



(12) 发明专利申请

(10) 申请公布号 CN 105242877 A

(43) 申请公布日 2016.01.13

(21) 申请号 201510630416.6

(22) 申请日 2015.09.29

(71) 申请人 浪潮电子信息产业股份有限公司
地址 250101 山东省济南市高新区浪潮路
1036 号

(72) 发明人 白云峰 侯昌帅

(74) 专利代理机构 济南信达专利事务有限公司
37100

代理人 姜明

(51) Int. Cl.

G06F 3/06(2006.01)

权利要求书4页 说明书5页

(54) 发明名称

一种 Linux 下硬件 RAID 管理方法

(57) 摘要

本发明提供一种 Linux 下硬件 RAID 管理方法,涉及 RAID/Linux 技术领域,本发明主要从三个方面进行 RAID 组建管理及维护:首先,通过查看物理磁盘信息及客户需求提出 RAID 组建策略,包括 WT/WB/NORA/RA/ADRA/Cached 和 Direct 设置;其次,通过脚本创建/删除 RAID,可以创建包括 RAID0/1/5/6/10 等几种不同形态,同时可以指定热备盘等相关备用盘符,阵列创建完后,会有一个初始化同步块的过程,之后完成 RAID 组建;最后,磁盘信息查看及维护策略,包括磁盘使用占比及当发生掉盘后的维护。该方案在 RACK 生产实践中可大大降低因硬盘故障需而重建 RAID 人力及时间,不必耗时耗力在 WEBBIOS 下通过手动一步一步操作建立,仅需 Linux 下几条命令可自动重建,大大提高了生产效率。

1. 一种 Linux 下硬件 RAID 管理方法,其特征在於,主要从三个方面进行 RAID 组建管理及维护:

首先,通过查看物理磁盘信息及客户需求提出 RAID 组建策略,包括 WT/WB/NORA/RA/ADRA/Cached 和 Direct 设置;

其次,通过创建 / 删除 RAID,同时可以指定热备盘相关备用盘符,阵列创建完后,会有一个初始化同步块的过程,之后完成 RAID 组建;

最后,磁盘信息查看及维护策略,包括磁盘使用占比及当发生掉盘后的维护。

2. 根据权利要求 1 所述的方法,其特征在於,

RAID 组建策略

通过查看所有物理磁盘信息

```
MegaCli -PDList -aALL Adapter #0 Enclosure Number: 1 Slot Number: 5 Device
Id: 5 Sequence Number: 2 Media Error Count: 0 Other Error Count: 0 Predictive
Failure Count: 0 Last Predictive Failure Event Seq Number: 0 Raw Size: 140014MB
[0x11177328 Sectors] Non Coerced Size: 139502MB [0x11077328 Sectors] Coerced
Size: 139392MB [0x11040000 Sectors] Firmware state: Hotspare SAS Address(0):
0x5000c50008e5cca9 SAS Address(1): 0x0 Inquiry Data: SEAGATE ST3146855SS
S5273LN4Y1X0 .....
```

查看磁盘缓存策略

```
MegaCli -LDGetProp -Cache -L0 -a0 Adapter 0-VD 0: Cache Policy:WriteBack,
ReadAheadNone, Direct or
```

```
MegaCli -LDGetProp -Cache -L1 -a0 Adapter 0-VD 1: Cache Policy:WriteBack,
ReadAheadNone, Direct or
```

```
MegaCli -LDGetProp -Cache -LALL -a0 Adapter 0-VD 0: Cache Policy:WriteBack,
ReadAheadNone, Direct Adapter 0-VD 1: Cache Policy:WriteBack, ReadAheadNone,
Direct or
```

```
MegaCli -LDGetProp -Cache -LALL -aALL Adapter 0-VD 0: Cache
Policy:WriteBack, ReadAheadNone, Direct Adapter 0-VD 1: Cache
Policy:WriteBack, ReadAheadNone, Direct or
```

```
MegaCli -LDGetProp -DskCache -LALL -aALL Adapter 0-VD 0: Disk Write Cache
: Disk's Default Adapter 0-VD 1: Disk Write Cache : Disk's Default
```

设置磁盘缓存策略

磁盘缓存设置包括 WT /WB/NORA/RA/ADRA/Cached/Direct ;具体设置为:

```
MegaCli -LDSetProp WT|WB|NORA|RA|ADRA -L0 -a0 or
```

```
MegaCli -LDSetProp -Cached|-Direct -L0 -a0 or
```

enable / disable disk cache

```
MegaCli -LDSetProp -EnDskCache|-DisDskCache -L0 -a0。
```

3. 根据权利要求 1 所述的方法,其特征在於,

创建 RAID

在 Redhat 和 CentOS 系统下执行 rpm -ivh xxxx.rpm 来安装

在 Ubuntu 下应该执行 `sudo apt-get install packagename` 来安装

1) 安装库文件：

```
#alien Lib_Utills-1.00-09.noarch.rpm
```

```
#dpkg -i lib-utills_1.00-10_all.deb
```

使用 `alien` 命令把 `Lib_Utills-....noarch.rpm` 转换为 `deb` 包, 使用 `dpkg` 命令安装；

2) 解压 MegaCLI:

```
#rpm2cpio MegaCli-8.02.16-1.i386.rpm | cpio -idmv
```

安装完成后会出现 `opt/` 目录, 可以根据不同系统使用如下命令查看配置信息.

```
./opt/MegaRAID/MegaCli/MegaCli -AdpAllInfo -aALL
```

```
./opt/MegaRAID/MegaCli/MegaCli64 -AdpAllInfo -aALL
```

3) 创建阵列, 以 RAID5 为例: 由物理盘 2, 3, 4 构成, 该阵列的热备盘是物理盘 5

```
MegaCli -CfgLdAdd -r5 [1:2,1:3,1:4] WB Direct -Hsp[1:5] -a0
```

创建阵列, 不指定热备

```
MegaCli -CfgLdAdd -r5 [1:2,1:3,1:4] WB Direct -a0
```

删除阵列

```
MegaCli -CfgLdDel -L1 -a0
```

4) 在线添加磁盘

```
MegaCli -LDRecon -Start -r5 -Add -PhysDrv[1:4] -L1 -a0
```

5) 查看阵列初始化信息

阵列创建完后, 会有一个初始化同步块的过程, 可以看看其进度；

```
MegaCli -LDInit -ShowProg -LALL -aALL
```

或者以动态可视化文字界面显示

```
MegaCli -LDInit -ProgDsply -LALL -aALL
```

查看阵列后台初始化进度

```
MegaCli -LDBI -ShowProg -LALL -aALL
```

或者以动态可视化文字界面显示

```
MegaCli -LDBI -ProgDsply -LALL -aALL
```

6) 创建全局热备, 指定第 5 块盘作为全局热备：

```
MegaCli -PDHSP -Set [-EnclAffinity] [-nonRevertible] -PhysDrv[1:5] -a0
```

也可以指定为某个阵列的专用热备：

```
MegaCli -PDHSP -Set [-Dedicated [-Array1]] [-EnclAffinity] [-nonRevertible] -PhysDrv[1:5] -a0
```

7) 删除全局热备：

```
MegaCli -PDHSP -Rmv -PhysDrv[1:5] -a0
```

8) 将某块物理盘下线 / 上线：

```
MegaCli -PDOffline -PhysDrv [1:4] -a0
```

```
MegaCli -PDOnline -PhysDrv [1:4] -a0
```

9) 查看物理磁盘重建进度：

```
MegaCli -PDRbld -ShowProg -PhysDrv [1:5] -a0
```

或者以动态可视化文字界面显示：

```
MegaCli -PDRbld -ProgDsply -PhysDrv [1:5] -a0。
```

4. 根据权利要求 1 所述的方法,其特征在于,

RAID 维护策略

当发生掉盘后,应该更换新的硬盘,更换硬盘后,在 SAS 卡 2008 下会默认为 JBOD 模式,或者硬盘本身带有 foreign 信息,此时应该执行命令把它转换为 Unconfigure good 模式并清除 foreign 信息再做 Raid;

1) 首先查找到该卡子的 Dvice ID,记录

```
#!/MegaCli64 -PDlist -aALL | grep "Dvice ID"
```

```
Enclosure Device ID: 64,
```

记录该 En Device ID,如果由两个卡子可以分别执行参数 -a0 和 -a1 来取代 -aALL , 查找看有没有 foreign 信息 #./MegaCli64 -PDlist -aALL | grep "Foreign State" 得到如下信息:

```
Foreign State: None
```

2) 在执行该命令前应该看下 slot 号是否按顺序排列

#./MegaCli64 -PDlist -aALL | grep "Slot",如果不按顺序则应找到该硬盘对应的 Slot 号,经查看如上的信息按 0-7 顺序排列;该信息显示 slot0-7 的 foreign 信息,可以看到 slot1 的硬盘状态存在 foreign

3) 修改硬盘状态信息为 unconfigue good

把改 Dvice ID 64 卡子的 slot 1 状态改为 unconfigue good

```
#!/MegaCli64 -PDMakeGood -PhysDrv[64:1] -a0
```

```
Adapter: 0: EnclId-64 SlotId-1 state changed to Unconfigured-Good. Exit Code: 0x00
```

4) 清除 foreign 信息 扫描该卡子上的硬盘看是否存在 foreign 信息

```
./MegaCli64 -CfgForeign -Scan -a0
```

不存在就不用再执行下面的命令了;如果在 a0 上存在则执行,清除

```
#!/MegaCli64 -CfgForeign -Clear [a0]
```

```
Foreign configuration 0 is cleared on controller 0. Exit Code: 0x00
```

5) 重新做 Raid

En DeviceID64 的 slot1 的单盘做 Raid0,则

```
#!/MegaCli64 -CfgLdAdd-r0 [64:1] WT Direct -a0
```

```
Adapter 0: Created VD 1
```

```
Adapter 0: Configured the Adapter!!
```

```
Exit Code: 0x00
```

在该卡子上 slot4 和 slot5 制作 Raid1

```
./MegaCli64 -CfgLdAdd-r1 [64:4,64:5] WT Direct -a0
```

```
Adapter 0: Created VD 4
```

```
Adapter 0: Configured the Adapter!!
```

Exit Code: 0x00

当 Raid1 掉线后更换新硬盘,使用参数 -PDMakeGood -PhysDrv[64:4] 把硬盘配置为 unconfigue good 后,拔出该盘,重新插入,则 Raid 卡会自动处于 Rebuild状态;或者得到掉线的 PD 信息

```
#./MegaCli64 -Pdgetmissing -a0
```

Adapter 0 - Missing Physical drives Exit Code: 0x00

替代该掉线硬盘

#MegaCli64 -PdReplaceMissing -physdrv[64:4] -array4, -row0 -a0 重启后将自动重建。

一种 Linux 下硬件 RAID 管理方法

技术领域

[0001] 本发明涉及 RAID/Linux 技术领域,尤其涉及一种 Linux 下硬件 RAID 管理方法。

背景技术

[0002] 磁盘阵列(Redundant Arrays of Independent Disks, RAID)是由很多价格较便宜的磁盘,组合成一个容量巨大的磁盘组,利用个别磁盘提供数据所产生加成效果提升整个磁盘系统效能。利用这项技术,将数据切割成许多区段,分别存放在各个硬盘上。同时磁盘阵列还能利用同位检查(Parity Check)的观念,在数组中任意一个硬盘故障时,仍可读出数据,在数据重构时,将数据经计算后重新置入新硬盘中。

[0003] 传统 RAID 配置需要如下步骤:

开机按 <Ctrl><H> 进入到 LSI 阵列卡的图形界面 > 进入到阵列卡的 WEBBIOS 界面 > 进行 RAID 的设置 > 选择 New configuration > Manual configuration > Redundancy when possible > 选择磁盘并 Add To Array > Accept DG > Next > 建立的 RAID 级别(包括条带及写策略) > Accept > Yes > rebeoot

上述步骤不仅繁琐,而且容易出错,耗时耗力不能保证机器完整正确设置。

发明内容

[0004] 为了解决以上问题,本文提出了一种 Linux 下硬件 RAID 管理方法,采用 Megacli 管理软件策略可以在 Linux 系统下便捷的管理维护 RAID,效率大大的提升,适合大规模的机器部署,减少人工操作出错的概率。

[0005] 本发明提出了一种 Linux 下硬件 RAID 管理方法. 主要从三个方面进行 RAID 组建及管理维护: 首先通过查看物理磁盘信息及客户需求提出 RAID 组建策略,可以根据用户资源如硬盘形态数量及缓冲等设置选项调优 RAID,综合出最优 RAID 组建方案。其次,通过创建/删除 RAID,可以创建包括 RAID0/1/5/6/10 等几种不同形态,其中 RAID0 为非冗余高性能存储方式,RAID1 模式为数据同步备份容量减半存方式,同时可以指定热备盘等备用盘符便于后期重建,阵列创建完后,会有一个初始化同步块的过程,可直观观察并完成 RAID 组建;最后是磁盘信息查看及维护策略,包括磁盘使用占比及当发生掉盘后的重建及维护。

[0006] RAID 组建策略

通过查看所有物理磁盘信息

```
MegaCli -PDList -aALL
```

```
Adapter #0 Enclosure Number: 1 Slot Number: 5 Device Id: 5 Sequence  
Number: 2 Media Error Count: 0 Other Error Count: 0 Predictive Failure  
Count: 0 Last Predictive Failure Event Seq Number: 0 Raw Size: 140014MB  
[0x11177328 Sectors] Non Coerced Size: 139502MB [0x11077328 Sectors] Coerced  
Size: 139392MB [0x11040000 Sectors] Firmware state: Hotspare SAS Address(0):  
0x5000c50008e5cca9 SAS Address(1): 0x0 Inquiry Data: SEAGATE ST3146855SS
```

S5273LN4Y1X0

查看磁盘缓存策略

```
MegaCli -LDGetProp -Cache -L0 -a0 Adapter 0-VD 0: Cache Policy:WriteBack,
ReadAheadNone, Direct or
```

```
MegaCli -LDGetProp -Cache -L1 -a0 Adapter 0-VD 1: Cache Policy:WriteBack,
ReadAheadNone, Direct or
```

```
MegaCli -LDGetProp -Cache -LALL -a0 Adapter 0-VD 0: Cache Policy:WriteBack,
ReadAheadNone, Direct Adapter 0-VD 1: Cache Policy:WriteBack, ReadAheadNone,
Direct or
```

```
MegaCli -LDGetProp -Cache -LALL -aALL Adapter 0-VD 0: Cache
Policy:WriteBack, ReadAheadNone, Direct Adapter 0-VD 1: Cache
Policy:WriteBack, ReadAheadNone, Direct or
```

```
MegaCli -LDGetProp -DskCache -LALL -aALL Adapter 0-VD 0: Disk Write Cache
: Disk's Default Adapter 0-VD 1: Disk Write Cache : Disk's Default
```

设置磁盘缓存策略

磁盘缓存设置包括 WT (Write through)/WB(Write back)/NORA (No read ahead)/RA (Read ahead)/ADRA (Adaptive read ahead)/Cached/Direct 等。具体设置见下：

```
MegaCli -LDSetProp WT|WB|NORA|RA|ADRA -L0 -a0 or
```

```
MegaCli -LDSetProp -Cached|-Direct -L0 -a0 or
```

enable / disable disk cache

```
MegaCli -LDSetProp -EnDskCache|-DisDskCache -L0 -a0
```

创建 RAID

在 Redhat 和 CentOS 等系统下执行 `rpm -ivh xxxx.rpm` 来安装

在 Ubuntu 下应该执行 `sudo apt-get install packagename` 来安装

1) 安装库文件：

```
#alien Lib_Utils-1.00-09.noarch.rpm
```

```
#dpkg -i lib-utils_1.00-10_all.deb
```

使用 alien 命令把 Lib_Utils-....noarch.rpm 转换为 deb 包，使用 dpkg 命令安装。

[0007] 2) 解压 MegaCLI:

```
#rpm2cpio MegaCli-8.02.16-1.i386.rpm | cpio -idmv
```

安装完成后会出现 opt/ 目录,可以根据不同系统使用如下命令查看配置信息.

```
./opt/MegaRAID/MegaCli/MegaCli -AdpAllInfo -aALL
```

```
./opt/MegaRAID/MegaCli/MegaCli64 -AdpAllInfo -aALL
```

3) 创建阵列,以 RAID5 为例:由物理盘 2,3,4 构成,该阵列的热备盘是物理盘 5

```
MegaCli -CfgLdAdd -r5 [1:2,1:3,1:4] WB Direct -Hsp[1:5] -a0
```

创建阵列,不指定热备

```
MegaCli -CfgLdAdd -r5 [1:2,1:3,1:4] WB Direct -a0
```

删除阵列

```
MegaCli -CfgLdDel -L1 -a0
```

4) 在线添加磁盘

```
MegaCli -LDRecon -Start -r5 -Add -PhysDrv[1:4] -L1 -a0
```

5) 查看阵列初始化信息

阵列创建完后,会有一个初始化同步块的过程,可以看看其进度。

```
[0008] MegaCli -LDInit -ShowProg -LALL -aALL
```

或者以动态可视化文字界面显示

```
MegaCli -LDInit -ProgDsply -LALL -aALL
```

查看阵列后台初始化进度

```
MegaCli -LDBI -ShowProg -LALL -aALL
```

或者以动态可视化文字界面显示

```
MegaCli -LDBI -ProgDsply -LALL -aALL
```

6) 创建全局热备,指定第 5 块盘作为全局热备:

```
MegaCli -PDHSP -Set [-EnclAffinity] [-nonRevertible] -PhysDrv[1:5] -a0
```

也可以指定为某个阵列的专用热备:

```
MegaCli -PDHSP -Set [-Dedicated [-Array1]] [-EnclAffinity]
[-nonRevertible] -PhysDrv[1:5] -a0
```

7) 删除全局热备:

```
MegaCli -PDHSP -Rmv -PhysDrv[1:5] -a0
```

8) 将某块物理盘下线/上线:

```
MegaCli -PDOffline -PhysDrv [1:4] -a0
```

```
MegaCli -PDOnline -PhysDrv [1:4] -a0
```

9) 查看物理磁盘重建进度:

```
MegaCli -PDRbld -ShowProg -PhysDrv [1:5] -a0
```

或者以动态可视化文字界面显示:

```
MegaCli -PDRbld -ProgDsply -PhysDrv [1:5] -a0
```

RAID 维护策略

当发生掉盘后,应该更换新的硬盘,更换硬盘后,在 SAS 卡 2008 下会默认为 JBOD 模式,或者硬盘本身带有 foreign 信息,此时应该执行命令把它转换为 Unconfigure good 模式并清除 foreign 信息再做 Raid。

```
[0009] 1) 首先查找到该卡子的 Dvice ID,记录
```

```
#. /MegaCli64 -PDlist -aALL | grep "Dvice ID"
```

```
Enclosure Device ID: 64,
```

记录该 En Device ID,如果由两个卡子可以分别执行参数 -a0 和 -a1 来取代 -aALL, 查找看有没有 foreign 信息#. /MegaCli64 -PDlist -aALL | grep "Foreign State" 得到如下信息:

```
Foreign State: None
```

2) 在执行该命令前应该看下 slot 号是否按顺序排列

```
#. /MegaCli64 -PDlist -aALL | grep "Slot",如果不按顺序则应找到该硬盘对应的
```


Slot 号,经查看如上的信息按 0-7 顺序排列。该信息显示 slot0-7 的 foreign 信息,可以看到 slot1 的硬盘状态存在 foreign

3) 修改硬盘状态信息为 unconfigue good

把改 Dvice ID 64 卡子的 slot 1 状态改为 unconfigue good

```
#./MegaCli64 -PDMakeGood -PhysDrv[64:1] -a0
```

```
Adapter: 0: EnclId-64 SlotId-1 state changed to Unconfigured-Good. Exit Code: 0x00
```

4) 清除 foreign 信息 扫描该卡子上的硬盘看是否存在 foreign 信息

```
./MegaCli64 -CfgForeign -Scan -a0
```

不存在就不用再执行下面的命令了。如果在 a0 上存在则执行,清除

```
#./MegaCli64 -CfgForeign -Clear [a0]
```

```
Foreign configuration 0 is cleared on controller 0. Exit Code: 0x00
```

5) 重新做 Raid

En DeviceID64 的 slot1 的单盘做 Raid0,则

```
#./MegaCli64 -CfgLdAdd-r0 [64:1] WT Direct -a0
```

```
Adapter 0: Created VD 1
```

```
Adapter 0: Configured the Adapter!!
```

```
Exit Code: 0x00
```

在该卡子上 slot4 和 slot5 制作 Raid1

```
./MegaCli64 -CfgLdAdd-r1 [64:4,64:5] WT Direct -a0
```

```
Adapter 0: Created VD 4
```

```
Adapter 0: Configured the Adapter!!
```

```
Exit Code: 0x00
```

当 Raid1 掉线后更换新硬盘,使用参数 -PDMakeGood -PhysDrv[64:4] 把硬盘配置为 unconfigue good 后,拔出该盘,重新插入,则 Raid 卡会自动处于 Rebuild 状态。或者得到掉线的 PD 信息

```
#./MegaCli64 -Pdgetmissing -a0
```

```
Adapter 0 - Missing Physical drives Exit Code: 0x00
```

替代该掉线硬盘

```
#MegaCli64 -PdReplaceMissing -physdrv[64:4] -array4, -row0 -a0
```

重启后将自动重建。

[0010] 本发明的有益效果是。

[0011] 在 RACK 生产实践中可大大降低因硬盘故障需而重建 RAID 人力及时间,不必耗时耗力在 WEBBIOS 下通过手动一步一步操作建立,仅需 Linux 下几条命令可自动重建,大大提高了生产效率。

具体实施方式

[0012] 下面对本发明的内容进行更加详细的阐述：

硬件 RAID 管理方法主要从三个方面进行 RAID 组建管理及维护：

首先,通过查看物理磁盘信息及客户需求提出 RAID 组建策略,可以根据用户资源如硬盘形态数量及缓冲等设置选项调优 RAID,综合出最优 RAID 组建方案。

[0013] 通过查看所有物理磁盘信息 :MegaCli -PDList -aALL

查看磁盘缓存策略 :MegaCli -LDGetProp -Cache -L0 -a0

其次,通过创建 / 删除 RAID,可以创建包括 RAID0/1/5/6/10 等几种不同形态,其中 RAID0 为非冗余高性能存储方式, RAID1 模式为数据同步备份容量减半存方式,同时可以指定热备盘等备用盘符便于后期重建,阵列创建完后,会有一个初始化同步块的过程,可直观观察并完成 RAID 组建 ;

创建阵列,以 RAID5 为例 :由物理盘 2,3,4 构成,该阵列的热备盘是物理盘 5 :

MegaCli -CfgLdAdd -r5 [1:2,1:3,1:4] WB Direct -Hsp[1:5] -a0

创建阵列,不指定热备 :MegaCli -CfgLdAdd -r5 [1:2,1:3,1:4] WB Direct -a0

在线添加磁盘 :MegaCli -LDRecon -Start -r5 -Add -PhysDrv[1:4] -L1 -a0

查看阵列初始化信息 :MegaCli -LDInit -ShowProg -LALL -aALL

以动态可视化文字界面显示 :MegaCli -LDInit -ProgDsply -LALL -aALL

创建全局热备 :MegaCli -PDHSP -Set [-EnclAffinity] [-nonRevertible] -PhysDrv[1:5] -a0

最后,磁盘信息查看及维护策略,包括磁盘使用占比及当发生掉盘后的重建及维护。在查找到该卡子的 Dvice ID 后修改硬盘状态信息,之后清除 foreign 信息并重建 RAID。