

IBM POWER 7 multicore server processor

POWER7 的目的是改善 socket 级别、核心级和线程级别的性能表现明显超过 POWER 6，因为 POWER 7 使用了如下技术：

1. 减少核心区域和功率，使得与 POWER6 相同的电路功率下扩展成为 4 倍核心数量。在 POWER7 处理器频率降低，而更高的性能是通过更强调微体系结构的改进，如积极的乱序执行、先进四同步多线程(SMT)、先进的分支预测、高级预取、缓存和内存减少延迟和带宽的改善。
2. POWER7 芯片使用与 POWER6 同一套接字，SMP 和 I/O 总线。内存总线已被重新设计以更高的频率运行，使用不同信号支持八个高性能核心处理器芯片的高内存带宽要求。
3. 删除外部三级(L3)缓存芯片，为了降低系统成本功耗和节省大量芯片 I/O 带宽。
4. 显著的性能改进的高性能计算(HPC)、加倍 POWER7 核心的浮点能力，与 POWER6 相比，产生 8 倍浮点操作每拍每核心。

POWER 7 核心主要增强分支预测、指令预取和乱序执行的能力。最多支持 4 个线程，线程之间可以动态切换。

Power 7 CPU 内核 除此之外，在 Power 7 每个内核中都拥有 12 个执行单元。每个执行单元包含 2 个整数运算单元、2 个存储/读取单元、4 个双精度浮点运算单元、1 个支持 VSX 的矢量执行单元、1 个十进制浮点运算单元、1 个分支单元、1 个寄存器单元。在 Power 7 CPU 中每个内核都有 32KB 一级指令和数据缓存、256KB 二级缓存。所有核心共享 32MB eDRAM 三级缓存。Power 7 初期频率就能达到 3GHz~4.14GHz，并内置了两个 4 通道 DDR3 内存控制器，以提供最大 100GB/s 的内存带宽。值得一提的是，IBM Power 7 中三级缓存是用 eDRAM 实现的，而非传统 CPU 中的 SRAM。和 SRAM 相比，eDRAM 能极大的节省晶体管数量和降低芯片面积，并且提供直逼 SRAM 的传输带宽——在微软 XBox 360 游戏机的 GPU 上，eDRAM 就轻松提供了 256GB/s 的传输率。由于 eDRAM 的引入，Power 7 在搭载 32MB 三级缓存时芯片面积也没有过度增长。和前辈 Power 6 相比，Power 7 的主频已经从 5GHz 降低到了 4.14GHz。但由于 Power 7 拥有更多的内核、更强的并发多线程能力，所以在性能上 Power 7 可以实现倍数的跨越。Power 7 是一颗聪明的处理器，IBM 工程师为它赋予了更多的智能技术，在性能、功耗上取得更大的突破。能自动优化核心 在 Intel 推出 Core i7 CPU 的时候，TurboBoost 功能成了产品重要卖点不断宣传。Intel 的 TurboBoost 技术可以让 CPU 根据负载，在不同的内核负载下，自动提升单核频率以及整体频率以实现更快的运行速度。在 Power 7 上，IBM 更进一步，提出了智能核心的概念。Power 7 处理器拥有最多 8 个内核，因此 IBM 为 Power 7 设计了 TurboCore 和 MaxCore 两大运行模式。TurboCore 模式可以对数据库或工作负载进行高度优化，同时采用 4 个内核运行，并把芯片中所有 8 个内核大部分资源放到 4 个运行的内核中，以提供更大的缓存和内存带宽，提升时钟频率，提高单核性能。

Power 7 CPU 内核架构 当不采用 TurboCore 模式时，所有 Power 7 处理器在 MaxCore 模式下，拥有多达 8 内核、每内核 4 线程，共计 32 线程的并行处理能力。你千万别以为智能内核只是简单的超频、关闭打开内核那么容易。在 TurboCore 模式下，虽

然有 4 个内核会被关闭，但这些内核所占据的缓存和各种执行资源都会被释放，由工作中的 4 个内核统一调度，以实现资源利用的最大化。夸张的 8 核 32 线程并行处理能力 在 Power 7 中，IBM 引入的 SMT4 技术可以在单一内核上实现 4 路同步多线程功能。这样一次就能吃下 4 条线程，极大地提升并行能力。但是所谓的 SMT 同步多线程，并不是真正的多内核多线程。只是通过不断的状态切换，以提升内核利用率的一种方式。这样的设计在并行度很高的应用中能获得立竿见影的效果。但在数据库等应用中往往会出现性能下降。过去我们大多采用手动打开、关闭 SMT 同步多线程的方法“因地制宜”。

Power 7 提供 4 路同步多线程功能 在 Power 7 上，IBM 引入了智能线程 (Intelligent Threads)，可以根据工作负载要求进行设置不同的多线程模式，系统可以自动选择，也可以由管理员进行手动设置。这样就能在并发线程和执行效率中获得良好的平衡。必须指出的是，智能线程功能是需要操作系统支持的。只有在 2010 年 4 月问世的 IBM AIX 6.1 TL05 操作系统支持此功能——在现阶段的 Linux 等操作系统中，Power 7 的 SMT 功能将完全无法使用。这些操作系统会把 Power 7 当作 8 内核普通处理器进行管理。更智能的缓存和内存控制器 从自动控制内核开关，到自动判断 SMT 功能是否打开，打开多少。在处理器运算能力调度方面，Power 7 已经技压群雄。不过，IBM 似乎还不满足于运算能力的灵活调度，他们想让缓存和内存系统也变得更加智能高效。Power 7 处理器的 L3 缓存和 Intel Core i7 有许多相似之处——例如每个内核都在 L3 缓存中有自己的高速本地 L3 缓存区 (Fast Local Region of L3 Cache, FLR-L3)。但每个内核之间，却依然能通过 L3 缓存共享数据。根据 IBM 的说法，Power 7 32MB L3 Cache 中，有 4MB 缓存的速度快和延迟低，提供的性能介于 L1 和 L2 之间，由此来确保处理器在众多内核并行工作时仍然有较好的性能。

Power 7 内存访问机制 为了压倒 Intel 在 Core i7 中集成的 3 通道 DDR3 内存控制器，Power 7 干脆直接集成了两个 4 通道 DDR3 内存控制器，这意味着每个 Power 7 处理器最多能支持 256GB 的 DDR3 内存。在内存控制器内部，Power 7 专门设计了芯片内 DDR3 内存缓冲区。这样的设计极大地增加了每个 Power 7 内核所能管理的内存容量，并且实现了更高级别的冗余扩展和电源管理。由于采用了 4 通道内存设计，所以 Power 7 已经取消了对单根低容量 DDR3 内存的支持。在 Power 7 上用户最少需要安装两条 4GB DDR3 内存才能正常启动。

Power 7 内存控制器 值得一提的是，IBM Power 7 中引入了新的低电压差分信号传输方式，这样让 Power 7 在支持海量内存和 8 个以上 DIMM 内存插槽时，主板布线不至于太过复杂。强大的智能功耗管理功能 几乎所有新的 CPU 都在电源管理上狠下功夫。Power 7 也不例外。在 Power 7 上，IBM 提出了智能功耗 (Intelligent Energy) 的概念，并在 Power 7 中延续了 Power 6 CPU 上大获好评的 EnergyScale 功能。Power 7 处理器内部构建的 EnergyScale 单元能不断地搜集整个电脑的功耗数据，然后将其汇报给 IBM Systems Director Active Energy Manager 功耗管理软件。在 Power 7 运行的时候，IBM 的功耗管理软件就能即时提供功耗数据和运行状态。用户甚至可以直接设定整部 Power 7 电脑最多能使用的功耗，以及最少能使用的功耗，以调整机房的耗电量。在智能功耗功能的帮助下，我们甚至可以让 Power 7 在夜晚以 50% 的功耗工作，在白天繁忙时段才满负荷工作。而 Power 7 处理器在空闲时，也将会自动进入“Nap”状态，关闭执行单元的供电、降低频率和电压以实现更低的功耗。

IBM 电源管理工具 IBM Systems Director Active Energy Manager 根据 IBM 的测试，Power 7 在引入一系列“智能”技术后，每瓦性能大幅提高，Power 7 比相似的 Intel x86 系统提升 2 倍，比 Sun SPARC 服务器和相似的 HP 安腾服务器分别高出 4 倍和 8 倍。无需操作系统支持，无需做任何硬件改动。可用内存容量就能多 50%？这可不是什么内存清理软件的广告词，而是 IBM 在 Power 7 上引入的 Active Memory Expansion 内存实时压缩技术的魔力。Active Memory Expansion 功能可以通过 Power 7 电脑管理程序打开或者关闭。在打开 Active Memory Expansion 功能后，CPU 将会实时压缩内存中的数据，以获得更大的可用内存空间。在 IBM 的测试数据中，Power 7 打开 Active Memory Expansion 功能后，将会在 SAP 程序里面多出 50% 的可用内存空间，从而极大的提升系统性能。

打开 AME 功能后，内存会被分成压缩数据和非压缩数据两大区域 Power 7 将会自动压缩不被反复调用的数据，以实现可用内存空间加大。数据被压缩之后，将无法由程序实时调用，因此 Active Memory Expansion 内存压缩功能并无法适应所有类型的程序。

AME 打开和关闭后的性能对比 由于 Active Memory Expansion 功能完全由硬件实现，因此该功能对性能的影响微乎其微。由于打开内存压缩之后，可用内存有所增加，所以在高负载情况下，Active Memory Expansion 对性能提升有明显的帮助。Active Memory Expansion 可以针对每个内存逻辑分区打开或者关闭，操作系统将会自动分析哪些数据能被压缩后放入压缩数据池，哪些数据无法压缩。现阶段，要打开 Power 7 的 AME 功能，必须使用 IBM AIX 6.1 TL4 SP2 或更高版本的操作系统。PowerVM—1000 个虚拟机的威力 在服务器和高性能计算领域，虚拟化已经成了标准配备。AMD 和 Intel CPU 大多只能提供不到 10 个虚拟机同时运行的硬件虚拟化支持。而在 Power 7 上，IBM 祭出的 PowerVM 虚拟化技术能让 Power 7 CPU 每个内核都能硬件支持 10 个虚拟机镜像。目前的 8 路 64 核系统可支持单系统最高 640 个虚拟机同时运行。