

OCFS2 共享存储实践指南

iSCSI 存储


版本 1.10

日期 2017-03-15

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 OCFS2 共享文件系统实践过程进行说明。

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”
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
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
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
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

中英文术语对照

管理节点	Management Node
物理机	Host
云主机	VM Instance
镜像服务器（备份服务器）	Backup Storage
主存储	Primary Storage
镜像	Image
云盘	Volume
集群	Cluster
区域	Zone
二层网络	L2 Network
三层网络	L3 Network
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

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 1.10（2017-03-15）

第一次正式发布。

目录

第一章 概述.....	7
第二章 环境准备.....	8
2.1 iSCSI 存储	8
2.2 物理服务器	10
2.3 存储网络拓扑.....	11
第三章 安装 OCFS2	12
3.1 升级系统内核.....	12
3.2 配置 iSCSI 发起.....	13
3.3 初始化集群	17
3.4 创建存储设备.....	23
第四章 运维管理.....	26
4.1 添加裸设备	26
4.2 移除裸设备	28
4.3 添加服务器	30
4.4 移除服务器	33

第一章 概述

OCFS2 是 Oracle 发起的开源共享文件系统项目，以开源版权协议 GPLv2 发布。

管理员通过安装 OCFS2 软件，把各节点共享访问的裸设备（Raw Device）以共享文件系统形式呈现。类似网络文件系统（NFS），应用程序可直接读写 OCFS2 目录而无需更改。

OCFS2 支持 iSCSI、FC 和 SAS 共享访问协议。OCFS2 常用于 Oracle 数据库集群和虚拟化存储环境。本文将描述虚拟化存储环境下，基于 iSCSI 安装 OCFS2 共享文件系统。

第二章 环境准备

2.1 iSCSI 存储

在安装 OCFS2 前，管理员需检查目标的存储是否支持 iSCSI 访问协议。通过存储设备出厂时候的产品说明书可快速查看 iSCSI 相关技术参数。若当前不具备存储设备，管理员可采用 OpenFiler、FreeNAS 或 NAS4Free 等基于 x86 服务器的软件存储方案，可提供 iSCSI 存储访问方式。

确认存储支持 iSCSI，记录存储 iSCSI IP 和控制器识别名。管理员需在存储中预先创建以下裸设备：

存储	IP 地址	控制器识别码	设备名	容量 (MB)	LUN ID
OpenFiler	172.20.12.221	iqn.2015-05.org.zstack:tsn.200001	zstack-sto-1	20000	0
OpenFiler	172.20.12.221	iqn.2015-05.org.zstack:tsn.200001	zstack-sto-2	25024	1
OpenFiler	172.20.12.221	iqn.2015-05.org.zstack:tsn.200001	zstack-sto-3	30016	2

以上服务每个控制识别码 CHAP 验证名为 **zstack**，CHAP 验证密码为 **password**。

注意：本文档使用 OpenFiler 构建 iSCSI 存储服务，并配置单个 IP 地址。管理员在生产环境中，务必使用具备商业服务的存储系统，并具备双控制器和多路径，整体架构上容忍模块单点故障，从而达到高可用。

上述的裸设备中，zstack-sto-1、zstack-sto-2 和 zstack-sto-3 将自我实现心跳检测，从而判断目录挂载节点的存储链路是否正常。同时，zstack-sto-1、zstack-sto-2 和 zstack-sto-3 是实际使用的存储设备，管理员可将这些存储设备通过 OCFS2 挂载后提供 ZStack 使用。

Target Configuration

LUN Mapping

Network ACL

CHAP Authentication

LUNs mapped to target: iqn.2015-05.org.zstack:tsn.200001

LUN Id.	LUN Path	R/W Mode	SCSI Serial No.	SCSI Id.	Transfer Mode	Unmap LUN
0	/dev/volume-1/zstack-sto-1	write-thru	LzedfI-cZI0-qKej	LzedfI-cZI0-qKej	blockio	Unmap
1	/dev/volume-1/zstack-sto-2	write-thru	FVnSC7-xfj8-rwq3	FVnSC7-xfj8-rwq3	blockio	Unmap
2	/dev/volume-1/zstack-sto-3	write-thru	21IqYi-RghT-0AE6	21IqYi-RghT-0AE6	blockio	Unmap

图一 OCFS2 挂载存储设备-存放数据

至此，iSCSI 存储端裸设备就绪。

2.2 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 6G 内存	172.20.12.249	hostnamectl set-hostname ZStack-mn
ocfs2-host1	4 核 6G 内存	172.20.11.107	hostnamectl set-hostname ocfs2-host1
ocfs2-host2	4 核 6G 内存	172.20.11.244	hostnamectl set-hostname ocfs2-host2
ocfs2-host3	4 核 6G 内存	172.20.12.241	hostnamectl set-hostname ocfs2-host3

以上服务器均使用 ZStack 社区定制 ISO 光盘版本 1.3 安装（或大于 1.3 版本）。

操作系统安装完成，配置防火墙规则：

```
[root@ocfs2-host1 ~]# vim /etc/sysconfig/iptables
...
-A INPUT -s 172.20.0.0/16 -j ACCEPT
...
[root@ocfs2-host1 ~]# service iptables restart

[root@ocfs2-host2 ~]# vim /etc/sysconfig/iptables
...
-A INPUT -s 172.20.0.0/16 -j ACCEPT
...
[root@ocfs2-host2 ~]# service iptables restart

[root@ocfs2-host3 ~]# vim /etc/sysconfig/iptables
...
-A INPUT -s 172.20.0.0/16 -j ACCEPT
...
[root@ocfs2-host3 ~]# service iptables restart
```

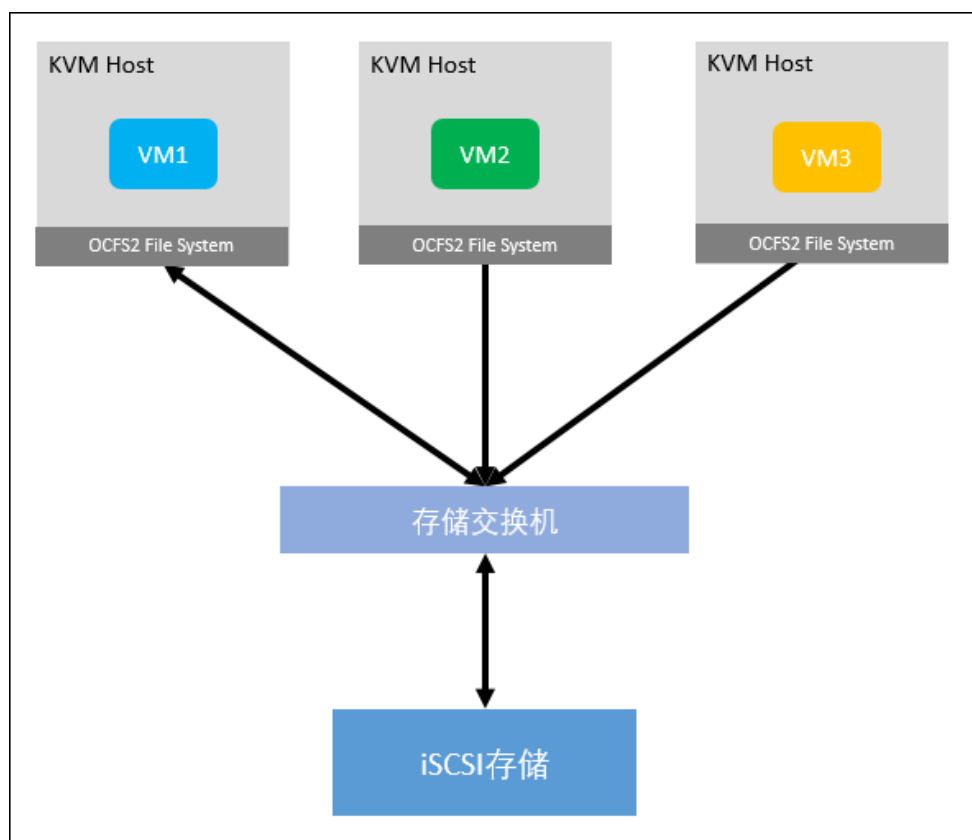
配置全局主机名解析：

```
[root@ocfs2-host1 ~]# cat /etc/hosts
...
172.20.11.107    ocfs2-host1
172.20.11.244    ocfs2-host2
172.20.12.241    ocfs2-host3
...
```

至此，物理服务器准备就绪。

2.3 存储网络拓扑

物理服务器、存储设备和存储交换机，网络拓扑如下：



图二 OCFS2 与 iSCSI 存储网络拓扑图

上图展示物理服务器与 iSCSI 存储的网络拓扑。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 OCFS2 共享文件系统，iSCSI 存储必须双控，且配置存储网络为高可用冗余模式。

第三章 安装 OCFS2

3.1 升级系统内核

使用 OCFS2 共享文件系统，需要 Linux 内核预编译 ocfs2.ko 等相关模块。目前，由于 RHEL/CentOS 7.2 中尚未预先编译，推荐使用 Oracle Linux UEK 内核。ZStack-Community-x86_64-DVD-1.3.0.iso 已包含 UEK 内核和 OCFS2 工具，管理员可通过以下操作，安装 UEK 内核和 OCFS2 工具：

```
# 升级软件包
[root@localhost ~]# yum -y --disablerepo=* --enablerepo=zstack-local,uek4-ocfs2 update

# 安装 UEK 内核和 OCFS2 工具
[root@localhost ~]# yum --disablerepo=* --enablerepo=zstack-local,uek4-ocfs2 \
install kernel-uek kernel-uek-devel kernel-uek-doc kernel-uek-firmware \
dtrace-modules ocfs2-tools ocfs2-tools-devel iscsi-initiator-utils \
device-mapper-multipath device-mapper-multipath-sysvinit

# 默认从 UEK 内核引导
[root@localhost ~]# grub2-set-default "CentOS Linux (4.1.12-37.2.2.el7uek.x86_64) 7 (Core)"

# 检查默认启动项
[root@localhost ~]# grub2-editenv list
```

完成以上步骤后，重启物理主机。

```
# 重启物理主机
[root@localhost ~]# reboot

# 重启后查看内核版本
[root@localhost ~]# uname -a
Linux localhost 4.1.12-37.2.2.el7uek.x86_64 #1 SMP Thu Mar 31 16:04:38 UTC 2016 x86_64
x86_64 x86_64 GNU/Linux
```

注意：以上安装 UEK 内核和 OCFS2 工具的过程，务必在每个运行云主机的物理主机配置。原则上，ZStack 管理服务器不运行云主机，无需安装 UEK 内核和 OCFS2 工具。

ocfs2-host1、ocfs2-host2 和 ocfs2-host3 均执行以上步骤。至此，物理主机已经安装 UEK 内核和 OCFS2 工具。

3.2 配置 iSCSI 发起

检查各物理服务器网络，确认可互相 Ping，也能 Ping 存储服务器。

配置物理服务器 iSCSI 发起配置：

```
# 配置 ocfs2-host1 iSCSI 发起名
[root@ocfs2-host1 ~]# vim /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.2016-05.org.zstack:ocfs2-host1

# 配置 ocfs2-host1 iSCSI CHAP
[root@ocfs2-host1 ~]# vim /etc/iscsi/iscsid.conf
...
node.session.auth.authmethod = CHAP
node.session.auth.username = zstack
node.session.auth.password = password
...

# 配置 ocfs2-host2 iSCSI 发起名
[root@ocfs2-host2 ~]# vim /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.2016-05.org.zstack:ocfs2-host2

# 配置 ocfs2-host2 iSCSI CHAP
[root@ocfs2-host2 ~]# vim /etc/iscsi/iscsid.conf
...
node.session.auth.authmethod = CHAP
node.session.auth.username = zstack
node.session.auth.password = password
...

# 配置 ocfs2-host3 iSCSI 发起名
[root@ocfs2-host3 ~]# vim /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.2016-05.org.zstack:ocfs2-host3

# 配置 ocfs2-host3 iSCSI CHAP
[root@ocfs2-host1 ~]# vim /etc/iscsi/iscsid.conf
...
node.session.auth.authmethod = CHAP
node.session.auth.username = zstack
node.session.auth.password = password
...
```

扫描 iSCSI 存储：

```
[root@ocfs2-host1 ~]# iscsiadm -m discovery -t sendtargets -p 172.20.12.221
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.200001
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.100001
```

```
[root@ocfs2-host2 ~]# iscsiadm -m discovery -t sendtargets -p 172.20.12.221
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.200001
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.100001
```

```
[root@ocfs2-host3 ~]# iscsiadm -m discovery -t sendtargets -p 172.20.12.221
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.200001
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.100001
```

记录 iSCSI 存储控制器标识。在 3 个物理服务器登陆到 iSCSI 存储：

```
[root@ocfs2-host1 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.100001 --portal 172.20.12.221 --login
[root@ocfs2-host1 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.200001 --portal 172.20.12.221 --login
```

```
[root@ocfs2-host2 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.100001 --portal 172.20.12.221 --login
[root@ocfs2-host2 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.200001 --portal 172.20.12.221 --login
```

```
[root@ocfs2-host3 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.100001 --portal 172.20.12.221 --login
[root@ocfs2-host3 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.200001 --portal 172.20.12.221 --login
```

检查识别的裸设备：

```
[root@ocfs2-host1 ~]# fdisk -l | grep sd
Disk /dev/sda: 21.0 GB, 20971520000 bytes, 40960000 sectors
Disk /dev/sdb: 26.2 GB, 26239565824 bytes, 51249152 sectors
Disk /dev/sdc: 31.5 GB, 31474057216 bytes, 61472768 sectors
[root@ocfs2-host1 ~]# lsscsi
[1:0:0:0]    cd/dvd  QEMU      QEMU DVD-ROM    1.5.  /dev/sr0
[2:0:0:0]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sda
[2:0:0:1]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdb
[2:0:0:2]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdc

[root@ocfs2-host2 ~]# fdisk -l | grep sd
Disk /dev/sda: 21.0 GB, 20971520000 bytes, 40960000 sectors
Disk /dev/sdb: 26.2 GB, 26239565824 bytes, 51249152 sectors
Disk /dev/sdc: 31.5 GB, 31474057216 bytes, 61472768 sectors

[root@ocfs2-host2 ~]# lsscsi
[1:0:0:0]    cd/dvd  QEMU      QEMU DVD-ROM    1.5.  /dev/sr0
[2:0:0:0]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sda
[2:0:0:1]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdb
[2:0:0:2]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdc

[root@ocfs2-host3 ~]# fdisk -l | grep sd
Disk /dev/sda: 21.0 GB, 20971520000 bytes, 40960000 sectors
Disk /dev/sdb: 26.2 GB, 26239565824 bytes, 51249152 sectors
Disk /dev/sdc: 31.5 GB, 31474057216 bytes, 61472768 sectors

[root@ocfs2-host3 ~]# lsscsi
[1:0:0:0]    cd/dvd  QEMU      QEMU DVD-ROM    1.5.  /dev/sr0
[2:0:0:0]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sda
[2:0:0:1]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdb
[2:0:0:2]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdc
```

另外，可以通过 lsscsi 带参数 -st, -i, 查看更多信息：

```
[root@ocfs2-host1 ~]# lsscsi -st
[1:0:0:0]    cd/dvd  ata:                                /dev/sr0      -
[2:0:0:0]    disk    iqn.2015-05.org.zstack:tsn.200001,t,0x1  /dev/sda      20.9GB
[2:0:0:1]    disk    iqn.2015-05.org.zstack:tsn.200001,t,0x1  /dev/sdb      26.2GB
[2:0:0:2]    disk    iqn.2015-05.org.zstack:tsn.200001,t,0x1  /dev/sdc      31.4GB

[root@ocfs2-host1 ~]# lsscsi -i
[1:0:0:0]    cd/dvd  QEMU      QEMU DVD-ROM      1.5.  /dev/sr0      -
[2:0:0:0]    disk    OPNFILER VIRTUAL-DISK  0      /dev/sda      14f504e46494c4552524d394b6a462d52696e482d4777564e
[2:0:0:1]    disk    OPNFILER VIRTUAL-DISK  0      /dev/sdb      14f504e46494c45525462666848472d5763654a2d31677058
[2:0:0:2]    disk    OPNFILER VIRTUAL-DISK  0      /dev/sdc      14f504e46494c45524b6f747464792d476531332d3359426d
```

至此，3 个物理服务器发现 iSCSI 存储设备。另外，如果 iSCSI 存储是 2 个以上的链路，可配置多路径访问方式：

```
[root@ocfs2-host1 ~]# modprobe dm-multipath
[root@ocfs2-host1 ~]# modprobe dm-round-robin

[root@ocfs2-host1 ~]# service multipathd start
[root@ocfs2-host1 ~]# mpathconf --enable

# 显示多路径信息
[root@ocfs2-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
```


3.3 初始化集群

物理服务器发现 iSCSI 裸设备后，方可进行 OCFS2 集群初始化（若使用多路径，则可以根据 SCSI ID 设定裸设备别名，防止因为服务器重启后盘符变更导致挂载失效）。

```
# 创建 DLM 目录
[root@ocfs2-host1 ~]# mkdir -p /dlm
[root@ocfs2-host2 ~]# mkdir -p /dlm
[root@ocfs2-host3 ~]# mkdir -p /dlm

# 创建 OCFS2 集群配置
[root@ocfs2-host1 ~]# o2cb add-cluster zstackstorage
[root@ocfs2-host1 ~]# o2cb add-node zstackstorage ocfs2-host1 --ip 172.20.11.107
[root@ocfs2-host1 ~]# o2cb add-node zstackstorage ocfs2-host2 --ip 172.20.11.244
[root@ocfs2-host1 ~]# o2cb add-node zstackstorage ocfs2-host3 --ip 172.20.12.241

[root@ocfs2-host1 ~]# o2cb heartbeat-mode zstackstorage local
```

```
# 查看 OCFS2 配置
[root@ocfs2-host1 ~]# cat /etc/ocfs2/cluster.conf
cluster:
    heartbeat_mode = local
    node_count = 3
    name = zstackstorage

node:
    number = 0
    cluster = zstackstorage
    ip_port = 7777
    ip_address = 172.20.11.107
    name = ocfs2-host1

node:
    number = 1
    cluster = zstackstorage
    ip_port = 7777
    ip_address = 172.20.11.244
    name = ocfs2-host2

node:
    number = 2
    cluster = zstackstorage
    ip_port = 7777
    ip_address = 172.20.12.241
    name = ocfs2-host3
```

执行 OCFS2 集群初始化：

```
[root@ocfs2-host1 ~]# /etc/init.d/o2cb configure
Configuring the O2CB driver.
```

This will configure the on-boot properties of the O2CB driver.
The following questions will determine whether the driver is loaded on boot. The current values will be shown in brackets ('[]'). Hitting <ENTER> without typing an answer will keep that current value. Ctrl-C will abort.

```
Load O2CB driver on boot (y/n) [n]: y
Cluster stack backing O2CB [o2cb]:
Cluster to start on boot (Enter "none" to clear) [ocfs2]: zstackstorage
Specify heartbeat dead threshold (>=7) [31]: 121
Specify network idle timeout in ms (>=5000) [30000]: 30000
Specify network keepalive delay in ms (>=1000) [2000]: 5000
Specify network reconnect delay in ms (>=2000) [2000]: 4000
Writing O2CB configuration: OK
checking debugfs...
Loading filesystem "ocfs2_dlmfs": OK
Mounting ocfs2_dlmfs filesystem at /dlm: OK
Setting cluster stack "o2cb": OK
Registering O2CB cluster "zstackstorage": OK
Setting O2CB cluster timeouts : OK
Starting local heartbeat for cluster "zstackstorage": OK
```

创建自启动：

```
[root@ocfs2-host1 ~]# systemctl enable o2cb.service
[root@ocfs2-host1 ~]# systemctl enable ocfs2.service
```

检查 ocfs2-host1 上 o2cb 状态：

```
[root@ocfs2-host1 ~]# /etc/init.d/o2cb online
[root@ocfs2-host1 ~]# /etc/init.d/o2cb status
Driver for "configfs": Loaded
Filesystem "configfs": Mounted
Stack glue driver: Loaded
Stack plugin "o2cb": Loaded
Driver for "ocfs2_dlmfs": Loaded
Filesystem "ocfs2_dlmfs": Mounted
Checking O2CB cluster "zstackstorage": Online
  Heartbeat dead threshold: 121
  Network idle timeout: 30000
  Network keepalive delay: 5000
  Network reconnect delay: 4000
  Heartbeat mode: Local
Checking O2CB heartbeat: No Active
```

状态显示正常。【Nodes in O2CB cluster: 0】，显示只有 ID 为 0 的服务器加载。添加 Kernel

Panic 参数：

```
[root@ocfs2-host1 ~]# vim /etc/sysctl.conf
...
kernel.panic = 30
kernel.panic_on_oops = 1
...

[root@ocfs2-host1 ~]# sysctl -p
```

增加 ocfs2-host2 节点到 OCFS2 集群：

```
[root@ocfs2-host2 ~]# mkdir -p /etc/ocfs2/

[root@ocfs2-host1 ~]# scp /etc/ocfs2/cluster.conf ocfs2-host2:/etc/ocfs2/
[root@ocfs2-host1 ~]# scp /etc/sysconfig/o2cb ocfs2-host2:/etc/sysconfig/

[root@ocfs2-host2 ~]# systemctl start o2cb.service
checking debugfs...
Loading filesystem "ocfs2_dlmfs": OK
Mounting ocfs2_dlmfs filesystem at /dlm: OK
Setting cluster stack "o2cb": OK
Registering O2CB cluster "zstackstorage": OK
Setting O2CB cluster timeouts : OK
Starting local heartbeat for cluster "zstackstorage": OK

# 查看状态
[root@ocfs2-host2 ~]# /etc/init.d/o2cb status
Driver for "configfs": Loaded
Filesystem "configfs": Mounted
Stack glue driver: Loaded
Stack plugin "o2cb": Loaded
Driver for "ocfs2_dlmfs": Loaded
Filesystem "ocfs2_dlmfs": Mounted
Checking O2CB cluster "zstackstorage": Online
  Heartbeat dead threshold: 121
  Network idle timeout: 30000
  Network keepalive delay: 5000
  Network reconnect delay: 4000
  Heartbeat mode: Local
Checking O2CB heartbeat: No Active

[root@ocfs2-host2 ~]# systemctl enable o2cb.service
[root@ocfs2-host1 ~]# systemctl enable ocfs2.service
[root@ocfs2-host2 ~]# /etc/init.d/o2cb online
```

ocfs2-host1 和 ocfs2-host2 已经添加到集群。

增加 ocfs2-host3 节点到 OCFS2 集群：

```
[root@ocfs2-host3 ~]# mkdir -p /etc/ocfs2/

[root@ocfs2-host1 ~]# scp /etc/ocfs2/cluster.conf ocfs2-host3:/etc/ocfs2/
[root@ocfs2-host1 ~]# scp /etc/sysconfig/o2cb ocfs2-host3:/etc/sysconfig/

[root@ocfs2-host2 ~]# /etc/init.d/o2cb start
checking debugfs...
Loading filesystem "ocfs2_dlmfs": OK
Mounting ocfs2_dlmfs filesystem at /dlm: OK
Setting cluster stack "o2cb": OK
Registering O2CB cluster "zstackstorage": OK
Setting O2CB cluster timeouts : OK
Starting local heartbeat for cluster "zstackstorage": OK

# 查看状态
[root@ocfs2-host3 ~]# /etc/init.d/o2cb status
Driver for "configfs": Loaded
Filesystem "configfs": Mounted
Stack glue driver: Loaded
Stack plugin "o2cb": Loaded
Driver for "ocfs2_dlmfs": Loaded
Filesystem "ocfs2_dlmfs": Mounted
Checking O2CB cluster "zstackstorage": Online
  Heartbeat dead threshold: 121
  Network idle timeout: 30000
  Network keepalive delay: 5000
  Network reconnect delay: 4000
  Heartbeat mode: Local
Checking O2CB heartbeat: No Active

[root@ocfs2-host2 ~]# systemctl enable o2cb.service
[root@ocfs2-host1 ~]# systemctl enable ocfs2.service
[root@ocfs2-host2 ~]# /etc/init.d/o2cb online
```

ocfs2-host1、ocfs2-host2 和 ocfs2-host3 已经添加到集群。

3.4 创建存储设备

通过 mkfs.ocfs2 工具可对共享裸设备进行格式化（只需要在其中一个 host 执行）：

```
[root@ocfs2-host1 ~]# mkfs.ocfs2 --cluster-stack=o2cb -C 256K -J size=128M -N 16 -L
ocfs2-disk1 --cluster-name=zstackstorage --fs-feature-level=default -T vmstore /dev/sda
mkfs.ocfs2 1.8.6
Cluster stack: o2cb
Cluster name: zstackstorage
Stack Flags: 0x1
NOTE: Feature extended slot map may be enabled
Filesystem Type of vmstore
Label: ocfs2-disk1
Features: sparse extended-slotmap backup-super unwritten inline-data strict-journal-super
xattr indexed-dirs refcount discontig-bg
Block size: 4096 (12 bits)
Cluster size: 262144 (18 bits)
Volume size: 20971520000 (80000 clusters) (5120000 blocks)
Cluster groups: 3 (tail covers 15488 clusters, rest cover 32256 clusters)
Extent allocator size: 4194304 (1 groups)
Journal size: 134217728
Node slots: 16
Creating bitmaps: done
Initializing superblock: done
Writing system files: done
Writing superblock: done
Writing backup superblock: 3 block(s)
Formatting Journals: done
Growing extent allocator: done
Formatting slot map: done
Formatting quota files: done
Writing lost+found: done
mkfs.ocfs2 successful

# 挂载文件系统
[root@ocfs2-host1 ~]# mkdir -p /opt/smp/disk1/
[root@ocfs2-host1 ~]# mount.ocfs2 /dev/sda /opt/smp/disk1/

[root@ocfs2-host2 ~]# mkdir -p /opt/smp/disk1/
[root@ocfs2-host2 ~]# mount.ocfs2 /dev/sda /opt/smp/disk1/

[root@ocfs2-host3 ~]# mkdir -p /opt/smp/disk1/
[root@ocfs2-host3 ~]# mount.ocfs2 /dev/sda /opt/smp/disk1/
```

至此 3 个节点已经可以共享访问/opt/smp/disk1/目录。同理，/dev/sdb 和/dev/sdc 也通

过此方式分别挂载到目录/opt/smp/disk2/和/opt/smp/disk3/。

挂载完毕，3 个节点的目录挂载情况：

```
[root@ocfs2-host1 ~]# df -Th
```

Filesystem	Type	Size	Used	Avail	Use%	Mounted on
devtmpfs	devtmpfs	1.8G	0	1.8G	0%	/dev
tmpfs	tmpfs	1.9G	0	1.9G	0%	/dev/shm
tmpfs	tmpfs	1.9G	8.4M	1.8G	1%	/run
tmpfs	tmpfs	1.9G	0	1.9G	0%	/sys/fs/cgroup
/dev/vda1	xfs	38G	3.4G	35G	9%	/
tmpfs	tmpfs	370M	0	370M	0%	/run/user/0
/dev/sda	ocfs2	20G	2.1G	18G	11%	/opt/smp/disk1
/dev/sdb	ocfs2	25G	2.2G	23G	9%	/opt/smp/disk2
/dev/sdc	ocfs2	30G	2.2G	28G	8%	/opt/smp/disk3

```
[root@ocfs2-host2 ~]# df -Th
```

Filesystem	Type	Size	Used	Avail	Use%	Mounted on
devtmpfs	devtmpfs	1.8G	0	1.8G	0%	/dev
tmpfs	tmpfs	1.9G	0	1.9G	0%	/dev/shm
tmpfs	tmpfs	1.9G	8.3M	1.8G	1%	/run
tmpfs	tmpfs	1.9G	0	1.9G	0%	/sys/fs/cgroup
/dev/vda1	xfs	38G	3.4G	35G	9%	/
tmpfs	tmpfs	370M	0	370M	0%	/run/user/0
/dev/sda	ocfs2	20G	2.1G	18G	11%	/opt/smp/disk1
/dev/sdb	ocfs2	25G	2.2G	23G	9%	/opt/smp/disk2
/dev/sdc	ocfs2	30G	2.2G	28G	8%	/opt/smp/disk3

```
[root@ocfs2-host3 ~]# df -Th
```

Filesystem	Type	Size	Used	Avail	Use%	Mounted on
devtmpfs	devtmpfs	1.8G	0	1.8G	0%	/dev
tmpfs	tmpfs	1.9G	0	1.9G	0%	/dev/shm
tmpfs	tmpfs	1.9G	8.4M	1.8G	1%	/run
tmpfs	tmpfs	1.9G	0	1.9G	0%	/sys/fs/cgroup
/dev/vda1	xfs	38G	3.4G	35G	9%	/
tmpfs	tmpfs	370M	0	370M	0%	/run/user/0
/dev/sda	ocfs2	20G	2.1G	18G	11%	/opt/smp/disk1
/dev/sdb	ocfs2	25G	2.2G	23G	9%	/opt/smp/disk2
/dev/sdc	ocfs2	30G	2.2G	28G	8%	/opt/smp/disk3

至此，OCFS2 集群创建完毕，可用存储目录/opt/smp/disk1、/opt/smp/disk2 和 /opt/smp/disk3。

为了实现服务器重启后可自动挂载，建议在/etc/rc.local 添加挂载项：

```
[root@ocfs2-host1 ~]# vim /etc/rc.local
...
sleep 10
mount.ocfs2 /dev/sda /opt/smp/disk1/
...
[root@ocfs2-host1 ~]# chmod +x /etc/rc.local
```

注意：由于服务器节点重启后，iSCSI 导出的裸设备盘符可能发生变化，强烈建议通过 /dev/disk/by-uuid/ 显示的 UUID 软连接挂载存储设备，以免挂载路径不匹配。

第四章 运维管理

4.1 添加裸设备

本节描述如何对线上的 OCFS2 集群增加裸设备。

管理员需在 iSCSI 存储创建一定容量大小的裸设备（LUN），例如下图：

Target Configuration

LUN Mapping

Network ACL

CHAP Authentication

LUNs mapped to target: iqn.2015-05.org.zstack:tsn.200001

LUN Id.	LUN Path	R/W Mode	SCSI Serial No.	SCSI Id.	Transfer Mode	Unmap LUN
0	/dev/volume-1/zstack-sto-1	write-thru	LzedfI-cZIO-qKej	LzedfI-cZIO-qKej	blockio	Unmap
1	/dev/volume-1/zstack-sto-2	write-thru	FVnSC7-xfj8-rwq3	FVnSC7-xfj8-rwq3	blockio	Unmap
2	/dev/volume-1/zstack-sto-3	write-thru	21IqYi-RghT-0AE6	21IqYi-RghT-0AE6	blockio	Unmap
3	/dev/volume-1/zstack-sto-4	write-thru	GgYkGT-GBUu-00J1	GgYkGT-GBUu-00J1	blockio	Unmap

图三 iSCSI 存储添加裸设备

通过以下方式在线扫描并发现磁盘：

```
[root@ocfs2-host1 ~]# echo "- - -" > /sys/class/scsi_host/host0/scan
[root@ocfs2-host1 ~]# echo "- - -" > /sys/class/scsi_host/host1/scan
[root@ocfs2-host1 ~]# echo "- - -" > /sys/class/scsi_host/host2/scan
[root@ocfs2-host1 ~]# echo "- - -" > /sys/class/scsi_host/host3/scan

# /dev/sdg 为新发现设备
[root@ocfs2-host1 ~]# ls SCSI -s
[1:0:0:0]   cd/dvd   QEMU      QEMU DVD-ROM    1.5.  /dev/sr0        -
[2:0:0:0]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sda        20.9GB
[2:0:0:1]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdb        26.2GB
[2:0:0:2]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdc        31.4GB
[2:0:0:3]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdg        41.9GB
```

以上操作需在 OCFS2 集群服务器执行，即 ocfs2-host1、ocfs2-host2 和 ocfs2-host3。

```
# /dev/sdg 未显示为 ocfs2 标记磁盘
[root@ocfs2-host1 ~]# mount.ocfs2 -d
```

Device	Stack	Cluster	F	UUID	Label
/dev/sda	o2cb	zstackstorage	G	9924FC4C94BE403BAB24BB6C5D4417FB	ocfs2-disk1
/dev/sdb	o2cb	zstackstorage	G	6311E734E8AC453C9E860C7A06E85DE8	ocfs2-disk1
/dev/sdc	o2cb	zstackstorage	G	D6E82EB382A54BD3BE84F101770F07D5	ocfs2-disk1

格式化新增加裸设备/dev/sdg：

```
# /dev/sdg 未显示为 ocfs2 标记磁盘
[root@ocfs2-host1 ~]# mkfs.ocfs2 --cluster-stack=o2cb -C 256K -J size=128M -N 16 -L
ocfs2-disk1 --cluster-name=zstackstorage --fs-feature-level=default -T vmstore /dev/sdg

# 执行挂载操作
[root@ocfs2-host1 ~]# partprobe
[root@ocfs2-host1 ~]# mkdir -p /opt/smp/disk4/
[root@ocfs2-host1 ~]# mount.ocfs2 /dev/disk/by-uuid/cfe295f-07f5-4c5d-8963-
e0f887c0c8dd /opt/smp/disk4/

[root@ocfs2-host2 ~]# partprobe
[root@ocfs2-host2 ~]# mkdir -p /opt/smp/disk4/
[root@ocfs2-host2 ~]# mount.ocfs2 /dev/disk/by-uuid/cfe295f-07f5-4c5d-8963-
e0f887c0c8dd /opt/smp/disk4/

[root@ocfs2-host3 ~]# partprobe
[root@ocfs2-host3 ~]# mkdir -p /opt/smp/disk4/
[root@ocfs2-host3 ~]# mount.ocfs2 /dev/disk/by-uuid/cfe295f-07f5-4c5d-8963-
e0f887c0c8dd /opt/smp/disk4/
```

至此，新裸设备/dev/sdg 已经添加到 OCFS2 集群，作为共享目录可读写。

4.2 移除裸设备

如果需要移除裸设备，管理员需确保该设备的挂载目录应用是否已经停止。

注意：当需要移除裸设备，确保其挂载目录上的云主机必须已经停止或数据已经迁移。

在 ZStack 管理界面上，针对该共享目录，需停止云主机并迁移数据，然后删除该共享目录。

管理员确认该共享目录已经从 ZStack 中删除后，执行以下移除操作：

```
[root@ocfs2-host1 ~]# umount /opt/smp/disk4/

[root@ocfs2-host2 ~]# umount /opt/smp/disk4/

[root@ocfs2-host3 ~]# umount /opt/smp/disk4/

# 从 SCSI 中移除/dev/sdg

[root@ocfs2-host1 ~]# ls SCSI -s
[1:0:0:0]   cd/dvd   QEMU      QEMU DVD-ROM    1.5.  /dev/sr0        -
[2:0:0:0]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sda      20.9GB
[2:0:0:1]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdb      26.2GB
[2:0:0:2]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdc      31.4GB
[2:0:0:3]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdg      41.9GB

# 移除 [2:0:0:3] SCSI 设备
[root@ocfs2-host1 ~]# echo "scsi remove-single-device 2 0 0 3" >/proc/scsi/scsi

[root@ocfs2-host2 ~]# ls SCSI -s
[1:0:0:0]   cd/dvd   QEMU      QEMU DVD-ROM    1.5.  /dev/sr0        -
[2:0:0:0]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sda      20.9GB
[2:0:0:1]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdb      26.2GB
[2:0:0:2]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdc      31.4GB
[2:0:0:3]   disk     OPNFILER  VIRTUAL-DISK    0      /dev/sdg      41.9GB

# 移除 [2:0:0:3] SCSI 设备
[root@ocfs2-host2 ~]# echo "scsi remove-single-device 2 0 0 3" >/proc/scsi/scsi
```

```
[root@ocfs2-host3 ~]# lsscsi -s
[1:0:0:0]    cd/dvd   QEMU      QEMU DVD-ROM    1.5.  /dev/sr0        -
[2:0:0:0]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sda        20.9GB
[2:0:0:1]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdb        26.2GB
[2:0:0:2]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdc        31.4GB
[2:0:0:3]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdg        41.9GB
[3:0:0:0]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdd        838MB
[3:0:0:1]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sde        973MB
[3:0:0:2]    disk    OPNFILER  VIRTUAL-DISK    0      /dev/sdf        1.07GB

# 移除 [2:0:0:3] SCSI 设备
[root@ocfs2-host3 ~]# echo "scsi remove-single-device 2 0 0 3" >/proc/scsi/scsi
```

至此，/dev/sdg 已经从 ocfs2-host1、ocfs2-host2 和 ocfs2-host3 中移除。管理员需在 iSCSI 存储移除/dev/sdg 对应的裸设备：

Target Configuration	LUN Mapping	Network ACL	CHAP Authentication			
LUNs mapped to target: iqn.2015-05.org.zstack:tsn.200001						
LUN Id.	LUN Path	R/W Mode	SCSI Serial No.	SCSI Id.	Transfer Mode	Unmap LUN
0	/dev/volume-1/zstack-sto-1	write-thru	LzedfI-cZI0-qKej	LzedfI-cZI0-qKej	blockio	<input type="button" value="Unmap"/>
1	/dev/volume-1/zstack-sto-2	write-thru	FVnSC7-xfj8-rwq3	FVnSC7-xfj8-rwq3	blockio	<input type="button" value="Unmap"/>
2	/dev/volume-1/zstack-sto-3	write-thru	21IqYi-RghT-0AE6	21IqYi-RghT-0AE6	blockio	<input type="button" value="Unmap"/>
Map New LUN to Target: "iqn.2015-05.org.zstack:tsn.200001"						
Name	LUN Path	R/W Mode	SCSI Serial No.	SCSI Id.	Transfer Mode	Map LUN
ocfs2-heartbeat-1	/dev/volume-1/ocfs2-heartbeat-1	<input type="button" value="write-thru"/>	RM9KJF-RinH-GwVN	RM9KJF-RinH-GwVN	<input type="button" value="blockio"/>	<input type="button" value="Map"/>
ocfs2-heartbeat-2	/dev/volume-1/ocfs2-heartbeat-2	<input type="button" value="write-thru"/>	TbfhHG-WceJ-1gpX	TbfhHG-WceJ-1gpX	<input type="button" value="blockio"/>	<input type="button" value="Map"/>
ocfs2-heartbeat-3	/dev/volume-1/ocfs2-heartbeat-3	<input type="button" value="write-thru"/>	Kottdy-Ge13-3YBm	Kottdy-Ge13-3YBm	<input type="button" value="blockio"/>	<input type="button" value="Map"/>
zstack-sto-4	/dev/volume-1/zstack-sto-4	<input type="button" value="write-thru"/>	GgYkGT-GBUu-00J1	GgYkGT-GBUu-00J1	<input type="button" value="blockio"/>	<input type="button" value="Map"/>

图四 iSCSI 存储取消裸设备映射

4.3 添加服务器

若 ZStack 所管理的物理资源不足，需要添加服务器，则需将服务器添加到 OCFS2 集群。

添加步骤如下：

```
[root@ocfs2-host1 ~]# vim /etc/hosts
...
172.20.11.221    ocfs2-host4
...

[root@ocfs2-host1 ~]# scp /etc/hosts ocfs2-host2:/etc/
[root@ocfs2-host1 ~]# scp /etc/hosts ocfs2-host3:/etc/
[root@ocfs2-host1 ~]# scp /etc/hosts ocfs2-host4:/etc/

# 设置主机名
[root@localhost ~]# hostnamectl set-hostname ocfs2-host4

# 安装 UEK 内核和 OCFS2 工具
[root@ocfs2-host4 ~]# yum --disablerepo=* --enablerepo=zstack-local,uek4-ocfs2 update

# 安装 UEK 内核和 OCFS2 工具
[root@ocfs2-host4 ~]# yum --disablerepo=* --enablerepo=zstack-local,uek4-ocfs2 \
install kernel-uek kernel-uek-devel kernel-uek-doc kernel-uek-firmware \
dtrace-modules ocfs2-tools ocfs2-tools-devel iscsi-initiator-utils \
device-mapper-multipath device-mapper-multipath-sysvinit

# 默认从 UEK 内核引导
[root@ocfs2-host4 ~]# grub2-set-default "CentOS Linux (4.1.12-37.2.2.el7uek.x86_64) 7
(Core)"

# 检查默认启动项
[root@ocfs2-host4 ~]# grub2-editenv list

# 配置防火墙，允许网络访问
[root@ocfs2-host4 ~]# vim /etc/sysconfig/iptables
...
-A INPUT -s 172.20.0.0/16 -j ACCEPT
...

[root@ocfs2-host4 ~]# service iptables restart

# 重启 ocfs2-host4 服务器
[root@ocfs2-host4 ~]# reboot
```

```
# 配置 ocfs2-host4 iSCSI 发起名
[root@ocfs2-host4 ~]# vim /etc/iscsi/initiatorname.iscsi
InitiatorName=iqn.2016-05.org.zstack:ocfs2-host4

# 配置 ocfs2-host4 iSCSI CHAP
[root@ocfs2-host1 ~]# vim /etc/iscsi/iscsid.conf
...
node.session.auth.authmethod = CHAP
node.session.auth.username = zstack
node.session.auth.password = password
...

# 扫描 iSCSI 存储裸设备
[root@ocfs2-host4 ~]# iscsiadm -m discovery -t sendtargets -p 172.20.12.221
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.200001
172.20.12.221:3260,1 iqn.2015-05.org.zstack:tsn.100001

# 登陆 iSCSI 存储设备
[root@ocfs2-host4 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.100001 --portal 172.20.12.221 --login
[root@ocfs2-host4 ~]# iscsiadm --mode node --targetname iqn.2015-05.org.zstack:tsn.200001 --portal 172.20.12.221 --login
[root@ocfs2-host4 ~]# mkdir -p /etc/ocfs2/
[root@ocfs2-host4 ~]# mkdir -p /d1m

# 增加 ocfs2-host4 到 OCFS2 集群配置文件, 在 ocfs2-host1 设置
[root@ocfs2-host1 ~]# o2cb add-node zstackstorage ocfs2-host4 --ip 172.20.11.221

# 同步配置文件到各节点
[root@ocfs2-host1 ~]# scp /etc/sysconfig/o2cb ocfs2-host4:/etc/sysconfig/

# 动态添加节点信息
[root@ocfs2-host1 ~]# o2cb_ctl -C -i -n ocfs2-host4 -t node -a number=3 -a ip_address=172.20.11.221 -a ip_port=7777 -a cluster=zstackstorage
[root@ocfs2-host2 ~]# o2cb_ctl -C -i -n ocfs2-host4 -t node -a number=3 -a ip_address=172.20.11.221 -a ip_port=7777 -a cluster=zstackstorage
[root@ocfs2-host3 ~]# o2cb_ctl -C -i -n ocfs2-host4 -t node -a number=3 -a ip_address=172.20.11.221 -a ip_port=7777 -a cluster=zstackstorage

# 复制集群配置文件
[root@ocfs2-host1 ~]# scp /etc/ocfs2/cluster.conf ocfs2-host4:/etc/ocfs2/
```

关于多路径简单配置：

```
[root@ocfs2-host1 ~]# modprobe dm-multipath
[root@ocfs2-host1 ~]# modprobe dm-round-robin

[root@ocfs2-host1 ~]# service multipathd start
[root@ocfs2-host1 ~]# mpathconf --enable

# 显示多路径信息
[root@ocfs2-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
```

启动 OCFS2 集群管理服务：

```
# ocfs2-host4 启动 o2cb 服务
[root@ocfs2-host4 ~]# /etc/init.d/o2cb start
checking debugfs...
Mounting configfs filesystem at /sys/kernel/config: OK
Loading stack plugin "o2cb": OK
Loading filesystem "ocfs2_dlmfs": OK
Mounting ocfs2_dlmfs filesystem at /dlm: OK
Setting cluster stack "o2cb": OK
Registering O2CB cluster "zstackstorage": OK
Setting O2CB cluster timeouts : OK
Starting local heartbeat for cluster "zstackstorage": OK

[root@ocfs2-host4 ~]# /etc/init.d/o2cb online
[root@ocfs2-host4 ~]# systemctl enable o2cb.service
[root@ocfs2-host4 ~]# systemctl enable ocfs2.service

# 创建 OCFS2 磁盘的挂载点
[root@ocfs2-host4 ~]# mkdir -p /opt/smp/disk1 /opt/smp/disk2 /opt/smp/disk3

# 挂载 OCFS2 磁盘
[root@ocfs2-host4 ~]# mount /dev/disk/by-uuid/9924fc4c-94be-403b-ab24-bb6c5d4417fb
/opt/smp/disk1/
[root@ocfs2-host4 ~]# mount /dev/disk/by-uuid/6311e734-e8ac-453c-9e86-0c7a06e85de8
/opt/smp/disk2/
[root@ocfs2-host4 ~]# mount /dev/disk/by-uuid/d6e82eb3-82a5-4bd3-be84-f101770f07d5
/opt/smp/disk3/
```

至此, ocfs2-host4 节点已经成功添加到 OCFS2 集群, 并挂载共享磁盘, 接下来可在 ZStack 界面上添加此服务器作为虚拟化的物理主机。

4.4 移除服务器

若需要移除某个物理主机, 管理员先在 ZStack 界面上, 把该物理主机设置为【维护模式】, 此时将会触发该物理主机上的云主机**动态迁移**。该物理主机上的云主机迁移完毕, 该物理主机则进入维护模式状态。此时, 管理员可在界面上【删除】该物理主机。

物理主机删除后, 管理员需要 SSH 进入该物理主机, 执行以下 OCFS2 集群移除步骤:

```
# 卸载共享目录
[root@ocfs2-host4 ~]# umount /opt/smp/disk1
[root@ocfs2-host4 ~]# umount /opt/smp/disk2
[root@ocfs2-host4 ~]# umount /opt/smp/disk3

# 关闭 ocfs2-host4 节点
[root@ocfs2-host4 ~]# halt -p

# 在 ocfs2-host1, ocfs2-host2 和 ocfs2-host3 上移除 ocfs2-host4 信息
[root@ocfs2-host1 ~]# o2cb remove-node zstackstorage ocfs2-host4
[root@ocfs2-host2 ~]# o2cb remove-node zstackstorage ocfs2-host4
[root@ocfs2-host3 ~]# o2cb remove-node zstackstorage ocfs2-host4
```

至此, ocfs2-host4 已经从 OCFS2 集群中移除。

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