

Storage Technology Foundations

存储技术基础

刘凯 刘博 编著



西安电子科技大学出版社
XIDIAN UNIVERSITY PRESS

存储技术基础

刘 凯 刘 博 编著

西安电子科技大学出版社

内 容 简 介

本书主要介绍计算机系统中与存储相关的技术,包括存储器的组成结构、工作原理、存储管理系统以及网络存储等内容。

全书共七章,第一章主要介绍当代存储技术发展的基本情况、存储需求以及三种常用的存储环境,即 DAS、NAS 和 SAN;第二章主要介绍半导体存储器的基本原理,并针对常用的半导体存储器,如 SRAM 等给出了应用实例;第三章重点介绍有关磁盘存储的内容,包括磁介质存储原理、磁盘工作原理和数据组织结构等;第四章介绍主机环境、连接关系、磁盘阵列和智能磁盘存储系统等内容;第五章介绍各种类型的存储器如何组成计算机中不同层次的存储系统;第六章介绍网络连接存储和存储区域网络两部分内容;第七章重点介绍在多处理器计算机系统中如何设计存储的内容,包括并行处理系统的类型和存储设计实例、数据一致性协议和机群系统存储设计等。

本书可以作为相关专业本科生或研究生的教材,也可供对存储技术感兴趣的专业人士参考。

图书在版编目(CIP)数据

存储技术基础 / 刘凯, 刘博编著. —西安: 西安电子科技大学出版社, 2011.1

ISBN 978-7-5606-2498-3

I. ① 存… II. ① 刘… ② 刘… III. ① 数据存储 IV. ① TP333

中国版本图书馆 CIP 数据核字(2010)第 212638 号

策 划 臧延新

责任编辑 任倍萱 臧延新

出版发行 西安电子科技大学出版社(西安市太白南路 2 号)

电 话 (029)88242885 88201467 邮 编 710071

网 址 www.xduph.com 电子邮箱 xdupfb001@163.com

经 销 新华书店

印刷单位 陕西光大印务有限责任公司

版 次 2011 年 2 月第 1 版 2011 年 2 月第 1 次印刷

开 本 787 毫米×1092 毫米 1/16 印 张 14.75

字 数 345 千字

印 数 1~2000 册

定 价 27.00 元

ISBN 978-7-5606-2498-3/TP · 1244

XDUP 2790001-1

如有印装问题可调换

本社图书封面为激光防伪覆膜,谨防盗版。

前 言

当今消费类电子产品，如数码相机、数码摄像机、移动电话、用于快速上网的计算机等，已经得到了广泛的应用，这些数字电子产品都具有产生大量数据的特征。实际上，人们现在不仅处于信息时代，而且也处于信息爆炸时代，这就需要有能够管理、平衡并优化存储和信息的专业人员。本书就是针对这样的需求而编写的。

以往，存储技术作为计算机专业某门课程的一个方面出现，如“计算机组成与体系结构”、“微型计算机原理及接口技术”等，或者作为微电子专业中集成电路课程的一部分出现，很难找到对存储技术从计算机角度和微电子角度进行研究而单独出现的情况。目前的技术要求不能够仅从某一方面考虑信息及其管理方面的内容，还需要以较为独立的观点对存储技术进行深入描述。为此，我们设计并编写了这本关于存储技术方面的书，主要目的是希望以存储技术为主题，将涉及到存储的内容统一起来，使读者能够以较新的视角考虑存储问题，为应对日益增长的信息管理和存储需求提供有效的解决方案。

本书主要围绕从集成电路到大型网络存储设备这一主题，对相关存储技术进行描述。全书共七章，下面给出各个章节的内容安排。

第一章，存储技术概述。本章主要说明当前数据存储的需求，以及数据存储带来的商业价值。本章还给出了数据存储演变的过程和常见的存储解决方案，并重点描述了 DAS、NAS 和 SAN 这三种常用基本存储环境的基本结构和特点。在本章最后，给出了常见的如磁带和磁盘之类存储设备的概况，包括其基本形式和技术指标等。

第二章，半导体存储器。本章从集成电路的角度说明了构成基本半导体存储单元的工作原理，包括随机读/写存储器、只读存储器和闪速存储器等。在该章中，我们还从工程角度对静态随机读/写存储器和动态随机读/写存储器以实例进行说明，并且给出了存储器验证、汉明码纠错保护等以硬件描述语言编写的实例。读者能够从原理和应用两个角度学习和掌握有关半导体存储器的内容。最后，结合 CF 存储卡，本章给出了有关典型存储卡等内容，并在附录中给出了相关控制代码实例。

第三章，磁介质存储器与光存储器。本章重点介绍有关磁盘存储的内容，包括磁介质存储原理、磁盘工作原理和数据组织结构等，最后对光盘存储也作了一定的说明。

第四章，存储系统结构。本章给出了关于主机环境、连接关系、磁盘阵列和智能磁盘存储系统等内容。重点描述了以 RAID 形式给出的磁盘阵列内容，包括 RAID 分级、RAID 软/硬件实现和性能指标等内容。

第五章，计算机系统存储管理。本章描述由各种类型的存储器组成计算机中不同层次的存储系统的方法。在引入局部性原理的基础上，通过具体实例介绍了高速缓冲存储器的三种映像方式、替换算法和写策略。虚拟存储器主要介绍了分页模型、地址翻译机制，并通过一个小系统进行实例讲解。本章最后介绍了 Pentium 处理器的存储管理机制。

第六章，网络存储。本章重点介绍了网络连接存储和存储区域网络两部分内容。网络连接存储主要描述网络连接存储的软/硬件组成及 NFS、CIFS 两种网络文件系统。存储区域网络主要描述光纤通道各个层次的技术要点、存储区域网络的三种拓扑结构、端口类型和端口地址，并在此基础上通过具体实例描述了如何进行存储区域网络设计，实例中还涉及到了业务连续性、备份与恢复、远程容灾等相关概念。

第七章，并行处理系统中的存储设计。本章重点介绍了在多处理器计算机系统中如何设计存储的内容，包括并行处理系统的类型和存储设计实例、数据一致性协议和机群系统存储设计等。最后着重说明了衡量存储系统性能方面的两个基本定律，即 Little 和 Amdahl 定律，并给出了分析研究存储系统的基本方法，以及存储发展的情况。

本书在编写过程中参考了较多的书籍和网络资料，为了尊重他人成果，我们在书中给出了具体引用的说明，并在参考资料部分列出了主要参考书目和资料名称。本书的编写得到了以下单位的支持：西安电子科技大学计算机学院、EMC 公司教育资源部门、西安电子科技大学出版社，特别是本书受到了“西安电子科技大学教材建设基金资助项目”和国家自然科学基金项目(编号：60802076)的资助。

在此，感谢马建峰教授、王泉教授、裘雪红教授等对本书的关心与帮助。同时，作者对出版社的同志也表示衷心的感谢。

本书第一、二、三、四、七章由刘凯编写，其余章节由刘博编写，部分图由许超提供。全书由刘凯统稿。由于作者自身水平有限，书中难免存在不妥之处，恳请有关专业人士及读者给予批评指正，在此表示感谢。

E-mail: kailiu@mail.xidian.edu.cn, liuboletter@hotmail.com。

编 者
2010 年 12 月

目 录

第一章 存储技术概述.....	1
1.1 当代数据存储需求.....	1
1.2 存储数据的价值.....	2
1.3 数据的类型及存储演变.....	3
1.4 数据增长率和管理数据的挑战.....	3
1.5 数据存储解决方案.....	4
1.6 存储环境.....	6
1.7 存储设备的基本情况.....	8
第二章 半导体存储器.....	11
2.1 半导体存储技术基本原理.....	11
2.1.1 半导体存储器概述.....	11
2.1.2 半导体存储器基本结构.....	13
2.1.3 静态随机读/写存储.....	15
2.1.4 动态随机读/写存储.....	17
2.1.5 只读存储器.....	19
2.2 静态随机读/写存储器.....	24
2.2.1 典型芯片.....	24
2.2.2 应用实例分析.....	29
2.3 动态随机读/写存储器.....	41
2.3.1 典型芯片.....	41
2.3.2 SDRAM 工作模式设定.....	44
2.3.3 SDRAM 的操作命令.....	48
2.3.4 SDRAM 的操作时序.....	50
2.3.5 动态随机读/写存储器控制器设计.....	53
2.4 只读存储器.....	53
2.4.1 CF 存储卡的基本情况.....	53
2.4.2 CF 存储卡标准结构.....	54
2.4.3 CF 存储卡读/写时序.....	59
2.4.4 CF 存储卡的配置.....	65
2.4.5 I/O 传输功能.....	68
2.4.6 通用内存传输功能.....	68
2.4.7 True IDE 模式 I/O 传输功能.....	69

2.4.8 典型芯片及应用实例分析	69
2.5 其他类型的半导体存储器	71
小结	76
第三章 磁介质存储器与光存储器	77
3.1 磁介质存储技术概述	77
3.1.1 磁介质存储器的基本原理	77
3.1.2 磁盘的基本结构、读/写过程及信息记录方式	79
3.1.3 磁盘的数据组织和格式化	81
3.1.4 磁盘的物理特性	82
3.2 硬盘驱动器基本结构	85
3.2.1 硬盘简介	85
3.2.2 硬盘驱动器组织结构	86
3.2.3 硬盘驱动器主要技术指标	88
3.2.4 硬盘接口整体模型	90
3.3 硬盘驱动器 IDE 接口	91
3.3.1 IDE 接口概述	91
3.3.2 IDE 适配器	93
3.3.3 IDE 物理接口	93
3.3.4 IDE 协议	98
3.3.5 IDE 磁盘驱动器模型	105
3.4 磁带存储系统	115
3.5 光存储器	117
3.5.1 概述	117
3.5.2 常见光存储器	118
3.6 ATAPI 接口	121
3.6.1 概述	121
3.6.2 ATAPI 任务文件	122
3.6.3 ATAPI 传输协议	124
小结	126
第四章 存储系统结构	128
4.1 主机环境	128
4.2 连接关系	133
4.3 磁盘阵列	137
4.4 磁盘存储系统	147
小结	153
第五章 计算机系统存储管理	154
5.1 局部性原理	154

5.2 存储器层次结构.....	155
5.3 高速缓冲存储器.....	155
5.3.1 直接映像.....	156
5.3.2 全相联映像.....	158
5.3.3 组相联映像.....	160
5.3.4 替换算法.....	162
5.3.5 写策略.....	163
5.4 虚拟存储器.....	164
5.4.1 分页模型.....	164
5.4.2 地址翻译.....	166
5.4.3 小系统举例.....	168
5.5 Pentium 处理器.....	170
小结.....	172
第六章 网络存储.....	173
6.1 直连式存储.....	173
6.2 网络连接存储.....	174
6.2.1 NAS 的硬件构成.....	174
6.2.2 NAS 的软件构成.....	175
6.2.3 NAS 产品介绍.....	179
6.2.4 其他新技术.....	179
6.3 光纤通道.....	180
6.3.1 传输介质.....	181
6.3.2 连接器.....	182
6.3.3 8B/10B 编码.....	183
6.3.4 有序集.....	185
6.3.5 光纤通道帧格式.....	186
6.3.6 序列和交换.....	187
6.3.7 服务类型.....	188
6.3.8 流量控制.....	188
6.4 存储区域网络.....	189
6.4.1 拓扑结构.....	189
6.4.2 端口类型.....	191
6.4.3 端口地址.....	192
6.4.4 存储区域网络设计.....	193
小结.....	197
第七章 并行处理系统中的存储设计.....	198
7.1 并行处理系统及其存储设计.....	198

7.1.1 并行处理系统分类.....	198
7.1.2 对称多处理机及存储设计	199
7.2 数据一致性协议.....	204
7.2.1 Cache 一致性协议的软件方法.....	205
7.2.2 Cache 一致性协议的硬件方法.....	205
7.2.3 MESI 协议	206
7.3 机群及其存储系统.....	208
7.3.1 机群系统配置.....	209
7.3.2 机群系统和 SMP 系统的区别	211
7.4 非均匀存储器存取.....	211
7.5 存储系统性能分析.....	213
7.5.1 Amdahl 定律	213
7.5.2 Little 定律	215
小结.....	219
附录一 CF 卡与 80C51 连接原理图	220
附录二 80C51 汇编代码.....	223
参考文献.....	227

第一章 存储技术概述

随着信息技术尤其是计算机技术的不断发展和应用，每时每刻都有新的数据产生，比如移动电话通信数据、气象数据、银行交易数据等等。在不断产生的数据中，有些是瞬时的，即不需要进行存储的，而有些则需要保存一段时期，更有一些数据甚至需要长期保存。那么，如何有效地保存日益增长的数据就成为存储技术面临的主要问题。本章将从存储需求的角度切入，对存储技术的基础进行简要的论述，为后面章节展开论述奠定基础。

1.1 当代数据存储需求

数据存储的需求在不断增长，那么这种增长由什么因素推动呢？也就是说，产生数据的主体是什么？对所产生数据总量的估计如何进行？伴随着数据增长以及对存储管理数据的相应需求，工业界迫切需要理解数据存储和价值的专业人员。这些人可以称为存储技术专家。这样的称呼有利于考察数据的产生和组成，以及数据产生的类型和数据变为信息的时刻。

平均而言，数据产生的年增长率已超过了 50%，对于需要长期保存且易于读取数据的需求也在不断增长。同时，信息技术的预算也在不断增长。这其中主要包括服务器、网络、存储以及个人终端的花费。据估计，在整个信息技术的投资预算中存储的相关部分约占四成。

数据主要由个人和各种实体(如公司企业、政府机关、学校、社会团体等)产生。例如，个人产生的数据包括私人照片、私人文档、电子数据表以及音/视频资料等。保存这些个人数据的现有存储设备主要有：数码相机、MP3 播放器、硬盘、CD ROM/DVD 或者 U 盘等。如何管理这些数据设备并使其能够有效利用，这将是一个巨大的挑战。那么，对于像公司企业这样的实体又会产生什么样的数据呢？这里简单地将其分为四类，说明如下：第一类数据包括产品清单、说明书、报价、可用情况、销售数量及预测；第二类数据与客户相关的，包括订单、运输细节说明等；第三类数据与银行相关，包括各种账户信息、存取款记录等；第四类数据与医疗相关，比如医院的健康记录、保险信息等。当然，还有许多类型的数据没有给出，但是根据以上描述可见，相比较个人数据，企业等实体产生的数据量要庞大得多。企业数据一般存储在诸如工作站、服务器、磁盘阵列、磁带、CD ROM/DVD 等大容量的存储器上，各个存储器甚至可以组成网络形式来提供存储服务。而这些数据在重要性、可用性、安全性方面的要求也比个人数据高许多。这也就造成了对这些数据的管理和应用的困难，因此必须采用一些新的数据存储技术为人们提供更有效的

存储服务。譬如，我们所要涉及的存储区域网络以及磁盘阵列技术都是解决大规模存储的有效技术。

1.2 存储数据的价值

日益增长的数据存储需求，使得存储技术得以长足发展。那么，我们为什么要保存诸多的数据呢？对于企业而言，海量数据中蕴含着许多重要的信息，企业保存数据就是为了从中挖掘出信息。这些信息可以转换为企业的效益或者提升企业的管理水平。例如，零售商可以从客户的数据中，找到客户购买习惯和购买模式的重要信息；快递公司的 GPS 系统数据可以提供目前运输车的位置信息；医院可以从病人的数据中得到患者的相关信息，有助于治疗过程；银行的信用卡数据可以提供用户使用的交易信息等。数据中信息的价值可以体现在发现新的商业机遇、找到改变现有商业的模式、提高竞争力等三个主要方面。如果不具有一定的保存数据和抽取信息的能力，那么企业的生存和发展将受到很大限制。对于个人数据而言，如果没有有效保存数据的能力，也会给我们的生活带来很大的困难，甚至造成混乱。

具有创新能力的企业总是以创新的方式利用数据。比如，以个人简历为例，为了能找到更广范围的雇主，我们就必须使得简历能够正确投递到雇主手中。职业介绍公司就可以非常有效地收集我们的数据并把这些简历放到一个公众可以获取的平台上以供雇主查阅。这种业务就暗含着存储需求。如果每份简历的大小为 4 KB，要保存 100 万份简历的话，所需的存储容量就要 4 GB。如果再加上雇佣公司提供的工作岗位信息的话，存储容量还会增加。这时就需要通过相关软件采用关键字查询的方式找到合适的公司和人选。这样，工作搜索引擎就实现了把个人数据转换为有用的信息的目的。

为了更进一步说明信息的价值，我们以计算机系统死机对各个行业带来的实际经济损失为例给出直观的数据。图 1.1 为具体以百万美元计的每小时平均经济损失数据。

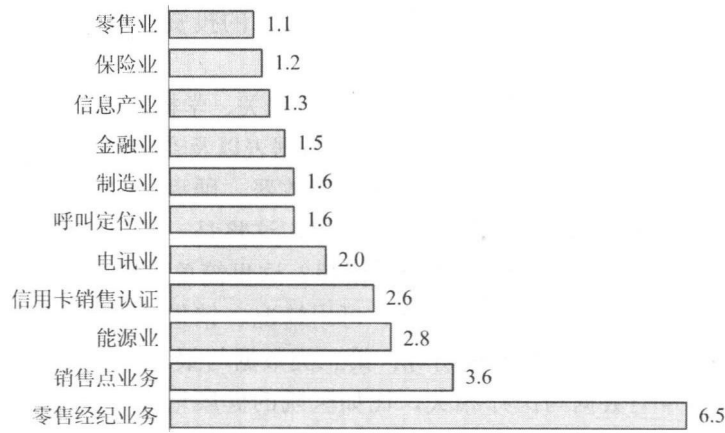


图 1.1 系统死机带给各行业每小时的平均经济损失数据(单位：百万美元)

由图 1.1 可见，信息系统死机将影响到诸多行业，而且带来的经济损失也非常巨大，同时印证了信息本身带来的价值。

1.3 数据的类型及存储演变

存储数据可以分为结构化和非结构化两种基本类型。结构化数据是指形式上整齐且有良好的组织的数据，这些数据一般存储于数据库或者电子表单中。例如，我们经常使用的 Excel 表格形式的数据文件、各种名单等。而非结构化数据则没有一定的形式，例如 XML 标记的内容，通常以文本方式给出。非结构化数据是没有任何组织的，通常作为整个文档存储为不同形式的文件，或者保存于内容管理的文件系统中。根据研究机构的调查，目前大约 80% 以上的企业信息都是非结构化数据。常见的非结构化数据有论文、音/视频资料、账单、网页、支票、电子邮件等。

数据的存储实际上经历了一个演变过程。在计算机系统发展的早期，对于数据的计算和存储只能选择集中化模型。在这种模型情况下，处理和获取数据都集中在适当的位置。该模型使得新的应用开发与利用得不到较快发展。系统对数据的存取可以依据商业需求进行预测，在这种系统下，通常认为计算能力比实时存取数据能力更加重要。在集中化系统中，用户通过终端连接到主机上并通过主机访问内部或外部的存储设备。随着网络技术的进步，客户端/服务器模型开始流行，逐步形成了分散式的存储模型。在这种分散式存储模型中，企业内的业务单元能够访问自己的服务器和存储器。这样，应用程序就不再排队等待数据的获取和程序的执行了。但是，分布式存储导致了信息的分段化。这样也会对信息处理和分布管理的均衡性带来一定困难。面对这些分布式存储的需求，网络存储的概念也随之出现。简单地讲，在网络存储系统中生产部门拥有自己的客户端、服务器和存储器集合，财务部门也拥有自己的客户端、服务器和存储器，人力资源部门同样具有自己的客户端、服务器和存储器。网络存储实际上也可以看做集中化存储方式，而且这种方式是目前最好的信息存储形式。网络存储解决方案的主要优点体现在每一种情况下，数据都是集中定位并保存在磁盘存储系统上的。网络存储实现了连接多个计算机到一个中央位置以利用数据的存储和获取，数据能实现更容易的管理、共享和保护，而且数据具有非常高的可用性。

1.4 数据增长率和管理数据的挑战

全球范围内信息的快速增长比率是显而易见的，图 1.2 给出了截至 2005 年磁盘阵列数据的年增长比率。2005 年前信息年平均增长约 60%，而 2005 年后增长率超过了七成。

如此快速的增长给信息管理带来了严重的困难。数据存储整合到集中的阵列上仅仅是整体信息管理的一部分。信息管理主要面临以下几方面的挑战：

(1) 存储容量的规划。信息的增长是持续不断的，为了应对不断增长的信息而采用的解决方法应能适应不断增长的需求。

(2) 数据的分类。信息价值会随着时间而不断改变。数据和信息应根据对业务贡献的作用而进行分类。我们将会看到存储阵列是由不同类型和成本的设备构成的。对数据的分类将使得每一类数据都能具有正确的存储设备。

(3) 数据可用性支持。信息比以往更具战略性，为此确保信息的可用性将是管理工作的

关键。

(4) 信息的安全性。

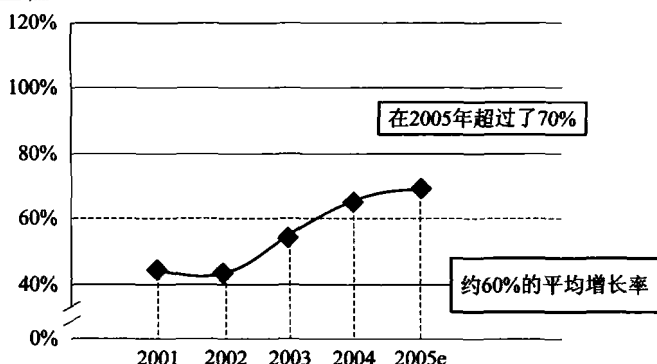


图 1.2 信息年增长率

1.5 数据存储解决方案

大量的数据存储于硬盘、磁带以及光盘上。每一种存储介质都提供进行特定数据存取需要的解决方案。包含磁盘阵列和网络技术的网络存储已经成为大多数商业数据存储需求的主要解决方案。接下来我们将逐一说明这些解决方案，特别是磁盘系统和互联技术。

在硬盘系统演变的早期，磁带是主要的数据存储介质。有些用户或许曾见过具有数个磁带驱动器的计算中心的照片。这些磁带驱动器排列在支架上，操作员则忙于将成卷的磁带安装或者卸载到这些驱动器上。通常我们把磁带驱动器和磁带的集合称为带库，而把光盘和光驱的集合称为自动唱片点唱机(Jukeboxes)，硬盘的集合称为磁盘阵列。目前，在主要的数据存储方案中，磁带往往作为二级存储使用。光盘集合则用于存储那些长期内容不变的数据。磁盘阵列存储可以立即获取在线数据。

相比较磁盘阵列而言，磁带非常便宜。在记录信息方面，磁带使用读/写磁头将单位信息记录在磁带表面的磁介质上，并且记录工艺技术也在不断改进，提供了更高的存储容量、读/写可靠性和性能。比如磁带从早期的大块卷系统已经变为更小的压缩盒形式并且能够自动加载到存储支架上。随着时间的推移，磁带的记录格式也使得能够在单个卷上容纳更多数据并且提供了更高的数据率，而其功能始终没有发生变化。值得一提的是，磁带上的数据是以串行的形式一个字节接一个字节记录的。因为存储数据与磁带的长度呈线性关系，所以对特定数据位的随机读/写往往比较慢。为了读取磁带上的某个文件，磁带驱动器需要不断地定位磁头到特定文件下。然后，将文件信息读入计算机内存系统中的这个过程非常费时。如果是对磁带信息从头至尾进行读取并写入其他存储位置上，例如存储在硬盘上时，那么磁带系统工作的效率则比较高。正是由于此项弱点限制了磁带用于实时快速的数据存储场合，而通常作为备份或数据压缩时使用的低成本存储方案。磁带系统不能被多用户或程序共享使用。有趣的是，在 CD-ROM 出现之前，许多操作系统和应用软件的生产商往往将他们的软件使用磁带盒方式发行。

随着光盘的出现，人们拥有了随机数据获取以及数据写保护的手段。大量光盘装配到类似自动点唱机式的设备中，这样就能够为集中获取的存储提供大容量阵列式解决方案。

而单张光盘的容量仍然与磁带和磁盘的容量有巨大差距。光盘数据存储具有以下应用特征:

- (1) 在小型单用户的计算环境中非常流行。
- (2) 个人经常用光盘作为存储共享数据甚至备份的解决方案。
- (3) 光盘有时也被应用程序用作分布式介质,或者作为将少量数据从一个自我包含的系统传送到其他系统的方式。

磁盘驱动器自出现起就是备受欢迎的数据存储方式。数据的急速增长导致了各种介质存储阵列的发展。阵列能使数据并发地由多个服务器获取。磁盘阵列甚至智能磁盘阵列快速成长,逐渐成为有效的大容量数据存储解决方案。按照磁盘发展的历史情况,我们主要考察的磁盘存储方案有:直接获取存储设备(DASD)、磁盘簇系统(JBOD)、磁盘阵列以及智能磁盘阵列。

图 1.3 给出了早期直接获取存储设备的结构。

DASD 系统由 IBM 于 1956 年设计采用,是从主机中获得磁盘数据最老的技术。磁盘直接由主机进行读/写操作,并且与主机环境紧密耦合。我们可以认为目前个人计算机中的硬盘驱动器就是一种 DASD 系统的实例。可以认为 DASD 是一种在服务器/计算机与其磁盘驱动器间的一对一映射关系。在 DASD 系统中,所有磁盘上的数据都必须经由服务器/计算机进行路由使用。在网络技术出现以前,这种路由涉及到向操作员提交任务申请的过程。这种系统中的服务器/计算机有能力直接获取数据而个人却无法做到直接读取数据。除了像磁带系统那样不断定位装载盘卷的不便外,DASD 磁盘包必须为特定运行的任务交换磁盘。虽然这种方式速度慢,但它提供了比磁带系统更快的切换方法。在 DASD 系统中如果某个磁盘出现故障,其上所有数据有可能丢失或者最近使用的数据有可能不可靠。

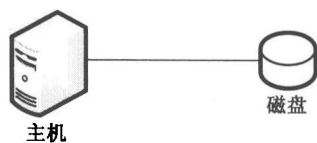


图 1.3 DASD 结构

磁盘簇系统(JBOD)是将多个物理磁盘安装到外部的机柜中,并通过线路连接到唯一的服务器上。JBOD 系统能够随着内部磁盘数目的增加而提供更大的存储容量,但磁盘上的数据并不受到 JBOD 的保护。JBOD 系统结构如图 1.4 所示。

由于 JBOD 存在的局限性,磁盘阵列系统则提供了更高的性能和可靠性。在磁盘阵列中,具有多个与主机相连的接口,使得多个主机能够共享访问相同的磁盘阵列。阵列管理软件可以实现阵列资源的分割,以使单个或者多个磁盘能够分配给系统中的主机。

一般在磁盘阵列中都具有硬件磁盘阵列控制器实现对独立磁盘冗余的计算和系统中读/写操作的优化处理。

智能磁盘阵列具有独立的操作环境。该操作环境可视为阵列的操作系统。在智能磁盘阵列中分布有大量的高速缓冲存储器(Cache),并采用复杂的算法对 Cache 进行管理优化来自主机的读/写请求任务。大容量的磁盘驱动器可以分割为多个小单元。这些小单元又可以

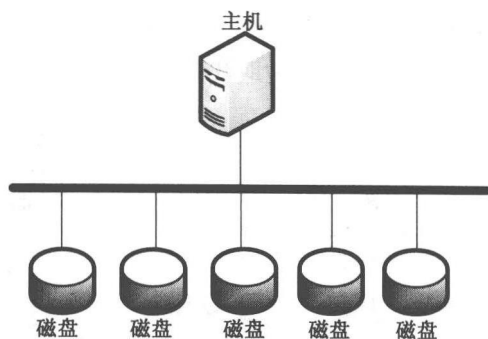


图 1.4 JBOD 系统结构

在主机方以单独的磁盘驱动器形式出现。阵列管理软件可以实现多个主机通过相同的 I/O 通道读/写阵列存储。操作系统能够确保每个主机仅能访问到分配的磁盘资源。磁盘阵列系统结构如图 1.5 所示。

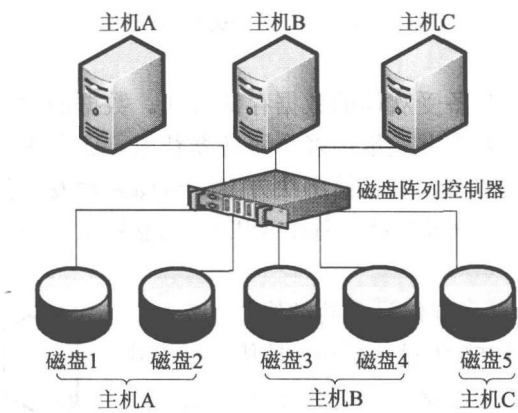


图 1.5 磁盘阵列系统结构

1.6 存储环境

目前对于存储而言,存在两种不同的存储环境,即直接互联存储(Direct Attached Storage, DAS)和网络存储(Network Storage, NS)。其中,网络存储的主要形式包括存储区域网络(Storage Area Network, SAN)和网络互联存储(Network Attached Storage, NAS)。

图 1.6 给出了典型的 DAS 系统结构。在 DAS 环境中,服务器直接通过 SCSI 接口与磁盘阵列连接,在服务器之间相同的磁盘阵列端口不能共享,客户端则通过局域网方式连接到服务器上,服务器和磁盘阵列的距离受到 SCSI 线路的限制。随着网络技术和光纤技术的发展,DAS 已经逐步被取代了。

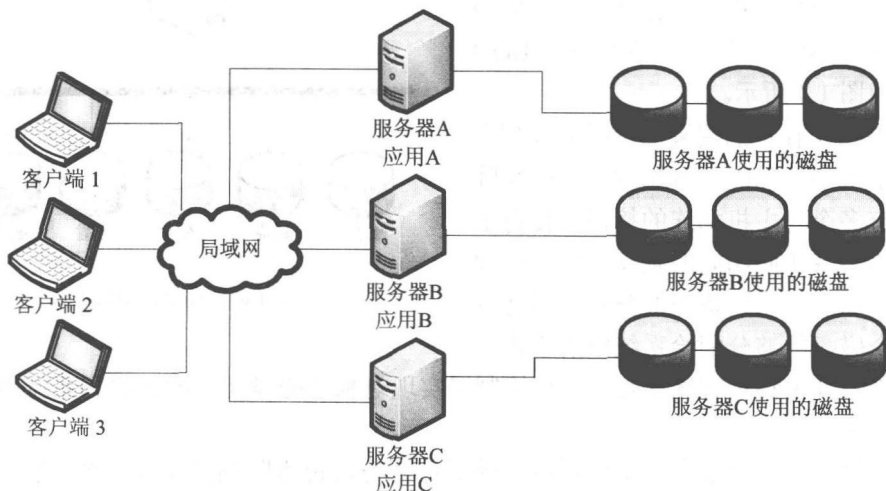


图 1.6 DAS 系统结构

图 1.7 给出了 SAN 系统结构。在 SAN 环境中，服务器采用专用的网络，即 SAN 网络实现对磁盘阵列的读/写。SAN 网络包括光纤通道交换设备以提供服务器和磁盘阵列间的互联。在 SAN 模型中，多个服务器可以访问磁盘阵列上相同的光纤通道端口，服务器和磁盘阵列间的距离比 DAS 环境中的 SCSI 接口方式远得多。同样，客户端也采用局域网方式访问服务器。

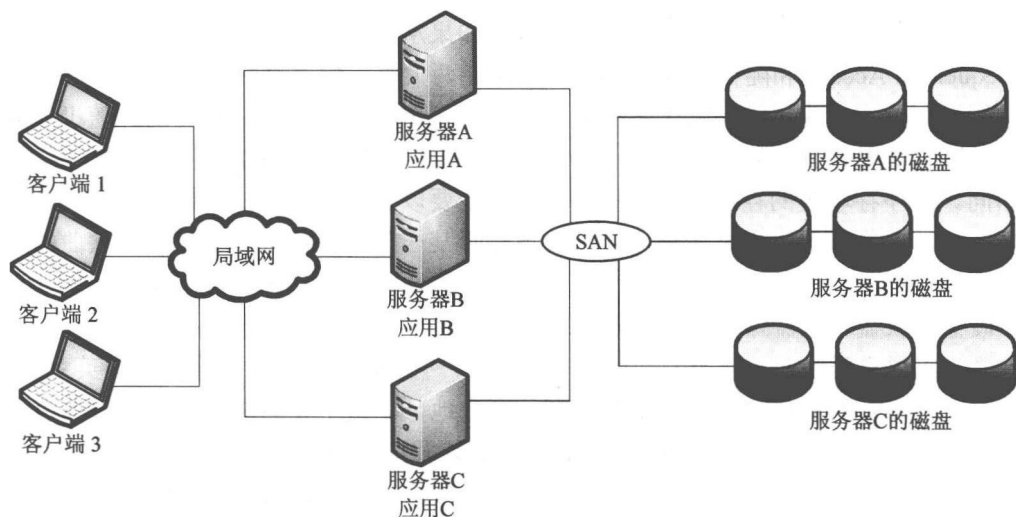


图 1.7 SAN 系统结构

图 1.8 给出了 NAS 系统结构。在 NAS 环境中，NAS 设备通过直接连接或者外部连接对磁盘进行读/写操作。NAS 头实现了文件服务的优化，称为 NAS 客户端的服务器采用局域网读取文件系统的方式运行应用程序，而客户端也是通过局域网连接到服务器上的。

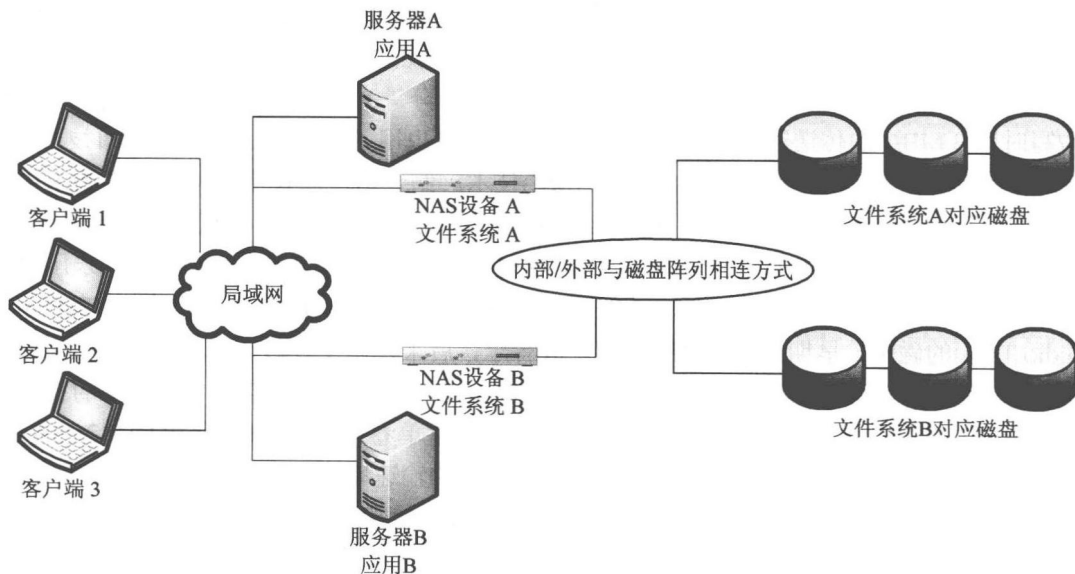


图 1.8 NAS 系统结构

1.7 存储设备的基本情况

存储设备在计算机的接口设备中具有非常重要的作用,在前面我们主要给出了从逻辑层面看到的存储基本问题,接下来则从物理层面说明存储设备的基本情况。

存储设备从不同角度可以分成不同的类型。如果从产品本身的特性来分,可分为顺序存取(Sequential Access)和随机存取(Random Access)两大类。顺序存取类读/写数据时按照一定的顺序进行,有时为了访问某个数据必须经过若干个数据后才能达到目的。随机存取类存储设备则对其中存放的数据同样进行读/写处理,即无论地址顺序如何,对任意位置读/写形式相同。顺序存取类的存储设备有哪些呢?它主要以磁带类为主,其外形及结构是带状的存储介质,就像一般家用录音带一样,如果要听某一首歌曲,必须要先卷到歌曲的所在位置才能开始播放。而随机存取类的存储设备的种类比较多,且多为盘状或碟状的外形,比如,硬盘、软盘、磁盘阵列、磁光驱、只读式光驱等,皆属于这类产品。

首先,以磁带类为例。磁带的种类也非常丰富,有 DLT、8 mm、DAT、1/2 英寸及 1/4 英寸。目前磁带在计算机存储设备中仍占有一席之地,其最主要的原因是它的每单位储存成本较为低廉,所以,这类产品较适合做成备份装置(Backup Devices)或层次存储管理(Hierarchical Storage Management, HSM)的底层存储设备。DLT(Digital Linear Tape)最早是由 Digital Equipment Corporation 发展出来的,使用在 VAX 计算机系统上,至 1993 年,全球销售出了仅 37 000 台。之后,由于计算机的迅速发展,数据量大增,对于存储设备的容量需求也明显增加,这才使 DLT 有了进一步发展,从 1994 年到 1995 年,销售额增长了 93.5%,年度销售数量由 43 400 增长到 84 000 台。最新的 DLT7000 的存储容量可达单卷未压缩 35 GB,传输速率为 5 MB/s。8 mm 指的是其磁带宽度为 8 mm 的计算机存储设备,这一类设备是利用回旋式(Helical Scan)读/写头进行读/写操作的。所谓回旋式读/写头,指磁头快速旋转着划过磁带表面,而磁带则以缓慢的速度前进着。在此以 Exabyte 公司为例,它们的最新一代 8 mm 磁带被称为 Mammoth。其读/写头以每分钟 5660 转的速度高速旋转,使得读/写时以 3 MB/s 的传送速率来传送数据。而它的磁带也用了最新式的磁性物涂布技术,即金属蒸发式(Advanced Metal Evaporated, AME)技术。结合这些新技术,使得单卷的未压缩容量可以高达 20 GB。另外,还有一种 8 mm 的新技术,称为 AIT(Advanced Intelligent Tape),它是结合 AME 技术,再加上类似 DDS 数据格式,并加以改良,使得 AIT 的单卷未压缩容量高达 25 GB,传输速率也达 3 MB/s。DDS(Digital Data Storage)也就是常被称为 DAT (Digital Audio Tape)的磁带,早期,这种磁带用于存放音乐,后来才用于存放计算机数据。而 DDS3 是指第三代 DDS 标准,最早于 1992 年由 DDS Manufacturers Group 制定了该标准的初稿,一直到 1995 年完整的版本才被制定发布。它的容量可达单卷未压缩 12 GB。和 DDS2 的 4 GB 未压缩容量比较起来,DDS3 的容量增大了两倍,其传输速率达到 1 MB/s 的水平。

随机存取类的存储设备主要是使用盘状或是碟状存储媒体来制作的设备。其读/写装置可以自由且快速地移动到文件存放位置上,并立即进行读/写数据的动作。目前,这类设备按照材质的特性可分为:可多次读/写设备(Rewritable)、单写多读式(Write-Once-Read-Many)及只读式(Read Only)设备。按照实际产品可制作成磁盘阵列(Disk Array)、磁光驱(Magneto