

# 高手进阶，终极内存技术指南——完整/进阶版

赵效民

作为电脑中必不可少的三大件之一（其余的两个是主板与 CPU），内存是决定系统性能的关键设备之一，它就像一个临时的仓库，负责数据的中转、暂存……

不过，虽然内存对系统性能的至关重要，但长期以来，DIYer 并不重视内存，只是将它看作是一种买主板和 CPU 时顺带买的“附件”，那时最多也就注意一下内存的速度。这种现象截止于 1998 年 440BX 主板上市后，PC66/100 的内存标准开始进入普通 DIYer 的视野，因为这与选购有着直接的联系。一时间，有关内存时序参数的介绍文章大量出现（其中最为著名的恐怕就是 CL 参数）。自那以后，DIYer 才发现，原来内存也有这么多的学问。接下来，始于 2000 年底/2001 年初的 VIA 芯片组 4 路交错(4-Way Interleave) 内存控制和部分芯片组有关内存容量限制的研究，则是深入了解内存的一个新开端。本刊在 2001 年第 2 期上也进行了 VIA 内存交错控制与内存与模组结构的详细介绍，并最终率先正确地解释了这一类型交错（内存交错有多种类型）的原理与容量限制的原因。从那时起，很多关于内存方面的深入性文章接踵而至，如果说那时因此而掀起了一股内存热并不夸张。大量的内存文章让更多的用户了解了内存，以及更深一层的知识，这对于 DIY 当然是一件好事情。然而，令人遗憾的是这些所谓的内存高深技术文章有不少都是错的（包括后来的 DDR 与 RDRAM 内存的介绍），有的甚至是很低级的错误。在这近两年的时间里，国内媒体上优秀的内存技术文章可谓是寥若晨星，有些媒体还编译国外 DIY 网站的大篇内存文章，但可惜的是，外国网站也不见得都是对的（这一点，似乎国内很多作者与媒体似乎都忽视了）。就这样，虽然打开了一个新的知识领域，可“普及”的效果并不那么好，很多媒体的铁杆读者高兴地被带入内存深层世界，但也因此被引向了新的误区。

不过，从这期间（2001 年初至今）各媒体读者对这类文章的反映来看，喜欢内存技术的玩家大有人在且越来越多，这是各媒体“培养”的成果。这些用户已经不能满足如何正确的使用内存，他们更渴望深入的了解这方面原来非常贫乏的知识，这些知识可能暂时不会对他们在内存使用过程中有什么帮助，但会大大满足他们的求知欲。在 2001 年初，我们揭开 VIA 芯片组 4 路交错内存控制和部分芯片组有关内存容量限制之迷时，还是主要围绕着内存使用的相关话题来展开，而且在这期间有关内存技术的话题，《电脑高手》也都是一笔带过。但在今天，在很多人希望了解内存技术而众多媒体的文章又“力不从心”时，我们觉得有必要再次站出来以正视听，也就是说，我们这次的专题不再以内存使用为中心，更多的是纯技术性介绍，并对目前现存的主要内存技术误区进行重点纠正。

在最后要强调的是，本专题以技术为主，由于篇幅的原因，不可能从太浅的方面入手，所以仍需要有一定的技术基础作保证，而对内存感兴趣的读者则绝不容错过，这也许是您最好的纠正错误认识的机会！

在本专题里，当讲完内存的基本操作之后，我们会给大家讲一个仓库的故事，从中相信您会更了解内存这个仓库是怎么工作的，希望您能喜欢。

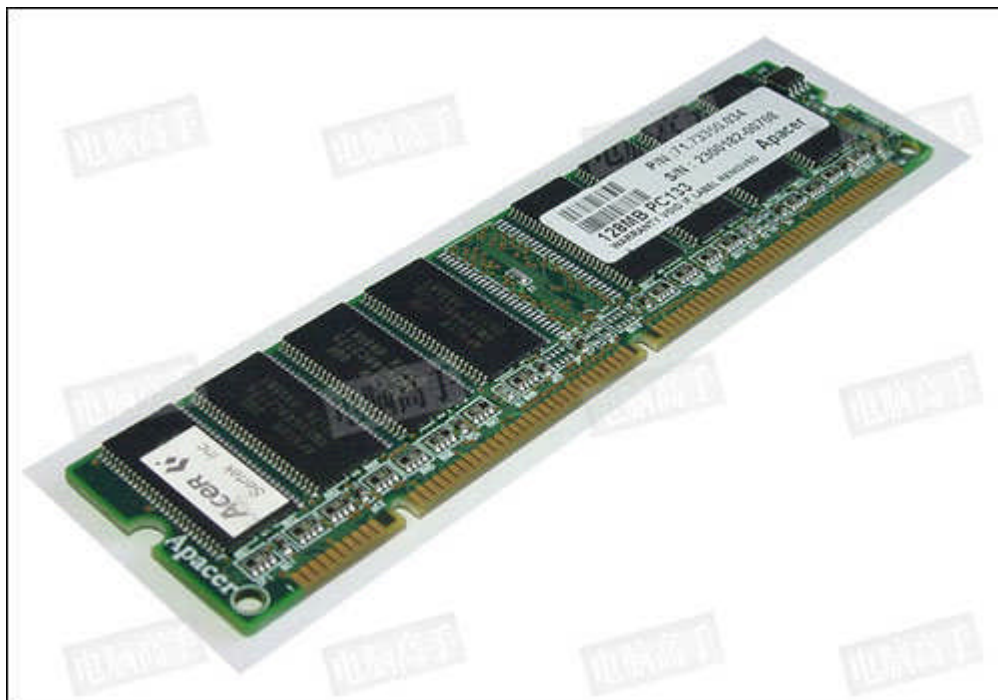
## SDRAM 与内存基础概念（一）

虽然有关内存结构与时序的基础概念，在本刊 2001 年第 2 期的专题中就已有阐述，但在这里为了保证专题的可读性，我们需要再次加强这方面的系统认识。正确并深刻理解内存的基础概念，是阅读本专题的第一条件。因为即使是 RDRAM，在很多方面也是与 SDRAM 相似的，而至于 DDR 与 DDR-II、QBM 等形式的内存更是与 SDRAM 有着紧密的联系。

### 一、SDRAM 内存模组与基本结构

我们平时看到的 SDRAM 都是以模组形式出现，为什么要做成这种形式呢？这首先要接触到两个概念：

物理 Bank 与芯片位宽。



### 1、物理 Bank

传统内存系统为了保证 CPU 的正常工作，必须一次传输完 CPU 在一个传输周期内所需要的数据。而 CPU 在一个传输周期能接受的数据容量就是 CPU 数据总线的位宽，单位是 bit（位）。当时控制内存与 CPU 之间数据交换的北桥芯片也因此将内存总线的数据位宽等同于 CPU 数据总线的位宽，而这个位宽就称之为物理 Bank（Physical Bank，下文简称 P-Bank）的位宽。所以，那时的内存必须要组织成 P-Bank 来与 CPU 打交道。资格稍老的玩家应该还记得 Pentium 刚上市时，需要两条 72pin 的 SIMM 才能启动，因为一条 72pin-SIMM 只能提供 32bit 的位宽，不能满足 Pentium 的 64bit 数据总线的需要。直到 168pin-SDRAM DIMM 上市后，才可以使用一条内存开机。下面将通过芯片位宽的讲述来进一步解释 P-Bank 的概念。

不过要强调一点，P-Bank 是 SDRAM 及以前传统内存家族的特有概念，在 RDRAM 中将以通道（Channel）取代，而对于像 Intel E7500 那样的并发式多通道 DDR 系统，传统的 P-Bank 概念也不适用。

#### 提示：SDRAM、SIMM、DIMM、pin 的含义

SDRAM: Synchronous Dynamic Random Access Memory, 同步动态随机存储器。同步是指其时钟频率与 CPU 前端总线的系统时钟频率相同，并且内部的命令的发送与数据的传输都以它为基础；动态是指存储阵列需要不断的刷新来保证数据不丢失；随机是指数据不是线性依次存储，而是自由指定地址进行数据的读写。

pin: 模组或芯片与外部电路联接用的金属引脚，而模组的 pin 就是常说的“金手指”。

SIMM: Single In-line Memory Module, 单列内存模组。内存模组就是我们常说的内存条，所谓单列是指模组电路板与主板插槽的接口只有一列引脚（虽然两侧都有金手指）。

DIMM: Double In-line Memory Module, 双列内存模组。所谓双列是指模组电路板与主板插槽的接口有两列引脚，模组电路板两侧的金手指对应一列引脚。

## 2、 芯片位宽

上文已经讲到 SDRAM 内存系统必须要组成一个 P-Bank 的位宽，才能使 CPU 正常工作，那么这个 P-Bank 位宽怎么得到呢？这就涉及到了内存芯片的结构。

每个内存芯片也有自己的位宽，即每个传输周期能提供的数据量。理论上，完全可以做出一个位宽为 64bit 的芯片来满足 P-Bank 的需要，但这对技术的要求很高，在成本和实用性方面也都处于劣势。所以芯片的位宽一般都较小。台式机市场所用的 SDRAM 芯片位宽最高也就是 16bit，常见的则是 8bit。这样，为了组成 P-Bank 所需的位宽，就需要多颗芯片并联工作。对于 16bit 芯片，需要 4 颗（ $4 \times 16\text{bit} = 64\text{bit}$ ）。对于 8bit 芯片，则就需要 8 颗了。

### 提示：内存芯片与颗粒

很多时候，经常听人们说到“内存颗粒”，其实这是港台地区对内存芯片的一种称呼（仅对内存，其他的芯片，港台则称为“晶片”），两者的意思是一样的。具体怎么说，就看个人喜好了，就笔者而言，而倾向于用“内存芯片”来表述。

以上就是芯片位宽、芯片数量与 P-Bank 的关系。P-Bank 其实就是一组内存芯片的集合，这个集合的容量不限，但这个集合的总位宽必须与 CPU 数据位宽相符。随着计算机应用的发展，一个系统只有一个 P-Bank 已经不能满足容量的需要。所以，芯片组开始可以支持多个 P-Bank，一次选择一个 P-Bank 工作，这就有了芯片组支持多少（物理）Bank 的说法。而在 Intel 的定义中，则称 P-Bank 为行（Row），比如 845G 芯片组支持 4 个行，也就是说它支持 4 个 P-Bank。另外，在一些文档中，也把 P-Bank 称为 Rank（列）。

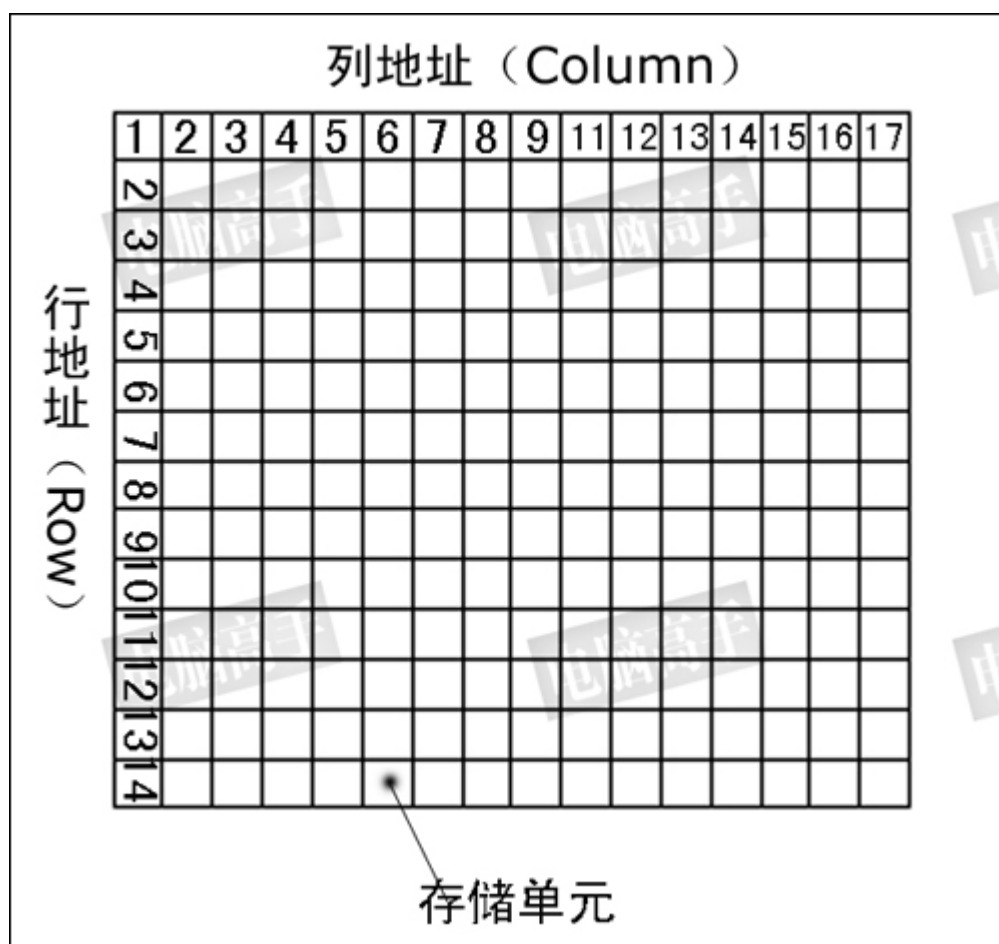
回到开头的话题，DIMM 是 SDRAM 集合形式的最终体现，每个 DIMM 至少包含一个 P-Bank 的芯片集合。在目前的 DIMM 标准中，每个模组最多可以包含两个 P-Bank 的内存芯片集合，虽然理论上完全可以在一个 DIMM 上支持多个 P-Bank，比如 SDRAM DIMM 就有 4 个芯片选择信号（Chip Select，简称片选或 CS），理论上可以控制 4 个 P-Bank 的芯片集合。只是由于某种原因而没有这么去做。比如设计难度、制造成本、芯片组的配合等。至于 DIMM 的面数与 P-Bank 数量的关系，在 2001 年 2 月的专题中已经明确了，面数  $\neq$  P-Bank 数，只有在知道芯片位宽的情况下，才能确定 P-Bank 的数量，大度 256MB 内存就是明显一例，而这种情况在 Registered 模组中非常普遍。有关内存模组的设计，将在后面的相关章节中继续探讨。

## 二、 SDRAM 内存芯片的内部结构

### 1、逻辑 Bank 与芯片位宽

讲完 SDRAM 的外在形式，就该深入了解 SDRAM 的内部结构了。这里主要的概念就是逻辑 Bank。简单地讲，SDRAM 的内部是一个存储阵列。因为如果是管道式存储（就如排队买票），就很难做到随机访问了。

阵列就如同表格一样，将数据“填”进去，你可以想象成一张表格。和表格的检索原理一样，先指定一个行（Row），再指定一个列（Column），我们就可以准确地找到所需要的单元格，这就是内存芯片寻址的基本原理。对于内存，这个单元格可称为存储单元，那么这个表格（存储阵列）叫什么呢？它就是逻辑 Bank（Logical Bank，下文简称 L-Bank）。



L-Bank 存储阵列示意图

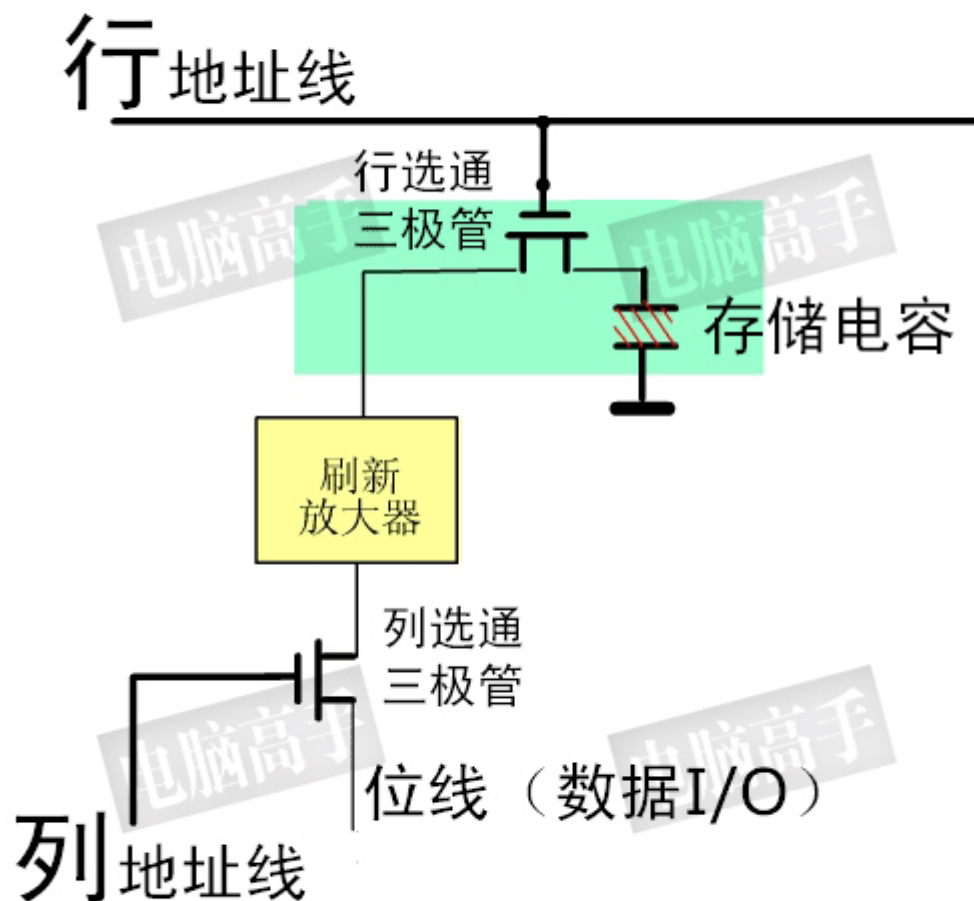
由于技术、成本等原因，不可能只做一个全容量的 L-Bank，而且最重要的是，由于 SDRAM 的工作原理限制，单一的 L-Bank 将会造成非常严重的寻址冲突，大幅降低内存效率（在后文中将详细讲述）。所以人们在 SDRAM 内部分割成多个 L-Bank，较早以前是两个，目前基本都是 4 个，这也是 SDRAM 规范中的最高 L-Bank 数量。到了 RDRAM 则最多达到了 32 个，在最新 DDR-II 的标准中，L-Bank 的数量也提高到了 8 个。

这样，在进行寻址时就要先确定是哪个 L-Bank，然后再在这个选定的 L-Bank 中选择相应的行与列进行寻址。可见对内存的访问，一次只能是一个 L-Bank 工作，而每次与北桥交换的数据就是 L-Bank 存储阵列中一个“存储单元”的容量。在某些厂商的表述中，将 L-Bank 中的存储单元称为 Word（此处代表位的集合而不是字节的集合）。

从前文可知，SDRAM 内存芯片一次传输率的数据量就是芯片位宽，那么这个存储单元的容量就是芯片的位宽（也是 L-Bank 的位宽），但要注意，这种关系也仅对 SDRAM 有效，原因将在下文中说明。

提示：DRAM 的存储原理

L-Bank 中的存储单元是基本的存储单位，它的容量是若干 Bit(对于 SDRAM 而言，就是芯片的位宽)，而每个 Bit 则是存放于一个单独的存储体中。这些存储体就是内存中最小的存储单元。你可以用硬盘操作中的簇与扇区的关系来理解内存中的存储形式。扇区是硬盘中的最小存储单元（相当于内存中的存储体），而每个簇则包含有多个扇区（相当于 L-Bank 中的存储单元），数据的交换都是一个簇为单位进行（一次传输一个存储单元的数据）。



DRAM 的存储原理示意图：行选与列选信号将使存储电容与外界间的传输电路导通，从而可进行放电（读取）与充电（写入）。另外，图中刷新放大器的设计并不固定，目前这一功能被并入读出放大器（Sense Amplifier，简称 S-AMP），具体操作在下文中详细讲述。

## 2、内存芯片的容量

现在我们应该清楚内存芯片的基本组织结构了。那么内存的容量怎么计算呢？显然，内存芯片的容量就是所有 L-Bank 中的存储单元的容量总合。计算有多少个存储单元和计算表格中的单元数量的方法一样：

存储单元数量=行数×列数（得到一个 L-Bank 的存储单元数量）×L-Bank 的数量

在很多内存产品介绍文档中，都会用 M×W 的方式来表示芯片的容量（或者说是芯片的规格/组织结

构)。M 是该芯片中存储单元的总数，单位是兆（英文简写 M，精确值是 1048576，而不是 1000000），W 代表每个存储单元的容量，也就是 SDRAM 芯片的位宽（Width），单位是 bit。计算出来的芯片容量也是以 bit 为单位，但用户可以采用除以 8 的方法换算为字节（Byte）。比如 8M×8，这是一个 8bit 位宽芯片，有 8M 个存储单元，总容量是 64Mbit（8MB）。

提示：bit、Byte、Word 的关系
bit：位。二进制数中，一个 0 或 1 就是一个 bit。
Byte：字节。8 个 bit 为一个字节，这与 ASCII（American Standard Code for Information Interchange，美国标准信息交换代码）的规定有关，ASCII 用 8 位二进制数来表示 256 个信息代码，所以 8 个 bit 定义为一个字节。
Word：字：两个字节为一个字，这里的 Word 不是指 L-Bank 中存储单元，此外还有双字（DWords，Double Words，4 个字节）和四字（QWord，Quad Words，8 个字节）的表示法。目前一个 P-Bank 的位宽就是 QWord，这在很多 CPU 与芯片组的介绍中经常用到。

不过，M×W 是最简单的表示方法。下图则是某公司对自己内存芯片的容量表示方法，这可以说是最正规的形式之一。

2,097,152-WORDS × 4BANKS × 16-BITS
4,194,304-WORDS × 4BANKS × 8-BITS
8,388,608-WORDS × 4BANKS × 4-BITS

我们可以计算一下，结果可以发现这三个规格的容量都是 128Mbits，只是由于位宽的变化引起了存储单元的数量变化。从这个例子就也可以看出，在相同的总容量下，位宽可以采用多种不同的设计。

### 3、与芯片位宽相关的 DIMM 设计

为什么在相同的总容量下，位宽会有多种不同的设计呢？这主要是为了满足不同领域的需要。现在大家已经知道 P-Bank 的位宽是固定的，也就是说当芯片位宽确定下来后，一个 P-Bank 中芯片的个数也就自然确定了，而前文讲过 P-Bank 对芯片集合的位宽有要求，对芯片集合的容量则没有任何限制。高位宽的芯片可以让 DIMM 的设计简单一些（因为所用的芯片少），但在芯片容量相同时，这种 DIMM 的容量就肯定比不上采用低位宽芯片的模组，因为后者在一个 P-Bank 中可以容纳更多的芯片。比如上文中那个内存芯片容量标识图，容量都是 128Mbit，合 16MB。如果 DIMM 采用双 P-Bank+16bit 芯片设计，那么只能容纳 8 颗芯片，计 128MB。但如果采用 4bit 位宽芯片，则可容纳 32 颗芯片，计 512MB。DIMM 容量前后相差出 4 倍，可见芯片位宽对 DIMM 设计的重要性。因此，8bit 位宽芯片是桌面台式机上容量与成本之间平衡性较好的选择，所以在市场上也最为普及，而高于 16bit 位宽的芯片一般用在需要更大位宽的场所，如显卡等，至于 4bit 位宽芯片很明显非常适用于大容量内存应用领域，基本不会在标准的 Unbuffered 模组设计中出现。