

多路径设置

TYWD00060_v1.0

2017.05

文档说明：

- 北京同有飞骥科技股份有限公司（简称“同有科技”）版权所有 2017 年®。
- 如未事先得到北京同有飞骥科技股份有限公司的书面许可，本文档中任何部分都不得进行复制，或以任何形式、任何手段进行转载。
- 北京同有飞骥科技股份有限公司对本文档未作任何形式的担保，包括对具体用途的商品性和适用性的隐含担保。
- 北京同有飞骥科技股份有限公司对本文档中的全部内容有最终解释权，如有变更，恕不另行通知。
- 本文档中涉及第三方的品牌和名称是他们相应的拥有者的产权。
- NetStor®为北京同有飞骥科技股份有限公司的注册商标。

版权声明

北京同有飞骥科技股份有限公司版权所有，并保留对本文档及本声明的最终解释权和修改权。本文档中出现的任何文字叙述、格式、插图、图片、方法等内容，除另有特别注明外，其著作权均属于北京同有飞骥科技有限公司。未经本公司书面许可，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。

产品声明

本文档是针对北京同有飞骥科技股份有限公司 NetStor® iSUM 系列存储设备产品的使用指导说明手册，除非另有约定，本文档仅作为使用指导。北京同有飞骥科技股份有限公司确信本文档在发布之日内容准确无误，由于产品版本升级或其他原因，本文档中的内容可随时更改而不另行通知。

【Linux 环境 iSCSI 存储及多路径功能配置】

一. 概述

本文详细介绍在 Linux OS 环境下 iSCSI 存储设备连接安装配置使用方法和步骤，特别是 Linux 操作系统提供的 Mutipath 多路径功能作为 Linux 平台下多路径主要可选方案之一，在本文中有重点介绍和讲述。

二. 安装配置步骤和方法

2.1 客户端安装 iSCSI 包

- (1) 服务器安装 `iscsi initiator` 包。安装包从 OS 安装光盘中找到。是否安装成功将可以通过以下命令行确认

```
root>rpm -qa |grep iscsi  
  
iscsi-initiator-utils-6.2.0.868-0.7.el5
```

- (2) 在 `/etc/iscsi/` 目录下 `/etc/iscsi/initiatorname.iscsi`

查看此文件可发现主机端的 `iqn` 号码。在磁阵管理软件中添加 HOST 时需用到。

2.2 划分磁盘阵列磁盘空间

基于磁盘阵列设备手册在磁盘阵列上做好 RAID 配置后，并成功将 LUN 映射到服务器，需要根据应用需要对磁盘空间进行合理规划使用。

磁盘阵列配置步骤方法不同设备会存在差异，需要依据相应设备手册配置使用。

2.3 启动 iSCSI 设备

- (1) 在服务器端，启动 ISCSI 服务：

```
root>service iscsi start
```

- (2) 查询 ISCSI 设备 (HP storageworks mpx100) target 的 `iqn` 号码：(必须)

```
root>iscsiadm -m discovery -t sendtargets -p 192.168.1
```

4.1

```
192.168.14.1:3260,0 iqn.1986-03.com.hp:fcgw.mpx10  
0.0.50014380025bad30.50014380025bad38
```

(3) 登陆到 iSCSI 存储设备

```
root#iscsiadm -m node -T iqn.1986-03.com.hp:fcgw.mpx1  
00.0.50014380025bad30.50014380025bad38 -p 192.  
168.14.1:3260 - l
```

上面命令是连续的，其中 iqn 号码为上面查询得到的号码，ip 为 iscsi 存储中 iscsi 口对外映射的地址。

(4) 查看磁盘：

```
[root@localhost ~]# fdisk -l  
Disk /dev/sda: 8795 MB, 8795105280 bytes  
255 heads, 63 sectors/track, 1069 cylinders  
Units = cylinders of 16065 * 512 = 8225280 bytes  
  
Device Boot Start End Blocks Id System  
/dev/sda1 * 1 941 7558551 83 Linux  
/dev/sda2 942 1068 1020127+ 82 Linux swap / Solaris
```

```
Disk /dev/sdb: 10.4 GB, 10487232000 bytes  
255 heads, 63 sectors/track, 1275 cylinders  
Units = cylinders of 16065 * 512 = 8225280 bytes
```

```
Device Boot Start End Blocks Id System
```

得到新的盘符

(5) 同样的方法添加另外一个 iSCSI 路径的 target 。

(说明：一个 iSCSI 磁阵的一个 iSCSI 端口有一个 iqn 号和一个 ip 地址)

添加成功后 fdisk -l 应该可以看到两个盘符。他们的容量都一样。这两个盘符对应的是同一个磁盘阵列下的同一个 LUN 。

(6) 映射完成之后使用下面命令查看映射结果：

```
root>iscsiadm -m node  
192.168.14.1:3260,0 iqn.1986-03.com.hp:fcgw.mpx100.0.5  
0014380025bad30.50014380025bad38  
192.168.14.3:3260,0 iqn.1986-03.com.hp:fcgw.mpx100.0.5  
0014380025bad30.50014380025bad3c
```

2.4 安装 DM-multipath 软件包

由于需要实现存储设备的多路径访问及故障切换，故需要运行 multipath 服务，这个在 os 中已经安装好了。如没有执行下面的命令。Multipath 服务原理为 iSCSI 磁阵中的一个真实的 virtual disk 通过与 2 个冗余阵列控制器连接的端口对外提供映射，故服务器上可以看到两个物理磁盘，但此磁盘对应一个真实的一个真实的 virtual disk，故这两条路径间可以实现故障切换和负载均衡。

```
rpm -ivh device-mapper-1.02.13-6.9.i686.rpm
```

```
rpm -ivh multipath-tools-0.4.7-34.18.i686.rpm
```

安装完成后，使用命令 modprobe dm_multipath 来加载相应的模块，可以使用以下命令来检查安装是否正常：

```
lsmod |grep dm_multipath
```

2.5 配置文件的修改和创建

(1) /etc/multipath.conf 是多路径软件的配置文件，其中大部分配置是注释掉的，可以将他保存为备用，然后新建一个 multipath.conf 文件，编辑新的配置文件，实际有用的就是下面这些参数：其余参数都可以注释掉：

```
defaults {  
    user_friendly_names yes  
    udev_dir /dev  
    path_grouping_policy multibus  
    failback immediate  
    no_path_retry fail
```

```
}
```

保存退出

(2) 启动服务:

`Service multipathd restart` 启动成功后, 我们会在 `dev` 目录下看到下面的目录:

`/dev/mapper/mpathn,`

`/dev/mpath/mpathn,`

`/dev/dm-n.`

说明:

①`/dev/mapper/mpathn` 是软件虚拟出来的多路径设备, 这个可以被我们用来挂载使用。

②`/dev/mpath/mpathn` 这个是 `udev` 设备管理器创建的, 不能用来挂载。

③`/dev/dm-n` 这个是软件自身使用的, 不能被软件以外使用。不可挂载。

(3) 用 `multipath -ll` 命令查看到两条活跃路径, 他们之间互为 **A/A** 关系。断掉其中一根线路, 那么系统自动切换到另外一条。

```
root>multipath -ll
```

```
mpath2 (3600508b4000a5bfd0000b00000200000) dm-2 HP,HSV300
```

```
[size=2.0T][features=0][hwandler=0]
```

```
\_ round-robin 0 [prio=2][active]
```

```
\_ 1:0:0:2 sdd 8:48 [active][ready]
```

```
\_ 2:0:0:2 sdf 8:80 [active][ready]
```

```
mpath1 (3600508b4000a5bfd0000b000001a0000) dm-1 HP,HSV300
```

```
[size=2.0T][features=0][hwandler=0]
```

```
\_ round-robin 0 [prio=2][active]
```

```
\_ 1:0:0:1 sdb 8:16 [active][ready]
```

```
\_ 2:0:0:1 sdc 8:32 [active][ready]
```

```
mpath3 (3600508b4000a5bfd0000b00000350000) dm-3 HP,HSV300
```

```
[size=2.0T][features=0][hwandler=0]
\_ round-robin 0 [prio=2][active]
\_ 1:0:0:3 sde 8:64 [active][ready]
\_ 2:0:0:3 sdg 8:96 [active][ready]
```

说明：上面蓝色字体是磁阵生产厂商和产品型号标识，不同厂商和不同型号产品相关信息显示将会不同。

- (4) 在多路径设备创建后，我们就可以像使用实际的物理设备样使用多路径设备了。前提是需要通过下面的命令将其标记为物理卷。

```
pvcreate /dev/mapper/mpath1
```

- (5) 然后对磁盘进行分区和格式化

```
fdisk /dev/mapper/mpath1
```

分区之后，会在 dev/mapper/目录下创建新的块设备，

/dev/mapper/mpath1p1 这个表示 mpath1 设备下面的分区 1.

如果没有看到或者不匹配，需要使用 `multipath -F` 命令清除多路径设备的缓存信息后，再用命令 `multipath -v3` 重新进行加载。

分区之后使用 `fdisk -l` 命令查看磁盘会看到 `dm-1` 磁盘下面已经有分区的信息了。

```
*****
```

```
**
```

```
isk /dev/dm-1: 2197.9 GB, 2197949513728 bytes
```

```
255 heads, 63 sectors/track, 267218 cylinders
```

```
Units = cylinders of 16065 * 512 = 8225280 bytes
```

```
Device Boot Start End Blocks Id System
```

```
/dev/dm-1p1 1 121577 976567221 83 Linux
```

```
mkfs.ext3 /dev/mapper/mpath0p1
```

```
*****
```

- (6) 最后将磁盘 `mount` 挂载，就可以正常使用了。

(7) 将 multipathd 服务设置成开机自启动

```
chkconfig multipathd on
```

(8) 其他注意事项:

①删除失效的 iscsi 连接: 如果因为错误配置等原因创建了 iscsi 对应, 系统不会自动删除, 需手动删除。

例如:

```
[root]# iscsiadm -m node 192.168.1.51:3260,1 iqn.2006-01.com.openfiler:tsn.59dc8fc04fa2
```

我们看到上面有一个 iscsi 对应, 但实际中这个对应已经失效, 系统不会自动删除。

②删除命名:

```
iscsiadm -m node -o delete -T iqn.2006-01.com.openfiler:tsn.59dc8fc04fa2 -p 192.168.1.51:3260  
iscsiadm -m node -o delete -T iqn.2003-06.com.rocketdivision.starwind.subenhai-pc.imagefile0 -p 192.168.50.107:3260
```

③如果其中一条链路出现故障, 那么在日志里面我们可以看到如下类似的信息:

```
root@bhnfs1.baihui.com~>tail /var/log/messages
```

```
Dec 12 14:03:15 bhnfs1 kernel: iscsi: cmd 0x28 is not queued (8)
```

```
Dec 12 14:03:15 bhnfs1 kernel: iscsi: cmd 0x28 is not queued (8)
```

```
Dec 12 14:03:15 bhnfs1 multipathd: sdf: readsector0 checker reports path is down
```

```
Dec 12 14:03:15 bhnfs1 multipathd: sdg: readsector0 checker reports path is down
```

```
Dec 12 14:03:20 bhnfs1 kernel: iscsi: cmd 0x28 is not queued (8)
```

Dec 12 14:03:20 bhnfs1 multipathd: sdc: readsector0 checker reports path is down

Dec 12 14:03:20 bhnfs1 kernel: iscsi: cmd 0x28 is not queued (8)

Dec 12 14:03:20 bhnfs1 multipathd: sdf: readsector0 checker reports path is down

Dec 12 14:03:20 bhnfs1 kernel: iscsi: cmd 0x28 is not queued (8)

Dec 12 14:03:20 bhnfs1 multipathd: sdg: readsector0 checker reports path is down

查看多链路状态:

```
root@bhnfs1.baihui.com/etc> multipath -ll
```

```
sdc: checker msg is "readsector0 checker reports path is down"
```

```
sdf: checker msg is "readsector0 checker reports path is down"
```

```
sdg: checker msg is "readsector0 checker reports path is down"
```

```
mpath2 (3600508b4000a5bfd0000b00000200000) dm-2 HP,HSV300
```

```
[size=2.0T][features=0][hw_handler=0]
```

```
  \_ round-robin 0 [prio=1][active]
```

```
    \_ 1:0:0:2 sdd 8:48 [active][ready]
```

```
    \_ 2:0:0:2 sdf 8:80 [failed][faulty]
```

```
mpath1 (3600508b4000a5bfd0000b000001a0000) dm-1 HP,HSV300
```

```
[size=2.0T][features=0][hw_handler=0]
```

```
  \_ round-robin 0 [prio=1][active]
```

```
    \_ 1:0:0:1 sdb 8:16 [active][ready]
```



```
\_ 2:0:0:1 sdc 8:32 [failed][faulty]
mpath3 (3600508b4000a5bfd0000b00000350000) dm-3 HP,
HSV300
[size=2.0T][features=0][hwhandler=0]
\_ round-robin 0 [prio=1][active]
\_ 1:0:0:3 sde 8:64 [active][ready]
\_ 2:0:0:3 sdg 8:96 [failed][faulty]
```

三. 总结

随着 iSCSI 存储技术的发展和普及，iSCSI 存储产品在电信行业应用越来越多，掌握 iSCSI 存储在不同应用环境下的配置和使用非常必要，而 Linux 提供的多路径方案作为一种开放的应用解决方案也为用户应用方案配套提供了更灵活的选择。

机器重启后 LVM 和 LUN 识别问题：

1、修改initrd初始化模块队列，添加FC驱动模块：

```
linux:~ # grep MOD /etc/sysconfig/kernel
INITRD_MODULES="sata_nv mptsas reiserfs qla2xxx qla2400"
MODULES_LOADED_ON_BOOT=""
```

2、重新制作initrd：

```
# cd /boot/
linux:/boot # cp initrd-2.6.5-7.244-smp initrd-2.6.5-7.244-smp-
bak20090416
linux:/boot # mkinitrd -k vmlinuz-2.6.5-7.244-smp -i initrd-2.6.5-
7.244-smp
```

3、重启系统自动识别LVM等LUN设备：

```
linux:/boot # reboot
```