

Linux 系统下 LV 的磁盘故障与修复

一【实验目标】

- 学习并掌握 Linux 系统下创建 LV

二【实验环境】

- 实验机环境：Centos 6.6
- 目标机环境：Centos 6.6
- 实验拓扑：如图 1 所示。



图 1 实验拓扑

三【实验原理】

在 Linux 的 LAM 管理模式下，当磁盘出现故障时需要进行修复。

四【实验步骤】

- 1、新加磁盘启动系统后，查看现有磁盘使用情况

(1)命令：df -h

```
[root@localhost yangbin]# df -h
Filesystem      Size  Used Avail Use% Mounted on
/dev/sda1        9.5G  3.4G  5.7G  37% /
tmpfs            491M   80K  491M   1% /dev/shm
```

图 2

(2)命令：fdisk -l

```
[root@localhost yangbin]# fdisk -l
```

```
Disk /dev/sda: 21.5 GB, 21474836480 bytes
255 heads, 63 sectors/track, 2610 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x0007475b
```

Device	Boot	Start	End	Blocks	Id	System
/dev/sda1	*	1	1275	10240000	83	Linux
/dev/sda2		1275	1537	2097152	82	Linux swap / Solaris

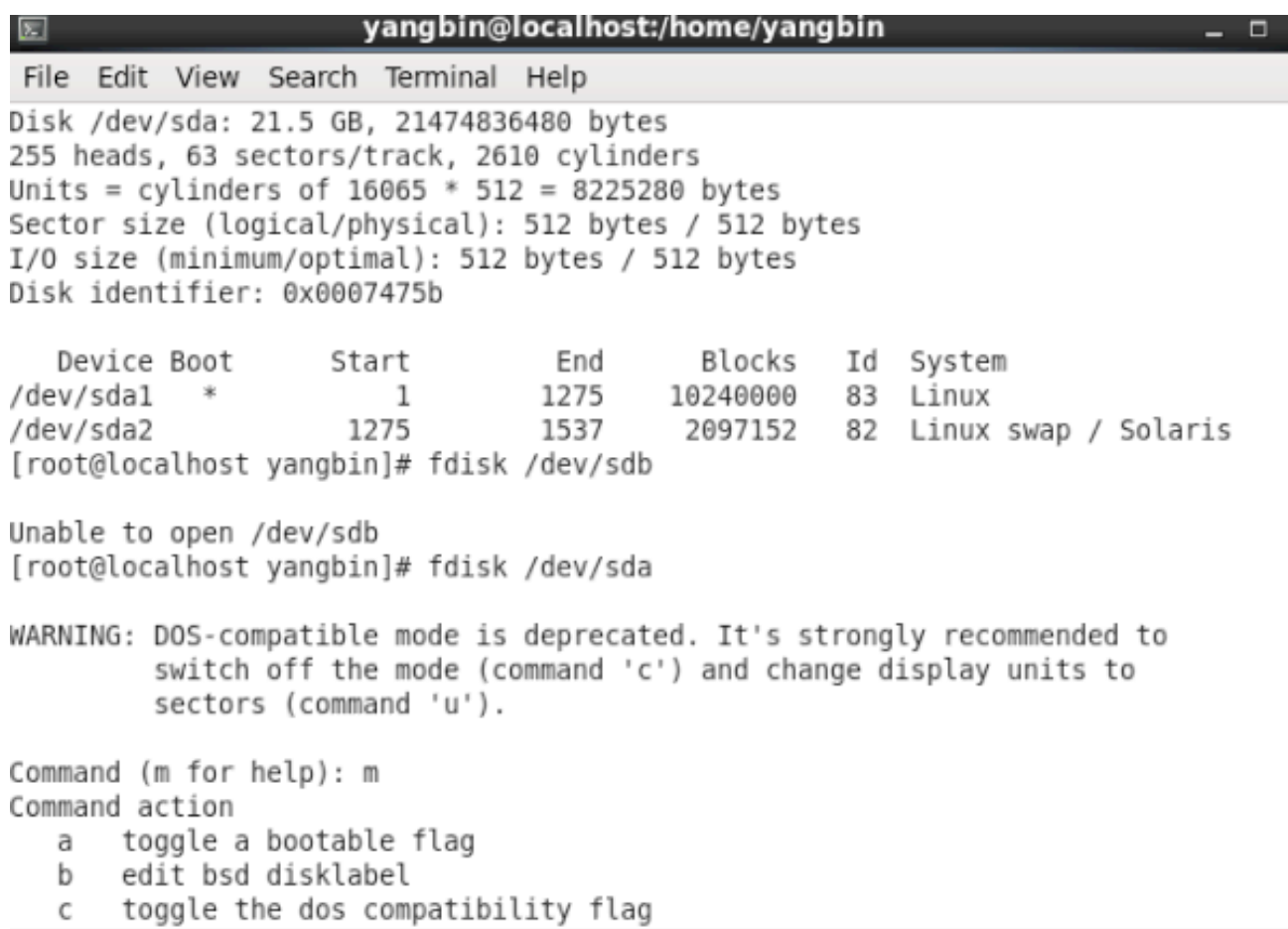
```
[root@localhost yangbin]#
```

图 3

2、 对磁盘 sda 进行分区

如果不知道 fdisk 里面的具体操作，可输入 m 进行帮助。最常用的是 n（新建）d（删除）p（打印）q（退出）t（修改系统标识符）w（写入并退出）。

命令：fdisk /dev/sda



```
yangbin@localhost:/home/yangbin
File Edit View Search Terminal Help
Disk /dev/sda: 21.5 GB, 21474836480 bytes
255 heads, 63 sectors/track, 2610 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x0007475b

   Device Boot      Start         End      Blocks   Id  System
/dev/sda1  *          1         1275     1024000    83   Linux
/dev/sda2             1275         1537      2097152    82   Linux swap / Solaris
[root@localhost yangbin]# fdisk /dev/sdb

Unable to open /dev/sdb
[root@localhost yangbin]# fdisk /dev/sda

WARNING: DOS-compatible mode is deprecated. It's strongly recommended to
switch off the mode (command 'c') and change display units to
sectors (command 'u').

Command (m for help): m
Command action
  a   toggle a bootable flag
  b   edit bsd disklabel
  c   toggle the dos compatibility flag
```

图 4

(1) 建立新的分区

输入 **p** 打印现有分区情况（还没有分区）

输入 **n** 新建分区

输入 **p** 为建立主分区（此时的 p 是在 n 后的，不是打印）

输入 **1** 为建立第一个主分区

分区起始位置可以直接回车，默认是 1

输入 **p** 打印分区情况，发现已建立一个分区 **/dev/sda3**，但是 此分区为 Linux 格式

```

yangbin@localhost:/home/yangbin
File Edit View Search Terminal Help
Command (m for help): n
Command action
  e   extended
  p   primary partition (1-4)
3
Invalid partition number for type `3'
Command action
  e   extended
  p   primary partition (1-4)
p
Partition number (1-4): 3
First cylinder (1537-2610, default 1537): 1537
Last cylinder, +cylinders or +size{K,M,G} (1537-2610, default 2610): 2200

Command (m for help): p

Disk /dev/sda: 21.5 GB, 21474836480 bytes
255 heads, 63 sectors/track, 2610 cylinders
Units = cylinders of 16065 * 512 = 8225280 bytes
Sector size (logical/physical): 512 bytes / 512 bytes
I/O size (minimum/optimal): 512 bytes / 512 bytes
Disk identifier: 0x0007475b

   Device Boot      Start         End      Blocks   Id  System

```

图 5

Disk identifier: 0x0007475b

Device	Boot	Start	End	Blocks	Id	System
/dev/sda1	*	1	1275	10240000	83	Linux
/dev/sda2		1275	1537	2097152	82	Linux swap / Solaris
/dev/sda3		1537	2200	5333324	83	Linux

图 6

(2)改变系统标识符:

输入 t 改变分区 1 的属性

输入 L 查看有个属性对应的命令

输入 8e 改变分区 1 为 Linux LVM 格式

输入 p 打印分区情况，发现建立的分区 /dev/sda3 为 Linux LVM 格式

tip:再次使用 **fdisk -l** 查看系统内磁盘情况发现 /dev/sda 上已有一个 Linux LVM 格式的 /dev/sda3 分

区

```
yangbin@localhost:/home/yangbin
File Edit View Search Terminal Help
Command (m for help): t
Partition number (1-4): 3
Hex code (type L to list codes): l

0 Empty          24 NEC DOS       81 Minix / old Lin bf Solaris
1 FAT12          39 Plan 9        82 Linux swap / So c1 DRDOS/sec (FAT-
2 XENIX root     3c PartitionMagic 83 Linux           c4 DRDOS/sec (FAT-
3 XENIX usr      40 Venix 80286    84 OS/2 hidden C:  c6 DRDOS/sec (FAT-
4 FAT16 <32M     41 PPC PReP Boot  85 Linux extended  c7 Syrix
5 Extended       42 SFS           86 NTFS volume set da Non-FS data
6 FAT16          4d QNX4.x         87 NTFS volume set db CP/M / CTOS / .
7 HPFS/NTFS      4e QNX4.x 2nd part 88 Linux plaintext de Dell Utility
8 AIX            4f QNX4.x 3rd part 8e Linux LVM       df BootIt
9 AIX bootable   50 OnTrack DM     93 Amoebe          e1 DOS access
a OS/2 Boot Manag 51 OnTrack DM6 Aux  94 Amoebe BBT      e3 DOS R/O
b W95 FAT32       52 CP/M          9f BSD/OS          e4 SpeedStor
c W95 FAT32 (LBA) 53 OnTrack DM6 Aux a0 IBM Thinkpad hi eb BeOS fs
e W95 FAT16 (LBA) 54 OnTrackDM6     a5 FreeBSD        ee GPT
f W95 Ext'd (LBA) 55 EZ-Drive      a6 OpenBSD        ef EFI (FAT-12/16/
10 OPUS          56 Golden Bow    a7 NeXTSTEP       f0 Linux/PA-RISC b
11 Hidden FAT12   5c Priam Edisk   a8 Darwin UFS     f1 SpeedStor
12 Compaq diagnost 61 SpeedStor     a9 NetBSD         f4 SpeedStor
14 Hidden FAT16 <3 63 GNU HURD or Sys ab Darwin boot    f2 DOS secondary
16 Hidden FAT16   64 Novell Netware af HFS / HFS+      fb VMWare VMFS

16 Hidden FAT16   64 Novell Netware af HFS / HFS+      fb VMWare VMFS
17 Hidden HPFS/NTF 65 Novell Netware b7 BSDI fs        fc VMWare VMKCORE
18 AST SmartSleep  70 DiskSecure Mult b8 BSDI swap      fd Linux raid auto
1b Hidden W95 FAT3 75 PC/IX          bb Boot Wizard hid fe LANstep
1c Hidden W95 FAT3 80 Old Minix       be Solaris boot   ff BBT
1e Hidden W95 FAT1

Hex code (type L to list codes): 8e
Changed system type of partition 3 to 8e (Linux LVM)

Command (m for help): w
The partition table has been altered!

Calling ioctl() to re-read partition table.

WARNING: Re-reading the partition table failed with error 16: Device or resource
busy.
The kernel still uses the old table. The new table will be used at
the next reboot or after you run partprobe(8) or kpartx(8)
Syncing disks.
.. .. .
```

图 7

(3)使 kernel 重新读取分区表

命令: **partprobe**

对于 /dev/sda 的警告不予理会

```
[root@localhost yangbin]# partprobe
Warning: WARNING: the kernel failed to re-read the partition table on /dev/sda (
Device or resource busy). As a result, it may not reflect all of your changes u
ntil after reboot.
[root@localhost yangbin]#
```

图 8

3、 创建 PV:

扫描系统 PV: **pvscan**

创建 PV: **pvcreate /dev/sdb1**

查看 PV: **pvdisplay**

这样我们就创建了一个 5.09G 的 PV，注意 Allocatable 为 NO

```
[root@localhost yangbin]# pvcreate /dev/sda3
Physical volume "/dev/sda3" successfully created
[root@localhost yangbin]# pvdisplay
"/dev/sda3" is a new physical volume of "5.09 GiB"
--- NEW Physical volume ---
PV Name                /dev/sda3
VG Name
PV Size                5.09 GiB
Allocatable            NO
PE Size                0
Total PE               0
Free PE                0
Allocated PE           0
PV UUID                ruWsoH-HldF-jzaJ-az5S-l10U-ehzb-YEZmft

[root@localhost yangbin]#
```

图 9

4、 创建 VG:

扫描系统 VG: **vgscan**

创建 VG: **vgcreate vg_test /dev/sdb1**

查看 VG: **vgdisplay**

这样我们就创建了一个 5.09G（1301 个 PE，要记住这个数字）的 VG（名字为 **vg_test**）

```
[root@localhost yangbin]# vgcreate vg_test /dev/sda3
Volume group "vg_test" successfully created
[root@localhost yangbin]# vgsdisplay
--- Volume group ---
VG Name                vg_test
System ID
Format                 lvm2
Metadata Areas         1
Metadata Sequence No   1
VG Access               read/write
VG Status               resizable
MAX LV                 0
Cur LV                 0
Open LV                0
Max PV                 0
Cur PV                 1
Act PV                 1
VG Size                 5.08 GiB
PE Size                 4.00 MiB
Total PE                1301
Alloc PE / Size         0 / 0
Free PE / Size          1301 / 5.08 GiB
VG UUID                kl6Lqw-Tpyc-kjsj-E88H-7AoV-cMtZ-NiMCuf
```

图 10

5、 创建 LV:

扫描系统 LV: `lvscan`

创建 LV: `lvcreate -l 1301 -n lv_test vg_test` (1301 是 VG 中 PE 的个数)

查看 LV: `lvdisplay`

这样我们就创建了一个名字为 `lv_test` 的 LV

```
[root@localhost yangbin]# lvcreate -l 1301 -n lv_test vg_test
Logical volume "lv_test" created
[root@localhost yangbin]# lvsdisplay
--- Logical volume ---
LV Path                /dev/vg_test/lv_test
LV Name                 lv_test
VG Name                 vg_test
LV UUID                 Yot8QP-eKUE-x9yD-auho-T4o4-ha2q-XLQz0Q
LV Write Access         read/write
LV Creation host, time localhost.localdomain, 2016-04-15 05:16:59 +0800
LV Status                available
# open                  0
LV Size                 5.08 GiB
Current LE              1301
Segments                1
Allocation              inherit
Read ahead sectors      auto
- currently set to     256
Block device            253:0

[root@localhost yangbin]# █
```

图 12

6、 格式化 LV:

命令: `mkfs -t ext4 /dev/vg_test/lv_test`

```
yangbin@localhost:/home/yangbin
File Edit View Search Terminal Help
[root@localhost yangbin]# mkfs -t ext4 /dev/vg_test/lv_test
mke2fs 1.41.12 (17-May-2010)
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
333248 inodes, 1332224 blocks
66611 blocks (5.00%) reserved for the super user
First data block=0
Maximum filesystem blocks=1367343104
41 block groups
32768 blocks per group, 32768 fragments per group
8128 inodes per group
Superblock backups stored on blocks:
    32768, 98304, 163840, 229376, 294912, 819200, 884736

Writing inode tables: done
Creating journal (32768 blocks): done
Writing superblocks and filesystem accounting information: done

This filesystem will be automatically checked every 35 mounts or
180 days, whichever comes first.  Use tune2fs -c or -i to override.
[root@localhost yangbin]#
```

7 挂载目录并创建:

命令: `mkdir /test`

`mount /dev/vg_test/lv_test /test`

```
[root@localhost yangbin]# mkdir /test
[root@localhost yangbin]# mount /dev/vg_test/lv_test/test
mount: can't find /dev/vg_test/lv_test/test in /etc/fstab or /etc/mtab
[root@localhost yangbin]# mount /dev/vg_test/lv_test /test
[root@localhost yangbin]# df -h
Filesystem              Size  Used Avail Use% Mounted on
/dev/sdal                9.5G  3.3G  5.8G  37% /
tmpfs                   491M   80K  491M   1% /dev/shm
/dev/mapper/vg_test-lv_test
                        4.9G   11M  4.7G   1% /test
[root@localhost yangbin]#
```

8 设置开机挂载

命令: `vim /etc/fstab`

加入: `/dev/mapper/vg_test-lv_test /test ext4 defaults 1 2`


```
yangbm@localhost:~/home/yangbm
File Edit View Search Terminal Help

#
# /etc/fstab
# Created by anaconda on Fri Apr 15 23:43:17 2016
#
# Accessible filesystems, by reference, are maintained under '/dev/disk'
# See man pages fstab(5), findfs(8), mount(8) and/or blkid(8) for more info
#
JUID=26d5b60f-30ce-430b-8224-f5b1bb78cab5 / ext4 defau
lts 1 1
JUID=3dac0d69-8051-41e5-98d1-a1685ba2847b swap swap defau
lts 0 0
tmpfs /dev/shm tmpfs defaults 0 0
devpts /dev/pts devpts gid=5,mode=620 0 0
sysfs /sys sysfs defaults 0 0
proc /proc proc defaults 0 0
/dev/mapper/vg_test-lv_test /test ext4 defaults 1 2
~
~
~
~
~
-- INSERT -- 15,79 All
```

五【LAM 灾难修复】

1 PV 损坏与替换

模拟错误并进行恢复操作

(1) 检查磁盘及文件系统状况

命令: pvs

vgs

lvs

lvs -o +devices

mount |grep '/dev/mapper'

```

[root@localhost yangbin]# pvs
PV          VG          Fmt  Attr  PSize  PFree
/dev/sda3  vg_test  lvm2  a--   5.08g    0
[root@localhost yangbin]# vgs
VG          #PV #LV #SN Attr   VSize VFree
vg_test     1   1   0 wz--n- 5.08g    0
[root@localhost yangbin]# lvs
LV          VG          Attr      LSize Pool Origin Data%  Meta%  Move Log Cpy%Sync
Convert
lv_test    vg_test  -wi-ao---- 5.08g

[root@localhost yangbin]# lvs -o +devices
LV          VG          Attr      LSize Pool Origin Data%  Meta%  Move Log Cpy%Sync
Convert Devices
lv_test    vg_test  -wi-ao---- 5.08g
           /dev/sda3(0)
[root@localhost yangbin]# mount |grep '/dev/mapper'
/dev/mapper/vg_test-lv_test on /test type ext4 (rw)
[root@localhost yangbin]# █

```

(2) 备份文件

命令: `/etc/lvm/backup:# cp * /testback/`

(4) 如若文件系统损坏, 采用故障卷替换的方法。

命令: `pvccreate --restorefile /etc/lvm/backup/test \n`

(3) 同步 metadata

命令: `vgchange -an test`

(5) 修复文件

```

命令: reiserfsck /dev/test/lv0 --check
      reiserfsck /dev/test/lv0 --rebuild-sb
      reiserfsck /dev/test/lv0 --check
      reiserfsck /dev/test/lv0 --rebuild-tree
      reiserfsck /dev/test/lv0 - - check

```

2 通过备份修复文件

(1) 提示操作错误的情况类似:

```

# vgscan
Reading all physical volumes. This may take a while...
WARNING: Inconsistent metadata found for VG vg_test - updating to use version 18
Removing PV /dev/sdc (DHmMDP-bqQy-TalG-2GLa-sh6o-fyVW-3XQ3gp) that no longer
belongs to VG vg_test

Found volume group "vg_test" using metadata type lvm2
#pvs
PV          VG          Fmt  Attr  PSize    PFree
/dev/sdb    vg_test  lvm2  a-    200.00m 100.00m
/dev/sdc                    lvm2  --    204.00m 204.00m

```

(2) 通过恢复配置还原:

```
# vgcfgrestore -f /etc/lvm/archive/vg_test_01564.vg vg_test
Cannot restore Volume Group vg_test with 1 PVs marked as missing.
Restore failed.
```

(3) 手动修改配置

```
...
    pv1 {
        id = "DHmMDP-bqQy-TalG-2GLa-sh6o-fyVW-3XQ3gp"
        device = "unknown device"
        flags = ["MISSING"]
    }
...
```

(4) 再次恢复

```
# vgcfgrestore -f vg_test_edited.vg vg_test
Restored volume group vg_test
```

六【实验思考】

LAM 故障的类别有很多，解决的方式也很多但是恢复的话很多时候数据并不能完全恢复，所以做好备份以及防患于未然是非常重要的。