

# 操作系统课程报告 3：中央处理器与存储器系统速度及影响因素的探究

张昊 1927405160

苏州大学计算机科学与技术学院 19 级图灵班

## 1. 引言

操作系统是方便用户管理和控制计算机软硬件资源的程序集合，在用户和计算机硬件间充当中介的角色。操作系统在计算机系统之上完成了进程管理、内存管理、文件管理和 I/O 设备管理这四大任务，其中与计算机硬件设备中的 CPU 和存储器系统密切相关。因此，了解中央处理器与存储器系统速度及影响因素，以及对于不平衡的速度差距是如何解决的是必要的。

本文主要讨论了以下三个内容：

- 中央处理器的性能及其影响因素；
- 存储器系统的速度及其影响因素；
- 计算机各部件速度差异的匹配。

简要地概括一下后续章节的内容，第2章从以 12 代酷睿 i9 处理器为例介绍了中央处理器性能的影响因素，第5章和第4章介绍了存储器系统中内存、硬盘等组成部件的速度差异以及影响因素，第5章从存储器的层次结构出发，介绍了现代计算机系统是如何解决速度不同的部件之间平衡速度差异的方法，并在第6章进行了总结。

## 2. 中央处理器的性能及其影响因素

中央处理器 (Central Processing Unit, CPU) 是计算机的主要设备之一，功能主要是解释计算机指令以及处理计算机软件中的数据 [2]，其性能和速度对计算机的整体运行速度有着较为显著的影响。

本文主要讨论消费级 CPU，这类 CPU 通常不是无限功耗 (即可以随意超频) 或者有无限的芯片面积，而关注在特定功率范围和特定芯片尺寸下，CPU

的性能表现 [8]。因此，暂且不考虑处理器的超频带来的性能上的提升。

以 2021 年第四季度英特尔公司发布的 12 代酷睿 i9 处理器 i9-12900K 为例，其采用了 Alder Lake-S 架构。该架构基于 Intel 7 制程 (10nm Enhanced SuperFin) 工艺打造，并首次将性能混合架构<sup>1</sup>(Intel Hybrid Technology) 引入 x86 桌面平台，首次集成了两种内核类型：性能核心 (Performance Core) 和能效核心 (Efficiency Core)，以带来跨越所有工作负载类型的性能提升<sup>2</sup>。酷睿 i9-12900K 采用 16 核心，24 线程设计，包括 8 颗性能核心和 8 颗能效核心。(由于能效核心不支持超线程，因此总计为 24 个线程。) 性能核心的基准频率 3.2GHz，最高睿频频率 5.2GHz；能效核心的基准频率为 2.4GHz，最高睿频频率 3.9GHz。[10]

接下来将对如上一些性能指标作简单介绍。

### 2.1. CPU 单核性能的影响因素

从 CPU 的性能参数来讲，CPU 的性能和速度取决于时钟频率 (一般以 GHz 来衡量) 和每周期指令数 (IPC, 每个时钟周期执行的平均指令数)，两者合并起来就是每秒可处理的指令 (IPS)。下面具体解释一下有关的评价指标：

- 时钟频率  $f$ ，又称主频，是指 CPU 内部晶振的频率，反映了 CPU 的基本工作节拍。

<sup>1</sup>英特尔的混合架构并不是简单的“大小核”，与 ARM 移动平台功耗优先的设计理念不同，是以性能为优先的。

<sup>2</sup>此外英特尔还开发了硬件线程调度器 (Intel Thread Director)，内嵌在 CPU 架构中，专门为合理地分配性能核心和能效核心的负载而设计，同操作系统紧密配合，以确保两种全新内核微架构无缝衔接。

- 时钟周期  $T$ ，为主频的倒数。

$$T = \frac{1}{f}$$

- 机器周期:  $m \times T$ ，一个机器周期包含若干个 ( $m$  个) 时钟周期。
- 指令周期:  $m \times T \times n$ ，执行一条指令所需要的时间，一般包含若干个 ( $n$  个) 机器周期。
- CPI (clock per instruction)，平均每条指令的平均时钟周期个数。

$$\begin{aligned} CPI &= \frac{\text{程序中所有指令的时钟周期数之和}}{\text{程序指令总数}} \\ &= \text{各类指令CPI} \times \text{该类指令比例} \end{aligned}$$

- IPC (instruction per clock)，每 (时钟) 周期运行多少个指令。
- MIPS，每秒执行百万条指令数。

$$\begin{aligned} MIPS &= \frac{1}{CPI \times T \times 10^6} \\ &= \frac{1}{\text{指令周期} \times 10^6} \\ &= \frac{f}{CPI \times 10^6} \\ &= \frac{f \times IPC}{10^6} \end{aligned}$$

- MFLOPS, 每秒百万浮点运算次数, 代表了 CPU 处理浮点运算的能力 (与 MIPS 不能相互转换)。

指令周期通常用若干个机器周期来表示，在机器语言中，使用执行一条指令所需要的机器周期数来说明指令执行的速度。而机器周期又包含若干个时钟周期。时钟周期是最基本的操作单位。

从 CPU 设计的角度来看 CPU 性能衡量标准 (单核性能)，主要由以下几个方面决定：

- 半导体工艺，也就是我们通常所说的 5/7/14 nm 制程<sup>3</sup>；
- 先进的微架构，采用的技术比如流水线、分支预测、乱序执行等等；
- 高速缓存架构。

<sup>3</sup>制程，即集成电路 (IC) 内电路与电路之间的距离。距离愈小，密度愈高，意味着在同样大小面积的 IC 中，可以拥有密度更高、功能更复杂的电路设计。

此外，我们常听到“睿频”这一专业名词。睿频是指 CPU 根据实际运行程序的需求，动态的增加处理器的运行频率来提高处理器的性能，同时保持处理器继续稳定运行在规定的功耗、电流、电压和温度范围内。睿频是 CPU 自动实现的，无需人工设置，并且 CPU 运行稳定。

可以看到，在硬件，微架构，缓存架构，芯片驱动软件等方面都有太多地方可以对 CPU 晶体管利用率，能耗比，极限性能，芯片极限面积，甚至芯片软件开发时间等方面做出优化，这些优化都会对传统意义上的“CPU 性能”产生积极影响。[8]

## 2.2. CPU 的多核性能

随着半导体工艺的不断发进步，芯片上的晶体管密度不断提高。而得益于晶体管密度的提升，芯片设计师们有足够的空间，在保证合理的功率和能量范围的前提下，同时获得复杂的架构和更高的频率。以这样一种简单粗暴的思路，甚至可以只依靠晶体管的小型化简单粗暴添加更多内核，然后通过制程进步带来的更多的核心来提升 CPU 性能，也间接地降低了计算机的价格。

此外，在这里解释一下“超线程”的概念。超线程 (HT, Hyper-Threading) 是英特尔研发的一种技术，于 2002 年发布。超线程技术把多线程处理器内部的两个逻辑内核模拟成两个物理芯片，让单个处理器就能使用线程级的并行计算，进而兼容多线程操作系统和软件。超线程技术充分利用空闲 CPU 资源，在相同时间内完成更多工作。虽然采用超线程技术能够同时执行两个线程，当两个线程同时需要某个资源时，其中一个线程必须让出资源暂时挂起，直到这些资源空闲以后才能继续。因此，超线程的性能并不等于两个 CPU 的性能。而且，超线程技术的 CPU 需要芯片组、操作系统和应用软件的支持，才能比较理想地发挥该项技术的优势。

CPU 的多核发展至今，有以下两类架构：

- 对称多处理 (Symmetric Multiprocessing, SMP)
  - 每个处理器运行操作系统的相同副本；
  - 许多进程可以立即运行不会降低性能；
  - 多数现代操作系统支持 SMP。

- 非对称多处理 (Asymmetric Multiprocessing, ASMP)

- 各个处理器不对等；
- 分为主处理器和从处理器；
- 主处理器为从处理器安排任务；
- 一般用于特定的场景。[13]

在很早之前，CPU 的多核趋势就已经越来越明显。但是在半导体制程进步逐渐缓慢时，将不同的任务分别交给不同的核心可以更有效地利用 CPU 的性能，从而达到更高的能效比。但是，多核架构分为对称和非对称两种，其中各个核心如何分工，以及处理多核心间的通信延迟是目前提升 CPU 性能需要重点考虑的两个问题。一个普遍的观点是，上层的算法和调度方法会成为 CPU 性能的一个重要影响因素。[8]

### 3. 主存储器的速度及其影响因素

在现代计算机中，当前计算机正在执行的程序和数据均存放在内存中，主存储器 (内存) 处于全机中心地位。我们现在使用的内存，绝大多数都是随机存储器 (random access memory, RAM)，其最大的特点是可以通过指令可以随机地、个别地对各个存储单元进行访问 (读写)，但是停电会造成信号丢失。其中，按存储元件在运行中能否长时间保存信息分为静态存储器 (SRAM) 和动态存储器 (DRAM) 两种。

#### 3.1. 主存储器的主要技术指标

主存储器的主要技术指标主要为容量、存取时间和存取周期。考虑到本文是有关速度的探究，所以主要关注后两者。

- 存储器存取时间 (Memory Access Time): 又称存储器访问时间，是指启动一次存储器操作到完成该操作所经历的时间。
- 存储周期 (Memory Cycle Time): 是指连续启动两次独立的存储器操作 (例如连续两次读操作) 所需间隔的最小时间。

通常来说，存储周期略大于存取时间。但是一

般具有合适价格的主存储器能提供信息的速度总跟不上 CPU 的处理速度<sup>4</sup>。

#### 3.2. 常见影响因素

**频率.** 内存主频和 CPU 主频一样，用来表示内存的速度，它代表着该内存所能达到的最高工作频率，以 MHz 为单位来计量的。内存主频越高在一定程度上代表着内存所能达到的速度越快。

内存本身并不具备晶体振荡器，因此内存工作时的时钟信号是由主板芯片组的北桥或直接由主板的时钟发生器提供的，也就是说内存无法决定自身的工作频率，其实际工作频率是由主板来决定的。

**带宽.** 内存带宽是处理器可以从内存读取数据或将数据存储到内存的速率。内存带宽通常以字节/秒为单位表示，但对于自然数据大小不是 8 位的倍数的系统，这可能会有所不同。

从参考资料 [3] 中我们可以获得如下解释。

对于商用 DDR SDRAM，DDR2 SDRAM 和 DDR3 SDRAM 内存，总带宽是以下值之乘积：

- **基本 DRAM 时钟频率**
- **每时钟的数据传输次数:** 在“双倍数据速率” (DDR、DDR2、DDR3、DDR4) 内存的情况下是两次
- **内存总线 (接口) 带宽:** 每个 DDR、DDR2 或 DDR3 内存接口都是 64 位。有时也被称为 1 个“行”
- **接口数量:** 现代 PC 通常使用两个内存接口 (双通道模式) 来实现有效的 128 位总线带宽

例如，一个计算机有两个通道的内存，每通道都配备运行在 400MHz 时钟频率的 DDR2-800 模块，则其理论最大内存带宽为：每秒 400,000,000 个时钟 × 每个时钟 2 个通道 × 每行 64 个位 × 2 个接口 = 每秒 102,400,000,000 (1024 亿) 比特 (也就是 12,800 MB/s 或 12.8 GB/s)。但是该理论最大内存带宽被称为“突发速率”，是不可持续的。

DDR，DDR2 和 DDR3 模块的命名约定指定了最大速度 (例如，DDR2-800) 或最大带宽 (例如，

<sup>4</sup>这一点将在第5章作详细介绍。

PC2-6400)。速度等级 (800) 不是最大时钟速度，而是它的两倍 (1600，因为数据速率是加倍的)。带宽等级 (6400) 是使用 64 位每秒传输的最大兆字节数。在双通道模式中，这实际上是 128 位。因此，示例中的内存配置可以简化为：两个 DDR2-800 模块以双通道模式运行。

**时序。** 内存时序是描述内存条性能的一种参数，一般存储在内存条的 SPD 中。一般数字“A-B-C-D”分别对应的参数是“CL-tRCD-tRP-tRAS”，它们的含义依次为：

- CAS Latency(简称 CL 值) 内存 CAS 延迟时间，它是内存的重要参数之一，某些牌子的内存会把 CL 值印在内存条的标签上；
- RAS-to-CAS Delay(tRCD)，内存行地址传输到列地址的延迟时间；
- RAS Precharge Delay(tRP)，内存行地址选通脉冲预充电时间；
- Row Active Delay(tRAS)，内存行地址选通延迟。[11]

可以看出，这些数字都表示延迟，也就是内存的反应时间。当内存接收到 CPU 发来的指令后，通常需要几个时钟周期来处理它，比如访问某一块数据，这就对应时序参数。当然，处理时间越短，内存性能越好。下面引用参考资料 [9] 中的内容对上述参数作具体解释。

内存收到 CPU 的指令之后，要先确定数据具体在哪一行，时序的第二个参数 tRCD 就代表这个时间，意思就是内存控制器接收到行的指令后，需要等待多长时间才能访问这一行。由于每一行都有多个数据，内存并不能哪一个才是 CPU 需要找的，所以 tRCD 只是一个估值，这也就是为什么小幅改动这个数值，并不会显著改变内存的性能。

内存确定了数据所在的行之后，要想找出数据，还得确定列。时序的第一个数字也就是 CL(CAS)，就表示内存确定了行数之后，还得等待多长时间，才能访问具体的列数 (时间周期)。行列必然产生交点，也就是说确定了行数和列数之后，就能准确找到目

标数据，所以 CL 是一个准确的值，任何改动都会影响目标数据的位置，所以它在时序当中是最关键的一个参数，对内存性能的发挥着举足轻重的作用。

内存时序的第三个参数 tRP，就是如果我们已经确定了一行，还要再确定另外一行所需要等待的时间 (时间周期)。

第四个参数 tRAS，可以简单理解成内存写入或者读取数据的一个时间，它一般接近于前三个参数的总和。

所以，在保障稳定性的前提下，内存时序越低越好，而高频率和低时序是个矛盾体，一般频率上去了，时序就得有所牺牲，要想足够低的时序，频率又很难拔高。时序改变后，内存延迟的变化比内存读写速度的变化更加明显，这也说明时序对内存的影响，更侧重在延迟方面。

**接口类型。** 起初，内存是一块块的 IC，焊接到主机板上才能正常使用，一旦某一块内存 IC 坏了，必须焊下来才能更换，这样很不方便。后来，出现了模块化的条装内存，上面集成了多块内存 IC，在主板上设计了内存插槽，内存条就可随意拆卸，内存的维修和扩充变得非常方便。(当然，目前许多笔记本电脑仍使用了板载内存，其中苹果公司 M1 系列芯片的统一内存即为一个例子。) 由此，内存也就有了相应的接口类型。

386 时代，计算机开始使用内存条。内存条的接口类型从最初的 FPM(快页式内存)，经历了 EDO(extended data out，扩展数据输出)DRAM，SDRAM 内存 (synchronous dynamic RAM，同步动态内存)，RAMBUS DRAM 多种类型。由于 RDRAM 的工艺复杂，价格过高，而且 RAMBUS 公司还要收取相应的版权费，再加上闹得沸沸扬扬的 i820 回收事件。其它厂家出于市场考虑，根据 RAMBUS 双向脉冲的特点，提出了 DDR SDRAM，也就是我们现在最主流的内存条。

DDR SDRAM(Dual data rate SDRAM) 简称 DDR 内存，也就是“双倍速率 SDRAM”的意思。DDR 内存存在时钟信号上升沿与下降沿各传输一次数据，这使得 DDR 的数据传输速度为传统 SDRAM

的两倍。DDR 内存的引脚数为 184 线，工作电压仅为 2.5V，因而功耗亦随之减少。

虽然 DDR5 在内存市场上还比较少见，但是在显卡上 DDR5 已经被广泛地使用了。DDR5 是 JEDEC 在 2020 年 7 月公布的新一代内存标准，目前最新的英特尔酷睿 12 代平台支持该标准，与目前广泛使用的 DDR4 内存相比性能更强，功耗更低，频率起步 4800MHz，电压从 1.2V 降低到 1.1V，同时集成了集成电源管理芯片 PMIC，支持内存纠错机制 ECC。[1]

#### 4. 辅助存储器的速度及其影响因素

当前市场上流行的辅助存储器主要有磁表面存储器（主要代表为磁盘）以及采用半导体存储介质和传统磁盘接口的存储器（即固态硬盘），他们的速度要远远慢于内存。我们都知道，CPU 读取位于不同位置的数据所耗费的代价是不同的。有一个很直观的数据可以说明这个问题：如果程序需要的数据存储在 CPU 的寄存器内，那么在指令的执行期间，在 0 个周期内就可以访问到；如果存储在高速缓存（cache）中，需要 4-75 个周期；如果存储在主存中，需要上百个周期；如果存储在磁盘上，需要大约几千万个周期 [12]。

接下来将以磁盘和固态硬盘为代表，简要说明其速度及其影响因素。

##### 4.1. 磁盘（机械硬盘）

**寻址时间。** 磁盘采取直接存取方式，寻址时间包括两部分：

- $t_s$ ：磁头寻找目标磁道所需的找道时间
- $t_w$ ：找到磁道以后，磁头等待所需要读写的区段旋转到它的下方所需要的等待时间。

寻找相邻磁道和从最外面磁道找到最里面磁道所需的时间不同，磁头等待不同区段所花的时间也不同，取它们的平均值，称作平均寻址时间  $T_a$ ，由平均找道时间  $t_{sa}$  和平均等待时间  $t_{wa}$  组成：

$$T_a = T_{sa} + T_{wa} = \frac{t_{smax} + t_{smin}}{2} + \frac{t_{wmax} + t_{wmin}}{2}$$

**数据传输率。** 磁表面存储器在单位时间内与主机之间传送数据的位数或字节数，叫数据传输率  $D_r$ 。为确保主机与磁表面存储器之间传输信息不丢失，传输率与存储设备和主机接口逻辑两者有关。

- 从设备方面考虑，传输率等于记录密度  $D$  和记录介质的运动速度  $V$  的乘积。
- 从主机接口逻辑考虑，应有足够快的传送速度接收/发送信息，以便主机与辅存之间的传输正确无误。

**转速。** 当硬盘运行时，磁头沿着盘片的半径方向运动，盘片以每分钟上千转的速度旋转，磁头就可以在盘片的指定位置上进行数据的读写操作。因此，机械硬盘的转速指的是内部电机主轴的旋转速度，衡量转速的单位为 RPM (Revolutions Per Minute，即转/分)，硬盘转速越快，访问数据的时间就越短，硬盘的数据传输率一般就越快。

目前市面上常见的机械硬盘转速有 5400rpm 和 7200rpm。单从转速上看，7200rpm 硬盘的盘片每转一圈耗时 8.33 毫秒左右，5400rpm 硬盘每转一圈耗时 11.11 毫秒左右，前者领先 15% 左右，不过这一性能提升在日常使用中并不会十分显著的体现（因为日常使用的场景中对硬盘的顺序读写性能要求远没有对随机读写能力那么高）。

随着转速的提高，不仅是速度在提升，噪音和发热量也会随之显著增加。目前来说，常见的消费级硬盘多为 5400rpm 或 7200rpm，转速过万的产品更多时候出现在企业中。

##### 4.2. 固态硬盘

固态硬盘 (SSD) 采用闪存作为存储介质，速度相对机械硬盘更快。固态硬盘厂商大多会宣称自家的固态硬盘持续读写速度超过了 500 MB/s，近年来的 NVMe 固态硬盘可达到 2000 MB/s 左右，甚至 4000 MB/s 以上。固态硬盘的快不仅仅体现于持续读写上，还体现在随机读写速度快。此外，固态硬盘不需要磁头，读取速度更快。最常见的 7200 转机械硬盘的寻道时间一般为 12-14 毫秒，而固态硬盘可以轻易达到 0.1 毫秒甚至更低。

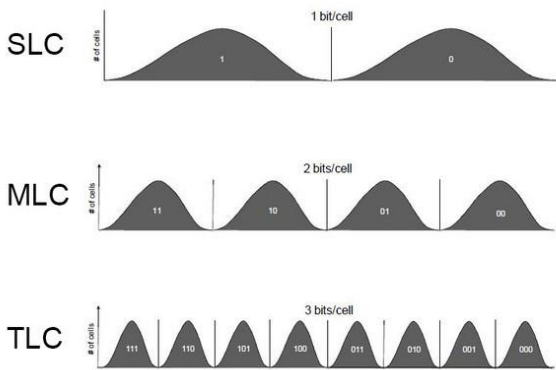


图 1. SLC MLC TLC 示意图 [7]。

影响固态硬盘性能的几个因素主要在于主控芯片、闪存介质以及是否设置有缓存。在上述条件相同的情况下，采用何种接口也可能会影响 SSD 的性能。

**主控芯片.** 市面上比较常见的固态硬盘有多种主控芯片，如 Sandisk、Samsung、Intel 等。主控芯片可以看作是固态硬盘的大脑，其作用一是合理调配数据在各个闪存芯片上的负荷，二则是承担了整个数据中转，连接闪存芯片和外部接口。不同的主控之间能力相差非常大，在数据处理能力、算法，对闪存芯片的读取写入控制上会有非常大的不同，直接会导致固态硬盘产品在性能上差距高达数倍。

**闪存介质.** 现在主流的固态硬盘大多采用 NAND 闪存为核心。在固态硬盘的电路板上，闪存颗粒占据了大部分的空间。

NAND 闪存芯片又分为 SLC、MLC、TLC、QLC 这四种规格。如图1展示了 SLC、MLC、TLC 浮动栅存储容量的关系。

- SLC(Single-Level Cell, 单层单元), 即 1bit/cell, 利用正、负两种电荷, 一个浮动栅存储 1 个 bit 的信息, 约 10 万次擦写寿命。速度快, 寿命长, 价格贵 (约 MLC 3 倍以上的价格)。
- MLC(Multi-Level Cell, 双层单元), 即 2bit/cell, 利用不同电位的电荷, 一个浮动栅存储 2 个 bit 的信息, 约 5000-10000 次擦写寿命。速度一般,

寿命一般, 价格一般。

- TLC(Trinary-Level Cell, 三层单元), 即 3bit/cell, 利用不同电位的电荷, 一个浮动栅存储 3 个 bit 的信息, 约 500-1000 次擦写寿命。也有 Flash 厂家叫 8LC, 速度慢, 寿命短, 价格便宜。
- QLC(Quad-Level Cell, 四层单元), 即 4bit/cell。支持 16 充电值, 速度最慢寿命最短, 目前中技术上在研发阶段, 但是 intel、三星电子等厂商都已经取得了不错的进展。但在 SSD 应用中目前仍不现实。

相对于 SLC 来说, MLC 的容量大了 100%, 寿命缩短为 SLC 的 1/10。相对于 MLC 来说, TLC 的容量大了 50%, 寿命缩短为 MLC 的 1/20。

**缓存.** 主控芯片旁边是缓存颗粒, 固态硬盘和传统硬盘一样需要高速的缓存芯片辅助主控芯片进行数据处理。有一些廉价固态硬盘方案为了节省成本, 省去了缓存芯片, 这对于性能会有一些影响, 尤其是小文件的读写性能和使用寿命上。

**接口与协议.** 对于固态盘来说, 不仅仅是颗粒 (介质) 阻碍了速度, 实际上硬盘采用了何种协议、使用了哪种总线也可能阻碍了固态理论上能达到的速度。我们常说的 SATA、PCIe、NVMe、M.2 等名词, 都或多或少对硬盘的速度有影响。

- PCIe(Peripheral Component Interconnect Express, PCI-Express) 是继 ISA 和 PCI 总线之后的第三代 I/O 总线。PCIe 协议是一种端对端的互连协议, 提供了高速传输带宽的解决方案。目前 PCIe 已经发展到第四代 PCIe4.0(目前市面上可见的最高为 PCIe 4.0), 每一代的发展, 最明显的特征就是速率翻倍。典型 PC 主板上的 PCIe 4.0 扩展槽有四种不同类型: PCIe x1、PCIe x4、PCIe x8 和 PCIe x16。这些数字表示每个 PCIe 插槽有多少个通道。PCIe 插槽的通道越多, 传输带宽就越高, 具体速率如表1所示。
- SATA(Serial ATA) 是一种电脑总线, 负责主板和大容量存储装置 (如硬盘及光驱) 之间的数据传输。SATA 总线的速率如表2所示。

版本	推出时间	原始传输率	带宽 ×1	带宽 ×2	带宽 ×4	带宽 ×8	带宽 ×16
1	2003	2.5 GT/s	250 MB/s	0.50 GB/s	1.0 GB/s	2.0 GB/s	4.0 GB/s
2	2007	5.0 GT/s	500 MB/s	1.0 GB/s	2.0 GB/s	4.0 GB/s	8.0 GB/s
3	2010	8.0 GT/s	984.6 MB/s	1.97 GB/s	3.94 GB/s	7.88 GB/s	15.8 GB/s
4	2017	16.0 GT/s	1969 MB/s	3.94 GB/s	7.88 GB/s	15.75 GB/s	31.5 GB/s

表 1. PCI Express 总线性能 [4]。

版本	带宽	理论速度
SATA Express	16 Gb/s	1969 MB/s
SATA 3.0	6 Gb/s	600 MB/s
SATA 2.0	3 Gb/s	300 MB/s
SATA 1.0	1.5 Gb/s	150 MB/s

表 2. SATA 总线性能 [5]。

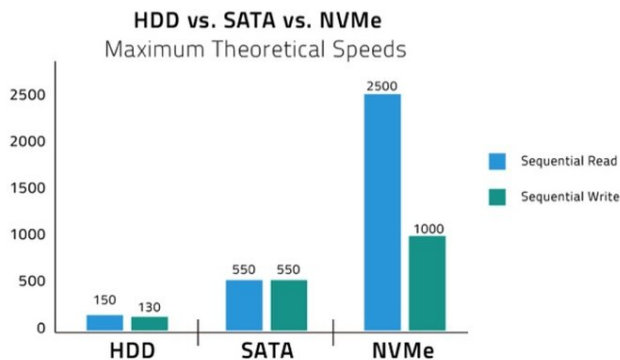


图 2. HDD 能达到的最高理论速度，与 SATA、NVMe 最高提供的理论速度对比 [6]。

- M.2 接口是一种新的主机接口方案，可以兼容多种通信协议，如 SATA、PCIe 等。M.2 接口是为超极本身身定做的新一代接口标准，以取代原来的 mSATA 接口。无论是更小巧的规格尺寸，还是更高的传输性能，M.2 都远胜于 mSATA。
- NVM Express(NVMe)，或称非易失性内存主机控制器接口规范，是一个逻辑设备接口规范，专为 NAND 闪存等非易失性存储设计。NVMe 在很大程度上实现了并行性，极大地提高了吞吐量，能够使固态硬盘运行得更快，如图2所示为 HDD、SATA、NVMe 最高提供的理论速度对比。

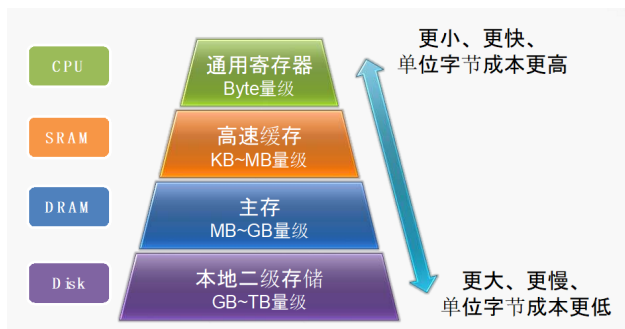


图 3. 存储层次。

## 5. 计算机各部件速度差异的匹配

CPU 发展迅速，计算能力不断提高，速度越来越快，但是外围设备却无法运行的更快，于是产生了高速缓存 (cache)、主存储器、固态硬盘、磁盘等一系列的硬件，这些硬件的速度分布呈金字塔形 (如图3所示)，由它们组成了存储器系统。

如图4所示，我们对比 CPU 和存储器的特性以及价格可以发现，随着技术水平的提高，速度、容量均在提高而 (每兆) 价格均在降低，且存储器的发展速度是远大于中央处理器的。

实际上，存储器系统是一个具有不同容量、成本和访问时间的存储设备组成的层次结构。CPU 的寄存器保存着最常用的数据。靠近 CPU 的小的、快速的高速缓存存储器作为一部分存储在相对慢速的主存储器中数据和指令的缓冲区域。主存缓存存储在容量较大的、慢速磁盘上的数据，而磁盘常常又作为存储在网络连接的其他机器的磁盘或磁带上的数据的缓冲区域。[12]

存储器层次结构是可行的，这是因为与下一个更低层次的存储设备相比来说，一个编写良好的程序倾向于更频繁地访问某一个层次上的存储设备。所以，下一层的存储设备可以更慢速一点，也因此

	1980	1990	2000	2010	2010:1980	
	8080	386	Pentium II	Core i7	/	
CPU	Clock rate(MHz)	1	20	600	2,500	2,500
	Cycle time(ns)	1,000	50	1.6	0.4	2,500
	Cores	1	1	1	4	4
	Effective Cycle time(ns)	1,000	50	1.6	0.1	10,000
SRAM	\$/MB	19,200	320	100	60	320
	access time(ns)	300	35	3	1.5	200
DRAM	\$/MB	8,000	100	1	0.06	130,000
	access time(ns)	375	100	60	40	9
	typical size(MB)	0.064	4	64	8,000	125,000
Disk	\$/MB	500	8	0.01	0.0003	1,600,000
	access time(ms)	87	28	8	3	29
	typical size(MB)	1	160	20,000	1,500,000	1,500,000

图 4. CPU 和存储器的特性以及价格比较。

可以更大，每个比特位更便宜。整体效果是一个大的存储器池，其成本与层次结构底层最便宜的存储设备相当，但是却以接近于层次结构顶部存储设备的高速率向程序提供数据。[12]

事实上，高速缓存也分了三类型，称为一级缓存 (L1)、二级缓存 (L2)、三级缓存 (L3)，三级缓存在 CPU 和内存之间建立了数据通路，使得数据可以很快地在 CPU(的寄存器) 和内存中交换。最接近 CPU(中的寄存器) 的就是 L1，它提供了足够高的速度来满足 CPU 运行的需求，但是其发热很大，无法做的更大。于是 L1 周围就有了 L2，降速提高容量，同时降低内存延迟。L3 其实和内存的地位接近，有时候数据可以不经过 L3(即 L3 缓存不是必需的)，不过如果数据在 L3 中则可以提高性能，进一步降低内存延迟。一般来说，L1 可分为数据缓存和指令缓存，为单个核心所有；L1 和 L2 都可以整合至 CPU 内部；L3 则是所有的核心共享一份缓存。硬盘等之所以称之为“外存”，是因为 CPU 需要先把外存的数据和指令加载到内存中，才可以使用。这一级级的缓存，就是速度的降低，容量的增大，同时也是计算机系统解决速度不匹配的一个重要手段。

这种缓存技术不仅仅存在于 CPU 和主存之间，而是广泛存在于速度、容量、价格存在差距的地方。上面第章提到的 SSD 的缓存正是这一思路，可以很好地减少等待时间；在 NAS 中，使用多块大容量的机械硬盘作为主要的存储介质，增加一个固态硬盘作为缓存可以显著提升访问速度，也是缓存技术的一个应用。

## 6. 总结

本文针对中央处理器与存储器系统速度及影响因素，以及计算机系统对于不平衡的速度差距是如何解决的这两个问题进行了总结和整理。总体来说，作为程序员，理解存储器的层次结构，认识到缓存技术在计算机系统中的重要作用是十分重要的。这是一种基本而持久 [12] 的思想，对于如何使得自己的程序更为高效，更好地理解操作系统中的有关概念都有十分重要的帮助。

## 参考文献

- [1] JEDEC Memory Workshops: DDR5, NVDIMM-P, DRAM Tutorial. <https://www.jedec.org/memory-workshops-2017>. 5
- [2] Wikipedia: Central processing unit. [https://en.wikipedia.org/wiki/Central\\_processing\\_unit](https://en.wikipedia.org/wiki/Central_processing_unit). 1
- [3] Wikipedia: Memory bandwidth. [https://en.wikipedia.org/wiki/Memory\\_bandwidth](https://en.wikipedia.org/wiki/Memory_bandwidth). 3
- [4] Wikipedia: PCI Express. [https://zh.wikipedia.org/wiki/PCI\\_Express](https://zh.wikipedia.org/wiki/PCI_Express). 7
- [5] Wikipedia: SATA. <https://zh.wikipedia.org/wiki/SATA>. 7
- [6] 什么是 NVMe? 一篇文章理清它的前生今世. <https://storage.it168.com/a2018/0921/5045/000005045252.shtml>. 7
- [7] 固态硬盘的 SLC、MLC、TLC 和 QLC 的区别. <https://blog.csdn.net/liuyukuan/article/details/85872743>. 6
- [8] 知乎: CPU 的性能是不是快到天花板了? 为什么? <https://www.zhihu.com/question/376567574/answer/1954682510>. 1, 2, 3
- [9] 知乎: 内存时序是什么? 对性能影响有多大? <https://zhuanlan.zhihu.com/p/382912420>. 4
- [10] 英特尔酷睿 i9-12900K 评测. <https://xw.qq.com/cmsid/20211105A02YEA00>. 1
- [11] 英睿达: 内存时序. <https://www.crucial.cn/articles/about-memory/what-is-the-memory-timing-sequence>. 4
- [12] Bryant and Randal E. Computer systems : a programmer's perspective / Third edition. 2016. 5, 7, 8
- [13] P. B. Galvin, G. Gagne, A. Silberschatz, et al. Operating system concepts. 3