[Top](http://tts.tmooc.cn/ttsPage/LINUX/NSDTN201904/PROJECT1/DAY03/CASE/01/index.html" \l "page_top_case)

# NSD Project1 DAY03

1. [案例1：Keepalived高可用](http://tts.tmooc.cn/ttsPage/LINUX/NSDTN201904/PROJECT1/DAY03/CASE/01/index.html" \l "case1)
2. [案例2：部署Ceph分布式存储](http://tts.tmooc.cn/ttsPage/LINUX/NSDTN201904/PROJECT1/DAY03/CASE/01/index.html" \l "case2)

## 1 案例1：Keepalived高可用

### 1.1 问题

部署两台代理服务器，实现如下效果：

* 利用keepalived实现两台代理服务器的高可用
* 配置VIP为192.168.4.80
* 修改对应的域名解析记录

### 1.2 方案

实验拓扑如图-1所示，做具体实验前请先配置好环境。

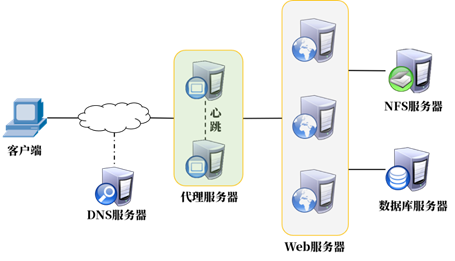
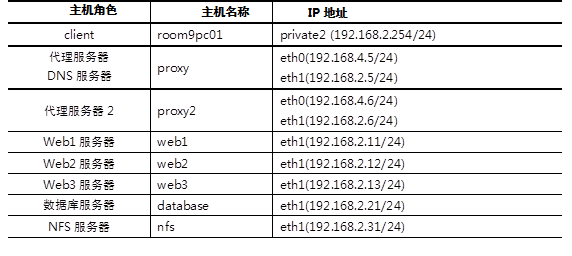


图-1

备注：实际操作中DNS服务代理服务器部署在同一台主机上（节约虚拟机资源）。

主机配置如表-1所示。

表-1



### 1.3 步骤

实现此案例需要按照如下步骤进行。

步骤一：配置第二台代理服务器

1）部署HAProxy

安装软件，手动修改配置文件，添加如下内容。

1. [root@proxy2 ~]# yum -y install haproxy
2. [root@proxy2 ~]# vim /etc/haproxy/haproxy.cfg
3. listen wordpress \*:80
4. balance roundrobin
5. server web1 192.168.2.11:80 check inter 2000 rise 2 fall 3
6. server web2 192.168.2.12:80 check inter 2000 rise 2 fall 3
7. server web3 192.168.2.13:80 check inter 2000 rise 2 fall 3
8. [root@proxy2 ~]# systemctl start haproxy
9. [root@proxy2 ~]# systemctl enable haproxy

步骤二：为两台代理服务器配置keepalived

1）配置第一台代理服务器proxy（192.168.4.5）。

1. [root@proxy ~]# yum install -y keepalived
2. [root@proxy ~]# vim /etc/keepalived/keepalived.conf
3. global\_defs {
4. router\_id proxy1                        //设置路由ID号
5. vrrp\_iptables //不添加任何防火墙规则
6. }
7. vrrp\_instance VI\_1 {
8. state MASTER                      //主服务器为MASTER（备服务器需要修改为BACKUP）
9. interface eth0                    //定义网络接口
10. virtual\_router\_id 51
11. priority 100                 //服务器优先级,优先级高优先获取VIP（实验需要修改）
12. advert\_int 1
13. authentication {
14. auth\_type pass
15. auth\_pass 1111                      //主备服务器密码必须一致
16. }
17. virtual\_ipaddress { //谁是主服务器谁获得该VIP（实验需要修改）
18. 192.168.4.80
19. }
20. }
21. [root@proxy ~]# systemctl start keepalived

！！！重要！！！

在全局配置global\_defs{}中手动添加vrrp\_iptables，即可解决防火墙的问题。

2）配置第二台代理服务器proxy（192.168.4.6）。

1. [root@proxy2 ~]# yum install -y keepalived
2. [root@proxy2 ~]# vim /etc/keepalived/keepalived.conf
3. global\_defs {
4. router\_id proxy2                        //设置路由ID号
5. vrrp\_iptables //不添加任何防火墙规则
6. }
7. vrrp\_instance VI\_1 {
8. state BACKUP                      //主服务器为MASTER（备服务器需要修改为BACKUP）
9. interface eth0                    //定义网络接口
10. virtual\_router\_id 51
11. priority 50                     //服务器优先级,优先级高优先获取VIP
12. advert\_int 1
13. authentication {
14. auth\_type pass
15. auth\_pass 1111                      //主备服务器密码必须一致
16. }
17. virtual\_ipaddress { //谁是主服务器谁获得该VIP
18. 192.168.4.80
19. }
20. }
21. [root@proxy2 ~]# systemctl start keepalived

！！！重要！！！

在全局配置global\_defs{}中手动添加vrrp\_iptables，即可解决防火墙的问题。

步骤三：修改DNS服务器

1）修改网站域名对应的解析记录，解析到新的VIP地址。

192.168.4.5为DNS服务器。

1. [root@proxy ~]# vim /var/named/lab.com.zone
2. $TTL 1D
3. @ IN SOA @ rname.invalid. (
4. 0 ; serial
5. 1D ; refresh
6. 1H ; retry
7. 1W ; expire
8. 3H ) ; minimum
9. @ NS dns.lab.com.
10. dns A 192.168.4.5
11. www A 192.168.4.80

2）重启DNS服务

1. [root@proxy ~]# systemctl restart named

## 2 案例2：部署Ceph分布式存储

### 2.1 问题

部署Ceph分布式存储，实现如下效果：

* 使用三台服务器部署Ceph分布式存储
* 实现Ceph文件系统共享
* 将网站数据从NFS迁移到Ceph存储

### 2.2 方案

实验拓扑如图-2所示，做具体实验前请先配置好环境。

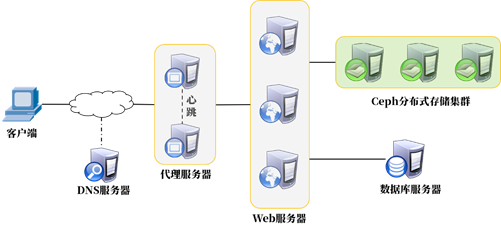


图-2

备注：实际操作中DNS服务代理服务器部署在同一台主机上（节约虚拟机资源）。

主机配置如表-2所示。

表-2



### 2.3 步骤

实现此案例需要按照如下步骤进行。

步骤一：准备实验环境

1）物理机为所有节点配置yum源服务器。

提示：ceph10.iso在/linux-soft/02目录。

1. [root@room9pc01 ~]# mkdir /var/ftp/ceph
2. [root@room9pc01 ~]# mount ceph10.iso /var/ftp/ceph/

2）在node1配置SSH密钥，让node1可用无密码连接node1,node2,node3

1. [root@node1 ~]# ssh-keygen -f /root/.ssh/id\_rsa -N ''
2. [root@node1 ~]# for i in 41 42 43
3. do
4. ssh-copy-id 192.168.2.$i
5. done

3)修改/etc/hosts域名解析记录（不要删除原有的数据），同步给所有ceph节点。

1. [root@node1 ~]# vim /etc/hosts
2. 192.168.2.41    node1
3. 192.168.2.42     node2
4. 192.168.2.43    node3
5. [root@node1 ~]# for i in 41 42 43
6. do
7. scp /etc/hosts 192.168.2.$i:/etc
8. done

4）为所有ceph节点配置yum源，并将配置同步给所有节点

1. [root@node1 ~]# cat /etc/yum.repos.d/ceph.repo
2. [mon]
3. name=mon
4. baseurl=ftp://192.168.2.254/ceph/MON
5. gpgcheck=0
6. [osd]
7. name=osd
8. baseurl=ftp://192.168.2.254/ceph/OSD
9. gpgcheck=0
10. [tools]
11. name=tools
12. baseurl=ftp://192.168.2.254/ceph/Tools
13. gpgcheck=0
14. [root@node1 ~]# yum repolist                #验证YUM源软件数量
15. 源标识            源名称                    状态
16. Dvd                redhat                    9,911
17. Mon                mon                        41
18. Osd                osd                        28
19. Tools            tools                    33
20. repolist: 10,013
21. [root@node1 ~]# for i in 41 42 43
22. do
23. scp /etc/yum.repos.d/ceph.repo 192.168.2.$i:/etc/yum.repos.d/
24. done

5）所有节点主机与真实主机的NTP服务器同步时间。

提示：默认真实物理机已经配置为NTP服务器。

1. [root@node1 ~]# vim /etc/chrony.conf
2. … …
3. server 192.168.2.254 iburst
4. [root@node1 ~]# for i in 41 42 43
5. do
6. scp /etc/chrony.conf 192.168.2.$i:/etc/
7. ssh 192.168.2.$i "systemctl restart chronyd"
8. done

6）使用virt-manager为三台ceph虚拟机添加磁盘。

每台虚拟机添加2块20G的磁盘。

步骤二：部署ceph集群

1）给node1主机安装ceph-deploy，创建工作目录，初始化配置文件。

1. [root@node1 ~]# yum -y install ceph-deploy
2. [root@node1 ~]# mkdir ceph-cluster
3. [root@node1 ~]# cd ceph-cluster

2）给所有ceph节点安装ceph相关软件包

1. [root@node1 ceph-cluster]# for i in node1 node2 node3
2. do
3. ssh $i "yum -y install ceph-mon ceph-osd ceph-mds"
4. done

3）初始化mon服务

1. [root@node1 ceph-cluster]# ceph-deploy new node1 node2 node3
2. [root@node1 ceph-cluster]# ceph-deploy mon create-initial
3. [root@node1 ceph-cluster]# ceph -s                    #查看结果
4. cluster 9f3e04b8-7dbb-43da-abe6-b9e3f5e46d2e
5. health HEALTH\_ERR
6. monmap e2: 3 mons at
7. {node1=192.168.2.41:6789/0,node2=192.168.2.42:6789/0,node3=192.168.2.43:6789/0}
9. osdmap e45: 0 osds: 0 up, 0 in

4）准备磁盘分区，创建journal盘，并永久修改设备权限。

1. [root@node1 ceph-cluster]# for i in node1 node2 node3
2. do
3. ssh $i "parted /dev/vdb mklabel gpt"
4. ssh $i "parted /dev/vdb mkpart primary 1 100%"
5. done

提示：下面的步骤在所有主机都需要操作（node1，node2，node3）

1. #临时修改权限：
2. [root@node1 ceph-cluster]# chown ceph.ceph /dev/vdb1
3. #永久修改权限：
4. [root@node1 ceph-cluster]# vim /etc/udev/rules.d/70-vdb.rules
5. ENV{DEVNAME}=="/dev/vdb1",OWNER="ceph",GROUP="ceph"

4）使用ceph-deploy工具初始化数据磁盘（仅node1操作）。

1. [root@node1 ceph-cluster]# ceph-deploy disk zap node1:vdc
2. [root@node1 ceph-cluster]# ceph-deploy disk zap node2:vdc
3. [root@node1 ceph-cluster]# ceph-deploy disk zap node3:vdc

5）初始化OSD集群。

1. [root@node1 ceph-cluster]# ceph-deploy osd create \
2. node1:vdc:/dev/vdb1
3. //创建osd存储设备，vdc为集群提供存储空间，vdb1提供JOURNAL缓存，
4. //一个存储设备对应一个缓存设备，缓存需要SSD，不需要很大
5. [root@node1 ceph-cluster]# ceph-deploy osd create \
6. node2:vdc:/dev/vdb1
7. [root@node1 ceph-cluster]# ceph-deploy osd create \
8. node3:vdc:/dev/vdb1
9. [root@node1 ceph-cluster]# ceph -s                 #查看集群状态
10. cluster 9f3e04b8-7dbb-43da-abe6-b9e3f5e46d2e
11. health HEALTH\_OK
12. monmap e2: 3 mons at {node1=192.168.4.11:6789/0,node2=192.168.4.12:6789/0,node3=192.168.4.13:6789/0}
13. election epoch 6, quorum 0,1,2 node1,node2,node3
14. osdmap e45: 3 osds: 3 up, 3 in
15. flags sortbitwise
16. pgmap v25712: 64 pgs, 1 pools, 86465 kB data, 2612 objects
17. 508 MB used, 60 GB / 60 GB avail
18. 64 active+clean

步骤三：部署ceph文件系统

1）启动mds服务

1. [root@node1 ceph-cluster]# ceph-deploy mds create node3

2）创建存储池（文件系统由inode和block组成）

1. [root@node1 ceph-cluster]# ceph osd pool create cephfs\_data 128
2. [root@node1 ceph-cluster]# ceph osd pool create cephfs\_metadata 128
3. [root@node1 ceph-cluster]# ceph osd lspools
4. 0 rbd,1 cephfs\_data,2 cephfs\_metadata

3）创建文件系统

1. [root@node1 ceph-cluster]# ceph fs new myfs1 cephfs\_metadata cephfs\_data
2. [root@node1 ceph-cluster]# ceph fs ls
3. name: myfs1, metadata pool: cephfs\_metadata, data pools: [cephfs\_data ]

步骤四：迁移网站数据到ceph集群

1）卸载web1，web2，web3的NFS共享。

暂停服务防止有人实时读写文件。

1. [root@web1 ~]# /usr/local/nginx/sbin/nginx -s stop
2. [root@web2 ~]# /usr/local/nginx/sbin/nginx -s stop
3. [root@web3 ~]# /usr/local/nginx/sbin/nginx -s stop
4. [root@web1 ~]# umount /usr/local/nginx/html
5. [root@web2 ~]# umount /usr/local/nginx/html
6. [root@web3 ~]# umount /usr/local/nginx/html
7. [root@web1 ~]# vim /etc/fstab
8. #192.168.2.31:/web\_share/html /usr/local/nginx/html/ nfs defaults 0 0
9. [root@web2 ~]# vim /etc/fstab
10. #192.168.2.31:/web\_share/html /usr/local/nginx/html/ nfs defaults 0 0
11. [root@web3 ~]# vim /etc/fstab
12. #192.168.2.31:/web\_share/html /usr/local/nginx/html/ nfs defaults 0 0

2）web服务器永久挂载Ceph文件系统（web1、web2、web3都需要操作）。

在任意ceph节点，如node1查看ceph账户与密码。

1. [root@node1 ~]# cat /etc/ceph/ceph.client.admin.keyring
2. [client.admin]
3. key = AQA0KtlcRGz5JxAA/K0AD/uNuLI1RqPsNGC7zg==

/etc/rc.local是开机启动脚本，任何命令放在该文件中都是开机自启。

1. [root@web1 ~]# mount -t ceph 192.168.2.41:6789:/ /usr/local/nginx/html/ \
2. -o name=admin,secret=AQA0KtlcRGz5JxAA/K0AD/uNuLI1RqPsNGC7zg==
3. [root@web1 ~]# echo 'mount -t ceph 192.168.2.41:6789:/ /usr/local/nginx/html/ \
4. -o name=admin,secret=AQA0KtlcRGz5JxAA/K0AD/uNuLI1RqPsNGC7zg==' >> /etc/rc.local
5. [root@web1 ~]# chmod +x /etc/rc.local
6. [root@web2 ~]# mount -t ceph 192.168.2.41:6789:/ /usr/local/nginx/html/ \
7. -o name=admin,secret=AQA0KtlcRGz5JxAA/K0AD/uNuLI1RqPsNGC7zg==
8. [root@web2 ~]# echo 'mount -t ceph 192.168.2.41:6789:/ /usr/local/nginx/html/ \
9. -o name=admin,secret=AQA0KtlcRGz5JxAA/K0AD/uNuLI1RqPsNGC7zg==' >> /etc/rc.local
10. [root@web2 ~]# chmod +x /etc/rc.local
11. [root@web3 ~]# mount -t ceph 192.168.2.41:6789:/ /usr/local/nginx/html/ \
12. -o name=admin,secret=AQA0KtlcRGz5JxAA/K0AD/uNuLI1RqPsNGC7zg==
13. [root@web3 ~]# echo 'mount -t ceph 192.168.2.41:6789:/ /usr/local/nginx/html/ \
14. -o name=admin,secret=AQA0KtlcRGz5JxAA/K0AD/uNuLI1RqPsNGC7zg==' >> /etc/rc.local
15. [root@web3 ~]# chmod +x /etc/rc.local

另一种解决方案，还可以通过fstab实现永久挂载。

提示：如果希望使用fstab实现永久挂载，客户端需要额外安装libcephfs1软件包。

1. [root@web1 ~]# yum -y install libcephfs1
2. [root@web1 ~]# vim /etc/fstab
3. … …
4. 192.168.4.11:/ /usr/local/nginx/html/    ceph defaults,\_netdev,name=admin,secret=AQCVcu9cWXkgKhAAWSa7qCFnFVbNCTB2DwGIOA== 0 0

第三种挂载方案：对于高可用的问题，可以在mount时同时写入多个IP。

1. 临时命令：
2. [root@web1 ~]# mount -t ceph \
3. 192.168.4.11:6789,192.168.4.12:6789,192.168.4.13:6789:/ /usr/local/nginx/html \
4. -o name=admin,secret=密钥
5. 永久修改：
6. [root@web1 ~]# vim /etc/fstab
7. 192.168.4.11:6789,192.168.4.12:6789,192.168.4.13:6789:/ /usr/local/nginx/html/ \
8. ceph defaults,\_netdev,name=admin,secret=密钥 0 0

3)迁移NFS服务器中的数据到Ceph存储

登陆NFS服务器备份数据，将备份数据拷贝给web1或web2或web3，tar备份数据时注意使用-f选项保留文件权限。

1. [root@nfs ~]# cd /web\_share/html/
2. [root@nfs html]# tar -czpf /root/html.tar.gz ./\*
3. [[root@nfs](mailto:root@nfs) html]# scp /root/html.tar.gz 192.168.2.11:/usr/local/nginx/html/

登陆web1将数据恢复到Ceph共享目录

1. [root@web1 html]# tar -xf html.tar.gz
2. [root@web1 html]# rm -rf html.tar.gz

4）恢复web服务

1. [root@web1 ~]# /usr/local/nginx/sbin/nginx
2. [root@web2 ~]# /usr/local/nginx/sbin/nginx
3. [root@web3 ~]# /usr/local/nginx/sbin/nginx

附加知识（常见面试题）

1) 如何使用awk查看TCP连接状态？

答：ss -ant |awk '{print $1}'

netstat -ant |awk '{print $6}'

2) 有个txt文件内容如下：

http://a.domain.com/l.html

http://b.domain.com/l.html

http://c.domain.com/l.html

http://a.domain.com/2.html

http://b.domain.com/2.html

http://a.domain.com/3.html

要求：编写脚本获取主机名域名，并统计每个网址出现的次数，并排序。

答：

1. #！/bin/bash
2. awk -F"[/.]" '{print $3}' txt #单独获取主机名
3. awk -F"[/]" '{print $3}' txt #获取完整域名
4. awk -F"[/]" '{IP[$3]++} END{for(i in IP){print IP[i],i}}' txt | sort -n

3) 至少说出一种linux下实现高可用的方案名称？

答：keepalived，HeartBeat

4）简述下负载均衡与高可用的概念？

答：

LB（Load\_balancing）: 多台服务器平均响应客户端的多次连接请求。

HA（High Availability）: 主备模式，主服务器宕机后，备用服务器才接替工作。

5）列举几种你知道的LVS调度算法？

答：

轮询（Round Robin）

加权轮询（Weighted Round Robin）

最少连接（Least Connections）

加权最少连接（ Weighted Least Connections ）

源地址哈希值（source hash）

6）如果你们公司的网站访问很慢，你会如何排查？

答：

查看流量(Zabbix,ifconfig,sar,ping延迟… …)

系统负载(Zabbix,uptime,sar,top,ps,free查看CPU和内存)

日志（数据库日志-慢查询日志、web服务器日志、ELK）

DNS解析；ss端口状态、并发量；本机时间（时间错误会导致服务器故障）

浏览器F12（开发者工具）

7）你会用什么方法来查看某个应用服务的流量使用情况?

答：

ifconfig eth0（查看网卡整体流量）

iftop（需要安装iftop软件包，实时查看具体IP、端口的流量，iftop -P）

iptraf-ng (需要安装iptraf-ng软件包，实时查看IP、端口的流量)

sar -n DEV（需要安装sysstat软件包，查看历史网卡流量，或者实时查看流量）

nethogs eth0（需要安装nethogs软件包，实时查看进程流量）

查看网站的访问日志（利用awk统计资源的大小并求和）

通过zabbix查看软件流量