载到当前区域。

区域在进行操作时需注意：

- 禁用区域：禁用区域会导致此区域下的所有集群及集群内所有物理机均被禁用，此时本区域的所有集群内均不可再申请新资源，需再次启用才可申请。
- 启用区域：启用区域会导致此区域下的所有集群及集群内所有物理机均会被启用。
- 删除区域：删除区域会删除此区域下所有资源包括其子资源，包括集群、二层网络、主存储及其子资源物理机、云主机、云盘、快照、三层网络、各网络服务等，此操作非常危险，请极其慎重操作！这里的删除所有资源，只是移除所有资源在ZStack中的记录，并不删除真实的数据。
- UI为便于管理镜像服务器和区域的关系，特别设置了一个镜像服务器只能对应一个Zone。在删除区域的同时会直接删除与挂载此区域的镜像服务器。

8.2 集群

集群：一组物理机的逻辑集合，在数据中心中，一个集群一般对应了一个机架。

目的：

- 用于提高整体系统的计算能力、IO能力；
- 提高服务的可靠性；

- 提升系统的规模扩展能力，降低系统的运维成本。

集群内物理机要求：

- 所有的物理机必须拥有相同的操作系统；
- 所有的物理机必须拥有相同的网络配置；
- 所有的物理机必须能够访问相同的主存储。

集群相关的约束条件：

- 集群与主存储的挂载关系如下：
 - 一个集群可以挂载一个或多个本地主存储。
 - 一个集群可以挂载一个或多个NFS主存储。
 - 一个集群可以挂载一个本地主存储和一个NFS主存储。
 - 一个集群可以挂载一个本地主存储和一个Shared Mount Point主存储。
 - 一个集群只能挂载一个Ceph主存储，除此外不能再挂载新的存储。
 - 一个集群只能挂载一个Fusionstor主存储，除此外不能再挂载新的存储。
 - 一个主存储可以挂载到多个集群。
- 一个集群可以挂载多个二层网络。
- 一个二层网络也可以挂载到多个集群。
- 停用集群后，会停用此集群下所有的物理机，停用后，此集群就不能作为候选用于申请新资源，原本运行的云主机不受影响，但这些运行的云主机停止后就不能再次在此集群启动。
- 启用集群后，会同时启用此集群下所有的物理机。
- 删除使用共享存储的集群，相关的云主机会被停止，如果这些已停止的云主机设置了高可用标签，那么可以在其他可用的集群上的物理机再次启动。
- 删除使用本地存储的集群，会删除集群内所有的物理机，也会删除所有的VM，但不删除物理机内所有的云盘和快照真实数据。
- 如果不同集群挂载了相同网络和相同主存储，则支持云主机跨集群迁移。
- 从集群卸载网络，ZStack会卸载掉集群内所有云主机的网卡，此时云主机没有网卡，运行中的云主机一旦停止，将无法再次开机，需要再次给集群加载网络后，再给云主机添加网卡，才可正常开机。
- 从集群卸载主存储，ZStack会停止集群内所有的VM，只有再次加入挂载此主存储，这些云主机才可正常启动。
- 存在多个本地主存储时，创建云主机会优先选择可用容量大的存储来创建云盘。

- 存在多个NFS主存储时，创建云主机会随机选择主存储来创建云盘。
- 针对本地主存储+NFS/Shared Mount Point类型主存储，默认会选择与当前根云盘不在同一个主存储的存储来创建。
- 存在多个主存储时，可以指定主存储来创建云盘。
- 云主机不支持在多个主存储之间进行迁移。
- 云主机的数据盘所在的主存储进入维护模式，云主机会停止且无法启动。
- 在VDI场景，创建集群时，如果已经部署了VDI单独使用的网络，可以指定VDI网络的CDIR，此网络用于VDI连接的协议流量，例如输入172.20.0.0/16，则在物理机上以此网络进行VDI协议流量的传输，如果输入错误或不存在的网络CIDR，则会默认使用管理节点添加物理机时的管理网络。

8.3 物理机

物理机：为云主机实例提供计算、网络、存储的物理主机。

添加物理机注意：

- 管理节点的管理IP需正常访问物理机的SSH 端口以部署代理程序。
- 物理机的用户名须拥有sudo权限。
- SSH端口需开启防火墙或iptables允许。
- 需支持Intel VMX|AMD SVM的硬件虚拟化特性。
- 添加或更改物理机的信息，需与实际环境相匹配。

ZStack的物理机有一个字段managementIp，用于表示如果此物理机有多个IP的情况下，添加的IP地址用于管理节点对此物理机进行连接部署代理程序的IP。

ZStack管理节点会周期性的向物理机发送进行ping以检查命令通道的健康状态；一旦某个物理机响应失败，或者ping命令超时，此物理机就会变为失联状态，失联后物理机上运行的云主机会变成unknown状态，待物理机恢复连接后，会再次同步其上云主机的状态。

ZStack管理节点会提供ntpd时间服务器对各物理机进行时间同步。

物理机在使用时需注意：

- 物理机维护模式，主要用于管理员对物理机打算进行关机实施维护操作，例如添加硬盘，更换网卡等。

- 物理机维护模式，若主存储为本地存储，则会停止此物理机上所有云主机，若主存储为共享存储，则会自动迁移其上云主机到可用物理机上，如果其他物理机资源不足，则会停止这些云主机。
- 物理机停用后，原本运行的云主机依然保持运行状态，如果停止后将不能使用此物理机启动。
- 物理机停用或维护后，此物理机不能再作为候选用于申请新资源，只有再次启用后，才可正常使用。
- 物理机失联或者进入维护模式后，才可删除物理机。
- 删除本地存储上的物理机会直接彻底删除此物理机上所有的云主机或云盘。这里的删除，只是移除物理机、云主机、云盘等信息在ZStack中的记录，并不删除真实的数据。
- 删除共享存储上的物理机，不会直接删除其上的云主机或云盘，这些资源可在其他物理机上重新启用或使用。
- 删除物理机后，即使再次添加回来，系统也将无法识别删除的VM，需慎重删除。
- 如果打算使用GPU透传功能，物理机的BIOS需开启VT-d功能，添加物理机时，需勾选IOMMU选项，以扫描IOMMU设备。IOMMU启用状态为启用，IOMMU就绪状态为可用时，才可使用GPU设备透传功能。

物理机运维出现异常时，需检查：

- 管理节点日志：在管理节点检查`/usr/local/zstack/apache-tomcat/logs/management-server.log`是否存在异常。
- 物理机agent进程：在物理机检查`ps -ef|grep kvmagent`，检查kvmagent代理进程是否存在异常。
- 物理机日志：在物理机检查`/var/log/zstack/zstack-kvmagent.log`是否存在异常。
- 物理机libvirt状态：在物理机检查`service libvirtd status`状态是否存在异常。
- 物理机系统日志：在物理机检查`/var/log/messages`是否存在异常。
- 检查物理机是否与主存储连接是否出现异常。

8.3.1 GPU设备透传

GPU设备透传：将物理机的GPU显卡设备，直接透传给云主机使用，云主机独占显卡设备，适用于视频渲染、游戏、人工智能等场景。

使用GPU设备透传需注意：

- 需在物理机上准备GPU设备，支持AMD全系列显卡，支持NVIDIA计算卡。
- 在BIOS开启VT-d或IOMMU支持。

- 添加物理机时，需勾选扫描物理机的IOMMU设备，以支持GPU设备透传，勾选此项后，会遍历物理机可用的GPU设备，扫描后，需重启物理机以使得IOMMU配置在内核生效。
- IOMMU启用状态：表示IOMMU设备直通的状态，勾选表示已在内核启用IOMMU配置，需系统重启后生效。
- IOMMU就绪状态：确保物理机IOMMU为就绪状态才可正常使用GPU设备透传。如果不就绪，表示物理机BIOS或者内核的IOMMU选项未打开。
- 使用GPU透传后物理机的GPU设备透传给云主机后，接入的显示器只显示云主机的界面，物理机的界面将不再显示。
- GPU设备透传需先建立云主机后，再动态加载GPU设备到相应的云主机上，加载后，需安装相应的显卡驱动，才可正常显示。
- 拥有GPU透传设备的云主机不支持迁移功能，需卸载后迁移。
- 拥有GPU透传设备的云主机关机后，只能从原本拥有GPU设备的物理机启动，不支持跨物理机启动。
- 共享存储上的云主机，在关机情况下加载GPU设备时，可以选择任意物理机，但启动时会从选定的物理机启动，如果此物理机资源不足，会启动失败。
- 单个云主机可以挂载多个GPU设备，但是这些GPU设备只能来源于同一个物理机。

8.4 主存储

主存储：用于存储云主机的磁盘文件(根云盘、数据盘、根云盘快照、数据盘快照、镜像缓存等)的存储服务器。

ZStack支持的存储类型包括：

- 本地存储
- NFS
- Ceph
- Fusionstor
- Shared Mount Point

主存储使用时需注意：

- 本地存储、NFS、Shared Mount Point主存储基于qcow2。
- Ceph存储和Fusionstor存储使用的是分布式块存储。
- 以虚拟容量来计算云主机的根云盘和数据云盘消耗。

- 主存储停用后，将不能使用此主存储申请新资源，但不影响当前运行中的云主机，只是不能新建云主机和新建数据盘。
- 主存储维护后，将不能使用此主存储申请新资源，原本运行中的VM也会被停止。
- 本地存储、NFS、Shared Mount Point主存储匹配镜像仓库和Sftp镜像服务器使用。
- Ceph主存储匹配同一Ceph镜像服务器使用。
- Fusionstor主存储匹配同一Fusionstor镜像服务器使用。
- 主存储的使用基于瘦身模式，支持超分，超分比例设置过大时，需留意真实物理容量的使用。
- ZStack创建云盘时使用的是虚拟大小，如果虚拟大小不足以提供需要的容量，此时创建云盘会提示失败。
- 删除主存储会删除主存储所有的云主机、云盘和快照。需谨慎操作！这里的删除，只是移除主存储、云主机、云盘和快照等信息在ZStack中的记录，并不删除真实的数据。
- 共享存储可以设置独立的存储网络与管理网络隔离，管理节点依据此存储网络来判断云主机健康状态，如果不存在独立的存储网络，可以留空不填，表示采用管理网络来判断云主机健康状态。

8.4.1 本地存储

本地存储：使用各物理机的本地硬盘上指定的目录作为主存储。

本地存储的特性：

- 使用物理机的本地存储空间，存放云盘数据。
- 云主机数据保护需依赖节点RAID阵列技术。
- 投资低、性能高、大规模，但云主机热迁移存在问题，需要业务层高可靠。
- 适用于软件开发测试环境、分布式应用、应用与数据分离的场景。

使用本地主存储时需注意：

- 不能使用以下/、/dev/、/proc/、/sys/、/usr/bin、/bin等系统目录，使用这些目录可能会导致物理机异常。
- 本地主存储的状态一般为Connected。
- 本地主存储容量由各物理机的主存储目录的容量累加。
- 镜像缓存在添加镜像后，会自动分发到各物理机的主存储目录进行缓存，使用此镜像的不同云主机共享使用此镜像缓存作为base。
- 因为存储的隔离性，使用本地存储的云主机，不支持高可用，不支持指定物理机启动。
- 使用本地存储的云主机默认支持冷迁移。

- 使用本地存储的云主机支持在线热迁移，默认关闭，需要在全局设置打开，本地存储的云主机不支持带数据盘的热迁移，Windows操作系统的云主机不支持热迁移。
- 本地存储的镜像缓存在删除镜像后，如果没有任何云主机使用，此镜像缓存会自动删除。
- 本地存储一般都处于可连接状态，若某物理机失联，则此物理机上的本地存储容量将被扣除，一旦物理机再次连接，容量会被再次计算进来。
- 多个本地主存储，创建云主机时，可指定任意主存储；若未指定，系统自动选择可用容量最充足的本地主存储。

8.4.2 NFS

NFS主存储：使用网络文件系统作为主存储，需提供NFS的共享文件目录。

NFS使用时添加的URL格式为 `192.168.0.1:/nfs_root`

NFS主存储的特性：

- 云盘存放在集中存储，数据需通过存储双控和RAID技术保护。
- 需要构建万兆存储网络，购买集中存储，云主机受高可用保护。
- 特点：构建了存储网、云主机支持高可用、数据支持安全保护、IO性能有瓶颈（8~10节点/一个存储）。
- 场景：适用于传统应用、应用数据紧密、SLA高可用保证、数据安全场景。

使用NFS主存储需注意：

- NFS的挂载目录需提供读写权限。
- NFS的挂载参数，需NFS Server端支持。
- ZStack会在所有的物理机上自动挂载此目录作为主存储。
- 只要集群内任一物理机可访问NFS主存储，则主存储就处于可连接状态。
- NFS属于共享存储，支持VM在线迁移。
- NFS主存储是所有物理机共享带宽访问同一目录，效率和性能具有瓶颈。
- NFS主存储只会拷贝一份镜像缓存到主存储的共享目录。
- NFS存储支持输入存储网络CIDR，表示基于此存储网络来判断云主机的健康状态。
- 多个NFS主存储，创建云主机时，可指定任意主存储，若未指定，系统随机分配可用主存储。
- 一个集群支持挂载一个NFS主存储和一个本地主存储，在创建云主机时，可指定任意主存储，如果带数据云盘，则需指定数据云盘所使用的主存储；如果不指定主存储，则使用本地主存储来创建云主机。

- 使用NFS+本地主存储的云主机，进行迁移前，需卸载其本地主存储上的云盘。这里的迁移仅限于同一主存储内部，不可跨主存储迁移。

8.4.3 Ceph

Ceph主存储：Ceph集群提供的分布式块存储作为主存储。

Ceph分布式存储的特性：

- 云盘存放在分布式存储，数据通过多副本和纠删码机制提供保护。
- 需要构建存储网络，基于x86通用服务器提供存储服务，云主机受高可用保护。
- 特点：构建了存储网，云主机支持高可用，数据支持多副本，计算和存储都可以横向扩展。
- 适用场景：租用平台、通用型应用、应用数据紧密、SLA高可用保证、大规模环境。

使用Ceph主存储时需注意：

- 需输入Mon IP节点信息以部署相关的代理程序。
- 如果指定了镜像缓存、根云盘、数据云盘的Ceph存储池名，则这些存储池需提前在Ceph集群自行创建。
- 存储池名为可选项，如果不指定，则ZStack会自动创建以uuid相关命名的存储池，如果指定了存储池名，但是实际不存在，添加Ceph主存储会提示失败。
- 添加Ceph主存储，需添加同一Ceph集群存储作为镜像服务器。
- Ceph主存储可指定多个Mon IP，需至少一个Mon处于连接状态。
- Ceph存储支持共享云盘，要求使用VirtioSCSI类型作为前提。
- Ceph集群工作异常时，主存储会失联，可通过`ceph -s`检查状态。
- Ceph存储支持输入存储网络CIDR，表示基于此存储网络来判断云主机的健康状态。
- Ceph存储支持添加扩展池，用于存储数据云盘，可表示不同的性能的云盘。可添加或创建Ceph存储池，添加表示使用Ceph集群中已经存在的Pool，如果Ceph存储池已经存在，需进行添加，如果不存在，可指定存储池名来创建，创建时，会采用默认的CRUSH Map。
- 删除扩展池，会删除Pool下所有的数据盘。这里的删除只移除Ceph Pool在ZStack的记录，并不删除真实数据，Ceph存储集群依然可看到此Pool及其上数据。
- 拥有独立存储网络的Ceph存储，节点管理IP表示管理节点连接存储节点的IP，Mon IP表示存储网络的监控节点IP，可以通过`ceph -s`读取。

8.4.4 Fusionstor

Fusionstor主存储：采用Fusionstor分布式块存储作为主存储。

Fusionstor分布式存储的特性：

- 云盘存放在分布式存储，数据通过多副本和纠删码机制提供保护。
- 需要构建存储网络，基于x86通用服务器提供存储服务，云主机受高可用保护。
- 特点：构建了存储网，云主机支持高可用，数据支持多副本，计算和存储都可以横向扩展。
- 适用场景：租用平台、通用型应用、应用数据紧密、SLA高可用保证、大规模环境。

使用Fusionstor主存储时需注意：

- 需输入Mon IP节点信息以部署相关的代理程序。
- 如果指定了镜像缓存、根云盘、数据云盘的Fusionstor存储池名，则这些存储池需提前在Fusionstor集群自行创建。
- 存储池名为可选项，如果不指定，则ZStack会自动创建以uuid相关命名的存储池，如果指定了存储池名，但是实际不存在，添加Fusionstor主存储会提示失败。
- 添加Fusionstor主存储，需添加同一Fusionstor集群存储作为镜像服务器。
- Fusionstor主存储可指定多个Mon IP，需至少一个Mon处于连接状态。
- 使用Fusionstor存储，需添加Fusionstor存储集群内的存储节点作为计算节点，以超融合方式部署。
- Fusionstor集群工作异常时，主存储会失联，可通过lich stat检查状态。

8.4.5 Shared Mount Point

Shared Mount Point主存储：采用共享文件系统存储为主存储，支持SAN存储，格式化为MooseFS，GlusterFS，OCFS2，GFS2等文件系统。

Shared Mount Point主存储的特性：

- 云盘存放在集中存储，数据需通过存储双控和RAID技术保护。
- 需要构建万兆存储网络，购买集中存储，云主机受高可用保护。
- 特点：构建了存储网、云主机支持高可用、数据支持安全保护，IO性能有瓶颈（8~10节点/一个存储）。
- 场景：适用于传统应用、应用数据紧密、SLA高可用保证、数据安全场景。

使用Shared Mount Point主存储时需注意：

- 需在各物理机提前挂载各挂载点，挂载到相同目录。
- 在添加到ZStack时，需提供挂载到物理机的本地目录。
- 支持在线迁移，对应挂载目录需提供读写权限。

- 一个集群支持挂载一个Shared Mount Point和本地主存储，创建云主机时，可指定任意主存储，如果带数据云盘，则需指定数据云盘所使用的主存储；如果不指定主存储，则使用本地主存储来创建云主机。
- 使用Shared Mount Point+本地主存储的云主机，进行迁移前，需卸载其本地主存储上的云盘。这里的迁移仅限于同一主存储内部，不可跨主存储迁移。

8.5 镜像服务器

镜像服务器：用于保存云主机的镜像模板或ISO的存储服务器。

使用镜像服务器时需注意：

1. Sftp：

- 以文件方式存储镜像文件。
- 支持关机创建镜像。创建的镜像可以在镜像服务器上，以对应的镜像路径访问，拷贝到其他云环境可直接使用。

2. 镜像仓库：

- 以镜像切片方式存储镜像文件。
- 支持VM的在线快照和在线镜像、克隆；支持导出镜像。

3. Ceph镜像服务器：

- 支持块存储方式存储镜像文件。
- 支持在线快照，支持在线克隆和关机克隆。
- 导出镜像需在镜像服务器上导出。使用假定镜像路径为`ceph://bak-t-c9923f9821bf45498fdf9cdfa1749943/61ece0adc7244b0cbd12dafbc5494f0c`，则需镜像服务器执行：

```
rbd export -p bak-t-c9923f9821bf45498fdf9cdfa1749943 --image 61ece0adc7244b0cbd12dafbc5494f0c /root/export-test.image
```

```
# bak-t-c9923f9821bf45498fdf9cdfa1749943表示镜像所在的pool的名字  
# 61ece0adc7244b0cbd12dafbc5494f0c表示镜像的名字  
# /root/export-test.image表示导出的目标文件名字
```

4. Fusionstor镜像服务器：

- 支持块存储方式存储镜像文件。
- 支持在线快照，不支持克隆。

- 导出镜像需要在镜像服务器上执行类似命令：

```
lichbd export bak-t-8e694c40cf214db1af9e5d641b2e792d/8f1e0debfcae042e5ae074  
133a59c0622 /root/test.img -p nbd
```

5. 开源版本支持Sftp、Ceph和Fusionstor类型，企业版本支持镜像仓库、Ceph和Fusionstor类型。
6. 删除镜像服务器需注意：
 - 删除镜像服务器会删除所有镜像文件，且无法恢复，需谨慎操作。这里的删除，只是移除镜像服务器和镜像在ZStack中的记录，并不删除真实的数据。
7. UI为便于管理镜像服务器和区域的关系，特别设置了一个镜像服务器只能对应一个Zone。添加镜像服务器时，默认挂载到了当前区域。在删除区域的同时会直接删除挂载此区域的镜像服务器。

关于镜像服务器支持克隆和快照的关系需注意：

- 在线克隆：镜像仓库、Ceph支持；
- 关机克隆：镜像仓库、Ceph支持；
- 在线快照：镜像仓库、Ceph、Fusionstor支持；
- 关机快照：镜像仓库、Sftp、Ceph、Fusionstor均支持；
- 在线创建镜像：镜像仓库和Ceph支持；
- 关机创建镜像：镜像仓库、Sftp、Ceph、Fusionstor均支持。

9 网络

9.1 二层网络

二层网络：二层广播域，进行二层相关的隔离，一般以设备编号进行识别。

支持以下类型：

- **L2NoVlanNetwork**

NoVlanNetwork 类型表示相关的物理机对应的网络设备不设置Vlan，如果交换机端口设置了Vlan，则需在交换机端配置Access模式。如果交换机端口没有设置Vlan，则无须特别设置。

- **L2VlanNetwork**

VlanNetwork类型表示相关的物理机对应的网络设备需设置Vlan，从逻辑上划分虚拟局域网，支持1- 4094个子网，此类型需在物理机接入的交换机端进行trunk设置。

- **VxlanNetworkPool**

VxlanNetworkPool类型表示使用UDP进行报文封装的VXLAN类型的集合，是基于IP网络组建的大二层网络，可满足于大规模云计算中心的需求，最大支持16M个逻辑子网。

- **VxlanNetwork**

VxlanNetwork类型表示使用VXLAN的子网进行网络配置，需要先建立VxlanNetworkPool，再建立VxlanNetwork。

使用二层网络时需注意：

- NovlanNetWork和VlanNetwork作为一组物理定义的二层网络，主要差异在于是否支持Vlan，在添加时，指定的是网卡设备号，与物理机网卡端口在交换机端的设置应相同。
- VxlanNetworkPool和VxlanNetwork作为一组基于三层网络上叠加的逻辑定义的二层网络，是专门提供VxlanNetwork使用，在使用VxlanNetwork前，需要先建立VxlanNetworkPool。与集群发生关系的是三层VTEP，基于VTEP进行VXLAN报文的封装和解封装。创建完毕VxlanNetworkPool后，可指定或随机选择Vni来创建VxlanNetwork。最后基于VxlanNetwork来创建三层网络，再创建云主机。
- NovlanNetWork、VlanNetwork、VxlanNetworkPool类型的网络可以挂载集群，VxlanNetwork只是Vni Range的子集，不能挂载集群。
- VxlanNetworkPool挂载集群时，需检查集群内所有的物理机在对应网段是否存在可用IP来匹配VTEP，如果不存在，则此VxlanNetworkPool不能正确挂载集群

- 如果集群已挂载某二层网络，但物理机不存在L2设备，则此物理机不能添加进入对应集群。
- 如果集群未挂载某二层网络，但物理机不存在L2设备，则集群则不能挂载此二层网络。
- 如果某物理机存在某二层网络对应的设备，但设备接线与其他物理机接线不一致，则创建出的云主机IP不能正常工作。
- 删除二层网络，会删除其下的三层网络，相关的云主机的网络也会被删除，请慎重操作。
- 删除二层网络，会删除使用此二层网络的云路由设备和云路由规格。
- 在添加NovlanNetWork和VlanNetwork时，需要输入网卡设备名称，在CentOS7系列系统中，ethx格式的网卡名称会在系统重启后导致网卡顺序随机改变，建议将各计算节点的网卡设备名称修改成非ethx格式，例如，可修改成em01格式。尤其是带多网卡的虚拟机环境中。
- 当使用Vlan设备时，无需在物理机上创建带有Vlan号的网卡，例如eth0.2016，用户只需要输入设备名为eth0，Vlan号为2016，ZStack就会自动创建eth0.2016。如果已经存在eth0.2016网络设备时，用户又输入了Vlan号为2016，ZStack会创建一个新的名为eth0.2016.2016的网络设备。

9.2 三层网络

三层网络：云主机使用的网络配置，包含了IP地址范围、网关、DNS、网络服务等。

- IP地址范围包含起始和结束IP地址、子网掩码、网关等，例如可指定172.20.12.2到172.20.12.255，子网掩码指定255.255.0.0，网关指定172.20.0.1。也可使用CIDR无域间路由来表示，例如192.168.1.0/24。
- DNS用于设置云主机网络的DNS解析服务。

ZStack中相关网络的基本概念：

- 公有网络：可直接连通互联网的网络，在云路由网络中可以提供网络服务，可用于扁平网络创建使用公网的云主机，也可用于云路由网络环境，单独创建使用公网的云主机。
- 系统网络：管理节点用于特定用途的网络，可用于部署配置相关资源的管理网络，例如部署物理机、主存储、镜像服务器、云路由等资源；也可用于云主机迁移的迁移网络，如果网络资源不足，可与公有网络共用，独立的系统网络仅用于特定用途，例如云路由设备。系统网络不能用于创建普通云主机。
- 管理网络：作为系统网络的一种，用于管理控制对应的物理资源，例如物理机、镜像服务器、主存储等需提供IP进行访问的资源时使用的网络；创建云路由设备时需要云路由设备存在管理节点互通的IP，以便部署agent及agent代理消息返回。

- 私有网络：也可称之为业务网络或接入网络，云主机使用的网络，一般情况下设置为私网。私有网络指定为云主机使用的网络，支持两种网络架构模型：扁平网络和云路由网络。
- 存储网络：特指在进行分布式存储部署时，底层存储系统通信使用的网络。在添加主存储时，可标识存储网络的CIDR，表示使用此网络来判断云主机健康状态。
- VDI网络：在创建集群时，可以指定VDI网络的CDIR，此网络用于VDI连接的协议流量。

ZStack支持的网络服务包括DHCP、DNS、SNAT、弹性IP、端口转发、负载均衡、IPsec、安全组、UserData等，网络架构的类型决定了此网络提供的网络服务类型。

- 扁平网络支持以下网络服务：DHCP、弹性IP、安全组、UserData等，扁平网络的网络服务采用分布式的DHCP、分布式的EIP结构。扁平网络的DHCP服务也包含了DNS的功能。ZStack的初始化引导设置使用的网络模型就是采用了扁平网络。
- 云路由网络支持以下网络服务：DHCP、DNS、SNAT、弹性IP、端口转发、负载均衡、IPsec、安全组等，云路由网络主要使用特定的Linux云主机作为路由设备提供网络服务。

创建云主机时，可指定多个网络，即可指定多个扁平网络，又可指定多个云路由网络，还可指定扁平网络和云路由网络的混合使用。

ZStack支持多级网络，而且多级网络的L2可以实际通信，需要特别避免IP地址空间冲突的问题。

9.3 扁平网络

扁平网络具备以下特性：

- 物理机和云主机均处于同一个二层广播域。
- 提供DHCP、Userdata (ZStack使用SystemTag来使用Userdata注入ssh key)、EIP 等服务。
- 分布式DHCP、分布式弹性IP，可规避DHCP服务器的单点故障，高并发时，可有效提高系统整体并发性。

扁平网络在每台物理机上均独立运行DHCP服务器进行云主机IP地址的分配。

扁平网络使用dnsmasq来提供DHCP服务。

扁平网络使用Linux的namespace和iptables来实现弹性IP的转发。

扁平网络运维时需注意：

- 如果某物理机上所有的云主机都无法获取IP，此时需要在物理机上检查dnsmasq进程及/var/lib/zstack/dnsmasq/对应二层网卡目录下的dnsmasq.log日志。
- 扁平网络的弹性IP需要云主机存在通向私网网关的默认路由。

9.4 云路由网络

云路由网络：通过定制的Linux云主机作为路由设备来提供DNS、SNAT、弹性IP、端口转发、负载均衡、IPsec、安全组等网络服务，支持分布式DHCP服务。

依赖：公有网络、管理网络、云路由镜像、云路由规格、私有网络。

云路由网络具备以下特性：

- 使用Linux云主机作为路由设备提供相关网络服务。
- 使用云路由设备提供DNS、SNAT、弹性IP、端口转发、负载均衡、IPsec、安全组等网络服务。
- 支持分布式DHCP服务。
- 可实现管理网络、公有网络、私有网络三网隔离的环境，适用于构建大型专有云网络环境。

云路由网络使用时注意：

- 需先建立公有网络、再建立管理网络，再建立云路由规格，再创建云路由私有网络，在创建云主机时，会自动创建云路由设备。
- 创建云路由设备至少需要一个集群，同时挂载了管理网络、公有网络、私有网络，而且此集群要求有可用物理机来创建云路由设备。
- 创建云路由规格时，管理网络可从系统网络和公有网络里面选择，如果公有网络和管理网络合一，需选择公有网络作为管理网络；如果存在独立的管理网络，则需在系统网络选择。
- 云路由设备的默认IP为公有网络的IP，会从公有网络的网络段中选择IP地址，提供各种网络服务。
- 云路由设备的私有网络段地址如果是类似192.168.0.0/24，那么云路由设备的私网IP默认为192.168.0.1，并以此作为使用此云路由网络的普通云主机的默认网关。普通云主机使用此网关，在云路由设备里，进行路由转发以访问互联网。
- 云路由设备中公有网络、管理网络、私有网络或其他网络的网络段之间，其CIDR不可重叠，例如，假定私有网络网络段为10.0.0.0/8，云路由设备加载网络时，新的网络段就不可为10.20.0.0/16，后者作为前者的子集，存在IP地址冲突的风险。

云路由设备支持加载网卡，并支持路由表和路由条目，使用时需注意：

- 加载网卡一般用于多网互联的场景，例如混合云互通。
- 云路由设备支持加载路由表，加载后，路由表内的路由条目会添加至云路由设备，对使用云路由的私有网络进行相应的路由功能。

- 添加路由条目时，需指定目标网段，并输入下一跳的地址，需确保下一跳可达。路由类型一般设置静态路由。为防止环路，也可设置黑洞路由，表示丢弃匹配的数据包。路由优先级的数字越大，表示优先级越低。

云路由设备创建失败时，需检查：

- 云路由规格定义的镜像：如果存放在Imagestore或SFTP镜像服务器，则应存在挂载了本地存储、NFS或者SharedMountPoint主存储的集群；如果存放在Ceph镜像服务器，则应存在挂载了此Ceph镜像服务器对应的Ceph主存储的集群；如果存放在Fusionstor镜像服务器，则应存在挂载了此Fusionstor镜像服务器对应的Fusionstor主存储的集群。
- 需存在可用集群同时加载了公有网络、管理网络和私有网络，此集群内有可用的物理机满足云路由规格定义的CPU、内存大小。

9.5 VXLAN

VXLAN网络具备以下特性：

- 建立在物理IP网络之上的虚拟网络，使用UDP封装五十个字节的报文头。
- 使用24位VXLAN网络标识符，最大支持16M个逻辑网络。
- VXLAN可跨越物理三层网络。
- 使用UDP封装在三层物理网络上建立的二层逻辑网络。
- 使用IP多播封装广播和多播报文。

VxlanNetworkPool和VxlanNetwork共同提供了VXLAN网络类型的配置，使用VxlanNetwork需先创建VxlanNetworkPool，VxlanNetwork对应了VxlanNetworkPool里面的一个虚拟网络。

- VxlanNetworkPool最大可支持16777216（16M）个的虚拟网络。其Vni（Vxlan网络ID）范围可从1-16777216设置。
- 在创建VxlanNetworkPool时，如果需要加载到相应集群，则需设置相应的VTEP（Vxlan隧道端点），VTEP一般对应于集群内计算节点中的某一网卡的IP地址。ZStack对VTEP的设置基于相应的CIDR进行配置。



注：

- 例如，假定计算节点某网卡的IP为10.12.0.8，子网掩码为255.0.0.0，网关为10.0.0.1，则VTEP输入的CIDR应为10.0.0.1/8；
- 假定计算节点某网卡的IP为172.20.12.13，子网掩码为255.255.0.0，网关为172.20.0.1，则VTEP输入的CIDR应为172.20.0.1/16。

- VxlanNetworkPool与集群进行挂载时，检查的是VTEP相关的IP地址，与物理的二层设备无关。
- 只有VxlanNetworkPool 可以挂载集群和卸载集群的操作，VxlanNetwork作为VxlanNetworkPool的子资源，不可进行挂载集群和卸载集群的操作。

9.6 网络服务

ZStack for Alibaba Cloud的网络支持两种架构模型：扁平网络和云路由网络。

扁平网络：

扁平网络支持以下网络服务：

- DHCP：分布式DHCP实现的动态获取IP地址。
- 弹性IP：分布式EIP实现的弹性IP地址，可通过公有网络访问内部私有网络。
- 安全组：使用iptables进行云主机防火墙的安全控制。
- Userdata：使用cloud-init进行云主机开机加载并执行特定的用户数据，例如ssh-key注入。



注：其中DHCP服务包含了DNS的功能。

云路由网络：

云路由网络支持以下网络服务：

- DHCP：分布式DHCP实现的动态获取IP地址。
- DNS：使用云路由设备提供DNS服务。
- SNAT：云主机使用SNAT可以直接访问外部互联网。
- 弹性IP：使用云路由设备可通过公有网络访问云主机的私有网络。
- 端口转发：提供将指定公有网络的IP地址端口流量转发到云主机对应协议的端口。
- 负载均衡：将公网地址的访问流量分发到一组后端的云主机上，并自动检测并隔离不可用的云主机。
- IPsec：使用IPsec隧道协议实现虚拟私有网络（VPN）的连接。
- 安全组：使用云路由进行云主机防火墙的安全控制。

9.6.1 安全组

扁平网络和云路由网络均支持安全组服务。使用方法均相同。

安全组主要给云主机提供三层网络防火墙控制，控制TCP/UDP/ICMP等数据包进行有效过滤，对指定网络的指定云主机按照指定的安全规则进行有效控制。

使用安全组的基本流程为：选择三层网络，设置相应的防火墙规则，选择指定的云主机加入规则中。

安全组支持白名单机制，即设置的所有规则均为允许机制，一旦对指定端口设置了允许机制，那么没有被允许的端口就无法通过。

安全组规则对数据包的流向有两种：

- Ingress：代表数据包从外部进入云主机。
- Egress：代表数据包从云主机往外部发出。

安全组规则对通信协议的支持以下类型：

- TCP：支持1-65536端口。
- UDP：支持1-65536端口。
- ICMP：默认起始结束端口均为-1，表示支持全部的ICMP协议。
- ALL：表示涵盖所有协议类型，此时不能指定端口。

安全组规则支持对数据源的限制，目前源可以设置为CIDR和安全组。

- CIDR作为源：仅允许指定的CIDR才可通过
- 安全组作为源：仅允许指定的安全组内的云主机才可通过



注：如果两者都设置，只取两者交集。

安全组存在以下约束条件：

- 安全组可以挂载到多个云主机，它们会共享相同的安全组规则。
- 安全组，默认为允许规则，即设置的所有规则均为允许机制。
- 新建安全组时，默认配置了两条规则（即：协议类型为ALL的进口规则和出口规则），用于显示设置组内互通。用户可以删除这两条默认规则，取消组内互通。
- 新建安全组时，如果没有配置任何规则，则默认所有的外部访问均禁止进入安全组内的云主机，安全组内云主机访问外部不受限制。添加出口规则（默认规则除外），出口规则之外的都限制。

9.6.2 虚拟IP

虚拟IP定义了桥接网络环境中，使用虚拟IP地址来提供弹性IP、端口转发、负载均衡、IPsec隧道等网络服务，数据包会被发送到虚拟IP，再路由至云主机网络。

虚拟IP一般是将可以访问互联网的公有IP地址，路由到云主机的私有网络。

扁平网络下的虚拟IP服务仅用于弹性IP网络服务。

云路由网络下的虚拟IP服务可用于弹性IP、端口转发、负载均衡、IPsec隧道。

使用虚拟IP的有两种方法：

- 创建虚拟IP地址后，在弹性IP、端口转发、负载均衡、IPsec隧道界面，可选择使用已有虚拟IP。
- 在弹性IP、端口转发、负载均衡、IPsec隧道界面使用新建虚拟IP的方法。

9.6.3 弹性IP

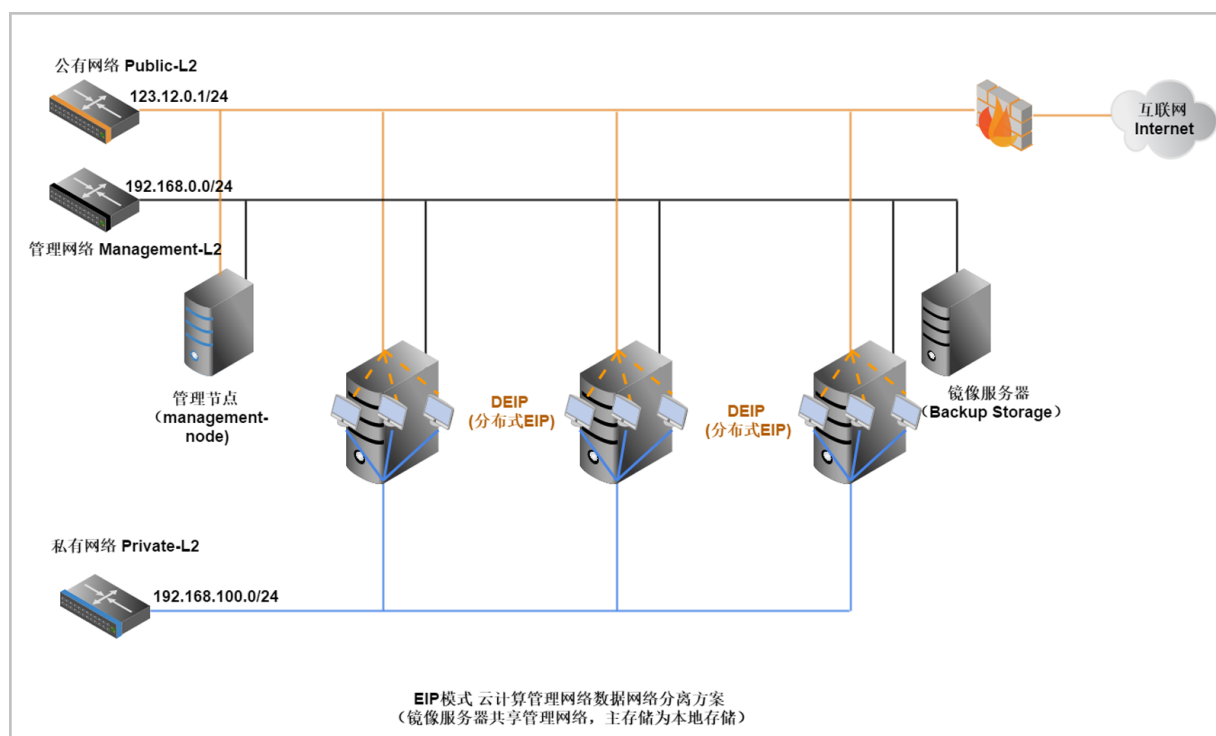
弹性IP定义了通过公有网络访问内部私有网络的方法。

内部私有网络是隔离的网络空间，不能直接被外部网络访问。在扁平网络的私有隔离网络中，内部网络也无法直接访问外部网络。

通过弹性IP可以将对公网的访问直接关联到内部私网的云主机IP。

扁平网络下的弹性IP的应用场景如图 77: 扁平网络下弹性IP的应用场景所示：

图 77: 扁平网络下弹性IP的应用场景



- 123.12.0.1/24的公有网络，可以通过防火墙连接到互联网。

- 管理节点通过管理网络控制内部的专有云管控系统。
- 各个计算节点所在的私有网络提供给云主机内部的私有网络192.168.100.0/24的IP地址。此IP地址默认情况下无法连接到外部的互联网。
- ZStack for Alibaba Cloud在每个计算节点分别部署分布式的EIP。可分布独立实现公有网络与私有网络的绑定。

9.6.4 端口转发

端口转发是基于云路由提供的三层转发服务，可提供将指定公共网络的IP地址端口流量转发到云主机对应协议的端口。在公共网络地址紧缺情况下，通过端口转发可以提供多个云主机对外服务，以节省公共网络地址。

端口转发可以将一个公网节点的网络端口转发到一个私网节点的网络端口。让用户从外部经过云路由访问一个私有网络IP的端口上（访问云主机的某个端口）。

端口转发功能需要使用虚拟IP来实现，虚拟IP对应于公有网络的IP地址范围中的一个可用IP。

端口转发使用虚拟IP有两种方法：

- 创建虚拟IP地址后，在端口转发界面，可选择使用已有虚拟IP。
- 在端口转发界面使用新建虚拟IP的方法。

端口转发指定端口映射也有两种方法：

- 单个端口到单个端口的映射，例如将公网IP的22端口转发到私网IP的80端口。
- 端口区间的映射，例如指定22-80端口，将公网IP的22到80端口按照顺序，映射到私网IP的22到80端口。

端口转发功能存在以下约束条件：

- 端口转发功能要求云主机内部的防火墙策略对指定的转发端口开放。
- 同一个虚拟IP，在提供端口服务时，虚拟IP所用的端口之间不可重复。
- 同一个虚拟IP，在提供端口服务时，可对同一个三层网络上的多个云主机网卡的不同端口提供端口转发服务。
- 虚拟IP地址在解除绑定关系后，再次加载云主机时，只能选择解除绑定关系前的同一个三层网络上的云主机网卡。
- 同一个云主机，只能使用一个虚拟IP来提供端口转发服务。
- 端口转发区间需一一对应，例如设置了源端口22到80端口的端口区间，在云主机私网，默认也选择22到80端口。

9.6.5 负载均衡

负载均衡是将公网地址的访问流量分发到一组后端的云主机，并支持自动检测并隔离不可用的主机，从而提高业务的服务能力和可用性。

可以根据实际情况，随时添加或删减负载均衡规则中的云主机来调整服务能力，而且这些操作不会影响业务的正常访问。负载均衡器支持HTTP/TCP 两种协议。此外，负载均衡器还支持灵活配置多种转发策略，实现高级转发控制功能。

使用负载均衡的基本流程：

1. 选择可用的虚拟IP来创建负载均衡器。
2. 创建监听器来指定公网端口到云主机端口的对应关系，规则及算法等设置。
3. 选择指定三层网络的云主机网卡加入到监听器里，使负载均衡器生效。

负载均衡功能存在以下约束条件：

- 一个虚拟IP可以支持多个监听器。
- 一个虚拟IP支持的监听器指定的云主机网卡必须在同一个三层网络，如果所有的监听器均被删除。

9.6.6 IPsec隧道

IPsec隧道：透过对IP协议的分组进行加密和认证来保护IP协议的网络传输数据，实现站点到站点（site-to-site）的虚拟私有网络（VPN）连接。

ZStack for Alibaba Cloud支持的IPsec特性如下：

- **IPsec连接模式：**
基于安全考虑，只支持主动模式（Main Mode），不支持积极模式（Aggressive Mode）；仅支持ESP封装协议。
- **IPsec传输模式：**
只支持站点到站点的隧道模式，不支持PC点对点模式（基于云端网络模型考虑），不支持两端存在NAT网络。
- **IPsec路由模型：**
只支持基于对端网段配对模型，只支持路由配对模式，不支持路由转发模式（不支持OSPF或BGP等动态路由协议）。

IPsec隧道的典型场景如下：

- 在两套相互隔离的ZStack for Alibaba Cloud环境中，使用云路由的方案。
- 两套环境中云主机的私有网络无法直接通信。
- 使用IPsec可以实现两套云主机的私有网络互相通信。

两套隔离的ZStack for Alibaba Cloud环境中建立IPsec的流程：

1. 在第一套环境中，创建IPsec通道，指定第一套网络的本地公网IP，输入第二套网络指定的公网IP作为远端IP。
2. 在第二套环境中，创建IPsec通道，指定第二套网络的本地公网IP，输入第一套网络指定的公网IP作为远端IP。

10 vCenter

10.1 基础资源

vCenter的基础资源主要涉及vCenter自身的管理，包括添加、删除vCenter的操作。

- 添加vCenter后，ZStack for Alibaba Cloud会自动同步vCenter的集群、物理机、镜像服务器、网络、存储、云主机等资源，相关资源界面均可显示。
- ZStack for Alibaba Cloud支持添加多个vCenter并进行管理。

ZStack for Alibaba Cloud支持vCenter的虚拟化资源，必须满足以下特性：

- vCenter必须建立数据中心、集群和物理机的资源结构；
- 支持显示已经添加的本地存储和共享存储，包括FC、iSCSI和NFS存储；
- 目前不支持存储集群（Datastore Cluster）模式，建议分离使用；
- vCenter需配置分布式交换机（dvSwitch）的端口组信息（目前不支持标准交换机）；
- vCenter已有的模板虚拟机，需要转换为【模板】类型；

vCenter基础资源支持的操作：

- 添加vCenter：添加vCenter至ZStack for Alibaba Cloud，将会导入vCenter已经存在的群集、物理机、虚拟机和模板信息到ZStack for Alibaba Cloud中。



注：

vCenter物理机导入ZStack for Alibaba Cloud支持过滤，只有添加到dvSwitch中的物理机导入ZStack for Alibaba Cloud；未添加到dvSwitch中的物理机不能导入ZStack for Alibaba Cloud。

- 删除：删除vCenter只会删除vCenter相关资源在ZStack for Alibaba Cloud上的记录，并不会删除vCenter上的真实资源。
- 查看集群：查看vCenter相关的集群信息。
- 查看主存储：查看vCenter相关的主存储信息。
- 查看镜像服务器：查看vCenter相关的镜像服务器信息。
- 查看物理机：查看vCenter相关的物理机信息。

10.2 云主机

添加vCenter后，vCenter的云主机会自动同步至ZStack for Alibaba Cloud；vCenter也支持新建云主机。

vCenter云主机支持的操作：

- 停止：停止云主机。
- 启动：将停止状态的云主机启动。
- 重启：将云主机重启。
- 暂停：暂停云主机。
- 恢复：从暂停状态恢复云主机。
- 关闭电源：将云主机电源直接断电。
- 打开控制台：打开云主机的控制界面，可以登录控制云主机。
- 设置控制台密码：设置云主机的控制台密码。
- 取消控制台密码：取消云主机的控制台密码。
- 删除：删除云主机。
- 恢复：从删除状态恢复云主机。
- 彻底删除：将删除状态的云主机彻底删除。

10.3 网络

ZStack for Alibaba Cloud支持在vCenter环境中配置云路由网络。

vCenter云路由网络搭建流程：

1. 在vCenter中创建公有网络（包括二层公有网络和三层公有网络）。
2. 在创建vCenter云路由网络前，需提前在云路由界面添加云路由镜像和云路由规格。
3. 创建vCenter云路由网络。

vCenter网络支持的操作：

- vCenter公有网络：添加网络段、添加DNS、删除

vCenter云路由网络：添加网络段、添加DNS、加载云路由规格、卸载云路由规格、删除

10.4 镜像

vCenter支持添加vmdk格式的镜像。

vCenter中镜像支持的操作：

- 添加：添加vmdk格式的镜像到vCenter。
- 启用：启用后，此镜像可作为候选使用。
- 停用：停用后，此镜像不可再作为候选使用。
- 删除：删除此镜像。

11 平台管理

11.1 性能统计

ZStack for Alibaba Cloud平台管理提供对已运行云主机、云路由、物理机的实时性能统计。

- 在性能统计界面，默认按照创建日期排序。可以点击CPU、内存、磁盘等按钮按照相应的资源的使用率进行排序。
- 在性能统计界面可选择不同的时间段，来显示最近一段时间内的资源使用情况，可选值为1分钟，1小时，1天，2周。
- 其中磁盘为读写速度的总和，网络为上下行速度总和，CPU、内存、磁盘、网络的读出值为用户所选时间单位内的平均值。

11.2 报警

ZStack for Alibaba Cloud无需额外安装软件，就可以自动监控云主机、物理机等资源的CPU/内存/磁盘IO/网络IO等多项指标。用户可以自行创建报警规则，以邮件方式发送监控报警信息。使用这些分析结果，用户可以快速做出反应，保证业务的稳定运行，并配合监控数据回顾资源的历史信息，对未来资源进行合理规划。

目前报警机制如下：

- 监控的资源类型包括：云主机、物理机
- 监控的条目包括：CPU使用率、内存使用率、磁盘使用率、网络使用率
- 监控周期：
 - 在监控周期内，系统从创建报警器时刻起，每隔10秒采集监控目标当相关资源条目持续到达或超出报警阈值，就会触发邮件报警。
 - 用户按需设置监控周期，ZStack for Alibaba Cloud建议监控周期不要设置过小，以免频繁接收报警邮件。

11.3 计费

ZStack for Alibaba Cloud平台管理的计费界面，支持显示各账户指定时间内使用根云盘、数据云盘、云主机（包含：CPU与内存）的各费用信息以及总额信息。

- 默认指定的时间段为当前时间过去的一个月，可以根据需求进行适当的调整。
- 计费详情页可以显示某账户在指定时间内的根云盘、云主机、数据云盘具体费用信息。

- 在计费单价变化的过程中，也会分阶段显示相关的费用信息。

11.4 定时器

定时器功能非常适用于长时间运行的操作，例如，为某个云主机定时创建快照。ZStack for Alibaba Cloud将定时器和定时任务解耦，用户更方便的创建不同的定时任务到不同规则的定时器上。此外，定时任务可以选择性的停用/启用定时任务，灵活处理生产环境中的特殊情况。定时任务的执行也会完整的进入审计中。

关于定时器和定时任务：

1. 定时器执行策略：包括重复执行和按次数执行

- 选择重复执行：定时任务按周期无限重复执行
- 选择按次数执行：定时任务按周期有限次执行，需设置执行次数

2. 目前支持的定时任务类型包括：

- 启动云主机
- 停止云主机
- 重启云主机
- 创建云主机快照
- 创建云盘快照

3. 定时器与定时任务松耦合。

- 在定时器上，既支持创建并启用定时任务，也支持停用、删除定时任务。
- 删除定时器，连级删除定时器上的定时任务
- 在定时器详情页或云主机详情页，均支持相关定时器上的定时任务或相关云主机的定时任务的查看。
- 对于周期内有限次执行的定时器，当定时任务执行完后，定时器状态将显示为已完成。

11.5 控制台服务

ZStack for Alibaba Cloud平台管理支持显示当前控制台代理的信息，即打开云主机的控制台时使用的代理信息。

- 默认代理显示的IP为管理节点的IP地址。
- 显示类型为ManagementServerConsoleProxy。
- 仅有状态为启用和已连接时，打开云主机的控制台才可正常访问到云主机。

支持重连操作。一般发生在云主机控制台打开失败时，进行重连操作。重连后状态显示为启用和已连接时，代表控制台可以正常打开。

11.6 用户管理

用户管理主要提供了用户对系统资源的访问控制，可实现以细粒度对资源归属及权限控制的划分。

用户管理提供账户、用户组、用户的管理，同时涉及相关资源的所有者更改。

相关定义

- 账户：

作为资源拥有的基本单位，对作用域的资源可以进行创建、删除、分享、召回等操作。账户分为admin管理员账户和普通账户。

- 用户：

用户账户创建，用于实现更细粒度的权限控制。admin创建的用户，也称之为admin用户，拥有和admin账户相同的全部权限。

- 用户组：

账户可以通过创建用户组对一组用户进行批量的权限控制

- 资源配额：

简称配额，是admin账户对普通账户的资源总量进行控制的衡量标准。

- 主要包括云主机数量、CPU数量、内存容量、最大数据云盘数目和所有云盘最大容量等。
- admin账户可修改以上各参数对各个普通账户进行资源总额的控制。当资源删除后，但还未彻底删除时，会占用主存储资源和云盘数量

相关约束

1. **admin**管理员账户：也称之为admin账户，不受权限控制，拥有超级权限，通常由IT系统管理员拥有。

- admin账户可以共享计算规格、云盘规格、网络、镜像等其他资源给普通账户，而普通账户只能操作属于自己的资源。admin账户同时也可以对相关资源进行召回，不再共享。
- admin账户可以通过修改配额对普通账户进行资源总量控制。
- admin账户创建的admin用户，和admin账户一样，拥有全局的控制权限。
- admin账户不能够修改别的账户的普通用户的权限。普通用户的权限应该由该用户所属的账户管理。

- admin账户不支持创建用户组，也不支持对其他账户的用户和用户组进行跨越管理。但可以修改普通账户、普通用户的用户名、密码和简介。
- admin账户创建VxlanNetworkPool后，普通账户可以基于VxlanNetworkPool创建VxlanNetwork。
- 只支持删除admin用户，不支持删除admin账户。
- 更改云主机所有者会更改云主机的EIP所有者属性。

2. 普通账户：由admin管理员账户创建。

- 普通账户拥有对自己创建的云主机、镜像、云盘、安全组、用户组和管理权限。普通账户可以对admin账户共享的资源进行读操作，但不可以进行删除操作。
- 普通账户可以通过权限控制来操控属于自己的用户或用户组。
- 普通账户可以使用用户组对批量用户进行权限控制。
- 删除普通账户会导致此账户下的所有资源被删除，例如，云主机、云盘、镜像、名下用户和用户组等信息。
- 普通用户默认只拥有对普通账户资源的只读权限。
- 普通用户不占有资源，经授权后，可共享并使用自己所属账户下的资源。
- 删除普通用户只会删除普通用户的自身信息，其所创建的云主机、镜像、云盘均会保留在自己所属的账户名下。
- 普通账户名称不可重复。同一账户下的用户和用户组名称不可重复。
- 普通用户的名字、简介和密码可以通过admin账户修改，也可通过所属账户进行修改。
- 同一用户可加入多个不同用户组。
- 账户登录只需输入账户名和密码，用户登录需要输入账户名、用户名和密码。
- 普通账户首页看到的资源是admin账户分配的资源配额的上限。
- 普通账户创建云主机前，需要admin账户提前共享计算规格、网络和云盘规格等资源，否则不可创建云主机。
- 普通账户可以添加自有的镜像文件，也可由admin账户提前共享。
- 用户权限受到用户权限设置页以及该用户所属用户组权限设置页共同控制。只要用户权限设置页，或者该用户所属任意用户组权限设置页授予了某资源的权限，即代表该用户拥有该权限对应的操作。如果需要禁止该用户对某资源的操作权限，需要禁止该用户权限页，以及该用户所属所有用户组权限页相关资源的操作权限。

11.7 设置

11.7.1 全局设置

11.7.1.1 基础设置

全局设置包括基础设置与高级设置。目前基础设置支持十六项，具体介绍如下：

- 云主机高可用：

可以打开或者关闭云主机高可用功能。如果关闭此选项，则云主机不支持设置高可用。

- CPU超分率：

- 默认值为10，主要用于设置可分配的虚拟CPU个数。
- 如果物理机的CPU为四核八线程，ZStack会将物理主机CPU的总数(8)以处理器超分率的倍数(10)计算可分配的虚拟CPU的个数，即总量为80。
- 虚拟CPU的总数可以在首页的处理器器的总量。

- 会话超时时间：

设定ZStack图形界面登录后会话多长时间失效，默认为2小时，即7200秒。当用户登录会话失效后，用户将需要重新登录。

- 物理机保留内存：

物理主机保留内存：所有KVM主机上保留的内存容量。建议可用单位：T/G/M。



注：

例如：512M：表示为系统预留512M内存，当系统剩下700M的内存的时候，用户又想启动一个内存为512M的云主机是启动不了的。

- 云主机缓存模式：

云主机缓存模式设置,可选模式为：writethrough, none, writeback。

1. Writethrough: 物理机的页面缓存工作在透写模式，数据完全写入云主机存储设备后，才返回成功。
2. None: 云主机不使用物理机的页面缓存，直接访问存储，不带cache。
3. Writeback: 云主机使用了物理机的页面缓存机制，数据写入物理机页面缓存即报告给云主机返回成功。

- 云主机CPU模式：

- 选择云主机的CPU类型是否与物理机的CPU类型一致。

- 当选择host-model类型时，云主机的CPU类型将与物理机的CPU类型相符（例如都显示为Haswell的Intel CPU）。
- 当选择host-passthrough类型，云主机的CPU类型将与物理机的CPU完全一致。
- 当选择host-model或者host-passthrough类型时，云主机可以支持嵌套虚拟化，但可能导致云主机在不同型号CPU的物理机之间迁移失败。

- 本地存储迁移：

本地存储迁移全局设置打开或关闭。打开此开关，则支持本地热迁移。

- 内存超分率：

- 默认值为1.0。如果物理内存为4G，设置为2，那么ZStack会认为系统可以分配8G的内存给云主机使用。
- 允许范围 ≥ 1 ，该数值为经验数值，需要根据不同系统和应用的需求来进行配置。通常不应设置的过大，否则会影响云主机的性能。
- 在生产环境中，如果用户打算采用物理内存超分，建议在安装系统时，配置相应的swap分区。



注：

例如，如果物理机内存为100G，并打算设置物理内存超分为2，建议设置swap分区也为100G，这样的配置在实际使用中，有足够的swap空间供内存超分使用。

- 主存储超分率：

- 默认值为1.0。如果主存储可用空间为2T，设置为2，那么ZStack会认为系统可以分配4T的空间给云主机使用。
- 允许范围 ≥ 1 。该数值不应设置的过大。
- 用户需完全理解该设置的含义并设置正确的主存储使用阈值后，才能进行相应的设置。否则会有严重的数据丢失风险！

- 主存储使用阈值：

- 默认值为0.9。为了防止系统过度使用主存储空间（尤其是当设置了主存储的超分比例后，过度分配云盘有可能使存储溢出，从而导致云主机存储失效崩溃），需要设置主存储阈值。



注：

例如，阈值为0.9，如果当前主存储空间实际使用率到达总容量的90%，整个系统将不能新建云主机或者云盘。此时用户需添加更多的云盘给主存储，并手动重新连接计算节点后，即可进行新建云盘操作。

- 主存储使用阈值允许范围是0~1范围内的小数。当使用大于1的主存储超分率后，该阈值应设定为0.6或更小，以确保可及时添加主存储。用户需完全理解该设置的含义，才能进行相应的设置。否则会有严重的数据丢失风险！

- 云路由管理员密码：

登录云路由设备的密码。admin可以直接在这里设置任意长度字符串，该操作对所有云路由生效。



注：停止云路由后，系统自动再次启动后生效，手动在控制台内重启无效。

- 云主机**Network Anti-Spoofing**：

防 IP/MAC 伪造和 ARP 欺骗。

- 删除策略：

可以选择三种策略，立刻删除（Direct），延时删除（Delay）和永不删除（Never）。该策略设置会设定用户在删除云主机，云盘以及镜像相关内容的删除规则。

1. 立刻删除（Direct）：

当设置为立刻删除时，如果用户删除云主机或者其他资源，这些相关资源会被立刻删除。

2. 延时删除（Delay）：

当设置为延时删除（默认）时，如果用户删除云主机或者其他资源，这些资源会被标记为已删除，显示在对应资源的已删除栏，等彻底延时删除时延（默认时延24小时，也就是86400秒）超时后或用户手动强制删除时，才会彻底删除相关资源。

3. 永不删除（Never）：

当设置为永不删除时，当用户删除云主机或者其他资源，这些资源永远不会被系统自动删除。

- 彻底删除时延：

当删除策略为延时删除时，可以选择延时多久彻底删除资源。ZStack默认时延24小时，也就是86400秒。

- **NUMA**：

- 打开NUMA选项可支持在线修改CPU内存，此操作支持CentOS7.2、CentOS6.6、Ubuntu14.04、Ubuntu16.04等云主机操作系统，不建议在生产环境中对Windows云主机执行在线修改CPU、内存的操作。
- 建议对Windows云主机关机后再修改配置。
- 云主机显卡类型：

管理员可以进入基础设置页面更改云主机启动时默认的显卡类型。更改后可通过 `ps -ef | grep qemu` 命令查看 `-device` 后的显示类型，检查qemu终端设置是否与对应的VM一致。

ZStack 提供更改以下三种显卡类型：

1. cirrus：

cirrus提供一种简单的显卡类型，但对某些操作系统，无法提供更好的显示支持；（`-device cirrus-vga`）

2. vga：

vga是一种提供更好分辨率的显卡类型；（`-device VGA`）

3. qxl：

qxl显卡类型在SPICE协议下能够表现出更好的性能。（`-device qxl-vga`）或通过 `virsh dumpxml $domainID` 指令查看中的是否对应。`$domainID`指云主机UUID。



注：

该选项改变后，只针对新创建的云主机和停止后再启动的云主机生效。

11.7.1.2 高级设置

ZStack支持上百种高级设置，用户可根据自身的需求搜索并且修改对应的设置。目前高级设置支持以下选项，具体介绍如下：

- 云路由设备**agent**自动启动：
 - 默认为false，默认情况下，云路由设备已集成了agent程序，支持开机自动启动，如果设置并部署新的agent，并打算使其生效，此选项需设置为true。
 - 设置为true后，在云路由设备在每次创建、启动或重启时，管理节点均会部署新的agent程序。
- 云路由设备连接超时时间：

默认为300，单位为秒，用于设置管理节点与云路由设备上运行的agent程序的连接超时上限，如果管理节点在指定时间内没有连接成功agent程序，则云路由设备会启动失败并被删除。

- 云路由设备**ssh**超时时间：

默认为300，单位为秒，用于设置管理节点与云路由设备进行ssh连接的超时上限，如果管理节点在指定时间内没有ssh成功云路由设备，则云路由设备会启动失败并被删除。

- 镜像服务器**Ping**间隔：

默认为60，单位为秒，用于设置管理节点在指定的时间间隔去检查镜像服务器，如果检查成功，表示镜像服务器是连接状态。

- 镜像服务器**Ping**最大并发度：

默认为50，用于设置管理节点同时检查多个镜像服务器时，最大的并发度。

- 镜像服务器保留容量：

默认为1G，用于设置镜像服务器在使用时保留使用的容量。

- **Ceph**镜像下载超时时间：

默认为3600，单位秒，用于设置Ceph镜像服务器下载镜像时的超时时间，如果超过此时间，镜像还未下载完成，此镜像会添加失败。

- **Ceph**镜像服务器**Mon**自动重连：

默认为true，用于设置Ceph镜像服务器的Mon IP在检查失败时，是否开启自动重连。

- **Ceph**镜像服务器**Mon**重连延迟：

默认为30，单位秒，用于设置Ceph镜像服务器的Mon IP在一次检查失败后，再次重连的延迟间隔。

- **Ceph**镜像缓存清理间隔：

默认为43200，单位秒，用于设置Ceph镜像已删除后，也无云主机使用的前提下，Ceph主存储里面不再使用的镜像缓存被执行清理的时间间隔。

- **Ceph**主存储删除**Pool**选项：

默认为false，在删除Ceph主存储时，是否删除所有的ceph存储池，默认关闭，打开此选项，非常危险，可能导致数据丢失，需谨慎操作。

- **Ceph**主存储**Mon**自动重连：

默认为true，用于设置Ceph主存储的Mon IP在检查失败时，是否开启自动重连。

- **Ceph**主存储**Mon**重连延迟：

默认为30，单位秒，用于设置Ceph主存储的Mon IP在一次检查失败后，再次重连的延迟间隔。

- 配置信息的私钥：

和公钥一起作为一个密钥对，用于远程连接各物理资源时无密码登录，与用户在系统设置的密钥无关。

- 配置信息的公钥：

和私钥一起作为一个密钥对，用于远程连接各物理资源时无密码登录，与用户在系统设置的密钥无关。

- 控制台agentPing间隔：

默认为60，单位秒，用于设置检查控制台agent程序的时间间隔。

- 控制台agent空闲超时：

默认为60，单位秒，用于设置空闲的控制台agent等待的最大时间。

- Fusionstor镜像下载超时时间：

默认为3600，单位秒，用于设置Fusionstor镜像服务器下载镜像时的超时时间，如果超过此时间，镜像还未下载完成，此镜像会添加失败。

- Fusionstor镜像服务器Mon自动重连：

默认为true，用于设置Fusionstor镜像服务器的Mon IP在检查失败时，是否开启自动重连。

- Fusionstor镜像服务器Mon重连延迟：

默认为30，单位秒，用于设置Fusionstor镜像服务器的Mon IP在一检查失败后，再次重连的延迟间隔。

- Fusionstor镜像缓存清理间隔：

默认为43200，单位秒，用于设置Fusionstor镜像已删除后，也无云主机使用的前提下，Fusionstor主存储里面不再使用的镜像缓存被执行清理的时间间隔。

- Fusionstor主存储删除Pool选项：

默认为false，在删除Fusionstor主存储时，是否删除所有的Fusionstor存储池，默认关闭，打开此选项，非常危险，可能导致数据丢失，需谨慎操作。

- Fusionstor主存储Mon自动重连：

默认为true，用于设置Fusionstor主存储的Mon IP在检查失败时，是否开启自动重连。

- Fusionstor主存储Mon重连延迟：

默认为30，单位秒，用于设置Fusionstor主存储的Mon IP在一次检查失败后，再次重连的延迟间隔。

- 孤儿GC清理间隔：

默认为60，单位秒，在管理节点以指定时间间隔扫描孤儿GC，若有则进行接管。

- 云主机高可用全局开关：

默认为true，用于设置云主机高可用功能。如果关闭此选项，则云主机不支持设置高可用。（强调危险，关闭则整体关闭VM HA功能）

- 高可用异常物理机检查间隔：

默认为5，单位为秒，用于设置在指定的时间间隔去检查一个异常的物理机。

- 高可用判定物理机失联尝试次数：

默认为12，用于设置在指定的最大次数去重连物理机，如果在指定最大次数内连接此物理机均失败，则判定物理机失联。

- 高可用判定物理机成功间隔：

默认为5，用于判定物理机成功连接的时间间隔，如果物理机在此间隔内成功返回，则表示连接成功。

- 高可用判定物理机成功尝试几率：

小数，默认为0.5，用于设置在指定连接次数的既有成功又有失败的情况下，以多大几率来判定物理机连接成功。

- 高可用判定物理机成功次数：

默认为3，用于设置在物理机连接成功的判定时，需要至少建立几次成功的连接。

- 高可用物理机自检间隔：

默认为5，设置物理机状态自检的间隔。

- 高可用物理机自检最大尝试次数：

默认为12，用于设置在指定的最大次数去自检物理机，在最大的尝试测试次数均失败时，会判定物理机网络异常。

- 高可用物理机自检主存储检查超时时间：

默认为15，单位为秒，用于设置在物理机自检连接主存储的超时时间。

- 高可用永不停止云主机尝试的延迟时间：

默认为60，单位为秒，针对永不停止的云主机，如果上一次尝试启动失败，到下一次尝试时中间的延迟时间。

- 高可用永不停止云主机扫描间隔：

默认为60，单位为秒，针对永不停止的云主机，如果启动失败，下次扫描的时间间隔。

- 物理机Ping不通时自动重连：

默认为true，管理节点连不通物理机时，会自动重连。

- 物理机自动重连：

默认为true，管理节点在启动服务过程中，是否全部重连物理机。

- 物理机重连并发度：

默认为100，管理节点在重连物理机过程中，全部重连物理机时的最大并发度。

- 物理机维护模式忽略错误：

默认为false，物理机维护模式，忽略相关错误，会返回成功。

- 物理机Ping间隔：

默认为60，单位为秒，管理节点在指定间隔对物理机进行检查，来物理机是否连接。

- 物理机Ping并发度：

默认为100，管理节点在检查物理机是否连接时，支持的最大并发度。

- 会话清理间隔：

默认为3600，用于设置在指定的时间间隔清理哪些已经超时的会话。

- 会话最大会话数：

默认为500，用于设置管理节点可接受的最大会话数量，超过此数值，将不能建立新的会话。

- 会话超时时间：

默认为7200，当前会话登录超过该会话时间后，系统将不可用，需重新登录。

- 镜像删除策略：

可选类型为直接删除、延迟删除、永不删除，用于在删除镜像时，管理节点设置的删除策略：

1. 直接删除：直接从数据库删除和文件级别删除镜像；
2. 延迟删除：将此镜像的状态变为已删除，在镜像彻底删除到期后，会从数据库删除和文件级别删除镜像；
3. 永不删除：从数据库删除镜像，但是永不删除其文件。

- 镜像彻底删除间隔：

默认为3600，单位为秒，用于设置执行彻底删除镜像任务的时间间隔。

镜像彻底删除周期：

默认为86400，单位为秒，用于设置将一个已删除的镜像从数据库移除，并从镜像服务器移除的执行周期。

- 镜像启用重置密码：

默认为true，用于设置镜像是否启用重置密码。

- **KVM**云主机云盘的最大数量：

默认为24，用于设置一个云主机可支持的最大数据盘数量。

- **KVM**物理机并发级别：

默认为10，用于设置在物理机执行命令的并发数量，如果设置的并发数量过大，可能导致物理机负载过高。

- **KVM**物理机保留内存：

默认为1G，用于设置为物理机上保留的内存容量。

- **KVM**云主机缓存模式：云主机缓存模式设置,可选模式为：writethrough, none, writeback。

- writethrough: 物理机的页面缓存工作在透写模式，数据完全写入云主机存储设备后，才返回成功。
- None: 云主机不使用物理机的页面缓存，直接访问存储，不带cache。
- writeback: 云主机使用了物理机的页面缓存机制，数据写入物理机页面缓存即报告给云主机返回成功。

- **KVM**云主机控制台模式：

支持vnc和spice协议，默认为vnc，用于设置云主机控制台连接的协议类型。

- **KVM**云主机**CPU**模式：

- 选择云主机的CPU类型是否与物理机的CPU类型一致。
- 当选择host-model类型时，云主机的CPU类型将与物理机的CPU类型相符（例如都显示为Haswell的Intel CPU）
- 当选择host-passthrough类型，云主机的CPU类型将与物理机的CPU完全一致。
- 当选择host-model或者host-passthrough类型时，云主机可以支持嵌套虚拟化，但可能导致云主机在不同型号CPU的物理机之间迁移失败。

- **KVM云主机并发迁移数量：**

默认为2，用于设置物理机维护模式时，同时迁移云主机的数量。

- **KVM物理机Ping时云主机状态同步：**

默认为true，用于设置物理机在连接时，是否同步云主机的状态。

- **负载均衡默认算法：**

可选方法为轮询法、最少连接法和源地址法，默认为轮询，用于设置负载均衡时，采用哪种算法。

- **负载均衡空闲连接超时时间：**

默认为60，在没有数据传输时，设置触发负载均衡器终止服务器和客户端连接的超时时间。

- **负载均衡健康检查间隔：**

默认为5，用于设置对负载均衡器内的云主机进行健康检查的时间间隔。

- **负载均衡健康检查目标：**

健康检查的协议类型，默认为TCP，支持TCP, HTTP, HTTPS, or SSL，用于设置对负载均衡器内的云主机进行健康检查的协议类型，对应的有效端口是1到65535。

- **负载均衡健康检查超时时间：**

默认为2，单位为秒，用于设置在指定时间内如果负载均衡检查没有相应则判定检查失败。

- **负载均衡健康边界：**

默认为2，在执行多少次成功检查后，将云主机移至健康状态。

- **负载均衡最大连接：**

默认为5000，用于设置负载均衡器同时建立的最大连接数。

- **负载均衡不健康边界：**

默认为2，在执行多少次失败检查后，将云主机移至不健康状态。

- **本地存储迁移：**

默认为true，本地存储迁移全局设置打开或关闭。打开此开关，则支持本地热迁移。

- **日志选项：**

默认为true，设置是否开启资源的日志记录。

- **日志语言：**

默认为zh_CN，设置记录相关日志的地域，此设置影响相关的日志记录的地域时间。

- 管理节点心跳检测间隔：

默认为5，单位为秒，用于设置 管理节点将心跳写入数据库的时间间隔。

- **ZStack**分发镜像：

默认为true，在添加镜像后，设置是否分发镜像到本地存储。

- **ZStack**分发镜像并发度：

默认为2，在添加镜像后，同时分发镜像到本地存储的并发度。

- 内存超分率：

- 默认值为1.0。如果物理内存为4G，设置为2，那么ZStack会认为系统可以分配8G的内存给云主机使用。允许范围 ≥ 1 ，该数值为经验数值，需要根据不同系统和应用的需求来进行配置。
- 通常不应设置的过大，否则会影响云主机的性能。在生产环境中，如果用户打算采用物理内存超分，建议在安装系统时，配置相应的swap分区。



注：

例如，如果物理机内存为100G，并打算设置物理内存超分为2，建议设置swap分区也为100G，这样的配置在实际使用中，有足够的swap空间供内存超分使用。

- 主存储超分率：

- 默认值为1.0。如果主存储可用空间为2T，设置为2，那么ZStack会认为系统可以分配4T的空间给云主机使用。允许范围 ≥ 1 。该数值不应设置的过大。
- 用户需完全理解该设置的含义并设置正确的主存储使用阈值后，才能进行相应的设置。否则会有严重的数据丢失风险！

- 主存储使用阈值：

- 默认值为0.9。为了防止系统过度使用主存储空间（尤其是当设置了主存储的超分比例后，过度分配云盘有可能使存储溢出，从而导致云主机存储失效崩溃），需要设置主存储阈值。



注：

例如，阈值为0.9，如果当前主存储空间实际使用率到达总容量的90%，整个系统将不能新建云主机或者云盘。此时用户需添加更多的云盘给主存储，并手动重新连接计算节点后，即可进行新建云盘操作。

- 主存储使用阈值允许范围是0~1范围内的小数。当使用大于1的主存储超分率后，该阈值应设定为0.6或更小，以确保可及时添加主存储。用户需完全理解该设置的含义，才能进行相应的设置。否则会有严重的数据丢失风险！

- 监控物理机间隔：

默认为1，单位为秒，收集物理机监控数据的时间间隔。

- 监控云主机间隔：

默认为1，单位为秒，收集云主机监控数据的时间间隔。

- 虚拟网络DHCP MTU设置：

默认为1500，用于设置虚拟网络的默认最大传输单元的字节数。

- L2NoVlan网络DHCP MTU设置：

默认为1500，用于设置L2NoVlan网络的最大传输单元的字节数。

- L2Vlan网络DHCP MTU设置：

默认为1500，用于设置L2Vlan网络的最大传输单元的字节数。

- L2Vxlan网络DHCP MTU设置：

默认为1500，用于设置Vxlan网络的最大传输单元的字节数。

- NFS主存储GC清理时间间隔：

默认为3600，单位为秒，用于设置NFS主存储GC清理云盘和快照的时间间隔。

- 主存储镜像缓存的GC时间间隔：

默认为86400，单位为秒，用于设置主存储定期扫描那些孤儿的需要被清理的镜像缓存的时间间隔。

- 主存储Ping间隔：

默认为60，单位为秒，管理节点在指定间隔对主存储进行检查是否连接。

- 主存储Ping并发度：

默认为100，管理节点在检查多个主存储时，支持的最大并发度。

- 主存储的预留容量：

默认为1G，用于设置主存储的预留容量。

- 进度条清理线程间隔：

默认为300，单位为秒，用于设置清理过期的进度条记录的时间间隔。

- 进度条开启：

默认true，用于设置是否开启进度条功能。

- 进度条记录保存时间：

默认为86400，默认为秒，用于设置进度条数据在数据库中进保留的最长时间。

- 弹性IP数量默认配额：

默认为20，设置账户的弹性IP数量默认配额。

- 镜像数量默认配额：

默认为20，设置账户的镜像IP数量默认配额。

- 镜像容量默认配额：

默认为10T，设置账户的镜像容量默认配额。

- 三层网络数量默认配额：

默认为20，设置账户的三层网络数量默认配额。

- 负载均衡器数量默认配额：

默认为20，设置账户的负载均衡器配额。

- 端口转发数量默认配额：

默认为20，设置账户的端口转发数量默认配额。

- 定时任务数量默认配额：

默认为80，设置账户的定时任务数量默认配额。

- 安全组数量默认配额：

默认为20，设置账户的安全组数量默认配额。

- 云盘快照数量默认配额：

默认为200，设置账户的云盘快照配额。

- 虚拟IP数量默认配额：

默认为20，设置账户的虚拟IP数量默认配额。

- CPU数量默认配额：

默认为60，设置账户的CPU数量默认配额。

- 内存大小默认配额：

默认为80G，设置账户的内存大小默认配额。

- 运行中的云主机数量默认配额：

默认为100，设置账户的运行中的云主机数量默认配额。

- 数据云盘容量默认配额：

默认为10T，设置账户的数据云盘容量默认配额。

- 数据云盘数量默认配额：

默认为40，设置账户的数据云盘数量默认配额。

- 安全组出口策略：

默认为允许，没有安全组规则时，默认的出口规则是允许。

- 安全组进入策略：

默认为禁止，没有安全组规则时，默认的进入规则是禁止。

- **Shared Mount Point**删除GC间隔：

默认为3600，单位为秒，用于设置Shared Mount Point主存储的GC去清理过期云盘以及快照的时间间隔。

- 云路由设备**agent**自动启动：

- 默认为false，默认情况下，云路由设备已集成了agent程序，支持开机自动启动，如果设置并部署新的agent，并打算使其生效，此选项需设置为true。
- 设置为true后，在云路由设备在每次创建、启动或重启时，管理节点均会部署新的agent程序。

- 云路由命令并发度：

默认为100，用于设置云路由可以执行命令的最大并发度。

- 云路由重启**dnsmasq**边界：

默认为500，在发送多少次SIGUSER1后重启DNSMASQ，因为如果有非常多的SIGUSER1的话DNSMASQ会在重新加载配置文件的时候卡住。

- 云路由**Ping**间隔：

默认为60，单位为秒，用于设置管理节点在指定的时间间隔去检查云路由设备，如果检查成功，表示云路由设备是连接状态。

- 云路由**Ping**最大并发度：

默认为500，单位为个，用于设置管理节点在同时去检查多个云路由设备时，最大的并发度。

- 云路由设备连接超时时间：

- 默认为300，单位为秒，用于设置管理节点与云路由设备上运行的agent程序的连接超时上限，如果管理节点在指定时间内没有连接成功agent程序，则云路由设备会启动失败并被删除。

- 云路由设备**ssh**超时时间：
 - 默认为300，单位为秒，用于设置管理节点与云路由设备进行ssh连接的超时上限，如果管理节点在指定时间内没有ssh成功云路由设备，则云路由设备会启动失败并被删除。
- 云路由命令并发度：
 - 默认为100，用于设置云路由可以执行命令的最大并发度。
- 云路由**ssh**端口：
 - 默认为22，云路由的ssh登录端口。
- 云路由用户名：
 - 默认为vyos，登录云路由设备的用户名。
- 云路由管理员密码：
 - 默认为virtual123# 云路由管理员账户是：vyos。用户设定密码后需要通过UI重启云路由设备，该密码才会生效。
- 云主机**globalConfig.vm_bootMenu**：
 - globalConfig VM引导菜单：默认为true，用于设置云主机启动时是否启用引导菜单，少数不支持虚拟机引导菜单的服务器可以设置为false。
- 云主机**Network Anti-Spoofing**：
 - 默认为false，该功能防止 IP/MAC 伪造和 ARP 欺骗。
- 云主机声卡：
 - 默认为false，用于设置是否使用AC97的声卡。
- 云主机删除策略：
 - 删除云主机的策略。可选策略为:立刻删除，延时删除，永不删除。
- 云主机彻底删除间隔：
 - 默认为3600，单位为秒，执行彻底删除云主机任务运行的时间间隔。
- 云主机彻底删除时延：
 - 默认为864000，单位为秒，云主机删除策略为延迟删除时，针对已删除的云主机，选择延时多久彻底删除云主机。
- 计算规格删除后设置为空：默认为true，用于设置在删除计算规格时，是否更新云主机的计算规格列表为NULL。
- **NUMA**：

默认为false，表示云主机的架构采用NUMA结构。

- 云主机显卡类型：

默认为cirrus，设置云主机启动时默认的显卡类型。可选择类型为:cirrus, vga, qxl。

- 快照增量的最大数目：

默认为128，一个云盘快照链的最大长度。当一个云盘的快照链的长度达到这个数字时，下一个云盘快照会是一个完整的快照。

- 云盘删除策略：

删除数据云盘有三种策略：立刻删除，延时删除，永不删除。

1. 立刻删除：被删除的云盘会从数据库和主存储中删除；
2. 延时删除：被删除的云盘会在数据库中被标记为已删除，等彻底延时删除时延超时后，才会彻底从数据库与主存储中删除云盘；
3. 永不删除：当设置为永不删除时，当用户删除云盘时，这些资源永远不会被系统自动删除。

- 云盘规格删除时设置为空：

默认为true，删除磁盘规格时，是否更新磁盘的磁盘规格列表为NULL。

- 云盘彻底删除间隔：

默认为3600，单位为秒，执行彻底删除云盘的任务运行的时间间隔。

- 云盘彻底删除时延：

默认为864000，单位为秒，云盘删除策略为延迟删除时，针对已删除的云盘，选择延时多久彻底删除云盘。

12 管理节点高可用

12.1 管理节点高可用方案

管理节点高可用方案采用共享存储，使用云主机作为管理节点来管理云平台。

分为两种类型：

- 超融合基础架构：基于Ceph分布式块存储，存储与计算融合部署，管理节点以云主机方式运行，管理融合节点的计算、存储、网络资源。
- 共享文件系统高可用：基于NFS或Shared Mount Point共享文件系统存储，独立共享存储，管理节点以云主机方式运行，管理计算节点的计算、存储、网络资源。

ZStack for Alibaba Cloud支持的高可用支持以下三个层级的高可用：

1. 管理节点的高可用：由zsha和consul守护进程来提供管理节点高可用，保证管理节点云主机的高可用运行状态。
2. 业务云主机的高可用：通过对云主机所在的物理网络或节点进行检测，在失效时，会自动切换云主机到其他健康节点，以提供业务的可用性。
3. 存储的高可用：分布式块存储的提供的多副本机制和故障域隔离机制；共享文件系统存储由底层存储提供的高可用保障机制，例如，Raid、主备等机制。

存储可支持Ceph分布式块存储，也支持Shared Mount Point存储，也可使用NAS转化为Shared Mount Point存储。

管理节点高可用方案在部署时，需提前规划网络，支持管理网络和存储网络合并，也支持管理网络和存储网络分离，使用独立的存储网络，ZStack使用存储网络作为基准来判定管理节点主机状态。

13 基本运维

本章主要针对ZStack基本操作异常给出处理办法。

13.1 安装部署

安装、升级部署ZStack失败需检查的因素：

1. 假定安装操作选择了管理节点模式，但是安装系统后，没有安装ZStack。

原因：

安装操作系统时，没有设置合适的IP地址，ZStack管理节点必须要求有可用IP。

解决方案：

- 此时需要执行`bash /opt/zstack*installer.bin -E` 安装ZStack企业版；
- 或执行`bash /opt/zstack*installer.bin` 安装ZStack社区版。

2. 升级时遇到ZStack ISO版本不匹配，会提示类似以下错误信息。

现象：

```
Reason: The current local repo is not suitable for ZStack installation.  
Please download ZStack-x86-64-DVD-2.0.1.iso from http://www.zstack.io/product_downloads/ and run
```

解决方案：

此时需要执行以下步骤：

1. 在管理节点下载最新的ISO，例如在管理节点执行下载ISO：

```
wget http://cdn.zstack.io/product_downloads/iso/ZStack-x86_64-DVD-2.2.x.iso
```

2. 执行以下命令进行升级，升级ISO里面的yum repo同步到本地，并且直接升级管理节点。

```
zstack-upgrade ZStack-x86_64-DVD-2.2.x.iso
```

3. 安装ZStack时出现如下错误信息。

现象：

```
Fail Reason: /usr/local/zstack is existing. Please delete it manually before installing a new  
ZStack
```

原因：

此时系统已经存在ZStack，不允许再次安装。

解决方案：

- 如果打算升级 直接使用-u参数升级即可。
 - 如果打算重新安装，则需执行`rm -rf /usr/local/zstack; bash zstack-installer.bin -D`进行重新安装。
4. 安装部署失败，失败信息会打印到屏幕上。

解决方案：

可查看`/tmp/zstack_installation.log`，根据错误提示，尝试解决安装部署问题。

13.2 启动服务

ZStack相关的服务在管理节点重启后，会自动启动。

启动ZStack服务失败需检查的因素：

1. 数据库应正常运行：

可通过`systemctl status/start/stop/restart mariadb`进行状态/启动/停止/重启检查。

2. Rabbitmq应正常运行：

可通过`systemctl status /start/stop/restart rabbitmq-server`进行状态/启动/停止/重启检查。

3. 出现如下报错：

现象：

```
ERROR: failed to connect to the mysql server[hostname:10.0.0.18, port:3306, user:zstack, password:zstack.password]
```

解决方案：

- 需确保`/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties`里面关于数据库的IP、用户名、密码均正确，且可使用`mysql`命令正常访问。
 - 例如，可通过`mysql -u zstack -pzstack.password -h 10.0.0.18`直接访问。
4. 出现如下报错：

现象：

```
ERROR: unable to connect to the rabbitmq server[ip:10.0.0.18] with username/password configured in zstack.properties
```

原因及解决方案：

需检查rabbitmq状态，有可能是hostname发生变更引起rabbitmq无法连接，此时可执行`zstack-ctl reset_rabbitmq`

5. 出现如下报错：

现象：

```
8080 is occupied by some process. Please use netstat to find out and stop it
```

解决方案：

执行`netstat -an|grep 8080`找到8080占用程序将其停止再启动ZStack。

6. 出现如下报错：

现象：

```
MN status: Unknown, the management node seems to become zombie as it stops responding APIs
```

原因及解决方案：

一般是因为内存不足导致的消息总线不能正确处理，需扩大内存，此时不建议再将管理节点作为计算节点添加到ZStack使用。

13.3 zstack-ctl命令

ZStack提供`zstack-ctl`对服务或者资源进行命令行控制。

常用的`zstack-ctl`命令：

- 状态检查：`zstack-ctl status`
- 服务控制：`zstack-ctl stop/start/stop_ui/start_ui/restart_node/`
- 日志收集：`zstack-ctl collect_log`
- 备份数据库：`zstack-ctl dump_mysql`
- 恢复数据库：`zstack-ctl restore_mysql -f back.gz --mysql-root-password password`
- 管理节点IP变更：`zstack-ctl change_ip`
- 重置rabbitmq：`zstack-ctl reset_rabbitmq`
- 检查配置：`zstack-ctl show_configuration`
- 安装授权：`zstack-ctl install_license -f license_key.txt`
- 重置admin账户密码：`zstack-ctl reset_password`

- 修改数据库密码：`zstack-ctl change_mysql_password --root-password zstack.mysql.password --user-name zstack --new-password password`

13.4 zstack-cli命令

ZStack提供**`zstack-cli`**命令行交互界面对ZStack内部各资源进行控制。

1. **`zstack-cli`**登录：

```
LoginByAccount accountName=admin password=password
```

初始密码为password，如果有修改，则需按照正确的密码登录。

2. **`zstack-cli`**支持的关键字包括：

- 资源关键字：
zone cluster host vm primarystorage image L2 L3 backupstorage volume offering
- 操作关键字：
add reconnect start stop destroy delete create get set update attach detach

3. 使用**`zstack-cli`**进行控制的技巧：

- 输入资源关键字和操作关键字。
- 善用tab键。例如，输入集群的关键字cluster，点击两次tab键后，会列出cluster可支持的所有操作，如下：

```
AttachL2NetworkToCluster
AttachPrimaryStorageToCluster
ChangeClusterState
CreateCluster
DeleteCluster
DetachL2NetworkFromCluster
DetachPrimaryStorageFromCluster
GetCandidateZonesClustersHostsForCreatingVm
GetVmStartingCandidateClustersHosts
QueryCluster
QueryVCenterCluster
UpdateCluster
```

此时假定打算再创建集群，根据上面列出的Cluster操作，那么输入CreateCluster后，再点击两次tab键，cli将列出CreateCluster参数。

4. 使用**zstack-cli**创建云主机的方法：

```
zstack-cli CreateVmInstance name=ceph instanceOfferingUuid=$instanceoffer I3NetworkUids=$L3uuid imageUuid=$image
```

5. 使用**zstack-cli**查询运行中的云主机：

```
zstack-cli QueryVmInstance state= "Running" name~=test limit=3 field=uuid
```

名字包含test，只显示3条记录，只显示其uuid。

13.5 重连物理机

重连物理机会重新部署代理程序进行相关检查。

重连物理机失败需检查的因素：

1. 管理节点无法正常访问物理机22端口，例如以下，需先解决管理节点可以正常ssh到物理机的问题。

```
ssh: connect to host 192.168.0.65 port 22: Network is unreachable
```

2. 物理机应可访问管理节点，以将agent代理的结果返回给管理节点。
3. 没有使用ZStack定制版ISO安装计算节点，导致升级失败，出现类似以下软件包安装冲突的错误，需要使用ZStack定制版ISO来重新安装系统，才可解决。

```
ERROR: command pkg_list=`rpm -q openssh-clients qemu-kvm bridge-utils wget qemu-img libvirt-python libvirt nfs-utils vconfig libvirt-client net-tools iscsi-initiator-utils lighttpd dnsmasq iproute sshpass rsync nmap | grep \"not installed\" | awk '{ print $2 }'` && for pkg in $pkg_list ; do yum --disablerepo=* --enablerepo=zstack-local install -y $pkg; done;failed!
```

4. 物理机的libvirt服务状态应正常：

可通过systemctl status/start/stop/restart libvirtd进行状态/启动/停止/重启检查。

5. 重连物理机会检查其所属集群的网络挂载情况，假如集群挂载的二层网络指定的设备是eth1，但是物理机的网卡设备号发生变化，不存在eth1，那么此物理机将无法重连成功。
6. 如果物理机系统盘的存储容量已经用完，此时物理机将无法重连成功。
7. 对于要求挂载NFS主存储的物理机，如果挂载失败，物理机将无法重连成功，此时需要检查物理机到NFS主存储的链路，例如以下错误，就需检查物理机到NFS存储的连接。

```
failed to execute shell command: mount 192.168.18.242:/home/Storage/zhu /opt/zstack/nfsprimarystorage/prim-751565713a6340da884c33a8668a8916
```

8. 使用分布式块存储的物理机，如果存在存储网络，则要求物理机可直接访问存储网络。

9. 如果使用VXLAN网络，其挂载集群时指定的VTEP IP应该存在，如果不存在，物理机将无法重连成功。

13.6 重连主存储

重连主存储会重新部署代理程序进行相关检查。

重连主存储失败需检查的因素：

1. 本地存储一般处于可连接状态。
2. NFS主存储保证有可用物理机可以连接此主存储。重连失败时，需检查NFS Server的IP地址、NFS共享目录、NFS挂载参数是否正确。
3. Ceph主存储需检查分布式块存储是否正常。重连失败时，需检查Ceph IP地址、Ceph SSH 端口、用户名、密码、用户名是否有sudo权限，Ceph集群状态是否正常。
4. Fusionstor主存储需检查分布式块存储是否正常。重连失败时，需检查Fusionstor IP地址、Fusionstor SSH 端口、用户名、密码、用户名是否有sudo权限、Fusionstor集群状态是否正常。
5. Shared Mount Point主存储需在物理机手动挂载此目录，建议在`/etc/rc.local`里进行定义，开机自启动挂载。

13.7 重连镜像服务器

重连镜像服务器会重新部署代理程序进行相关检查。

重连镜像服务器失败需检查的因素：

1. Sftp镜像服务器和镜像仓库需检查SSH端口号、用户名、密码、用户名是否拥有sudo权限。
2. Ceph镜像服务器需检查分布式块存储是否正常。重连失败时，需检查Ceph IP地址、Ceph SSH 端口、用户名、密码、用户名是否有sudo权限，Ceph集群状态是否正常。
3. Fusionstor镜像服务器需检查分布式块存储是否正常。重连失败时，需检查Fusionstor IP地址、Fusionstor SSH 端口、用户名、密码、用户名是否有sudo权限，Fusionstor集群状态是否正常。

13.8 云主机异常

云主机相关异常处理：

异常信息	异常原因及解决办法
<p>"code" : " HOST.1005" , " description " : " Failed to start vm on hypervisor" , "details" : " failed to start, Libvirt error : internal error no supported architecture for os type 'hvm' " 或 "details": "failed to start vm on kvm host, because unable to start vm, libvirt error: invalid argument : could not find capabilities for domain type =kvm "</p>	<ul style="list-style-type: none"> 原因 : ZStack需要物理机支持并打开了硬件虚拟化 (如vmx或者svm) 的支持。出现该错误的原因通常是因为BIOS中没有打开CPU虚拟化。 解决方案 : 重启系统, 并进入BIOS打开CPU相关虚拟化支持。
<p>"details" : "cannot find either 'vmx' or 'svm' in /proc/cpuinfo, please make sure you have enabled virtualization in your BIOS setting"</p>	<ul style="list-style-type: none"> 原因 : ZStack需要物理机支持并打开了硬件虚拟化 (如vmx或者svm) 的支持。出现该错误的原因是用户添加的物理机的CPU不支持硬件虚拟化功能。 通常是由于用户添加的物理机是一个 “云主机” 导致的, 需要打开对应 “云主机” 的嵌套虚拟化功能。 解决方案 : 具体的打开办法, 请联系 “云主机” 提供方。
<p>"details": "the local primary storage has no hosts with enough disk capacity[xxx bytes] required by the vm[uuid:xxx]"</p>	<ul style="list-style-type: none"> 原因 : 当前系统中 (或者任何单独一台物理机) 没有足够的主存储容量 (磁盘空间)。 ZStack在分配云主机和云盘时使用的是thin clone 模式, 也就是云主机使用多少, 才分配多少。但是在创建云主机的时候, 会按照云主机最大申请的使用空间来扣除系统主存储的容量, 以防最后云主机使用的空间超过系统可用空间。 此外, 如果主存储为本地存储, 由于云主机的磁盘不能够跨物理机存储, 所以如果每台物理机的剩余空间都不足创建一台新的云主机, 也会遇到此错误。 解决方案 : 删除一些不需要使用的云主机、云盘, 或者调整云主机使用的镜像大小。

异常信息	异常原因及解决办法
	<p>在理解主存储超分的原理和使用方法的前提下，也可以增大主存储超分比例以获得更多超分空间。</p>
<p>"details": "unable to allocate hosts; due to pagination is enabled, there might be several allocation failures happened before ; the error list is [{no host having cpu[x HZ], memory [xxx bytes] found}]"</p>	<ul style="list-style-type: none"> 原因： 当前系统中（或者任何单独一台物理机）没有足够的内存。 解决方案： 删除一些不需要使用的云主机，或者调整云主机使用计算规格的大小。在理解内存超分的原理和使用方法的前提下，也可以增大存储超分比例以获得更多超分容量。
<p>"details": "failed to start vm[uuid:xxx name :xxx] on kvm host[uuid:xxx, ip:x.x.x.x], because unable to start vm[uuid:xxx, name :xxx], libvirt error: internal error: early end of file from monitor: possible problem : Cannot set up guest memory 'pc.ram ': Cannot allocate memory" set up guest memory 'pc.ram':Cannot allocate memory"</p>	<ul style="list-style-type: none"> 原因： 当前系统中实际可用物理内存不足以分配给需要创建的云主机。当这个错误出现时，就算提高内存超分比例，也没有办法创建更大容量的内存。 解决方案： 减小云主机的计算规格，增加物理主机的内存，或者增加物理主机的swap分区（会降低系统性能）。
<p>"failed to migrate vm[uuid:xxx] from kvm host[uuid:xxx, ip:xxx] to dest host[ip:xxx], unable to migrate vm[uuid:xxx] to qemu +tcp://xxx/system, Unsafe migration: Migration may lead to data corruption if disks use cache != none"</p>	<ul style="list-style-type: none"> 原因： 当前的基础设置中，将云主机的磁盘缓存模式设置为其他模式，例如，writethrough或writeback，系统认为带缓存模式的迁移是不安全的。 解决方案： 需要将此基础设置的缓存模式修改为None模式才可迁移。
<p>"description": "Failed to start vm on hypervisor", libvirt error: Unable to create tap device vnic15.0: Device or resource busy</p>	<ul style="list-style-type: none"> 原因： 此计算节点可能被重复添加，里面已经有使用了vnic15.0的云主机在运行，导致设备出现冲突。 解决方案：

异常信息	异常原因及解决办法
	建议添加物理机时，需保持一个干净的环境，或清理掉其上已经运行的云主机。
Permission denied: '/opt/zstack/nfsprimarystorage/prim-cd8075752a0b4c669afa79acfc433ca5/dataVolumes'	<ul style="list-style-type: none"> 原因： 创建云主机或数据盘时，NFS主存储的目录没有读写的权限， 解决方案： 需在NFS 服务器端开启读写权限。
internal error: unable to execute QEMU agent command 'guest-set-user-password': failed to set password: The user name could not be found.	<ul style="list-style-type: none"> 原因： 修改云主机密码时，指定的用户名不存在。 解决方案： 需要指定已存在的用户名。
internal error: unable to execute QEMU agent command 'guest-set-user-password': The command guest-set-user-password has not been found	<ul style="list-style-type: none"> 原因： 修改云主机密码时，在云主机内部没有安装qemu-ga或者代理程序版本过低。 解决方案： 可执行qemu-ga -V检查版本。 <ul style="list-style-type: none"> 在CentOS7.2里版本是2.3.0 在Windows 2012里版本是0.12.1
"description": "An operation failed","details": "unable to live migrate with local storage . The vm[uuid:0837761acd4ffab2f75798cdb65c38] has volumes on local storage,to protect your data, please stop the vm and do the volume migration"	<ul style="list-style-type: none"> 解决方案： 本地存储的云主机在线迁移时，需要先卸载相关的数据云盘。
"libvirt error: error from service: ListActivatableNames: Connection is closed"或"libvirt error: Activation of org.freedesktop.machine1 timed out"	<ul style="list-style-type: none"> 原因： 这类错误通常是由系统DBUS进程死掉或者DBUS运行异常引起的。 解决方案： <ul style="list-style-type: none"> 方法1：重启ZStack物理机，DBUS会自动恢复。 方法2：在物理机上使用如下命令：

异常信息	异常原因及解决办法
	<ol style="list-style-type: none"> 1. service systemd-machined restart 2. service libvirtd restart 3. 完成后用UI或者CLI重连出问题的物理机。
<pre>{ "error": { "code": "HOST.1009", "description": "Failed to migrate vm on hypervisor", "details": "failed to migrate vm[uuid:90bb61f55e774f5f89c9e1ea46db7661] from kvm host[uuid:e5105c61114a4efe8bcc025f744226bd, ip:10.0.247.244] to dest host[ip:10.0.197.238], unable to migrate vm[uuid:90bb61f55e774f5f89c9e1ea46db7661] to qemu+tcp://10.0.197.238/system, internal error: Attempt to migrate guest to the same host localhost.domain.com" }, "createTime": 1488170137857, "type": { "_name": "key.event.API.API_EVENT" }, "id": "e6bcd7c33ad9446b919e11cd647fc8dc" }</pre>	<ul style="list-style-type: none"> 原因： 这是因为两台物理机hostname相同，KVM在线迁移目前依靠hostname进行确认不同的物理机，如果hostname相同，则会被认定为相同的物理机。 解决方案： 请确保物理机hostname不能相同。可以通过hostnamectl set-hostname your_new_hostname来设定。
all mons failed to execute http call[/ceph/primarystoragevolume/clone	<ul style="list-style-type: none"> 解决方案： 重连主备存储。
<p>为什么冷迁移会失败？"SYS.1006"</p> <p>简介：</p> <p>An operation failedfailed to execute bash [rsync -a --relative /data/rootVolumes/acct-36c27e8ff05c4780bf6d2fa65700f22e/vol-3f1225391f414d84b1b025761440b297/snapshots/c0bb52f90d194a55a0ae47468d503816.qcow2 --rsh="/usr/bin/sshpass -p mevoc0_-03 ssh -o StrictHostKeyChecking=no -l mevoco" 10.20.40.32:/], return code: 23, stdout: , stderr: Warning: Permanently added '10.20.40.32' (ECDSA) to the list of known hosts. rsync: failed to set times on "/data": Operation not permitted (1) rsync: failed to set times on "/data/rootVolumes": Operation not permitted (1) rsync: failed to set times on "/data/rootVolumes/acct-36c27e8ff05c4780bf6d2</p>	<ul style="list-style-type: none"> 原因： 权限属组问题。 Operation not permitted (1) rsync: recv_generator: mkdir "/data/rootVolumes/acct-36c27e8ff05c4780bf6d2fa65700f22e/vol-3f1225391f414d84b1b025761440b297" failed: Permission denied (13) *** Skipping any contents from this failed directory *** rsync error: some files/attrs were not transferred 解决方案： 需要重新给正确的属组，确认要迁移的VM文件是不是root qemu组。

异常信息	异常原因及解决办法
<p>fa65700f22e": Operation not permitted (1) rsync: recv_generator: mkdir "/data/rootVolumes/acct-36c27e8ff05c4780bf6d2fa65700f22e/vol-3f1225391f414d84b1b025761440b297" failed: Permission denied (13) *** Skipping any contents from this failed directory *** rsync error: some files/attrs were not transferred (see previous errors) (code 23) at main.c(1052) [sender=3.0.9]</p> <p>详情:</p> <p>failed to execute bash[rsync -a --relative /data/rootVolumes/acct-36c27e8ff05c4780bf6d2fa65700f22e/vol-3f1225391f414d84b1b025761440b297/snapshots/c0bb52f90d194a55a0ae47468d503816.qcow2 --rsh="/usr/bin/sshpass -p mevoc0-_-03 ssh -o StrictHostKeyChecking=no -l mevoco" 10.20.40.32:/], return code: 23, stdout: , stderr: Warning: Permanently added '10.20.40.32' (ECDSA) to the list of known hosts. rsync: failed to set times on "/data": Operation not permitted (1) rsync: failed to set times on "/data/rootVolumes": Operation not permitted (1) rsync: failed to set times on "/data/rootVolumes/acct-36c27e8ff05c4780bf6d2fa65700f22e": Operation not permitted (1) rsync: recv_generator: mkdir "/data/rootVolumes/acct-36c27e8ff05c4780bf6d2fa65700f22e/vol-3f1225391f414d84b1b025761440b297" failed: Permission denied (13) *** Skipping any contents from this failed directory *** rsync error: some files/attrs were not transferred (see previous errors) (code 23) at main.c(1052) [sender=3.0.9]</p>	
<p>failed to start vm[uuid:4ef5a9e4ae6441c7b046c384c3ae6f7f name:stack_local] on kvm host[uuid:d0d9800a09c34da4891e0cc624a0d349, ip:192.168.1.166], because Traceback (most recent call last): File "/"</p>	<ul style="list-style-type: none"> 原因： 这是因为管理平台把host-passthrough开启而导致。 解决方案：

异常信息	异常原因及解决办法
var/lib/zstack/virtualenv/kvm/lib/python2.7/site-packages/kvmagent/plugins/vm_plugin.py", line 2721, in start_vm self._start_vm (cmd) File "/var/lib/zstack/virtualenv/kvm/lib/python2.7/site-packages/kvmagent/plugins/vm_plugin.py", line 2664, in _start_vm 'unable to start vm[uuid:%s, name:%s]', libvirt error: %s' % (cmd.vmInstanceId, cmd.vmName, str(e))) KvmError: unable to start vm[uuid:4ef5a9e4ae6441c7b046c384c3ae6f7f, name:stack_local], libvirt error: internal error: early end of file from monitor, possible problem: qemu-kvm: /root/rpmbuild/BUILD/qemu-2.6.0/target-i386/kvm.c:1736: kvm_put_msrs: Assertion 'ret == n' failed.	关闭此选项即可。
"code": "SYS.1006", description: An operation failed, details: failed to check physical network interfaces on lvm	<ul style="list-style-type: none"> 原因： 嵌套虚拟化。 解决方案： 建议云主机CPU模式设置成passthrough。
description": "A message or a operation timeout", "details": "[Async Http Timeout] url: http://10.10.10.253:7272/init, timeout after 300000[MILLISECONDS], command : {\"uuid\": \"78183b0a46094e30bc8a6128b30ee8cb\", \"restartDnsmasqAfterNumberOfSIGUSER1\": 0}"	<ul style="list-style-type: none"> 原因： 很可能是管理节点的IP地址无法访问出去。 一般出现这个情况是多网卡的服务器在安装管理服务时指定了第一个默认路由的网卡地址。 解决方案： 打开/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties 修改management.server.ip = xxx.xxx.xxx.xxx这个IP地址，修改为可以访问出去的IP地址。
failed : Error in connection establishment: net::ERR_CONNECTION_REFUSED	<ul style="list-style-type: none"> 解决方案： 打开/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties

异常信息	异常原因及解决办法
	将 consoleProxyOverriddenIp=xxx.xxx.xxx.xxx 改为管理节点的IP地址，修改完后要重启节点才生效。
Could not access KVM kernel module: No such file or directory failed to initialize KVM : No such file or directory.	<ul style="list-style-type: none"> 原因： 这个虚拟机不支持虚拟化技术。 解决方案： 需要先开启嵌套虚拟化。

13.9 日志分析

相关日志记录地址：

类型	日志路径	日志所在机器
管理节点日志	/usr/local/zstack/apache-tomcat/logs/management-server.log	管理节点
管理节点UI日志	/usr/local/zstack/apache-tomcat/logs/zstack-ui.log	管理节点
控制台代理日志	/var/log/zstack/zstack-console-proxy.log	管理节点
管理节点部署日志	/var/log/zstack/deploy.log	管理节点
shell命令日志	/var/log/zstack/zstack.log	物理机
zstack kvmagent 日志	/var/log/zstack/zstack-kvmagent.log	物理机
VM对应的libvirt 日志	/var/log/libvirt/qemu/vm_uuid.log	物理机
系统基本日志	/var/log/messages	管理节点/物理机等
镜像仓库日志	/var/log/zstack/zstack-store/zstore.log	镜像仓库
ceph 主存储日志	/var/log/zstack/ceph-primarystorage.log	Ceph Mon节点
ceph 备份存储日志	/var/log/zstack/ceph-backupstorage.log	Ceph Mon节点
Fusionstor 主存储日志	/var/log/zstack/fusionstor-primarystorage.log	Fusionstor Mon节点
Fusionstor 备份存储日志	/var/log/zstack/fusionstor-backupstorage.log	Fusionstor Mon节点
云路由设备的日志	/home/vyos/zvr/zvr.log	云路由设备

类型	日志路径	日志所在机器
扁平网络的配置及日志	<code>/var/lib/zstack/dnsmasq/</code>	物理机

日常运维出现错误时，进行错误分析的一般步骤如下：

1. 根据UI界面提示的错误信息进行简单分析。

- 例如，创建云主机失败时提示的信息是no host found, 那么此时可能的原因就是选择的镜像所在的镜像服务器，可选的集群加上选择的网络进行筛选后，没有可用的物理机。

此时需要检查集群内是否有可用的物理机，是否挂载相应的网络，集群是否挂载了相应的主存储。

2. 分析管理节点`/usr/local/zstack/apache-tomcat/logs/management-server.log`关于UI界面提示的错误信息。

- 查询提示错误的信息，根据上下文进行分析，分析发生异常时的工作流。
- 也可根据执行的操作API进行查询分析，例如创建云主机调用的API是APICreateVmInstance，在管理节点搜索这个API，往下逐步分析错误原因。

3. 再根据错误的具体信息，查看相关代理日志的错误信息。

- 例如，在物理机的`/var/log/zstack/zstack-kvmagent.log`查看对云主机进行生命周期控制时的代理错误日志。

13.10 网络脚本

ZStack定制的ISO提供了以下网络脚本方便配置网络：

1. `zs-show-network`用于显示网络状态，无须额外参数，显示网络连接及链路聚合状态。

```
[root@localhost ~]# zs-show-network
1: lo: <LOOPBACK,UP,LOWER_UP> mtu 65536 qdisc noqueue state UNKNOWN
   link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00
   inet 127.0.0.1/8 scope host lo
       valid_lft forever preferred_lft forever
   inet6 ::1/128 scope host
       valid_lft forever preferred_lft forever
2: vmnic0: <BROADCAST,MULTICAST,UP,LOWER_UP> mtu 1500 qdisc pfifo_fast state
   UP qlen 1000
   link/ether fa:84:ba:34:92:00 brd ff:ff:ff:ff:ff:ff
   inet 172.20.54.194/16 brd 172.20.255.255 scope global vmnic0
       valid_lft forever preferred_lft forever
   inet6 fe80::f884:baff:fe34:9200/64 scope link
       valid_lft forever preferred_lft forever
```

Bond Name	SLAVE	BONDING_OPTS

2. zs-network-setting用于配置物理网卡网络命令。用法有两种：

a. 方法一：

```
# 参数-i，意为对接口 ( interface ) 配置网络地址

zs-network-setting -i [interface] [ipaddress] [netmask] [gateway]
                    接口      地址      掩码      网关

# 例子1：zs-network-setting -i eth0 192.168.1.10 255.255.255.0 192.168.1.1
# 例子2：zs-network-setting -i eth0.10 192.168.1.10 255.255.255.0 192.168.1.1
```

b. 方法二：

```
# 参数-b，意为基于接口 ( interface ) 创建网桥 ( bridge ) 并配置网络地址

zs-network-setting -b [interface] [ipaddress] [netmask] [gateway]
                    接口      地址      掩码      网关

# 例子1：zs-network-setting -b eth0 192.168.1.10 255.255.255.0 192.168.1.1
# 例子2：zs-network-setting -b eth0.10 192.168.1.10 255.255.255.0 192.168.1.1
```

3. zs-change-nic用于修改网卡名命令，修改网卡名后，脚本会将配置到/etc/udev/rules.d/70-persistent-net.rules，物理主机重启后命名仍保持新的命名。

```
# 参数-c，意为修改 ( change ) 物理网口名字

zs-change-nic -c [old-nic-name] [new-nic-name]
                当前网卡名    新的网卡名

# 例子：zs-change-nic -c eth0 em01
```

4. zs-vlan用于配置VLAN接口。有两种用法：

a. 方法一：创建VLAN

```
# 参数-c，意为创建 ( create ) VLAN接口

zs-vlan -c [nic-name] [vlan]
          指定网卡  VLAN号

# 例子：zs-vlan -c zsnic0 10
```

b. 方法二：删除VLAN

```
# 参数-d，意为删除 ( delete ) VLAN接口

zs-vlan -d [nic-name] [vlan]
          指定网卡  VLAN号

# 例子1：zs-vlan -d zsnic0 10
```

```
# 例子2：zs-change-nic -c eth0 em01
```



注:

- 创建VLAN接口后，将会在/etc/sysconfig/network-scripts/创建对应的网络配置文件，物理主机重启后VLAN接口仍生效。删除VLAN接口后，对应网络配置将会删除。
- 创建VLAN接口后，若基于此端口与其他设备通信，需要参照网络设备厂商的VLAN配置手册，以允许该VLAN接口流量传输。
- 以下提供常见网络设备VLAN设定参考：
 - 华为：点击[这里](#)。
 - 华三：点击[这里](#)。
 - 思科：点击[这里](#)。
 - 瞻博：点击[这里](#)。

5. `zs-bond-lacp`用于创建删除Bonding链路聚合命令。有两种用法：

a. 方法一：创建Bonding 链路聚合

```
# 参数-c，意为创建（create）链路聚合接口  
zs-bond-lacp -c [bond-name]  
聚合接口  
  
# 例子：zs-bond-lacp -c bond0
```

b. 方法二：删除链路聚合

```
参数-d，意为删除（delete）链路聚合接口  
zs-bond-lacp -d [bond-name]  
聚合接口  
  
# 例子：zs-bond-lacp -d bond0
```



注:

本命令创建的接口是基于IEEE 802.3ad动态链路聚合协议（LACP），该接口需要管理员设定物理主机接入层网络交换机的端口组（Port-Group）。

6. `zs-nic-to-bond`用于配置链路聚合命令。

a. # 参数-a，意为加载（attach）物理接口到聚合接口

```
zs-nic-to-bond -a [bond-name] [nic-name]  
聚合接口 物理接口
```

```
# 例子：zs-nic-to-bond -a bond0 em1
```

b. # 参数-d，意为从聚合接口卸载（detach）物理接口

```
zs-nic-to-bond -d [bond-name] [nic-name]
                聚合接口    物理接口
```

```
# 例子：zs-nic-to-bond -d bond0 em1
```



注:

该命令执行后，会反馈执行接口，并显示聚合接口与物理接口的关系，例如：

```
-----
| Bond Name | SLAVE(s) | BONDING_OPTS |
-----
| bond0 | enp1s0f1 | miimon=100 mode=4 xmit_hash_policy=layer2+3 |
|       | enp1s0f0 |                               |
-----
```

- 其中，bond0是聚合接口，enp1s0f0和enp1s0f1是物理接口，显示其绑定关系。
- 此外，enp1s0f0和enp1s0f1对应的接入网络设备的物理端口组成端口组（Port-Group），管理员需依照网络设备型号配置端口组。
- 以下提供常见网络设备端口组设定参考：
 - 华为：点击[这里](#)。
 - 华三：点击[这里](#)。
 - 思科：点击[这里](#)。
 - 瞻博：点击[这里](#)。

14 其他配置运维

本章主要针对高级用户，提供一些场景下的配置运维指南。

14.1 安装多个管理节点

自动安装

如果用户希望搭建一个高可用的产品级云环境，那么可以部署多个管理节点以保证管理节点的高可用。如有对数据库和消息总线有高可用的需求，可参考ZStack 高可用集群快速配置进行配置。

用户通过[安装部署](#)已经安装完成一个管理节点，使用`zstack-ctl add_multi_management`命令安装。使用方法如下：

```
zstack-ctl add_multi_management --host-list root:passwd1@host1_ip root:passwd2@host2_ip
```



注：例如

```
zstack-ctl add_multi_management --host-list root:password@172.20.12.47 root:password@172.20.13.216
```

如图 78: ZStack多管理节点安装界面所示,172.20.14.154为第一个管理节点，使用改命令添加172.20.12.47和172.20.13.216两个管理节点。

安装完毕后，ZStack会使用这三台主机作为管理节点同时提供服务，任一管理节点服务中断，不影响其他管理节点的控制服务。



注：此时Mysql、Rabbitmq的服务依然存在在第一台管理节点上。如有需求，可以将其分别配置。

图 78: ZStack多管理节点安装界面

```
[root@172-20-14-154 ~]# zstack-ctl add_multi_management --host-list root:password@172.20.12.47 root:password@172.20.13.216
Checking system and init environment: ... PASS
Add public key to host 172.20.12.47: ... PASS
Deploy management node to host 172.20.12.47: ... PASS
Config management node on host 172.20.12.47: ... PASS
Install UI on host 172.20.12.47: ... PASS
Start management node on host 172.20.12.47: ... PASS
Add public key to host 172.20.13.216: ... PASS
Deploy management node to host 172.20.13.216: ... PASS
Config management node on host 172.20.13.216: ... PASS
Install UI on host 172.20.13.216: ... PASS
Start management node on host 172.20.13.216: ... \
All management nodes add successfully
```

手动安装

当管理大量物理机的时候，为了提高可用性，用户可以安装多个ZStack管理节点。多个ZStack管理节点之间关系平等，相互支撑。用户可以使用如下的方法安装：

```
zstack-ctl install_management_node --host=ip_of_machine_to_install_node_2 --yum=alibase,aliepel
```

当新节点成功安装后，请在新节点上配置新管理节点的IP地址：

```
zstack-ctl configure management.server.ip=ip_of_management_node2  
zstack-ctl save_config
```

可以重复这一步来安装更多的管理节点。如果启动了多个管理节点，在批量创建云主机时，需要修改安装了mysql的管理节点的配置文件`/etc/my.conf`，将其中的最大连接数修改为适合的值，来增加可接受的最大并发数量。假如有三个管理节点，启动一百个云主机，最大并发数量至少需高于300，下例设置为500，可作为参考：

```
vim /etc/my.conf  
max_connections = 500
```

如果用户购买了ZStack使用授权，请联系销售厂商给每一个管理节点单独安装一个授权协议。

多管理节点自动升级

ZStack一条命令即可完成多节点的升级：

```
zstack-ctl upgrade_multi_management_node --installer-bin zstack-installer.bin
```

14.2 配置独立的MySQL服务器

ZStack在安装过程中会配置一个本地的数据库。用户也可以安装独立的MySQL数据库。

安装独立的MySQL数据库后，用户需使用下面的ZStack的工具来初始化数据库，并进行对应的设置。

- MYSQL_ROOT_PASSWORD是MySQL的根用户密码
- MYSQL_ZSTACK_PASSWORD是用户希望设置的ZStack用户的MySQL密码
- MYSQL_SERVER_IP是mysql数据库的物理机IP地址

输入命令后，会要求用户输入MYSQL服务器的ssh密码用于远程登录：

```
sudo zstack-ctl deploydb --root-password="MYSQL_ROOT_PASSWORD" --zstack-password="MYSQL_ZSTACK_PASSWORD" --host=" MYSQL_SERVER_IP"
```

14.3 配置独立的RabbitMQ服务器

同样可以使用`zstack-ctl`来完成RabbitMQ服务的安装：

```
sudo zstack-ctl install rabbitmq  
--host=RABBITMQ_MACHINE_IP
```

当成功的安装了RabbitMQ之后，还需要为远程的RabbitMQ访问创建加密访问（需要在RabbitMQ的服务器上做如下操作）：

```
rabbitmqctl add_user zstack ZSTACK_USER_PASSWORD  
rabbitmqctl set_user_tags zstack administrator  
rabbitmqctl change_password zstack ZSTACK_USER_PASSWORD  
rabbitmqctl set_permissions -p / zstack ".*" ".*" ".*"
```

另外还需要把RabbitMQ的加密访问配置记录到`zstack.properties`（在ZStack管理节点做如下步骤）：

```
zstack-ctl configure CloudBus.rabbitmqUsername=zstack  
zstack-ctl configure \  
CloudBus.rabbitmqPassword= ZSTACK_USER_PASSWORD  
zstack-ctl save_config
```

当用户安装有多个管理节点的时候，需要到每个管理节点上配置正确的RabbitMQ访问方式。

14.4 zstack.properties核心配置文件

`zstack.properties`是ZStack的核心配置文件。它会存放在每一个管理节点中。

`zstack.properties`文件中会存放诸如数据库URL，用于数据库访问用户名密码，RabbitMQ的IP地址等等。每一个管理节点上的`zstack.properties`文件的内容基本上是一致的。它的路径可以通过`zstack-ctl status`来获得。

如果是默认安装的话，它会存放在`/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties`

14.5 ZStack计算节点密码变更

ZStack的物理机列表中的相应物理主机的密码，在发生变更后，需要通知ZStack系统进行变更维护，否则可能导致此台计算节点无法继续正常工作。

变更计算节点的密码可以在点击计算节点界面的物理机IP后的修改按钮即可修改。

14.6 ZStack管理节点更改物理环境

ZStack的授权与用户的物理机的环境——匹配。如果用户更新了管理节点物理机硬件配置，可能会导致ZStack的授权失效，而启动失败。

这个时候请联系ZStack的销售和技术支持获取帮助。

14.7 ZStack数据库备份与恢复

为保护数据安全，防止数据丢失，ZStack现已主动定期进行数据库备份。默认使用`crontab`进行控制。

用户使用`crontab -l`命令可查看管理节点相关的自动定时备份任务。

用户也可以使用`crontab -e`的界面里修改备份策略，默认为每天的凌晨和中午12:30进行备份，并且最多保留最新的14次备份。

默认的数据库备份文件存放路径：`/var/lib/zstack/mysql-backup/`

建议用户使用异地备份策略来提高安全性。只要数据库文件存在，即使ZStack的相关机器挂掉，均可使用数据库文件进行恢复。

14.8 本地存储空间扩容

背景信息

在使用ZStack产品的过程中，如果存储空间容量不足，可以通过以下方式进行存储空间扩容。

以下为系统主存储空间`/zstack_ps`扩容的样例：

操作步骤

1. 设置待扩容的物理机进入维护模式。
2. 插入新的大容量磁盘至物理机，并进行分区格式化，假如设备名称为`/dev/sdc`
3. 创建新目录`/new_volume`，挂载此磁盘`mount /dev/sdc1 /new_volume`
4. 执行`rsync -a /zstack_ps /new_volume/`拷贝原始的磁盘文件至新磁盘
5. 添加以下内容至`/etc/fstab`，将物理磁盘挂载点与`/zstack_ps`目录绑定

```
/dev/sdc1 /new_volume auto defaults 0 0
```

```
/new_volume /zstack_ps none defaults,bind 0 0
```

6. 重启物理机，使用`df -h /zstack_ps`检查扩容后的空间增量

```
Filesystem Size Used Avail Used% Mounted on  
/dev/vdb1 4.0T 304G 3.5T 8% /zstack_ps
```

7. 重启ZStack服务

```
zstack-ctl stop  
zstack-ctl start
```

8. 重连Host继续使用ZStack



注：备份存储的扩容方法与此类似，只是无须对Host进行维护模式及重连的操作。

15 常见问题

1.

Q：ZStack支持哪些硬件环境？

A：ZStack支持带硬件虚拟化特性的X86服务器、工作站或台式机，支持主流的Intel或AMD的处理器。

2.

Q：ZStack支持哪些操作系统？

A：ZStack目前只支持ZStack定制的CentOS7.2系统。

3.

Q：ZStack安装部署需要几台机器？

A：最小安装只需要一台机器，可同时作为管理节点和计算节点，但在生产环境中建议分离部署。ZStack只需一台管理节点，就可以管控ZStack内所有的硬件资源。

4.

Q：ZStack安装部署机器的最小配置要求？

A：4CPU，8G内存，主存储200G，备份存储200G空间。最小配置仅适用于测试环境。

5.

Q：ZStack支持使用标准的CentOS7.2系统安装ZStack么？

A：不支持，需使用ZStack定制的CentOS7.2 ISO来安装引导系统。

6.

Q：如何检查一个机器或者虚拟机是否支持硬件虚拟化？

A：在终端中执行`egrep "vmx|svm" /proc/cpuinfo`，如果有输出代表支持硬件虚拟化。

7.

Q：如何在BIOS打开硬件虚拟化支持？

A：BIOS中默认是开启硬件虚拟化支持。如果有修改，则需在BIOS打开Intel Virtual Technology或Secure Virtual Machine选项，保存并退出。

8.

Q：ZStack定制版ISO可以安装哪些模式及其差异？

A：定制版ISO支持五种模式：

- 企业版管理节点：安装ZStack定制版CentOS7.2并安装ZStack企业版管理节点。
- 社区版管理节点：安装ZStack定制版CentOS7.2并安装ZStack社区版管理节点。
- 计算节点：安装ZStack定制版CentOS7.2并安装计算节点必备软件包。
- 存储节点：安装ZStack定制版CentOS7.2使用UEK内核及OCFS2组件。
- 专家模式：安装ZStack定制版CentOS7.2并配置本地源，用户可自定义系统用途。

9.

Q：如何安装ZStack企业版？

A：下载ZStack定制版ISO，在安装系统时，选择企业版管理节点模式，安装完毕操作系统重启后，会自动安装相关的管理节点。

- 如果由于缺少网卡配置或者其他原因导致导致ZStack安装失败，都会退出到终端。
- 如安装ISO时，如果服务器没有可用IP，会导致无法正常安装ZStack管理节点。可以配置网络后，再执行以下命令进行ZStack的安装。



注：

- ZStack企业版安装执行`bash /opt/zstack*installer.bin -E`

- ZStack社区版安装执行`bash /opt/zstack*installer.bin`

10.

Q：如何升级ZStack企业版？

A：升级方法有两种途径：

1. 下载ZStack定制版ISO和ZStack升级脚本，执行`bash zstack-upgrade ZStack-x86_64-DVD-2.2.x.iso`直接升级本地的软件仓库和ZStack管理节点。
2. 下载ZStack定制版ISO和ZStack升级脚本，执行`bash zstack-upgrade -r ZStack-x86_64-DVD-2.2.x.iso`先升级本地仓库，再执行`bash ZStack-installer-2.2.x.bin -u`升级管理节点。

11.Q：如何从专家模式或计算节点升级到ZStack企业版管理节点？

A：执行`yum install zstack-manager; bash /opt/zstack*installer.bin -E`

12.

Q：如何从ZStack社区版升级到ZStack企业版管理节点？

A：获取最新的企业版许可证书，安装并加载许可证书即可。

13.

Q：如何从ZStack企业版变更到ZStack社区版？

A：在管理节点删除`/var/lib/zstack/license/license.txt`许可证书，安装并加载许可证书即可变更到社区版功能。

14.

Q：ZStack支持哪些浏览器？

A：支持Chrome和Firefox浏览器。

15.

Q：如何处理修改hostname后ZStack管理节点服务启动失败？

A：修改hostname会导致rabbitmq不可用，需执行`zstack-ctl reset_rabbitmq`来重置rabbitmq。

16.

Q：UI无法登录怎么办？

A：需检查：

- 如果提示用户名密码不正确，则需输入正确的用户名和密码。
- 如果提示无法连接管理节点，则需确保管理节点服务运行正常，可通过`zstack-ctl status`检查服务均运行正常。
- 如果依然无法连接，需使用`zstack-cli`在cli里来登录。
- 检查相关的Java进程是否出现僵尸导致程序无法运行，此时需要重启管理节点机器。

17.

Q：ZStack支持哪些速率的网络接口？

A：支持千兆、万兆、40G等，云平台环境建议千兆以上。

18.

Q：ZStack的DHCP服务和目前机房的DHCP是否会出现冲突？

A：ZStack的DHCP服务与目前机房已经配置的DHCP服务不冲突，但ZStack里面的云主机使用的网络段不可与当前环境内已经使用的IP范围重叠，如果发生重叠，则可能发生IP地址冲突。

19.

Q：ZStack的计算节点支持的Qemu和Libvirt版本是？

A：可在计算节点执行`virsh version`进行检查，Qemu支持Qemu 2.6.0, Libvirt支持1.3.3。

```
Compiled against library: libvirt 1.3.3
Using library: libvirt 1.3.3
Using API: QEMU 1.3.3
```

Running hypervisor: QEMU 2.6.0

20.

Q：管理节点重启后，为何物理机长时间处于连接中？

A：管理节点机器或者服务重启后，需要重新建立与各硬件资源的连接，此过程会重新部署相关代理程序。

21.

Q：如何修改控制台代理IP？

A：修改`/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties`里面的`consoleProxyOverriddenIp`后，再重启管理节点服务。

22.

Q：扁平网络云主机hostname会变成与IP地址相关的？

A：如果镜像里面原本的hostname是localhost，DHCP会主动将hostname修改成与IP地址相关，例如，云主机IP地址为192.168.12.3，那么云主机hostname会变成192-168-12-3。

23.

Q：`zstack-cli`里面设置扁平网络的hostname为何不生效？

A：如果镜像里面原本的hostname不是localhost,那么DHCP将不会分配新的hostname，而是使用原本的hostname。

24.

Q：同一个三层网络是否支持设置不同的子网掩码和网关？

A：ZStack不支持在同一个三层网络里设置不同的子网掩码和网关，这种配置会造成网络异常。

25.

Q：控制台代理出现"description": "Cannot do an IO operation(e.g. network IO error, file IO error)"如何处理？

A：

- 现象：

```
"description": "Cannot do an IO operation(e.g. network IO error, file IO error)",
"details": "I/O error on POST request for \"http://127.0.0.1:7758/console/establish\":\nConnect to 127.0.0.1:7758 [/127.0.0.1] failed: Connection refused;\nnested exception is org.apache.http.conn.HttpHostConnectException: Connect to 127.0.0.1:7758 [/127.0.0.1]\nfailed: Connection refused", "$$hashKey": "object:1344"
```

- 解决方案：

需要删除`/var/lib/zstack/pickledb/`下面的consoleProxy文件，再次重连控制台代理。

26.

Q：管理节点IP地址变化了怎么办？

A：管理节点IP变更，可以通过执行 `zstack-ctl change_ip --ip 172.20.12.88`来将当前的管理节点IP地址变更。执行过程中会重置rabbitmq。也可单独修改消息总线和数据库的IP地址，可根据`zstack-ctl change_ip -h`的提示进行变更。

27.

Q：如何更改UI的Admin密码？

A：有两种方法：

1. 直接在UI界面点击Admin信息进行修改；

2. 在管理节点上运行如下命令（如果password已经更改，请用改过的密码替换password）：

- `zstack-cli LoginByAccount accountName=admin password=password`
- 然后运行如下命令（NEW_PASSWORD 就是新要使用的密码）：
- `zstack-cli UpdateAccount uuid=36c27e8ff05c4780bf6d2fa65700f22e password=NEW_PASSWORD`

28.Q：如何手动延长ZStack管理节点启动时间？

A：

- 如果用户在虚拟机中运行ZStack，很可能会遇到启动超时失败，例如系统提示：

```
no management-node-ready message received within 120 seconds, \
please check error in log file /tmp/zstack_installation.log
```

- 如果之后运行`zstack-ctl status`又显示ZStack为Running的状态那么可以试用如下的参数直接增加timeout的时间：

```
zstack-ctl start_node --timeout 300
```

29.

Q：Vmware ESXi设定嵌套虚拟化的虚拟机方法？

A：

1. 方法一：

登陆VMware ESXi控制台打开VMware ESXi 5.0的ssh服务，然后用ssh登陆 VMware ESXi 后在config文件（`vi /etc/vmware/config`）中最后加入`vhv.enable = "TRUE"`一行。

2. 方法二：

如果不能重启整改ESXi，就通过vSphere下载想要设置嵌套虚拟化的虚拟机的config文件（xxx.vmx），然后在该文件最后加上`vhv.enable = "TRUE"`一行。做这个操作前需要停止这个虚拟机。加上设置后，再把修改的config拷贝覆盖原有配置文件，重新启动虚拟机。

30.

Q：如何重新安装ZStack？

A：删除千万小心，如果需要彻底重装ZStack，只需要两步：

1. `rm -rf /usr/local/zstack`
2. ZStack-installer.bin是用户预先下载的ZStack 安装包

```
bash ZStack-installer.bin -E #安装ZStack企业版
bash ZStack-installer.bin #安装ZStack社区版
```

31.

Q：在VMWare的嵌套虚拟化的虚拟机里创建ZStack的VR VM失败怎么办？

A：需要在VMware的VSwitch设备上打开混杂模式，以及Vlan号。

32.

Q：Cli中Windows VM怎么设置使用VirtIO？

A：

- ZStack 默认创建Windows虚拟机的根云盘使用IDE，而数据云盘使用VirtIO。当用户给Windows虚拟机安装了VirtIO，驱动后也可以让Windows虚拟机的根云盘使用VirtIO设备。

- 方法：给该虚拟机添加一条系统标签：

```
zstack-cli CreateSystemTag resourceType=VmInstanceVO \
resourceUuid=TARGET_VM_UUID tag=windows::virtioVolume
```

33.

Q：为什么ZStack报告主、备份存储容量和物理机上看到的内容不一致？

A：

- ZStack使用的是thin clone模式（copy on write技术），所以VM可以很快被创建。在KVM环境下，不论VM的镜像文件是10G还是100G，VM创建的时候只有一个很小的qcow2的独立文件被创建出来。这个qcow2文件和原始的镜像文件一起共同组成了新的VM的硬盘。当有新的数据产生的时候，该qcow2文件的大小会不断增加。文件大小的上限为原始VM的镜像文件的配置上限（例如10G，20G）也同VM的操作系统启动后看到的硬盘的实际大小相一致。
- 由于ZStack默认不支持资源超分（超卖），所以ZStack在计算空间的时候，会按照VM使用空间的上限来扣除可用空间的数量。于是用户在系统上用df命令看到的可用空间可能还有很大，但是ZStack已经把VM未来可能会占用的所有空间都已经计算在内了。于是就会导致用户可能看到硬盘上还有很多空间，但是不能创建虚拟机。

34.

Q：如何修改UI界面默认的5000端口？

A：

1. 先停止UI `zstack-ctl stop_ui`
2. 假定待使用的端口号为8889，则以下命令可在8889端口访问UI

```
zstack-ctl start_ui --server-port 8889
```

35.

Q：如何指定非默认路由的IP地址启动ZStack服务？

A：

1. 先用下面的命令设置管理节点的IP地址：

```
zstack-ctl configure management.server.ip=YOUR_EXPECTED_IP_ADDRESS
```

2. 然后重启zstack即可。

36.

Q：如何增加ZStack UI用户和zstack-cli的session过期时间？

A：

- ZStack UI界面全局配置下找到 session 对应的 timeout 设置，调整会话过期时间。
- 也可通过 `zstack-cli` 命令修改。例如下例把 `zstack account` 登陆的过期时间改成200个小时：

```
zstack-cli LoginByAccount accountName=admin password=password \
zstack-cli UpdateGlobalConfig name=session.timeout category=identity value=720000 zstack-
cli LogOut
```

37.

Q：希望从IP Range中保留一个IP地址,不让ZStack分配给云主机该怎么办？

A：

- 目前ZStack还没有提供 `ReserverIpRange`的API，如果我们希望ZStack从已经设置的IP Range中保留几个特定的IP地址，可以在UI界面创建指定的VIP方式来保留IP使用。

- 也可以用CreateVip这个API，例如下面的命令将会把a.b.c.d的IP地址从指定的L3网络上用创建VIP的方法保留住：

```
zstack-cli LoginByAccount accountName=admin password=password
zstack-cli CreateVip l3NetworkUuid=YOU_L3_NETWORK_UUID name=for_reserver require
dlp=a.b.c.d zstack-cli Logout
```

38.

Q：如何批量修改一批云主机的计算规格？

A：

- 在UI界面可以批量选中一批云主机，进行批量修改；
- 也可采用shell脚本配合**zstack-cli**命令进行批量修改，例如，下面是将一批名字里包含Win7的云主机的计算规格全部修改为名字是Win7-Instance-Offering的计算规格。

```
which jq || (echo "you need to install jq" && exit 1)
zstack-cli LoginByAccount accountName=admin password=password instance_offering_uuid
=zstack-cli \
QueryInstanceOffering name=Win7-Instance-Offering |jq '["inventories"]|[0].uuid' target_vms
=zstack-cli zstack-cli \
QueryVmInstance name=~Win7|jq '["inventories"]|[0].uuid' for vm in $target_vms; do zstack-cli
ChangeInstanceOffering \
instanceOfferingUuid=$instance_offering_uuid vmInstanceUuid=$vm echo "change vm: $vm
instance offering to \
$instance_offering_uuid" zstack-cli StopVmInstance uuid=$vm zstack-cli StartVmInstance uuid
=$vm done zstack-cli Logout
```

39.

Q：如何解决qemu版本不匹配问题？

A：

- 现象：

在使用过程中，启动虚拟机时可能遇到类似这样的错误信息：

```
uses a qcow2 feature which is not supported by this qemu version: QCOW version 3
```

- 原因：

主要原因是QCOW版本不一致，原始的qcow2创建版本使用的qemu-img为较新版本，现在创建时使用的为较旧版本，旧版本不支持新版本。

- 解决方案：

可在拥有较新版本的qemu-img里面进行兼容性转换，例如执行以下命令进行转换，转换完毕后，再重新添加镜像：

```
qemu-img convert -o compat=0.10 -f qcow2 -O qcow2 centos6-cloud-init.qcow2 centos-st-ssh-
key.qcow2
```

40.

Q：如何释放Flat Network Service Provider占用的IP？

A：

- 原因：

Flat Network Service Provider因为提供DHCP服务，所以会占用一个IP地址。

- 解决方案：

当用户不想使用该服务的时候，可以删除对应L3 network但是该provider所占用的IP地址并不会被主动释放。可用以下方法回收该IP地址，并消除对应的影响。

在所有的物理节点上执行：

1. `ip netns`
2. 对所有输出的namespace执行`ip netns delete xxxx`，xxxx是步骤一输出的对应的namespace
3. 执行`ebtables -F`
4. `pkill dnsmasq`，把所有DHCP server杀死

41.

Q：如何快速修改管理节点、数据库、消息总线的IP地址？

A：

修改更新ZStack的配置文件之前可以手动修改`zstack.properties`。

- 从ZStack 1.1开始，用户可以使用`zstack-ctl`命令修改IP地址：

```
zstack-ctl change_ip --ip MY.NEW.IP.ADDRESS
```

- 用户也可用上述命令单独替换MySQL服务器的IP地址和RabbitMQ的IP地址。具体的参数请使用如下命令查看：

```
zstack-ctl change_ip -h
```



注：用户需要把MY.NEW.IP.ADDRESS替换成需要修改的IP地址。

- 如果主机的IP地址发生了变化，可用如下命令更新IP地址：（调用前需要先用`zstack-cli`登陆管理员权限）

```
zstack-cli LoginByAccount accountName=admin password=password
zstack-cli UpdateKvmHost uuid=HOST_UUID managementIp=NEW.HOST.IP.ADDRESS
zstack-cli LogOut
```

- 如果镜像服务器的IP地址发生了变化，可用如下命令更新IP地址：（调用前需要先用`zstack-cli`登陆管理员权限）

```
zstack-cli LoginByAccount accountName=admin password=password
zstack-cli UpdateSftpBackupStorage uuid=BACKUP_STORAGE_UUID hostname=NEW.HOST.IP.ADDRESS
zstack-cli LogOut
```

42.

Q：计算节点root用户名密码修改后，如何通知ZStack修改？

A：

- 当ZStack的计算节点root用户密码修改后，可以在UI界面直接修改计算节点的IP地址。
- 也可使用UpdateKVMHost的API进行更新。更新方法如下：
- 1. 登录 `zstack-cli`：

```
zstack-cli
```

```
>>>LoginByAccount accountName=admin password=ZSTACK_PASSWORD
```

2. 更新密码，需要先从UI上获得需要修改的host的 UUID：

```
>>>UpdateKVMHost uuid=ed232959bce5471eae5145751f86176e password=NEW_PASSWORD
```

3. 退出登录。

43.

Q：如何把ZStack管理节点从一台物理机迁移到另一台物理机？

A：

1. 原管理端备份数据库，备份文件提示导出至 `/var/lib/zstack/mysql-backup/`

```
[root@old-zstack ~]#zstack-ctl dump_mysql
Backup mysql successful! You can check the file at /var/lib/zstack/mysql-backup/zstack-backup
-db-2016-07-05_18-13-11.gz
```

2. 暂停原管理端服务。

```
[root@old-zstack ~]#zstack-ctl stop
```

3. 将备份数据库转移至新服务器，并在新管理端导入数据库。

```
[root@new-zstack ~]#zstack-ctl restore_mysql --from-file=/root/zstack-backup-db-2016-07-05_18-13-11.gz --mysql-root-password=zstack.mysql.password
the management node has been stopped
Starting recover data ...
Recover data successfully! You can start node by: zstack-ctl start
```

4. 启动新管理端服务。

```
[root@new-zstack ~]#zstack-ctl start
```

5. 这时可以直接登录ZStack的web进行host和storage的管理，如果原有host的IP有变更可以直接在web端修改并重新连接。

44.

Q：使用root安装ZStack为何启动服务提示非root？

A：

- 现象：

安装ZStack时，有时使用root登录，但是启动服务时，提示非root用户需使用 `sudo`

```
ERROR: please use sudo or root user FAIL
Reason: failed to start zstack
```

在Python命令行环境下执行以下命令会发现Python认定用户并非root用户

```
import getpass
user = getpass.getuser()
print user
```

- 原因：

在之前由普通用户切换root用户时，使用了su并非使用sudo su，su只是shell认为切换了root用户，但是相关的环境变量并没有切换。

- 说明：

概述相关区别：

- sudo允许命令在当前用户下以root的超级权限来执行。
- su允许切换为root用户，但是相关的用户环境并没有切换到root下去。
- sudo su以root权限切换到root用户下。

45.

Q：如何修改计算节点IP地址？

A：

- 计算节点IP地址变更后，可以在UI界面直接修改。
- 也可以通过`zstack-cli`API进行修改。
- 举例，将原本的172.20.11.206 IP地址修改成172.20.11.207，具体步骤如下：
 1. 执行 `zstack-cli`进入cli交互界面
 2. 执行 `LoginByAccount accountName=admin password=password` 登录
 3. 执行 `QueryHost fields=uuid,managementIp`查询uuid与物理机IP地址信息。例如：

```
>>>QueryHost fields=uuid,managementIp
{
  "inventories": [
    {
      "managementIp": "172.20.11.206",
      "uuid": "cc6a6f3df9e9424fa2b71b17cddda97f"
    }
  ],
  "success": true
}
```

4. 期望将IP 172.20.11.206修改成172.20.11.207，那么执行以下命令即可：

```
>>>UpdateHost uuid="cc6a6f3df9e9424fa2b71b17cddda97f" managementIp="172.20.11.207"
{
  "inventory": {
    "availableCpuCapacity": 32,
    "availableMemoryCapacity": 3975249920,
    "clusterUuid": "1e819b164911472ab0136b1622d06e54",
    "createDate": "May 25, 2016 9:38:24 AM",
    "description": "HOST-NEW",
    "hypervisorType": "KVM",
    "lastOpDate": "Jul 18, 2016 11:46:08 AM",
    "managementIp": "172.20.11.207",
    "name": "HOST-2",
    "sshPort": 22,
    "state": "Enabled",
    "status": "Disconnected",
    "totalCpuCapacity": 32,
    "totalMemoryCapacity": 3975249920,
    "username": "root",
    "uuid": "cc6a6f3df9e9424fa2b71b17cddda97f",
    "zoneUuid": "a8d1aae19d6648d68a484b0f85b7548a"
  },
  "success": true
}
```

5. 可在UI界面执行重连物理机操作，或在cli里面执行 `ReconnectHost` 即可重连物理机。至此，物理机IP地址更新完毕。

46.

Q：ZStack只认eth0？为何不支持em01？

A：ZStack是根据用户的配置来确认网卡的名称，要求同集群内所有的计算节点网卡编号需要一致。

- 如果原本已添加eth0的二层网络，但是添加新的计算节点的网卡名称只有em01的网卡，却没有eth0的网卡。
- 如果在添加二层网络时，选择了eth0，那么所有的计算节点要求使用同样的eth0来设置网络。
- 用户也可以设置em01作为二层网络来添加，但是同样的要求所有的计算节点均配置em01的网络。
- 用户可以删除L2重新配置。

47.

Q：云盘可以建立在共享存储上，且可以进行任意挂载吗？

A：ZStack可以给一个Zone内不同Cluster添加不同主存储，主存储可以是本地存储或者共享存储。

- 如果主存储为共享存储，则云盘创建在共享存储上，云盘和物理主机无关，可以挂载到任意物理节点的云主机上。即：云盘的挂载不受物理机的相关约束，而且不支持云盘迁移。
- 如果主存储为本地存储，挂载一个卸载过的云盘时，则需要保证该云盘和目标云主机在相同的物理机上。如果该云盘和目标云主机不在相同的物理机上，就需要用户将云盘和云主机迁移到相同的物理机上。用户可以选择迁移云主机（需要首先停止云主机），也可以选择迁移云盘。

48.

Q：包含系统盘和数据盘两个qcow2文件时，导入时如何添加数据盘？

A：

- Add Image的时候，可以add data volume template。
- 创建volume的时候，可以从Data Volume template创建一个volume。使用 **zstack-cli**：

```
>>>CreateDataVolumeFromVolumeTemplate
```

49.

Q：VM上挂载新的L3网络时，为何只显示一个？

A：

- 动态挂载只是KVM的Qemu给VM一个虚拟网卡，使用lspci可以看到此网卡，但是云主机的操作系统使用时，需要手动执行dhclient获取
- 也可参考虚拟机里面的 `/etc/sysconfig/network-scripts/ifcfg-eth0` 对新增的eth1可重新设置下，再重启网络即可看到多网络路由。

50.

Q：管理网络和数据网络放在一起会有影响吗？

A：建议分离，安全性会更高，网络负载分配更合理。如果网络资源不足，可以放在一起。

51.

Q：QueryVolume actualSize怎么跟实际文件大小不一致？

A：随着磁盘使用过程，这个 **actualsize** 可能会不断增加，此时需要使用 SyncVolumeSize uuid= 进行同步。

52.

Q：可以添加两个二层网络吗？

A：可以，先添加L2使用eth0，然后添加L3，使用这个L2；然后添加第二个。分配给VM IP时候，也需要分别选择。

53.

Q：如何新增一块网卡跑数据网络？

A：千兆以上，要规划相应的网络段提供给VM使用，与管理口一样需要填加 onboot=yes。

54.

Q：ZStack如何自动多节点安装？

A :

- 部分用户会用到多管理节点做高可用等场景。多节点手动安装教程：点击[这里](#)。
- 以下主要对多管理节点的自动化安装做介绍，也是ZStack 1.7版本以后支持的一个功能：
 - 具体命令如下：

```
zstack-ctl add_multi_management --host-list root:passwd1@host1_ip root:passwd2@host2_ip
```

- 如果用户已经使用ssh_key来连接各个节点，可以用以下命令添加：--ssh-key此处为保存ssh_key文本文件的绝对路径。

55.

Q：关闭计算节点后，如何实现VM高可用？

A：VM高可用在ZStack上，使用共享存储，设置云主机高可用，需要先将VM 状态置为stopped, 再次启动此VM，可以在其他节点启动。

56.

Q：Detach L3网络，重连后云主机IP会变化吗？

A：在DHCP情况下会随机重新分配IP。

57.

Q：机房准备断电维护，上电后如何自动恢复VM的业务？

A：可参考以下步骤：

1. 设置全部云主机高可用为NeverStop模式（目前此功能ZStack企业版支持，社区版不支持）。
2. 全部物理机进入维护模式。
3. 断电维护。
4. 上电后，物理机启用，所有的云主机会自动恢复运行。

58.

Q：ZStack支持连接spice协议吗？

A：支持，可以在全局设置的基本设置中，设置云主机控制台的显示方式从vnc修改为spice，后续云主机的创建均会使用spice协议。

59.

Q：如何读取其中一个物理节点上的U盘？

A：目前ZStack正在研发，可以手动设定 **passthrough**，或者通过 **usb over tcp** 的方案解决。

60.

Q：创建image的占用空间是怎样的？

A：

- 以举例说明，一个host有800G的空间，创建了一个VM 400G，假定文件是A，然后创建了一镜像B，B会上传到镜像服务器，这个时候B的实际大小应该不是400G，比如只有10G。
- 用B来创建一个新的VM C的时候，B会从镜像服务器上下载到host上的一个cache目录。这个下载会消耗主存储的空间，等到下载成功，host的虚拟可用容量就小于400G了，这样创建VM C就失败了。

61.

Q：ZStack有根据uuid查询信息的API吗？

A：可以先看看/usr/local/zstack/apache-tomcat/logs/management-server.log里面这个uuid的类型，再用**zstack-cli**直接查询该属性的uuid。

62.

Q：如何快速找到需要的API？

A：遇到相关的问题，例如你打算了解Snapshot的API，可以在`cli`里面输入**snap**接着连续点击两次tab，即可显示所有的可用的API，相关的含义从字面意思就可以理解了。

63.

Q：能在ZStack的VM上安装ZStack吗？

A：可以，需要在ZStack全局变量中设置CPU模式为host-passthrogh，透传物理机的虚拟化特性。

64.

Q：云主机磁盘写满，无法启动怎么办？

A：可以扩大磁盘后启动，建议先备份镜像。

65.

Q：ZStack的管理节点是不是有优先分配资源？

A：对于Linux kernel来说没有特别优先，除非通过 `cgroup`手动配置，建议将管理和计算分开。

66.

Q：ZStack给KVM的mac地址，可以写死吗？

A：不能，mac地址由dnsmasq随机分配。

67.

Q：如何为host设置保留内存？

A：在全局配置里修改物理机预留内存的值即可。

68.

Q：如何修改虚拟机的删除时延？

A：删除VM的根云盘会根据全局设置的删除策略来配置，可以在UI界面的全局设置中设置云盘的删除时延，也可以在zstack-cli里面调整，例如执行：`UpdateGlobalConfig category=vm name=expunge Period value=3600`，设置成3600秒触发一次，删除操作。

69.

Q：创建镜像时提示URL相关错误一般有哪些原因？

A：

- ZStack允许输入的URL格式为 `http://`或`file:///`
- 例如：`http://192.168.0.10/test.qcow2`或`file:///tmp/test-CentOS 7.qcow2`



注：

- 建议用户创建一个http服务器来下载镜像。
- 如果使用本地镜像，注意file后必须为绝对路径。

70.

Q：重新安装管理节点，如何恢复原来配置？

A：备份之前管理节点的数据库，使用以下命令：`zstack-ctl dump_mysql --keep-amount 14`，在 `/var/lib/zstack/mysql-backup`目录下查看。

71.

Q：安装好ZStack之后，如何修改MySQL密码？

A：

- 用户一键安装好ZStack后，默认密码如下：
 - MySQL默认的root用户：

```
user = root
```

```
password = zstack.mysql.password
```
 - 默认ZStack用户：

```
user = zstack
```

```
password = zstack.password
```

- 部分用户有修改密码的需求，可以通过以下步骤进行更新。
 - 输入原本的root密码，加上待修改的--user-name，加上待修改的密码，执行后就直接在数据库修改了对应的密码，

```
zstack-ctl change_mysql_password --root-password zstack.mysql.password --user-name zstack --new-password password
```

- 修改完成后，将ZStack的MySQL用户信息写入：`/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties`

例如：

```
DB.user = zstack DB.password = zstack.password
```

72.

Q：当服务器有多个网卡时，安装ZStack报错怎么办？

A：在有多网卡的物理机上，直接使用命令安装，默认会选择一块网卡。如果出现安装错误，可以使用“-i 网卡IP”指定一块网卡。

73.

Q：一键安装ZStack后，启动ZStack时间过长怎么办？

A：

- 原因：

管理节点启动时间与节点配置资源有关，用户往往是在虚拟机上安装ZStack才出现时间过长的情况。

- 解决方案：

建议用户如果用虚拟环境，配置最好CPU>4,内存>8G，时间过长不代表安装启动失败，在time out报错后请先使用 `zstack-ctl status` 查看状态，如果为stopped状态，使用`zstack-ctl start`启动即可；如果为unknown状态，需要先stop，再start。

74.

Q：ZStack能为已经运行的虚拟机挂载ISO吗？

A：目前ZStack不支持在UI上挂载ISO，可以通过 `zstack-cli`中的 `AttachIsoToVmInstance`命令手动挂载。

75.

Q：多节点部署可以有节点使用虚拟机吗？

A：可以，但是网络需要连通，建议虚拟机的节点使用桥接。虚拟机节点运行管理节点更为合适，配置不低于4核CPU，4G内存。

76.

Q：为何我的IP Range中会少一个IP？DHCP服务的IP是多少？

A：

- 对于前一个问题，这个IP是用来分配给云路由（vRouter）或者DHCP服务了（无vRouter情况）虚拟路由的IP可以直接看到。
- 对于DHCP服务的IP，在登录后，可以查询DHCP Server占用的IP地址：

```
QuerySystemTag resourceType="L3NetworkVO"
```

这个IP是DHCP Server占用不应该释放的，是用于分配DHCP IP地址。

77.

Q：ZStack如何修改L3的IP Range呢？

A：需要在L3网络中，添加新的IP Range然后删除原来的，VM重启后会在新的IP Range里重新分配IP。

78.

Q：如何在一个VM上添加多块网卡？

A：

- 现象：

使用ZStack创建VM时，可能会遇到一个虚拟机挂载多个网络的情况，这时进入console可看到多块网卡，但没有发现IP。

- 原因：

其实每个网络对VM来说就是一块网卡，把网络加载到VM上，等价于把一块物理网卡插到一台Server上。因此使用 `ifconfig` 命令就能看到eth设备了。如果没有看到IP，说明网卡插上了，但没有配置好。

- 解决方案：

请仿照eth0的配置文件，为这个新网卡创建一个配置文件：

1. 把eth0拷贝一份。（以下示例是Centos 7.2环境，其他版本Linux路径和文件可能有变化）

```
# 将配置文件中的eth0替换为eth1,并删除其中的uuid信息
cp /etc/sysconfig/network-scripts/ifcfg-eth0 /etc/sysconfig/network-scripts/ifcfg-eth1
sed -i 's/eth0/eth1/g' ifcfg-eth0
sed -i '/UUID*/d' /etc/sysconfig/network-scripts/ifcfg-eth1
```

2. 将网络重启生效：

```
/etc/init.d/network restart
```

3. 用ifconfig查看，第二块网卡也有IP了，如果在虚拟机上挂载多个网络，重复以上操作即可。

79.

Q：如何修改BackupStorage的IP地址？

A：

1. 可以在UI界面修改sftp、镜像仓库的IP地址。
2. 也可以在`zstack-cli`设置，使用 `QueryBackupStorage` 查找待修改BackupStorage的uuid；
3. 然后使用以下命令即可：（其中BackupStorage的类型假设为sftp，hostname为目标IP）

```
QuerySftpBackupStorage uuid="3f6c52df04cb4c4d913320853e0bcc4b" hostname="172.20.12.56"
```

80.

Q：如何使用Tag将Instance offering/Disk Offering绑定主存储？

A：

1. 首先获取想要创建VM的Instance Offering和Primary Storage的uuid。

```
>>>QueryInstanceOffering name=large
{
  "inventories": [
    {
      "allocatorStrategy": "Mevoco",
      "cpuNum": 4,
      "cpuSpeed": 1,
      "createDate": "Aug 17, 2016 10:27:10 AM",
      "description": "",
      "lastOpDate": "Aug 17, 2016 10:27:10 AM",
      "memorySize": 8589934592,
      "name": "large",
      "sortKey": 0,
      "state": "Enabled",
      "type": "UserVm",
      "uuid": "a1faec292ec848be8f38e5519623cd84"
    }
  ],
  "success": true
}
```

```
>>>QueryPrimaryStorage
{
  "inventories": [
    {
      "attachedClusterUids": [
        "facc2e1f277440a7958008625b5dd515"
      ],
      "availableCapacity": 100520272896,
      "availablePhysicalCapacity": 100556562432,
      "createDate": "Aug 16, 2016 1:45:41 PM",
      "description": "",
      "lastOpDate": "Aug 17, 2016 10:08:27 AM",
      "mountPath": "/zstack_ps",
      "name": "PS-1",
      "state": "Enabled",
      "status": "Connected",
      "systemUsedCapacity": 4609003520,
      "totalCapacity": 105175318528,
      "totalPhysicalCapacity": 105175318528,
      "type": "LocalStorage",
      "url": "/zstack_ps",
      "uuid": "a8d8010f36a74ac2b9086281fbfd9084",
      "zoneUuid": "b924468082be4019bcc6e3492f428b3c"
    }
  ],
  "success": true
}
```

2. 然后为Primary Storage创建一个UserTag，为Tag命名为forWebTierVM。

```
>>>CreateUserTag tag=forWebTierVM resourceType=PrimaryStorageV0 resourceUuid=a8d8010f36a74ac2b9086281fbfd9084
{
  "inventory": {
    "createDate": "Aug 17, 2016 10:47:16 AM",
    "lastOpDate": "Aug 17, 2016 10:47:16 AM",
    "resourceType": "PrimaryStorageV0",
    "resourceUuid": "a8d8010f36a74ac2b9086281fbfd9084",
    "tag": "forWebTierVM",
    "type": "User",
    "uuid": "12f2817d5a9c49f99d6c6213f5034323"
  },
  "success": true
}
```

3. 为Instance Offering创建SystemTag，绑定刚才创建的UserTag。

```
>>>CreateSystemTag tag=primaryStorage::allocator::userTag::forWebTierVM resourceType=InstanceOfferingV0
resourceUuid=alfaec292ec848be8f38e5519623cd84
{
  "inventory": {
    "createDate": "Aug 17, 2016 10:47:55 AM",
    "inherent": false,
    "lastOpDate": "Aug 17, 2016 10:47:55 AM",
    "resourceType": "InstanceOfferingV0",
    "resourceUuid": "alfaec292ec848be8f38e5519623cd84",
    "tag": "primaryStorage::allocator::userTag::forWebTierVM",
    "type": "System",
    "uuid": "5a7cblafd1464f6aa22c2b05722deba8"
  },
  "success": true
}
```

4. 在创建VM时指定Instance Offering，会默认使用绑定的Primary Storage创建Root Volume。

5. 使用Disk Offering创建Volume也可仿照上述方法绑定Primary Storage。

81.Q：ZStack如何设置静态IP？

A：

- 现象及原因：

不少用户会习惯性到虚拟机系统里修改配置文件来设置它的IP，ZStack的DHCP Server无法侦测用户的自定义行为，因此会出现IP地址不一致的情况。

- 解决方案：

- 可以直接在UI界面，在云主机关机时设置静态IP。
- 也可以在`zstack-cli`里设置，可以通过以下命令来设置：`SetVmStaticIp ip= I3NetworkUuid= vmInstanceUuid=`

82.

Q：如果物理机宕机，如何恢复虚拟机？

A：

- 用户在使用非本地存储的主存储（例如Ceph和NFS）时，如果发现VM所在的host宕机，可通过以下方法将VM在其他host上恢复：（xxx是该VM Instance的uuid）

- ```
把需要迁移的VM的状态从unknown改成stopped
zstack-cli UpdateVmInstance uuid=xxx state=Stopped
启动VM
zstack-cli StartVmInstance uuid=xxx
```

#### 83.

Q：万兆网卡只能识别一个模块是什么原因？

A：用户服务器新添加万兆网卡，一个网卡有两个光电模块。必须保证这两个光电模块的品牌型号一致，且与网卡匹配，才能同时识别两个。

#### 84.

Q：ZStack如何更改安装目录？

A：在安装ZStack的时候，加上“-r目标目录”。例如：

```
bash zstack-installer.bin -r /home/zstack
```

#### 85.

Q：物理机维护模式和停用有什么区别？

A：

- 进入维护模式：

表示对物理机进行系统维护，可对物理机进行停机、故障修复等操作。确认物理机进入维护模式后，如果主存储类型为本地存储，ZStack会主动停止该物理机上所有正在运行的云主机；如果存储类型为共享存储，云主机会自动迁移至其他物理机上。

- 停用：

停止使用某个物理机，后续云主机不能在此物理机上创建，且已关闭的云主机不能再启动，目前允许中的云主机不受影响。

86.

Q：物理机连不上有哪些原因？

A：

- 物理机本身操作系统与管理节点不一致。

解决方法：所有节点统一使用CentOS 7.2。

- 物理机环境不干净，导致ZStack管理节点无法自动安装KVM。

解决方法：重装物理机操作系统，且不要自己安装软件。

- 确保连接信息正确，包括IP、用户名、密码。

87.

Q：如何将KVM上的云主机迁移到ZStack中？

A：解决方案：

1. 通过libvirt导出云主机的镜像（.qcow2文件）。
2. 将镜像上传至一个http服务器中。
3. 使用ZStack镜像服务器导入镜像。
4. 通过镜像创建云主机。



注：

如果原来的云主机挂载了数据盘，如何将数据盘也迁移过来呢？

- 方法一：与上述操作类似，需要先将数据盘生成镜像，同理导入，ZStack使用该镜像创建Data Volume，再将Data Volume挂载到云主机即可。
- 方法二：先在ZStack中创建一个相同大小的Data Volume，找到对应的路径，将原云盘数据直接复制到新的路径下，最后再挂载到云主机上。

88.

Q：管理节点更改了主机名后，如何重连rabbitmq？

A：解决方案：

1. 在 /etc/hosts 中添加新hostname解析。
2. 执行 `zstack-ctl reset_rabbitmq`
3. 上述命令也可通过手动修改来实现：
  - 调整 `/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties`，其中 `CloudBus.serverIp.0 = XX.XX.XX.XX`，填写管理网络IP。
  - 创建用户：

```
rabbitmqctl add_user zstack zstack.password
rabbitmqctl set_user_tags zstack administrator
rabbitmqctl set_permissions -p / zstack ".*" ".*" ".*"
```



注：

hostname中如果带.，例如：hostname命名为"www.test.com"，那么rabbitmq会解析此hostname为"www."，导致rabbitmq-server一直无法启动，需要将其hostname修改为"www-test-com"，再重置rabbitmq-server即可。

89.

Q：ZStack如何设置云主机启动顺序？

A：

- 目前ZStack可以在UI界面设置启动顺序。
- 也可以在cli中切换启动顺序，具体操作如下：
- 1. 用户在云主机创建之后挂载ISO，例如：

```
AttachIsoToVmInstance isoUuid=dff9e4db6ba84633be729f92530dbc21\
vmInstanceUuid=7b06433bc5834b7b8001e0ca90920250
```

- 2. 云主机默认的启动顺序是：第一从硬盘引导，第二从ISO引导。我们需要将启动顺序修改，例如：

```
SetVmBootOrder bootOrder=CdRom uuid=7b06433bc5834b7b8001e0ca90920250
```

- 3. 再启动云主机，就会优先从ISO启动。

90.

Q：ssh连接超时如何解决？

A：

- 现象：

在ZStack中，很多服务需要ssh连接，包括添加物理机、存储、搭建管理节点HA等。由于用户网络等问题，这些服务的操作可能因为ssh连接超时而提示time out的错误。

- 原因：

要是由于对IP地址进行DNS校验而产生的延时。

- 解决方案：

主以下以搭建管理节点HA为例解决该问题：

- 在需要ssh连接的第二个管理节点上，打开/etc/ssh/sshd\_config
- 将 UseDNS yes这一行使用#注释掉，这样在连接ssh时，就直接使用IP地址了。

91.

Q：如何在创建VM时指定静态IP？

A：

- 在大多数情况下，VM的IP由DHCP自动分配。对于需要静态IP的用户，可以使用SetVmStaticIp设置。
- 如果要求在创建VM时指定静态IP，目前ZStack企业版UI界面支持该功能，也可在ZStack命令行里使用SystemTag实现，具体操作如下：

1. 开启DHCP服务（VR或flat都可以）
2. 执行下面命令：

```
CreateVmInstance name=ceph-3
instanceOfferingUuid=7ec8f1148f14452aa359607112ce7bbe
l3NetworkUuids=f749ec8b1ea94676bda7a5968de6b947
imageUuid=b6ebcfa021e24bc3b08a27216fd589a7
systemTags=staticip::f749ec8b1ea94676bda7a5968de6b947::172.20.58.3
```

92.

Q：在镜像服务器中添加镜像超时怎么办？

A：

- 原因：

有时在下载镜像时，因为镜像过大或者网络等原因导致镜像下载太慢而超时报错。

- 解决方案：

在ZStack中，如果执行了下载操作，那么会在BS（镜像服务器）上启用http服务。当UI上显示超时，甚至关闭管理节点时（BS不在管理节点所在的host），并不影响服务继续运行。

因此，如果时间过长不用担心。后台自动下载完成，下次继续添加该镜像时会迅速完成。

93.

Q：当管理节点出现unknown状态而无法用命令停止怎么办？

A：

- 当多人在虚拟机上运行ZStack时，建议虚拟机配置CPU至少2核，内存大于6G。
- 管理节点使用 `zstack-ctl status` 查看，出现unknown状态，而无法使用 `zstack-ctl stop` 停止，请用 `dmesg` 命令查看，可能会发现如下提示：

```
[66363.604813] INFO: task java:2907 blocked for more than 120 seconds.
```

这种情况一般是由于分配资源不足造成。调整管理节点所在虚拟机配置后，重启管理节点即可。

94.

Q：ZStack启动后提示MySQL is not running如何解决？

A：

- 现象：

使用 `zstack-ctl start` 启动ZStack时提示错误，`zstack-ctl status` 查看状态时显示 MySQL is not running

- 解决方案：

可以先尝试以下命令搜索：

```
grep -rn /usr/local/zstack/apache-tomcat/logs/management-server.log "Cannot execute statement: \
impossible to write to binary log since BINLOG_FORMAT = STATEMENT and at least one table uses a storage \
engine limited to row-based logging"
```

95.

Q：为什么有的VM关机时间比较久？

A：

- 原因：

这种情况一般出现在Guest OS为Linux的VM上，这是因为缺少ACPID服务（用于处理电源相关事件的守护进程）。KVM虚拟机安装Linux系统默认是没有安装ACPID服务的，所以并不会做处理，而在一分钟后被ZStack杀掉VM进程而强行关闭。

- 解决方案：

对于这种情况，用户只需在虚拟机里安装和启动ACPID服务即可，执行命令如下：

```
[root@kvm_client_00 ~]# yum install acpid -y
重新启动ACPID服务，安装后默认加入到开机启动。
```



```
[root@kvm_client_00 ~]# /etc/init.d/acpid restart
```

96.

Q：如何更改ZStack同时登录次数？

A：

1. 首先QueryGlobalConfig name~=session，可以看到关于最大会话数的限制为500。

```
{
 "category": "identity",
 "defaultValue": "500",
 "description": "Max number of sessions management server accepts. \
When this limit met, new session will be rejected",
 "name": "session.maxConcurrent",
 "value": "500"
}
```

2. 假如希望改成1000，需要修改全局配置，可在UI界面修改全局设置中设置最大会话数，也可以在cli中执行以下命令：

```
UpdateGlobalConfig category=identity name="session.maxConcurrent" value=1000。
```

97.

Q：出现cannot acquire state change lock的错误怎么办？

A：

- 原因：

出现这个错误属于小概率事件，是由libvirt引发的。出现这个问题后，该VM上的后续操作只要调用到libvirtd就会失败。

- 解决办法：

```
killall -9 libvirtd
rm /var/run/libvirtd.pid
```

完成后重连物理机。

98.

Q：本地存储如何迁移Volume？

A：

- 对于一些使用Local Storage的用户，有迁移云主机或者数据云盘的需求，可以使用以下方法进行迁移：
- 1. 关闭待迁移云主机，如果是Data Volume需要先卸载。
- 2. 执行LocalStorageMigrateVolume destHostUuid= volumeUuid=
- 3. 在目标物理机上启动云主机/挂载Data Volume即可。

99.

Q：CPU未开启虚拟化，添加host报错怎么办？

A：

1. 用户安装ZStack后，添加的物理机报错。提示：

```
ERROR:could not insert 'kvm_intel':Operation not supportednnnstndout
```

2. 查看开启虚拟化情况，如果有返回值，则表示开启成功。

```
cat /proc/cpuinfo | grep "vmx|svm"
```

3. 开启后再次添加该物理机失败，会出现libvirt相关错误。使用 `systemctl status libvirtd.service` 查看，提示：

```
Failed to start Virtualization daemon
```

4. 这时需要将物理机的`/etc/libvirt/libvirtd.conf`修改成以下内容：

```
listen_tls = 0
listen_tcp = 1
tcp_port = "16509"
auth_tcp = "none"
lock_manager = "lockd"
```

## 100.

Q：ZStack的无状态是什么？

A：在无状态的服务情况下，请求者不再需要询问何处发送请求；当新的服务实例加入，或者旧的服务实例脱离的时候，服务也不再需要交换状态。



注：

例如，假如某个操作删除了物理机的br\_eth0（即：物理机连接ZStack L2网络的网卡），由于ZStack里所有的物理机均无状态，只要重连物理机即可。其实物理机硬盘上没有固定的配置文件，它在内存中。重连物理机后会重新生成配置文件。

## 101.

Q：ZStack的存储如何计算？

A：ZStack的存储分为主存储和备份存储。（以下讨论均不涉及ZStack企业版的超分情况）

- 备份存储的计算相对简单，其总容量和真实容量对应实际的物理值（即与`df -h`命令结果相同）
- 主存储分为物理总容量、物理可用容量和总容量、可用容量。其中物理总容量、物理可用容量与备份存储的技术方法相同，为物理真实容量。这里需要理解的主存储的总容量（Total Capacity）和可用容量（Available Capacity）。总容量与实际总容量相同，可用容量为总容量减去云盘的虚拟容量（Virtual Size）。



注：

- 例如，主存储为500G，使用一个大小为100G的镜像创建云主机，实际大小为3G。在主存储上看到的可用容量为500G-100G=400G，而它的物理真实容量为500G-3G=497G。
- ZStack主存储URL指定到某个分区的目录上，主存储的总容量为目录所属分区的总容量。
- ZStack只关心自己使用的容量（镜像、云盘、快照），其他用户在这个分区使用的空间不计入已使用的容量中。
- 因此，可用容量=所属分区的总容量 - 自己使用的容量。其中，主存储上的镜像和快照为真实容量，云盘为虚拟容量。
- 每次容量校准需要重连物理机。

## 102.

Q：修改Windows云主机密码报错怎么办？

A：

- 现象：

如果在修改Windows云主机密码时出现如下报错：

```
return code: 1 stdout: stderr: error: internal error: unable to execute QEMU agent command
'guest-set-user-password':
The command guest-set-user-password has not been found
```

- 原因：

可能原因是使用了错误的Windows Virtio驱动。

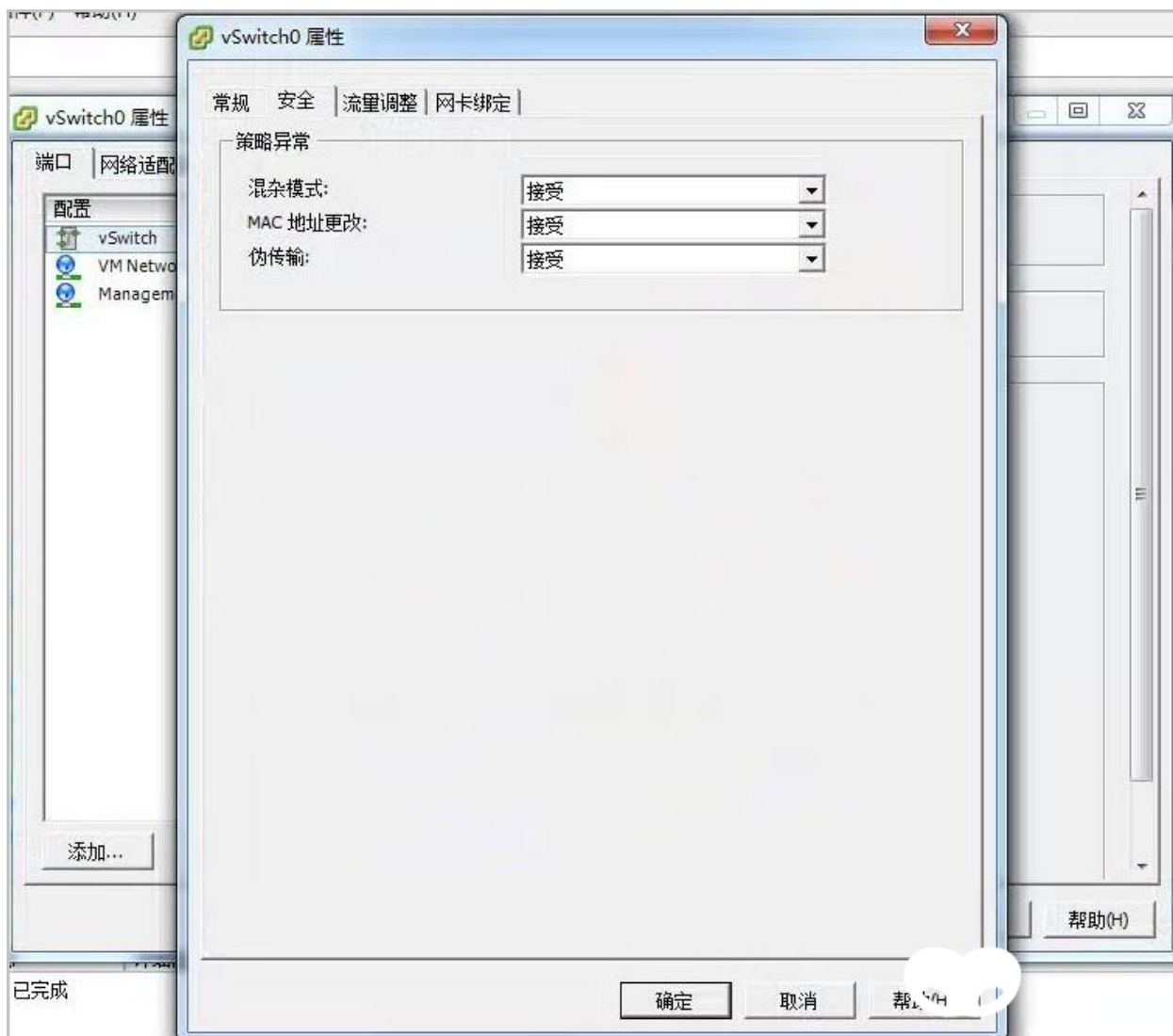
解决方案：

目前ZStack定制ISO里已经集成了可用的Windows Virtio驱动，解压此ISO，添加此Windows Virtio驱动作为ISO镜像到镜像服务器中，挂载此ISO，安装对应的qemu-ga即可。

### 103.

Q：VM无法ping通网关？

A：当使用vCenter创建的VM作为ZStack的管理节点和计算节点时，需要把vSwitch的属性安全 > 混杂模式修改为接受。



### 104.

Q：误删除ZStack ssh key后如何恢复？

A :

- ZStack的ssh key默认存放在目录：`/usr/local/zstack/apache-tomcat-7.0.35/webapps/zstack/WEB-INF/classes/ansible/rsaKeys`
- 如果用户不小心删除了ZStack ssh key，可以通过以下方式恢复：

1. 首先运行以下命令生成ssh key：

```
ssh-keygen -f /usr/local/zstack/apache-tomcat-7.0.35/webapps/zstack/WEB-INF/classes/ansible/rsaKeys/id_rsa -N "
```

2. 生成ssh key后，运行以下命令确保ZStack用户有权限访问：

```
chown -R zstack.zstack /usr/local/zstack/apache-tomcat-7.0.35/webapps/zstack/WEB-INF/classes/ansible/rsaKeys
```

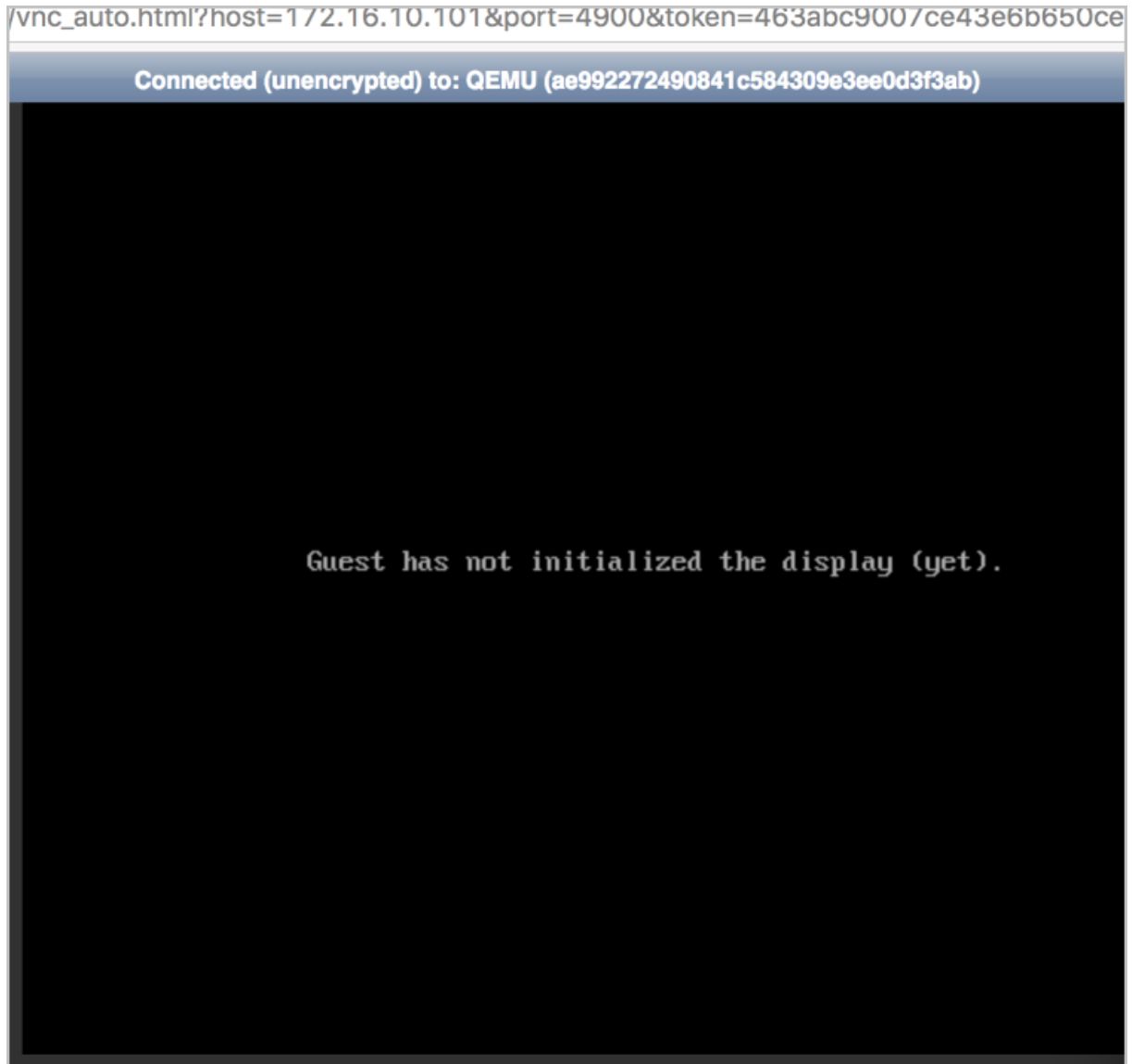
#### 105.

Q：因CPU型号过早而无法创建云主机怎么办？

A：

- 现象及原因：

使用利旧的服务器作为计算节点，可能会遇到因CPU型号过早，无法创建云主机或者创建的云主机不能显示console界面的问题，如下图所示：



- 解决方案：

出现这种情况，可以尝试如下设置：

1. kvm-nested.conf里面第一行加上ept=0，如：

```
options kvm_intel nested=1 ept=0
#vim /etc/modprobe.d/kvm-nested.conf
```

2. 关闭运行状态的VM，然后移除kvm\_intel：

```
#rmmod kvm_intel
```

3. 最后重新加载kvm\_intel：

```
#modprobe kvm-intel ept=0 unrestricted_guest=0
```

#### 106.

Q：CPU型号与Passthrough(支持嵌套虚拟化)的说明。

A：

- 可以通过改变全局配置让虚拟机获取同物理机CPU相同的型号：

```
UpdateGlobalConfig category=kvm name=vm.cpuMode value=host-model
```

- 支持三种模式：
  - **none**：CPU型号为QEMU模拟器
  - **host-model**：CPU类型为物理机CPU类型
  - **host-passthrough**：CPU型号跟物理机CPU完全一样



注：

- 当模式为**host-model**和**host-passthrough**时，虚拟机可以获得嵌套虚拟化功能。
- 使用该功能可能会影响虚拟机热迁移。因为热迁移时会检查虚拟机CPU型号，如果两台物理机的CPU型号不同，会导致虚拟机热迁移失败。

## 107.

Q：报错restartDnsmasqAfterNumberOfSIGUSER1怎么办？

A：

- 现象：

在创建VM云路由网络时出现如下错误消息：

```
"description": "A message or a operation timeout",
 "details": "[Async Http Timeout] url: http://10.10.10.253:7272/init, timeout after 300000
[MILLISECONDS], command: {\"uuid\":\"78183b0a46094e30bc8a6128b30ee8cb\", \"restart
DnsmasqAfterNumberOfSIGUSER1\":0}"
}
```

- 原因：

同时满足以下情况就会出现该问题：

1. 管理节点上有多块网卡；
2. 第一块网卡置了默认路由；
3. 第一块网卡设置的IP为内网IP，无法访问出去。

- 解决方案：

修改 `/usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties`

增加或修改参数`management.server.ip=xxx.xxx.xxx.xxx`这个IP地址，修改为可以访问出去的IP地址。

## 108.

Q：升级时，遇到ZStack ISO版本不匹配怎么办？

A：

- 现象：

在升级时遇到ZStack ISO版本不匹配，会提示以下错误信息：

```
Reason: The current local repo is not suitable for ZStack installation.
Please download from http://www.zstack.io/product_downloads/ and run
#wget http://cdn.zstack.io/product_downloads/iso/
#sh zstack-upgrade
For more information, see http://www.zstack.io/support/tutorials/upgrade/
```

- 原因：

当前系统使用的ISO版本不适合升级此安装包，需要升级对应的ISO才可继续升级ZStack。

- 解决方案：

请执行以下部署：

1. 在管理节点下载最新的ISO，例如在管理节点执行下载ISO：

```
wget http://cdn.zstack.io/product_downloads/iso/
```

2. 执行升级，升级ISO里面的yum repo同步到本地，并且直接升级管理节点。

```
sh zstack-upgrade
```

109.

Q：Windows安装Virtio驱动时，磁盘应该选择哪个驱动？

A：安装viosstor存储控制器和vioscsi磁盘控制器。

110.

Q：ZStack的DHCP服务由谁提供？

A：扁平网络：各物理机的dnsmasq；云路由网络：虚拟路由器的dhcp。

111.

Q：ZStack备份需要备份哪些？

A：ZStack本身的备份只需备份数据库即可。

112.

Q：ZStack如何备份数据库？

A：

- 执行zstack-ctl dump\_mysql 就会执行一次备份；
- 执行crontab -l可以看到，每天凌晨以及中午12：30均会备份一次数据库，保留最新的14个备份，可根据需求进行调整，为安全考虑建议进行异地远程备份。

```
30 0,12 * * * zstack-ctl dump_mysql --keep-amount 14
```

113.

Q：云主机被变成Paused状态？

A：云主机内存过低时，在创建后或者迁移后可能会变成Paused状态。

114.

Q：ZStack能和LDAP连接吗？只能用API接口控制用户吗？社区版都能做什么？

A：可以，设置页面有LDAP设置；企业版提供了RESTful API的调用文档，不限制语言。社区版也可适用于生产环境，只是UI开放功能与企业版相比相对简单，建议使用API或cli操作，企业版基于社区版定制，提供了全功能UI界面以及社区版不具备的一系列增值功能。

115.

Q：升级之后为什么MySQL的连接数会增加？

A：不会增加，但是遇到三个节点，只能启动两个，第三个报Too many connections错误，可以通过修改MySQL配置，确认host的libvirtd是否正常。如果突发异常，可能会出现此错误。

116.

Q：ZStack社区版和企业版的区别是什么？

A：后者提供企业级功能和完整技术支持。

117.

Q：添加新物理机时出现这个报错如何处理？

A：

- 现象：

添加新物理机时出现如下报错：

```
{
 "org.zstack.header.host.APIAddHostEvent": {
 "success": false,
 "apild": "75b0774fc7b94007ac1ba274e6ed5eb3",
 "headers": {
 "schema": {}
 },
 "error": {
 "code": "HOST.1000",
 "cause": {
 "code": "HOST.1003",
 "cause": {
 "code": "SYS.1006",
 "description": "An operation failed",
 "details": "failed to check physical network interfaces[names : enp2s0] on kvm host[uuid:fab977f38e024e6e86f048a4090845f8, ip:192.168.10.11]",
 "description": "An error happened when connecting to host",
 "details": "connection error for KVM host[uuid:fab977f38e024e6e86f048a4090845f8, ip:192.168.10.11]",
 "description": "Unable to add host",
 "details": "connection error for KVM host[uuid:fab977f38e024e6e86f048a4090845f8, ip:192.168.10.11]",
 "createTime": 1489036864392,
 "type": {
 "_name": "key.event.API.API_EVENT",
 "id": "8f1d551677c44788953eae7255eda54b"
 }
 }
 }
 }
 }
}
```

- 原因及解决方案：

没有enp2s0网卡，一个集群里的hosts的网卡和存储应该一致。

#### 118.

Q：基于“ZStack + Ceph”部署，创建vm，zstack-mn，qemu-kvm-ev-mn，离线安装时出现这个报错如何处理？

A：

- 现象：

报错如下：

```
"org.zstack.header.vm.APICreateVmInstanceEvent": {
 "success": false,
 "apild": "9ebc212f00ea45bbaa91e1e420230921",
 "headers": {
 "schema": {}
 },
 "error": {
 "code": "SYS.1006",
 "cause": {
 "code": "HOST.1005",
 "description": "Failed to start vm on hypervisor",
 "details": "failed to start vm[uuid:2da983e8b5a34a459bb71a0d705b0bba name:vm-qge5] on kvm host[uuid:b7a14605c6b042e0b2da17ae329ea70a, ip:10.79.186.30], because Traceback (most recent call last):\n File \"/var/lib/zstack/virtualenv/kvm/lib/python2.7/site-packages/kvmagent/plugins/vm_plugin.py", line 2721, in start_vm\n self._start_vm(cmd)\n File \"/var/lib/zstack/virtualenv/kvm/lib/python2.7/site-packages/kvmagent/plugins/vm_plugin.py", line 2664, in _start_vm\n 'unable to start vm [uuid:%s, name:%s], libvirt error: %s' % (cmd.vmlInstanceUuid, cmd.vmName, str(e))\nKvmError: unable to start vm[uuid:2da983e8b5a34a459bb71a0d705b0bba, name:vm-qge5], libvirt error: internal error: early end of file from monitor, possible problem: 2017-03-12T10:49:18.661399Z qemu-kvm: -drive file=rbd:pri-v-r-a5486c5605a8499dbe3ca79ee3ca87c4/a28d86db519c48ef9e4c50a3f8e27305:id=zstack:key=AQD4HsVYwfgaMBAAxBNUtDw1wzhitURDrUMM9g==:auth_supported=cephx\\;none:mon_host=10.79.186.112\\:6789,format=raw,if=none,id=drive-virtio-disk0: error reading header from a28d86db519c48ef9e4c50a3f8e27305: Function not implemented\n",
 "description": "An operation failed",
 "details": "failed to start vm[uuid:2da983e8b5a34a459bb71a0d705b0bba name:vm-qge5] on kvm host[uuid:b7a14605c6b042e0b2da17ae329ea70a, ip:10.79.186.30], because Traceback (most recent call last):\n File \"/var/lib/zstack/virtualenv/kvm/lib/python2.7/site-packages/kvmagent/plugins/vm_plugin.py", line 2721, in start_vm\n self._start_vm(cmd)\n File \"/var/lib/zstack/virtualenv/kvm/lib/python2.7/site-packages/kvmagent/plugins/vm_plugin.py", line 2664, in _start_vm\n 'unable to start vm [uuid:%s, name:%s], libvirt error: %s' % (cmd.vmlInstanceUuid, cmd.vmName, str(e))\nKvmError: unable to start vm[uuid:2da983e8b5a34a459bb71a0d705b0bba, name:vm-qge5], libvirt error: internal error: early end of file from monitor, possible problem: 2017-03-12T10:49:18.661399Z qemu-kvm: -drive file=rbd:pri-v-r-a5486c5605a8499dbe3ca79ee3ca87c4/a28d86db519c48ef9e4c50a3f8e27305:id=zstack:key=AQD4HsVYwfgaMBAAxBNUtDw1wzhitURDrUMM9g==:auth_supported=cephx\\;none:mon_host=10.79.186.112\\:6789,format=raw,if=none,id=drive-virtio-disk0: error reading header from a28d86db519c48ef9e4c50a3f8e27305: Function not implemented\n",
 "createTime": 1489315754576,
 "type": {
 "_name": "key.event.API.API_EVENT",
 "id": "bb86a69de1c04043a9ae218ee510dae0"
 }
 }
 }
}
```

- 原因及解决方案：

rbid无法读取镜像文件头文件，很可能是Ceph存储本身的问题，可google：libvirt error: internal error: early end of file from monitor

#### 119.



Q：出现这个IO报错怎么处理？

A：

- 现象：

报错如下：

```
"description": "Cannot do an IO operation(e.g. network IO error, file IO error)", "details": "I/O error on POST request for \"http://127.0.0.1:7758/console/establish\": Connect to 127.0.0.1:7758 [/127.0.0.1] failed: Connection refused; nested exception is org.apache.http.conn.HttpHostConnectException: Connect to 127.0.0.1:7758 [/127.0.0.1] failed: Connection refused", "$hashKey": "object:1344"
```

- 解决方案：

需要删除掉zstack/pickledb/下面的consoleProxy文件。

120.

Q：是否只有扁平化网络才能实现物理机与云主机互通？

A：基本上是这样。扁平网络，意味着对于一个企业内部全部用户的机器都在一个二层网络上，分配的172.16.x.y的IP地址，相互都可以联通。VM在扁平网络上，因此对于任何用户而言，就像是office其他机器一样。层级之间互通需要NAT。

121.

Q：为什么同一个网络段之间的VM相互之间ping不通？

A：如果本物理机的VM IP互通，而且VM所在网卡的物理机之间也互通，那么就应该是跨物理机的VM之间不通，很可能交换机出现问题。

122.

Q：添加主机时出现这个报错如何处理？

A：

- 现象：

报错如下：

```
unhandled exception happened when calling private coin org.zstack.core.ansible.Ansible
FacadeImp.1.run(org.zstack.header.core.Completion)
```

- 原因及解决方案：

kvm\_intel硬件虚拟化特性没有打开；请在BIOS上打开硬件虚拟化特性。

123.

Q：出现No available host found报错如何处理？

A：

- 现象：

报错如下：

```
No available host found details:unable to allocate hosts; due to pagination is enabled, there
might be several allocation failures happened before; the error list is {no host found in zones
that attaches to backup storage where image is on}
```

- 原因：

镜像服务器没有挂在到Zone上。

124.

Q：使用ZStack企业版时，镜像服务器始终添加不了主机？

A：尝试换一个路径，例如/zstack\_ps，不能挂载到整个LVM下面，要有一个确定的目录。

125.

Q：安装ZStack后出现cannot connect to all RabbitMQ servers报错如何处理？

A：

- 现象：

报错如下：

```
cannot connect to all RabbitMQ servers[ip:['CloudBus.serverIp.0,' local host']],port:5672]
defined in /usr/local/zstack/apache-tomcat/webapps/zstack/WEB-INF/classes/zstack.properties,
please reset rabbitmq by:" zstack-ctl reset_rabbitmq"
```

- 解决方案：

先把**CloudBus.serverIp.0**修改为用户实际使用的IP地址，修改hostname，执行**zstack-ctl reset\_rabbitmq**。

126.

Q：ZStack如何修改虚拟机网卡的mac地址？

A：mac地址是dnsmasq随机分配的，要修改需要将当前网络卸载再重新加载。

127.

Q：数据云盘是建立在主存储还是备份存储上？为什么我建立的数据云盘在Windows上总是发现不了？

A：数据云盘是建立在主存储上，需要在云主机里安装Virtio驱动。

128.

Q：1.10.0.192版本升级时出现报错The Operating System version is not suitable for zstack-enterprise installation.怎么办？

A：请使用最新的ZStack ISO安装升级。

129.

Q：ZStack对Loadbalance是如何规划的？目前支持像Netscaler这样的商业方案吗？

A：ZStack使用Vyos 提供负载均衡服务，主要是访问公网IP相关的服务，实际由后台一系列云主机提供服务。Netscaler更贴近于Web服务应用交付相关。

130.

Q：出现报错ip is not available in the l3network如何处理？

A：

- 在zstack-cli里面执行类似这样的命令**CheckIpAvailability ip="10.58.20.99" l3NetworkUuid="deca2a8f2c9147878751272a1660bc9b"**进行检查。
- 也可使用**GetFreeIp l3NetworkUuid="deca2a8f2c9147878751272a1660bc9b"**查询可用IP。

131.

Q：ZStack创建二层网络时，可以为一个物理网卡划分多个Vlan吗？用该物理网卡接口去连接交换机的trunk接口可以吗？

A：可以，1- 4094这些端口可用。

132.

Q：ZStack是否支持VRRP协议？希望虚拟机加Keepalived服务，两台虚拟机"飘"IP，提升应用的高可用性。

A：可以尝试一下在两个云主机内部安装Keepalived服务。

133.

Q：Vlan设置完成后不能通信，开了iptables，是不是要加什么配置才能让Vlan互相通信？

A：Vlan互相通信要到交换机上去配置。

134.

Q：如果网内没有自建DNS服务器，云主机怎么才能使用电信的DNS服务器IP地址来解析域名上网？

A：请在创建网络时指定DNS地址。

135.

Q：部署Shared Mount Point主存储对文件系统有什么要求？

A：Shared Mount Point主存储要求你的计算节点都需要挂载同一个nfs服务，并且挂载到本地计算节点相同的目录。

136.

Q：ZStack基于模版创建虚拟机，可否实现母盘+增量盘的模式？

A：ZStack是采用增量盘的模式。镜像从镜像服务器下发到主存储作为镜像缓存，以此为base创建VM。这些VM均使用同一base。

137.

Q：在Ceph环境下创建快照1、2、3，然后删除1，但2和3没有删除，这是bug吗？还是和后端的存储有关？创建模板时会自动生成一个快照，创建VM时是把模板快照恢复到VM上吗？

A：Ceph类型是无树状，快照点之间没有依赖关系，恢复快照会丢失当前状态，使用了Ceph的Snap Clone方法。

138.

Q：在ZStack部署vCenter环境报错"No dvSwitch found"怎么办？

A：

- 现象：

报错如下：

```
"org.zstack.vmware.APIAddVCenterMsg": {
 "domainName": "192.168.0.16",
 "name": "VC",
 "username": "administrator@vsphere.local",
 "password": "*****",
 "https": "true",
 "zoneUuid": "631998da879f4b63b484876138c87a6a",
 "session": {
 "uuid": "5cec7c53f71344858a30186dd0d8280a",
 "callid": "api-mMa45Vw7"
 }
}
}
{
 "success": false,
 "apild": "476e659189ef4097b98dcf0ed4adc51c",
 "headers": {
 "api-id": "476e659189ef4097b98dcf0ed4adc51c",
 "schema": {}
 },
 "session": {
 "callid": "api-mMa45Vw7",
 "uuid": "5cec7c53f71344858a30186dd0d8280a"
 },
 "error": {
 "code": "SYS.1006",
 "description": "An operation failed",
 "details": "No dvSwitch found",
 "$$hashKey": "object:672"
 },
}
```

```

"createdTime": 1496300098050,
"type": {
 "_name": "key.event.API.API_EVENT"
},
"id": "872c5d4ab75e4fd88a69bec0f90b41fe"
}

```

- 原因及解决方案：

vCenter环境需要先创建分布式交换机：VC CLUSTER HOST dvswitch。如果需要ZStack来接管vCenter，vCenter必须先创建dvswitch；ZStack创建云路由网络也基于dvswitch。

### 139.

Q：Ceph计算节点出现could not get value of the secret for username zstack错误如何处理？

A：

- 现象：

当计算节点使用Ceph存储时，在创建、启动、迁移云主机时，出现以下错误：

```

内部错误：could not get value of the secret for username 'zstack' using uuid 'ee5b6375-4564-446b-ac53-c94ea4e13904'

```

- 原因：

一般是计算节点连接Ceph时，丢失密钥，需要删除密钥，重连物理机。

- 解决方案：

需要进行以下步骤进行修复：

1. 在出现问题的计算节点执行virsh secret-list以获取其uuid。

例如：获取uuid为03b5ee0d-df21-46d2-9927-7c930333cb70

```

root@ceph-host ~]# virsh secret-list
UUID Usage

03b5ee0d-df21-46d2-9927-7c930333cb70 ceph 03b5ee0d-df21-46d2-9927-7c930333cb70

```

2. 执行virsh secret-undefine 03b5ee0d-df21-46d2-9927-7c930333cb70以删除其密钥。

```

[root@ceph-host ~]# virsh secret-undefine 03b5ee0d-df21-46d2-9927-7c930333cb70
Secret 03b5ee0d-df21-46d2-9927-7c930333cb70 deleted

```

3. 执行/etc/init.d/zstack-kvmagent stop停止管理节点agent。

```

root@ceph-host ~]# /etc/init.d/zstack-kvmagent stop
2017-05-08 10:07:57,590 DEBUG [zstacklib.utils.shell] /sbin/iptables-save
2017-05-08 10:07:57,596 DEBUG [zstacklib.utils.iptables] removed empty chains:[]
2017-05-08 10:07:57,598 DEBUG [zstacklib.utils.shell] /sbin/iptables-restore < /tmp/
tmpBW4mWX
Stop Daemon...
Stop Daemon Successfully
stop zstack kvm agent SUCCESS

```

4. 在UI界面再次重连此计算节点，重连成功后，创建、启动、迁移云主机即可正常。

目前ZStack在UI已提供关闭CephX密钥认证选项。

## 16 授权

### 关于界面

如图 79: 关于界面所示，显示了当前ZStack for Alibaba Cloud软件的协议、版本、授权状态和请求码等等信息。

图 79: 关于界面



### 使用条款与软件政策

使用条款与软件政策为ZStack for Alibaba Cloud软件的免责条款，请用户务必仔细阅读，明确相关法律责任，同意本条款相关约定，合理合法的使用本软件。

### 使用协议

ZStack for Alibaba Cloud的具体协议授权范围参见下表中的混合云版：

| 授权协议类型   | 授权协议说明                          |
|----------|---------------------------------|
| 社区版      | 无限期使用社区版所有功能                    |
| 无限期免费试用版 | 可以免费添加1台物理主机，并无限期使用全部功能         |
| 试用版      | 可以添加许可证指定的台数的物理机，并在有效时间内使用全部功能  |
| 付费正式版    | 可以添加已购买数量的物理机，并且在许可证有效时间内使用全部功能 |

| 授权协议类型 | 授权协议说明                                           |
|--------|--------------------------------------------------|
| 混合云版   | 可以添加已购买数量CPU的物理机，并且在许可证有效时间内使用ZStack企业版和混合云的全部功能 |

- 欢迎您直接与**ApsaraStack\_Support@service.aliyun.com**咨询购买ZStack for Alibaba Cloud的相关事宜。
- 如您需要升级版本或更新许可证，请点击请求码后面的

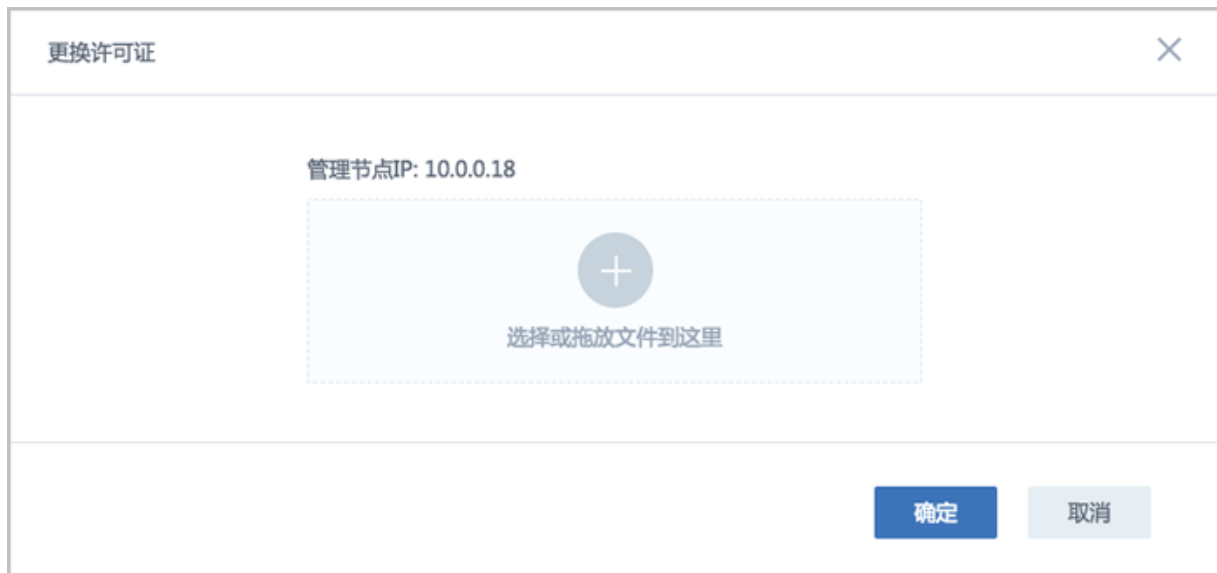


将您的请求码和升级需求（希望升级的版本号）发送电子邮件至**ApsaraStack\_Support@service.aliyun.com**，我们将尽快与您联系。

#### 加载许可

如果您获得了ZStack的新的许可证，可点击更换许可证，弹出更换许可证界面，直接将获得的新的许可证本地上传即可，如图 80: 许可证支持本地上传所示：

图 80: 许可证支持本地上传



可点击协议标题后面的



来重新加载许可证。

# 术语表

---

## 区域 ( Zone )

ZStack中最大的一个资源定义，包括集群、二层网络、主存储等资源。

## 集群 ( Cluster )

一个集群是类似物理主机 ( Host ) 组成的逻辑组。在同一个集群中的物理主机必须安装相同的操作系统 ( 虚拟机管理程序, Hypervisor )，拥有相同的二层网络连接，可以访问相同的主存储。在实际的数据中心，一个集群通常对应一个机架 ( Rack )。

## 管理节点 ( Management Node )

安装系统的物理主机，提供UI管理、云平台部署功能。

## 计算节点 ( Compute Node )

也称之为物理主机 ( 或物理机 )，为云主机实例提供计算、网络、存储等资源的物理主机。

## 主存储 ( Primary Storage )

用于存储云主机磁盘文件的存储服务器。支持本地存储、NFS、Ceph、FusionStor、Shared Mount Point等类型。

## 镜像服务器 ( Backup Storage )

也称之为备份存储服务器，主要用于保存云主机的镜像模板文件。建议单独部署镜像服务器。

## 镜像仓库 ( Image Store )

镜像服务器的一种类型，可以为正在运行的云主机快速创建镜像，高效管理云主机镜像的版本变迁以及发布，实现快速上传、下载镜像，镜像快照，以及导出镜像的操作。

## 云主机 ( VM Instance )

运行在物理机上的虚拟机实例，具有独立的IP地址，可以访问公共网络，运行应用服务。

## 镜像 ( Image )

云主机使用的镜像模板文件，包含云主机的操作系统，也可以定制安装相应的软件。

## 云盘 ( Volume )

云主机的数据盘，给云主机提供额外的存储空间，共享云盘可挂载到一个或多个云主机共同使用。

## 计算规格 ( Instance Offering )

启动云主机涉及到的CPU数量、内存、网络设置等规格定义。

## 云盘规格 ( Disk Offering )

创建云盘容量大小的规格定义。

## 二层网络 ( L2 Network )

二层网络对应于一个二层广播域，进行二层相关的隔离。一般用物理网络的设备名称标识。

## 三层网络 ( L3 Network )

云主机使用的网络配置，包括IP地址范围、网关、DNS等。

## 公有网络 ( Public Network )

由因特网信息中心分配的公有IP地址或者可以连接到外部互联网的IP地址。

## 私有网络 ( Private Network )

云主机连接和使用的内部网络。

## L2NoVlanNetwork

物理主机的网络连接不采用Vlan设置。

## L2VlanNetwork

物理主机节点的网络连接采用Vlan设置，Vlan需要在交换机端提前进行设置。

## VXLAN网络池 ( VXLAN Network Pool )

VXLAN网络中的 Underlay 网络，一个 VXLAN 网络池可以创建多个 VXLAN Overlay 网络（即 VXLAN 网络），这些 Overlay 网络运行在同一组 Underlay 网络设施上。

## VXLAN网络 ( VXLAN )

使用 VXLAN 协议封装的二层网络，单个 VXLAN 网络需从属于一个大的 VXLAN 网络池，不同 VXLAN 网络间相互二层隔离。



## 云路由 ( vRouter )

云路由通过定制的Linux云主机来实现的多种网络服务。

## 安全组 ( Security Group )

针对云主机进行第三层网络的防火墙控制，对IP地址、网络包类型或网络包流向等可以设置不同的安全规则。

## 弹性IP ( EIP )

公有网络接入到私有网络的IP地址。

## 快照 ( Snapshot )

某一个时间点上某一个磁盘的数据备份。包括自动快照和手动快照两种类型。