linux基础

## 驱动程序

驱动程序就是将CPU的指令转换成对应机械设备所能完成的操作

如硬盘驱动程序,就是将cpu的指令(读写文件)转换成机械硬盘芯片对磁臂完成对文件存放的位置(扇区)的寻道,读取写入

CPU指令 **驱动程序** 硬盘芯片对磁臂的控制

一般驱动程序都是由硬件厂商提供,如果不是硬件厂商提供叫公版驱动,优化方面可能不如硬件厂商.

CPU信号和硬件设备芯片能够通信都是建立在**协议**的基础上的.

## 存储设备接口



按道理说并行比串行块,但是未必,并行要额外控制,且有干扰.

Mbps/8 = m/s

**IDE** 并行 传输的理论值 133Mbps

**SATA** 串行

SATA1 : 300Mbps

SATA2 : 600Mbps

SATA3 : 6Gbps

**USB**

usb3.0 480Mbps / 8

**SCSI**

SCSI : small computer system interface

SCSI硬盘,一般都在10000,15000转 RPM

有一个小的CPU,比SATA3还块1倍.

## 硬件RAID级别

1987 加州伯克利大学几个人发明

RAID : A Case for Redundant Arrays of Inexpensive Disks

廉价冗余磁盘阵列

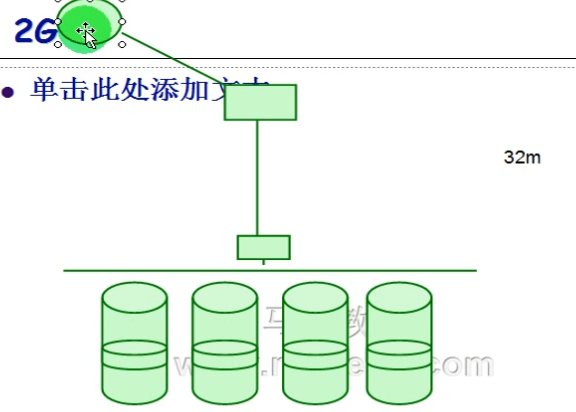
后来改为 A Case for Redundant Arrays of Independents Disks

冗余独立磁盘阵列

冗余 冗余在特定场合是非常有用的

RAID是今天事实上的解决存储方案的工业标准.

条带化存储



**0** **条带**

读写性能提升

没有冗余能力

空间利用率:ns

至少2块

**1** **镜像**

读性能提升,写性能有下降

有冗余能力

空间利用率: 1/2

至少2块

2

3

4

**5**

读写能力都提示

有冗余能力

空间利用率 : (n-1)/n

至少3块

**10 RAID10**

读写能力提升

有冗余能力

空间利用率 1/2

至少4块

**01 RAID01**

读写能力提升

有冗余能力

空间利用率 1/2

至少4块

**50**

读写能力提升

有冗余能力

空间利用率 (n-2)/n

至少6块

**jbod**

性能无提升

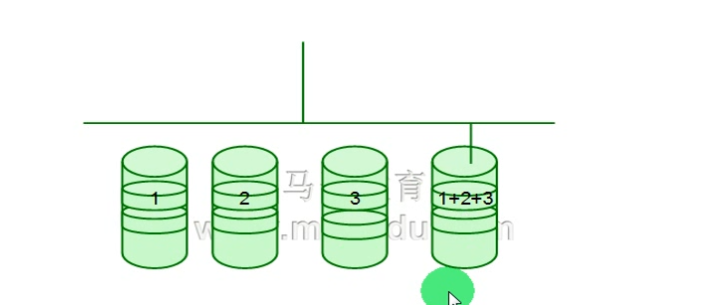
没有冗余能力

空间利用率 100%

至少2块

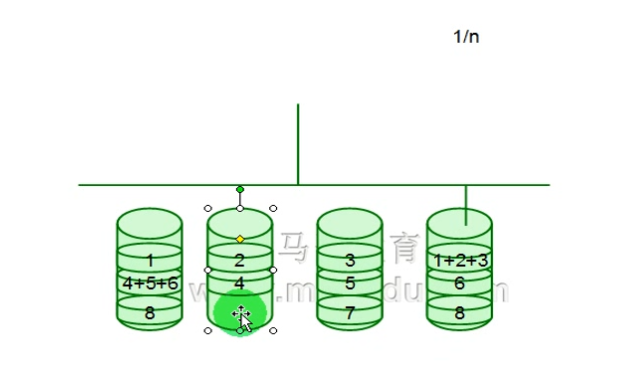
**RAID4**

检验码盘只有一个,有瓶颈,校验码盘的快慢决定了整个的快慢,压力都在最后一个盘,校验码盘不能坏,其他只能坏一个,所以有改进版本 RAID5



**RAID5**

每个盘轮流作为校验盘,只允许坏一个盘



**RAID6**

raid6用两块硬盘作为检验码盘,校验方式不同.用的不多

RAID 2 3 4 都没有用

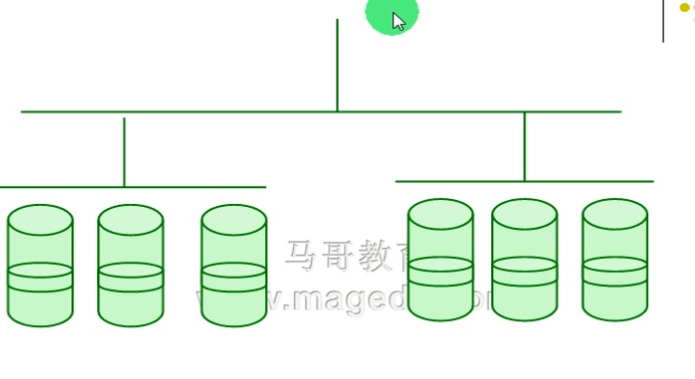
镜像浪费空间,1/2

检验码浪费空间1/n,允许一个出现问题,如果两个磁盘坏了,就over了

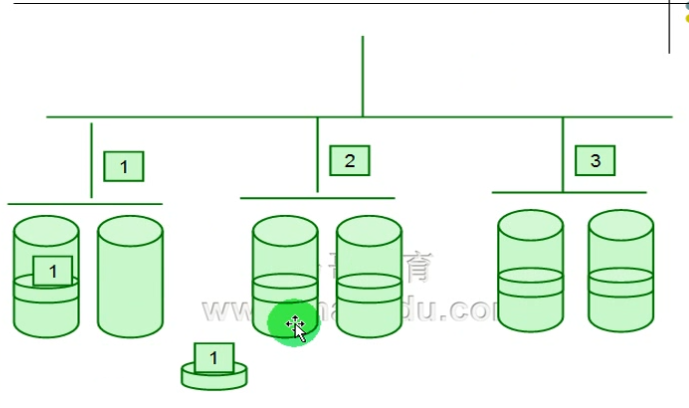
有钱的企业这么玩儿就是镜像和条带一起用

如果先做条带在做镜像,当中任何一个盘坏了,都会受到影响

**RAID 0+1**



**RAID 1+0**



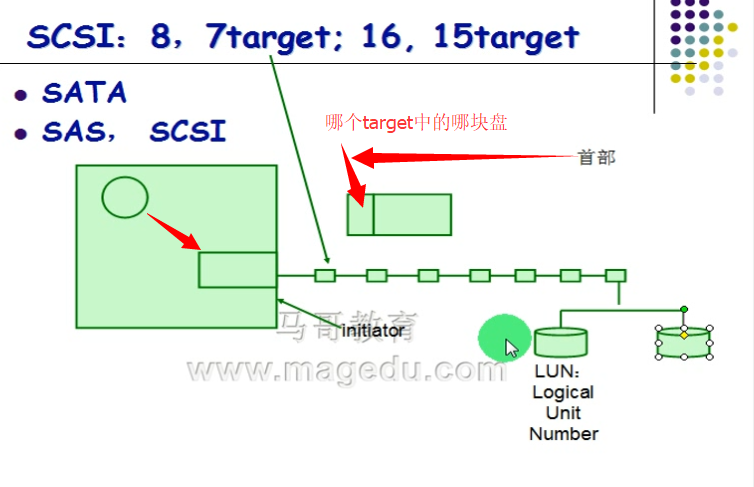
性能一样,但是一个盘坏了,只在镜像内部修复,可靠性要好很多,但是同组不能同时坏.

**jbod**

将多块小盘组合成一个大盘来用 , hadoop能用上,因为hadoop的HDFS在文件系统上就能冗余,在硬件层面就不用冗余了

早期组合 IDE USB

现在组合 SATA SAS



SAS 造价高 2.5寸 体积小 和移动硬盘差不多

SCSI 总线的扩展能力很强大,可以接很多盘

SATA 只能接一个盘

IDE 能接2个盘,主盘和从盘

硬件厂商提供的RAID芯片,操作系统不能识别,那么需要厂商额外提供驱动程序.

RAID芯片也比较贵，高端货不仅有CPU还自带#G内存,性能非常好,还自带电池,防止掉电数据丢失,但是贵,如果没钱那就软件级别RAID

## 软件级别RAID

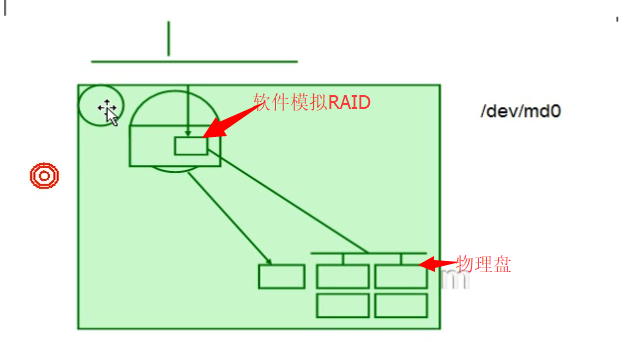
软件级别RAID也叫逻辑RAID

/dev/sda /dev/sdb

/dev/sdc /dev/sdd

/dev/md# /dev/md0 /dev/md1…

数据存储的时候先通过内核模拟的设备 /dev/md# , /dev/md# 在将数据分散到 sda sdb sdc sdd上



操作系统内核对物理盘标记为Fd 就是软RAID,如果故障,很难修复生成环境中并不建议使用,很多事情都是一个妥协和这种的过程

### mdadm

用户空间工具,管理工具

mdadm命令能将任何块设备做成RAID

**模式化命令**

创建模式 -c

专用选项

-l 级别

-n# 设备个数

-a 自动为其创建文件设备 -a yes

-c 指定chunk(数据块)大小 2^n 默认64K

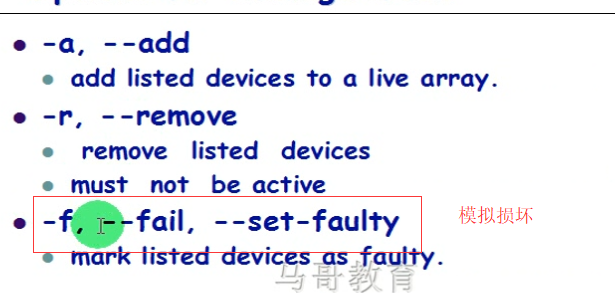
-x# 指定空闲盘的个数

**管理模式**

--add / -a 添加

--remove / -r 删除

--fail / -f 模拟损坏,不是真的损坏



**监控模式**

**-F**

**增长模式**

**-G**

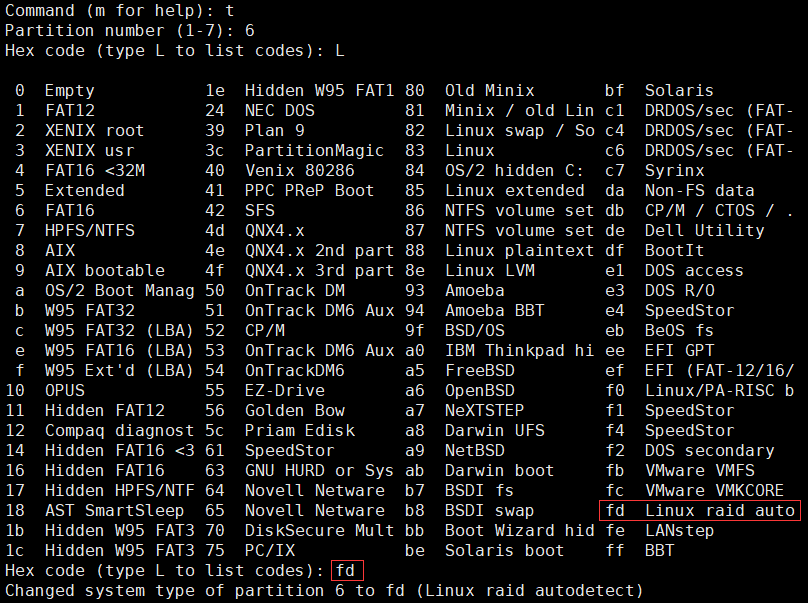
**装配模式**

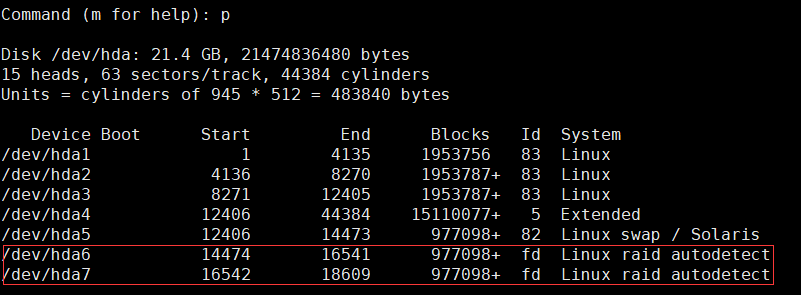
**-A**

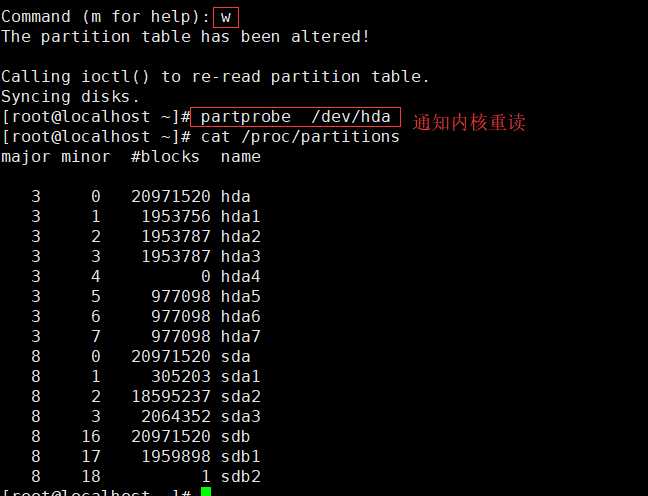
### demo

#### 1 用2G组成RAID0 (条带)

2个1G

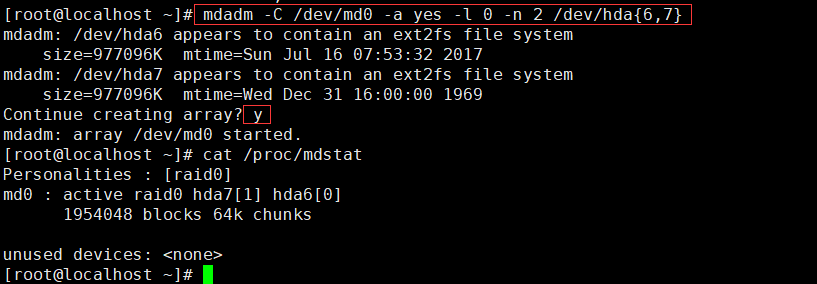




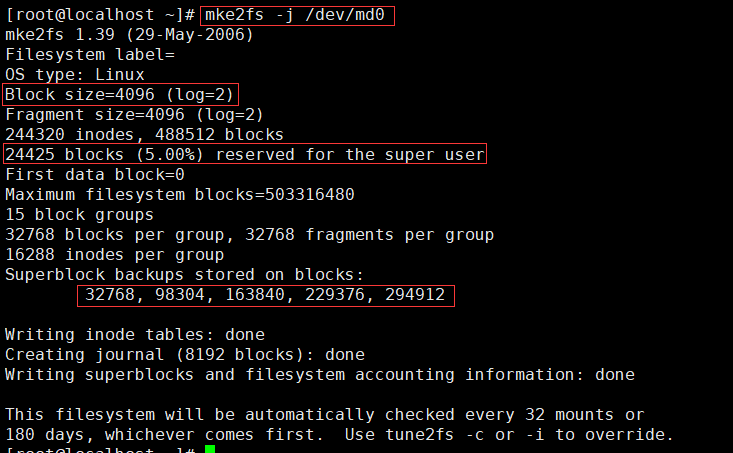


mdadm -C /dev/md0 -a yes -l 0 -n 2 /dev/sda{5,6}

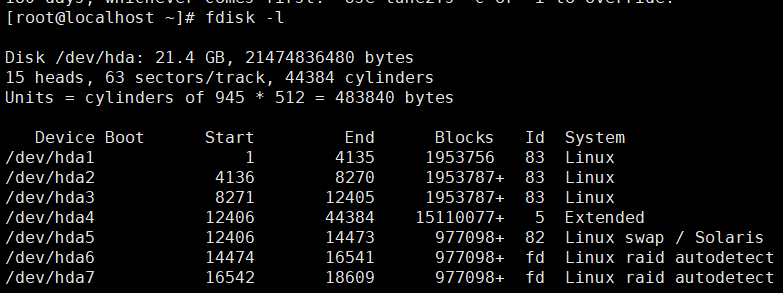
把/dev/sda5,/dev/sda6自动创建级别为0的设备md0

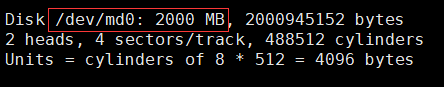


高级格式化

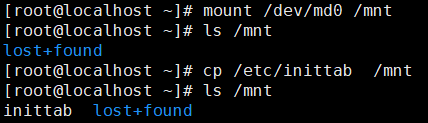


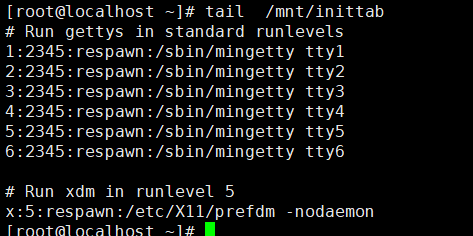
查看空间大小





就可以挂载使用





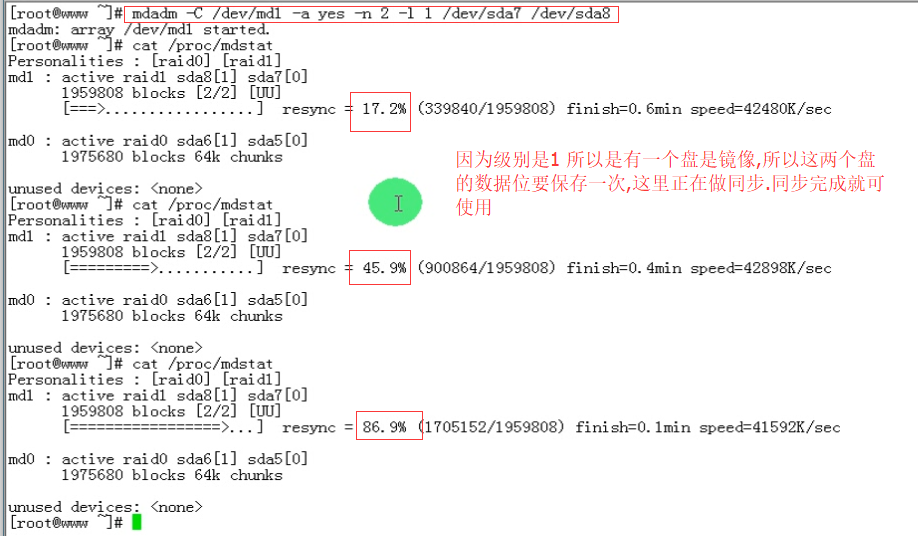
#### 2 用2G组成RAID1 (镜像)

利用率只有1G

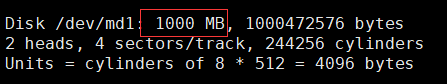
mdadm -C /dev/mdl -a yes -n 2 -l 1 /dev/sda7 /dev/sda8

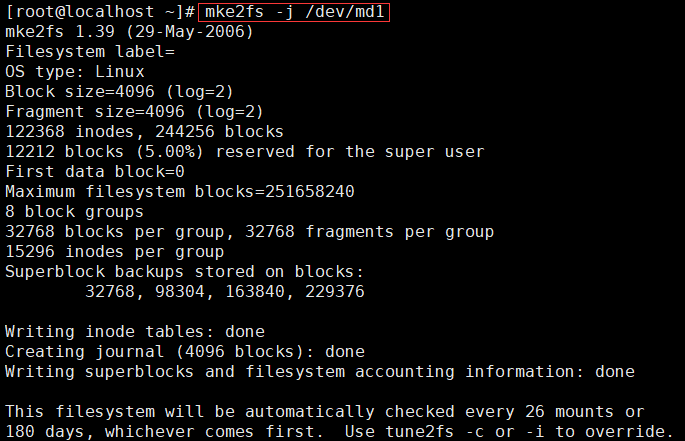


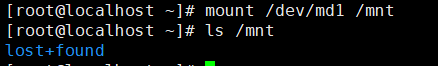
级别为1



RAID1利用率只有一半

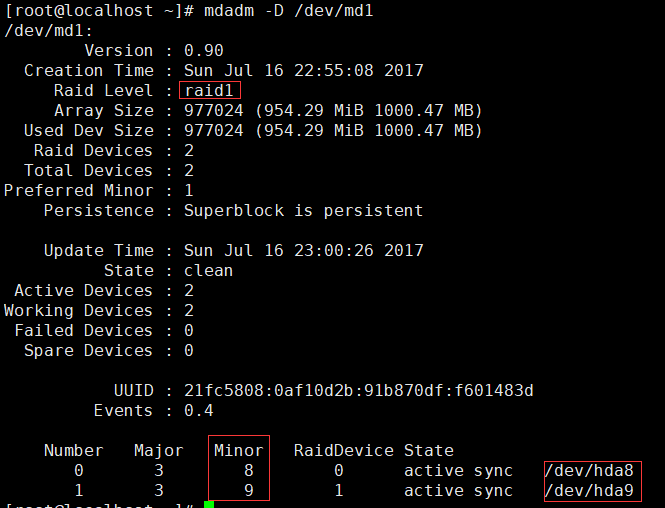






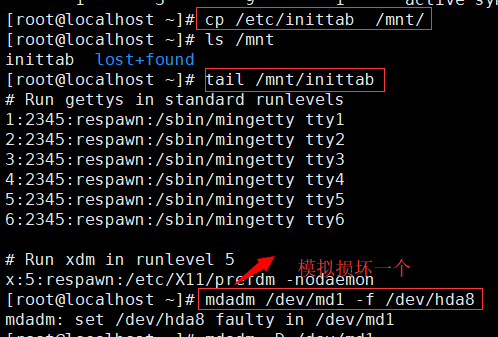
查看RAID阵列详细信息

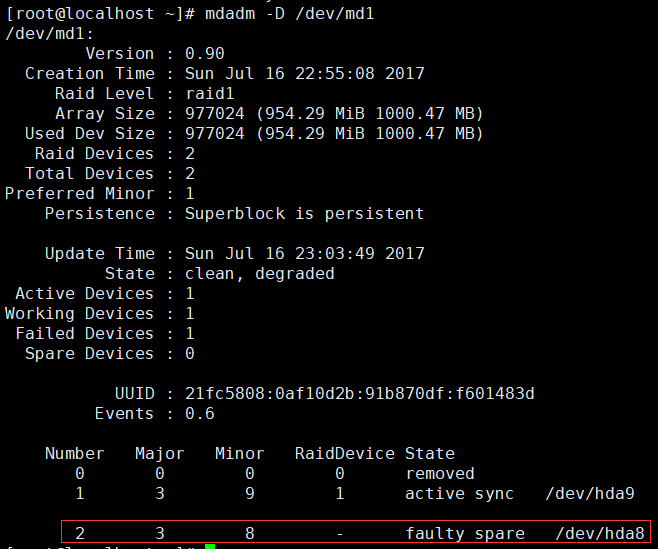
mdadm -D /dev/md1 == mdadm --detail /dev/md1



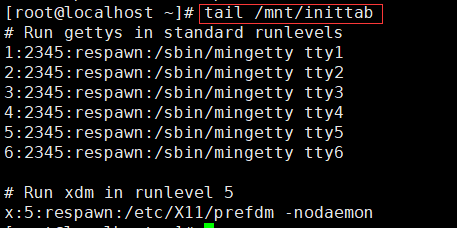
模拟损坏

mdadm /dev/md1 -f /dev/hda8

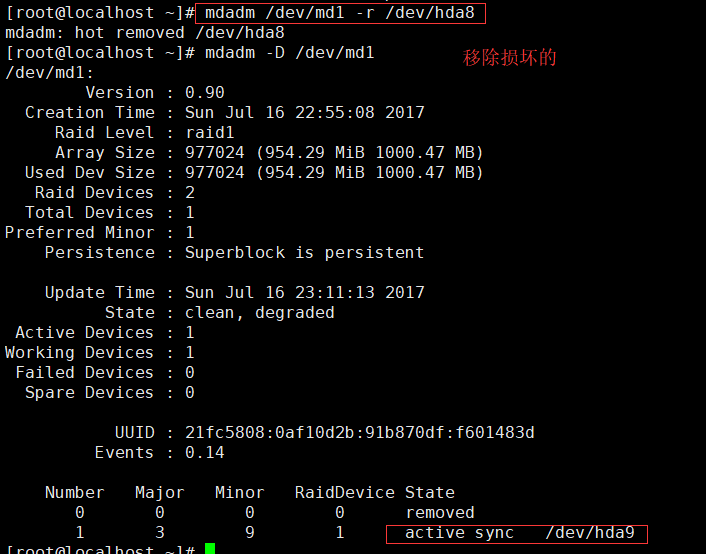




是可以访问的,因为这是RAID1 /dev/sda7是一个镜像



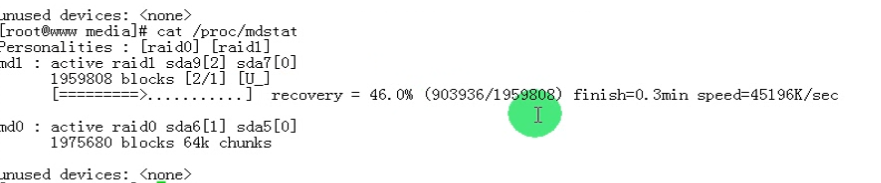
mdadm /dev/md1 -r /dev/hda8



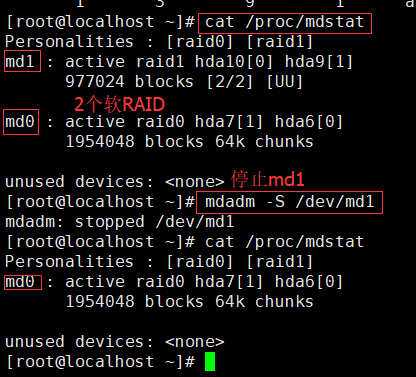
增加一个新的设备来替代损坏的设备,但是要和镜像盘保持一样



正在同步



所以镜像允许你设备损坏坏一个

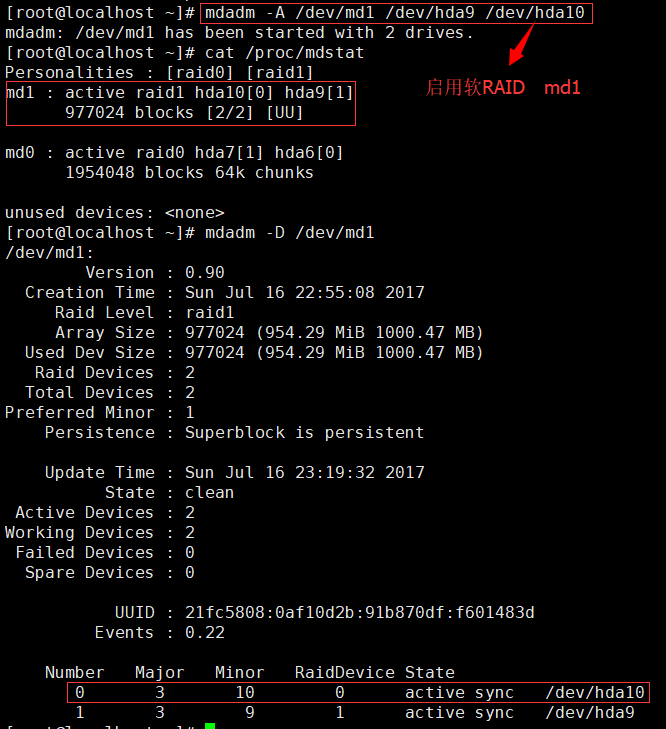


删除设备md1

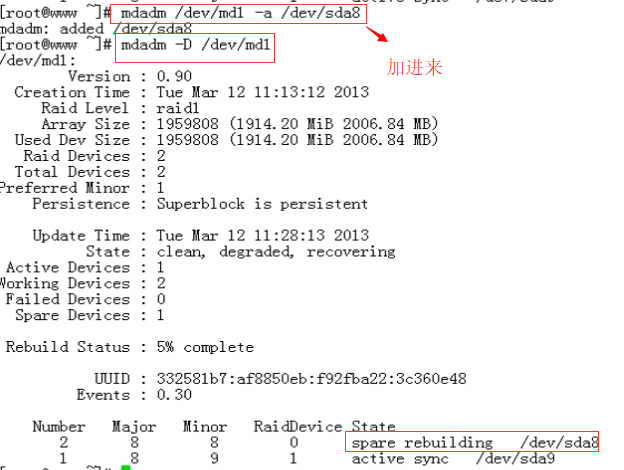


如果要启用

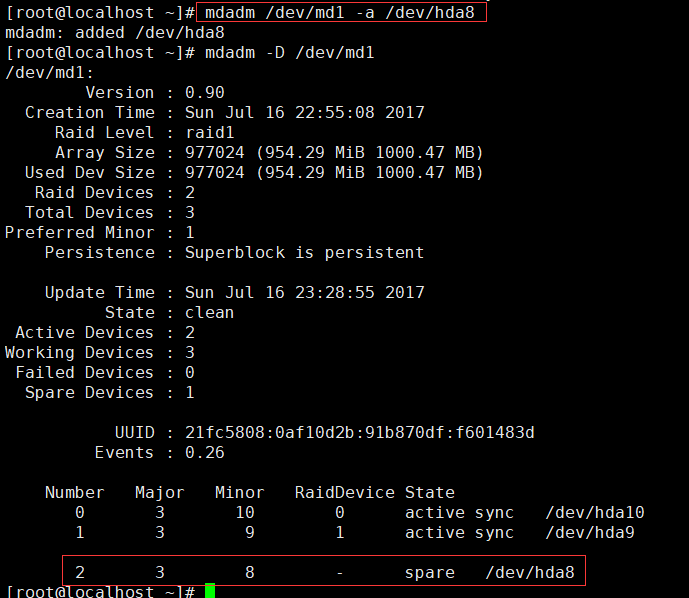
mdadm -A /dev/md1 /dev/sda7 /dev/sda9

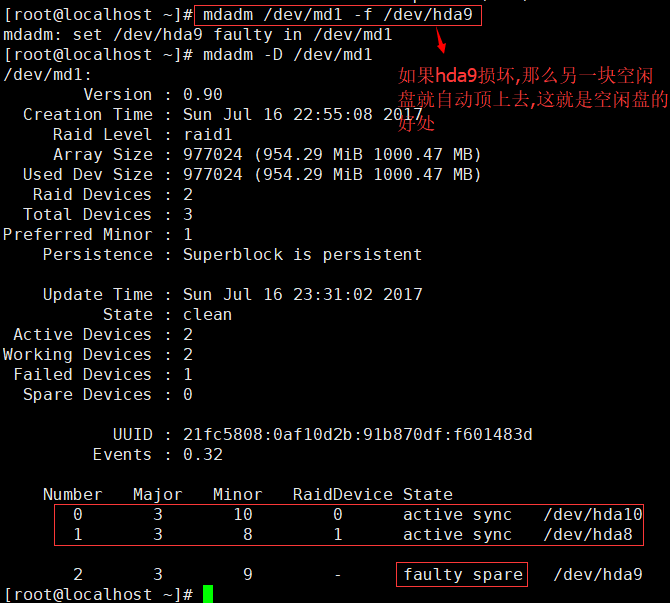


如果有一个是被移除了,那么加上来即可

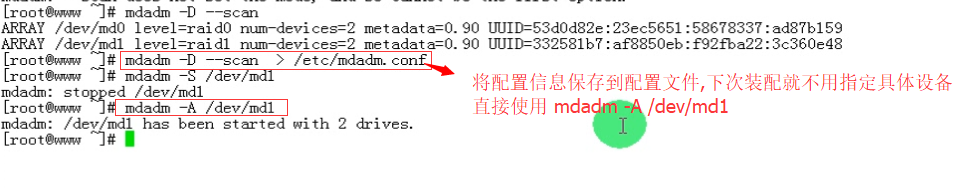


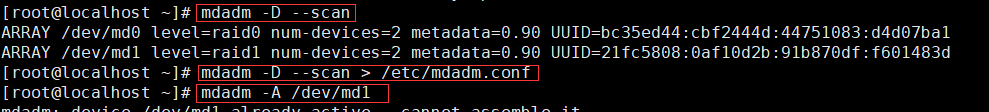
如果在装上一个/dev/hda8 那么hda8 就是一个空闲盘



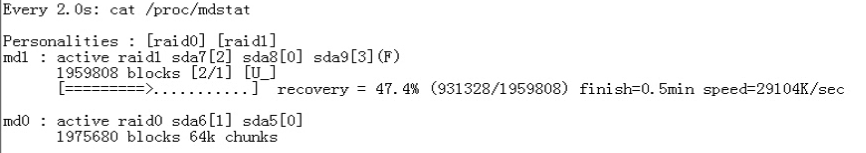


mdadm -D --scan 获取具体的设备装配信息





watch ‘cat /proc/mdstat’ 动态显示磁盘同步过程

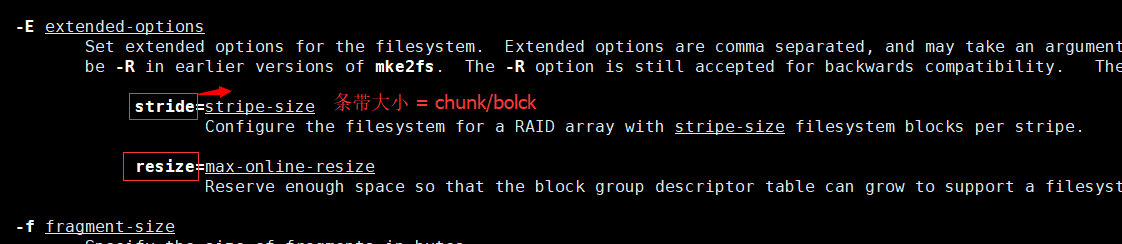


#### 3 用2G组成RAID5

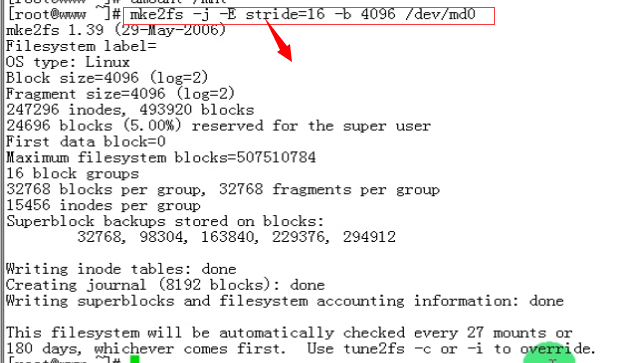
用法和上面相同 -l 5

man mke2fs 有一个选项 -E

stride 条带的大小 对应RAID0 RAID５是有意义的,需指chunk 和条带



指定条带大小,好处就是可以对磁盘优化



## watch

周期性地执行指定命令,并且以全屏的方式显示结果

-n # 指定周期长度默认单位为秒,默认为2

watch -n # ‘command’

## lsmod

列出模块类型

