

“一亿”是“一千万”的十倍吗？

原创 邓子平 [多物理场仿真技术](#)



笔者在 [一篇文章入门多物理场有限元（全篇）](#) 一文中提出过该问题。以下是原文内容：

-----1亿是1千万的10倍？-----

早期笔者使用矩量法求解线性方程组，在不使用快速多级方法，自由度达到3万时，台式机上已经无法求解出，8G的机器内存不够用。对于满秩矩阵的线性方程组，常规求解方法时间复杂度为 n^3 （ n 的3次方）。

对于自由度1千万以下稀疏矩阵的求解，好的台式机基本能应付，而当自由度达到1亿的时候，简单的将硬件乘以10倍完全不能满足要求。因为计算的空间复杂度，时间复杂度并不是线性，通常是 $N\log N$ ， N^2 或者更高。当自由度达到1亿时，不仅需要对硬件核心部件CPU，内存扩容，而且在磁盘阵列，I/O，并发计算，GPU，网络，带宽等方面都提出了更苛刻的要求。

在算法方面看，减少网格密度，在物理量梯度大的地方加密，无变化的地方将网格变稀疏，可以有效减少计算量；另外优化求解算法本身，使其更加易于并行化计算。

从目前来看，计算机硬件计算能力的更新速度，跟不上指数级求解规模的增加速度。这也是量子力学发展的最大推动力！

针对该问题，本文稍微展开一下。

这里的1亿，主要指仿真中最终求解线性方程组的自由度个数，就是刚度矩阵的规模，即在方程 $Kx=b$ 中， K 的规模为1亿*1亿的方阵。

从硬件角度看，性能瓶颈首先在内存。目前台式机的内存普遍在64G以下。对于1千万的稀疏矩阵，存储勉强够用，加上计算，一般都不够用，在windows系统上如果内存不够用，会启用虚拟内存，也就是用页面交换把硬盘当内存使用，频繁读写磁盘，这时候性能会大大降低。而对于1亿自由度的矩阵，可能刚度矩阵组装都不够用。对于一亿自由度的矩阵，首先要保证内存能够存储，一般的方法是使用具有超大内存的工作站服务器。

内存问题解决后，再就是CPU性能。前面讲过，通常矩阵计算的空间复杂度，时间复杂度并不是线性，而是 $N\log N$ ， N^2 或者更高。使用单核单线程，耗费的时间是条曲线上升趋势。虽然稀疏矩阵采用了CSR等方法压缩数据，但自由度的平方数据仍然还是海量数据。现在的CPU普遍多核多个逻辑线程。这也要求在进行迭代法求解方程组时，合理的进行矩阵分解，参考[一篇文章入门仿真软件性能优化](#)，采用分治的方法适应硬件的需求。

这里需要注意的是单核CPU的性能以及CPU的数量。根据实际情况进行分治！

在满足了单机CPU运行之后，再就是集群分布式计算。采用多台计算设备，比如刀片等专业服务器，超算等等，这里需要考量的是负载均衡和网络带宽，即保证单台设备的利用率和整体进度保持一致。对于超大量的计算，最好能软件实现动态规划，对于频繁交换数据的计算，要注意网络带宽的瓶颈。

再从软件方面来看，根据笔者的研发经验，相当一部分仿真的性能问题可以通过软件以及改进算法解决。在前处理中，尽量避免软件中高频操作，大数据拷贝，大量细小数据操作等。如果必须有，对操作进行采样性能评估，保证在可接受范围内。**在求解线性方程组前，提取矩阵特征，尽可能在最细度上根据矩阵特征，同时结合模型和业务的特点，选取最合适的求解方法。**

其实目前对于数值计算方法而言，其算法都是公开透明的，商业软件之间真正较量的是业务模型的准确性，算法的稳定性以及效率。而这也恰恰是每个软件可以改进和提升的地方

2004年，ANSYS解出了1亿自由度的模型，2008年，ANSYS解出了10亿自由度的模型。即使现在来看，我们认为自由度超过1千万的模型是个大模型，所以1亿自由度的模型仍然可以认为是超大模型，且只能在超大的服务器上运行，可以想象，随着硬件的提升和算法的改进，实际应用中会有越来越多的超1亿自由度的模型求解。

回到开始的问题，“一亿”是“一千万”的十倍吗？显然不是。

最后还是以业内的一句话作为结束语：

“如果你的仿真还没有受到硬件的限制，那说明你的仿真还没有入门”。