# RAID详解

一、RAID定义RAID(Redundant Array of Independent Disk 独立冗余磁盘阵列)技术是加州大学伯克利分校1987年提出，最初是为了组合小的廉价磁盘来代替大的昂贵磁盘，同时希望磁盘失效时不会使对数据的访问受损 失而开发出一定水平的数据保护技术。RAID就是一种由多块廉价磁盘构成的冗余阵列，在操作系统下是作为一个独立的大型存储设备出现。RAID可以充分发 挥出多块硬盘的优势，可以提升硬盘速度，增大容量，提供容错功能够确保数据安全性，易于管理的优点，在任何一块硬盘出现问题的情况下都可以继续工作，不会 受到损坏硬盘的影响。

二、 RAID的几种工作模式（仅讨论 RAID0，RAID1，RAID5，RAID10这四种，这四种比较典型）  
1、RAID0 （又称为Stripe或Striping－－分条）  
即Data Stripping数据分条技术。RAID 0可以把多块硬盘连成一个容量更大的硬盘群，可以提高磁 盘的性能和吞吐量。RAID 0没有冗余或错误修复能力，成本低，要求至少两个磁盘，一般只是在那些对数 据安全性要求不高的情况下才被使用。

特点：

|  |  |  |  |
| --- | --- | --- | --- |
| 容错性： | 没有 | 冗余类型： | 没有 |
| 热备盘选项： | 没有 | 读性能： | 高 |
| 随机写性能： | 高 | 连续写性能： | 高 |
| 需要的磁盘数： | 只需2个或2\*N个（这里应该是多于两个硬盘都可以） | 可用容量： | 总的磁盘的容量 |
| 典型应用： | 无故障的迅速读写，要求安全性不高，如图形工作站等。 | | |

#### RAID 0的工作方式：

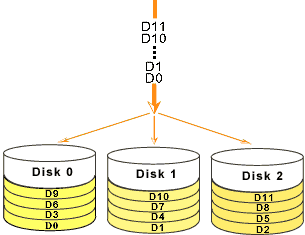


图1

如图1所示:系统向 三个磁盘组成的逻辑硬盘(RADI 0 磁盘组)发出的I/O数据请求被转化为3项操作，其中的每一项操作都对应于一块物理硬盘。我们从图中可以清楚的看到通过建立RAID 0，原先顺序的数据请求被分散到所有的三块硬盘中同时执行。  
从理论上讲，三块硬盘的并行操作使同一时间内磁盘读写速度提升了3倍。 但由于总线带宽等多种因素的影响，实际的提升速率肯定会低于理论值，但是，大量数据并行传输与串行传输比较，提速效果显著显然毋庸置疑。  
RAID 0的**缺点**是不提供数据冗余，因此一旦用户数据损坏，损坏的数据将无法得到恢复。  
RAID 0具有的**特点**，使其特别适用于对性能要求较高，而对数据安全不太在乎的领域，如图形工作站等。对于个人用户，RAID 0也是提高硬盘存储性能的绝佳选择。  
计算机技术发展迅速，但硬盘传输率也成了性能的瓶颈。怎么办？IDE RAID技术的成熟让我们轻松打造自己的超高速硬盘。在实际应用中，RAID 0硬盘阵列能比普通IDE 7200转ATA 133硬盘快得多，时至今日，在大多数的高端或者玩家主板上我们都能找到一颗PROMISE或者HighPoint的RAID芯片，同时发现它们提供的额 外几个IDE接口。没错，RAID已经近在眼前，难道你甘心放弃RAID为我们带来的性能提升吗？答案当然是否定的！  
实用 的IDE RAID  
RAID可以通过软件或硬件实现。像Windows 2000就能够提供软件的RAID功能，但是这样需要消耗不小的CPU资源，降低整机性能。而硬件实现则是一般由RAID卡实现的，高档的SCSI RAID卡有着自己专用的缓存和I/O处理器，但是对于家庭用户来说这样的开销显然是承受不了的，毕竟为了实现RAID买两个或者更多的HDD已经相当不 容易了。我们还有一种折中的办法——IDE RAID。或许这才是普通人最容易接受的方法。虽然IDE RAID在功能和性能上都有所折中，但相对于低廉的价格，普通用户看来并不在意。  
为什么要用RAID 0RAID 0至少需要两块硬盘才能够实现，它的容量为组成这个系统的各个硬盘容量之和，这几块硬盘的容量要相同，在家用IDE RAID中一般级联两块硬盘，一定要用同型号同容量的硬盘。RAID 0模式向硬盘写入数据的时候把数据一分为二，分别写入两块硬盘，读取数据的时候则反之，这样的话，每块硬盘只要负担一半的数据传输任务，得到的结果也就是 速度的增加。

#### 实现方式：

**（1）、**RAID 0最简单方式（我觉得这个方式不是它本意所提倡的）  
就是把x块同样的硬盘用硬件的形式通过智能磁盘控制器或用操作系统中的磁盘驱动程序以软件的方式串联在一起，形成一个独立的逻辑驱动器，容量是单独硬盘的 x倍,在电脑数据写时被依次写入到各磁盘 中，当一块磁盘的空间用尽时，数据就会被自动写入到下一块磁盘中，它的好处是可以增加磁盘的容量。  
速度与其中任何一块磁盘的速度相同，如果其中的任何一块磁盘出现故障，整个系统将会受到破坏，可靠  
性是单独使用一块硬盘的1/n。  
**（2）、**RAID 0的另一方式（常指的RAID 0就是指的这个）  
是用n块硬盘选择合理的带区大小创建带区集，最好是为每一块硬盘都配备一个专门的磁盘控制器,在 电脑数据读写时同时向n块磁盘读写数据,速度提升n倍。提高系统的性能。

2、RAID 1 （又称为Mirror或Mirroring－－镜像）RAID 1称为磁盘镜像：把一个磁盘的数据镜像到另一个磁盘上，在不影响性能情况下最大限度的保证系统的可靠性和可修复性上，具有很高的数据冗余能力，但磁盘利用 率为50%，故成本最高，多用在保存关键性的重要数据的场合。RAID 1的操作方式是把用户写入硬盘的数据百分之百地自动复制到另外一个硬盘上。

RAID 1有以下特点：  
（1）、RAID 1的每一个磁盘都具有一个对应的镜像盘，任何时候数据都同步镜像，系统可以从一组 镜像盘中的任何一个磁盘读取数据。  
（2）、磁盘所能使用的空间只有磁盘容量总和的一半，系统成本高。  
（3）、只要系统中任何一对镜像盘中至少有一块磁盘可以使用，甚至可以在一半数量的硬盘出现问题时系统都可以正常运行。  
（4）、出现硬盘故障的RAID系统不再可靠，应当及时的更换损坏的硬盘，否则剩余的镜像盘也出现问题，那么整个系统就会崩溃。  
（5）、更换新盘后原有数据会需要很长时间同步镜像，外界对数据的访问不会受到影响，只是这时整个系统的性能有所下降。  
（6）、RAID 1磁盘控制器的负载相当大，用多个磁盘控制器可以提高数据的安全性和可用性。

### RAID 1的工作方式：

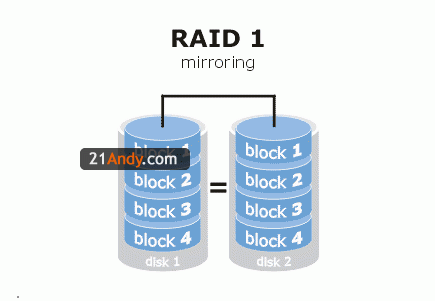


图2

如图2所 示：当读取数据时，系统先从RAID1的源盘读取数据，如果读取数据成功，则系统不去管备份盘上的数据；如果读取源盘数据失败，则系统自动转而读取备份盘 上的数据，不会造成用户工作任务的中断。当然，我们应当及时地更换损坏的硬盘并利用备份数据重新建立Mirror，避免备份盘在发生损坏时，造成不可挽回 的数据损失。  
raid 1的优缺点  
由于对存储的数据进行百分之百的备份，在所有RAID级别中，RAID 1提供最高的数据安全保障。同样，由于数据的百分之百备份，备份数据占了总存储空间的一半，因而Mirror（镜像）的磁盘空间利用率低，存储成本高。 Mirror虽不能提高存储性能，但由于其具有的高数据安全性，使其尤其适用于存放重要数据，如服务器和数据库存储等领域。

## 3、 RAID 5 （可以理解为是RAID 0和RAID 1的折衷方案，但没有完全使用RAID 1镜像理念，而是使用了“奇偶校验信息”来作为数据恢复的方式，与下面的RAID10不同。）

|  |  |  |  |
| --- | --- | --- | --- |
| 容错性： | 有 | 冗余类型： | 奇偶校验 |
| 热备盘选项： | 有 | 读性能： | 高 |
| 随机写性能： | 低 | 连续写性能： | 低 |
| 需要的磁盘数： | 三个或更多 | | |
| 可用容量： | （n-1）/n的总磁盘容量（n为磁盘数） | | |
| 典型应用： | 随机数据传输要求安全性高，如金融、数据库、存储等。 | | |

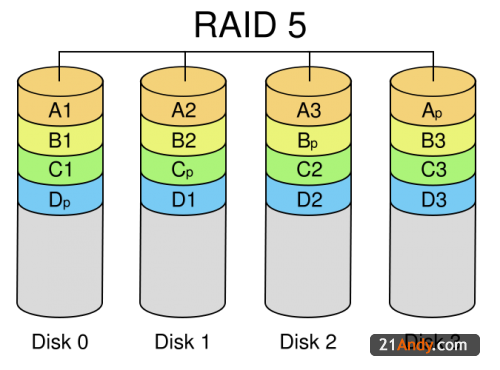
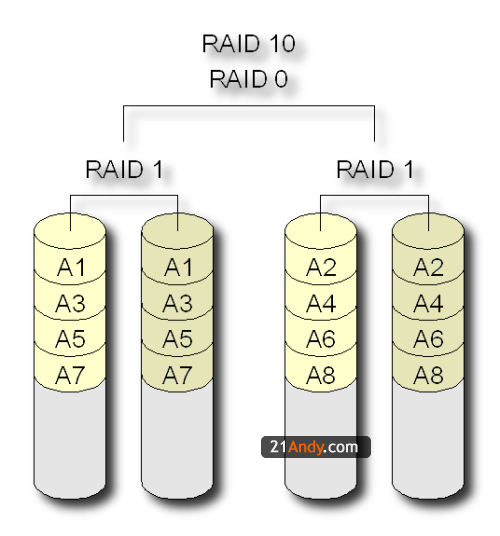


图3  
RAID 5 是一种存储性能、数据安全和存储成本兼顾的存储解决方案。 以四个硬盘组成的RAID 5为例，其数据存储方式如图4所示：图中，Ap为A1，A2和A3的奇偶校验信息，其它以此类推。由图中可以看出，RAID 5不对存储的数据进行备份，而是把数据和相对应的奇偶校验信息存储到组成RAID5的各个磁盘上，并且奇偶校验信息和相对应的数据分别存储于不同的磁盘 上。当RAID5的一个磁盘数据发生损坏后，利用剩下的数据和相应的奇偶校验信息去恢复被损坏的数据。  
RAID 5可以理解为是RAID 0和RAID 1的折衷方案。RAID 5可以为系统提供数据安全保障，但保障程度要比Mirror低而磁盘空间利用率要比Mirror高。RAID 5具有和RAID 0相近似的数据读取速度，只是多了一个奇偶校验信息，写入数据的速度比对单个磁盘进行写入操作稍慢。同时由于多个数据对应一个奇偶校验信息，RAID 5的磁盘空间利用率要比RAID 1高，存储成本相对较低。

## 4、 RAID 10 （可以理解为是RAID 0和RAID 1的折衷方案，但没有完全使用RAID 1镜像理念，而是使用了“奇偶校验信息”来作为数据恢复的方式）

  
图4

RAID10也被称为镜象阵列条带。象RAID0一样，数据跨磁盘抽取；象RAID1一样，每个磁盘都有一个镜象磁盘, 所以RAID 10的另一种会说法是 RAID 0+1。RAID10提供100%的数据冗余，支持更大的卷尺寸，但价格也相对较高。对大多数只要求具有冗余度而不必考虑价格的应用来说，RAID10提 供最好的性能。使用RAID10，可以获得更好的可靠性，因为即使两个物理驱动器发生故障（每个阵列中一个），数据仍然可以得到保护。RAID10需要4 + 2\*N 个磁盘驱动器（N >=0)， 而且只能使用其中一半(或更小, 如果磁盘大小不一)的磁盘用量, 例如 4 个 250G 的硬盘使用RAID10 阵列， 实际容量是 500G。

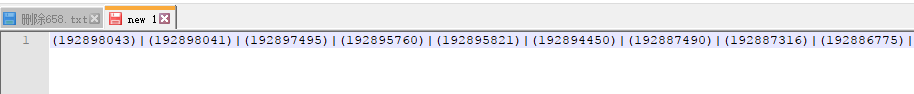
**RAID总结：**

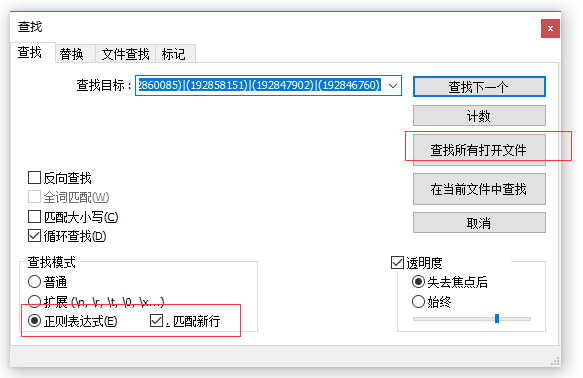
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 类型 | 读写性能 | 安全性 | 磁盘利用率 | 成本 | 应用方面 |
| RAID0 | 最好（因并行性而提高） | 最差（完全无安全保障） | 最高（100％） | 最低 | 个人用户 |
| RAID1 | 读和单个磁盘无分别，写则要写两边 | 最高（提供数据的百分之百备份） | 差（50％） | 最高 | 适用于存放重要数据，如服务器和数据库存储等领域。 |
| RAID5 | 读：RAID 5＝RAID 0（相近似的数据读取速度）    写：RAID 5<对单个磁盘进行写入操作（多了一个奇偶校验信息写入） | RAID 5<RAID 1 | RAID 5>RAID 1 | RAID 5<RAID 1 | 是一种存储性能、数据安全和存储成本兼顾的存储解决方案。 |
| RAID10 | 读：RAID10＝RAID0    写：RAID10＝RAID1 | RAID10＝RAID1 | RAID10＝RAID1（50％） | RAID10＝RAID1 | 集合了RAID0，RAID1的优点，但是空间上由于使用镜像，而不是类似RAID5的“奇偶校验信息”，磁盘利用率一样是50％ |

# Notepad++使用小技巧

## 1、查询多个条件

如我有100个数值，需要一次性搜索到，每个数值用括号括起来，然后用管道符号隔开如图





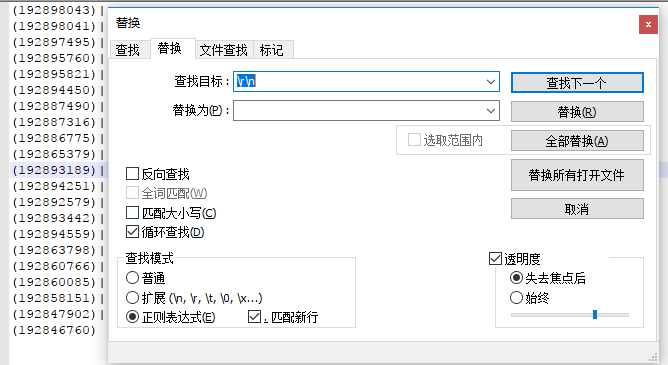
查到的结果如下



## 2、去掉换行

\r\n表示换行，其中\r表示回车，\s 空格

使用 \r\n 然后点替换就行



## 3、去掉空行、空格

去除行尾空格和空白行：按CTRL+H 选择正则表达式– 查找目标：\s+$ 替换为空

去除行首空格：按CTRL+H 选择正则表达式– 查找目标：^\s+ 替换为空

注： ^ 代表行首 $ 代表行尾

使用 ^\s+ 然后点击替换就行



## 4、去掉空格