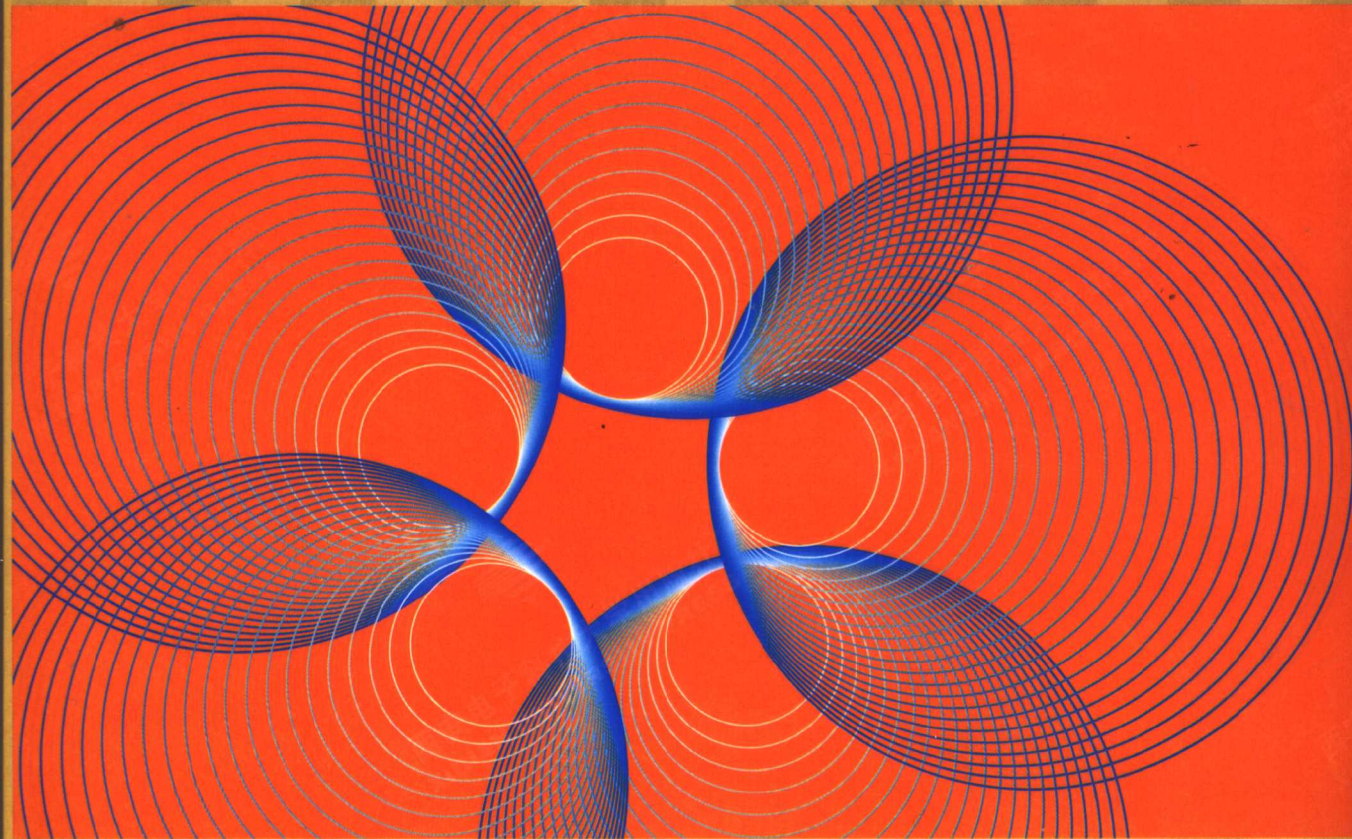


全国计算机技术与软件专业技术资格（水平）考试辅导用书



数据库系统工程师考试辅导

全国计算机技术与软件专业技术资格（水平）考试办公室组编

王亚平 主编 刘强 副主编

西安电子科技大学出版社
[http:// www.xduph. com](http://www.xduph.com)

全国计算机技术与软件专业技术资格(水平)考试辅导用书

数据库系统工程师考试辅导

全国计算机技术与软件专业技术资格水平考试办公室组编

王亚平 主 编

刘 强 副主编

西安电子科技大学出版社

2004

内 容 简 介

本书是根据《计算机技术与软件专业技术资格(水平)考试大纲(数据库系统工程师级)》编写的考试辅导书。全书共16章,主要内容包括:计算机系统知识、数据结构与算法、操作系统知识、程序语言基础知识、网络基础知识、多媒体基础、数据库技术基础、关系数据库、SQL语言、系统开发和运行知识、数据库设计、数据库运行与管理、网络与数据库、数据库发展趋势与新技术、知识产权基础和标准化基础。本书每章包括学习目标与要求、知识点概述、典型例题与分析、强化训练习题。

本书浓缩了考试复习内容,知识精练,重点突出,例题丰富,解答详细,既可作为计算机技术与软件专业技术资格(水平)考试的应试辅导教材,也可作为大专院校师生的教学参考书。

图书在版编目(CIP)数据

数据库系统工程师考试辅导/王亚平主编. —西安:西安电子科技大学出版社,2004.8

全国计算机技术与软件专业技术资格(水平)考试辅导用书

ISBN 7-5606-1433-7

I. 数... II. 王... III. 数据库系统—工程师—资格考核—自学参考资料 IV. TP311.13

中国版本图书馆 CIP 数据核字(2004)第 066462 号

策 划 臧延新 陈宇光

责任编辑 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

http://www.xduph.com

E-mail: xdupfxb@pub.xaonline.com

经 销 新华书店

印刷单位 陕西画报社印刷厂

版 次 2004年8月第1版 2004年9月第2次印刷

开 本 787毫米×1092毫米 1/16 印张 35.375

字 数 847千字

印 数 4 001~10 000册

定 价 50.00元

ISBN 7-5606-1433-7/TP·0764

XDUP 1704001-2

*** 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

全国计算机软件考试实施至今已经历了十多年，在社会上产生了很大的影响，对我国软件产业的形成和发展做出了重要的贡献。为适应我国信息化发展的需要，国家人事部和信息产业部决定将考试的级别拓展到计算机技术与软件的各个方面，以满足社会上对各种计算机信息技术人才的需求。

为了适应新的考试大纲要求，编者受信息产业部计算机软件专业技术资格(水平)考试办公室委托，在《数据库系统工程师教程》一书的基础上编写了这本关于数据库系统工程师的辅导用书。在考试大纲中，要求考生掌握计算机系统知识、数据库基础知识、数据库及数据库应用系统设计、数据库应用系统实施、数据库系统的运行和管理、网络与数据库、数据库发展及知识产权与标准等方面的知识，由于涉及到的内容多、知识面广，学习有一定的难度，因此作者在编写本书时，注重与原教程内容同步，对教程中的难点和重点知识进行补充。本书的每一章由学习目标与要求、知识点概述、典型例题与分析、强化训练习题等四部分组成。

本书共分16章，由王亚平担任主编，刘强担任副主编。第1章计算机系统知识由李伯成、王亚平编写，第2章数据结构与算法由王卫东、张淑平编写，第3章操作系统知识由王亚平编写，第4章程序语言基础知识由张淑平编写，第5章网络基础知识由张凤琴编写，第6章多媒体基础由刘强编写，第7章数据库技术基础由王亚平编写，第8章关系数据库和第9章SQL语言由王亚平编写，第10章系统开发和运行知识由褚华编写，第11章数据库设计和第12章数据库的运行与管理由苏向阳编写，第13章网络与数据库由胡圣明编写，第14章数据库发展趋势与新技术由王小兵编写，第15章知识产权基础和第16章标准化基础由刘强编写，最后由王亚平统稿。

在本书的编写过程中，参考了许多相关的书籍和资料，编者在此对这些参考文献的作者表示感谢。同时感谢西安电子科技大学出版社在本书出版过程中所给予的支持和帮助。

因水平有限，书中难免存在错漏和不妥之处，望读者指正，以利改进和提高。

编 者
2004年4月

目 录

第1章 计算机系统知识	1
1.1 学习目标与要求	1
1.2 知识点概述	1
1.2.1 计算机系统的组成	1
1.2.2 计算机基本工作原理	4
1.2.3 计算机体系结构	8
1.2.4 安全性、可靠性与系统性能 评测基础知识	21
1.3 典型例题与分析	31
1.4 强化训练习题	46
第2章 数据结构与算法	50
2.1 学习目标与要求	50
2.2 知识点概述	50
2.2.1 常用数据结构	50
2.2.2 常见算法设计方法	91
2.3 典型例题与分析	94
2.4 强化训练习题	101
第3章 操作系统知识	105
3.1 学习目标与要求	105
3.2 知识点概述	105
3.2.1 操作系统的基础知识	105
3.2.2 处理机管理	109
3.2.3 存储管理	122
3.2.4 设备管理	129
3.2.5 文件管理	135
3.2.6 作业管理	144
3.2.7 网络操作系统与嵌入式操作系统	148
3.2.8 操作系统实例	149
3.3 典型例题与分析	161
3.4 强化训练习题	168
第4章 程序语言基础知识	172

4.1 学习目标与要求	172
4.2 知识点概述	172
4.2.1 程序设计语言的基础知识	172
4.2.2 汇编程序的基本原理	182
4.2.3 编译程序的基本原理	182
4.2.4 解释程序的基本原理	200
4.3 典型例题与分析	201
4.4 强化训练习题	206
第5章 网络基础知识	210
5.1 学习目标与要求	210
5.2 知识点概述	210
5.2.1 网络概述	210
5.2.2 ISO/OSI 网络体系结构	211
5.2.3 网络的协议与标准	212
5.2.4 构建网络	215
5.2.5 Internet 及应用	216
5.2.6 Windows NT 系统及管理	218
5.2.7 网络安全	219
5.3 典型例题与分析	220
5.4 强化训练习题	228
第6章 多媒体基础	238
6.1 学习目标与要求	238
6.2 知识点概述	238
6.2.1 多媒体的基本概念	238
6.2.2 音频信息及处理	240
6.2.3 图形和图像	244
6.2.4 动画和视频	250
6.2.5 多媒体网络	255
6.2.6 多媒体计算机系统	256
6.2.7 虚拟现实的概念	259
6.3 典型例题与分析	261
6.4 强化训练习题	272

第7章 数据库技术基础	278	10.2.3 系统设计基础知识	417
7.1 学习目标与要求	278	10.2.4 系统实施基础知识	422
7.2 知识点概述	278	10.2.5 系统运行和维护知识	429
7.2.1 基本概念	278	10.3 典型例题与分析	432
7.2.2 数据模型	282	10.4 强化训练习题	448
7.2.3 DBMS的功能及特征	289		
7.2.4 数据库系统体系结构	290	第11章 数据库设计	454
7.2.5 数据库的控制功能	294	11.1 学习目标与要求	454
7.2.6 数据仓库和数据挖掘基础知识	300	11.2 知识点概述	454
7.3 典型例题与分析	307	11.2.1 数据库设计概述	454
7.4 强化训练习题	312	11.2.2 系统需求分析	455
		11.2.3 概念结构设计	456
		11.2.4 逻辑结构设计	457
		11.2.5 数据库的物理设计	458
第8章 关系数据库	314	11.2.6 应用程序设计	459
8.1 学习目标与要求	314	11.2.7 数据库系统的实现	459
8.2 知识点概述	314	11.2.8 系统实施与维护	460
8.2.1 概述	314	11.2.9 数据库的保护	460
8.2.2 关系运算	318	11.3 典型例题与分析	461
8.2.3 查询优化	333	11.4 强化训练习题	466
8.2.4 关系数据库设计基础理论	339		
8.3 典型例题与分析	351	第12章 数据库运行与管理	469
8.4 强化训练习题	359	12.1 学习目标与要求	469
		12.2 知识点概述	469
		12.2.1 数据库系统运行计划	469
第9章 SQL语言	364	12.2.2 数据库系统运行维护	470
9.1 学习目标与要求	364	12.2.3 数据库系统的管理	472
9.2 知识点概述	364	12.2.4 性能调整	473
9.2.1 数据库语言	364	12.2.5 用户支持	474
9.2.2 数据库定义	366	12.3 典型例题与分析	474
9.2.3 数据操作	370	12.4 强化训练习题	475
9.2.4 SQL中的授权	379		
9.2.5 创建触发器	384	第13章 网络与数据库	476
9.2.6 嵌入式SQL	386	13.1 学习目标与要求	476
9.2.7 SQL-99所支持的对象	389	13.2 知识点概述	476
关系模型	389	13.2.1 分布式数据库	476
9.3 典型例题与分析	403	13.2.2 Web与数据库	484
9.4 强化训练习题	408	13.2.3 XML与数据库	490
		13.3 典型例题与分析	493
		13.4 强化训练习题	493
第10章 系统开发和运行知识	412		
10.1 学习目标与要求	412	第14章 数据库发展趋势与新技术	496
10.2 知识点概述	412		
10.2.1 软件工程基础知识	412		
10.2.2 系统分析基础知识	415		

14.1	学习目标与要求	496	15.2.12	专利法概述	515
14.2	知识点概述	496	15.2.13	企业知识产权的保护	519
14.3	典型例题与分析	499	15.3	典型例题与分析	520
14.4	强化训练习题	501	15.4	强化训练习题	528
第 15 章	知识产权基础	504	第 16 章	标准化基础	533
15.1	学习目标与要求	504	16.1	学习目标与要求	533
15.2	知识点概述	504	16.2	知识点概述	533
15.2.1	知识产权的概念与特点	504	16.2.1	标准化的概念	533
15.2.2	计算机软件著作权的主体与 客体	506	16.2.2	标准化基本过程	534
15.2.3	计算机软件受著作权法保护的 条件	507	16.2.3	标准的分类	535
15.2.4	计算机软件著作权的权利	507	16.2.4	标准的编号	538
15.2.5	计算机软件著作权的行使	508	16.2.5	国际标准及国外先进标准	539
15.2.6	计算机软件著作权的保护期	509	16.2.6	采用国际标准和国外先进标准	539
15.2.7	计算机软件著作权的归属	509	16.2.7	标准化组织	540
15.2.8	计算机软件著作权侵权的鉴别	511	16.2.8	信息技术标准化	540
15.2.9	不构成计算机软件侵权的 合理使用行为	512	16.2.9	ISO9000:2000 标准	542
15.2.10	计算机软件著作权侵权的 法律责任	512	16.2.10	能力成熟度模型(CMM)	543
15.2.11	计算机软件的商业秘密权	513	16.2.11	过程评估标准 ISO/IEC 15504	544
			16.3	典型例题与分析	545
			16.4	强化训练习题	554

第1章 计算机系统知识

1.1 学习目标与要求

本章的学习目标是：掌握有关计算机的一些预备知识；掌握数制转换、数的编码表示及运算方法；学习并熟悉计算机的组成部分，包括 CPU、指令系统、内存与外存、总线等以及各部分的基本功能。

★ 通过本章的学习，要求掌握如下内容：

(1) 熟悉计算机系统结构的基本概念与计算机系统的分类，并掌握 CPU 的两大构成部件：运算器和控制器；熟悉算术逻辑运算的基本方法及构成运算器的基本思路；熟悉两种不同的控制器——组合逻辑控制器和微程序控制器的构成方法；掌握 CPU 的指令系统及寻址方式。

(2) 掌握计算机层次存储系统原理、存储设备的基本知识、内存的组成与连接；熟悉高速缓存 Cache 和虚拟存储器的基本概念。

(3) 掌握硬磁盘存储器的基本构成及其基本性能指标。

(4) 了解 RISC 的定义、特点与相关技术分析；了解指令流水原理、技术、相关性分析与流水线的概念以及流水线性能评价。

(5) 熟悉常见总线的特点、I/O 系统原理。

(6) 熟悉阵列处理机(SIMD 计算机)和并行处理机。

(7) 掌握有关计算机安全的基本知识；了解计算机系统的故障诊断与容错。

(8) 掌握计算机的可靠性概念及指标；熟悉计算机系统的性能评价技术。

1.2 知识点概述

1.2.1 计算机系统的组成

一、要求掌握的知识要点

(1) 熟悉构成计算机硬件的几个主要组成部分以及各部分的主要功能。

(2) 熟悉构成计算机软件的主要组成部分。

二、知识点概述

(一) 计算机发展概述

计算机的发明和应用是 20 世纪人类最重要的成就之一，标志着信息时代的开始。在过

去的 50 多年里, 计算机技术得到了飞速发展, 计算机及其应用已经渗透到社会的各个领域, 有力地推动了社会信息化的进程。目前, 一个国家计算机的应用水平直接标志着一个国家的科学现代化水平。

计算机的发展经历了以下五个重要的阶段。

1. 大型机阶段

1946 年美国宾州大学研制的第一台计算机 ENIAC (Electronic Numerical Integrator and Calculator) 被公认为大型机的鼻祖。ENIAC 采用电子管作为基本逻辑部件, 体积大, 耗电量多, 寿命短, 可靠性差, 成本高。由于采用电子射线管作为存储部件, 所以 ENIAC 的容量很小。

大型机 (Mainframe) 的发展经历了以下几代。

第一代: 采用电子管制作的计算机阶段;

第二代: 采用晶体管制作的计算机阶段;

第三代: 采用中、小规模集成电路制作的计算机阶段;

第四代: 采用大规模、超大规模集成电路制作的计算机阶段。其代表机型有 IBM 360/370/709/4300/9000 等等。

2. 小型机阶段

小型机 (Minicomputer) 或称小型电脑, 通常用以满足部门的需要, 被中小型企业事业单位使用。例如, DEC 公司的 VAX 系列机 (配备 UNIX 操作系统)。

3. 微型机阶段

微型机 (Microcomputer) 又称微电脑或个人电脑 (Personal Computer, 或 PC 机)。顾名思义, 该机是面向个人或家庭的, 它的价格与高档家用电器相当, 应用相当普及。例如 Apple II、IBM-PC 系列机。

4. 客户机/服务器阶段

1964 年美国航空公司建立了第一个联机订票系统, 将全美的 2000 个订票终端用电话线连在一起。订票中心的大型机 (即服务器) 用来处理订票事务, 而分散在各地的订票终端则称为客户机。从逻辑上来看, 这是早期的客户机/服务器模式。

早期的客户机/服务器模式主要是为客户机提供资源共享的磁盘服务器和文件服务器, 而现在的服务器主要是数据库服务器和应用服务器等。

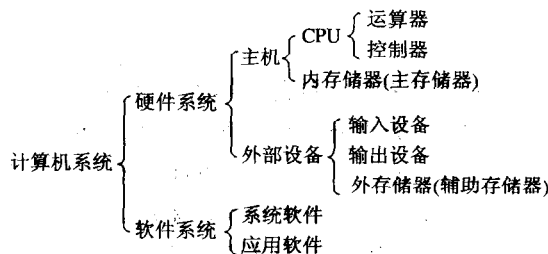
客户机/服务器 (Client/Server) 模式是对大型机的一次挑战。由于客户机/服务器模式结构灵活, 适应面广, 成本较低, 因此得到了广泛的应用。如果服务器的处理能力强, 客户机的处理能力弱, 则称为瘦客户机/胖服务器; 否则称之为胖客户机/瘦服务器。

5. 互联网阶段

自 1969 年美国国防部 ARPANET 网运行以来, 计算机广域网开始发展起来。1983 年 TCP/IP 传输控制与互联网协议正式成为 ARPANET 网的标准协议, 这使得网际互联有了突飞猛进的发展。以它为主干发展起来的因特网 (Internet) 到 1990 年已连接到 3000 多个网络和 20 万台计算机。进入 20 世纪 90 年代, 因特网继续以指数级迅猛扩展。进入 21 世纪, 全球已有上亿因特网用户。到 1994 年, 我国采用 TCP/IP 协议通过四大主干网接入因特网。目前全国的因特网用户已超过 3000 万。

计算机系统是由硬件系统和软件系统组成的。计算机硬件是计算机系统中看得见、摸

得着的物理装置, 计算机软件是程序、数据和相关文档的集合, 如下所示:



(二) 计算机硬件系统结构

1. 计算机的硬件组成

计算机硬件由运算器、控制器、存储器、输入设备和输出设备五大部件组成, 如图 1-1 所示。随着计算机技术的发展, 运算器、控制器等部件已被集成在一起统称为中央处理单元 (Central Processing Unit, CPU)。它是硬件系统的核心, 用于数据的加工处理, 能完成各种算术、逻辑运算及控制功能。存储器是计算机系统记忆设备, 分为内部存储器和外部存储器。前者速度快、容量小, 一般用以临时存放程序、数据及中间结果, 而后者容量大、速度慢, 可以长期

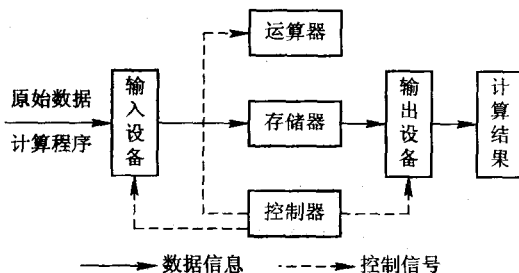


图 1-1

保存程序和数据。输入设备和输出设备合称为外部设备 (简称外设)。输入设备用于输入原始数据及各种命令, 而输出设备则用于输出计算机运行的结果。

2. 计算机硬件的典型结构

(1) 单总线结构。在单总线结构中用一组系统总线将计算机系统的各部件连接起来, 各部件之间可以通过总线交换信息。这种结构的优点是易于扩充新的 I/O 设备, 并且各种 I/O 设备的寄存器和主存储器的存储单元可以统一编址, 使 CPU 访问 I/O 设备更方便灵活; 其缺点是同一时刻只能允许挂在总线上的一对设备之间互相传送信息, 也即分时使用总线, 这就限制了信息传送的吞吐量。这种结构一般用在微型计算机和小型计算机中。

(2) 双总线结构。为了消除信息传送的瓶颈, 常设置多组总线, 最常见的是在主存和 CPU 之间设置一组专用的高速存储总线。双总线结构分为以 CPU 为中心的双总线结构和以存储器为中心的双总线结构两种。将连接 CPU 和外围设备的系统总线称为输入/输出 (I/O) 总线。这种结构的优点是控制线路简单, 对 I/O 总线的传送速率要求较低; 其缺点是 CPU 的工作效率较低, 因为 I/O 设备与主存之间的信息交换要经过 CPU 进行。在以存储器为中心的双总线结构中, 主存储器可通过存储总线与 CPU 交换信息, 同时还可以通过系统总线与 I/O 设备交换信息。这种结构的优点是信息传送速率高; 其缺点是需要增加硬件的投资。

(3) 采用通道的大型系统结构。为了扩大系统的功能和提高系统的效率, 在大、中型计算机系统中采用通道结构。在这种结构中, 一台主机可以连接多个通道, 一个通道可以连接一台或多台 I/O 控制器, 一台 I/O 控制器又可以连接一台或多台 I/O 设备, 所以它具有

有较大的扩展余地。另外,由通道来管理和控制 I/O 设备,减轻了 CPU 的负担,提高了整个系统的效率。

(三) 计算机软件

在计算机系统中如果仅有硬件系统,则它只具备了计算的功能,并不能真正运算,只有将解决问题的步骤编制成程序,并由输入设备输入到计算机内存中,由系统软件支持,才能完成运算。软件是指为管理、运行、维护及应用计算机所开发的程序和相关文档的集合。可见,计算机系统除了硬件系统,还必须有软件系统。软件系统是计算机系统中的重要组成部分,通常可将软件分为两大类:系统软件和应用软件。

1.2.2 计算机基本工作原理

一、要求掌握的知识要点

- (1) 掌握十进制数、二进制数、十六进制数、八进制数以及它们之间的相互转换方法。
- (2) 掌握二进制数的算术运算及逻辑运算的法则,数据在计算机中的表示方法。
- (3) 掌握 BCD 码、ASCII 码及汉字编码的概念。
- (4) 熟悉中央处理单元 CPU 的组成及内部主要部件的功能。

二、知识点概述

(一) 计算机中数据的表示

计算机最主要的功能是处理信息,如处理数值、文字、声音、图形和图像等。在计算机内部,各种信息都必须经过数字化编码后才能被传送、存储和处理,因此,掌握信息编码的概念与处理技术是至关重要的。所谓编码,就是采用少量的基本符号,选用一定的组合原则,以表示大量复杂、多样的信息。基本符号的种类和这些符号的组合规则是一切信息编码的两大要素。例如,用 10 个阿拉伯数码表示数字,用 26 个英文字母表示英文词汇等,都是编码的典型例子。

1. 进位计数制

在采用进位计数的数字系统中,如果只用 r 个基本符号(例如, $0, 1, 2, \dots, r-1$)表示数值,则称其为基 r 数制(Radix- r Number System), r 称为该数制的基(Radix)。对于不同的数制,它们的共同特点是:

- 每一种数制都有固定的符号集。例如,对于十进制数制,其符号有 10 个: $0, 1, 2, \dots, 9$; 对于二进制数制,其符号有两个: 0 和 1 。

- 都使用位置表示法。即处于不同位置的数符所代表的值不同,且与它所在位置的权值有关。例如,十进制数 1234.55 可表示为

$$1234.55 = 1 \times 10^3 + 2 \times 10^2 + 3 \times 10^1 + 4 \times 10^0 + 5 \times 10^{-1} + 5 \times 10^{-2}$$

可以看出,各种进位计数制中的权的值恰好是基数的某次幂。因此,对任何一种进位计数制表示的数都可以写成按权展开的多项式之和,即任意一个 r 进制数 N 可表示为

$$N = \sum_{i=-k}^{+n} D_i \times r^i$$

式中: D_i 是该数制采用的基本数符; r^i 是权; r 是基数,不同的基数表示不同的进制数。表 1-1 所示的是计算机中常用的几种进制数。

表 1-1 计算机中常用的几种进制数的表示

进位制	二进制	八进制	十进制	十六进制
规则	逢二进一	逢八进一	逢十进一	逢十六进一
基数	$r=2$	$r=8$	$r=10$	$r=16$
数符	0, 1	0, 1, 2, ..., 7	0, 1, 2, ..., 9	0, 1, 2, ..., 9, A, B, ..., F
权	2^i	8^i	10^i	16^i
形式表示	B	O	D	H

2. 算术逻辑运算

(1) 二进制加法。二进制加法与十进制加法相类似，所不同的是，二进制加法的规则是“逢二进一”，即

$$0+0=0 \quad 1+0=1 \quad 0+1=1 \quad 1+1=0(\text{有进位})$$

(2) 二进制减法。在二进制减法中，当不够减时需要借位，高位的1等于下一位的2，即“借一当二”，其运算法则如下：

$$0-0=0 \quad 1-0=1 \quad 1-1=0 \quad 0-1=1(\text{有借位})$$

(3) 二进制乘法。二进制乘法与十进制乘法是一样的，但因为二进制数只由0和1构成，因此，二进制乘法更简单，其运算法则如下：

$$0 \times 0=0 \quad 1 \times 0=0 \quad 0 \times 1=0 \quad 1 \times 1=1$$

(4) 二进制除法。二进制除法是二进制乘法的逆运算，其运算方法与十进制除法是一样的。

(5) 二进制与运算又称逻辑乘，其运算法则如下：

$$0 \wedge 0=0 \quad 0 \wedge 1=0 \quad 1 \wedge 0=0 \quad 1 \wedge 1=1$$

(6) 二进制或运算又称逻辑加，其运算法则如下：

$$0 \vee 0=0 \quad 0 \vee 1=1 \quad 1 \vee 0=1 \quad 1 \vee 1=1$$

(7) 二进制异或的运算法则如下：

$$0 \nabla 0=0 \quad 0 \nabla 1=1 \quad 1 \nabla 0=1 \quad 1 \nabla 1=0$$

3. 机器数和码制

各种数据在计算机中表示的形式称为机器数，其特点是数的符号用0、1表示，如“0”表示正号，“1”表示负号，小数点则隐含表示而不占位置。机器数对应的实际数值称为该数的真值。

机器数有无符号数和带符号数两种。无符号数表示正数，在机器数中没有符号位。对于无符号数，若约定小数点的位置在机器数的最低位之后，则是纯整数；若约定小数点的位置在机器数的最高位之前，则是纯小数。对于带符号数，机器数的最高位是表示正、负的符号位，其余二进制位表示数值。若约定小数点的位置在机器数的最低数值位之后，则是纯整数；若约定小数点的位置在机器数的最高数值位之前(符号位之后)，则是纯小数。

为了便于运算，带符号的机器数可采用原码、反码和补码等不同的编码方法，机器数的这些编码方法称为码制。

4. 汉字编码

汉字处理包括汉字的编码输入、汉字的存储和汉字的输出等环节。也就是说计算机处

理汉字，首先必须先将汉字代码化，即对汉字进行编码。

1) 输入码

(1) 数字编码。数字编码就是用数字串代表一个汉字的输入，常用的是国标区位码。国际区位码将国家标准局公布的 6763 个两级汉字分成 94 个区，每个区 94 位，实际上是把汉字表示成二维数组，区位和位码各两位十进制数字，因此，输入一个汉字需要按键四次。例如，“中”字位于第 54 区 48 位，区位码为 5448。

(2) 拼音码。拼音码是以汉语读音为基础的输入方法。由于汉字同音字太多，输入重码率很高，因此，按拼音输入后还必须进行同音字选择，影响了输入速度。

(3) 字形编码。字形编码是以汉字的形状确定的编码。汉字总数虽多，但都是由一笔一划组成，全部汉字的部件和笔划是有限的，因此，把汉字的笔划部件用字母或数字进行编码，按笔划书写的顺序依次输入，就能表示一个汉字。五笔字形、表形码等便是这种编码法。五笔字形编码是最有影响的编码方法。

2) 内部码

汉字内部码(简称汉字内码)是汉字在设备或信息处理系统内部最基本的表达形式，是在设备和信息处理系统内部存储、处理、传输汉字用的代码。在西文计算机中，没有交换码和内码之分。汉字数量多，用一个字节无法区分，采用国家标准局 GB2312-80 中规定的汉字国标码，两个字节存放一个汉字的内码，每个字节的最高位置“1”，作为汉字机内码。由于两个字节各用 7 位，因此可表示 16 384 个可区别的机内码。以汉字“大”为例，国标码为 3473H，两个字节的最高位置“1”，得到的机内码为 B4F3H。

为了统一地表示世界各国的文字，1993 年国际标准化组织公布了“通用多八位编码字符集”的国际标准 ISO/IEC 10646，简称 UCS(Universal Code Set)。UCS 包含了中、日、韩等国的文字，这一标准为包括汉字在内的各种正在使用的文字规定了统一的编码方案。

3) 字形码

汉字字形码是表示汉字字形的字模数据，通常用点阵、矢量函数等方式表示。用点阵表示字形时，汉字字形码指的就是这个汉字字形点阵的代码。字形码也称字模码，是用点阵表示的汉字字形码，它是汉字的输出方式。根据输出汉字的要求不同，点阵的多少也不同。简易型汉字为 16×16 点阵，高精度型汉字为 24×24 点阵、32×32 点阵、48×48 点阵等等。

(二) 中央处理机(CPU)

1. CPU 的组成

前面已经提到，CPU 主要由运算器、控制器组成。构成 CPU 的框图如图 1-2 所示。

1) 运算器

运算器是对数据进行加工处理的部件，它主要完成算术运算和逻辑运算，完成对数据的加工与处理。不同的计算机，运算器的结构也不同，但最基本的结构都是由算术/逻辑运算单元(ALU)、累加器(ACC)、寄存器组、多路转换器和数据总线等逻辑部件组成的。

2) 控制器

计算机能执行的基本操作叫做指令，一台计算机的所有指令组成指令系统。指令由操作码和地址码两部分组成，操作码指明操作的类型，地址码则指明操作数及运算结果存放的地址。

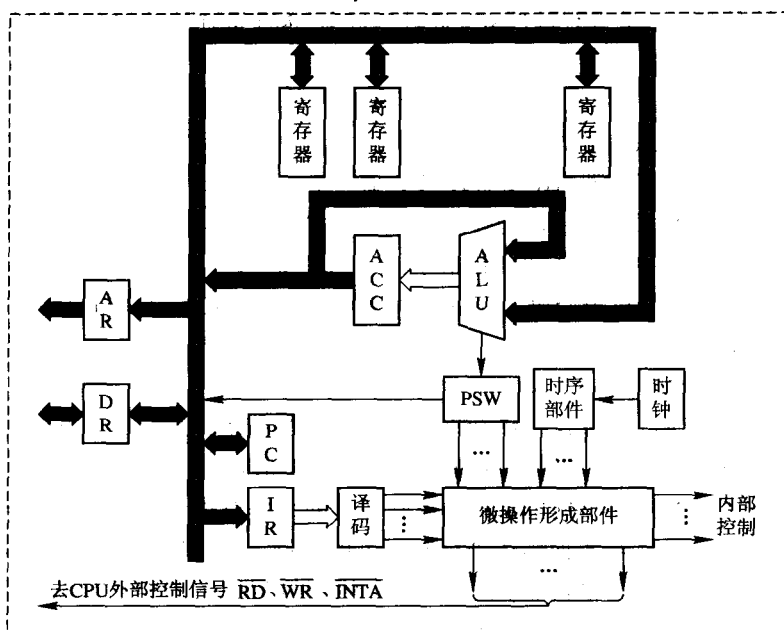


图 1-2 CPU 主要组成部件框图

控制器的主要功能是从内存中取出指令，并指出下一条指令在内存中的位置，将取出指令经指令寄存器送往指令译码器，经过对指令的分析发出相应的控制和定时信息，控制和协调计算机的各个部件有条不紊的工作，以完成指令所规定的操作。

控制器是由程序计数器(简称 PC)、指令寄存器、指令译码器、状态条件寄存器、时序产生器、微操作信号发生器组成，如图 1-3 所示。

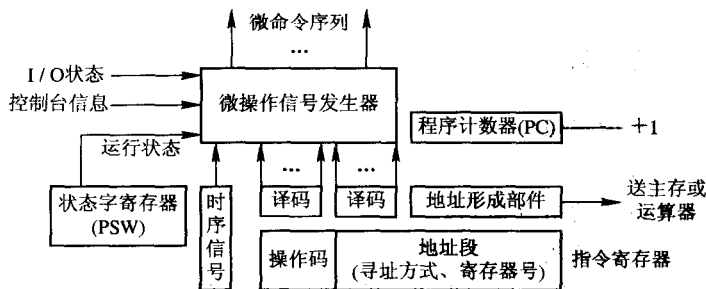


图 1-3 控制器组成框图

(1) 程序计数器。当程序顺序执行时，每取出一条指令，PC 内容自动增加一个值，指向下一条要取的指令。当程序出现转移时，则将转移地址送入 PC，然后由 PC 指向新的程序地址。

(2) 指令寄存器(IR)。用于存放当前要执行的指令。

(3) 指令译码器(ID)。用于对现行指令进行分析，确定指令类型、指令所要完成的操作以及寻址方式。

(4) 时序产生器。用于产生时序脉冲和节拍电位去控制计算机有序的工作。

(5) 状态/条件寄存器。用于保存指令执行完成后产生的条件码。例如，运算是否有溢

出, 结果为正还是为负, 是否有进位等。此外, 状态/条件寄存器还保存中断和系统工作状态等信息。

(6) 微操作信号发生器。把指令提供的操作信号、时序产生器提供的时序信号以及由控制功能部件反馈的状态信号等综合成特定的操作序列, 从而完成取指令的执行控制。

控制器一般由指令寄存器(IR)、程序计数器(PC)、时序部件、微操作形成部件和程序状态字寄存器(PSW)构成。控制器的作用是控制整个计算机的各个部件有条不紊地工作, 它的基本功能就是从内存取指令和执行指令。

执行指令有取指令、指令译码、按指令操作码执行、形成下一条指令地址四个步骤。

2. CPU 的功能

CPU 的基本功能如下。

(1) 程序控制。CPU 通过执行指令来控制程序的执行顺序, 这是 CPU 的重要职能。

(2) 操作控制。一条指令功能的实现需要若干操作信号来完成, CPU 产生每条指令的操作信号并将操作信号送往不同的部件, 控制相应的部件按指令的功能要求进行操作。

(3) 时间控制。CPU 对各种操作进行时间上的控制, 这就是时间控制。CPU 对每条指令整个的执行时间要进行严格控制。同时, 指令执行过程中的操作信号的出现时间、持续时间及出现的时间顺序都需进行严格控制。

(4) 数据处理。CPU 对数据以算术运算及逻辑运算等方式进行加工处理, 数据加工处理的结果为人们所利用。所以, 对数据的加工处理是 CPU 最根本的任务。

三、重点与难点分析

(1) 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, 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。

(2) 计算机组成是指计算机系统结构的逻辑实现, 计算机实现是指计算机组成的物理实现。

1.2.3 计算机体系结构

一、要求掌握的知识要点

(1) 计算机系统分类。

• 指令流: 机器执行的指令系列。

• 数据流: 由指令流调用的数据序列。

• Flynn 分类法: 单指令流单数据流, 指令流多数据流, 单指令流多数据流, 多指令流单数据流。

• 冯式分类法: 用最大并行度来对计算机体系结构进行分类。

(2) 计算机系统结构与计算机组成的区别, 计算机并行的发展。

计算机组成指计算机系统的逻辑结构, 包括机器内部数据流和控制流的组成以及逻辑设计等; 计算机实现是指计算机组成的物理实现。

(3) 存储器系统的层次结构。存储器系统由分布在计算机各个不同部件中的多种存储设备组成。主存储器由一片或多片存储芯片配以控制电路组成。辅助存储器包括磁表面存储器、光存储器。高速缓冲存储器(Cache 存储器)用于提高 CPU 数据输入/输出速度, 突破 CPU 与存储器间的数据传输带宽限制。在存储系统体系中, Cache 是访问速度最快的

层次。

二、知识点概述

(一) 计算机体系结构的发展

1. 计算机系统结构概述

计算机系统结构又称为计算机体系结构，就是计算机的属性及功能特征，即计算机的外特性。尽管不同的使用者所了解的计算机的属性有所不同，就通用计算机系统来说，计算机系统结构的属性应包括如下一些方面：

- 硬件所能处理的数据类型。
- 所能支持的寻址方式。
- CPU 的内部寄存器。
- CPU 的指令系统。
- 主存的组织与主存的管理。
- 中断系统的功能。
- 输入/输出设备及连接接口。
- 计算机体系结构分类。

1) Flynn 分类法

1966 年 Flynn 提出了如下定义。

指令流 (Instruction Stream)：机器执行的指令序列。

数据流 (Data Stream)：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 (Multiplicity)：在系统最受限制的元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。

按指令流和数据流的不同组织方式，把计算机体系结构分为如下四类：单指令流单数据流 (SISD)，单指令流多数据流 (SIMD)，多指令流单数据流 (MISD)，多指令流多数据流 (MIMD)。

2) 冯式分类法

1972 年冯泽云提出用最大并行度来对计算机体系结构进行分类。所谓最大并行度 P_m 是指计算机系统在单位时间内能够处理的最大的二进制位数。设每一个时钟周期 Δt_i 内能处理的二进制位数为 P_i ，则 T 个时钟周期内平均并行度为 $P_a = (\sum P_i) / T$ (其中 i 为 1, 2, ..., T)。平均并行度取决于系统的运行程度，与应用程序无关，所以，系统在周期 T 内的平均利用率为 $\mu = P_a / P_m = (\sum P_i) / (T * P_m)$ 。

图 1-4 所示为用最大并行度对计算机体系结构进行的分类。用平面直角坐标系中的一点表示一个计算机系统，横坐标表示字宽 (N 位)，即在一个字中同时处理的二进制位数；纵坐标表示位片宽度 (M 位)，即在一个位片中能同时处理的字数，则最大并行度 $P_m = N * M$ 。由此得出四种不同的计算机结构：

- ① 字串行、位串行 (简称 WSBS)。其中 $N=1, M=1$ 。
- ② 字并行、位串行 (简称 WPBS)。其中 $N=1, M>1$ 。
- ③ 字串行、位并行 (简称 WSBP)。其中 $N>1, M=1$ 。
- ④ 字并行、位并行 (简称 WPBP)。其中 $N>1, M>1$ 。

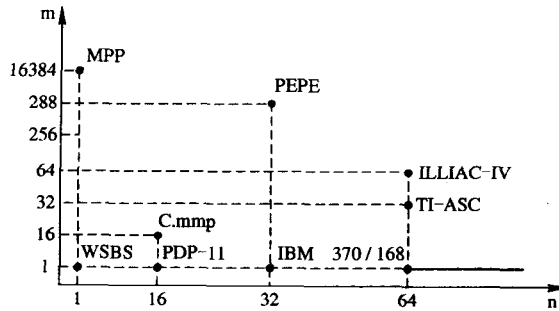


图 1-4 冯式分类法

2. 计算机系统结构与计算机组成的区别

计算机系统结构所解决的问题是计算机系统总体上、功能上需要解决的问题，而计算机组成要解决的是逻辑上如何具体实现的问题。

比如说，指令系统的确定属于计算机系统结构，而指令的具体实现则属于计算机组成。指令系统中要不要设置乘、除法指令是计算机系统结构要解决的问题，而一旦决定设置，具体用什么方法来实现乘、除法指令就属于计算机组成应解决的问题。

主存容量及编址方式的确定属于计算机系统结构，而具体构成主存则属于计算机组成。

可以想像，对于具有同样系统结构的计算机，但具体按此系统结构构成的计算机在实现方法、性能及价格上会有很大差别。

3. 系统结构中并行性的发展

1) 并行性

并行性包括两个方面：同时性和并发性。同时性是指两个或两个以上的事件在同一时刻发生，并发性是指两个或两个以上的事件在同一时间间隔内连续发生。

充分利用并行性实现计算机的并行处理，可以提高计算机的处理速度。

2) 并行处理

从计算机信息处理的步骤和阶段的角度，并行处理可分为：

- 存储器操作并行。
- 处理器操作步骤并行(流水线处理机)。
- 处理器操作并行(阵列处理机)。
- 指令、任务、作业并行(多处理机、分布处理系统、计算机网络)。

3) 并行性的发展

从 20 世纪 80 年代开始，计算机系统结构有了很大发展，相继出现了精减指令集计算机(RISC)、指令级上并行的超标量处理机、超级流水线处理机、超长指令计算机、多微处理机系统、数据流计算机等。20 世纪 90 年代以来，最主要的发展是大规模并行处理(MPP)，其中多处理机系统和多计算机系统是研究开发的热点。

(二) 存储系统

1. 存储器的层次结构

存储体系结构包括不同层次上的存储器，通过适当的硬件、软件有机地组合在一起形成计算机的存储体系结构。现在大多数人都将高性能计算机的存储体系结构描述成如图 1-5 所示的三层存储器层次结构。