

计算机组成原理与系统结构

第三章 信息编码与数据表示

<http://jpkc.hdu.edu.cn/computer/zcyl/dzkjdx/>





第三章 信息编码与数据表示

3.

数值数据的表示

3.

数据格式

3.3

定点机器数的表示

方法

3.4

浮点机器数的表示

方法

3.

非数值数据的表示

3.

校验码

3.7

现代计算机系统的数据表示

本章小结

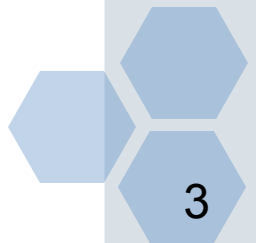
BACK



3.7 现代计算机系统的数据表示

❖ 以 Pentium 系列的微处理器为例：

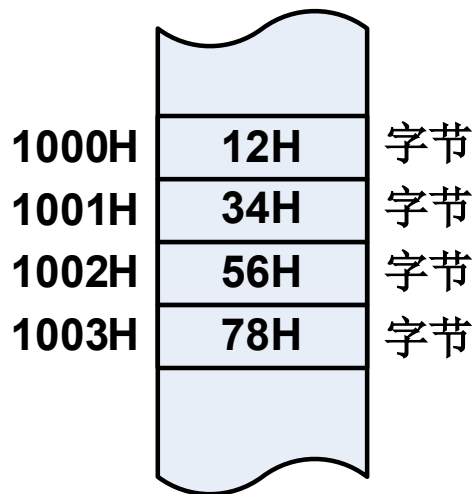
- 编址方式：X86CPU 是按字节编址的，即每 8 位就有一个编码地址。
- 数据存取的尺寸：字节数据、字数据、双字数据、四字数据等。
- 数据在内存中的存放位置：低字节（LSB）存放在低地址，高字节（MSB）存放在高地址。



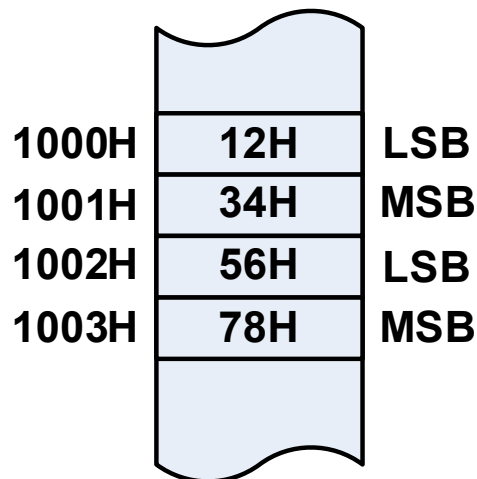


3.7 现代计算机系统的数据表示

❖ 字节数据、字数据、双字数据在内存中的存储

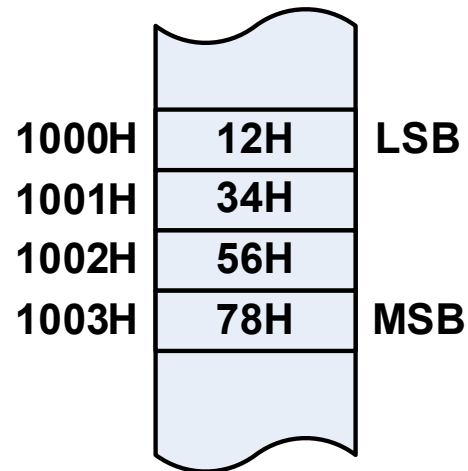


字节数据



字数据

MOV AX, [100
0H]
AX= 3412H



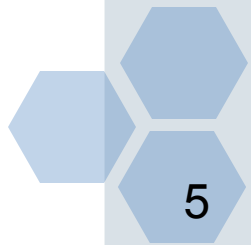
双字数据

MOV EAX, [10
00H]
EAX= 78563412
H



3.7 现代计算机系统的数据表示

- ❖ 在 Pentium 系列 CPU 中，无论是字节数据、字数据还是双字、四字数据，它们都可以是无符号数、带符号数、或者 BCD。
- ❖ 那么，如何区分内存某个单元存放的数据尺寸？如何分辨它们的数据类型？仍旧取决于访问它的指令的操作数尺寸和指令本身的功能。
- ❖ 当数据为有符号数时，默认它是补码表示的机器数。例如，假如某个有符号字节数据的二进制代码是 90H，则它的值为 -70H。





3.7 现代计算机系统的数据表示

❖ 几种类型的数据在 X86 系列 CPU 中的表示形式：

1、字符串

- 由字符的 ASCII 码（1B）或者文字的 Unicode 编码（2B）组成，按顺序存放在内存或寄存器中。
- Windows 下以 0 做结束符

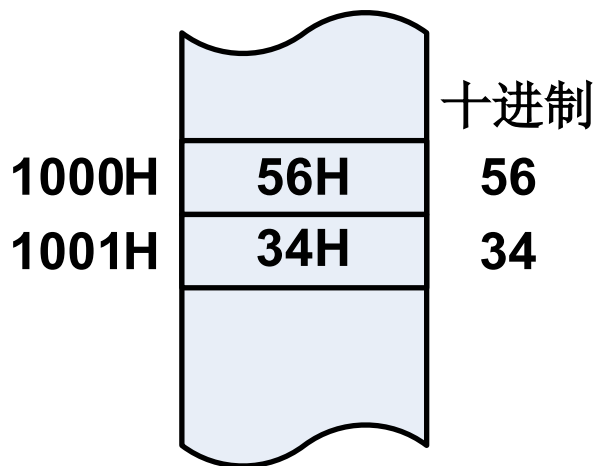
1000H	D0H	“中”
1001H	D6H	
1002H	4FH	“O”
1003H	4BH	“K”
1004H	21H	“!”



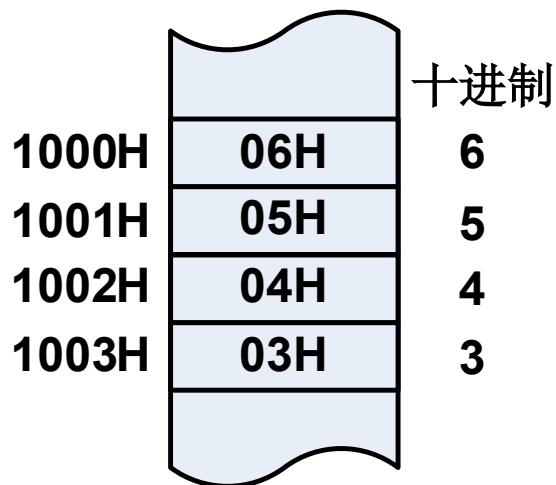
3.7 现代计算机系统的数据表示

2、BCD

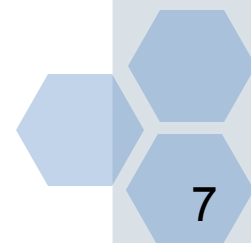
- BCD 数据分为压缩的 (packed) BCD 码 (4 位) 和非压缩 (unpacked) 的 BCD 码 (8 位) 两种。



(a) 压缩 BCD 码



(b) 非压缩 BCD 码





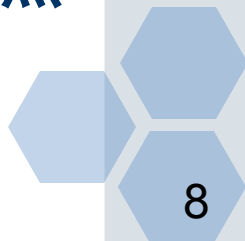
3.7 现代计算机系统的数据表示

3、指针

- 指针实际上是内存单元的地址，因此是无符号数据。IA 构架中定义了两种类型的指针：
 - **近指针**：32 位，用于定义段内偏移量和段内访问；
 - **全指针**：又称远指针，48 位，用于段间访问。

4、浮点数

- Pentium 系列 CPU 支持 IEEE 754 标准的 3 种浮点数格式：单精度、双精度和扩展精度浮点数。





本章小结

- ❖ **数制有两个要素**：基数 R 与位权 W 。计算机中的信息均由二进制来表示，即 $R=2$ ， $W=2^i$ 。用于表示十进制数值的二进制编码被称为 BCD 码，4 位二进制编码表示一个十进制数字。
- ❖ 机器数是数值数据在机器中的表示形式，根据小数点的位置是否浮动，可以分为**定点数和浮点数**。
 - 定点机器数根据小数点的隐含位置又分为**定点小数**和**定点整数**两种。
 - 定点机器数有四种表示形式：**原码、反码、补码和移码**。移码主要用于表示浮点数的阶码。



本章小结

- **浮点机器数由阶码 E 和尾数 M 两部分构成**，阶码是定点整数，尾数是定点小数；阶码 E（即指数）的底，一般隐含为 2。浮点机器数的小数点的位置随阶码数值而变化。
- **IEEE754 标准的浮点数**有单精度、双精度、临时浮点数 3 种格式，分别为 32 位、64 位和 80 位。
- 浮点数的规格化表示方法。



本章小结

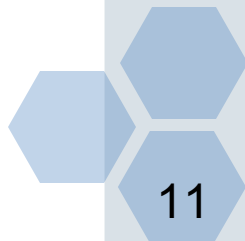
❖ 计算机中的非数值数据的表示：

- 字符数据通常采用 **7 位的 ASCII 码** 来表示。
- **汉字的输入编码** 用于使用西文标准键盘输入汉字，**汉字的机内码** 则用于汉字的存储、检索和处理，**汉字的字模码** 则用于汉字的显示和打印输出。

❖ 校验码：用于检错和纠错。

- **奇偶校验码** 是最简单的一种检错码，它可以检查出一位或奇数位错误。海明校验码是一种多重奇偶校验码，具有纠错能力，而 CRC 校验码则是一种目前广泛使用的纠错码，可以纠错一位。

❖ 本章重点为定点机器数和浮点机器数的表示方法。





作业

✦ P101 : 1 , 2 , 3 , 4 , 5 , 8 , 9
(1) ~ (4) , 14, 17



The End !