

中国计算机软件专业技术水平考试指定用书

中国计算机软件专业技术资格和水平考试中心组织编写

程 序 设 计

王春森 主编 (高级程序员级)



清华大学出版社

<http://www.tup.tsinghua.edu.cn>



中国计算机软件专业技术水平考试指定用书

程 序 设 计

(高级程序员级)

中国计算机软件专业技术资格和水平考试中心 组织编写

王春森 主编

北方工业大学

藏 书

图书馆

清华大学出版社

(京)新登字 158 号

内 容 简 介

本书是中国计算机软件专业技术资格和水平考试中心组织编写的考试指定用书,编写的依据是考试大纲,对高级程序员级考试应具备的知识和能力作了详尽的讲解。

本书的主要内容包括:计算机系统结构、语言处理程序、操作系统、软件工程、数据库、多媒体和网络等领域的基础知识,以及与程序设计有关的数据结构和基本算法。

本书是高级程序员资格和水平考试应试者的主要辅导读物,也可作为各类培训班、辅导班的教材,还可作为大专院校师生和广大工程技术人员学习相关知识的参考书。

版权所有,翻印必究。

本书封面贴有清华大学出版社激光防伪标签,无标签者不得销售。

图书在版编目(CIP)数据

程序设计:高级程序员级/王春森主编. —北京:清华大学出版社,1999

中国计算机软件专业技术水平考试指定用书

ISBN 7-302-03540-7

I. 程… II. 王… III. 程序设计 IV. TP311.11

中国版本图书馆 CIP 数据核字 (1999) 第 15138 号

出版者: 清华大学出版社(北京清华大学学研楼,邮编 100084)

<http://www.tup.tsinghua.edu.cn>

印刷者: 北京市清华园印刷厂

发行者: 新华书店总店北京发行所

开 本: 787×1092 1/16 **印张:** 27.75 **字数:** 653 千字

版 次: 1999 年 8 月第 1 版 1999 年 12 月第 2 次印刷

书 号: ISBN 7-302-03540-7/TP·1943

印 数: 31001~46000

定 价: 40.00 元

序

科学技术的日新月异,信息化时代的来临,使以计算机技术为基础的信息科学在经济和社会生活各个领域得到了极为广泛的应用,其发展水平成为衡量国家经济与科技实力的重要标志已是不争之事实。各国都把培养大量高水平计算机专业人才作为 21 世纪经济和科技发展的重要战略目标之一。一些经济发达国家通过开展计算机专业人才的教育培训,尤其是开展不同层次、不同规模的计算机水平测试,吸引、储备了大量计算机高级人才,为迎接日趋激烈的科技竞争奠定了坚实基础。这些国家的成功经验值得我们学习、借鉴。

中国计算机软件专业技术资格和水平考试自 1991 年开始实施至今已走过了 8 年的历程,共有近 40 万人参加考试,在国内外已产生较大影响。特别是 1999 年度中国计算机软件专业技术水平考试将由原来的一个专业扩展为程序设计、软件工程、数据库技术、计算机网络、多媒体技术 5 个专业,这无疑是中国计算机软件专业技术水平考试发展的一个质的飞跃,必将把中国计算机软件专业技术水平考试推向新的阶段。

我相信,新近编写出版的《中国计算机软件专业技术水平考试指定用书》能对广大应试者起到很好的指导作用。

我更加希望,在世纪之交,中国计算机软件专业技术水平考试能够抓住机遇,迎接竞争与挑战,为促进我国科教兴国战略的贯彻实施做出应有的贡献。

赵东伦

前 言

中国计算机软件专业技术资格和水平考试实施至今已经历了十余年历程,得到了社会的认可。为了满足我国信息技术发展和企业对计算机软件人才的需求,目前已将资格和水平考试的范围和内容扩大为程序设计、软件工程、计算机网络、多媒体技术和数据库五个方面,每个方面又分为若干级别,以适应社会上对各种软件人才的需求。

编者受中国计算机软件专业技术资格和水平考试中心委托,编写这套适用于程序设计方面考试要求的指定用书,分为初级程序员级、程序员级和高级程序员级三册。按照考试大纲要求,分别面向能熟练使用指定的常用软件和具有初步程序编制能力的初级程序员,能按照软件设计说明书编制程序的程序员,以及能按照软件需求规格说明书进行软件设计和编制程序的高级程序员。

按照考试大纲要求,对程序员级或高级程序员级所需的知识结构涉及的面很广,一个条目在大学里往往是一学期的课程,因此编写的难度很高。考虑到参加考试的人员应有一定的基础,所以书中只对考试大纲中所涉及到的知识领域的要点加以阐述,限于篇幅不能详细地展开,请读者谅解。再考虑到参加考试的人员一般都熟悉一种以上的程序设计语言,并已具有基本的编程能力,因此本书以数据结构和基本算法为主线,对读者原有的程序设计知识和编程能力起到总结、拓宽和提高的作用,使读者增强程序的设计和阅读能力,掌握常用数据结构和算法的原理和要点。考试大纲中要求应考者了解 CASL 汇编语言的程序编制,考生可参阅王世业编著、清华大学出版社出版的《CASL 汇编语言程序设计》一书。

全书共分八章。第一章计算机系统结构基础知识由丛锁、高传善编写,第二章语言处理程序基础知识由张守志编写,第三章操作系统基础知识由王编写,第四章软件工程基础知识由钱乐秋编写,第五章数据库基础知识由丁宝康编写,第六章多媒体基础知识由白雪峰编写,第七章网络基础知识由曹邦伟编写,第八章数据结构和基本算法由夏宽理编写。张作义、顾天珍、夏玲明和张培培等同志在本书编写过程中做了大量辅助性的工作,在此表示感谢。

因水平有限,书中难免存在错漏和不妥之处,望请读者指正,以利改进和提高。

编 者

1999 年 4 月于复旦大学

• II •

第1章 计算机系统结构基础知识

1.1 计算机系统结构概述

1.1.1 计算机系统结构、组成和实现

计算机系统结构(computer architecture)最初是由 G. M. Amdahl 等人在 1964 年提出的。Amdahl 在介绍 IBM360 系统时,把系统结构定义为从程序员的角度所看到的计算机的属性,即程序员要编写出能在机器上正确运行的程序所必须了解的概念性结构和功能特性。

按照计算机的层次结构,不同级别的程序员所看到的计算机属性显然不同。一般,低层机器的属性对高层机器程序员基本是透明(不可见)的,我们通常所说的计算机系统结构主要是指机器语言级机器(物理机器)的系统结构。

计算机系统结构主要研究软件、硬件功能分配,确定软件、硬件界面(机器级界面),即从机器语言程序员或编译程序设计者的角度所看到的机器物理系统的抽象。

计算机系统结构通常包括以下属性:

- 机内数据表示: 硬件能直接辨识和操作的数据类型和格式;
- 寻址方式: 最小可寻址单位、寻址方式的种类、地址运算;
- 寄存器组织: 操作数寄存器、变址寄存器、控制寄存器及专用寄存器的定义、数量和使用规则;
- 指令系统: 机器指令的操作类型、格式,指令间排序和控制机构;
- 存储系统: 最小编址单位、编址方式、主存容量、最大可编址空间;
- 中断机构: 中断类型、中断级别,以及中断响应方式等;
- 机器工作状态(管态、目态)的定义和切换;
- 输入输出结构: 输入输出的连接方式,处理机/存储器与输入输出设备间的数据交换的方式,数据交换过程的控制;
- 信息保护: 信息保护方式、硬件信息保护机制。

计算机组成(computer organization)指的是计算机系统结构的逻辑实现,包括机器内部数据流和控制流的组成以及逻辑设计等。其目标是合理地把各种部件、设备组成计算机,以实现特定的系统结构,同时满足所希望达到的性能价格比。一般而言,计算机组成研究的范围包括: 确定数据通路的宽度,确定各种操作对功能部件的共享程度,确定专用的功能部件,确定功能部件的并行度,设计缓冲和排队策略,设计控制机构和确定采用何种可靠性技术。

计算机实现(computer implementation)是指计算机组成的物理实现。包括处理机、主

存等部件的物理结构,器件的集成度和速度,器件、模块、插件、底板的划分与连接,专用器件的设计,信号传输技术,电源,冷却及装配等技术以及相关的制造工艺和技术。

计算机系统结构、计算机组成和计算机实现是三个不同的概念,但随着新技术、新器件的出现和应用,这三者之间的界限越来越模糊了。

1.1.2 计算机系统分类

1. Flynn 分类法

1966年,Michael. J. Flynn 提出根据指令流、数据流的多倍性(multiplicity)特征对计算机系统进行分类,定义如下:

指令流: 机器执行的指令序列;

数据流: 由指令流调用的数据序列,包括输入数据和中间结果;

多倍性: 在系统性能瓶颈部件上同时处于同一执行阶段的指令或数据的最大可能个数。

Flynn 根据不同的指令流—数据流组织方式把计算机系统分成四类:

(1) 单指令流单数据流 SISD (single instruction stream single data stream)

SISD 其实就是传统的顺序执行的单处理器计算机,其指令部件每次只对一条指令进行译码,并只对一个操作部件分配数据。如图 1.1(a)所示,流水方式的单处理机有时也被当作 SISD。

(2) 单指令流多数据流 SIMD(single instruction stream multiple data stream)

SIMD 以并行处理机(阵列处理机)为代表,其结构如图 1.1(b)所示。并行处理机包括多个重复的处理单元 $PU_1 \sim PU_n$,由单一指令部件控制,按照同一指令流的要求为它们分配各自所需的不同数据。相联处理机也属于这一类。

(3) 多指令流单数据流 MISD(multiple instruction stream single data stream)

如图 1.1(c)所示,具有 n 个处理单元,按 n 条不同指令的要求对同一数据流及其中间结果进行不同的处理。一个处理单元的输出又作为另一个处理单元的输入。这类系统实际上很少见到。有文献把流水线看作多个指令部件,称流水线计算机是 MISD。

(4) 多指令流多数据流 MIMD(multiple instruction stream multiple data stream)

如图 1.1(d)所示,MIMD 是指能实现作业、任务、指令等各级全面并行的多机系统。多处理机属于 MIMD。

2. 冯氏分类法

1972年,美籍华人冯泽云教授提出用最大并行度 P_m 来对计算机系统结构进行分类。 P_m 定义为计算机系统在单位时间内所能处理的最大二进制位数,平均并行度 P_a 定义为在一段时期内,计算机单位时间处理的二进制位数的平均值。

冯氏分类法把计算机系统分为四类:

(1) 字单位串 WSBS(word-serial and bit-serial): 字宽=1,位宽=1;

(2) 字并位串 WPBS(word-parallel and bit-serial): 字宽>1,位宽=1;

(3) 字单位并 WSBP(word-serial and hit-parallel): 字宽=1,位宽>1;

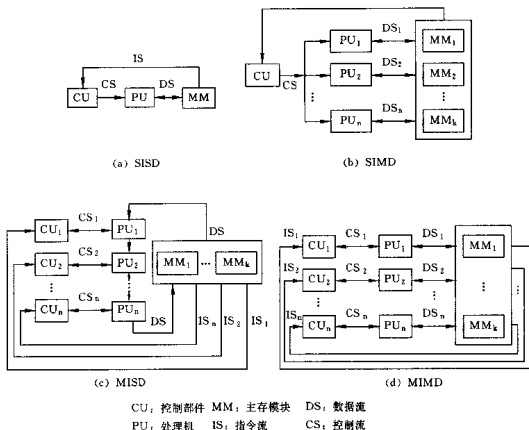


图 1.1 Flynn 分类法各类系统基本结构

(4) 字并位并 WPBP(word-parallel and bit-parallel): 字宽 >1 , 位宽 >1 。

其中, 字宽表示在一个字中同时处理的二进制位数, 位宽表示在一个位片中能同时处理的字数。

3. Handler 分类法

1977 年, Wolfigan Handler 根据并行度和流水线提出了另一种分类法, 把计算机硬件结构分为三个层次, 并分别考虑其可并行——流水处理的程度: 程序控制部件 PCU 的个数 K 、算术逻辑部件 ALU 或处理部件 PE 的个数 D 、每个算术逻辑部件包含基本逻辑线路 ELC 的套数 W 。一个计算机的系统结构可以表示为三元组 (K, D, W) 。

4. Kuck 分类法

1978 年, 美国的 David J. Kuck 提出了与 Flynn 法相似的分类法, 用指令流、执行流和多倍性来描述计算机系统特征:

- (1) 单指令流单执行流 SISE: 典型的单处理机;
- (2) 单指令流多执行流 SIME: 带多操作部件的处理机;
- (3) 多指令流单执行流 MISE: 带指令级多道程序的单处理机;
- (4) 多指令流多执行流 MIME: 多处理机。

1.2 存储器系统

存储器系统主要用于保存数据和程序,没有任何一种单一的技术能够完全优化地满足计算机存储系统的要求:高速存取、大容量和低成本。因为在存储器技术中,存在以下的制约关系:

- 存储器读写速率越高,每位的成本也越高;
- 存储器容量越大,每位的成本也越低;
- 存储器容量越大,读写速率越低。

解决这一难点的方法就是采用多级存储体系结构(memory hierarchy)。

1.2.1 存储器体系

1. 存储器系统的特征

(1) 位置: 存储器系统由分布在计算机各个不同部件的多种存储设备组成:位于 CPU 内部的寄存器,以及用于 CU 的控制存储器。内部存储器是可以被处理器直接存取的存储器,又称为主存储器(main memory);外部存储器(external memory)需要通过 I/O 模块与处理器交换数据,又称为辅助存储器(auxiliary memory)。除了正在执行的程序和处理的数​​据存放在主存储器外,其他所有的信息都保存在辅助存储器中,当需要时再从辅存中调入主存。

为了提高性能,弥补 CPU 处理速度和主存存取速度之间的差异,还设置了 cache(超高速缓冲存储器),其容量小但速度极快,位于 CPU 和主存储器之间,用于存放 CPU 正在执行的程序段和所需数据。

(2) 存储器单元: 衡量主存储器容量大小的单位通常是字节或字(word)(1 字=8, 16, 32 位),而外存储器的容量一般用字节表示。下面我们介绍三个与存储器单元相关的概念:

① 字(word): 字是存储器组织的基本单元,一个字可以是一个字节,也可以是多个字节。其实际大小,依赖于具体的机器实现。

② 可寻址单元(addressable unit): 通常,存储器的可寻址单元就是字,但有的存储器允许在更小的尺寸上(比如字节)进行寻址。如果某个系统的存储器地址宽度为 A 位,则该系统的可寻址单元数 $N=2^A$ 。

③ 传输单元(unit of transfer): 对主存而言,传输单元就是一次写入存储器或者从存储器读出的位数;对辅存而言,数据通常以更大的尺寸(如块)进行输入输出。

(3) 存取方式(method of accessing): 常用的存取方式有顺序存取、直接存取、随机存取和相联存取等四种。

① 顺序存取(sequential access): 存储器的数据以记录(record)的形式进行组织。对数据的访问必须按特定的线性顺序进行。使用一个共享的读写装置(R/W mechanism)对所有的数据进行访问。例如,假定当前访问的数据记录为 1,接着要访问第 N 个记录,就需要从第 1 个记录开始,顺序移动读写装置直到第 N 个记录,然后才能对它进行读写。访问

时间取决于读写装置访问存储单元的顺序。同样是对第 N 个记录进行读写,当读写装置位于不同的位置时,所需的时间也会不一样。磁带存储器的访问方式就是顺序存取。

② 直接存取(direct access):与顺序存取相似,直接存取也使用一个共享的读写装置对所有的数据进行访问。但是每个数据块(记录)都拥有唯一的地址标识,读写装置可以直接移动到目的数据块的所在位置进行访问。存取时间也是可变的。磁盘存储器就采用直接存取的方式。

③ 随机存取(random access):存储器的每一个可寻址单元都具有自己唯一的地址和读写装置,系统可以在相同的时间内对任意一个存储单元的数据进行访问,而与先前的访问序列无关。主存储器即是一例。

④ 相联存取(associative):这也是一种随机存取的形式,但是选择某一单元进行读写是取决于其内容而不是其地址。与普通的随机存取方式一样,每个单元都有自己的读写装置,读写时间也是一个常数。使用相联存取方式,可以对所有的存储单元的特定位进行比较,选择符合条件的单元进行访问。为了提高地址映射的速度,cache 可能采取该方式进行访问。

(4) 性能:存储器系统的性能主要由存取时间、存储器带宽、存储器周期和数据传输率等来衡量。

① 存取时间(access time):对于随机存取而言,存取时间就等于完成一次读或写所花的时间;对于非随机存取而言,存取时间就等于把读写装置移动到目的位置所需的时间。

② 存储器带宽(bandwidth):每秒钟能访问的位数。如果存储器周期是 500ns ,而每个周期可访问 4 字节,则带宽 $=64\text{Mbps}$ 。

③ 存储器周期(memory cycle time):主要是针对随机存取而言,一个存储器周期就等于两次相邻的存取之间所需的时间。

④ 数据传输率(transfer rate):每秒钟输入/输出的数据位数。对于随机存取而言,数据传输率 $R=1/\text{存储器周期}$;对于非随机存取而言, $T_N = T_A + N/R$, T_N :读写 N 位所需的平均时间, T_A :平均存取时间, R :数据传输率(bps),则数据传输率 $R=N/(T_N-T_A)$ 。

(5) 物理介质:常用的存储介质有半导体存储器(semiconductor memory)、磁介质存储器(magnetic surface memory)、光存储器三种。

(6) 物理特性

① 易失性存储器(volatile memory):当断电时,保存在易失性存储器里的内容就会丢失。

② 非易失性存储器(nonvolatile memory):保存在非易失性存储器里的内容不会因断电而丢失。除非对其进行写入操作,否则数据不会被破坏。

③ 不可改写存储器(nonerasable memory):数据一旦写入,就无法更改。

2. 多级存储器体系

计算机采用多级存储器体系的目的就在于能够获得尽可能高的存取速率,同时保持较低的成本。

如图 1.2 所示,从存储器体系顶端的寄存器和 cache 到底部的磁带存储器,其价格逐

渐降低;而容量和访问时间则逐渐增加。通过合理地设计存储器体系结构,使得经常访问的数据保存在高速存储器中,而较少访问的数据保存在低速存储器中,即从高到低,数据的访问频率逐渐减少。这样就可能组成一个访问速度尽可能快,容量尽可能大而价格也较便宜的存储系统。



图 1.2 计算机的存储体系

1.2.2 主存储器

1. 主存储器的种类

(1) RAM(random access memory): RAM 存储器既可以写入也可以读出,但断电后信息无法保存,因此只能用于暂存数据。RAM 又可分为 DRAM 和 SRAM 两种:

① DRAM(dynamic RAM): 信息会随时间逐渐消失,因此需要定时对其进行刷新(refresh)维持信息不丢失。

② SRAM(static RAM): 在不断电的情况下信息能够一直保持而不会丢失。

DRAM 的密度大于 SRAM 且更加便宜,但 SRAM 速度快,电路简单(无需刷新电路),然而容量小,价格高。

(2) ROM(read only memory): 只读存储器,信息已固化在存储器中。ROM 出厂时其内容由厂家用掩膜技术(mask)写好,只可读出,但无法改写。一般用于存放系统程序 BIOS 和用于微程序控制。

(3) PROM(programmable ROM): 可编程 ROM,只能进行一次写入操作(与 ROM 相同),但是可以在出厂后,由用户使用特殊电子设备进行写入。

(4) EPROM(erasable PROM): 可擦除的 PROM,其中的内容既可以读出,也可以写入。但是在一次写操作之前必须用紫外线照射 15~20 分钟以擦去所有信息,然后再写入。可以写多次。

(5) E²PROM(electrically EPROM): 电可擦除 EPROM,与 EPROM 相似,可以读出也可写入,而且在写操作之前,不需要把以前内容先擦去。能够直接对寻址的字节或块进行修改,只不过写操作所需的时间远远大于读操作所需时间(每字节需几百 ms),其集成度也较低。

(6) 闪速存储器(flash memory): 其性能介于



图 1.3 存储单元示意图

EPROM与 E²PROM 之间。与 E²PROM 相似,可使用电信号进行删除操作。整块闪存存储器可以在数秒内删除,速度远快于 EPROM;而且可以选择删除某一块而非整块芯片的内容,但还不能进行字节级别的删除操作。集成度与 EPROM 相当,高于 E²PROM。

2. 主存储器的组成

(1) 存储器单元(memory cell): 一个简单的存储器单元如图 1.3 所示。

① 存储器单元的基本性质: 具有两个稳定状态用以表示“1”和“0”;至少可以写入一次;可以读出(to sense the state)。

② 存储器单元的操作: select 信号用于选择该单元进行读写;control 信息用于表明执行读还是写操作;data 线用于数据的输入输出。

(2) 存储器芯片(memory chip): 由多个单元矩阵组成芯片,再由若干芯片组成存储体。现代计算机一般都采用半导体存储器。近来半导体工艺发展很快,以 DRAM 为例,集成度平均每年提高 60%,但价格却不断下降。当前使用最多的是 4Mb 和 16Mb 芯片。

我们以 DRAM 芯片为例,讲解存储器工作的基本原理,如图 1.4 所示。

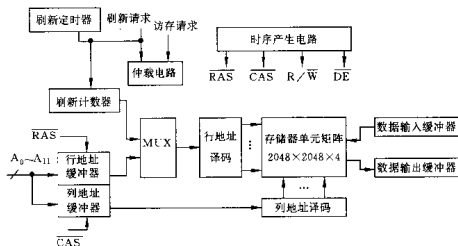


图 1.4 存储器的基本结构

① 存储单元矩阵: 由 4 个 2048×2048 单元矩阵组成存储体。

② 地址信号: 寻址 2048×2048 个字需要 22 条地址线 ($2^{22} = 2048 \times 2048$), 但为了节省芯片的引脚, 只使用 $A_0 \sim A_{11}$ 11 条地址线。由行选信号 $\overline{RAS}$ 和列选信号 $\overline{CAS}$ 分别控制。用 $\overline{RAS}$ 信号控制 $A_0 \sim A_{11}$ 从 2048 行中选出一行, 用 $\overline{CAS}$ 信号控制 $A_0 \sim A_{11}$ 再从 2048 列中选出一列。

③ 时序发生器: 在时钟信号与控制信号的控制下, 产生存储器读写所需的时序信号: $\overline{RAS}$ 、 $\overline{CAS}$ 、 $\overline{W}$ 和 $\overline{DE}$ 信号。

④ 刷新电路: 由刷新定时期定时产生刷新请求, 启动刷新计数器产生刷新所需的行地址。因为刷新时是将该行所有存储单元进行刷新, 因此只需提供刷新行地址。每发出一条刷新请求, 计数器值加 1, 以便刷新下一行。

假定 2ms 对全部 DRAM 刷新一次, 而共有 $2^{11} = 2K$ 行, 因此约 $1\mu s$ ($2ms/2K \approx 1\mu s$)

即需发出一次刷新请求。

⑤ 仲裁电路：当同时存在访存请求和刷新请求时，由仲裁电路依据机器的状态判断优先处理哪个请求。

(3) 存储器的组成：实际的存储器总是由一片或多片存储芯片配以控制电路组成的。其容量往往用 $W \times B$ 来表示。 W 表示该存储器的存储单元(word)的数量而 B 表示每一个 word 由多少 bit 组成。如果某一芯片规格为 $w \times b$ ，则用其组成 $W \times B$ 的存储器需要用 $\frac{W}{w} \times \frac{B}{b}$ 个芯片。

首先是 B/b 个芯片一组，每组芯片共有 w 个字，每个字则有 B 个位，一共用 W/w 个组来构造所需的存储器。

① 分组：如图 1.5 所示， B/b 个芯片组成一组，共用地址线、读写线 R/W 和片选线 CS ；它们的数据线组成了 B 位宽的数据。

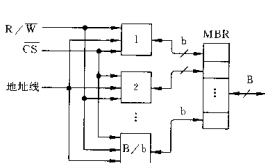


图 1.5 存储器芯片分组

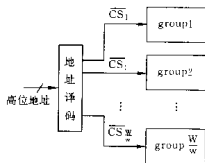


图 1.6 存储器芯片的组间互联

② 组间互联：如图 1.6 所示，用高位地址线译码产生组选择信号 CS ，低位地址线则用来在组内寻址。

1.2.3 辅助存储器

辅助存储器用于存放当前不需要立即使用的信息，一旦需要，再和主机成批交换数据，是主存储器的后备，因此称之为辅助存储器；它又是主机的外部设备，又称之为“外存储器”。辅助存储器的最大特点是存储器容量大、可靠性高、价格低。主要可分为磁表面存储器和光存储器两大类。

1. 磁表面存储器

磁表面存储器目前主要有磁带和磁盘存储器。

(1) 磁记录原理：磁表面存储器通过磁头来进行读写。磁头线圈中通入一定方向和大小的脉冲电流，磁化磁头的导磁体，从而建立起一定方向和强度的磁场。当磁介质从磁头下经过时，从磁头间隙处将其磁化。由于剩磁效应的存在，当磁场消失后，磁介质表面仍存在剩磁。线圈中电流方向不同，磁介质上被磁化的方向不同——用以代表“1”和“0”。随着电流的变化以及磁介质磁头间的相对运动，就可以把二进制信息序列转化为介质表面的磁化单元序列。读出过程则是逆过程，即把记录的磁化单元序列还原成电流脉冲序列。

(2) 磁带存储器：磁带存储设备是一种顺序存取的设备，存取时间较长，但存储容量大，便于携带，价格便宜，所以也是一种主要的辅助存储器。磁带的内容由磁带机进行读写（最便宜也最慢）。

① 工作原理：按磁带机的读写方式主要可以分为两种，启停式和数据流。

• 启停式磁带机：磁带按宽带可以分为 1/4 英寸 1/2 英寸和 1 英寸三种。以 1/2 英寸为例，其标准格式如图 1.7 所示。

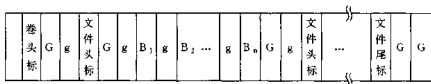


图 1.7 1/2 英寸磁带的记录格式

磁带上的信息以文件块的形式存放。整盘磁带的开始有一卷标标明，然后有一初始空白块，用以适应磁带从静止到稳定带速所需时间。文件记录以文件头标志和文件尾标志标识，一个文件由若干数据块组成，每一数据块又由若干记录组成（一个数据块所包括的记录条数叫块因子）。数据块之间以空白块(gap)进行分隔。文件之间也存在一段空隙 G，所有的文件都顺序地排列在磁带上，一个文件的长度不仅包括记录信息，也包括块间间隔。

磁带机每一次读写信息的位数与磁带表面并行记录信息的磁道数有关：如 7 道、9 道和 16 道，则分别有 7, 9, 16 个磁头并列，一次可以读写 7 位、9 位或 16 位。

• 数据流磁带机：数据流磁带机结构简单，价格低，数据传输速率快。数据流磁带机的数据格式见图 1.8。其记录格式是串行逐道记录信息，每次读写 1 位信息，数据连续地写在磁带上，数据块之间以空隙分隔。磁带机不能在块间启停。读写顺序如下：(4 个磁道)先从 0 道的首端(BOT)开始，到其末端(EOT)；然后第 1 道反向记录从 EOT 到 BOT，而 2 道又正向从 BOT 到 EOT，3 道再反向。



图 1.8 数据流磁带机的数据格式

② 性能参数

• 数据传输速率(B/s) = 磁带记录密度(B/mm) × 带速(mm/s)

• 数据块的长度 = B1(记录数据所需的长度) + B2(块间间隔)

• $B1 = (\text{字节数/记录}) \times \text{块因子/记录密度}$

• 读 N 条记录所需的时间： $T = S(\text{启停时间}) + R + D$

$R(\text{有效时间}) = (N \times \text{字节数/记录}) / \text{传输速率}$

$D(\text{间隔时间}) = \text{块间隔总长} / \text{带速} = [(N / \text{块化因子}) \times (\text{块间间隔})] / \text{带速}$

(3) 磁盘存储器

① 数据的组织：如图 1.9 所示，磁盘上的数据都存放于磁道上。磁道就是磁盘上的一组同心圆，其宽度与磁头的宽度相同。为了避免减少干扰，磁道与磁道之间要保持一定的

间隔(inter-track gap),沿磁盘半径方向,单位长度内磁道的数目称之为道密度(道/英寸, TPI),最外层为0道。

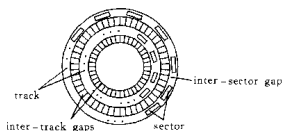


图 1.9 磁盘的数据组织

- 位密度: 沿磁道方向,单位长度内存储二进制信息的个数叫位密度。

为了简化电路设计,每个磁道存储的位数都是相同的,所以其位密度也随着从外向内而增加。

• 扇区: 磁盘的数据传输是以块为单位的,所以磁盘上的数据也以块的形式进行存放。这些块就称为扇区(sector),每个磁道通常包括10~100个扇区。同样为了避免干扰,扇区之间也相互留有空隙(inter-sector gap)。

• 柱面: 柱面是若干个磁盘组成的磁盘组,所有盘面上相同位置的磁道组称为一个柱面(每个柱面有n个磁道);若每个磁盘有m个磁道,则该磁盘组共有m个柱面。

② 数据记录格式: 为了让磁头读写定位机构能够确定磁盘上某一磁道的具体位置,因此在新磁盘使用之前,必须对其进行格式化,以写入控制数据供磁盘驱动器使用,该数据对用户不可见。从没有格式化过的盘叫做白盘,上面只是记录了有多少磁道、磁道密度等全球统一标准,还不能直接使用。Seagate ST506的数据格式如图1.10所示。

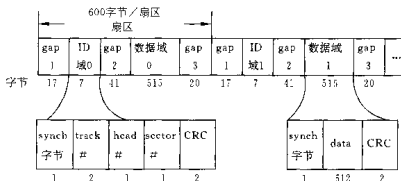


图 1.10 Seagate ST506 的数据记录格式

③ 磁盘存储器的特征

• 磁头: 固定磁头磁盘(fixed-head disk);磁头不沿半径方向运动,每一磁道都有一个用于读写数据的磁头。可移动头磁盘(movable-head disk);每一面只有一个磁头,可以沿径向运动,停留在任一磁道的上方。

• 磁盘驱动器：磁盘必须置于驱动器才能进行读写。不可替换盘(nonremovable disk)的盘片永久固定在驱动器内不可更换。可替换盘(removable disk)则可以替换驱动器内的盘片，因此可以在不同系统中交换盘片使用。

• 面：单面/ 双面磁盘。

• 磁盘组：由多个磁盘组成。

• 磁头机构：根据磁头与磁盘之间的间隔可以分为三类：

① 接触式(contact)：磁头在读写时与盘面接触，用于软磁盘(floppy disk)存储器；

② 固定空隙(fixed gap)：留有一个固定的空气隙(air gap)。为了对磁盘进行读写，磁头必须感应或产生相当强度的电磁场。为了提高数据密度就需要提高道密度，即缩小道的宽度——也就是需要减少磁头的尺寸，这样一来，磁头就必须更加接近盘面以满足读写需求。但是这样由于灰尘杂质的影响更加容易出错。

③ 温盘(winchester)：磁盘置于密封的驱动器内以避免灰尘杂质的影响，磁头更加接近盘面以获取更高的数据密度。当磁盘静止时磁头轻置于盘面上，当磁盘旋转时其产生的空气压力足以使磁头悬浮起来。因此可以使用更小尺寸的磁头。

硬盘的存储介质是铝合金等“硬”材料，并在其表面覆以磁性材料。软盘的介质是以塑料等软材料制成，其成本低、重量轻，易携带易保存，给计算机信息的转储、拷贝提供了极大的方便；但速度慢，易损坏，存储容量小。

④ 性能参数

• 存储容量(capacity)：磁盘格式化后能够存储有用信息的总量：

$$\text{存储容量} = n \times t \times s \times b$$

n：保存数据的总盘面数，t：每面磁道数，s：每道的扇区数，b：每个扇区存储的字节数。

• 存取时间(access time)：包括寻道时间和等待时间。寻道时间(查找时间，seek time)为磁头移动到目标磁道所需的时间(movable-head disk)，对于固定磁头磁盘而言，无需移动磁头，只需选择目标磁道对应的磁头即可。等待时间为等待读写的扇区旋转到磁头下方所用的时间。一般选用磁道旋转一周所用时间的一半作为平均等待时间。寻道时间由磁盘机的性能决定，目前主流硬盘典型的AST(average seek time)一般在10ms左右，而转速则有2400rpm, 5400rpm, 7200rpm等等。软盘转速较慢，一般只有360rpm(因为磁头与盘面接触性读写)。

• 数据传输速率：磁头找到地址后，单位时间写入或读出的字节数。

$R = TB / T$ ，其中：TB为一个磁道上记录的字节数，T为磁盘每转一圈所需的时间，R为数据传输速率。

(4) RAID存储器：因为辅助存储器性能改进的速度远远跟不上处理机以及主存储器性能的改进，所以存储系统性能的提高也就是改善整个计算机系统性能的关键。并行系统结构提供了一种采用低成本器件技术获得高性能的方法。

1988年美国加州大学伯克利分校D. A. Patterson教授提出了廉价磁盘冗余阵列RAID(redundant array of inexpensive disks)技术，旨在缩小日益扩大的CPU速度和磁盘存储器速度之间的差距。其策略是用多个较小的磁盘驱动器替换单一的大容量磁盘驱

驱动器,同时合理地多个磁头上分布存放数据以支持同时从多个磁盘进行读写,从而改善了系统的 I/O 性能。小容量驱动器阵列与大容量驱动器相比,具有成本低,功耗小,性能好等优势;低代价的编码容错方案在保持阵列的速度与容量优势的同时保证了极高的可靠性。同时也较容易扩展容量。但是由于允许多个磁头同时进行操作以提高 I/O 数据传输速度,因此不可避免地提高了出错的概率。为了补偿可靠性方面的损失,RAID 使用存储的校验信息(stored parity information)来从错误中恢复数据。最初, inexpensive 一词主要针对当时另一种技术 SLED(single large expensive disk)而言,但随着技术的发展, SLED 已是明日黄花, RAID 和 non-RAID 皆采用了类似的磁盘技术。因此 RAID 现在代表独立磁盘冗余阵列(redundant array of independent disks),用 independent 来强调 RAID 技术所带来的性能改善和更高的可靠性。

RAID 机制中共分 6 个级别,工业界公认的标准分别为 RAID0~RAID5,其共同特征如下:

- RAID 由若干个物理驱动器组成,但对操作系统而言仍是一个逻辑驱动器;
- 数据分布在阵列中多个驱动器上;
- 冗余的磁盘容量用以保存奇偶信息(parity),以便在磁盘失效时进行恢复(RAID0 不支持该特征)。

RAID 应用的主要技术有:

- 分块技术(striping):对主机请求读写的数据进行分块,使之分布于多台磁盘上;
- 交叉技术(interleaving):对存放在多台磁盘上数据的读写,采取交叉方式进行;
- 重聚技术(declustering):对多台磁盘上的存储空间进行重新编址,使数据按重新编址后的空间进行存放。

① RAID0 级(无冗余和无校验的数据分块):具有最高的 I/O 性能和最高的磁盘空间利用率,易管理,但系统的故障率高,属于非冗余系统,主要应用于那些关注性能、容量和价格而不是可靠性的应用程序。

所有的系统/用户数据表面上看起来存放在一个逻辑驱动器上,但实际上这个驱动器被分成若干块(strip,可能是 physical blocks, sectors...)这些 strip 轮转地映射到相邻的阵列磁盘中。

② RAID1 级(磁盘镜像阵列):由磁盘对组成,每一个工作盘都有其对应的镜像盘,上面保存着与工作盘完全相同的数据拷贝,具有最高的安全性,但磁盘空间利用率只有 50%。RAID1 主要用于存放系统软件、数据以及其他重要文件。它提供了数据的实时备份,一旦发生故障所有的关键数据即刻就可使用。

RAID1 通过简单的数据备份实现冗余,也使用了数据分块技术。任一个逻辑块都对应着两个(映射到)不同的物理驱动器上。对块的读请求可以由两个盘中任一个进行服务,往往是由寻址时间短的驱动器进行服务。对块的写请求,两个块都需要更新,但是可以并行进行。因为无需计算校验位,所以比 RAID2~5 的写入更快。很容易从失效中恢复。但缺点是代价很大。

③ RAID2 级(采用纠错海明码的磁盘阵列):采用了海明码纠错技术,用户需增加校验盘来提供单纠错和双纠错功能。对数据的访问涉及到阵列中的每一个盘。大量数据传