# 背景

某台机器上安装了2块IDE磁盘和4块SCSI磁盘，IDE硬盘直接连接到主板集成的IDE接口上，SCSI磁盘则是连接到一块PCI接口的SCSI卡上。在没有RAID参与的条件下，系统可以识别出6块磁盘，并且经过文件系统格式化以后，挂载到某个盘符或者目录下，供应用程序读写。

安装RAID后，通过配置，先将2块IDE磁盘做成一个RAID0系统，如果原来每块IDE磁盘是80G，做成RAID0后变成一个160GB的虚拟磁盘。然后将4块SCSI磁盘做成RAID5系统，如果每块SCSI盘大小73GB，则最后是3块磁盘，容量216GB。

因为RAID程序需要使用磁盘上的部分空间来存放一些RAID信息，所以实际容量将会变小，经过RAID程序的处理后，6块磁盘变为2块虚拟磁盘。

# 配置

Linux下RAID配置

必须划分一个/boot分区用来启动基本的操作系统内核，用第一块次磁盘sda的前100MB容量来创建这个分区

在创建/boot分区之后，将SDA磁盘剩余的分区以及所有剩余的物理磁盘，军配置为software RAID类型

选择create a RAIDdevice [default=/dev/md0]

# 软件RAID

## 概述

通过硬件阵列卡实现raid具有可靠性高，性能好等特点，但是对于一般的企业而言硬件阵列卡固然好，如果大规模应用的话动辄几千上万的费用也不是他们所能承受的，难道就没有既能保证数据安全，又能减少费用支持的IT方案吗？当然有，软raid就可以实现这一需求。

## 特点

软件RAID有3个缺点：

1. 占用内存空间
2. 占用CPU资源
3. 软件RAID程序无法将安装有操作系统的那个磁盘分区做成RAID模式

因为RAID程序是运行在操作系统之上的，所以在启动操作系统之前，是无法实现RAID功能的，也就是说系统损坏了，RAID程序也就无法运行了。

## 实践

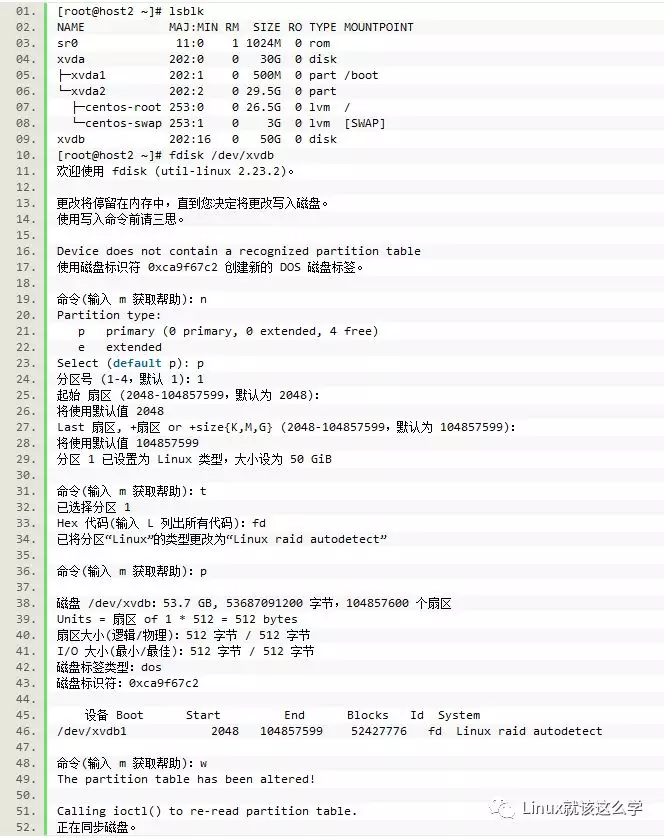
一台centos 7.2 VM，主机名分别为host1，host1上挂载两块50G的数据盘，我们的实验就在这两块数据盘上完成。

P.S:需要强调的是生产环境下两块组raid的硬盘必须是同品牌同型号同容量的，否则极容易出现软raid失效的情况。

1、确认操作系统是否安装了mdadm软件

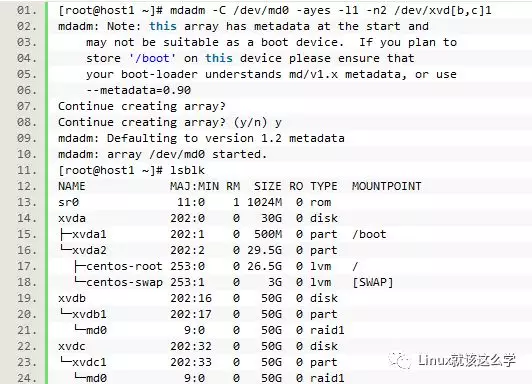
[root@host1 ~]# rpm -qa |grep mdadm

mdadm-3.3.2-7.el7.x86\_64

2、对两块数据盘进行分区，并设置分区类型为raid

fdisk命令只能对容量在2T以下的硬盘进行分区，如果超过2T则需要使用parted工具了。

在fdisk下raid的分区类型代码是fd，在parted工具下首先需要使用mklabel将磁盘格式由MBR改成GPT，然后才能使用mkpart命令进行分区，分完区之后使用set设置分区flag为raid即可。

3、使用mdadm命令创建raid1

可以看到创建时有个Note的提示，是说软raid不能当启动设备，这就是软raid比较鸡肋的地方了。

mdadm -C /dev/md0 -ayes -l1 -n2 /dev/xvd[b,c]1

命令说明：

-C创建阵列;

-a同意创建设备，如不加此参数时必须先使用mknod 命令来创建一个RAID设备，不过推荐使用-a yes参数一次性创建;

-l阵列模式;

-n阵列中活动磁盘的数目，该数目加上备用磁盘的数目应该等于阵列中总的磁盘数目;

/dev/md0阵列的设备名称，如果还有其他阵列组可以以此类推；

创建完成后可以使用cat /proc/mdstat查看阵列状态：

下图中第一次查看的时候提示resync完成了95.7%，第二次查询的时候两块盘才真正同步完。

也可以使用mdadm -D /dev/md0查看阵列组的状态

4、创建md0的配置文件

echo DEVICE /dev/sd{a,b}1 >> /etc/mdadm.conf

mdadm -Evs >> /etc/mdadm.conf

mdadm运行时会自动检查/etc/mdadm.conf 文件并尝试自动装配，因此第一次配置raid后可以将信息导入到/etc/mdadm.conf 中。

5、使用/dev/md0

在/dev/md0上创建文件，然后挂载进行使用。

mkfs.ext4 /dev/md0

# RAID卡

## 概述

RAID卡是一种利用独立硬件实现RAID功能的方法。对于硬件RAID操作系统无法感知，只能通过厂家提供的RAID卡管理软件来查看卡上所连接的物理磁盘。而且，在配置RAID卡的时候，也不能在操作系统上完成，而必须进入这个硬件完成（或者在操作系统改下通过RAID卡配置工具来设置），一般的RAID卡都是在开机自检的时候，进入他的ROM配置程序来配置各种RAID功能的。

**注：**在我们安装操作系统的时候，有时候找不到磁盘，这就是以为该服务器默认的是使用了硬件RAID，我们不去设置的话系统感知不到这个盘，我们必须在进入开机前设置RAID的功能才可以捕获这些磁盘。

RAID卡可以集成到SCSI卡或者主板的南桥（在南桥的称为板载RAID卡）。操作系统出了RAID卡驱动外不用安装任何额外的软件。

带CPU的RAID卡就是一个小型计算机系统，有自己的CPU、内存、ROM、总线和IO接口，只不过这个小计算机系统是为大计算机系统服务的。

SCSI RAID卡上一定要包含SCSI控制器，因为其后端链接的是SCSI物理磁盘，前端链接的是主机的PCI总线，所以一定要有一个PCI总线控制器来维护PCI总线的仲裁、数据发送、接收等功能。还需要有一个ROM，一般都是用Flash芯片，其中存放着初始化RAID卡必须的代码以及实现RAID功能所需的代码。

RAM的作用首先是数据缓存，提高性能，其次作为RAID卡上的CPU执行RAID计算所需要的内存空间。

## RAID与SCSI卡

RAID卡与SCSI卡的区别就在于RAID功能，其他的没有什么大的区别。如果RAID卡上有多个SCSI通道，那么就称为多通道RAID卡。目前，SCSI RAID卡最高有4通道的，其后端可以接入4条SCSI总线，所以最多可连接64个SCSI设备（16位总线）。

增加了RAID功能后，SCSI就成了RAID代码的傀儡，RAID让他干嘛就干嘛，SCSI对其下面的磁盘情况了解，那么RAID就通过它获取底层磁盘的情况，按照ROM中所设置的选项，比如RAID类型、条带大小等，对RAID程序代码做相应的调整，操控它的傀儡SCSI控制器向主机报告“虚拟”的逻辑盘，而不是所有的物理盘。

## LUN

条带化之后，RAID程序代码就操控SCSI控制器向OS层驱动程序代码提交一个虚拟化之后所谓的“虚拟盘”或者“逻辑盘”，称之为LUN。

LUN是SCSI ID更细一级的地址号，每个SCSI ID（Target ID）下面还可以有更多的LUN ID（视ID字段的长度而定），对于大型磁阵而言，可以生成几百或者几千个虚拟磁盘，为每个虚拟磁盘分配一个SCSI ID是不可能的，因为每个SCSI总线最多接入16个设备，LUN就是下一级的寻址ID，后来LUN概念转换为软件生成的虚拟磁盘，统一称为“卷”。

0通道RAID卡又称为RAID子卡，0通道的意思是这块卡的后端没有SCSI通道。

磁阵上会有多个外部SCSI接口，这个接口是为了链接多台主机使用的，每个由盘阵RAID控制器生成的逻辑磁盘，可以通过设置只分配（Assign/Map）到其中一个口。比兔LUN1被分配到了1号口，那么链接2号口的主机就看不到，月可以一个LUN分配给多个口，但是很容易造成数据不一致，除非使用集群文件系统或者高可用性系统软件的额参与。