阿里云

ZStack for Alibaba Cloud

Shared Block 部署教程（FC存储）

产品版本：V2.5.1

文档版本：20180710

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。	**禁止**： 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。	**说明**： 导出的数据中包含敏感信息，请妥善保存。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	**说明**： 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**。
courier字体	命令。	执行 **cd /d C:/windows** 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	**bae log list --instanceid** *Instance_ID*
[]或者[a|b]	表示可选项，至多选择一个。	**ipconfig** *[-all|-t]*
{}或者{a|b}	表示必选项，至多选择一个。	**swich** *{stand | slave}*

目录

法律声明.....I
通用约定.....I
1 概述.....1
2 环境准备.....2
2.1 物理服务器.....2
2.2 存储网络拓扑.....2
3 部署Shared Block主存储.....4
4 运维管理.....7
专有云术语表.....9
混合云术语表.....12

1 概述

Shared Block（共享块存储）是ZStack for Alibaba Cloud新支持的一种主存储类型，可以将用户在SAN存储上划分的LUN设备直接作为存储池，再提供给业务云主机使用。与之前Shared Mount Point（SMP）主存储类型不同，**Shared Block**具备便捷部署、灵活扩展、性能优异等优势。据实测数据显示，Shared Block可以完全发挥物理磁盘的性能。

ZStack for Alibaba Cloud Shared Block支持iSCSI、FC共享访问协议。

本文将描述虚拟化存储环境下，基于FC存储部署**Shared Block**共享块存储系统。

2 环境准备

2.1 物理服务器

本文预设服务器如下：

服务器	硬件配置	IP地址	配置主机名
zstack-mn	4核 8G内存	*172.20.12.249*	hostnamectl set-hostname zstack-mn
sharedblock-host1	4核 8G内存	*172.20.11.107*	hostnamectl set-hostname sharedblock-host1
sharedblock-host2	4核 8G内存	*172.20.11.244*	hostnamectl set-hostname sharedblock-host2
sharedblock-host3	4核 8G内存	*172.20.12.241*	hostnamectl set-hostname sharedblock-host3

以上服务器均使用ZStack for Alibaba Cloud定制版ISO安装。

- 所有服务器的操作系统版本需一致，即：全部基于c72 ISO安装或全部基于c74 ISO安装。
- 以**管理节点模式**安装其中一台服务器作为管理节点。
- 以**计算节点模式**或**专家模式**安装其它三台服务器。
- 操作系统安装完成，配置全局主机名解析：

```
[root@sharedblock-host1 ~]# cat /etc/hosts
...
172.20.11.107 sharedblock-host1
172.20.11.244 sharedblock-host2
172.20.12.241 sharedblock-host3
...
```

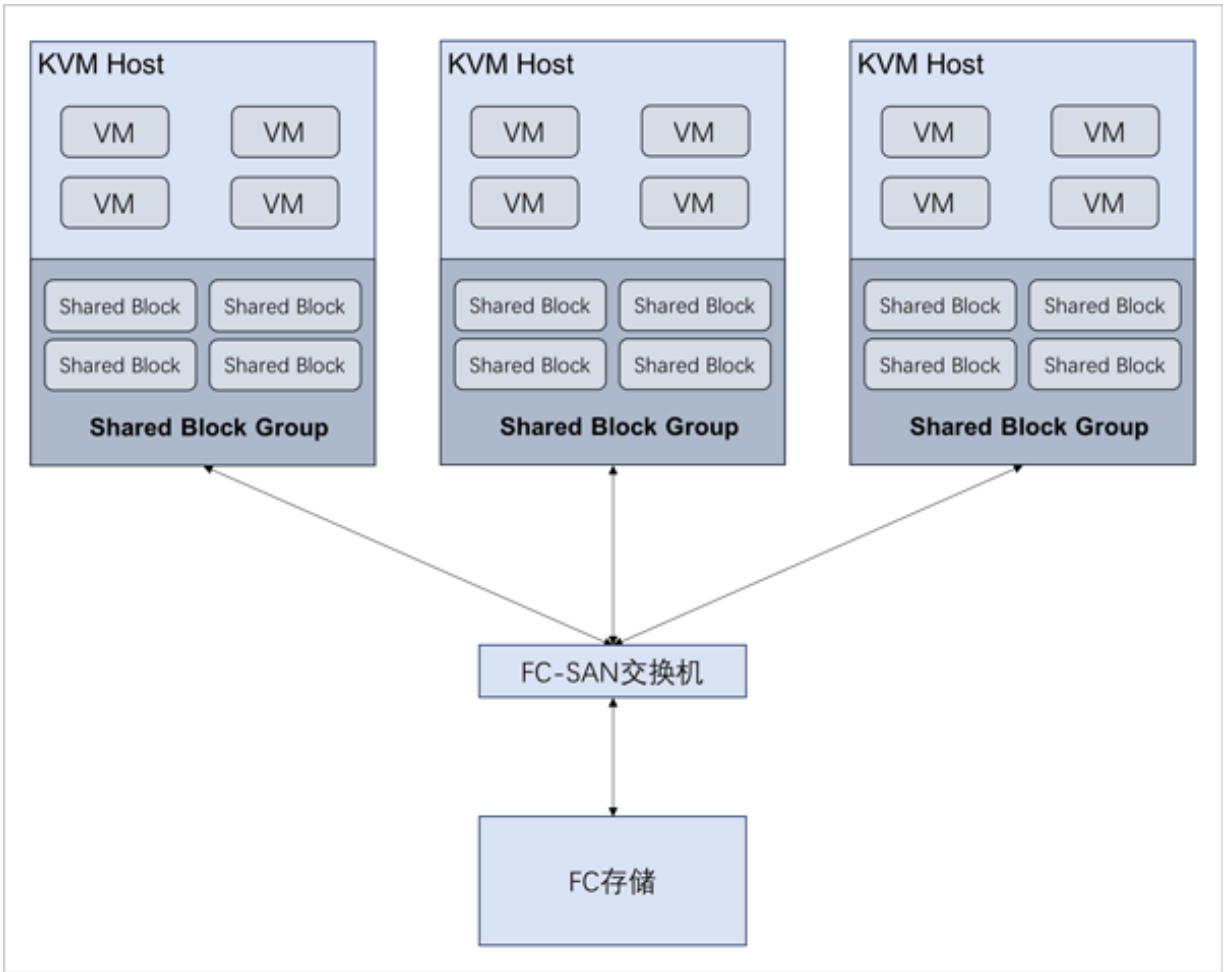
- 通过管理节点接管三台物理主机，物理主机自动部署安装LVM、Multipath等工具包。

至此，物理服务器准备就绪。

2.2 存储网络拓扑

物理服务器、存储设备和存储交换机，网络拓扑如*图 2-1: Shared Block与FC存储网络拓扑图*所示：

图 2-1: Shared Block与FC存储网络拓扑图



本文为了便于描述，此网络拓扑为单链路，若在生产环境部署Shared Block共享块存储系统，FC存储必须双控，且配置存储网络为高可用冗余模式。

3 部署Shared Block主存储

操作步骤

1. 添加裸设备

   管理员需在FC存储上划分出一定容量大小的裸设备（LUN设备），并通过以下方式在线扫描并发现磁盘：

```
[root@sharedblock-host1 ~]# echo "- - -" > /sys/class/scsi_host/host1/scan
[root@sharedblock-host2 ~]# echo "- - -" > /sys/class/scsi_host/host2/scan
[root@sharedblock-host3 ~]# echo "- - -" > /sys/class/scsi_host/host3/scan

# /dev/sdg为新发现设备
[root@sharedblock-host1 ~]# lsscsi -s
[0:0:0:3]    disk    DELL     MD32xx          0784  /dev/sdd   2.19TB
[0:0:1:3]    disk    DELL     MD32xx          0784  /dev/sdg   2.19TB
[0:0:2:3]    disk    DELL     MD32xx          0784  /dev/sdj   2.19TB
```

2. （可选）配置多路径访问

   如果FC存储是2个以上的链路，可配置多路径访问方式：

```
[root@sharedblock-host1 ~]# modprobe dm-multipath
[root@sharedblock-host1 ~]# modprobe dm-round-robin

[root@sharedblock-host1 ~]# service multipathd start
[root@sharedblock-host1 ~]# mpathconf --enable

# 显示多路径信息
[root@sharedblock-host1 ~]# multipath -ll

# 如果是多路径环境，将会显示/dev/mapper/mpatha，可直接使用该聚合设备
```

3. 获取磁盘唯一标识（例如：磁盘的WWID）

   - 如已配置多路径访问，请执行以下命令获取多路径设备*/dev/mapper/mpatha*（例如：dm-2）的WWID：

```
[root@sharedblock-host1 ~]# ll /dev/disk/by-id
total 0
lrwxrwxrwx 1 root root 10 May 25 17:52 dm-name-mpatha -> ../../dm-2
lrwxrwxrwx 1 root root 10 May 25 17:52 dm-uuid-mpath-36b083fe000daf018000
015505abbe00a -> ../../dm-2
lrwxrwxrwx 1 root root 10 May 25 17:52 lvm-pv-uuid-zUbBDF-i1Sg-YU3g-h2jN-m7hr-
H9Dh-qimsnF -> ../../dm-2
```

     由返回值可知，该多路径设备的WWID为**36b083fe000daf018000015505abbe00a**

   - 如未配置多路径访问，请执行以下命令获取获取LUN设备（例如：sde）的WWID：

```
[root@sharedblock-host1 ~]# ll /dev/disk/by-id
total 0
```

```
lrwxrwxrwx 1 root root  9 May 25 16:05 scsi-36b083fe000daf018000015505abbe00a -
> ../../sde
lrwxrwxrwx 1 root root  9 May 25 16:05 wwn-0x6b083fe000daf018000015505abbe00a -
> ../../sde
```

由返回值可知，该LUN设备的WWID为**scsi-36b083fe000daf018000015505abbe00a**或者**wwn-0x6b083fe000daf018000015505abbe00a**，均可作为磁盘唯一标识添加主存储。

4. 添加主存储

登录管理节点UI，在ZStack for Alibaba Cloud专有云主菜单，点击**硬件设施** > **主存储**按钮，进入**主存储**界面，点击**添加主存储**按钮，可根据以下示例输入相应内容：

- **名称**：输入主存储的名称
- **简介**：可选项，可留空不填
- **类型**：选择SharedBlock类型
- **清理块设备**：默认不勾选

> **说明**：
>
> - 勾选后将强制清理LUN设备中的文件系统、RAID或分区表中的签名，请谨慎选择。
> - 若LUN设备中未存放重要数据，可勾选此项。
> - 添加的LUN设备中不能有分区，否则会添加失败。

- **磁盘UUID**：输入磁盘唯一标识，例如：磁盘UUID、WWN、WWID；支持添加多个共享块设备
- **集群**：可选项，选择加载的集群

如*图 3-1: 添加Shared Block主存储*所示，点击**确定**按钮，添加Shared Block主存储。

图 3-1: 添加Shared Block主存储

支持以下两种场景：

- 一个集群支持添加多个Shared Block主存储。
- 一个集群支持添加1个Local Storage主存储+1个Shared Block主存储。

4 运维管理

Shared Block主存储支持以下操作：

- 启用：将处于停用状态的主存储启用。
- 停用：将主存储停用。停用主存储后，此主存储上的所有云盘被停用并且新的云主机、云盘、快照将无法创建。
- 重连：重新连接主存储。重连主存储会更新主存储相关的存储信息。

> **说明**：
> 如果有任意一台物理机正常连接到主存储，该主存储的就绪状态就会显示为**已连接**。

- 创建云盘：在主存储上创建云盘，此云盘为实例化云盘。
- 加载集群：将选中的主存储加载到指定的集群上。一个主存储支持挂载到多个集群；一个集群支持挂载一个Shared Block主存储，一个集群也支持挂载一个本地主存储和一个Shared Block主存储。
- 卸载集群：将主存储从指定的集群上卸载。
- 进入维护模式：主存储进入维护模式后，将停止所有使用该主存储的云主机（包括NeverStop云主机）。
- 删除：将主存储删除。执行删除操作前请从所有集群卸载该主存储，否则不能删除。删除主存储后，此主存储上的所有云主机和云盘都将被删除。

> **说明**：
> 删除主存储是非常危险的操作，此操作会直接删除该主存储上的所有云主机和云盘。即使重新添加此主存储，也无法自动识别原有的文件。

- 添加共享块：支持在线添加共享块设备。

添加共享块

Shared Block主存储支持在线添加共享块设备。

在**主存储**页面，选择某一SharedBlock主存储，进入其详情页，点击**共享块**，进入**共享块**子页面，点击**共享块**旁的**操作** > **添加**，弹出**添加共享块**界面，输入待添加的共享块设备的磁盘UUIID，可按需勾选是否强制清理块设备。

如图 *4-1:* 添加共享块所示，点击**确定**，Shared Block即时实现在线扩容。

图 4-1: 添加共享块

至此，基于FC存储部署**Shared Block**共享块存储系统介绍完毕。

专有云术语表

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）

云主机或云盘使用的镜像模板文件，镜像模板包括系统云盘镜像和数据云盘镜像。

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。

混合云术语表

访问密钥（AccessKey）

用于调用阿里云API或大河云联API的唯一凭证，AccessKey包括AccessKeyID（用于标识用户）和AccessKeySecret（用于验证用户密钥）。

数据中心（Data Center）

包含阿里云的地域和可用区等地域资源，用于匹配阿里云资源的地域属性。

地域（Region）

物理的数据中心，划分地区的基本单位，ZStack混合云的地域对应了阿里云端的地域。

可用区（Identity Zone）

在同一地域内，电力和网络互相独立的物理区域，ZStack混合云的可用区对应了阿里云端的可用区（Zone）。

存储空间（Bucket）

用于存储对象（Object）的容器，ZStack使用对象存储（OSS）里的Bucket来上传镜像文件。

ECS云主机（Elastic Compute Service）

阿里云端创建的ECS实例，可在ZStack混合云界面进行ECS云主机生命周期的管理。

专有网络VPC（Virtual Private Cloud）

用户基于阿里云构建的一个隔离的网络环境，不同的专有网络之间逻辑上彻底隔离。

虚拟交换机（VSwitch）

组成专有网络VPC的基础网络设备，可以连接不同的云产品实例。ZStack混合云的虚拟交换机对应了阿里云VPC下的虚拟交换机。

虚拟路由器（VRouter）

专有网络VPC的枢纽，可以连接专有网络的各个虚拟交换机，同时也是连接专有网络与其它网络的网关设备。ZStack支持查看VPC下的虚拟路由器。

路由表（Route Table）

虚拟路由器上管理路由条目的列表。

路由条目（Route Entry）

路由表中的每一项是一条路由条目。路由条目定义了通向指定目标网段的网络流量的下一跳地址。路由条目包括系统路由和自定义路由两种类型。ZStack支持自定义类型的路由条目。

安全组（Security Group）

针对云主机进行第三层网络的防火墙控制。ZStack混合云的安全组对应了阿里云端ECS云主机三层隔离的防火墙约束。

镜像（Image）

云主机使用的镜像模板文件，一般包括操作系统和预装的软件。ZStack支持上传本地镜像到阿里云，以及使用阿里云端镜像。

弹性公网IP（EIP）

阿里云端公有网络池中的IP地址，绑定弹性公网IP的ECS实例可以直接使用该IP进行公网通信。

VPN连接（VPN Connection）

通过建立点对点的IPsec VPN通道，实现企业本地数据中心的私有网络与阿里云端VPN网络进行通信。

VPN网关（VPN Gateway）

一款基于Internet，通过加密通道将本地数据中心和阿里云专有网络VPC安全可靠连接起来的服务。用户在阿里云VPC创建的IPsec VPN网关，与本地数据中心的用户网关配合使用。

VPN用户网关（Customer Gateway）

本地数据中心的VPN服务网关。可通过ZStack混合云创建VPN用户网关，并将VPN用户网关与VPN网关连接起来。

高速通道（Express Connect）

通过物理专线（即租用运营商的专线：电缆或光纤），连通本地数据中心到阿里云专线接入点，与阿里云VPC环境打通，实现云上云下不同网络间高速，稳定，安全的私网通信。

边界路由器（VBR）

用户申请的物理专线接入交换机的产品映射。用户在物理专线上可以创建边界路由器，边界路由器负责专线上的数据在阿里云上进行转发。通过边界路由器，用户数据可以直达阿里云VPC网络。

路由器接口（Router Interface）

一种虚拟的网络设备，可以挂载在路由器并与其他路由器接口进行高速通道互联，实现不同网络间的内网互通。