



大话 数据恢复

◎ 陈培德 编著

清华大学出版社


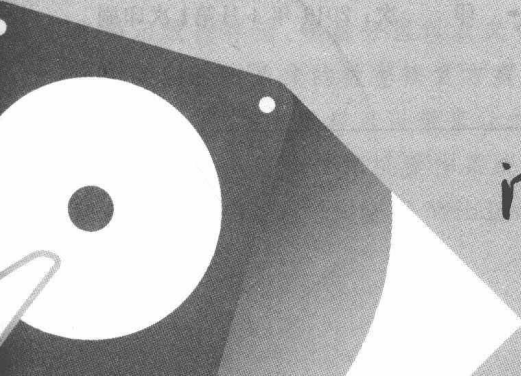




大话 数据恢复

◎ 陈培德 编著

清华大学出版社



清华大学出版社
北京

内 容 简 介

本书介绍了数据恢复的基本概念、硬盘相关知识、虚拟硬盘工具与磁盘编辑软件 WinHex 的使用；以实例形式详细讲解 MBR 分区与 GPT 分区的管理方式, FAT32 文件系统整体布局、文件及文件夹的管理方式等, NTFS 文件系统整体布局、元文件的作用、索引目录的管理等；以实例和案例形式介绍了数据恢复的基本思路、方法与步骤。

读者通过每章的学习并完成每章思考题后, 不仅加深了对每章知识的理解与掌握, 而且还增强了解决实际问题的能力。

本书内容丰富, 讲解由浅入深、通俗易懂、重点突出、示例翔实。在内容编排上, 系统全面、新颖实用、可读性强。

本书适用于高等院校计算机相关专业学生, 同时也适用于从事数据恢复、电子取证以及其他有关人员自学、参考等。

本书封面贴有清华大学出版社防伪标签, 无标签者不得销售。

版权所有, 侵权必究。侵权举报电话: 010-62782989 13701121933

图书在版编目(CIP)数据

大话数据恢复/陈培德编著. —北京: 清华大学出版社, 2019
ISBN 978-7-302-50693-5

I. ①大… II. ①陈… III. ①数据管理—安全技术 IV. ①TP309.3

中国版本图书馆 CIP 数据核字(2018)第 163544 号

责任编辑: 贾 斌

封面设计: 刘 键

责任校对: 李建庄

责任印制: 李红英

出版发行: 清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址: 北京清华大学学研大厦 A 座

邮 编: 100084

社 总 机: 010-62770175

邮 购: 010-62786544

投稿与读者服务: 010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈: 010-62772015, zhiliang@tup.tsinghua.edu.cn

课件下载: <http://www.tup.com.cn>, 010-62795954

印 装 者: 清华大学印刷厂

经 销: 全国新华书店

开 本: 185mm×260mm

印 张: 25.25

字 数: 648 千字

版 次: 2019 年 4 月第 1 版

印 次: 2019 年 4 月第 1 次印刷

印 数: 1~1500

定 价: 79.80 元

产品编号: 078661-01

FOREWORD

前

言

人们常用“硬盘有价、数据无价”来形容存储在外存储器中数据的重要性。但是由于各种原因(如:计算机病毒的破坏,用户误 Ghost、误删除文件、误分区、误格式化,外存储器物理损坏或者其他原因等),导致存储在外存储器中数据丢失的现象时有发生。

当数据丢失后,能否很好地保护现场,并找回丢失的数据就显得十分重要。于是“数据恢复”这个在国外已经使用了二十多年,而在国内却鲜为人知的名词和技术也逐渐被国人所接触和使用。与此同时,国内一些出版社已出版了有关数据恢复方面的书籍。这些书籍从不同的角度讲解了数据恢复的一些基本理论、方法和步骤。但是“数据恢复”要想走进高校课堂并成为计算机相关专业的课程仍然还有很长的路要走。

为此,经作者长时间的思考,认为有必要撰写一本有关“数据恢复”的书籍,以满足高校开设该课程的需要。如何撰写?根据作者长期的教学、实践经验,只有理论与实践相结合的书籍才最受读者欢迎。因此,本书在撰写过程中,始终坚持硬盘分区与文件系统的基本理论和数据恢复的实践紧密结合的原则。本书中所列举的每一个实例均能够找到相应的实践素材,读者在阅读本书过程中,可以按本书中的实例边阅读、边操作,这样可以在较短的时间内,加深对硬盘分区、文件系统基本理论的理解与掌握,本书为读者开展数据恢复工作提供了强有力的理论依据;通过数据恢复案例,不仅增强了读者的实际动手能力,而且提升了读者的操作技能。

俗话说“树有根、水有源”,任何形式的数据丢失,都有其原因所在,只有找到数据丢失的真正原因,才能对症下药,快速恢复所需数据。通过本书的学习,读者不仅能够掌握数据存储的基本结构、硬盘分区和文件系统的基本原理,而且还能以该原理为基础来查找数据丢失的真正原因,在较短的时间内制订数据恢复方案、以最好的方式恢复丢失的数据。

如果你是一位数据恢复的初学者,本书将带你步入数据恢复领域的殿堂,为你揭示数据存储、硬盘分区以及文件系统的神秘面纱,让你找到数据丢失的真正原因并制订出最佳的数据恢复方案,最终恢复所需数据。如果你已经是一位数据恢复的能手,本书同样可以带来让你万分惊喜的一些经验与技巧。

如果你使用的计算机操作系统是 Windows 7,你可以将本书所提供的素材复制到计算机的硬盘上,使用 Windows 7 操作系统计算机管理中的硬盘管理功能

将素材文件附加成为虚拟硬盘,即可按本书的章节来研究学习。

如果你使用的计算机操作系统是 Windows XP,可以下载并安装虚拟磁盘管理软件(如 InsPro Disk v2.0),将本书所提供的素材复制到计算机的硬盘上,并将素材文件的扩展名“.vhd”改为“.hdd”,使用虚拟磁盘管理软件将素材文件加载成为虚拟硬盘,同样也可以按本书的章节来研究学习。

全书共分为7章,第1~3章介绍了数据恢复的基本概念、硬盘相关知识、虚拟硬盘工具与磁盘编辑软件 WinHex 的使用等。第4章以实例的形式详细讲解了 MBR 分区与 GPT 分区的存储形式和管理方式等;第5章以实例的形式详细讲解了 FAT32 文件系统整体布局、文件及文件夹的管理方式等;第6章以实例的形式详细讲解了 NTFS 文件系统整体布局、元文件的作用、NTFS 对索引目录的管理等;第7章以实例和案例的形式介绍了数据恢复的基本思路、方法与步骤。

每章后均有大量的思考题,读者通过每章的学习并完成每章思考题后,不仅可加深对每章知识的理解与掌握,而且增强了实际动手能力和解决实际问题的能力。

第1~7章、各章思考题、第3~7章所需素材、思考题所需素材和思考题参考答案由陈培德老师完成;云南大学图书馆殷莉芬老师、云南大学信息学院刘洪涛老师对本书进行了部分排版及校对工作;全书由陈培德老师最后完成审稿及校对工作。

在本书的策划与撰写过程中,得到了电子科技大学计算机科学与工程学院计算机与网络取证与鉴定实验室主任刘乃琦教授,云南大学信息学院杨鉴、代红兵、岳昆、王丽清、吴建平、周永录、王云峰,云南大学软件学院姚绍文等专家学者的大力支持,并提出了许多建设性的意见和建议,在此表示由衷感谢。同时也感谢清华大学出版社的支持与帮助,使得本书得以顺利出版。

学习的道路是没有尽头的,作者非常愿意与广大读者进行交流,共同学习、共同进步、共同提高,同时也非常愿意为广大读者提供帮助和技术支持。

读者在阅读本书的过程中,有什么好的意见或建议,请通过 QQ(QQ 号:1814433586)告知作者,作者将不胜感激,并虚心接受。

由于作者水平有限,书中难免存在某些疏漏和不足,恳请读者批评、指正。

本书中各章所使用到的素材、思考题素材,读者可从清华大学出版社网站自行下载。网址:<http://www.tup.com.cn>。

如需各章思考题参考答案,请通过 QQ 直接与作者联系。

作者:陈培德
云南大学信息学院
2018年10月

CONTENTS

目

录

第 1 章 概述	1
1.1 数据恢复	1
1.1.1 数据恢复定义	1
1.1.2 常用数据恢复硬件和软件	1
1.1.3 数据恢复需要注意的事项	2
1.1.4 数据恢复应用领域	3
1.1.5 数据恢复从业人员	3
1.2 数据恢复相关知识	4
1.2.1 数据的表示方式	4
1.2.2 数据的存储形式	6
1.2.3 计算机的启动过程	7
思考题	8
第 2 章 硬盘相关知识	13
2.1 硬盘基础知识	13
2.1.1 硬盘物理结构与工作原理	13
2.1.2 硬盘主要接口技术	15
2.1.3 硬盘性能指标	17
2.1.4 盘面、磁道、柱面与扇区	18
2.1.5 硬盘寻址模式	20
2.1.6 硬盘故障	21
2.1.7 操作计算机时需要注意的事项	22
2.2 硬盘分区	23
2.2.1 硬盘分区作用	23
2.2.2 硬盘分区类型	23
2.3 文件系统	24
2.3.1 文件系统定义	24
2.3.2 常见文件系统	24
思考题	27

第 3 章 虚拟硬盘工具与 WinHex 的使用	28
3.1 虚拟硬盘工作原理	28
3.2 虚拟硬盘工具 InsPro Disk 使用介绍	28
3.2.1 安装 InsPro Disk	28
3.2.2 卸载 InsPro Disk	28
3.2.3 InsPro Disk 使用介绍	29
3.3 Windows 7 虚拟硬盘工具使用介绍	30
3.3.1 创建虚拟硬盘文件	31
3.3.2 附加虚拟硬盘文件	31
3.3.3 初始化虚拟磁盘	32
3.3.4 分离虚拟硬盘	33
3.3.5 虚拟硬盘文件的特点	33
3.4 WinHex 简介	34
3.4.1 安装 WinHex	34
3.4.2 启动 WinHex	35
3.4.3 WinHex 主界面	35
3.4.4 WinHex 菜单介绍	39
3.4.5 使用 WinHex 注意事项	46
思考题	46
第 4 章 Windows 平台下硬盘分区	49
4.1 硬盘分区与主引导扇区结构	49
4.1.1 低级格式化	49
4.1.2 初始化	49
4.1.3 建立分区	50
4.1.4 删除分区	51
4.1.5 高级格式化	51
4.1.6 读/写主引导扇区	52
4.1.7 主引导扇区结构	52
4.2 硬盘 MBR 分区	56
4.2.1 MBR 分区分类	56
4.2.2 MBR 分区表存储方式	56
4.2.3 MBR 分区表管理方式	58
4.3 硬盘 GPT 分区	71
4.3.1 硬盘 GPT 分区简介	71
4.3.2 在 GPT 磁盘上建立分区	71
4.3.3 硬盘 GPT 分区的整体结构	74
思考题	79

第 5 章 FAT32 文件系统	91
5.1 FAT 文件系统概述	91
5.2 FAT16 文件系统	91
5.2.1 FAT16 文件系统介绍	91
5.2.2 FAT16 文件系统组成	92
5.2.3 FAT16 逻辑扇区号与簇号	92
5.3 FAT32 文件系统	93
5.3.1 FAT32 文件系统介绍	93
5.3.2 FAT32 文件系统总体布局	93
5.3.3 FAT32 分区引导扇区	94
5.3.4 计算 FAT32 有关参数	98
5.3.5 FAT32 FSINFO 信息扇区	100
5.3.6 FAT 表	101
5.3.7 FAT32 目录项	104
5.3.8 根目录、子目录与回收站	113
5.3.9 删除文件对目录项、回收站等的影响	120
思考题	127
第 6 章 NTFS 文件系统	140
6.1 NTFS 文件系统概述	140
6.1.1 NTFS 文件系统简介	140
6.1.2 NTFS 总体布局	141
6.1.3 NTFS 引导扇区	144
6.1.4 有关 NTFS 容量计算公式	148
6.2 \$MFT 记录结构	150
6.2.1 \$MFT 概述	150
6.2.2 \$MFT 记录分类	150
6.2.3 \$MFT 记录结构	151
6.2.4 记录头结构	152
6.2.5 记录属性	154
6.2.6 记录属性分类	155
6.3 文件和文件夹记录常用属性	160
6.3.1 10H 属性	160
6.3.2 30H 属性	162
6.3.3 80H 属性	165
6.3.4 90H 属性	173
6.3.5 A0H 属性	186
6.3.6 B0H 属性	188
6.4 文件和文件夹记录不常用属性	188

6.4.1	20H 属性	188
6.4.2	40H 属性	189
6.4.3	50H 属性	189
6.4.4	60H 属性	190
6.4.5	70H 属性	191
6.4.6	C0H 属性	192
6.4.7	D0H 属性	192
6.4.8	E0H 属性	193
6.4.9	100H 属性	193
6.5	NTFS 文件系统的元文件	193
6.5.1	元文件 \$MFT	193
6.5.2	元文件 \$MFTMirr	203
6.5.3	元文件 \$LogFile	203
6.5.4	元文件 \$Volume	204
6.5.5	元文件 \$AttrDef	204
6.5.6	元文件“.”	204
6.5.7	元文件 \$Bitmap	208
6.5.8	元文件 \$Boot	213
6.5.9	元文件 \$BadClus	213
6.6	索引节点结构	215
6.6.1	索引节点介绍	215
6.6.2	索引节点分类	216
6.6.3	叶节点	216
6.6.4	非叶节点	220
6.7	NTFS 对索引目录的管理方式	224
6.7.1	B-树的定义及特征	224
6.7.2	B-树在 NTFS 索引目录管理中的应用	224
6.8	根目录的结构	228
6.9	回收站的结构	232
6.9.1	Windows XP 回收站结构	232
6.9.2	Windows 7 回收站结构	232
6.10	删除文件对元文件 \$MFT 和索引目录等的影响	233
	思考题	240
第 7 章	数据恢复	273
7.1	恢复分区	273
7.1.1	硬盘主引导扇区被破坏后的现象	273
7.1.2	恢复硬盘主引导记录	273
7.1.3	恢复硬盘分区表	274
7.1.4	恢复误 Ghost 后的数据	284

7.1.5	恢复硬盘 GPT 分区	293
7.2	恢复 DBR	298
7.2.1	DBR 被破坏后的现象	298
7.2.2	恢复 FAT32_DBR	300
7.2.3	恢复 NTFS_DBR	308
7.3	恢复 DBR 与分区表	313
7.3.1	恢复 FAT32_DBR 与分区表	314
7.3.2	恢复 FAT32_DBR、FAT32_DBR 备份与分区表	316
7.3.3	恢复 NTFS_DBR 与分区表	320
7.3.4	恢复 NTFS_DBR、NTFS_DBR 备份与分区表	321
7.4	文件被删除后的恢复	325
7.4.1	删除文件的基本方法	325
7.4.2	文件被删除后的基本情况(FAT32)	326
7.4.3	恢复已删除的文件(FAT32)	326
7.4.4	文件被删除后的基本情况(NTFS)	329
7.4.5	恢复已删除的文件(NTFS)	330
7.5	Windows XP 下(快速)格式化后的数据恢复	331
7.5.1	FAT32 被(快速)格式化成 NTFS 的恢复	331
7.5.2	FAT32 被(快速)格式化成 FAT32 的恢复	334
7.5.3	NTFS 被(快速)格式化成 FAT32 的恢复	338
7.5.4	NTFS 被(快速)格式化成 NTFS 的恢复	340
7.6	Windows 7 下快速格式化后的数据恢复	344
7.6.1	FAT32 被快速格式化成 NTFS 的恢复	344
7.6.2	FAT32 被快速格式化成 FAT32 的恢复	344
7.6.3	NTFS 被快速格式化成 FAT32 的恢复	349
7.6.4	NTFS 被快速格式化成 NTFS 的恢复	353
7.7	Windows 7 下格式化后的数据情况	356
7.8	数据恢复案例	356
	思考题	374
	参考文献	393

第 1 章

概 述

1.1 数据恢复

随着信息技术的飞速发展和无纸化办公时代的到来,计算机在人们的工作和生活中扮演着越来越重要的角色。企业、商家、银行、政府机关、事业单位等通过计算机来获取和处理信息,同时也将重要信息以数据的形式保存在计算机的外存储器上。这些数据一旦丢失,将给企业、商家、银行、政府机关、事业单位等造成无法挽回的损失。因此,数据丢失后,能否很好地保护现场并找回丢失的数据就显得十分重要。于是,数据恢复,这个在国外已经使用了二十多年,而在国内却鲜为人知的技术也逐渐被国内人所接触和使用。

1.1.1 数据恢复定义

什么是数据恢复呢?到目前为止,数据恢复还没有一个统一的定义,但有一个大家公认的提法。数据恢复是指外存储器硬件损坏或者用户误操作、误分区、误格式化、误删除文件、计算机病毒破坏等导致存储在外存储器中的数据无法通过正常方式进行存取,只有通过特殊的方式将所需要的数据恢复到正常状态,以便进行正常的存取或将其存储到其他外存储器的过程。

数据恢复一般分为“硬恢复”和“软恢复”两种。所谓“硬恢复”是指外存储器在物理上出现问题而导致数据无法正常读取的恢复;也就是说,由于外存储器出现物理问题所引起的故障,对此类故障进行的数据恢复,一般称为“硬恢复”。而“软恢复”则是指逻辑故障(如:用户误操作、误分区、误格式化、误删除文件、误 Ghost、硬件逻辑锁、操作过程中突然掉电、病毒破坏等原因)导致数据无法通过正常的方式进行存取,使得数据发生丢失,而存储介质不存在任何物理故障,对此类故障进行的数据恢复,一般称为“软恢复”。

“硬恢复”需要对存储介质的结构及工作原理相当了解;而“软恢复”则需要对硬盘分区和文件系统等有足够的认知。

1.1.2 常用数据恢复硬件和软件

随着数据恢复行业的逐渐兴起,一些数据恢复公司先后研发了一些数据恢复硬件产品,常

用的数据恢复硬件产品如下。

(1) DC 一体机计算机取证恢复专业设备：该设备是目前最先进的全球第四代专业数据恢复工具，也是全球首款全功能性数据恢复一体式设备，是一款高智能化专业数据恢复设备；该设备配备了 USB 接口，在移动性、便携性、功能性等方面表现强悍，其数据恢复以及计算机取证成功率较高。

(2) SD II 9000 服务器取证专业设备：该设备支持所有品牌的 SAS/SCSI 接口硬盘，支持所有的文件系统格式、各种服务器磁盘阵列类型、数据库类型等。

(3) SAS/SCSI 数据擦除一体机：该设备是 SAS/SCSI 存储介质数据擦除销毁设备巅峰之作，冠绝全球，兼容所有品牌 SAS/SCSI 接口硬盘，全面支持 IBM、HP、DELL、SUN、联想、浪潮、华硕、曙光、长城、清华同方、方正、天翱、Acer、AblestNet NE 等市面上所有品牌服务器，其特点是 Windows 界面、一体工控键盘设计、操作简单。

(4) FLASH 闪存数据恢复大师设备：该设备是一台专门针对 U 盘、CF 卡、记忆棒、录音笔等 FLASH 存储介质进行数据提取的专业 FLASH 数据恢复设备。

(5) 智能数据指南针(Data Compass)专业设备：该设备是一款专门针对硬盘逻辑层、固件层、物理层故障数据恢复和数据提取的高智能、高效率的专业设备。

(6) 硬盘复制机——DATA COPY KING：该设备是目前全球最先进的全领域硬盘复制产品，融合了硬盘高速复制、数据高速复制、安全擦除和故障自动检测的高性价比一体设备，硬盘复制机采用了效率源科技 2010 年最新技术，专为 TB 级大容量硬盘而设计，最大支持 131072TB。硬盘复制速度、擦除速度、对缺陷扇区的数据获取能力均超过市场同类硬盘复制产品。

与此同时，一些数据恢复公司也先后开发了一些数据恢复软件，如：EasyRecovery、Anedata、安易数据恢复软件、Get Data For FAT32/NTFS、WinHex、FinalData、R-studio、Recover my file、易我数据恢复向导、易我分区表医生、DiskGenius、顶尖数据恢复软件、效率源数据恢复软件，等等。

1.1.3 数据恢复需要注意的事项

在数据恢复过程中，应注意以下 6 点。

(1) 在数据恢复过程中最怕被误操作而造成二次破坏，导致数据恢复的难度陡增。因此，在数据恢复过程中，严禁再向要恢复数据的外存储器中写入新的数据。

(2) 严禁做磁盘检查：一般文件系统出现错误后，系统开机进入启动界面时，会自动提示——是否需要做磁盘检查？大约 10 秒后，开始进行磁盘检查；这种操作有时候可以修复一些比较小的损坏目录或文件，但是很多时候则会破坏数据链表。因为复杂的目录结构是无法修复的。当修复结束后，会在根目录下产生以“FOUND. XXX”(其中：XXX 为 000 至 999 之间的数字)命名的文件夹，文件夹里有大量的以“.CHK”为扩展名的文件。有时候这些文件重命名后就可以直接恢复，而有时候则不能，特别是比较大且不连续存储的文件。

(3) 严禁再次格式化逻辑盘或卷：如果再次对逻辑盘或卷进行格式化，将会给数据恢复带来更大的困难，数据可能无法恢复。

(4) 不要把数据直接恢复到源盘上：很多普通客户删除文件后，使用数据恢复软件将恢复出来的文件直接存储到原来的外存储器中，这样破坏原来数据的可能性非常大。因此，严禁

直接将数据恢复到源盘上。

(5) 最好不要使用分区工具重建分区：对分区原理不熟悉的数据恢复人员而言，如果分区被破坏后，最好不要使用分区工具重建分区，这样很容易破坏分区内的原来文件系统中的重要参数，从而导致数据恢复的难度大大增加。

(6) 服务器磁盘阵列丢失后不要重做磁盘阵列重组：在挽救服务器阵列的实践中遇到过有些网管员，在服务器崩溃后强行让阵列上线，即使掉线了的硬盘也强制上线，或者直接做 Rebuilding 命令。这些操作都是非常危险的，任何写入盘的操作都有可能破坏原来的数据。

总之，当数据丢失后，严禁向盘里存入任何新数据。建议关闭计算机，然后把硬盘卸下，连接到别的计算机上作为辅盘，先将该硬盘上的数据通过克隆的方式备份到新的硬盘上，再进行数据恢复操作。

1.1.4 数据恢复应用领域

电子证据第一次出现是在 1998 年，当时某公安机关网安部门在侦办某网络案件时对有关证据进行了提取，并被法院采纳。在具体法律规定方面，我国较发达国家而言相对晚一些。一方面在于可以用于证明案件事实的材料都是证据，与案件相关的电子证据自然属于证据范畴；另一方面在于刑事诉讼法中将证据种类限定为 7 种，并没有设定电子证据。2012 年 3 月 14 日，第十一届全国人民代表大会第五次会议通过了《关于修改〈中华人民共和国刑事诉讼法〉的决定》。根据该决定，电子证据成为法定证据类型，这适应了现代化技术的发展需要，同时也丰富了证据范围。

随着计算机犯罪数量的不断上升和犯罪手段的数字化，搜集电子证据的工作成为提供重要线索及破案的关键。恢复已被破坏的计算机数据并提供相关的电子资料证据就是电子取证。具体来说，电子取证就是利用计算机硬件和软件技术，以符合国家的法律、法规等方式对计算机入侵、破坏、欺诈、攻击等犯罪行为进行证据获取、保存、分析和出示的过程。从技术方面看，电子取证就是对受侵计算机系统进行扫描和破解，对入侵事件进行重建的过程。

因此，数据恢复不仅能为个人恢复已丢失的数据，同时也应用于公安、检察院、法院、司法等领域。

1.1.5 数据恢复从业人员

曾经有业内人士对从事数据恢复的人员按其所掌握的理论知识进行过分类，认为数据恢复从业人员可以分为 3 类，即数据恢复软件使用人员、理论知识与数据恢复软件使用相结合人员和数据恢复的“自由王国”人员。

1. 数据恢复软件使用人员

这类人员基本没有存储方面的理论基础，只会操作现有的数据恢复软件进行数据恢复，数据恢复的效果只能由所操作的数据恢复软件的功能来决定。

2. 理论知识与数据恢复软件使用相结合人员

这类人员具有深厚的存储知识理论功底，对文件系统环境及文件结构有相当的了解，熟悉

各种数据恢复软件参数设置的理论含义,可以针对不同的数据丢失情况,进行详细分析,并制定切实有效的数据恢复方案。在常用的数据恢复软件无法很好地完成恢复工作时,能够手工修改部分参数,为数据恢复软件创造一个良好的环境,从而最大限度地挽救丢失的数据。

3. 数据恢复的“自由王国”人员

具备第2类人员的基础,同时具有良好编程能力,在现有的数据恢复软件无法胜任恢复要求的情况下,随时可以自行编写实用的程序,以弥补现有数据恢复软件的不足,最大程度、最快地恢复数据。

成为数据恢复软件使用人员是很容易的,只要有一定的计算机操作经验即可,但这也是最危险的,因为数据恢复的成功率不仅取决于数据丢失后的情况,同时也取决于用户对数据丢失现场的保护程度,见参考文献[3,前言]。

1.2 数据恢复相关知识

1.2.1 数据的表示方式

计算机内的数据一般分为无符号数和带符号数两种,对于无符号数而言,整个数据均为数值部分,也就是说,该数据是正数或者是零;对于带符号的数据,在计算机中有3种表示方法:即原码、补码和反码;它们都是由符号位和数值两部分组成,数据最高位为符号位,符号位使用“0”表示正数,使用“1”表示负数;剩余位为数值部分。

1. 原码表示法

正数的符号位用“0”表示,而负数的符号位用“1”表示,数值部分按二进制的形式表示。

例 1.1 已知 $X=+42$, $Y=-42$,求: X 、 Y 的八位和十六位二进制数的原码并转换为对应的十六进制数。

解:

(1) X 和 Y 八位二进制数以及对应十六进制的原码

因为 $X=(+42)_{10}=(+010\ 1010)_2=(0010\ 1010)_2$

所以 $[X]_{\text{原}}=(0010\ 1010)_2=(2A)_{16}$

因为 $Y=(-42)_{10}=(-010\ 1010)_2=(1010\ 1010)_2$

所以 $[Y]_{\text{原}}=(1010\ 1010)_2=(AA)_{16}$

(2) X 和 Y 十六位二进制数以及对应十六进制的原码

因为 $X=(+42)_{10}=(+000\ 0000\ 0010\ 1010)_2=(0000\ 0000\ 0010\ 1010)_2$

所以 $[X]_{\text{原}}=(0000\ 0000\ 0010\ 1010)_2=(002A)_{16}$

因为 $Y=(-42)_{10}=(-000\ 0000\ 0010\ 1010)_2=(1000\ 0000\ 0010\ 1010)_2$

所以 $[Y]_{\text{原}}=(1000\ 0000\ 0010\ 1010)_2=(802A)_{16}$

原码表示很直观,与真值转换也很方便;但是原码进行加、减运算时,符号位不能视同数值一起参与运算,这时需要通过判断两数的符号来决定两数绝对值是做加法运算还是做减法运算,而且还要判断两数绝对值的大小,取绝对值大的数的符号作为结果的符号,这样运算规

则不仅复杂,而且运算时间长。

2. 反码表示法

正整数的反码表示与其原码表示相同;负整数的反码表示是将该数的原码除符号位以外其余各位取反。

例 1.2 已知 $X=+42, Y=-42$, 求: X, Y 的八位和十六位二进制数的反码并转换为对应的十六进制数。

解:

(1) X 和 Y 八位二进制数以及对应十六进制的反码

$$\text{因为 } X=(+42)_{10}=(+010\ 1010)_2=(0010\ 1010)_2$$

$$\text{所以 } [X]_{\text{反}}=(0010\ 1010)_2=(2A)_{16}$$

$$\text{因为 } Y=(-42)_{10}=(\text{---}010\ 1010)_2=(1010\ 1010)_2$$

$$\text{所以 } [Y]_{\text{反}}=(1101\ 0101)_2=(D5)_{16}$$

(2) X 和 Y 十六位二进制数以及对应十六进制的反码

$$\text{因为 } X=(+42)_{10}=(+000\ 0000\ 0010\ 1010)_2=(0000\ 0000\ 0010\ 1010)_2$$

$$\text{所以 } [X]_{\text{反}}=(0000\ 0000\ 0010\ 1010)_2=(002A)_{16}$$

$$\text{因为 } Y=(-42)_{10}=(\text{---}000\ 0000\ 0010\ 1010)_2=(1000\ 0000\ 0010\ 1010)_2$$

$$\text{所以 } [Y]_{\text{反}}=(1111\ 1111\ 1101\ 0101)_2=(FFD5)_{16}$$

3. 补码表示法

正整数的补码表示与其原码表示相同;负整数的补码表示为先求该数的反码,再在最低位加 1,即负整数的补码等于其反码加 1。

补码是计算机中用得最多的一种带符号数表示,因为计算机中最多的运算是加、减运算,补码的表示使符号位可以和有效数值部分一起直接参与加、减运算,无须像原码那样对符号位进行判断,从而简化了运算规则,提高了机器运算速度。因此,在计算机中对于带符号的数值一般是以补码表示的。

例 1.3 已知 $X=+42, Y=-42$, 求: X, Y 的八位和十六位二进制数的补码并转换为对应的十六进制数。

解:

(1) X 和 Y 八位二进制数以及对应十六进制的补码

$$\text{因为 } X=(+42)_{10}=(+010\ 1010)_2=(0010\ 1010)_2$$

$$\text{所以 } [X]_{\text{补}}=(0010\ 1010)_2=(2A)_{16}$$

求 Y 的补码,先求 Y 的反码:

$$\text{因为 } Y=(-42)_{10}=(\text{---}010\ 1010)_2=(1010\ 1010)_2$$

$$\text{所以 } [Y]_{\text{反}}=(1101\ 0101)_2=(D5)_{16}$$

由于负整数的补码等于反码加 1,

$$\text{因此, } [Y]_{\text{补}}=[Y]_{\text{反}}+(1)_2=(1101\ 0101+1)_2=(1101\ 0110)_2=(D6)_{16}$$

(2) X 和 Y 十六位二进制数以及对应十六进制的补码

因为 $X = (+42)_{10} = (+000\ 0000\ 0010\ 1010)_2 = (0000\ 0000\ 0010\ 1010)_2$

所以 $[X]_{\text{补}} = (0000\ 0000\ 0010\ 1010)_2 = (002A)_{16}$

求 Y 的补码, 先求 Y 的反码:

因为 $Y = (-42)_{10} = (-000\ 0000\ 0010\ 1010)_2 = (1000\ 0000\ 0010\ 1010)_2$

所以 $[Y]_{\text{反}} = (1111\ 1111\ 1101\ 0101)_2 = (\text{FFD5})_{16}$

由于负整数的补码等于反码加 1,

因此, $[Y]_{\text{补}} = [Y]_{\text{反}} + (1)_2 = (1111\ 1111\ 1101\ 0110)_2 = (\text{FFD6})_{16}$

1.2.2 数据的存储形式

数据的存储形式, 也就是数据在存储器中的存放顺序。在表示数值的大小时, 由于 1 字节最大只能表示到 255 (注: 无符号数), 如果要表示大于 255 的数据, 则需要 N 字节, 其中: N 为大于或者等于 2 的正整数, 这就存在 N 字节在存储器中存放顺序的问题; 在存储器中对 N 字节组成的数据有大头位序和小头位序两种存储形式。

1. 大头位序 (Big-Endian)

采用大头位序存储的数据, 在存储器中的存放顺序是: 从左到右为最高字节向最低字节依次存放, 即高字节存放在前(左)、低字节存放在后(右)。

假设某数据由 N 字节组成, 其中: N 为大于或者等于 2 的正整数; N 字节分别为“ $X_1, X_2, X_3, \dots, X_N$ ”, 如果采用大头位序存储, 在存储器中的存放顺序为“ $X_1\ X_2\ X_3\ \dots\ X_N$ ”; 则该数据的值为 $X_1X_2X_3\dots X_N$ 。

例 1.4 十进制数 143360, 转换成十六进制数为 23000; 在存储器中至少需要 3 字节来存储该数据。十进制数 143360 采用大头位序在存储器中分别占用 3 至 8 字节的存储形式见表 1.1 所列。

表 1.1 采用大头位序在存储器中的存储形式

分配给该数据的字节数	值									十六进制	十进制			
	存储形式													
3	02			30			00			23000	143360			
4	00		02		30		00			23000	143360			
5	00		00		02		30		00	23000	143360			
6	00		00		00		02		30	00	23000	143360		
7	00		00		00		00		02	30	00	23000	143360	
8	00		00		00		00		00	02	30	00	23000	143360

如果该数据占用 9 字节, 则在该数据前(左)添加 1 字节值“00”, 以此类推。

2. 小头位序 (Little-Endian)

采用小头位序存储的数据, 其数据在存储器中的存放顺序是: 从左到右为最低字节向最高字节依次存放, 即低字节存放在前(左)、高字节存放在后(右)。

假设某数据由 N 字节组成,其中: N 为大于或者等于 2 的正整数; N 字节分别为“ $X_1, X_2, X_3, \dots, X_N$ ”,如果采用小头位序存储,在存储器中的存放顺序为“ $X_1 X_2 X_3 \dots X_N$ ”;则该数据的值为 $X_N \dots X_3 X_2 X_1$ 。

例 1.5 十进制数 143360,转换成十六进制数为 23000;在存储器中至少需要 3 字节来存储该数据。十进制数 143360 采用小头位序在存储器中分别占用 3 至 8 字节的存储形式见表 1.2 所列。

表 1.2 采用小头位序在存储器中的存储形式

分配给该数据的字节数	值							十六进制	十进制
	存储形式								
3	00		30		02			23000	143360
4	00		30		02		00	23000	143360
5	00		30		02	00	00	23000	143360
6	00		30	02	00	00	00	23000	143360
7	00	30	02	00	00	00	00	23000	143360
8	00	30	02	00	00	00	00	23000	143360

如果该数据占用 9 字节,则在该数据后(右)添加 1 字节值“00”,以此类推。

不同的分区形式、文件系统,数据的存放形式不同。在 MBR 分区、GPT 分区、FAT32 和 NTFS 文件系统中,数据的存储形式采用小头位序;而在动态磁盘的 LDM 数据库中数据存储形式则是采用大头位序。

1.2.3 计算机的启动过程

计算机的启动过程主要由以下几个步骤组成。

(1) 开机, BIOS 加电自检, 如果自检正常, 则转到第 2 步; 自检不正常, 则出现错误提示或者响声并死机。

(2) 根据 CMOS 的设置开始启动, 将硬盘(假设 CMOS 的设置是硬盘为第一启动顺序)的 0 号扇区(即硬盘 0 磁头 0 柱面 1 扇区, 也就是主引导扇区)读入内存地址 0000:7C00 处, 并且从 0000:7C00 处开始执行。

(3) 检查 0000:7DFE 是否等于 0XAA55。若不等于则转去尝试其他介质; 如果没有其他启动介质, 则显示“NO ROM BASIC”, 然后死机。

(4) 主引导记录先将自己复制到 0000:0600 处, 然后继续执行。

(5) 在主分区表中搜索标志为活动的分区。如果发现没有活动分区或者不止一个活动分区, 则停止。

(6) 将活动分区的第一个扇区读入内存地址 0000:7C00 处。

(7) 检查 0000:7DFE 是否等于 0XAA55, 若不等于则显示“Missing Operating System”, 然后停止。

(8) 跳转到 0000:7C00 处继续执行特定系统的启动程序。

以上步骤是标准的硬盘主引导扇区, 多系统引导程序的引导过程与此不同; 多系统引导