使用OpenFiler模拟存储配置RAC中ASM共享盘及多路径（multipath）的测试

**第一章 本篇总览**

**本文介绍了OpenFiler、iSCSI和多路径的配置。**

**第二章 安装OpenFiler**

OpenFile是在rPath Linux基础上开发的，它能够作为一个独立的Linux操作系统发行。Openfiler是一款非常好的存储管理操作系统，开源免费，通过web界面对存储磁盘的管理，支持现在流行的网络存储技术IP-SAN和NAS，支持iSCSI、NFS、SMB/CIFS及FTP等协议。

本次安装OpenFiler锁需要的软件如下所示：

|  |  |  |
| --- | --- | --- |
| **序号** | **类型** | **内容** |
| 1 | openfiler | openfileresa-2.99.1-x86\_64-disc1.iso |

**2.1  安装**

OpenFiler的内存设置为1G大小或再小点也无所谓，磁盘选用IDE磁盘格式，由于后续要配置多路径，所以需要安装2块网卡。安装完成后，重新启动，界面如下所示：

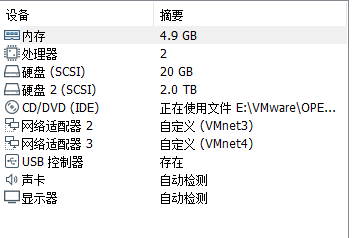


注意，方框中的内容，可以在浏览器中直接打开。可以用root用户登录进行用户的维护，若进行存储的维护则只能使用openfiler用户。openfiler是在远程使用Web界面进行管理的，小麦苗这里的管理地址是https://192.168.70.100:446，其管理的初始用户名是**openfiler**（小写的），密码是**password**，可以在登录之后，改这个密码。



**2.2  基本配置**

**2.2.1  网卡配置**



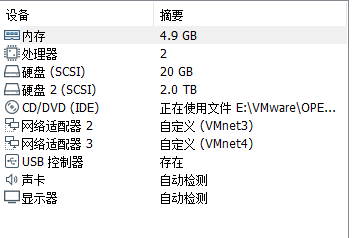
PS：网卡1和网卡2分别为70、71网段做为RAC接点连接服务器用，实现多路径。

查看网卡信息：

|  |
| --- |
| [root@localhost ~]# ip add  1: lo: <LOOPBACK,UP,10000> mtu 16436 qdisc noqueue  link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00  inet 127.0.0.1/8 scope host lo  inet6 ::1/128 scope host  valid\_lft forever preferred\_lft forever  2: eth1: <BROADCAST,MULTICAST,UP,10000> mtu 1500 qdisc pfifo\_fast qlen 1000  link/ether 00:0c:29:42:fc:45 brd ff:ff:ff:ff:ff:ff  inet 192.168.70.100/24 brd 192.168.70.255 scope global eth1  inet6 fe80::20c:29ff:fe42:fc45/64 scope link  valid\_lft forever preferred\_lft forever  3: eth2: <BROADCAST,MULTICAST,UP,10000> mtu 1500 qdisc pfifo\_fast qlen 1000  link/ether 00:0c:29:42:fc:4f brd ff:ff:ff:ff:ff:ff  inet 192.168.71.100/24 brd 192.168.71.255 scope global eth2  inet6 fe80::20c:29ff:fe42:fc4f/64 scope link  valid\_lft forever preferred\_lft forever |

**2.2.2  添加硬盘**

添加一块100G大小的IDE格式的硬盘作为存储。



 PS：配置2T硬盘，实际应用可调整。

**2.3  iscsi target配置**

为openfiler服务器配置了两块硬盘，其中10GB的硬盘已经用来安装openfiler操作系统，而200GB的硬盘则会用做数据存储。

**2.3.1  创建逻辑卷**

登录地址：**https://192.168.70.100:446**

初始用户名和密码：**openfiler**/**password**

在独立存储设备中，LUN（Logical Unit Number）是最重要的基本单位。LUN可以被SAN中的任何主机访问，不管是透过HBA或是iSCSI。就算是软件激活的iSCSI，也可以在不同的操作系统之下，在操作系统启动之后利用软件的iSCSI initiator访问LUN。在OpenFiler之下，LUN被称为Logical Volume（LV），因此在OpenFiler下创建LUN就是创建LV。

当你安装好OpenFiler之后，接下来就是要将OpenFiler下的磁盘分享出来给虚拟机或网络上的其他主机使用了。在标准的SAN之后，这些可以在RAID层面完成，但VG的好处及弹性是RAID无法比较的，下面看看OpenFiler下的VG是如何一步一步创建的。

 创建VG的步骤：

（1）进入OpenFiler的接口，并且选择要使用的实体硬盘。

（2）将要加入的实体硬盘格式化成Physical Volume格式。

（3）创建一个VG组，并且将格式化成为PV格式的实体硬盘加入。

（4）加入完毕之后，就成为一个大的VG组，被视为系统的一个大实体硬盘。

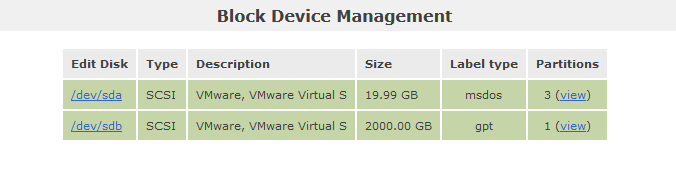
（5）在这个VG中添加逻辑分割区LUN，在OpenFiler中称为Logical Volume。

（6）指定LUN的文件格式，如iSCSI、ext3或是NFS，并且格式化。

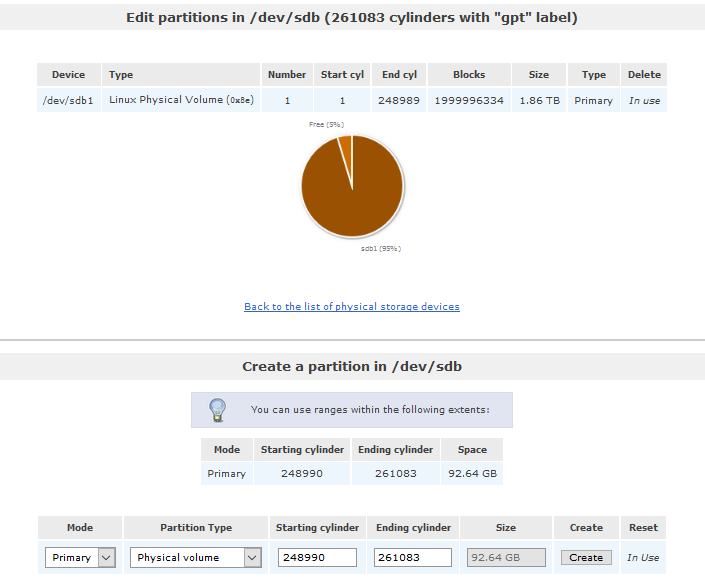
（7）如果是iSCSI则需要再配置，如果是其他文件格式，就可以用NAS的方式分享出去而

登录后，点击Volumes标签

为openfiler服务器配置了两块硬盘，其中20G的硬盘已经用来安装openfiler操作系统，而2T硬盘则会用做数据存储。



点击create new physical volumes后点击/dev/sdb

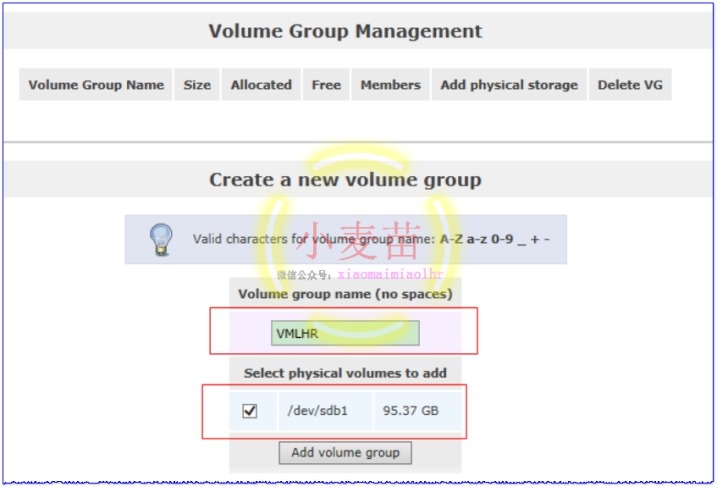


点击页面右下角Reset，然后点击Create。分区类型为Physical volume

[](http://images2015.cnblogs.com/blog/646850/201701/646850-20170123231412456-1207194035.jpg)

点击Volume Groups

[](http://images2015.cnblogs.com/blog/646850/201701/646850-20170123231415003-1692570007.jpg)

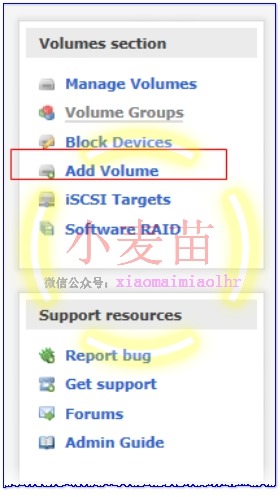
[](http://images2015.cnblogs.com/blog/646850/201701/646850-20170123231417206-1877236545.jpg)

输入名称，勾选复选框，单击Add volume group

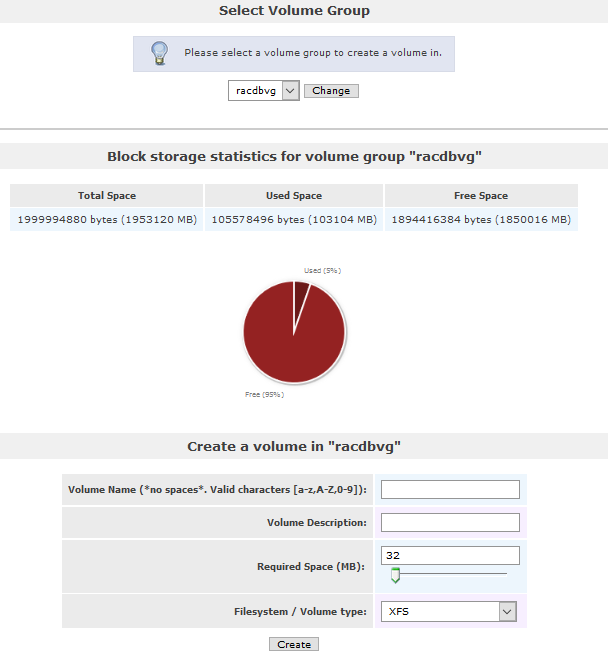
[](http://images2015.cnblogs.com/blog/646850/201701/646850-20170123231419409-1147492729.jpg)

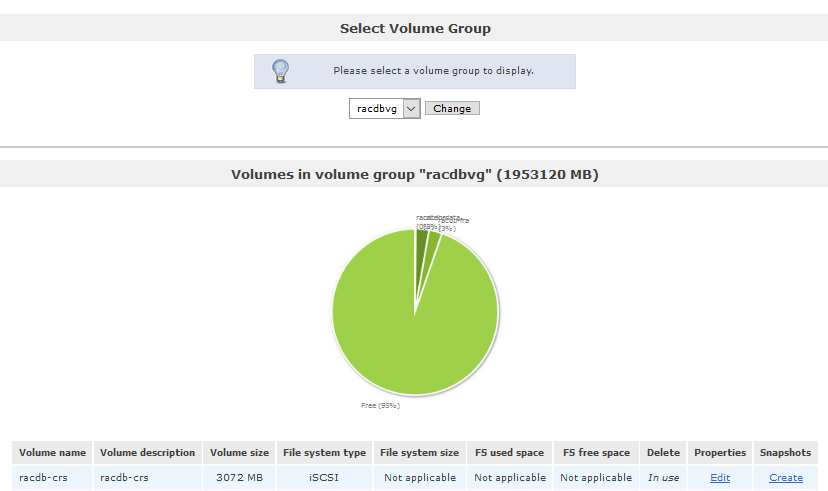
|  |
| --- |
| [root@OFLHR ~]# fdisk -l    Disk /dev/sda: 10.7 GB, 10737418240 bytes  255 heads, 63 sectors/track, 1305 cylinders, total 20971520 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk identifier: 0x000adc2c       Device Boot      Start         End      Blocks   Id  System  /dev/sda1   \*          63      610469      305203+  83  Linux  /dev/sda2          610470    17382329     8385930   83  Linux  /dev/sda3        17382330    19486844     1052257+  82  Linux swap / Solaris    WARNING: GPT (GUID Partition Table) detected on '/dev/sdb'! The util fdisk doesn't support GPT. Use GNU Parted.      Disk /dev/sdb: 107.4 GB, 107374182400 bytes  255 heads, 63 sectors/track, 13054 cylinders, total 209715200 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk identifier: 0x00000000       Device Boot      Start         End      Blocks   Id  System  /dev/sdb1               1   209715199   104857599+  ee  GPT  [root@OFLHR ~]# pvs    PV         VG    Fmt  Attr PSize  PFree    /dev/sdb1  vmlhr lvm2 a-   95.34g 95.34g  [root@OFLHR ~]# |

点击Add Volume

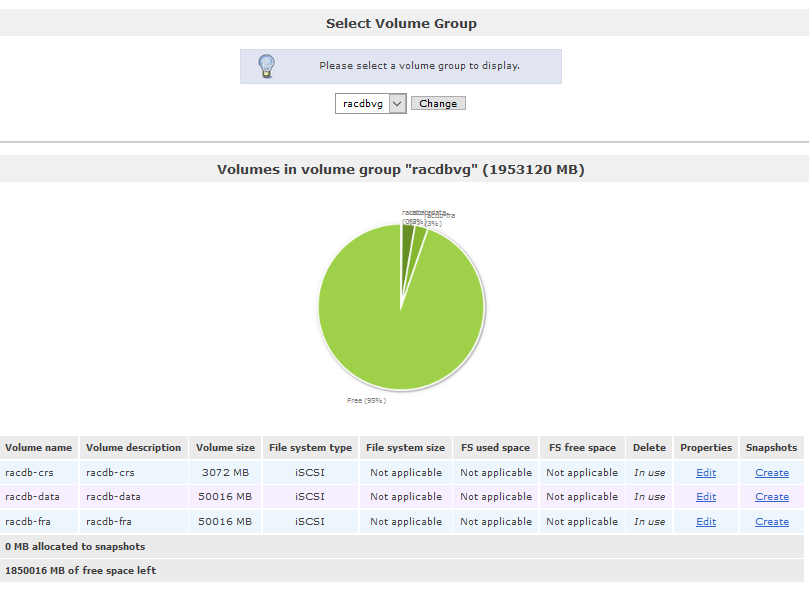
[](http://images2015.cnblogs.com/blog/646850/201701/646850-20170123231422253-1341569222.jpg)

输入内容，调整磁盘大小为10G，卷类型选择block（iSCSI，FC，etc）



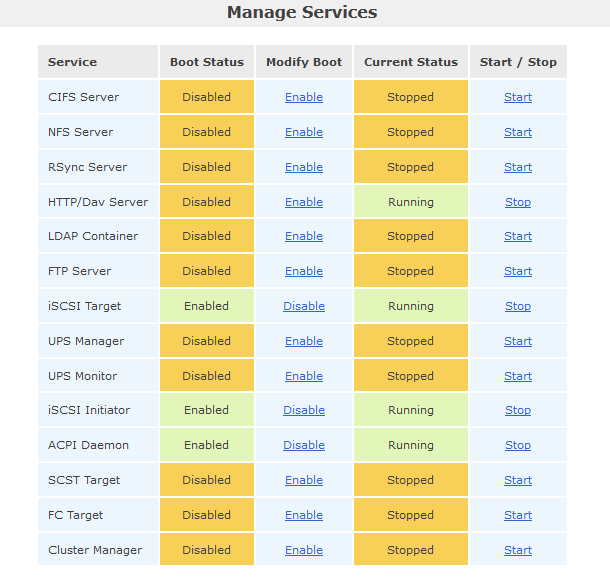


依次共创建3个逻辑卷：



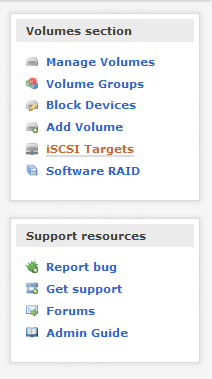
|  |
| --- |
| [root@localhost ~]# **vgs**  VG #PV #LV #SN Attr VSize VFree  racdbvg 1 3 0 wz--n- 1.86t 1.76t  [root@localhost ~]# **pvs**  PV VG Fmt Attr PSize PFree  /dev/sdb1 racdbvg lvm2 a- 1.86t 1.76t  [root@localhost ~]# **lvs**  LV VG Attr LSize Origin Snap% Move Log Copy% Convert  racdb-crs racdbvg -wi-ao 3.00g  racdb-data racdbvg -wi-ao 48.84g  racdb-fra racdbvg -wi-ao 48.84g  [root@localhost ~]# **fdisk -l**  Disk /dev/sda: 21.5 GB, 21474836480 bytes  255 heads, 63 sectors/track, 2610 cylinders, total 41943040 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk identifier: 0x000239ec  Device Boot Start End Blocks Id System  /dev/sda1 \* 63 610469 305203+ 83 Linux  /dev/sda2 610470 17382329 8385930 83 Linux  /dev/sda3 17382330 19486844 1052257+ 82 Linux swap / Solaris  WARNING: GPT (GUID Partition Table) detected on '/dev/sdb'! The util fdisk doesn't support GPT. Use GNU Parted.  Disk /dev/sdb: 2147.5 GB, 2147483648000 bytes  255 heads, 63 sectors/track, 261083 cylinders, total 4194304000 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk identifier: 0x00000000  Device Boot Start End Blocks Id System  /dev/sdb1 1 4194303999 2097151999+ ee GPT  Disk /dev/dm-0: 3221 MB, 3221225472 bytes  255 heads, 63 sectors/track, 391 cylinders, total 6291456 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk identifier: 0x644c8e40  Device Boot Start End Blocks Id System  /dev/dm-0p1 63 2088449 1044193+ 83 Linux  /dev/dm-0p2 2088450 4176899 1044225 83 Linux  /dev/dm-0p3 4176900 6281414 1052257+ 83 Linux  Disk /dev/dm-1: 52.4 GB, 52445577216 bytes  255 heads, 63 sectors/track, 6376 cylinders, total 102432768 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk identifier: 0x00000000  Disk /dev/dm-1 doesn't contain a valid partition table  Disk /dev/dm-2: 52.4 GB, 52445577216 bytes  255 heads, 63 sectors/track, 6376 cylinders, total 102432768 sectors  Units = sectors of 1 \* 512 = 512 bytes  Sector size (logical/physical): 512 bytes / 512 bytes  I/O size (minimum/optimal): 512 bytes / 512 bytes  Disk identifier: 0x00000000  Disk /dev/dm-2 doesn't contain a valid partition table |

**2.3.2  开启iSCSI Target服务**

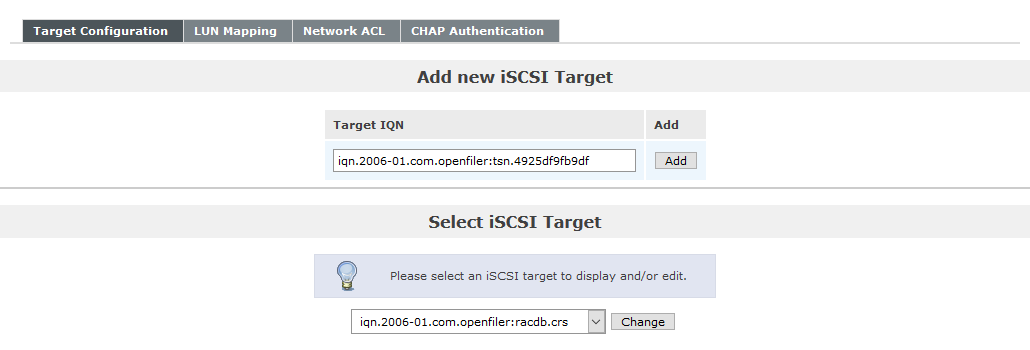


点击Services标签栏设置iSCSI Target 为Enable 开启服务Start。

**2.3.3  LUN Mapping操作**

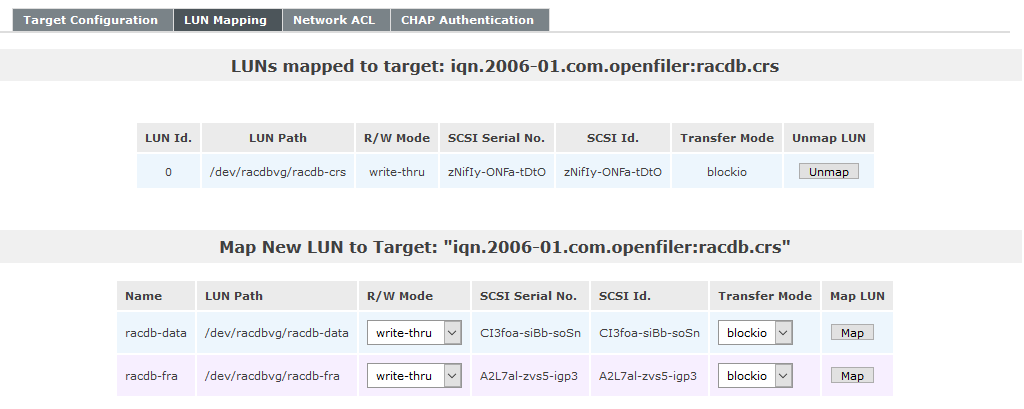


返回Volumes标签页，点击iSCSI Targets



点击Add

选择LUN Mapping标签 点击Map



**2.3.4  Network ACL**

由于iSCSI是走IP网络，因此我们要允许网络中的计算机可以透过IP来访问。下面就是OpenFiler中IP网络和同一网段中其他主机的连接方法。

1．进入OpenFiler中的System，并且直接拉到页面的下方。

2．在Network Access Configuration的地方输入这个网络访问的名称，如VM\_LHR。

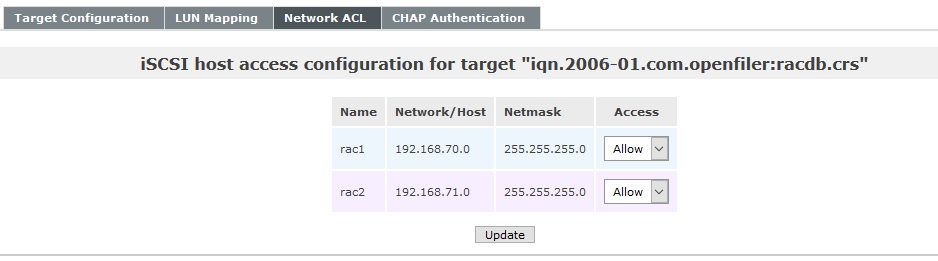
3．输入主机的IP段。注意不可以输入单一主机的IP，这样会都无法访问。我们在这边输入192.168.70.0/192.168.71.0，表示从192.168.70(71).1一直到192.168.70(71).254都能访问。

4．在Netmask中选择255.255.255.0，并且在Type下拉列表框中选择Share，之后即可以单击Update按钮。

选择完之后就更新

至此就可以在这个OpenFiler中看到被授权的网段了。

在iSCSI Targets中，点击 Network ACL 标签



设置Access为Allow 然后点击Update

到此存储的配置已经完成

**2.3.5  /etc/initiators.deny**

注释掉iqn.2006-01.com.openfiler: racdb.crs ALL：

|  |
| --- |
| [root@localhost ~]# vi /etc/initiators.deny  # PLEASE DO NOT MODIFY THIS CONFIGURATION FILE!  # This configuration file was autogenerated  # by Openfiler. Any manual changes will be overwritten  # Generated at: Sat Apr 20 0:40:11 EDT 2019  iqn.2006-01.com.openfiler:racdb.crs ALL  iqn.2006-01.com.openfiler:racdb.data ALL  iqn.2006-01.com.openfiler:racdb.fra ALL  # End of Openfiler configuration |

**第三章 RAC中配置共享**

**3.1  RAC节点配置iSCSI**

iSCSI（Internet Small Computer System Interface）。iSCSI技术由IBM公司研究开发，是一个供硬件设备使用的、可以在IP协议的上层运行的SCSI指令集，这种指令集合可以实现在IP网络上运行SCSI协议，使其能够在诸如高速千兆以太网上进行路由选择。iSCSI技术是一种新储存技术，该技术是将现有SCSI接口与以太网络(Ethernet)技术结合，使服务器可与使用IP网络的储存装置互相交换资料。iSCSI是一种基于 TCP/IP 的协议，用来建立和管理 IP存储设备、主机和客户机等之间的相互连接，并创建存储区域网络（SAN）。

iSCSI target：就是储存设备端，存放磁盘或RAID的设备，目前也能够将Linux主机模拟成iSCSI target了！目的在提供其他主机使用的『磁盘』；

iSCSI initiator：就是能够使用target的用户端，通常是服务器。也就是说，想要连接到iSCSI target的服务器，也必须要安装iSCSI initiator的相关功能后才能够使用iSCSI target提供的磁盘。

**3.1.1  iSCSI target**

|  |
| --- |
| 查看是否安装：  [root@ASM1 test]# **rpm -qa|grep iscsi**  没有发现，应该RAC两个节点都要安装 |

**3.1.2  iSCSI initiator**

**3.1.2.1  安装iSCSI initiator(RAC双节点运行)**

|  |
| --- |
| 安装：使用yum源安装  [root@ASM1 test]# **yum -y install iscsi\***  查看：  [root@ASM1 test]# **rpm -qa|grep scsi**  iscsi-initiator-utils-devel-6.2.0.873-27.el6\_9.x86\_64  iscsi-initiator-utils-6.2.0.873-27.el6\_9.x86\_64  启动：  [root@ASM1 test]# **service iscsid start**  [root@ASM1 test]# **chkconfig iscsid on**  //设置进程  [root@ASM1 test]# **chkconfig iscsi on**  **PS：**  **redhat 7设置开机启动：**  systemctl start iscsid.service  systemctl enable iscsid.service  systemctl enable iscsi.service |

**3.1.2.2  iscsiadm管理登陆及开机运行(RAC双节点运行)**

iscsi initiator主要通过iscsiadm命令管理，我们先查看提供服务的iscsi target机器上有哪些target:

|  |
| --- |
| 使用iscsiadm命令行接口发现网络存储服务器上的所有可用目标  [root@ASM1 test]# **iscsiadm -m discovery -t sendtargets -p 192.168.70.100 //IP为Openfiler地址**  正在启动 iscsid： [确定]  192.168.70.100:3260,1 iqn.2006-01.com.openfiler:racdb.fra  192.168.71.100:3260,1 iqn.2006-01.com.openfiler:racdb.fra  192.168.70.100:3260,1 iqn.2006-01.com.openfiler:racdb.data  192.168.71.100:3260,1 iqn.2006-01.com.openfiler:racdb.data  192.168.70.100:3260,1 iqn.2006-01.com.openfiler:racdb.crs  192.168.71.100:3260,1 iqn.2006-01.com.openfiler:racdb.crs  因为OpenFiler有双网卡，所以可以看到两个不同的设备。  [root@ASM1 test]# **ps -ef|grep iscsi**  root 21368 2 0 08:12 ? 00:00:00 [iscsi\_eh]  root 21390 1 0 08:12 ? 00:00:00 iscsiuio  root 21396 1 0 08:12 ? 00:00:00 iscsid  root 21397 1 0 08:12 ? 00:00:00 iscsid  root 21402 8955 0 08:19 pts/1 00:00:00 grep iscsi  所有服务都运行  PS：  discovery参数是报如下错误：  [root@node1 ~]# iscsiadm -m discovery -t sendtargets -p 192.168.12.100  iscsiadm: No portals found  [root@node1 ~]#  使用telnet命令检查openfiler端口开放情况发现3260端口已经开放  [root@node1 ~]# telnet 192.168.12.100 3260  Trying 192.168.12.100...  Connected to 192.168.12.100.  Escape character is '^]'.  2  解决方法  在openfiler12主机修改/etc/initiators.deny文件。将文件中全部行注释掉。如下：  [root@localhost ~]# cat /etc/initiators.deny  # PLEASE DO NOT MODIFY THIS CONFIGURATION FILE!  #       This configuration file was autogenerated  #       by Openfiler. Any manual changes will be overwritten  #       Generated at: Sun Jun 22 19:39:42 CST 2014    #iqn.2006-01.com.openfiler:tsn.99c16437c079 ALL    # End of Openfiler configuration    [root@localhost ~]#    做完上面操作后到node1运行iscsiadm discovery命令得到正确返回结果。  [root@node1 ~]# iscsiadm -m discovery -t sendtargets -p 192.168.12.100  192.168.12.100:3260,1 iqn.2006-01.com.openfiler:tsn.99c16437c079  192.168.70.100:3260,1 iqn.2006-01.com.openfiler:tsn.99c16437c079 |

到这一步就可以看出，你服务端创建的iSCSI Target 的编号和名称。这条命令只需记住-p后面跟iSCSI服务的地址就行了，也可以是主机名，都可以！3260是服务的端口号，默认的！

然后就可以登陆某个target了，登陆成功某个target后，这个target下的硬盘也就都共享过来了：

**登陆ISCSI远程磁盘：(用OpenFiler双网卡地址都登陆一次)**

[root@ASM1 test]# **iscsiadm -m node -T iqn.2006-01.com.openfiler:racdb.crs -p 192.168.70.100 -l**

Logging in to [iface: default, target: iqn.2006-01.com.openfiler:racdb.crs, portal: 192.168.70.100,3260] (multiple)

Login to [iface: default, target: iqn.2006-01.com.openfiler:racdb.crs, portal: 192.168.70.100,3260] successful.

[root@ASM1 test]# **iscsiadm -m node -T iqn.2006-01.com.openfiler:racdb.crs -p 192.168.71.100 -l**

Logging in to [iface: default, target: iqn.2006-01.com.openfiler:racdb.crs, portal: 192.168.71.100,3260] (multiple)

Login to [iface: default, target: iqn.2006-01.com.openfiler:racdb.crs, portal: 192.168.71.100,3260] successful.

**其他分区登陆如上，且两网卡IP都登陆一次。**

**设置开机自动连接ISCSI远程磁盘：**

[root@ASM1 test]# **iscsiadm -m node -T iqn.2006-01.com.openfiler:racdb.crs -p 192.168.70.100 --op update -n node.startup -v automatic**

[root@ASM1 test]# **iscsiadm -m node -T iqn.2006-01.com.openfiler:racdb.crs -p 192.168.71.100 --op update -n node.startup -v automatic**

**其他分区登陆如上，且两网卡IP都登陆一次。**

|  |
| --- |
| [root@ASM1 test]# **fdisk -l | grep dev**  Disk /dev/sda: 107.4 GB, 107374182400 bytes  /dev/sda1 \* 1 26 204800 83 Linux  /dev/sda2 26 13055 104651776 8e Linux LVM  Disk /dev/mapper/vg\_rac1-LogVol01: 98.6 GB, 98566144000 bytes  Disk /dev/mapper/vg\_rac1-LogVol00: 8594 MB, 8594128896 bytes  Disk /dev/sdb: 3221 MB, 3221225472 bytes  /dev/sdb1 1 130 1044193+ 83 Linux  /dev/sdb2 131 260 1044225 83 Linux  /dev/sdb3 261 391 1052257+ 83 Linux  Disk /dev/sdc: 3221 MB, 3221225472 bytes  /dev/sdc1 1 130 1044193+ 83 Linux  /dev/sdc2 131 260 1044225 83 Linux  /dev/sdc3 261 391 1052257+ 83 Linux  Disk /dev/sdd: 52.4 GB, 52445577216 bytes  Disk /dev/sde: 52.4 GB, 52445577216 bytes  Disk /dev/sdf: 52.4 GB, 52445577216 bytes  Disk /dev/sdg: 52.4 GB, 52445577216 bytes |

这里多出了6块盘，在openfiler中只map了3次，为什么这里是6块而不是3块呢？因为openfiler有2块网卡，使用两个IP登录两次iscsi target，所以这里有两块是重复的

要查看各个iscsi的信息：

|  |
| --- |
| [root@ASM1 test]# **iscsiadm -m session -P 3**  iSCSI Transport Class version 2.0-870  version 6.2.0-873.26.el6  Target: iqn.2006-01.com.openfiler:racdb.crs (non-flash)  Current Portal: 192.168.70.100:3260,1  Persistent Portal: 192.168.70.100:3260,1  \*\*\*\*\*\*\*\*\*\*  Interface:  \*\*\*\*\*\*\*\*\*\*  Iface Name: default  Iface Transport: tcp  Iface Initiatorname: iqn.1994-05.com.redhat:c9e851fffba1  Iface IPaddress: 192.168.70.101  Iface HWaddress: <empty>  Iface Netdev: <empty>  SID: 1  iSCSI Connection State: LOGGED IN  iSCSI Session State: LOGGED\_IN  Internal iscsid Session State: NO CHANGE  \*\*\*\*\*\*\*\*\*  Timeouts:  \*\*\*\*\*\*\*\*\*  Recovery Timeout: 120  Target Reset Timeout: 30  LUN Reset Timeout: 30  Abort Timeout: 15  \*\*\*\*\*  CHAP:  \*\*\*\*\*  username: <empty>  password: \*\*\*\*\*\*\*\*  username\_in: <empty>  password\_in: \*\*\*\*\*\*\*\*  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  Negotiated iSCSI params:  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  HeaderDigest: None  DataDigest: None  MaxRecvDataSegmentLength: 262144  MaxXmitDataSegmentLength: 131072  FirstBurstLength: 262144  MaxBurstLength: 262144  ImmediateData: No  InitialR2T: Yes  MaxOutstandingR2T: 1  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  Attached SCSI devices:  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  Host Number: 33 State: running  scsi33 Channel 00 Id 0 Lun: 0  Attached scsi disk sdb State: running  Current Portal: 192.168.71.100:3260,1  Persistent Portal: 192.168.71.100:3260,1  \*\*\*\*\*\*\*\*\*\*  Interface:  \*\*\*\*\*\*\*\*\*\*  Iface Name: default  Iface Transport: tcp  Iface Initiatorname: iqn.1994-05.com.redhat:c9e851fffba1  Iface IPaddress: 192.168.71.101  Iface HWaddress: <empty>  Iface Netdev: <empty>  SID: 2  iSCSI Connection State: LOGGED IN  iSCSI Session State: LOGGED\_IN  Internal iscsid Session State: NO CHANGE  \*\*\*\*\*\*\*\*\*  Timeouts:  \*\*\*\*\*\*\*\*\*  Recovery Timeout: 120  Target Reset Timeout: 30  LUN Reset Timeout: 30  Abort Timeout: 15  \*\*\*\*\*  CHAP:  \*\*\*\*\*  username: <empty>  password: \*\*\*\*\*\*\*\*  username\_in: <empty>  password\_in: \*\*\*\*\*\*\*\*  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  Negotiated iSCSI params:  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  HeaderDigest: None  DataDigest: None  MaxRecvDataSegmentLength: 262144  MaxXmitDataSegmentLength: 131072  FirstBurstLength: 262144  MaxBurstLength: 262144  ImmediateData: No  InitialR2T: Yes  MaxOutstandingR2T: 1  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  Attached SCSI devices:  \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  Host Number: 34 State: running  scsi34 Channel 00 Id 0 Lun: 0  Attached scsi disk sdc State: running  **其他略……** |

登陆之后要对新磁盘进行分区，格式化，然后在挂载即可

完成这些命令后，iscsi initator会把这些信息记录到/var/lib/iscsi目录下：

/var/lib/iscsi/send\_targets记录了各个target的情况，/var/lib/iscsi/nodes记录了各个target下的nodes情况。下次再启动iscsi initator时（service iscsi start)，就会自动登陆各个target上。如果想让重新手工登陆各个target，需要把/var/lib/iscsi/send\_targets目录下的内容和/var/lib/iscsi/nodes下的内容全部删除掉。

**3.2  多路径multipath**

**3.2.1  RAC的2个节点上分别安装multipath软件**

**1、安装多路径软件包：**

|  |
| --- |
| 查看多路径软件包，看是否有如下安装包，没有重新安装。  [root@ASM1 test]# **rpm -qa|grep device-mapper**  device-mapper-multipath-0.4.9-72.el6.x86\_64  device-mapper-persistent-data-0.2.8-2.el6.x86\_64  device-mapper-1.02.79-8.el6.x86\_64  device-mapper-event-libs-1.02.79-8.el6.x86\_64  device-mapper-event-1.02.79-8.el6.x86\_64  device-mapper-multipath-libs-0.4.9-72.el6.x86\_64  device-mapper-libs-1.02.79-8.el6.x86\_64  使用yum源安装：  [root@ASM1 test]# **yum -y install device-mapper-multipath\***  再次查看安装包：  [root@ASM1 test]# **rpm -qa|grep device-mapper**  device-mapper-event-libs-1.02.117-12.el6.x86\_64  device-mapper-1.02.117-12.el6.x86\_64  device-mapper-persistent-data-0.6.2-0.1.rc7.el6.x86\_64  device-mapper-multipath-libs-0.4.9-106.el6\_10.1.x86\_64  device-mapper-libs-1.02.117-12.el6.x86\_64  device-mapper-event-1.02.117-12.el6.x86\_64  device-mapper-multipath-0.4.9-106.el6\_10.1.x86\_64 |

**3.2.2  启动multipath**

将多路径软件添加至内核模块中

|  |
| --- |
| [root@ASM1 test]# **modprobe dm-multipath**  [root@ASM1 test]# **modprobe dm-round-robin** |

检查内核添加情况

|  |
| --- |
| [root@ASM1 test]# **lsmod | grep multipath**  dm\_multipath 17980 1 dm\_round\_robin  dm\_mod 102823 9 dm\_multipath,dm\_mirror,dm\_log |

将多路径软件multipath设置为开机自启动

|  |
| --- |
| [root@ASM1 test]# **chkconfig --level 2345 multipathd on**  [root@ASM1 test]# **chkconfig --list|grep multipathd**  multipathd 0:关闭 1:关闭 2:启用 3:启用 4:启用 5:启用 6:关闭 |

 启动multipath服务

|  |
| --- |
| [root@ASM1 test]# **service multipathd restart**  ux\_socket\_connect: No such file or directory  正在关闭multipathd 端口监控程序： [失败]  正在启动守护进程multipathd： [确定]  **PS：**  **redhat 7设置多路径：**  1、安装多路径软件  yum -y install device-mapper-multipath\*  2、创建配置文件  mpathconf –enable  3、启动多路径服务  systemctl start multipathd.service  4、查看多路径服务状态  systemctl status multipathd.service  5、创建自启动服务  systemctl enable multipathd.service  PS:生成配置文件是/etc/multipath.conf  6、查看HBA的WWN  FC挂载存储，寻找WWN号：  cat /sys/class/fc\_host/host\*/port\_name  7、查看存储分给服务器的lun  扫盘命令：echo "- - -" > /sys/class/scsi\_host/host0/scan 或 重启服务器  8、查看多路径中磁盘状态  multipath -ll  9、查看磁盘聚合结果  fdisk -l  PS：好像不用编辑multipath.conf,RAC可用UDEV直接绑定做 |

**3.2.3  配置多路径软件/etc/multipath.conf**

**1、配置multipath软件，编辑/etc/multipath.conf**

注意:默认情况下，/etc/multipath.conf是不存在的，需要用如下命令生成multipath.conf文件:

/sbin/mpathconf --enable --find\_multipaths y --with\_module y --with\_chkconfig y

|  |
| --- |
| [root@ASM1 test]# **multipath -ll**  Apr 20 08:50:19 | /etc/multipath.conf does not exist, blacklisting all devices.  Apr 20 08:50:19 | A sample multipath.conf file is located at  Apr 20 08:50:19 | /usr/share/doc/device-mapper-multipath-0.4.9/multipath.conf  Apr 20 08:50:19 | You can run /sbin/mpathconf to create or modify /etc/multipath.conf  创建multipath.conf文件  [root@ASM1 test]# **/sbin/mpathconf --enable --find\_multipaths y --with\_module y --with\_chkconfig y**  查看是否创建  [root@ASM1 test]# ll /etc/multipath.conf  -rw------- 1 root root 2775 4月 20 08:51 /etc/multipath.conf |

**2、查看并获取存储分配给服务器的逻辑盘lun的wwid信息**

|  |
| --- |
| [root@ASM1 test]# **multipath -v0**  [root@ASM1 test]# **more /etc/multipath/wwids**  # Multipath wwids, Version : 1.0  # NOTE: This file is automatically maintained by multipath and multipathd.  # You should not need to edit this file in normal circumstances.  #  # Valid WWIDs:  /14f504e46494c45527a4e696649792d4f4e46612d7444744f/  /14f504e46494c4552434933666f612d736942622d736f536e/  /14f504e46494c455241324c37616c2d7a7673352d69677033/ |

**PS：RAC节点1创建完成后，将文件/etc/multipath/wwids和/etc/multipath/bindings的内容覆盖节点2：**

|  |
| --- |
| [root@ASM2 test]# **multipath -v0**  [root@ASM2 test]# **more /etc/multipath/wwids**  # Multipath wwids, Version : 1.0  # NOTE: This file is automatically maintained by multipath and multipathd.  # You should not need to edit this file in normal circumstances.  #  # Valid WWIDs:  /14f504e46494c45527a4e696649792d4f4e46612d7444744f/  /14f504e46494c4552434933666f612d736942622d736f536e/  /14f504e46494c455241324c37616c2d7a7673352d69677033/  查看绑定是否一致  [root@ASM1 test]# **more /etc/multipath/bindings**  # Multipath bindings, Version : 1.0  # NOTE: this file is automatically maintained by the multipath program.  # You should not need to edit this file in normal circumstances.  #  # Format:  # alias wwid  #  mpatha 14f504e46494c45527a4e696649792d4f4e46612d7444744f  mpathb 14f504e46494c4552434933666f612d736942622d736f536e  mpathc 14f504e46494c455241324c37616c2d7a7673352d69677033   [root@ASM2 test]# **more /etc/multipath/bindings**  # Multipath bindings, Version : 1.0  # NOTE: this file is automatically maintained by the multipath program.  # You should not need to edit this file in normal circumstances.  #  # Format:  # alias wwid  #  mpatha 14f504e46494c45527a4e696649792d4f4e46612d7444744f  mpathb 14f504e46494c4552434933666f612d736942622d736f536e  mpathc 14f504e46494c455241324c37616c2d7a7673352d69677033 |

|  |
| --- |
| 对比信息：  [root@ASM1 test]# **multipath -ll**  mpathc (14f504e46494c455241324c37616c2d7a7673352d69677033) dm-4 OPNFILER,VIRTUAL-DISK  size=49G features='0' hwhandler='0' wp=rw  |-+- policy='round-robin 0' prio=1 status=active  | `- 37:0:0:0 sdf 8:80 active ready running  `-+- policy='round-robin 0' prio=1 status=enabled  `- 38:0:0:0 sdg 8:96 active ready running  mpathb (14f504e46494c4552434933666f612d736942622d736f536e) dm-3 OPNFILER,VIRTUAL-DISK  size=49G features='0' hwhandler='0' wp=rw  |-+- policy='round-robin 0' prio=1 status=active  | `- 35:0:0:0 sdd 8:48 active ready running  `-+- policy='round-robin 0' prio=1 status=enabled  `- 36:0:0:0 sde 8:64 active ready running  mpatha (14f504e46494c45527a4e696649792d4f4e46612d7444744f) dm-2 OPNFILER,VIRTUAL-DISK  size=3.0G features='0' hwhandler='0' wp=rw  |-+- policy='round-robin 0' prio=1 status=active  | `- 33:0:0:0 sdb 8:16 active ready running  `-+- policy='round-robin 0' prio=1 status=enabled  `- 34:0:0:0 sdc 8:32 active ready running  [root@ASM2 test]# **multipath -ll**  mpathc (14f504e46494c455241324c37616c2d7a7673352d69677033) dm-4 OPNFILER,VIRTUAL-DISK  size=49G features='0' hwhandler='0' wp=rw  |-+- policy='round-robin 0' prio=1 status=active  | `- 37:0:0:0 sdf 8:80 active ready running  `-+- policy='round-robin 0' prio=1 status=enabled  `- 38:0:0:0 sdg 8:96 active ready running  mpathb (14f504e46494c4552434933666f612d736942622d736f536e) dm-3 OPNFILER,VIRTUAL-DISK  size=49G features='0' hwhandler='0' wp=rw  |-+- policy='round-robin 0' prio=1 status=active  | `- 35:0:0:0 sdd 8:48 active ready running  `-+- policy='round-robin 0' prio=1 status=enabled  `- 36:0:0:0 sde 8:64 active ready running  mpatha (14f504e46494c45527a4e696649792d4f4e46612d7444744f) dm-2 OPNFILER,VIRTUAL-DISK  size=3.0G features='0' hwhandler='0' wp=rw  |-+- policy='round-robin 0' prio=1 status=active  | `- 33:0:0:0 sdb 8:16 active ready running  `-+- policy='round-robin 0' prio=1 status=enabled  `- 34:0:0:0 sdc 8:32 active ready running  对比查看磁盘信息  [root@ASM1 test]# **fdisk -l | grep dev**  Disk /dev/sda: 107.4 GB, 107374182400 bytes  /dev/sda1 \* 1 26 204800 83 Linux  /dev/sda2 26 13055 104651776 8e Linux LVM  Disk /dev/mapper/vg\_rac1-LogVol01: 98.6 GB, 98566144000 bytes  Disk /dev/mapper/vg\_rac1-LogVol00: 8594 MB, 8594128896 bytes  Disk /dev/sdb: 3221 MB, 3221225472 bytes  /dev/sdb1 1 130 1044193+ 83 Linux  /dev/sdb2 131 260 1044225 83 Linux  /dev/sdb3 261 391 1052257+ 83 Linux  Disk /dev/sdc: 3221 MB, 3221225472 bytes  /dev/sdc1 1 130 1044193+ 83 Linux  /dev/sdc2 131 260 1044225 83 Linux  /dev/sdc3 261 391 1052257+ 83 Linux  Disk /dev/sdd: 52.4 GB, 52445577216 bytes  Disk /dev/sde: 52.4 GB, 52445577216 bytes  Disk /dev/sdf: 52.4 GB, 52445577216 bytes  Disk /dev/sdg: 52.4 GB, 52445577216 bytes  Disk /dev/mapper/mpatha: 3221 MB, 3221225472 bytes  /dev/mapper/mpathap1 1 130 1044193+ 83 Linux  /dev/mapper/mpathap2 131 260 1044225 83 Linux  /dev/mapper/mpathap3 261 391 1052257+ 83 Linux  Disk /dev/mapper/mpathb: 52.4 GB, 52445577216 bytes  Disk /dev/mapper/mpathc: 52.4 GB, 52445577216 bytes  [root@ASM2 test]# **fdisk -l | grep dev**  Disk /dev/sda: 107.4 GB, 107374182400 bytes  /dev/sda1 \* 1 26 204800 83 Linux  /dev/sda2 26 13055 104651776 8e Linux LVM  Disk /dev/mapper/vg\_rac2-LogVol01: 98.6 GB, 98566144000 bytes  Disk /dev/mapper/vg\_rac2-LogVol00: 8594 MB, 8594128896 bytes  Disk /dev/sdb: 3221 MB, 3221225472 bytes  /dev/sdb1 1 130 1044193+ 83 Linux  /dev/sdb2 131 260 1044225 83 Linux  /dev/sdb3 261 391 1052257+ 83 Linux  Disk /dev/sdc: 3221 MB, 3221225472 bytes  /dev/sdc1 1 130 1044193+ 83 Linux  /dev/sdc2 131 260 1044225 83 Linux  /dev/sdc3 261 391 1052257+ 83 Linux  Disk /dev/sdd: 52.4 GB, 52445577216 bytes  Disk /dev/sde: 52.4 GB, 52445577216 bytes  Disk /dev/sdf: 52.4 GB, 52445577216 bytes  Disk /dev/sdg: 52.4 GB, 52445577216 bytes  Disk /dev/mapper/mpatha: 3221 MB, 3221225472 bytes  /dev/mapper/mpathap1 1 130 1044193+ 83 Linux  /dev/mapper/mpathap2 131 260 1044225 83 Linux  /dev/mapper/mpathap3 261 391 1052257+ 83 Linux  Disk /dev/mapper/mpathb: 52.4 GB, 52445577216 bytes  Disk /dev/mapper/mpathc: 52.4 GB, 52445577216 bytes |

**3.2.4  编辑/etc/multipath.conf**

用如下脚本： (也可用multipath -ll获取WWID)

|  |
| --- |
| 获取WWID值：  [root@ASM1 /]# **for i in a b c d e f g ;**  > **do**  > **echo "KERNEL==\"sd\*\", BUS==\"scsi\", PROGRAM==\"/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/\$name\", RESULT==\"`/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/mapper/mpath$i`\", NAME=\"asm-disk$i\", OWNER=\"grid\", GROUP=\"asmadmin\", MODE=\"0660\""**  > **done**  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="14f504e46494c45527a4e696649792d4f4e46612d7444744f", NAME="asm-diska", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="14f504e46494c4552434933666f612d736942622d736f536e", NAME="asm-diskb", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="14f504e46494c455241324c37616c2d7a7673352d69677033", NAME="asm-diskc", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="", NAME="asm-diskd", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="", NAME="asm-diske", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="", NAME="asm-diskf", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="", NAME="asm-diskg", OWNER="grid", GROUP="asmadmin", MODE="0660"  修改multipath.conf(红色的修改)  [root@raclhr-12cR1-N1 multipath]# **vi /etc/multipath.conf**  defaults {          find\_multipaths yes          user\_friendly\_names yes  }    //去掉前面的#号  blacklist {        wwid 3600508b1001c5ae72efe1fea025cd2e5        devnode "^hd[a-z]"        devnode "^sd[a-e]"        devnode "^sda"  }    //去掉前面的#号并做相应修改(有几块磁盘就多加几条)  multipaths {         multipath {                 wwid                    14f504e46494c45527a4e696649792d4f4e46612d7444744f                 alias                   RACcrs                 path\_grouping\_policy    multibus                 path\_selector           "round-robin 0"                 failback                manual                 rr\_weight               priorities                 no\_path\_retry           5        }         multipath {                 wwid                    14f504e46494c4552434933666f612d736942622d736f536e                 alias                   RACdata                 path\_grouping\_policy    multibus                 path\_selector           "round-robin 0"                 failback                manual                 rr\_weight               priorities                 no\_path\_retry           5         }         multipath {                 wwid                    14f504e46494c455241324c37616c2d7a7673352d69677033                 alias                   RACfra                 path\_grouping\_policy    multibus                 path\_selector           "round-robin 0"                 failback                manual                 rr\_weight               priorities                 no\_path\_retry           5         }  }  devices {         device {                 vendor                  "VMWARE"                 product                 "VIRTUAL-DISK"                 path\_grouping\_policy    multibus                 getuid\_callout          "/lib/udev/scsi\_id --whitelisted --device=/dev/%n"                 path\_checker            readsector0                 path\_selector           "round-robin 0"                 hardware\_handler        "0"                 failback                15                 rr\_weight               priorities                 no\_path\_retry           queue         }  }  PS：示例  multipath.conf主要包括blacklist、multipaths、devices三部份的配置  blacklist配置  blacklist {  　　devnode "^sda" #去掉本地sda盘  }  Multipaths部分配置multipaths和devices两部份的配置。  multipaths {  　　multipath {  　　　　wwid \*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\* 　　　　　　#此值multipath -v3可以看到 或multipath -ll  　　　　alias iscsi-dm0 　　　　　　　　#映射后的别名,可以随便取  　　　　path\_grouping\_policy multibus 　　#路径组策略  　　　　path\_checker tur 　　　　　　　　#决定路径状态的方法  　　　　path\_selector "round-robin 0" 　　#选择那条路径进行下一个IO操作的方法  　　}  }  Devices部分配置  devices {  　　device {  　　　　vendor "iSCSI-Enterprise" 　　　　　　　　　　#厂商名称  　　　　product "Virtual disk" 　　　　　　　　　　　　#产品型号 此值multipath -v3可以看到\:  　　　　path\_grouping\_policy multibus 　　　　　　　　#默认的路径组策略  　　　　getuid\_callout "/sbin/scsi\_id -g -u -s /block/%n"　#获得唯一设备号使用的默认程序  　　　　prio\_callout      "/sbin/acs\_prio\_alua %d" 　　　  #获取有限级数值使用的默认程序  　　　　path\_checker readsector0 　　　　　　　　　　#决定路径状态的方法  　　　　path\_selector "round-robin 0" 　　#选择那条路径进行下一个IO操作的方法  　　　　failback        immediate 　　　　#故障恢复的模式     　　　no\_path\_retry      queue 　　　　#在disable queue之前系统尝试使用失效路径的次数的数值    　　　　rr\_min\_io       100 　　　　　　#在当前的用户组中，在切换到另外一条路径之前的IO请求的数目  　　}  } |

重启动multipath配置

|  |
| --- |
| [root@ASM1 /]# **service multipathd restart**  ok  正在关闭multipathd 端口监控程序： [确定]  正在启动守护进程multipathd： [确定]  [root@ASM1 /]# **multipath -ll**  Apr 20 16:10:19 | /etc/multipath.conf line 69, invalid keyword: path\_checker  Apr 20 16:10:19 | /etc/multipath.conf line 80, invalid keyword: path\_checker  Apr 20 16:10:19 | /etc/multipath.conf line 91, invalid keyword: path\_checker  Apr 20 16:10:19 | /etc/multipath.conf line 125, duplicate keyword: blacklist  RACcrs (14f504e46494c45527a4e696649792d4f4e46612d7444744f) dm-2 OPNFILER,VIRTUAL-DISK  size=3.0G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active  |- 33:0:0:0 sdb 8:16 active ready running  `- 34:0:0:0 sdc 8:32 active ready running  RACfra (14f504e46494c455241324c37616c2d7a7673352d69677033) dm-4 OPNFILER,VIRTUAL-DISK  size=49G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active  |- 37:0:0:0 sdf 8:80 active ready running  `- 38:0:0:0 sdg 8:96 active ready running  RACdata (14f504e46494c4552434933666f612d736942622d736f536e) dm-3 OPNFILER,VIRTUAL-DISK  size=49G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active  |- 35:0:0:0 sdd 8:48 active ready running  `- 36:0:0:0 sde 8:64 active ready running  [root@ASM1 /]# **multipath -ll|grep RAC**  RACcrs (14f504e46494c45527a4e696649792d4f4e46612d7444744f) dm-2 OPNFILER,VIRTUAL-DISK  RACfra (14f504e46494c455241324c37616c2d7a7673352d69677033) dm-4 OPNFILER,VIRTUAL-DISK  RACdata (14f504e46494c4552434933666f612d736942622d736f536e) dm-3 OPNFILER,VIRTUAL-DISK |

启用multipath配置后,会在/dev/mapper下生成多路径逻辑盘

|  |
| --- |
| [root@ASM1 /]# **cd /dev/mapper**  [root@ASM1 mapper]# **ll**  总用量 0  crw-rw---- 1 root root 10, 58 4月 20 07:56 control  lrwxrwxrwx 1 root root 7 4月 20 16:09 RACcrs -> ../dm-2  lrwxrwxrwx 1 root root 7 4月 20 16:09 RACcrsp1 -> ../dm-5  lrwxrwxrwx 1 root root 7 4月 20 16:09 RACcrsp2 -> ../dm-6  lrwxrwxrwx 1 root root 7 4月 20 16:09 RACcrsp3 -> ../dm-7  lrwxrwxrwx 1 root root 7 4月 20 16:09 RACdata -> ../dm-3  lrwxrwxrwx 1 root root 7 4月 20 16:09 RACfra -> ../dm-4  lrwxrwxrwx 1 root root 7 4月 20 07:56 vg\_rac1-LogVol00 -> ../dm-1  lrwxrwxrwx 1 root root 7 4月 20 07:56 vg\_rac1-LogVol01 -> ../dm-0 |

至此，多路径multipath配置完成。

PS：记得更改另一个节点，

**3.2.5  配置multipath设备的权限**

在6.2之前配置multipath设备的权限只需要在设备配置里增加uid,gid,mode就可以

uid 1100 #uid

gid 1020 #gid

如：

|  |
| --- |
| multipath {                  wwid                    360050763008101d4e00000000000000a                  alias                   DATA03                  uid                     501                                               #uid                  gid                     501                                               #gid  } |

在6.2之后配置multipath配置文件里去掉uid,gid,mode这三个参数，需要使用udev使用,示例文件在/usr/share/doc/device-mapper-version中有一个模板文件,名为12-dm-permissions.rules,您可以使用它并将其放在 /etc/udev/rules.d 目录中使其生效。

|  |
| --- |
| [root@ASM1 mapper]# **ll /usr/share/doc/device-mapper-1.02.117/12-dm-permissions.rules**  -rw-r--r--. 1 root root 3186 2月 22 2016 /usr/share/doc/device-mapper-1.02.117/12-dm-permissions.rules  [root@ASM1 device-mapper-1.02.117]# **cd /etc/udev/rules.d/**  [root@ASM1 rules.d]# **ll**  总用量 16  -rw-r--r--. 1 root root 316 1月 3 2017 60-raw.rules  -rw-r--r--. 1 root root 803 4月 20 06:40 70-persistent-cd.rules  -rw-r--r--. 1 root root 915 4月 20 06:30 70-persistent-net.rules  -rw-r--r--. 1 root root 304 1月 20 2017 98-kexec.rules  [root@ASM1 rules.d]# **cp /usr/share/doc/device-mapper-1.02.117/12-dm-permissions.rules ./**  [root@ASM1 rules.d]# **ll**  总用量 20  -rw-r--r-- 1 root root 3186 4月 20 16:22 12-dm-permissions.rules  -rw-r--r--. 1 root root 316 1月 3 2017 60-raw.rules  -rw-r--r--. 1 root root 803 4月 20 06:40 70-persistent-cd.rules  -rw-r--r--. 1 root root 915 4月 20 06:30 70-persistent-net.rules  -rw-r--r--. 1 root root 304 1月 20 2017 98-kexec.rules  [root@ASM1 rules.d]# **vi /etc/udev/rules.d/12-dm-permissions.rules**  ENV{DM\_NAME}=="RAC\*", OWNER:="grid", GROUP:="asmadmin", MODE:="660"  **PS:** RAC\* 多路径定义的别名。  30S后查看权限有没有变化，如没变化则正常：  [root@ASM1 dev]# ll dm\*  brw-rw---- 1 root disk 253, 0 4月 20 18:40 dm-0  brw-rw---- 1 root disk 253, 1 4月 20 18:40 dm-1  brw-rw---- 1 grid asmadmin 253, 2 4月 20 18:47 dm-2  brw-rw---- 1 grid asmadmin 253, 3 4月 20 18:40 dm-3  brw-rw---- 1 grid asmadmin 253, 4 4月 20 18:40 dm-4  brw-rw---- 1 grid asmadmin 253, 5 4月 20 18:40 dm-5  brw-rw---- 1 grid asmadmin 253, 6 4月 20 18:48 dm-6  brw-rw---- 1 grid asmadmin 253, 7 4月 20 18:48 dm-7  crw-rw---- 1 root audio 14, 9 4月 20 18:40 dmmidi |

将文件/etc/udev/rules.d/12-dm-permissions.rules复制到节点2上。

**3.2.6  配置udev规则(这以上部分可以单独用)**

**用UDEV可以绑定多路径配置的磁盘，但绑定后子分区由于WWID是一样的，所以不建立用UDEV绑定的磁盘再划子分区。建议用UDEV最好是单独一个分区，这样ASM创建时能认出。**

**多路径下不能直接用/dev/mapper/下的设备，但可以有子分区，子分区可以由ASM创建，但没有grid权限，所以要用/dev/dm\* 。注意/dev/mapper/\*与/dev/dm\*的对应关系。**

**如不用多路径绑定，而直接用UDEV，由于多路径的设置WWID有重复，所以应该去掉文件/etc/udev/rules.d/99-oracleasm.rules中的重复的行。（这样多路径就不起作用）**

在节点1执行以下操作：(如果更换了别名，红色部分可能会有问题)

|  |
| --- |
| [root@ASM1 rules.d]# **for i in a b c ;**  > **do**  > **echo "KERNEL==\"sd\*\", BUS==\"scsi\", PROGRAM==\"/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/\$name\", RESULT==\"`/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/mapper/mpath$i`\", NAME=\"asm-disk$i\", OWNER=\"grid\", GROUP=\"asmadmin\", MODE=\"0660\"" >> /etc/udev/rules.d/99-oracleasm.rules**  > **done**  **PS：以上是LINUX6脚本，其他脚本如下：**  **Linux 5：**  **for i in b c d e f g h i ;**  **do**  **echo "KERNEL==\"sd\*\", BUS==\"scsi\", PROGRAM==\"/sbin/scsi\_id -g -u -s %p\", ESULT==\"`scsi\_id -g -u -s /block/sd$i`\", NAME=\"asm-disk$i\", OWNER=\"grid\", GROUP=\"asmadmin\", MODE=\"0660\""**  **done**  **/sbin/start\_udev**  **Linux6：**  **for i in b c d e f g ;**  **do**  **echo "KERNEL==\"sd\*\", BUS==\"scsi\", PROGRAM==\"/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/\$name\",**  **RESULT==\"`/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/sd$i`\", NAME=\"asm-disk$i\", OWNER=\"grid\",**  **GROUP=\"asmadmin\", MODE=\"0660\""      >> /etc/udev/rules.d/99-oracle-asmdevices.rules**  **done**  **/sbin/start\_udev**  **Linux 7：**  **for i in b c d e ;**  **do**  **echo "KERNEL==\"sd\*\", ENV{DEVTYPE}==\"disk\", SUBSYSTEM==\"block\", PROGRAM==\"/lib/udev/scsi\_id -g -u -d \$devnode\",RESULT==\"`/lib/udev/scsi\_id -g -u -d /dev/sd$i`\", SYMLINK+=\"asm-disk$i\", OWNER=\"grid\", GROUP=\"asmadmin\", MODE=\"0660\""      >> /etc/udev/rules.d/99-oracle-asmdevices.rules**  **done**  **#udevadm control --reload-rules #udevadm trigger**  **#systemctl restart systemd-udevd.service**  PS：多路径时注意更换红字部分。如下(绑定DM设备，可编辑99文件更改别名)：  **Linux 7：**  **for i in 2 3 4 5 6 7 8 9 10 ;**  **do**  **echo "KERNEL==\"dm-\*\", ENV{DEVTYPE}==\"disk\", SUBSYSTEM==\"block\", PROGRAM==\"/lib/udev/scsi\_id -g -u -d \$devnode\",RESULT==\"`/lib/udev/scsi\_id -g -u -d /dev/dm-$i`\", SYMLINK+=\"asm-disk$i\", OWNER=\"grid\", GROUP=\"asmadmin\", MODE=\"0660\""      >> /etc/udev/rules.d/99-oracle-asmdevices.rules**  **done** |

打开文件/etc/udev/rules.d/99-oracleasm.rules，如有重复，去掉WWID重复的行只保留一行即可（做过多路径略过）。

|  |
| --- |
| [root@ASM1 rules.d]# vi 99-oracleasm.rules  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="14f504e46494c45527a4e696649792d4f4e46612d7444744f", NAME="asm-diska", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="14f504e46494c4552434933666f612d736942622d736f536e", NAME="asm-diskb", OWNER="grid", GROUP="asmadmin", MODE="0660"  KERNEL=="sd\*", BUS=="scsi", PROGRAM=="/sbin/scsi\_id --whitelisted --replace-whitespace --device=/dev/$name", RESULT=="14f504e46494c455241324c37616c2d7a7673352d69677033", NAME="asm-diskc", OWNER="grid", GROUP="asmadmin", MODE="0660" |

将文件/etc/udev/rules.d/99-oracleasm.rules的内容拷贝到节点2，然后重启udev。

|  |
| --- |
| [root@ASM1 mapper]# **start\_udev**  正在启动 udev： [确定]  [root@ASM1 ~]# **ll /dev/asm-disk\***  brw-rw---- 1 grid asmadmin 8, 16 4月 20 17:27 /dev/asm-diskc  brw-rw---- 1 grid asmadmin 8, 96 4月 20 17:27 /dev/asm-diske  brw-rw---- 1 grid asmadmin 8, 64 4月 20 17:27 /dev/asm-diskg  [grid@raclhr-12cR1-N1 ~]ORACLE\_HOME/bin/kfod disks=all s=true ds=true  --------------------------------------------------------------------------------  Disk          Size Header    Path                                     Disk Group   User     Group  ================================================================================     1:       6144 Mb MEMBER    /dev/asm-diskc                           OCR          grid     asmadmin     2:      10240 Mb MEMBER    /dev/asm-diskd                           DATA         grid     asmadmin     3:      10240 Mb MEMBER    /dev/asm-diske                           FRA          grid     asmadmin     4:      10240 Mb CANDIDATE /dev/asm-diskf                           #            grid     asmadmin     5:      10240 Mb CANDIDATE /dev/asm-diskh                           #            grid     asmadmin     6:      10240 Mb CANDIDATE /dev/asm-diskj                           #            grid     asmadmin     7:      10240 Mb CANDIDATE /dev/asm-diskk                           #            grid     asmadmin  --------------------------------------------------------------------------------  ORACLE\_SID ORACLE\_HOME  ================================================================================       +ASM2 /u01/app/12.1.0/grid       +ASM1 /u01/app/12.1.0/grid  [grid@raclhr-12cR1-N1 ~]$ asmcmd    ASMCMD> lsdg  State    Type    Rebal  Sector  Block       AU  Total\_MB  Free\_MB  Req\_mir\_free\_MB  Usable\_file\_MB  Offline\_disks  Voting\_files  Name  MOUNTED  EXTERN  N         512   4096  1048576     10240     6487                0            6487              0             N  DATA/  MOUNTED  EXTERN  N         512   4096  1048576     10240    10144                0           10144              0             N  FRA/  MOUNTED  EXTERN  N         512   4096  1048576      6144     1672                0            1672              0             Y  OCR/  ASMCMD> lsdsk  Path  /dev/asm-diskc  /dev/asm-diskd  /dev/asm-diske  ASMCMD>  lsdsk --candidate -p  Group\_Num  Disk\_Num      Incarn  Mount\_Stat  Header\_Stat  Mode\_Stat  State   Path          0         1           0  CLOSED      CANDIDATE    ONLINE     NORMAL  /dev/asm-diskf          0         3           0  CLOSED      CANDIDATE    ONLINE     NORMAL  /dev/asm-diskh          0         2           0  CLOSED      CANDIDATE    ONLINE     NORMAL  /dev/asm-diskj          0         0           0  CLOSED      CANDIDATE    ONLINE     NORMAL  /dev/asm-diskk  ASMCMD> |

**3.3  利用新磁盘创建磁盘组**

CREATE DISKGROUP FRA external redundancy DISK '/dev/asm-diskf','/dev/asm-diskh' ATTRIBUTE 'compatible.rdbms' = '12.1', 'compatible.asm' = '12.1';

|  |
| --- |
| SQL> select path from v$asm\_disk;    PATH  --------------------------------------------------------------------------------  /dev/asm-diskk  /dev/asm-diskf  /dev/asm-diskj  /dev/asm-diskh  /dev/asm-diske  /dev/asm-diskd  /dev/asm-diskc    7 rows selected.    SQL> CREATE DISKGROUP TESTMUL external redundancy DISK '/dev/asm-diskf','/dev/asm-diskh' ATTRIBUTE 'compatible.rdbms' = '12.1', 'compatible.asm' = '12.1';    Diskgroup created.    SQL>  ASMCMD> lsdg  State    Type    Rebal  Sector  Block       AU  Total\_MB  Free\_MB  Req\_mir\_free\_MB  Usable\_file\_MB  Offline\_disks  Voting\_files  Name  MOUNTED  EXTERN  N         512   4096  1048576     10240     6487                0            6487              0             N  DATA/  MOUNTED  EXTERN  N         512   4096  1048576     10240    10144                0           10144              0             N  FRA/  MOUNTED  EXTERN  N         512   4096  1048576      6144     1672                0            1672              0             Y  OCR/  MOUNTED  EXTERN  N         512   4096  1048576     20480    20381                0           20381              0             N  TESTMUL/  ASMCMD>    [root@raclhr-12cR1-N1 ~]# crsctl stat res -t | grep -2 TESTMUL                 ONLINE  ONLINE       raclhr-12cr1-n1          STABLE                 ONLINE  ONLINE       raclhr-12cr1-n2          STABLE  ora.TESTMUL.dg                 ONLINE  ONLINE       raclhr-12cr1-n1          STABLE                 ONLINE  ONLINE       raclhr-12cr1-n2          STABLE  [root@raclhr-12cR1-N1 ~]# |

**3.3.1  测试磁盘组**

|  |
| --- |
| [oracle@raclhr-12cR1-N1 ~]$ sqlplus / as sysdba    SQL\*Plus: Release 12.1.0.2.0 Production on Mon Jan 23 16:17:28 2017    Copyright (c) 1982, 2014, Oracle.  All rights reserved.      Connected to:  Oracle Database 12c Enterprise Edition Release 12.1.0.2.0 - 64bit Production  With the Partitioning, Real Application Clusters, Automatic Storage Management, OLAP,  Advanced Analytics and Real Application Testing options    SQL> create tablespace TESTMUL datafile '+TESTMUL' size 10M;    Tablespace created.    SQL> select name from v$datafile;    NAME  --------------------------------------------------------------------------------  +DATA/LHRRAC/DATAFILE/system.258.933550527  +DATA/LHRRAC/DATAFILE/undotbs2.269.933551323  +DATA/LHRRAC/DATAFILE/sysaux.257.933550483  +DATA/LHRRAC/DATAFILE/undotbs1.260.933550575  +DATA/LHRRAC/DATAFILE/example.268.933550723  +DATA/LHRRAC/DATAFILE/users.259.933550573  +TESTMUL/LHRRAC/DATAFILE/testmul.256.934042679    7 rows selected.    SQL> |

将存储停掉一块网卡eth1：

|  |
| --- |
| [root@OFLHR ~]# ip a  1: lo: <loopback,up,10000> mtu 16436 qdisc noqueue      link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00      inet 127.0.0.1/8 scope host lo      inet6 ::1/128 scope host         valid\_lft forever preferred\_lft forever  2: eth0: <broadcast,multicast,up,10000> mtu 1500 qdisc pfifo\_fast qlen 1000      link/ether 00:0c:29:98:1a:cd brd ff:ff:ff:ff:ff:ff      inet 192.168.59.200/24 brd 192.168.59.255 scope global eth0      inet6 fe80::20c:29ff:fe98:1acd/64 scope link         valid\_lft forever preferred\_lft forever  3: eth1: <broadcast,multicast,up,10000> mtu 1500 qdisc pfifo\_fast qlen 1000      link/ether 00:0c:29:98:1a:d7 brd ff:ff:ff:ff:ff:ff      inet 192.168.2.200/24 brd 192.168.2.255 scope global eth1      inet6 fe80::20c:29ff:fe98:1ad7/64 scope link         valid\_lft forever preferred\_lft forever  [root@OFLHR ~]# ifconfig eth1 down  [root@OFLHR ~]# ip a  1: lo: <loopback,up,10000> mtu 16436 qdisc noqueue      link/loopback 00:00:00:00:00:00 brd 00:00:00:00:00:00      inet 127.0.0.1/8 scope host lo      inet6 ::1/128 scope host         valid\_lft forever preferred\_lft forever  2: eth0: <broadcast,multicast,up,10000> mtu 1500 qdisc pfifo\_fast qlen 1000      link/ether 00:0c:29:98:1a:cd brd ff:ff:ff:ff:ff:ff      inet 192.168.59.200/24 brd 192.168.59.255 scope global eth0      inet6 fe80::20c:29ff:fe98:1acd/64 scope link         valid\_lft forever preferred\_lft forever  3: eth1: <broadcast,multicast> mtu 1500 qdisc pfifo\_fast qlen 1000      link/ether 00:0c:29:98:1a:d7 brd ff:ff:ff:ff:ff:ff      inet 192.168.2.200/24 brd 192.168.2.255 scope global eth1  [root@OFLHR ~]# |

rac节点查看日志：

|  |
| --- |
| [root@raclhr-12cR1-N1 ~]# tail -f /var/log/messages  Jan 23 16:20:51 raclhr-12cR1-N1 iscsid: connect to 192.168.2.200:3260 failed (No route to host)  Jan 23 16:20:57 raclhr-12cR1-N1 iscsid: connect to 192.168.2.200:3260 failed (No route to host)  Jan 23 16:21:03 raclhr-12cR1-N1 iscsid: connect to 192.168.2.200:3260 failed (No route to host)  [root@raclhr-12cR1-N1 ~]# multipath -ll  VMLHRStorage003 (14f504e46494c4552614e35626c6f2d4e794d702d4c344a6c) dm-8 OPNFILER,VIRTUAL-DISK  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 5:0:0:3 sdm 8:192 failed faulty running    `- 4:0:0:3 sdl 8:176 active ready  running  VMLHRStorage002 (14f504e46494c455233384b7353432d52454b4c2d79506757) dm-9 OPNFILER,VIRTUAL-DISK  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 5:0:0:2 sdj 8:144 failed faulty running    `- 4:0:0:2 sdk 8:160 active ready  running  VMLHRStorage001 (14f504e46494c455242674c7079392d753750482d63734443) dm-7 OPNFILER,VIRTUAL-DISK  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 4:0:0:1 sdi 8:128 active ready  running    `- 5:0:0:1 sdh 8:112 failed faulty running  VMLHRStorage000 (14f504e46494c455232326c6c76442d4361634f2d4d4f4d41) dm-6 OPNFILER,VIRTUAL-DISK  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 4:0:0:0 sdf 8:80  active ready  running    `- 5:0:0:0 sdg 8:96  failed faulty running  [root@raclhr-12cR1-N1 ~]# |

表空间可以正常访问：

|  |
| --- |
| SQL> create table tt tablespace TESTMUL as select \* from dual;    Table created.    SQL> select \* from tt;    D  -  X    SQL> |

同理，将eth1进行up，而将eth0宕掉，表空间依然正常。**重启集群和存储后，集群一切正常。**

**第四章 测试多路径**

重新搭建一套多路径的环境来测试多路径。

最简单的测试方法，是用dd往磁盘读写数据，然后用iostat观察各通道的流量和状态，以判断Failover或负载均衡方式是否正常：

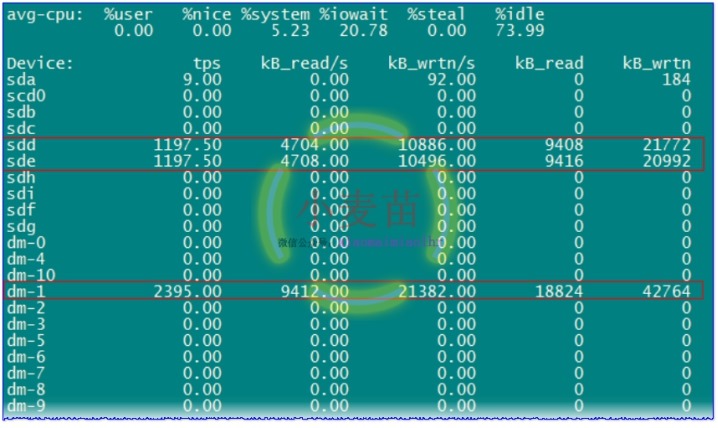
# dd if=/dev/zero of=/dev/mapper/mpath0

# iostat -k 2

|  |
| --- |
| [root@orcltest ~]# multipath -ll  VMLHRStorage003 (14f504e46494c4552674a61727a472d523449782d5336784e) dm-3 OPNFILER,VIRTUAL-DISK  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 35:0:0:2 sdf 8:80  active ready running    `- 36:0:0:2 sdg 8:96  active ready running  VMLHRStorage002 (14f504e46494c4552506a5a5954422d6f6f4e652d34423171) dm-2 OPNFILER,VIRTUAL-DISK  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 35:0:0:3 sdh 8:112 active ready running    `- 36:0:0:3 sdi 8:128 active ready running  **VMLHRStorage001 (14f504e46494c4552324b583573332d774e5a622d696d7334) dm-1 OPNFILER,VIRTUAL-DISK**  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 35:0:0:1 sdd 8:48  active ready running    `- 36:0:0:1 sde 8:64  active ready running  VMLHRStorage000 (14f504e46494c45523431576859532d643246412d5154564f) dm-0 OPNFILER,VIRTUAL-DISK  size=10G features='1 queue\_if\_no\_path' hwhandler='0' wp=rw  `-+- policy='round-robin 0' prio=1 status=active    |- 35:0:0:0 sdb 8:16  active ready running    `- 36:0:0:0 sdc 8:32  active ready running  [root@orcltest ~]# dd if=/dev/zero of=/dev/mapper/VMLHRStorage001 |

重新开一个窗口执行iostat -k 2可以看到：

|  |
| --- |
| avg-cpu:  %user   %nice %system %iowait  %steal   %idle             0.00    0.00    5.23   20.78    0.00   73.99    Device:            tps    kB\_read/s    kB\_wrtn/s    kB\_read    kB\_wrtn  sda               9.00         0.00        92.00          0        184  scd0              0.00         0.00         0.00          0          0  sdb               0.00         0.00         0.00          0          0  sdc               0.00         0.00         0.00          0          0  sdd            1197.50      4704.00     10886.00       9408      21772  sde            1197.50      4708.00     10496.00       9416      20992  sdh               0.00         0.00         0.00          0          0  sdi               0.00         0.00         0.00          0          0  sdf               0.00         0.00         0.00          0          0  sdg               0.00         0.00         0.00          0          0  dm-0              0.00         0.00         0.00          0          0  dm-4              0.00         0.00         0.00          0          0  dm-10             0.00         0.00         0.00          0          0  dm-1           2395.00      9412.00     21382.00      18824      42764  dm-2              0.00         0.00         0.00          0          0  dm-3              0.00         0.00         0.00          0          0  dm-5              0.00         0.00         0.00          0          0  dm-6              0.00         0.00         0.00          0          0  dm-7              0.00         0.00         0.00          0          0  dm-8              0.00         0.00         0.00          0          0  dm-9              0.00         0.00         0.00          0          0 |

[](http://images2015.cnblogs.com/blog/646850/201701/646850-20170123231449769-1979273361.jpg)

好了，有关使用OpenFiler来模拟存储配置RAC中ASM共享盘及多路径的测试就到此为止了，2016年结束了，今天是1月23日，明天是1月24日，小麦苗回家过年了，O(∩\_∩)O~。

**4.1  有关多路径其它理论知识**

用multipath生成映射后，会在/dev目录下产生多个指向同一条链路的设备：

|  |
| --- |
| /dev/mapper/mpathn  /dev/mpath/mpathn  /dev/dm-n |

但它们的来源是完全不同的：

**/dev/mapper/mpathn 是multipath虚拟出来的多路径设备，我们应该使用这个设备；/dev/mapper 中的设备是在引导过程中生成的。可使用这些设备访问多路径设备，例如在生成逻辑卷时。**

**/dev/mpath/mpathn 是udev设备管理器创建的，实际上就是指向下面的dm-n设备，仅为了方便，不能用来挂载；提供 /dev/mpath 中的设备是为了方便，这样可在一个目录中看到所有多路径设备。这些设备是由 udev 设备管理器生成的，且在系统需要访问它们时不一定能启动。请不要使用这些设备生成逻辑卷或者文件系统。**

**/dev/dm-n 是软件内部自身使用的，不能被软件以外使用，不可挂载。所有 /dev/dm-n 格式的设备都只能是作为内部使用，且应该永远不要使用。**

**简单来说，就是我们应该使用/dev/mapper/下的设备符。对该设备即可用fdisk进行分区，或创建为pv。**