

内存编址和关于按字寻址和按字节寻址的理解

我们先从一道简单的问题说起！

设有一个 **1MB** 容量的存储器，字长 **32** 位，问：按字节编址，字编址的寻址范围以及各自的寻址范围大小？

如果按字节编址，则

$$1\text{MB} = 2^{20}\text{B}$$

$$1 \text{ 字节} = 1\text{B} = 8\text{bit}$$

$$2^{20}\text{B}/1\text{B} = 2^{20}$$

地址范围为 $0 \sim (2^{20}) - 1$ ，也就是说需要二十根地址线才能完成对 **1MB** 空间的编码，所以地址寄存器为 **20** 位，寻址范围大小为 $2^{20} = 1\text{M}$

如果按字编址，则

$$1\text{MB} = 2^{20}\text{B}$$

$$1 \text{ 字} = 32\text{bit} = 4\text{B}$$

$$2^{20}\text{B}/4\text{B} = 2^{18}$$

地址范围为 $0 \sim 2^{18} - 1$ ，也就是说我们至少要用 **18** 根地址线才能完成对 **1MB** 空间的编码。因此按字编址的寻址范围是 2^{18}

以上题目注意几点：

1.区分寻址空间与寻址范围两个不同的概念，寻址范围仅仅是一个数字范围，不带有单位而寻址范围的大小很明显是一个数，指寻址区间的大小

而寻址空间指能够寻址最大容量，单位一般用 **MB**、**B** 来表示；本题中寻址范围为 $0 \sim (2^{20}) - 1$ ，寻址空间为 **1MB**。

2.按字节寻址，指的是存储空间的最小编址单位是字节，按字编址，是指存储空间的最小编址单位是字，以上题为例，总的存储器容量是一定的，按字编址和按字节编址所需要的编码数量是不同的，按字编址由于编址单位比较大（**1 字 = 32bit = 4B**），从而编码较少，而按字节编址由于编码单位较小（**1 字节 = 1B = 8bit**），从而编码较多。

3.区别 **M** 和 **MB**。

M 为数量单位。**1024 = 1K, 1024K = 1M**

MB 指容量大小。**1024B = 1KB, 1024KB = 1MB.**

2. 内存编址

前面我们知道了 DRAM 颗粒以及内存模块是如何扩展字长和容量的。一个内存可能是 8 位，也可能是 64 位，容量可能是 1M，也可能是 1G。那么内存是如何编地的呢？和地址总线，计算机字长之间又有什么关系呢？

2.1 字长

计算机在同一时间内处理的一组二进制数称为一个计算机的“字”，而这组二进制数的位数就是“字长”。通常称处理字长为 8 位数据的 CPU 叫 8 位 CPU，32 位 CPU 就是在同一时间内处理字长为 32 位的二进制数据。所以这里的字并不是我们理解的双字节（Word）而是和硬件相关的一个概念。一般来说计算机的数据线的位数和字长是相同的。这样从内存获取数据后，只需要一次就能把数据全部传送给 CPU。

2.2 地址总线

前面我们已经介绍过地址总线的功能。地址总线的数量决定了他最大的寻址范围。就目前来说一般地址总线先字长相同。比如 32 位计算机拥有 32 为数据线和 32 为地线，最大寻址范围是 4G（0x00000000 ~ 0xFFFFFFFF）。当然也有例外，Intel 的 8086 是 16 为字长的 CPU，采用了 16 位数据线和 20 位数据地线。

2.3 内存编址

从前面我们知道一个内存的大小和它芯片扩展方式有关。比如我们内存模块是采用 16M*8bit 的内存颗粒，那么我们使用 4 个颗粒进行位扩展，成为 16M*32bit，使用 4 个颗粒进行字容量扩展变为 64M*32bit。那么我们内存模块使用了 16 个内存颗粒，实际大小是 256MB。

我们需要对这个 256M 的内存进行编址以便 CPU 能够使用它，通常我们多种编址方式：

1. 按字编址：对于这个 256M 内存来说，它的寻址范围是 64M，而每个内存地址可以存储 32bit 数据。
2. 按半字编址：对于这个 256M 内存来说，它的寻址范围是 128M，而每个内存地址可以存储 16bit 数据。
3. 按字节编址：对于这个 256M 内存来说，它的寻址范围是 256M，而每个内存地址可以存储 8bit 数据。

对于我们现在的计算机来说，**主要都是采用按字节编址的方式**。所以我们可以把内存简单的看成一个线性数组，数组每个元素的大小为 8bit，我们称为一个存储单元。这一点很重要，因为后面讨论的所有问题内存都是以按字节编址的方式。这也是为什么对于 32 位计算机来说，能使用的最多容量的内存为 4GB。如果我们按字编地址，能使用的最大内存容量就是 16GB 了。

于是很容易想到一个问题，为什么我们要采用字节编址的方式呢？关于这个问题，我在网上基本没有找到答案，甚至都找不到问这个问题的。所以这里没法给出答案，为什么呢？麻烦知道的朋友告诉我哈。

另一方面的问题是，内存编址方式和 DRAM 芯片是否有关呢？我认为还是有一定关系。比如我 DRAM 的芯片是 8M*8bit，那么芯片最小的存储单位就是 8bit，那么我们内存编址就不能按照半个字节来编址。否则内存取出 8bit，根本不知道你要那 4bit 传给 CPU。也有一种说法是现在的 DRAM 芯片 cell 都是 8bit，所以采用按字节编址。另一方面应该也和数据总线位宽有关。