

ZStack 私有云实践指南

版本 1.10.0

日期 2017-03-10

版权所有©上海云轴信息科技有限公司 2017。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标说明



和其他云轴商标均为上海云轴信息科技有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受上海云轴公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，上海云轴公司对本文档内容不做任何明示或暗示的声明或保证。

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

上海云轴信息科技有限公司

地址：上海市闵行区紫竹科学园东川路 555 号 6 号楼

邮编：200241

网址：<http://www.zstack.io/>

客户服务邮箱：support@zstack.io

客户服务电话：400-962-2212

概述

本文档针对 ZStack 私有云实践过程进行说明。

读者对象

本文档适合以下工程师阅读：

- 运维工程师
- 测试工程师
- 存储工程师
- 预研工程师

术语定义

术语	概念
管理节点	安装ZStack系统的物理主机，提供UI管理、云系统部署功能
计算节点	也称之为物理主机，为云主机实例提供计算、内存、网络、存储的物理主机。运行KVM虚拟化的物理主机，简称“KVM主机”
集群	一个集群是类似主机（Host）组成的逻辑组。在同一个集群中的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）
云主机	ZStack 特制虚拟机，即运行在物理机上的虚拟机实例，具有独立的IP地址，可以安装部署应用服务
主存储	用于存储云主机的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	云主机所使用的镜像模板文件，包含了云主机的操作系统，也可以定制安装相应的软件
镜像服务器	也称之为备份存储服务器，存储云主机镜像文件的物理主机。虽然可以和管理节点或其他计算节点共享同一台物理服务器，但不建议在生

	产环境中这么部署
镜像仓库	可以为正在运行的云主机快速创建镜像，可以高效的管理虚拟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。
云盘	云主机的数据盘，给云主机提供额外的存储空间，一块云盘在同一时刻只能挂载到一个云主机。一个云主机最多可以挂载24块云盘
计算规格	启动云主机涉及到的CPU数量、内存大小、网络设置等规格定义
云盘规格	创建云盘容量大小的规格定义
安全组	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
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
二层网络	计算节点的物理网卡设备名称，例如eth0
三层网络	云主机需要使用的网络配置，包括IP地址范围，网关，DNS等
云路由	云路由通过定制的Linux云主机来实现的多种网络服务
管理网络	ZStack管理物理机和其他云资源的网络
公有网络	云主机连接和使用的网络
私有网络	云主机连接和使用的内部网络
弹性 IP	公有网络接入到私有网络的IP地址

中英文术语对照

管理节点	Management Node
区域	Zone
集群	Cluster
物理机	Host
云主机	VM Instance
镜像服务器（备份服务器）	Backup Storage
镜像仓库	Image Store
主存储	Primary Storage
镜像	Image

云盘	Volume
二层网络	L2 Network
三层网络	L3 Network
云路由	Virtual Router
安全组	Security Group
计算规格	Instance Offering
云盘规格	Disk Offering
扁平网络模式	Flat Network Mode
本地存储	Local Storage
弹性IP	EIP
无域间路由	CIDR

修改记录

修改记录积累了每次文档更新的说明。最新版本的文档包含以前所有文档版本的更新内容。

文档版本 1.0.0 (2016-05-05)

第一次正式发布。

文档版本 1.1.0 (2016-07-01)

第二次正式发布，添加 1.2.3 分布式存储。

文档版本 1.1.1 (2016-07-04)

第三次正式发布，勘误修复。

文档版本 1.10.0 (2017-03-10)

1.10 版本发布

目 录

1.1 基础环境.....	8
1.2 整体规划.....	9
1.2.1 本地存储.....	9
1.2.2 NFS 商业存储.....	11
1.2.3 融合型分布式存储.....	14

ZStack 可用于部署企业私有云，实现企业 IT 基础设施的云化，加速 IT 基础资源申请的响应时间，改变 IT 资源的交付方式。目前，ZStack 提供云计算基础设施即服务（Infrastructure as a Service, IaaS），即提供云主机服务。ZStack 提供便利的图形界面访问，也可提供命令行（zstack-cli）和 API 接口，方便与企业内部系统和业务系统对接，实现灵活的工作流程和业务伸缩能力。

当企业评估 ZStack 时，可通过单节点快速安装方式，即时体验其大部分功能。当企业计划使用 ZStack 支撑起业务环境时，需注意一些整体规划与实践操作，包括网络冗余，存储备份和灵活扩展等方面。本文档描述基于 ZStack 构建企业私有云实践。

1.1 基础环境

服务器硬件配置推荐如下：

服务器	CPU	内存	阵列卡	硬盘	网络	远程卡 IPMI	电源
Sever 1	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	NL SAS 2TB * 8	1GbE 网口 2 个	IDRAC/ILO	双电源
Sever 2	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	NL SAS 2TB * 8	1GbE 网口 2 个	IDRAC/ILO	双电源
Sever 3	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	NL SAS 2TB * 8	1GbE 网口 2 个	IDRAC/ILO	双电源

硬盘方面，根据容量和带宽，可选 SATA 和 SAS 硬盘。服务器上架首次启动后，设置阵列卡模式为 RAID 10 和缓存模式（Write-Back），并进入 BIOS 打开硬件虚拟化 CPU-VT 与超线程功能，关闭 CPU 节能模式。远程卡是服务器厂商提供的基于以太网远程管理带外卡，管理员通过浏览器即可方便管理服务器，包括上/下电操作和控制台访问。

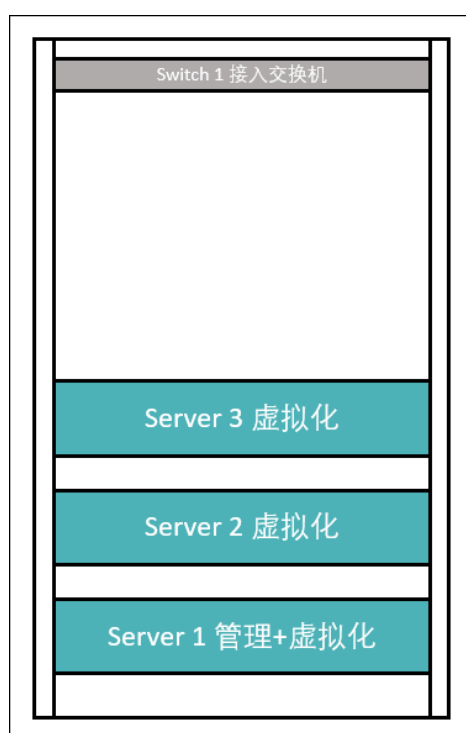
另外，企业 IDC 需准备千兆网络交换机，以便私有云环境接入。

1.2 整体规划

1.2.1 本地存储

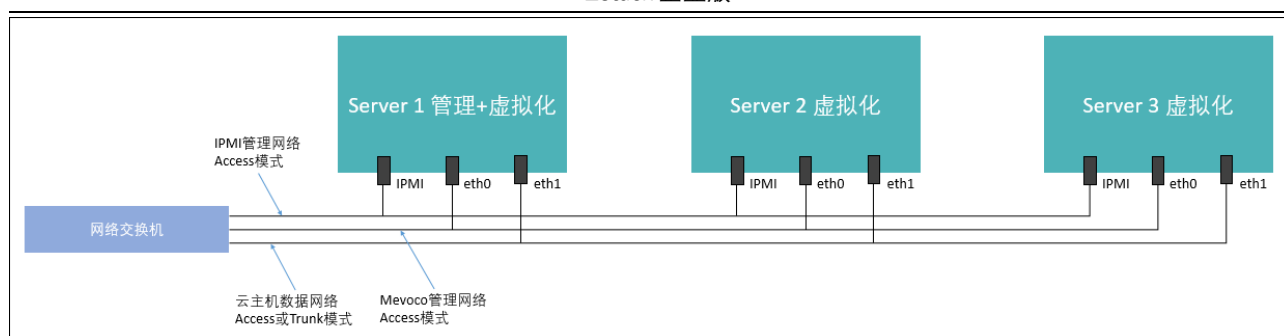
根据上一节服务器推荐配置，采用 ZStack 本地存储方案部署。

服务器阵列卡配置 RAID10 或 RAID6 模式。RAID10 相比 RAID6 的读写带宽更高；而 RAID6 采用纠删码技术，更高容量利用率和数据安全性。三个服务器以及千兆网络交换机的安装结构如下：



图一 ZStack 本地存储方案物理空间分布

其中 Server 1 将部署 ZStack 管理服务器和虚拟化环境，Server 2 和 Server 3 安装虚拟化环境。ZStack 管理服务核心数据存放在数据库 MariaDB，ZStack 将会对核心数据进行定期备份，管理员可将备份数据备份到存放点。三个服务器以及千兆网络交换机的连接拓扑如下：



图二 ZStack 本地存储方案网络拓扑图

上图显示，分别有三个网络：**IPMI 管理网络**、**ZStack 管理网络**和**云主机数据网络**。IPMI 管理网络用于承载服务器带外管理流量；ZStack 管理网络用于承载 ZStack 与其他虚拟化节点的通信流量，当前使用 eth0 网卡；云主机数据网络用于承载云主机之间以及面向访问用户的通信流量，当前使用 eth1 网卡。

IPMI 管理网络和 ZStack 管理网络，对应的交换机端口配置 Access 模式；云主机数据网络，若对应的交换机端口配置 Access 模式，在 ZStack 初始化界面的【三层网络】配置，选择【L2NoVlanNetwork】；若对应的交换机端口配置 Trunk 模式，在 ZStack 初始化界面的【三层网络】配置，选择【L2VlanNetwork】，并填写对应使用的 VLAN ID。

在【主存储】步骤，选择本地存储，填写路径/zstack_ps/；在【备份存储】步骤，选择 Sftp，填写 Server 1 IP 地址和路径/zstack_bs/。

按照以上网络拓扑图部署后，可进行 ZStack 管理节点和虚拟化节点的安装，详细安装步骤可参考官网相关文档：<http://www.zstack.io/>。

1.2.2 NFS 商业存储

ZStack 云管理平台存储方案可采用提供 NFS 协议的商业存储。采用 NFS 商业存储方案，虚拟化服务器本地无需更多的本地存储空间，同时需配置可靠的存储网络（增加 2 个以太网口），故服务器硬件配置变更如下：

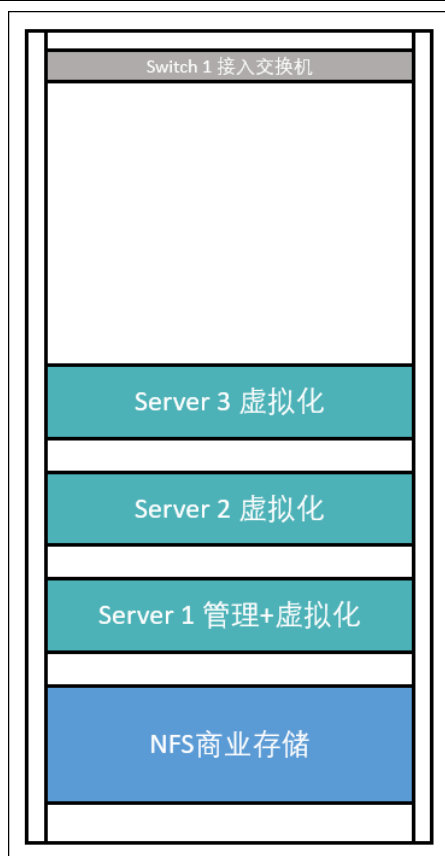
服务器	CPU	内存	阵列卡	硬盘	网络	远程卡 IPMI	电源
Sever 1	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	NL SAS 2TB * 2	1GbE 网口 4 个	IDRAC/ILO	双电源
Sever 2	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	NL SAS 500GB * 2	1GbE 网口 4 个	IDRAC/ILO	双电源
Sever 3	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	NL SAS 500GB * 2	1GbE 网口 4 个	IDRAC/ILO	双电源

Server 1 运行 ZStack 管理服务，包含关系型数据库 MariaDB 和非关系型数据库 Cassandra。MariaDB 存放 ZStack 核心运行数据，Cassandra 存放监控信息数据。由于监控信息较多，存放在 Server 1 的本地硬盘，所以配置 NL SAS 2TB 硬盘做 RAID1；Server 2 和 Server 3 运行虚拟化环境，不存放业务关键数据，配置 NL SAS 500GB 硬盘做 RAID1。

网络方面，增加 2 个千兆以太网口作为存储网络。由于存储网络承载所有的云主机云盘读写带宽，为了满足带宽容量和高可靠需要，需实施 2 个网口绑定操作。

关于 RHEL/CentOS 的网口绑定操作，可参考 RedHat 手册：https://access.redhat.com/documentation/enUS/Red_Hat_Enterprise_Linux/7/html/Networking_Guide/ch-Configure_Network_Bonding.html

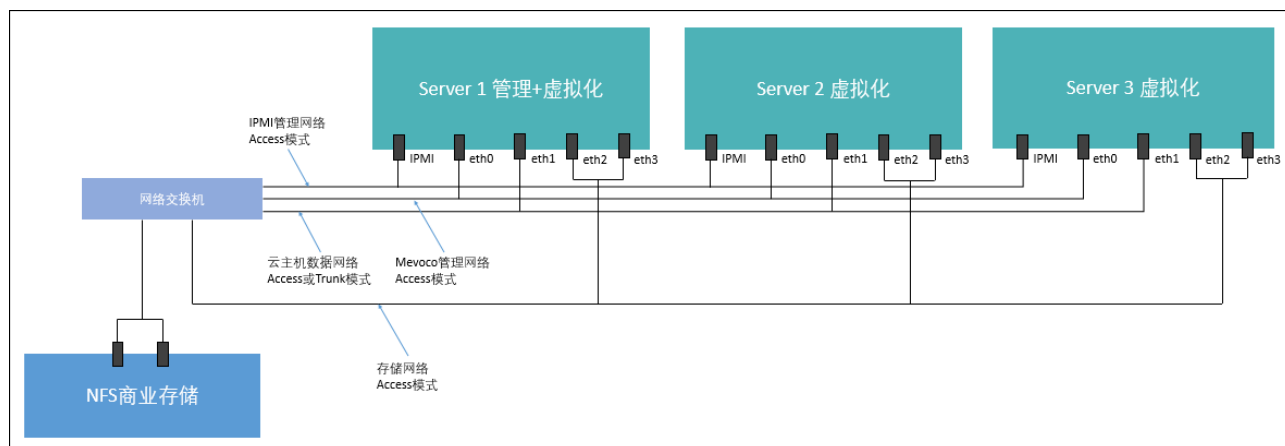
三个服务器、千兆网络交换机和 NFS 存储的安装结构如下：



图三 ZStack NFS 商业存储方案物理空间分布

存储方案采用 NFS 商业存储，在存储中创建三个 NFS 共享挂载池：**主存储池**、**备份存储池**和**数据库备份池**。主存储池用于存放云主机云盘数据，备份存储池用于存放镜像和模板，数据库备份池用于存放 MariaDB 导出备份信息。备份存储需预先挂载到 Server 1 的路径 /zstack_bs，在 ZStack 初始化过程中【备份存储】步骤便可填写此路径。

三个服务器、千兆网络交换机和 NFS 商业存储的连接拓扑如下：



图四 ZStack NFS 商业存储方案网络拓扑图

上图显示，分别有四个网络：**IPMI 管理网络**、**ZStack 管理网络**和**云主机数据网络**。IPMI 管理网络用于承载服务器带外管理流量；ZStack 管理网络用于承载 ZStack 与其他虚拟化节点的通信流量，当前使用 eth0 网卡；云主机数据网络用于承载云主机之间以及面向访问用户的通信流量，当前使用 eth1 网卡；存储网络用于承载云主机云盘读写的通信流量，当前使用 eth2 和 eth3。

相比 ZStack 本地存储方案，使用 NFS 商业存储方案，每个服务器增加配置 2 个以太网口作为存储网络。NFS 商业存储也进行 2 个以太网口绑定。

IPMI 管理网络、ZStack 管理网络和存储网络，对应的交换机端口配置 Access 模式；云主机数据网络，若对应的交换机端口配置 Access 模式，在 ZStack 初始化界面的【三层网络】配置，选择【L2NoVlanNetwork】；若对应的交换机端口配置 Trunk 模式，在 ZStack 初始化界面的【三层网络】配置，选择【L2VlanNetwork】，并填写对应使用的 VLAN ID。

在【主存储】步骤，选择 NFS 存储，填写 NFS 存储的 IP 地址以及主存储池目录；在【备份存储】步骤，选择 Sftp，填写 Server 1 IP 地址和路径/zstack_bs/（管理员预先挂载 NFS 存储的备份存储池目录到/zstack_bs/）。

关于 ZStack 数据库备份，管理员预先挂载 NFS 存储的数据库备份池到路径/mysql_bk/，

```
[root@zstack-1~]# crontab -e
...
0 0,12 * * * /usr/bin/rsync -avu /var/lib/zstack/mysql-backup/* /mysql_bk/ > /dev/null 2>&1
...
```

添加以下备份命令到 crontab：

由于 ZStack 核心数据备份到 NFS 存储，如果 Server 1 出现故障时，通过备份数据即可恢复 ZStack 管理服务。

1.2.3 融合型分布式存储

ZStack 云管理平台存储方案可采用分布式存储，例如 Ceph、FusionStor 等。分布式存储解决方案有两种类型，一种是融合型分布式存储；第二种是分离式分布式存储。融合型分布式存储主要是指提供分布式存储的物理机服务器也同时运行主机。融合型分布式存储与当前主流的超融合一体机的概念类似。采用融合型分布式存储的私有云最小需要三台物理服务器，之后可以根据需求无缝扩展。采用融合型分布式存储方案，虚拟化服务器需要提供高性能的 SSD 硬盘、SAS 盘和万兆网卡。故服务器硬件的推荐配置变更如下：

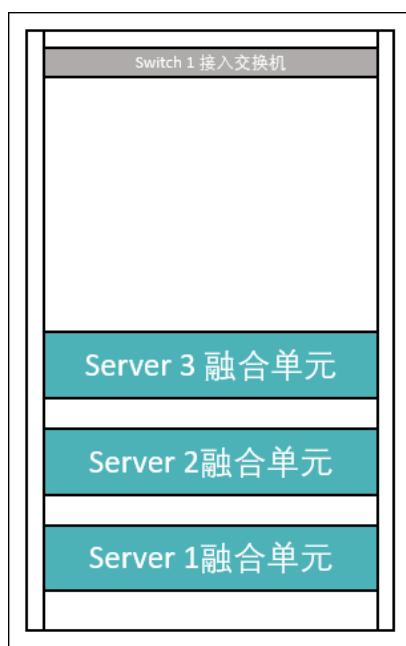
服务器	CPU	内存	阵列卡	硬盘	网络	远程卡 IPMI	电源
Sever 1	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	SSD 400G *4 NL SAS 1TB * 12	10GbE 网 口 2 个 1GbE 网卡 2 个	IDRAC/ILO	双电源
Sever 2	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	SSD 400G *4 NL SAS 1TB * 12	10GbE 网 口 2 个 1GbE 网卡 2 个	IDRAC/ILO	双电源
Sever 3	Intel Xeon E5-2680v3 2 颗	DDR4-2133 16GB 16 条	1GB 缓存 RAID10/5/6	SSD 400G *4 NL SAS 1TB * 12	10GbE 网 口 2 个 1GbE 网卡 2 个	IDRAC/ILO	双电源

Server 1 和 Server 2 运行 ZStack 集群管理服务，包含关系型数据库 MariaDB 和非关系型数据库 Cassandra。MariaDB 存放 ZStack 核心运行数据，Cassandra 存放监控信息数据。前两块硬盘使用阵列卡做 RAID1 作为系统盘，其余硬盘提供给分布式存储使用；Server1、Server 2 和 Server 3 运行超融合虚拟化环境。

网络方面，增加 2 个万兆以太网口作为存储网络。由于存储网络承载所有的云主机云盘读写带宽，为了满足带宽容量和高可靠需要，需实施 2 个网口绑定操作。

关于 RHEL/CentOS 的网口绑定操作，可参考 RedHat 手册：https://access.redhat.com/documentation/enUS/Red_Hat_Enterprise_Linux/7/html/Networking_Guide/ch-Configure_Network_Bonding.html

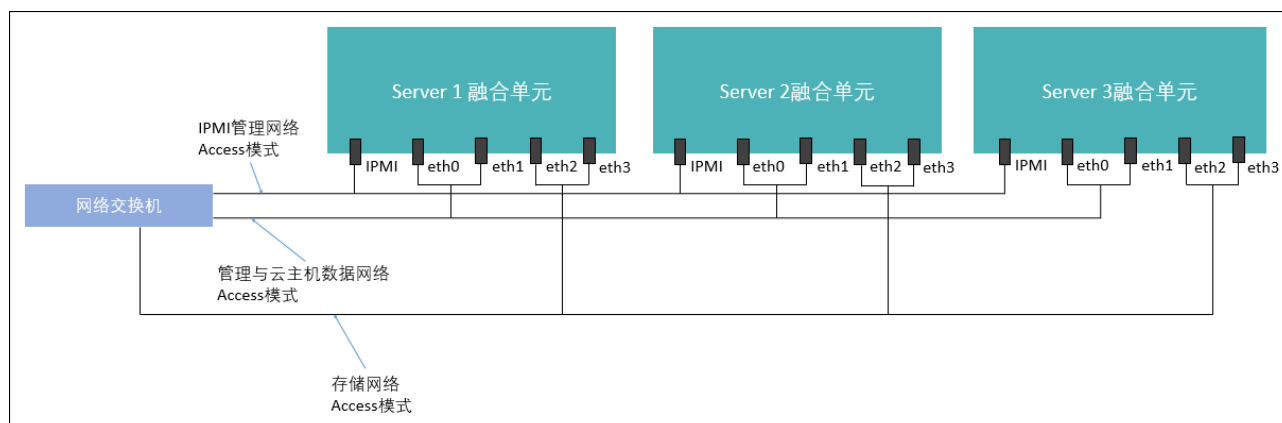
三个服务器、网络交换机的安装结构如下：



图五 ZStack 分布式存储方案物理空间分布

存储方案采用分布式存储，ZStack 使用该存储作为**主存储池**和**备份存储池**。主存储池用于存放云主机云盘数据，备份存储池用于存放镜像和模板。在 ZStack 初始化过程中【主存储】和【备份存储】需要填写对应分布式存储的 IP 地址。

三个服务器和网络交换机连接拓扑如下：



图六 ZStack 融合型分布式存储方案网络拓扑图

上图显示，分别有三个网络：**IPMI 管理网络**、**管理网络与云主机数据网络**和**存储网络**。IPMI 管理网络用于承载服务器带外管理流量；管理网络与云主机数据网络用于承载 ZStack 与其他虚拟化节点的通信流量，也承载云主机数据流量，使用 eth0 和 eth1 双网卡绑定；存储

网络用于承载云主机云盘读写的通信流量，使用 eth2 和 eth3 双网卡绑定。

相比前两个私有云存储方案，使用融合型分布式存储，每个服务器增加配置 2 个以太网口作为存储网络，数据将以副本形式分布在三个节点中（可配置 2 副本或 3 副本）。

IPMI 管理网络和存储网络，对应的交换机端口配置 Access 模式；管理网络和云主机数据网络，交换机端口配置 Access 模式，在 ZStack 初始化界面的【三层网络】配置，选择【L2NoVlanNetwork】；若对应的交换机端口配置 Trunk 模式，在 ZStack 初始化界面的【三层网络】配置，选择【L2VlanNetwork】，并填写对应使用的 VLAN ID。

在【主存储】步骤，选择 Ceph 或 FusionStor 存储，填写分布式存储的访问 IP 地址；在【备份存储】步骤，根据上一步选择存储的类型，默认候选该类型存储，并填写分布式存储的访问 IP 地址。

由于配置 ZStack 集群管理，整个**融合架构无单点故障**，通过高可用自愈能力，在发生节点失效后即进行自动恢复，满足业务连续性要求。

更多的配置与实践指导请访问 ZStack 官方网站 <http://www.zstack.io/>。