杭州应用工程技术学院学报,第 13 卷第 2 期,2001 年 6 月
Journal of Hangzhou Institute of Applied Engineering
Vol.13 No.2, Jun. 2001

利用冗余技术预防数据灾难

黄明达　　马永成
(杭州应用工程技术学院 现教中心　杭州 310012)　(杭州赛阳电子科技有限公司存储备份实验室)

摘　要　由各种原因导致的数据不可用所带来的重大损失称为数据灾难,数据灾难的预防是一门正在研究和完善的新技术,本文分析了利用 RAID 磁盘阵列冗余技术实现数据存储领域数据灾难预防的技术,并简要介绍了备份、SAN 等其它预防技术,作为进一步数据保护和灾难恢复技术讨论的敲门砖.

关键词　冗余　数据灾难　RAID　SAN
中图分类号　TP306.3

1　数据灾难

如果有一天你有急事手拿信用卡去银行取钱时,却被告知电脑系统故障,要修复后才能使用;如果有一天,你在家中的电脑由于儿子按了几个键,你花三个月时间开发的程序和长长的论著被删除;如果有一天身为网管的你,突然发现公司的服务器死机了,而公司的合同要打印,销售的定单要 mail,财务的现金要进帐,这时你的感觉是什么? 你一定感觉这是一场灾难,更准确的说这是一场数据灾难.

我们所讨论的数据灾难是指能引起数据丢失和操作故障的一切有害事件[1].从广义上讲数据灾难包括任何原因造成的数据丢失或不可用,如无意删除文件和记录的人为失误,重大的系统故障,电击和电力异常引起的电源故障,由洪水、雷击、火灾等自然现象或盗窃、阴谋活动、病毒人为原因造成的损害程度不一的故障等等."数据灾难"所造成的数据不可用,及其带来的巨大损失是十分明显的.而"数据灾难"的预防和恢复还不太为人所了解,大多数的计算机使用者面对严重的数据灾难束手无策,因此认识数据灾难,研究预防技术,对于保护劳动成果避免意外损失具有相当现实的意义.

2　数据灾难预防的有关术语与概念

数据灾难的预防是一个正在继续研究和不断完善的新课题,先了解一下有关的术语与概念.

收稿日期:2001 - 01 - 09

2.1 容错与容错类型

容错对于不同形式下有不同的理解,一般认为容错是在预定时间内从系统故障中恢复正常的能力.如停电到UPS供电时间是很短,而如要重装系统与恢复数据,要化几小时到几天的时间不等.因此容错决定于硬件与软件应用及恢复时间.从时间长短容错可分为实时、短时和长时.

实时容错是最有实际意义的.银行7天工作、电厂一周每天24 h不间断供电、证券交易时刻发生,1 s的工作停顿都会有不可估计的损失,所以在这种情况下,任何的故障(硬件或软件)一定要在"瞬间"恢复到正常状态.

短时容错是指系统故障后可以有1~2 h时间挂起.在此期间可用人工修复故障或将服务由另一机器来替代.

长时容错是指在出现故障后可以允许几小时到几天的时间来恢复.这种恢复往往是从最初原始的状态开始.

但是无论时间长短,容错中最为重要的原则是要保证数据完整、正确、可用,如不能保证这一点则容错就失去了意义.

2.2 冗余

冗余的意思是"不必要的重复",对于一个正常的系统来说,冗余的部件(如UPS)并不能带来好处,可以说是不必要的,冗余这个定义是十分恰当的.但从可靠性的角度来说冗余是必要的,有了UPS毕竟能让我们的工作不会因停电而出现问题.冗余现在已经深入到各个层次,如外部电源冗余(UPS),内部部件冗余(双电源,双CPU),存储器冗余(RAID等级阵列),软件冗余(安装并行系统),数据冗余(数据镜像)等.

2.3 备份

备份是一种古老的、安全的、简单有效的预防灾难的手段.将数据复制到磁盘或磁带中保存下来,出错时再复制回去.但这样简单的方法,许多人往往是经过一次"灾难"后才会去做.当然备份也存在容量小、速度慢的缺点.

2.4 群集

群集对于个人来说是一项比较新的技术,但就大型机而言这是一项成熟的技术.现在NT Enterprise Server (NT企业版)使用的是两路交换最简单的群集技术,在这种情况下,二台主机使用同一存储设备(一般为RAID5阵列),当其中一个系统出现故障时,另一台会在30 s内接管工作.它将使我们的服务器系统全天24 h安全的工作成为可能.

2.5 RAID

RAID(redundant arrays of inexpensive disks)是指冗余的廉价磁盘阵列.它是一种实现磁盘冗余的技术.它有许多等级使用于不同方面,但在实际中RAID1(磁盘镜像方式)和RAID5(磁盘条块化方式)在灾难预防与恢复中运用最为广泛.

3 利用冗余技术预防数据灾难

可能引起数据灾难的原因是多种多样的,如电源故障,自然原因,硬件错误,人为误操作,软件BUG,Internet病毒感染与安全侵害.对此人们有了不同的预防方法.如在机房中安装避雷电路,实现可靠接地等等来预防自然原因引起的数据损坏,而对于病毒与安全侵害更是形成了一个Internet安全的专门学科,而在数据存储领域对数据灾难的预防可以使用冗余技术.

3.1 硬件存储器冗余技术(RAID)

在目前硬件存储器最为主要的是硬盘.当硬盘出现I/O错误时,不用说大家也知道,数据的完

整性、可用性和操作性等各个方面都会受到影响；而在实际的操作过程中，这种错误是不可避免的，而且我们又不能预测它的发生，因此就有了 RAID 技术及标准的提出.

当然 RAID 技术并不是单纯的针对容错的，它包含的范围更广泛.主要有海量存储、高速 I/O、利用冗余实现容错 3 个方面.我们结合 RAID 级别来说一下 RAID 技术[2].

JBOD 它只是将驱动器简单地放在一起提供整个的存储空间，也可称为数据跨盘(海量存储).

RAID0 提供的所有并行的 I/O，也就是说一个数据流分成几个部分同时写入到几个驱动器中，在读的时候也是同时读出，这样随着驱动器数目的增加而读写性能也会飞速增加(高速 I/O).

RAID1 提供的是将数据同时写入两个驱动器，这样即使有一个驱动器坏了另一个驱动器也可以提供正确数据(冗余).

RAID3 提供将其中的一个驱动器用于存放校验数据，当有一个驱动器出错时可用校验方式修复数据.

RAID5 是一个折衷方式，它只不过是将校验数据存放在不同驱动器中以提高速度.

在实际工作中，RAID1 与 RAID5 这两种运用最为广泛.

RAID1 的实现如图 1 所示.当数据流 A 写入驱动器时实际存放的两组数据 A 和 A′，这样如果 A 中的数据单元 A1 出错，我们可以从 A1′中得到正确的数据，同理当 A1′中出错时可通过 A 中的数据来修复.所以只要 A 和 A′中相同数据单元不同时损坏数据就不会丢失.同时这项技术的实现是非常简单的，因此这种“镜像”技术被最早运用到实际当中.

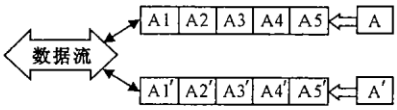

图 1　RAID1 实现模型

RAID5 的实现如图 2 所示，这里可以用一个简单的例子来说明：对于 A + B + C + D = E 这个式子来说，如果 E 出错我们无法直接得到，那么就可以通过 A，B，C，D 计算得到.同样 A 出错时，我们可以通过(E - B - C - D)来实现(这个过程被称为重建).这样我们知道 A，B，C，D，E 只要不是有两个同时出错，那么数据就是安全的，同时如图 2 它还将数据分别放于不同的驱动器中实现并行的传输.

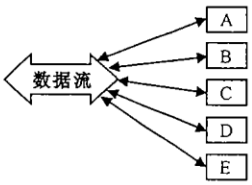

图 2　RAID5 实现模型

RAID 的具体实现有软件方式和硬件方式两种.硬件方式是通过专用的 RAID 卡或 RAID 阵列柜来实现，软件是通过 NT4.0 或 NT2000 中的实用程序来实现.从实际效果来说硬件方式可靠性要高于软件方式，所以在实时容错与短时容错系统中大多使用硬件方式.

3.2　系统的冗余技术(并行系统和群集)

系统的冗余技术主要针对的是当操作系统出错时的数据修复.一般来说，操作系统为了针对出错都会要求你制作应急启动盘.但这并不都是安全而有效的.因此就有了并行的操作系统.所谓的并行操作系统是指在同一主机上安装两个相同的操作系统.当其中的一个操作系统出现错误无法运行时，可用另一后备系统运行.这对于简单要求或单机系统来说具有一定的意义.

除了并行系统外另一种技术是群集技术，NT 的一个重要功能是提供备份域服务器(BDC)，它的作用是在主域服务器(PDC)出现问题时，可以被人工提升为主域服务器以完成管理域内的工作.但 BDC 本身不能实现接替 PDC 上的一些文件服务(如数据库).而在 NT 企业版中这种功能被自动实现，而不再需要人工干预.同时对 SQL 数据库还能实现数据库的服务切换，这对于实际的运用有十分重要的意义.虽然 NT 企业版对其它的数据库如(Oracle、Sybase、Informix)不支持，但有大量的第三方软件来实现数据库切换的功能.这种非监督的自动切换功能为实时的容错提供了可能.

3.3　数据冗余技术(备份)

数据冗余主要针对的是数据的破坏未被实时监控发现的情况.这时数据已经被破坏，又未能及

时得到恢复,事后也无法从磁盘镜像等方式得到修复,那么以上所有的技术都失去了作用.有效的方法只有建立数据备份,有了备份在数据灾难过后,你可能失去一些新的数据,但是你却有一个安全放心、可靠可用的数据.

备份从严格意义上来说并不能说是一种新的技术,但是备份技术也在发展前进.主要的新技术出现在软件上,备份主要难题是针对活动数据和操作系统相关的打开文件的备份,因为现在的大型数据库的数据是每时每刻变化的,在备份时如何保证数据的一致性就很困难.一般的作法是对数据进行锁定,在备份的过程中锁定数据不进行读写.新技术是进行"快照"方式,在备份前将数据快速的复制到另一个地方,再进行备份,这样不用锁定数据对正常工作的影响就减小了.此外由于数据备份总是迟于活动数据,所以在恢复到灾前数据时,会有一些数据丢失,如何减少丢失数据就成为备份的又一新课题.

3.4 SAN技术

SAN(Storage Area Network)存储区域网.在最纯粹的意义上,是一个单独的计算机网络,特点是基于光纤通道技术(Fibre Channel)的电缆,交换机和集线器,将很多的存储设备连接起来,再与有很多不同的服务器组成的网络相连接,以多点对多点的方式进行管理.SAN还可以让存储设备与存储设备直接相连,并基于新的集群技术和直接与网络连接的存储设备技术(Network Attached Storage),集成多台服务器与多台磁带库磁盘阵列,当网上的一台服务器出现问题,网上能实现同一功能的服务器就可代替工作,当一个存储设备出现问题,服务器可访问另一个有相同数据的存储设备,这样存储系统真正独立于主机系统外,实现了以数据为中心的独立保护系统,从而实现数据的保护.

4 结束语

以上只是介绍数据灾难预防的一些技术,现今还有一些技术正在发展之中,如数据异地镜像、数据集中处理、网络冗余等,为数据灾难预防又开辟一个新的天地.数据灾难预防的研究永不会停止.

参 考 文 献

1 John McMains, Bob Chronister. Windows NT backup & recovery. New York: McGraw-Hill, 1998.1 ~ 7

2 Marc Farley, Tom Stearns, Jeffrey Hsu. LAN times guide to security and data integrity. 李明之等译. 北京:机械工业出版社,1998.82 ~ 85

Preventing data-disaster with redundancy techniques

Huang Mingda Ma Yongcheng

(Hangzhou Institute of Applied Engineering, Hangzhou 310012) (Hangzhou San Electronic Technology CO.)

Abstract Data backup and recovery is a developing technique. This paper introduces appling redundancy to prevent data-disaster, and some preventing techniques, such as backup and SAN. It's also helpful for more researching in the field of data backup and recovery.

Key words redundancy data-disaster RAID SAN