第 35 卷　第 3 期　　电 子 科 技 大 学 学 报　　Vol.35　No.3
2006 年 6 月　　Journal of University of Electronic Science and Technology of China　　Jun. 2006

磁盘阵列RAID可靠性分析

陈华英

(中国民航飞行学院计算机学院　四川 广汉　618307)

【摘要】介绍了磁盘阵列的基本概念和常用的磁盘阵列种类，分析了RAID0、RAID3和RAID10三种磁盘阵列的可靠性值，同时与实验数据进行了比较。结果表明，磁盘阵列的可靠性值基于标准模型在一定程度上可以进行量化，能进行科学的计算，得出RAID10比相应其他的RAID的容量更大，可靠性更佳。该文的可靠性分析对磁盘阵列的进一步研究和生产能起到现实的指导作用。

关　键　词　磁盘阵列;　可靠性;　镜像;　磁盘阵列控制器
中图分类号　TP333　　文献标识码　A

Reliability Analysis of RAID

CHEN Hua-ying

(Dept. of Computer, Civil Aviation Flight University　Guanghan Sichuan　618307)

Abstract　This paper introduces the conception and common categories of Redundant Array of Inexpensive Disks (RAID). The reliability value for three RAIDs, RAID0, RAID3 and RAID10, is analyzed. The analysis results show that the value of the reliability for RAID base on standard model can be calculated, and RAID10 is better than other RAIDS on capability and reliability.

Key words　redundant array of inexpensive disks;　reliability;　mirror;　RAID controller

廉价冗余磁盘阵列(Redundant Array of Inexpensive Disks，RAID)是将一堆磁盘通过相关的技术连接起来，构成逻辑上的空间。RAID可以使诸多磁盘驱动器同时传输数据，在逻辑上它们又是一个磁盘驱动器，用RAID可以使单个的磁盘驱动器提高数倍的速率。因此，RAID技术正在变成支持文件共享、邮件交换和Web服务器等关键性服务的计算机系统所必备的一项标准功能[1]。磁盘阵列根据RAID控制器采用的工作模式和算法不同有不同的级别，因而具有不同的可靠性和性能。目前，常用的磁盘阵列有RAID 0、RAID1、RAID 0+1、RAID 3、RAID 5、RAID 6和RAID 10等[2-5]。本文将分析最具代表性的RAID 0，RAID 3和RAID 10三种磁盘阵列的可靠度。

1　磁盘阵列可靠度模型和可靠性值

1.1 磁盘阵列可靠度模型

一个磁盘阵列可以有包含多个不同的RAID级和磁盘容量，为简化模型，本文仅讨论基于8个相同RAID级硬盘驱动器组成的磁盘阵列，所有计算均假设在百分之百运转情况下3年内可靠性可以达到0.90的40 GB硬盘驱动器[2]。

1.2　RAID 0的可靠性值

磁盘阵列RAID 0 的结构图如图1所示，图中的磁盘阵列RAID 0由8个磁盘组成，因此磁盘阵列的可靠性值就是每一个单独的硬盘驱动器的可靠性的乘积。

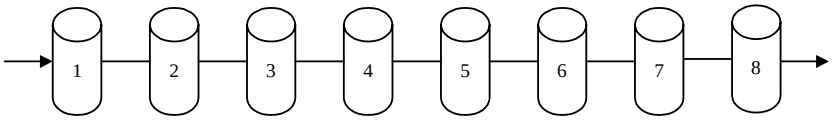

图1　RAID 0 的结构图

收稿日期：2006 – 01 – 26
作者简介：陈华英(1965 –)，女，副教授，主要从事数据库、计算机辅助教学等方面的研究.

由图1可知，1个磁盘阵列RAID 0 的可靠性值

$$R_{\mathrm{RAIDSET}} = \prod_{i=1}^{n}(R_{\mathrm{HDD}_i}) \tag{1}$$

根据1.1节的假设，一个单独的硬盘驱动器的可靠性值R_{HDD}是0.9，则8个磁盘组成的RAID0磁盘阵列的可靠性值

$$R_{\mathrm{RAIDSET}} = \prod_{i=1}^{8}(R_{\mathrm{HDD}_i}) = 0.9^8 = 0.430\ \ 5$$

计算表明，在3年内没有任何数据损失的可能性是43%，而发生数据损失的可能性是：

$$100\% - 43\% = 57\%$$

1.3 RAID 3的可靠性值

RAID 3属于并行访问磁盘阵列，其结构如图2所示。在RAID 3中，使用了专门的磁盘存放所有的校验值，并把数据分散到剩余的磁盘上，因而RAID 3需要至少3块盘阵来运行。RAID 3是使用专门校验盘的并行访问，其性能优势依赖于缓存技术和更高的磁盘转速。整个RAID 3系统只需要一个校验盘，就可以使硬盘的利用率提高。

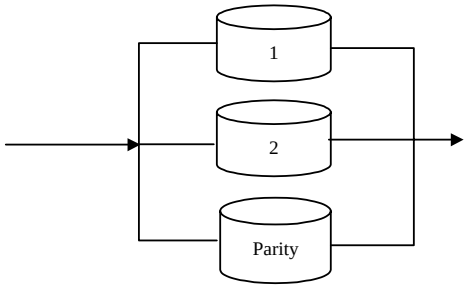

图2 RAID 3的结构图

RAID 3可以承受在由n个硬盘驱动器组成的磁盘阵列中的一个硬盘发生故障，例如：如果奇偶校验硬盘发生故障，其余用作存储数据用的硬盘不会被影响，但是冗余信息将丢失，如果其中一个数据硬盘发生故障，RAID控制器使用剩余的数据硬盘，并且奇偶校验硬盘重新计算丢失的数据，系统性能将有稍微的降低直到发生故障的硬盘驱动器被恢复，但是不会有数据丢失。如果在发生故障的硬盘恢复以前另一个硬盘发生故障，那么RAID组件内的所有的数据将会丢失。

图3中，配置了n块硬盘驱动器的RAID 3的可靠性值

$$R_{\mathrm{RAIDSET}} = \sum_{j=k}^{n}(n_j)R_{\mathrm{HDD}}{}^{j}(1-R_{\mathrm{HDD}})^{(n-j)} \tag{2}$$

因此，当配置了3块盘阵(共有8个相同RAID级硬盘驱动器，1个作为奇偶校验硬盘，7个作为数据盘)时，可靠性值

$$R_{\mathrm{RAIDSET}} = \sum_{J=7}^{8}(8_7)R_{\mathrm{HDD}}{}^{7}(1-R_{\mathrm{HDD}})^{(8-7)} = 0.813\ 104\ 73$$

计算表明，3年内没有任何数据损失的可能性是81%。因此，发生数据损失的可能性是：

$$100\% - 81\% = 19\%$$

1.4 RAID 10的可靠性值

RAID 10的结构如图3所示，图中RAID 10是将RAID 1的镜像功能和RAID 0的数据分割功能混合了起来，因此RAID 10需要至少4块硬盘驱动器(2块互为镜像盘，2块作为数据盘)来运行。其中数据是在被镜像的驱动器组之间分割，当一个硬盘驱动器在重建的时候，RAID 10系统将比奇偶校验RAID系统的性能更好，这是因为丢失的数据不是从奇偶校验信息来重新获得，而是从其余的驱动器上拷贝得到的。

在RAID 10磁盘阵列中，在镜像组里有一个硬盘驱动器发生故障不会导致任何数据丢失，只要每一个镜像组里有一个活动的驱动器，所有的数据就仍然是可用的，如果一个镜像组里的2个硬盘驱动器都发生了故障，那么所有的数据都会丢失。

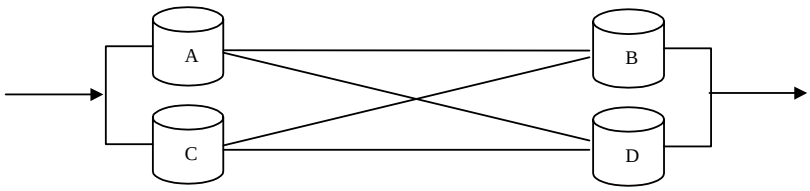

图3 RAID 10的结构图

图3中，RAID 10的一组阵列的可靠性值

$$R_{\mathrm{RAIDSET}} = \prod_{i=1}^{\#of\ RAIDsets} (R_{\mathrm{diskset}_i}) \tag{3}$$

当RAID 10分成4组时，

$$Let\, R_{\mathrm{diskset}} = P(AB \cup CD \cup AD \cup CB)$$

因此，1个RAID 10磁盘组的分组如下：

每次1组：$+P(AB)+P(CD)+P(AD)+P(CB)$

每次2组：$-P(ABCD)\ -P(ABD)\ -P(ABC)\ -P(BCD)\ -P(ACD)\ -P(ABCD)$

每次3组：$+P(ABCD)+P(ABCD)\ +P(ABCD)+P(ABCD)$

每次4组：$-P(ABCD)$

RAID 10的可靠性值是每组可靠性值之和，即：

$$R_{\mathrm{RAIDSET}} = [P(AB)+P(CD)+P(AD)+P(CB)-P(ABCD)-P(ABD)-P(ABC)-P(BCD)- P(ACD)-P(ABCD)\ +P(ABCD)+P(ABCD)\ +P(ABCD)\ +P(ABCD)-P(ABCD)]$$

将相同项相抵消，得到RAID 10的可靠性值

$$R_{\mathrm{RAIDSET}} = [P(AB)+P(CD)+P(AD)+P(CB)-P(ABD)-P(ABC)-P(BCD)-P(ACD)\ +P(ABCD)]$$

1个盘的可靠性值0.9，2个串联盘(如$P(AB)$)的可靠性值为：0.9^2=0.810 0；3个串联盘(如$P(ABD)$)的可靠性值为：0.9^3=0.729 0；4个串联盘(如$P(ABCD)$)的可靠性值为：0.9^4=0,656 1。因此，一个RAID 10的磁盘组的可靠性值

$$R_{\mathrm{RAIDSET}} =[0.810\ 0+0.810\ 0+0.810\ 0+0.810\ 0-0.729\ 0-0.729\ 0-0.729\ 0-0.729\ 0+ 0.656\ 1]=0.980\ 0$$

计算表明，1个RAID 10的磁盘组的可靠性值是0.980 0，对于一个由8块硬盘组成的RAID 10的可靠性值(也就是说包括2个RAID 10磁盘组)，其可靠性值是两组盘阵可靠性值的乘积，即：

$$R_{\mathrm{RAIDSET}} = (0.980\ 0)^2=0.960\ 4$$

计算表明，在3年内没有任何数据损失的可能性是96%，发生数据损失的可能性是：

$$100\% - 96\%=4\%$$

2　磁盘阵列可靠度对比

将设计的模型在实验室按照标准模型进行测试，结论和分析基本一致，由8块盘组成的RAID组可靠性对比数据如表1所示。

表1　RAID可靠性值实验数据

8块盘	3年内无任何数据损失的可能性/(%)	3年内发生数据损失的可能性/(%)
RAID 0	42.8	57.2
RAID 3	80.9	19.1
RAID 10	95.8	4.2

3　结　论

磁盘阵列有着广泛的应用，但它却又往往是整个网络存储系统安全运行的瓶颈，它的可靠性的量化计算越来越受到关注，因为有了量化的数据，就能为深层次地认识事物本质提供科学的依据，为多方案的优化比较奠定基础，进而做出正确的决定。本文在通过大量的理论研究及实验分析基础之上，得出RAID 0、RAID 3 和RAID 10三种磁盘阵列基于标准模型的可靠性值，其中RAID 10比相应其他RAID的容量更大，可靠性更佳。研究磁盘阵列的可靠性值，提出可靠性值的量化计算公式，在磁盘阵列理论研究和实际的开发应用中都具有重要指导意义。

本文得到了中国民航飞行学院科研基金“设备虚拟网(J2004-23)”的支持，在此表示感谢。

参　考　文　献

[1] 赵文辉. 网络存储技术[M]. 北京: 清华大学出版社, 2005.
[2] TOM C . 存储区域网络设计[M]. 邓劲生, 李宝峰, 李　蕾, 等译. 北京: 电子工业出版社, 2005.
[3] 周　可, 冯　丹, 王　芳, 等. 网络磁盘阵列流水调度研究[J]. 计算机学报, 2005(3): 319-325.
[4] 李　琼. 高可靠磁盘阵列的设计[J]. 计算机应用研究, 2003(3): 49 -51.
[5] 刘劲松. 关于存储系统性能的测试、仿真与评价的研究[D]. 武汉: 华中科技大学, 2004.

编　辑　徐安玉